

Taller Excel

Análisis de datos



Iván Santiago Rivera
isantiago@prits.pr.gov
versión 2.0
28 de noviembre, 2023

Tabla de contenido

Sobre este tutorial.....	3
MÓDULO 1: GRÁFICAS Y FILTROS.....	4
GRÁFICAS	4
GENERAR GRÁFICAS.....	4
CÓMO GENERAR UNA GRÁFICA EN EXCEL.....	5
ESCOGER UNA GRÁFICA APROPIADA.....	6
QUÉ TIPO DE GRÁFICA ELEGIR.....	6
SERIES Y CATEGORÍAS	11
EJERCICIO 1: GRÁFICAS.....	12
FORMATEAR DATOS	16
FILTRAR DATOS.....	17
SPARKLINES (MINI-GRÁFICAS).....	19
EJERCICIO 2: AÑADIR SPARKLINES	23
MÁS PRÁCTICAS CON GRÁFICAS	25
EJERCICIO 3: FORMATEAR DATOS.....	27
EJERCICIO 4: GRÁFICA DE DISPERSIÓN Y FILTRAR DATOS.....	28
EJERCICIO 5: GRÁFICA BOXPLOT	36
EJERCICIO 6: GRÁFICA TIPO RADAR	40
EJERCICIO 7: GRÁFICAS CON SERIES TEMPORALES	52
MÓDULO 2: EXCEL TABLAS Y FUNCIONES.....	69
BENEFICIOS DE LAS TABLAS EXCEL	69
ALGUNOS CONCEPTOS ANTES DE DISCUTIR SOBRE FUNCIONES.....	73
FUNCIONES.....	76
TABLAS EXCEL: RESUMIR DATOS	87
PRÁCTICAS CON TABLAS:	89
EJERCICIO 1: CREAR TABLAS.....	89
EJERCICIO 2: TRABAJAR CON TABLAS.....	90
EJERCICIO 3: ANÁLISIS DE DATOS.....	92
EJERCICIO 4: USO DE FUNCIONES SUMIF Y COUNTIF	94
EJERCICIO 5: FUNCIONES XLOOKUP Y LOOKUP	95
MÓDULO 3: PIVOTTABLES Y PIVOTCHARTS	100
DIMENSIONES Y MÉTRICAS EN PIVOTTABLES.....	100
CONSOLIDAR COLUMNAS (VECTORIZACIÓN)	115
PIVOTCHARTS	118

PRÁCTICAS: CREAR PIVOTTABLES Y PIVOTCHARTS	122
EJERCICIO 1: CREAR UNA TABLA PIVOT	122
EJERCICIO 2: VECTORIZACIÓN O CONSOLIDACIÓN (UNPIVOT) DE MÚLTIPLES COLUMNAS .	127
EJERCICIO 3: CREAR PIVOTCHART	141
EJERCICIO 4: CAMPOS CALCULADOS EN PIVOTTABLES.....	146
EJERCICIO 5: AÑADIR SLICERS Y TIMELINE.....	151
FUENTES.....	158

Sobre este tutorial

Este tutorial está diseñado para ofrecerse en persona por dos días. Proveerá instrucciones y recomendaciones sobre el análisis de datos usando Excel 365. No es un curso básico de Excel y requiere familiaridad con este programa. Los datos para prácticas serán provistos. Habrá poca o ninguna entrada manual de datos.

PRE-REQUISITOS:

- Conocimiento general del sistema operativo Windows 10
- Conocimiento básico de manejo de hojas de cálculo Excel

Contenido DE ESTE TUTORIAL

- **Módulo 1: Gráficas y filtros**
- **Módulo 2: Tablas Excel** y algunas **funciones** para Análisis de datos
- **Módulo 3: PivotTables y PivotCharts**, describe cómo producir y trabajar con PivotTables y PivotCharts

MÓDULO 1: GRÁFICAS Y FILTROS

Este módulo trata sobre **cómo convertir datos en gráficas**. Excel es muy utilizado, en parte por la facilidad de generar gráficas a partir de datos. En este módulo se considerarán algunas recomendaciones generales para generar gráficas más efectivas.

OBJETIVOS

- **Crear gráficas** a partir de datos
- **Filtrar datos** y formatear gráficas
- Crear **Sparklines**
- Escoger **el tipo de gráfica más apropiada** para los datos

GRÁFICAS

Las gráficas se utilizan para resumir y entender datos extensos. Una gráfica puede resumir múltiples líneas y columnas y hacer visibles, tendencias, proporciones y patrones imperceptibles en tablas.

OBJETIVOS

Al completar esta lección, podrá ser capaz de:

- **Crear gráficas** usando datos provistos
- **Entender los diferentes tipos de gráficas** disponibles en Excel
- **Seleccionar el tipo de gráfica** más apropiado para los datos que quiere representar

GENERAR GRÁFICAS

- Gráficas
 - Son representaciones visuales de los datos
 - Ayudan a visualizar patrones en los datos
 - Buenas para ver datos a un nivel más alto
 - No son para dar detalles de los datos
- Para crear una gráfica:
 - Sombrear (escoger) datos
 - Ir al **menú principal**, escoger **Insert** y seleccionar el tipo de gráfica
 - La gráfica se insertará en el worksheet

Una gráfica es una representación visual de los datos. Una gráfica usa líneas, barras, segmentos y otros elementos para mostrar relaciones entre los valores de los datos. Normalmente las gráficas tienen un eje horizontal X y otro vertical Y.

Ver datos gráficamente

A menudo se requiere ver los datos en diferentes **niveles de detalle**.

Nivel de detalle alto: para identificar tendencias, anomalías, diferencias entre clases y jerarquías, diferencias a través del tiempo.

Presentar datos usando gráficas ayuda a entender y ver ciertos patrones y anomalías rápidamente. Luego de ver patrones, tendencias y anomalías, se puede ir directamente a los datos para ver ciertos detalles.

Sin embargo, las gráficas no son buenas para ver valores precisos en los datos (para esto están las tablas de valores, si son pocos). A menudo, se requerirá incluir gráficas y tablas pequeñas en los reportes. Una manera efectiva de comunicar datos complejos es presentarlos desde un enfoque de arriba hacia abajo, (top-down) usando gráficas para el nivel más alto y tablas con algunos detalles necesarios.

CÓMO GENERAR UNA GRÁFICA EN EXCEL

Esto solo es una ilustración. Es solo para entender los pasos generales.

1. Sombrear los datos que se usarán para hacer la gráfica. Puede usar las teclas **CTRL+A** siempre y cuando esté dentro del *data range*:

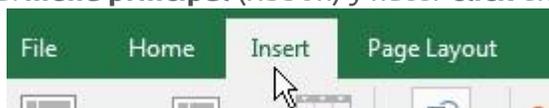
	A	B	C
1	Municipio	Población	2020
2	Culebra	1,792	
3	Maricao	4,755	
4	Vieques	8,249	
5	Las Marías	8,874	
6		89	
7		07	
8	Florida	11,692	
9	Guánica	13,787	
10	Jayuya	14,779	
11	Piñón	15,107	

Qué es un *Data Range*:

Es una extensión o grupo de celdas contiguas con valores. Cualquier espacio de una o más celdas vacías (NODATA) interrumpe o separa un *Data Range* de otro.

Esta es una lista de población por municipios, Censo de 2020.

2. Ir al **menú principal** (ribbon) y hacer **click** en el tab **Insert**

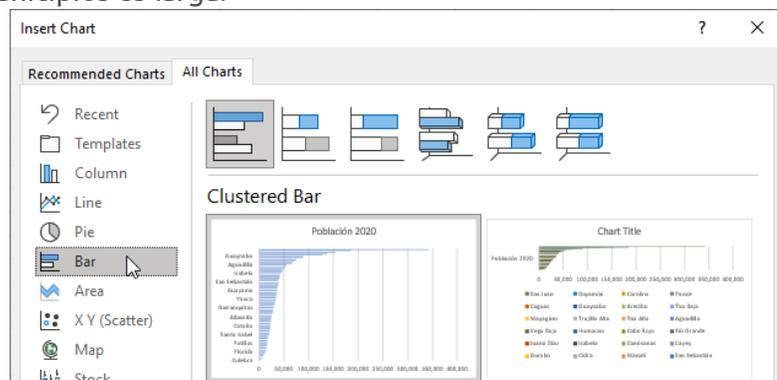


3. Hacer **click** en el botón **Recommended Charts**
Dependiendo de los datos, Excel le puede devolver un listado de gráficas.



4. Puede escoger también la gráfica entre la lista recomendada o alguna otra que aparecerá en el lado izquierdo de la forma **Insert Chart**.

Escoja el tab Bar. La opción **Clustered Bar** es la gráfica recomendada, ya que la lista de municipios es larga.



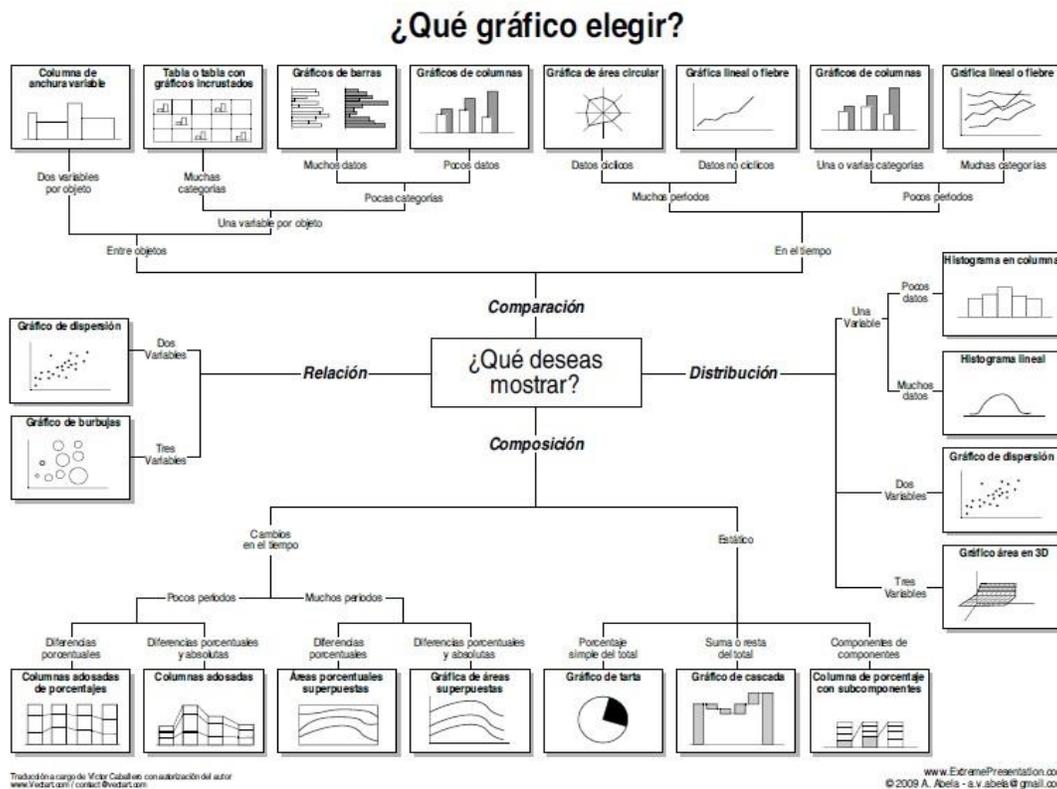
5. Hacer **click** en el botón **OK** para insertar la gráfica en la hoja de cálculo activa.

ESCOGER UNA GRÁFICA APROPIADA

Excel tiene una gran variedad de tipos de gráficas. Producir una gráfica es fácil en Excel; sin embargo, escoger el tipo correcto de gráfica requiere cierta consideración según el tipo de conjunto de datos y la audiencia. Excel incluye los siguientes tipos de gráfica: Líneas, columnas, barras, pie charts (pastel), áreas, diagrama de dispersión, stock (gráfica de líneas para acciones en el mercado), superficie, radar, treemap, sunburst, histograma, box and whisker, cascada.

QUÉ TIPO DE GRÁFICA ELEGIR

Este diagrama muestra cómo escoger el gráfico más apropiado, según los datos y el propósito de la gráfica.



Tomado de: <http://extremepresentation.typepad.com/files/que-grafico-elegir.pdf>

versión en inglés: <http://extremepresentation.typepad.com/files/choosing-a-good-chart-09.pdf>

La pregunta central es **¿QUÉ DESEA MOSTRAR?**

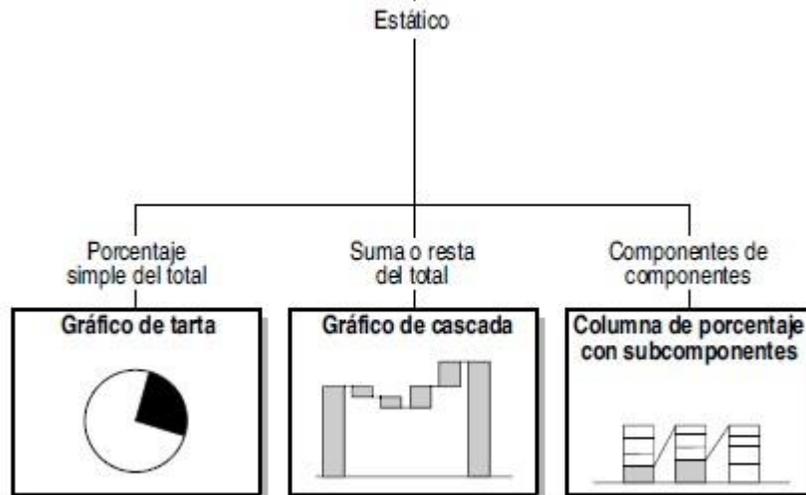
¿Comparación, Distribución, Composición, Relación?

Para cada una de estas acciones, habrá diferentes opciones para crear una o más gráficas para representar los datos.

GRÁFICAS DESCRIPTIVAS: Composición y Comparación:

Para visualizar o explorar datos mediante resumen. No supone necesariamente un análisis previo de los datos.

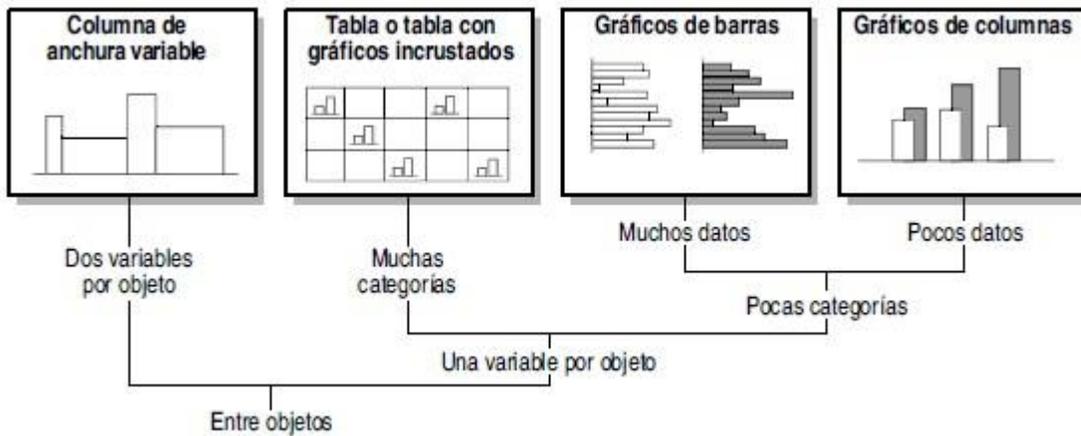
Composición: *Estático*



Composición: *Cambios en el tiempo*



Comparación: entre objetos



Comparación: en el tiempo

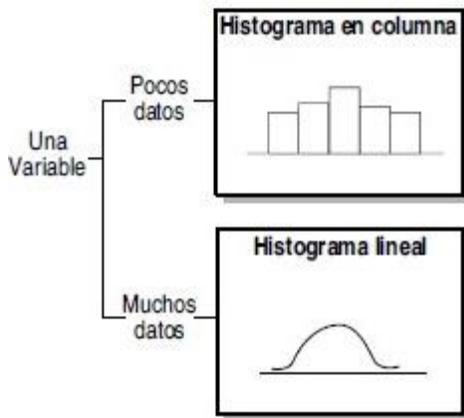


ANALÍTICAS:

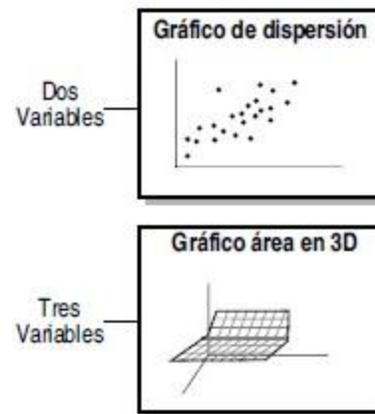
Distribución y Relación: Suponen que ha ocurrido exploración y análisis previo de los datos, mostrar patrones ocultos, reducir información.

Distribución:

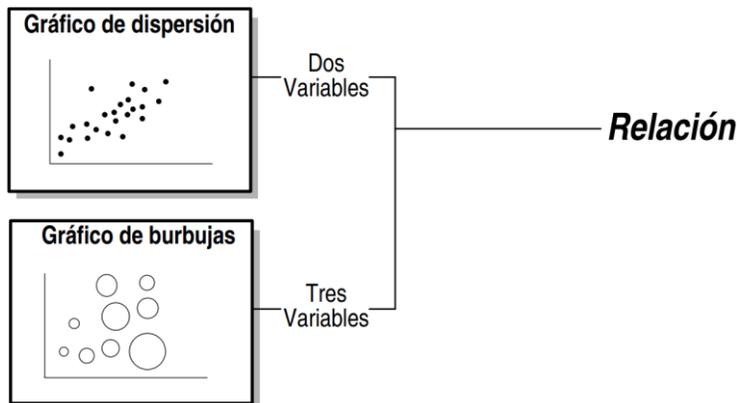
De una variable:



De dos o más variables:



Relación: entre una o más variables:



ALGUNAS RECOMENDACIONES:

PIE CHART:

Este gráfico se recomienda para comparar entre pocas categorías (3, 4, o 5) y las categorías no deben parecerse mucho. La vista humana no es buena distinguiendo porciones de áreas circulares (*wedges*). Si va a usar un pie chart, las [recomendaciones](#) son:

- No usar más de 5 categorías
- Asegúrese que los valores sumen 100%
- Ordene las porciones del pie chart por tamaño
- Las proporciones deben mostrar diferencias sustanciales

BARRA O COLUMNA: EL MÁS RECOMENDADO PARA REPRESENTAR MAGNITUDES

Al contrario de las gráficas pie chart, las gráficas de barras o columnas son más apropiadas para mostrar las magnitudes y proporciones. Para pocas categorías, es mejor usar un diagrama de columnas; para muchas categorías, es mejor usar un gráfico de barras.

LÍNEAS

Las gráficas de líneas muestran tendencias por periodos de tiempo. También pueden mostrar varias categorías en una sola gráfica. Se recomiendan para distribuciones de datos continuos. Se recomienda:

- Prefiera usar líneas sólidas
- Evite usar más de cuatro categorías
- Altura de línea adecuada: debe tomar más o menos 2/3 partes de la altura del eje Y.

MAPA PROPORCIONAL (TREEMAP)

Este tipo de gráfica ha sido utilizado generalmente para representar datos proporcionales por jerarquías (Grupos y subgrupos). Se utiliza por ejemplo para representar datos presupuestarios.



En este ejemplo, cada dato de población municipal estimada en 2016 está clasificado por zonas: Metro, Sur, Norte, etc. Los colores representan las zonas.

SERIES Y CATEGORÍAS

Para propósitos de hacer gráficas en Excel, los datos son clasificados en series y categorías. Las categorías se asocian a etiquetas (labels) y las series se asocian a datos numéricos que queremos representar en la gráfica.

Por ejemplo, tomemos este conjunto de datos

	A	B	C	D
1	Década	Bayamón	Ponce	Mayagüez
2	1930	29,524	87,604	58,270
3	2000	224,044	186,475	98,434
4	2020	185,187	137,491	73,077

La parte sombreada, son datos numéricos.
La parte no sombreada son categorías.

Las *décadas* podrían ser tanto numéricas como categorías. Depende de cómo las queremos representar en la gráfica.

¿Cuáles serían las series en este caso?

Puede prestarse a confusión, por ejemplo:

Esta secuencia de filas puede ser una serie

29,524	87,604	58,270
--------	--------	--------

, y podemos ponerle una categoría, "1930".

29,524
224,044
185,187

De forma alterna, **esta puede ser otra serie:**

185,187

 que representa la población en las décadas de 1930, 2000 y 2020 para el Municipio de Bayamón.

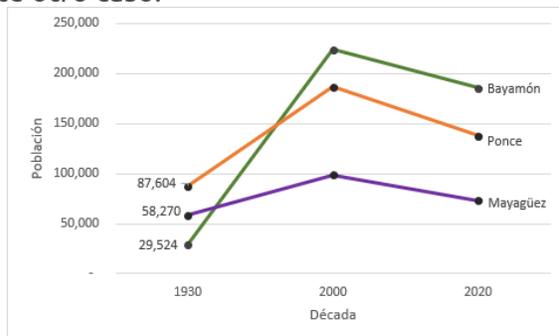
Este data range puede ser representado en una gráfica de varias maneras, por ejemplo, como una gráfica de columnas:



Ninguna de las dos gráficas está incorrecta. Todo depende de las preferencias y prioridades.

En ambas, Excel interpreta que los datos numéricos están en el eje vertical (Y o variable dependiente) y las categorías en el eje horizontal (X o variable independiente). Tanto puede ser los municipios como las décadas.

Vea este otro caso:



Esta es otra manera diferente de hacer la gráfica: **usando líneas**, donde las alturas cambiantes (**Eje Y** o dependiente) van a representar las **poblaciones** de municipios, mientras que las décadas van a estar en el eje horizontal.

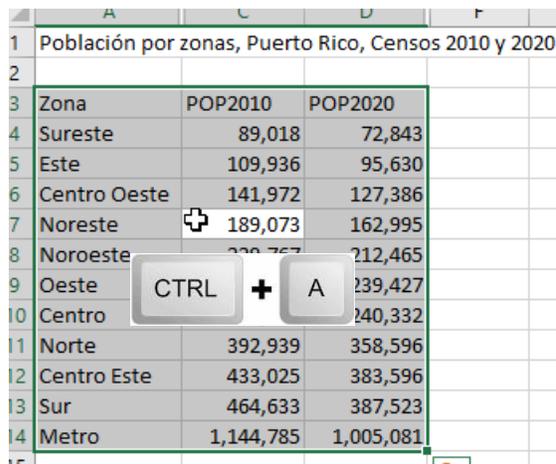
Series: poblaciones
Categoría: Municipios (color de línea)
Categoría: Décadas (en el eje X)

EJERCICIO 1: GRÁFICAS

- Abra Excel y busque el archivo **zonas_poblacion_municipios_censos2010-20.xlsx** en el folder **C:\Tutorial_Excel\DataAnalysis\Modulo_1**

Población por zonas, Puerto Rico, Censos 2010 y 2020			
Zona	POP2010	POP2020	
Este	83,216	72,244	
Sureste	89,018	72,843	
Noroeste	148,539	136,180	
Noreste	215,793	186,381	
Centro Oeste	223,988	202,354	
Oeste	237,235	208,551	
Norte	418,858	382,234	
Sur	490,386	410,857	
Centro Este	678,453	609,149	
Metro	1,144,785	1,005,081	

- Haga **click en cualquiera celda dentro del data range** y presione **CTRL+A**. Esto sombreadá el data range

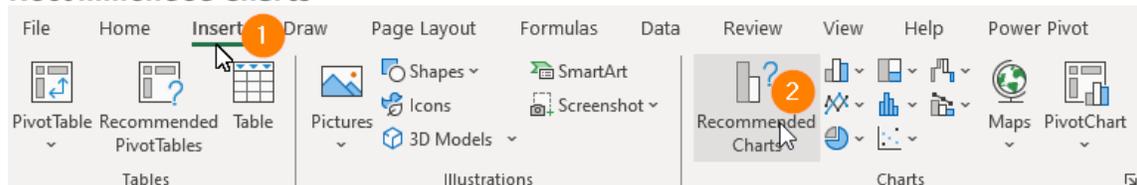


Población por zonas, Puerto Rico, Censos 2010 y 2020			
Zona	POP2010	POP2020	
Sureste	89,018	72,843	
Este	109,936	95,630	
Centro Oeste	141,972	127,386	
Noreste	189,073	162,995	
Noroeste	148,539	136,180	
Oeste	237,235	208,551	
Centro	223,988	202,354	
Norte	392,939	358,596	
Centro Este	433,025	383,596	
Sur	464,633	387,523	
Metro	1,144,785	1,005,081	

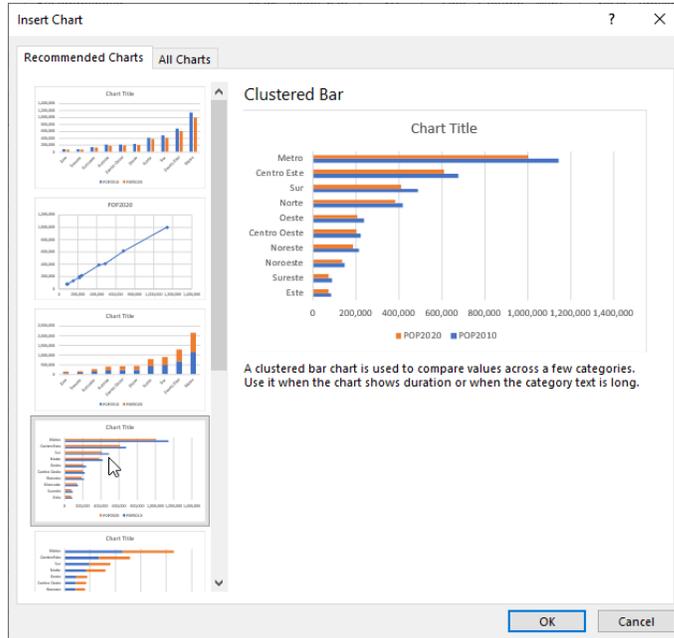
Puede usar esta combinación de teclas para evitar tener que sombread manualmente un data range.

Recuerde que al usar **CTRL+A**, **seleccionará todas las celdas contiguas** que tengan algún valor. Si hay alguna celda vacía, Excel no la incluirá.

- Una vez **sombreado el data range**, haga **click en el tab Insert**; luego en el botón **Recommended Charts**



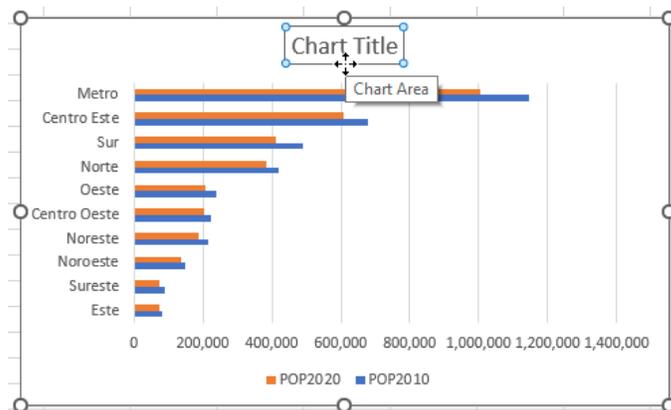
Aparecerá la forma **Insert Chart**.



- Haga **click** en las diferentes gráficas al lado izquierdo, las cuales son recomendadas y compararlas...
- Finalmente haga **click** en la opción **Clustered bar** y haga **click** en el botón **OK**.

MODIFICAR EL TÍTULO

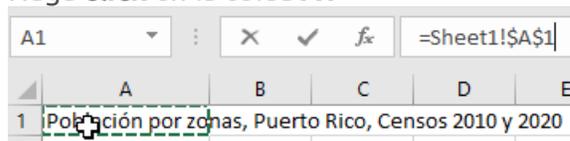
- En la gráfica, haga **click** en el **título** para activarlo.



- Haga **click** en la caja de texto **Formula Bar**, y escriba el signo de igualdad =



- Haga **click** en la celda **A1**



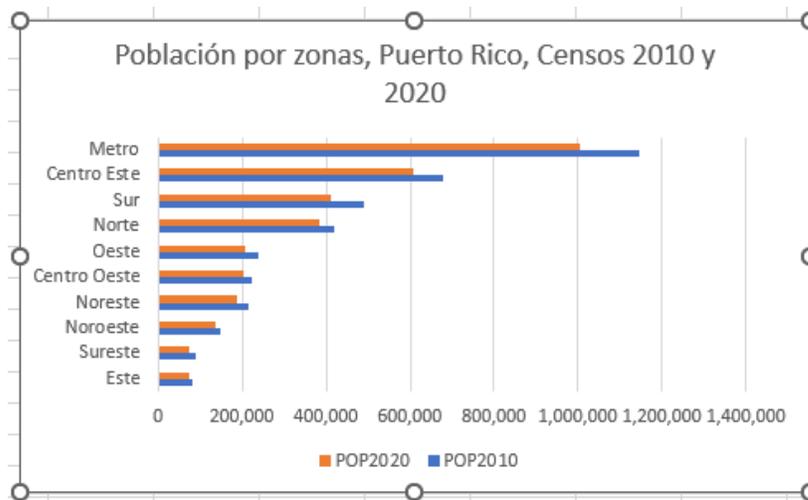
Notará que en **Formula Bar**, cambia el texto a **=Sheet1!\$A\$1**.

Los signos de dólar (\$) son referencias fijas o móviles a una o más celdas, pero esto se explicará más adelante.

- En la caja de texto **Formula Bar**, presione la tecla **Enter** para realizar este cambio.

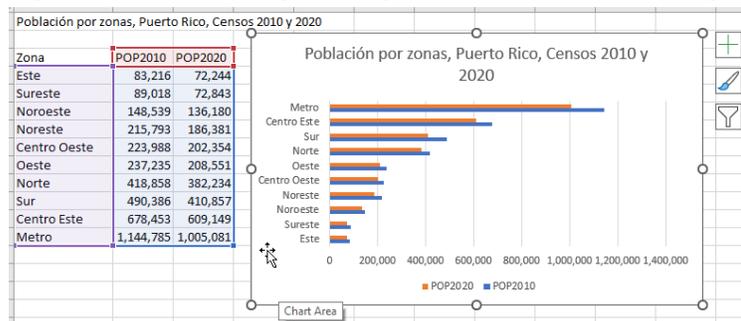


Podrá ver que el título **adoptó el contenido** de la celda **A1** de la hoja **Sheet1**.



MOVER LA GRÁFICA A OTRA HOJA

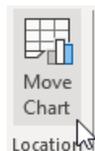
- Haga **otro click** en la gráfica, pero fuera del título. La gráfica deberá estar activada.



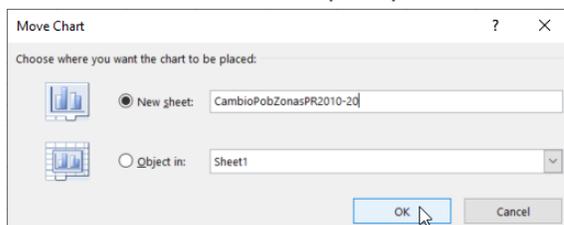
Note que, **al activar la gráfica completa, la hoja se activa y muestra los datos fuente** para esta gráfica.

Cuando la gráfica está activa, se habilita el tab **Chart Design**.

- A la derecha del conjunto de botones del Ribbon de Excel, vaya a la sección **Location** y haga **click** en el botón **Move Chart**



- En la forma **Move Chart** que aparecerá:



- Haga **click** en la opción **New Sheet**
- En la caja de texto escriba el nombre: **CambioPobZonasPR2010-20**
- Haga **click** en el botón **OK** para aceptar los cambios y cerrar la forma.

Note el **tab** abajo con el nombre de la gráfica (**CambioPobZonasPR2010-20**).

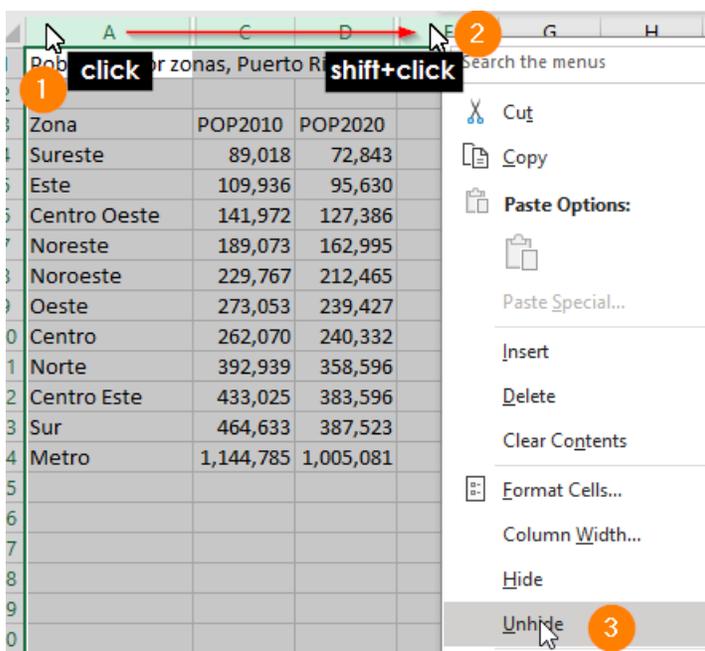


Pregunta: ¿Por qué se prefirió usar la gráfica de barras en lugar de la gráfica de columnas?

Respuesta en pág. 9

UNHIDE (REVELAR COLUMNAS)

En ocasiones, hay columnas que por conveniencia no aparecen, por ejemplo, para que no aparezcan esos datos en una gráfica.:

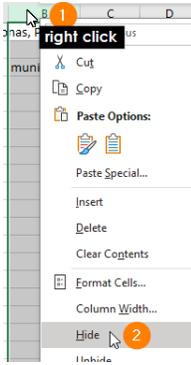


- Haga **click** en la **cabecera** de la **columna A**.
- Haga **shift+click** en la **cabecera** de la **columna F**.
- Haga **click** en la opción **Unhide** para revelar estas columnas.

ESCONDER (HIDE) COLUMNAS

Esconder columnas permite hacer gráficas sin necesidad de borrar ni hacer *cut & paste* de los datos. Esto ayuda a reducir errores. Al esconder columnas, los datos permanecen intactos y las columnas podrán ser mostradas posteriormente, de ser necesario.

Para esconder columnas:



- Seleccione/**sombree** una o más columna(s).
- Haga **right click** y escoja la opción **Hide**.

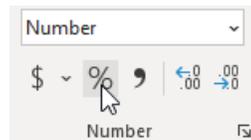
FORMATEAR DATOS

Un formato correcto hace que los números y las fechas puedan entenderse mejor.

Para formatear celdas:

2020	ΔPCT
2,843	-0.10
5,630	-0.07
7,386	-0.05
2,995	-0.07
2,465	-0.04
3,427	-0.07
1,332	-0.04
3,596	-0.05
3,596	-0.06
7,523	-0.09
5,081	-0.06

- Seleccione** la(s) **celda(s)** a formatear usando **click & drag**.
- Vaya al **menú principal** y escoja **Home** y vaya al **grupo Number**
- Haga **click** en el botón % porcentaje, ya que estos números representan porcentajes que no han sido multiplicados por 100 todavía



- Mantenga** solamente **un espacio decimal**. Use los botones **Increase o Decrease Decimal**



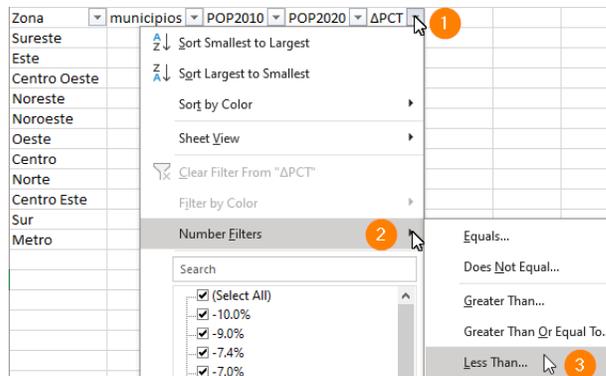
para que haya solo un lugar decimal.

ΔPCT
-10.0%
-7.0%

APLICAR UN FILTRO NUMÉRICO

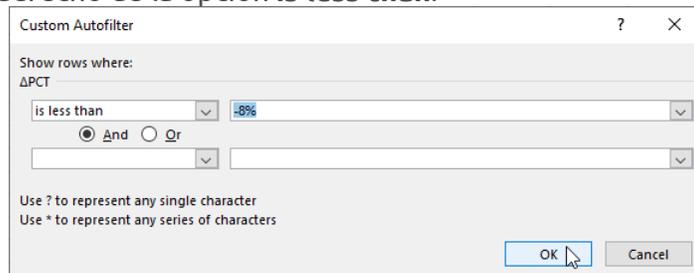
Vamos a filtrar para que aparezcan las zonas o regiones con **cambio poblacional menor que -8%**.

- Siga estos pasos:



- Haga **click** en el **botón drop down** de la columna **ΔPCT** (Cambio porcentual)
- **Click** en **Number Filters**.
- **Click** en la opción **Less Than...**

- En la forma **Custom Autofilter** que aparecerá, escriba **-8%** en la caja de texto al lado derecho de la opción **is less than**.



En este caso **NO debe escribir ni 8 ni -8 SIN el signo de por ciento**. Ambos criterios (8 o -8) están fuera del alcance de los datos de esta columna que va desde 3.9% a 10% (0.39 hasta 0.10 en términos decimales).

- Haga **click** en el botón **OK** para aplicar el filtro.

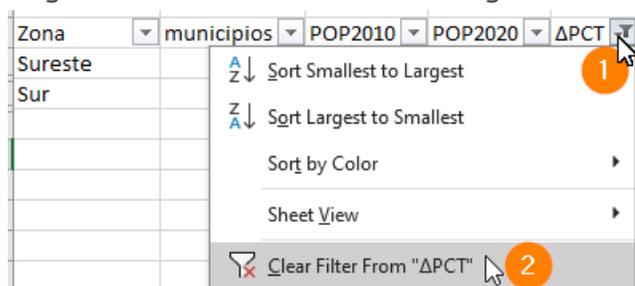
Podrá ver los casos con el filtro aplicado. Las zonas con mayor pérdida poblacional son la zona **Sureste** con -10% y la zona **Sur** con -9%

Zona	municipios	POP2010	POP2020	ΔPCT
Sureste	4	89,018	72,843	-10.0%
Sur	10	464,633	387,523	-9.0%

QUITAR FILTROS

Para quitar filtros:

- Haga **click** en el **botón de filtro**. Haga **click** en la opción **Clear Filter From**



- Cierre esta hoja de cálculo. No es necesario guardar el archivo.

SPARKLINES (MINI-GRÁFICAS)

Los *sparklines* son mini gráficas que se generan dentro de una celda. Es una manera de visualizar tendencias en datos voluminosos, sin tomar demasiado espacio.

OBJETIVOS

Al completar esta lección, usted podrá:

- Añadir sparklines
- Escoger el sparkline más apropiado
- Remover sparklines
- Formatear sparklines

Los sparklines son mini-gráficas que caben dentro de una celda

- Hay tres tipos: Líneas, columnas, Win/Loss
- Pueden añadirse y removerse
- Deben estar ubicadas cerca de los datos para un mejor entendimiento.

QUÉ ES UNA MINI-GRÁFICA SPARKLINE

Los sparklines no son objetos de una gráfica, aunque parecen pequeñas gráficas. Se trata de gráficos simples ubicados en el fondo de una celda. Esto significa que puede haber texto encima del espacio donde aparece el sparkline. Cada sparkline representa una fila de datos.

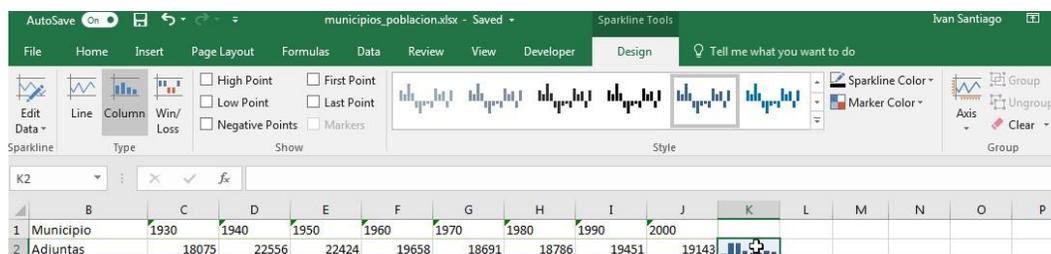
USO DE SPARKLINES

Los sparklines muestran patrones en los datos, haciendo que la vista se concentre en puntos específicos en los datos. Estas ayudan a mostrar más fácilmente los patrones existentes en los datos y fomentan que las personas miren dónde están los números que crean estos patrones.

FORMATEO DE SPARKLINES

Para hacer cambios a sparklines:

- **Seleccionar** la **celda** donde está el sparkline. Se activará el tab **Sparkline Tools | Design**

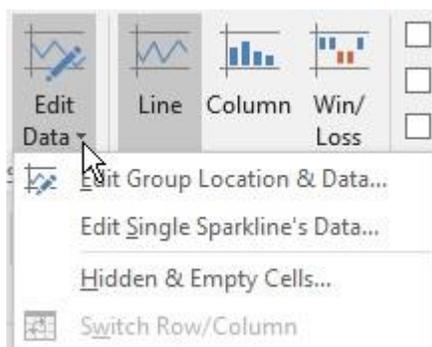


- **Edit data:** Cambiar el data range y la posición del sparkline
- **Tipo:** Línea, columna, win-loss
- **Show:** Resaltar lugares específicos de los datos
- **Style:** Colores y estilos
- **Group:** Crear o remover grupo, eliminar sparklines, definir configuración *Settings* de los ejes

Nota: Los sparklines pueden ser agrupados. Esto facilita modificar varios sparklines incluidos dentro del grupo.

MENÚ Y HERRAMIENTAS SPARKLINES

La opción **Edit Data** le permite **cambiar el data range** utilizado para generar el sparkline, así como la **ubicación del sparkline**. Al hacer **click** en la flecha, aparecerán las opciones:



Edit Group Location and Data: Para hacer cambios en cuál es el data range a usarse, así como la localización del sparkline.

Edit Single Sparkline's Data: Para hacer cambios (data range y ubicación) para una celda sparkline.

Hidden and Empty Cells: Para especificar cómo será el manejo de celdas vacías, además de especificar la inclusión o exclusión en el sparkline de celdas escondidas.

TIPOS DE SPARKLINES

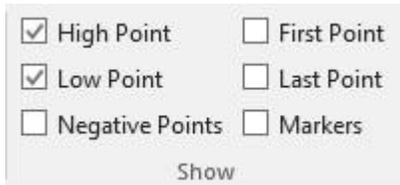
Línea, columna, Win/Loss



Estas pueden ser cambiadas al tipo apropiado, según los datos.

SHOW

Para llamar la atención a diferentes puntos de los datos. Por ejemplo:



High Point: Mostrar el punto más alto.

Low Point: Mostrar el punto más bajo.

Negative Points: Mostrar puntos negativos, si existen

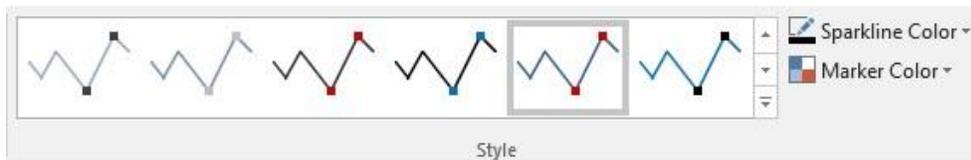
First Point: Primer punto

Last Point: Último punto

Markers: Marcadores (mostrar los vértices solo en líneas)

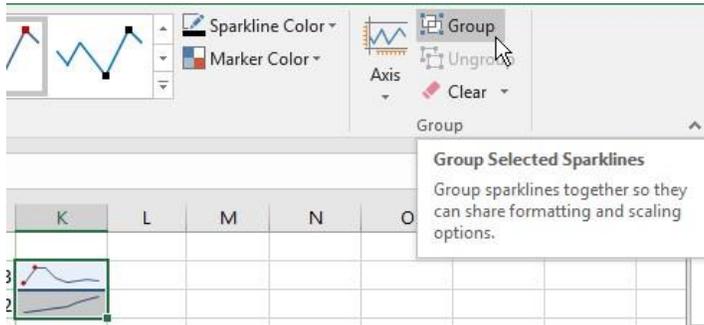
STYLE

Este grupo de botones sirve para cambiar los colores de las líneas de un grupo, además de los colores de los markers.



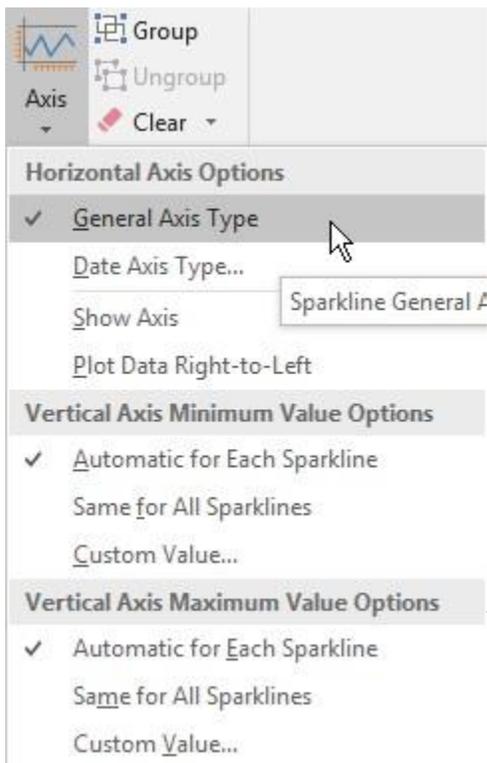
GROUP

El botón **Group** se usa para agrupar varios sparklines y tratarlos como una sola mini-gráfica al momento de cambiar el formato.



AXIS

El menú Axis contiene múltiples opciones.



El menú Axis contiene múltiples opciones.

Axis permite hacer **cambios en los ejes** de la mini gráfica o en grupos de mini gráficas

Por ejemplo, podemos **establecer el mismo punto de origen y extensión del eje** para **estandarizar las escalas**. De esta manera se podrán hacer comparaciones de manera más efectiva.

El eje horizontal puede cambiarse a fecha y además invertirlo

Opciones del eje vertical para valor mínimo

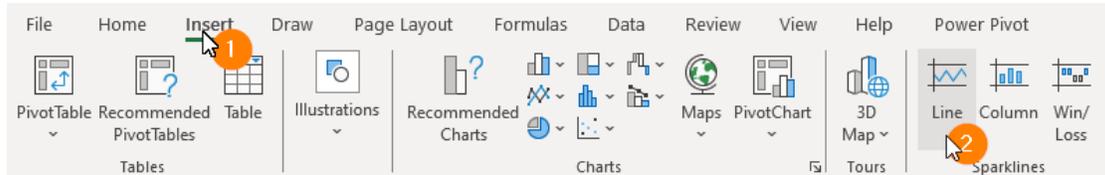
Opciones del eje vertical para valor máximo

EJERCICIO 2: AÑADIR SPARKLINES

- Para añadir Sparklines en Excel, abra el workbook **municipios_poblacion_1930-2020.xlsx**, localizado en el folder **C:\Tutorial_Excel\DataAnalysis\Modulo_1**
- Haga **click** en la celda **L2**

K	L	M
2010	2020	
19483	18020	
41050	38136	

- Vaya al **menú principal** y haga **click** en **Insert**, luego vaya a la derecha, a la sección **Sparklines**, y haga **click** en el botón **Line**.



Aparecerá la forma (dialog box) **Create Sparklines**.

Create Sparklines ? X

Choose the data that you want

Data Range:

Choose where you want the sparklines to be placed

Location Range:

OK Cancel

- Haga **click** en el botón  para **escoger la serie de datos** (aunque lea "data range").

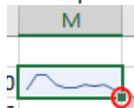
- Haga **click** en la celda **C2** y **arrastre el pointer** (la cruz) **hasta la celda L2**

	B	C	D	E	F	G	H	I	J	K	L	M
	Municipio	1930	1940	1950	1960	1970	1980	1990	2000	2010	2020	
1	Adjuntas	18075	22556	22424	19658	18691	18786	19451	19143	19483	18020	
3	Aguada	14670	17923	20743	23344	25550	21557	25911	49000	1959	38136	
5	Aguadilla	28319	34956	44357	41344	41344	41344	41344	41344	3949	55101	
7	Aguas Buenas	12885	14671	15565	14111	14111	14111	14111	14111	3659	24223	
9	Aibonito	16361	16819	18191	14111	14111	14111	14111	14111	5900	24637	

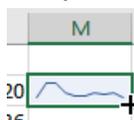
- Presione** la tecla **Enter**.

- De regreso a la forma **Create Sparklines**, haga **click** en el botón **OK** para aceptar los cambios y cerrar la forma.

- Estando ubicado en la celda **M2**, **fijese** en el pequeño **cuadro** (punto) **verde** que aparece en la esquina inferior derecha de la misma:



- Ubique el cursor en ese punto verde y haga **click encima del punto verde**.



- Mantenga el punto presionado y **arrastre** el cursor **hasta la celda M79**

78	72151	Yabucoa	21914	27438	28810	29782	30165	31425	36483	39246	37941	30426	
79	72153	Yauco	27787	30533	33708	34780	35103	37742	42058	46384	42043	34172	

Podrá ver los Sparklines en la columna M

B	C	D	E	F	G	H	I	J	K	L	M
Municipio	1930	1940	1950	1960	1970	1980	1990	2000	2010	2020	
Adjuntas	18075	22556	22424	19658	18691	18786	19451	19143	19483	18020	
Aguada	14670	17923	20743	23234	25658	31567	35911	42042	41959	38136	
Aguadilla	28319	34956	44357	47864	51355	54606	59335	64685	60949	55101	
Aguas Buenas	12885	14671	15565	17034	18600	22429	25424	29032	28659	24223	
Aibonito	16361	16819	18191	18360	20044	22167	24971	26493	25900	24637	
Añasco	14276	15701	17235	17200	19416	23274	25234	28348	29261	25596	

Verá que **la mayoría de los municipios aumentaron su población de 1930 a 2000**. Además, podrá notar la pendiente de cada línea, la cual le dará idea de la magnitud del aumento o descenso de la población municipal.

AÑADIR ELEMENTOS A LOS SPARKLINES

- Haga **click** en la celda **M2**. Verá que esta celda **activa** también **el resto de las celdas contiguas hacia abajo** hasta la celda **K80**. **Esto quiere decir que los sparklines de estas celdas están en el mismo grupo**.

L	M
2020	
18020	
38136	
55101	
24223	

- Vaya al grupo **Show** y **escoja check** las opciones **High Point** y **Low Point**.

<input checked="" type="checkbox"/> High Point	<input type="checkbox"/> First Point
<input checked="" type="checkbox"/> Low Point	<input type="checkbox"/> Last Point
<input type="checkbox"/> Negative Points	<input type="checkbox"/> Markers
Show	

Podrá ver los puntos mínimo y máximo en cada sparkline:

B	C	D	E	F	G	H	I	J	K	L	M
Municipio	1930	1940	1950	1960	1970	1980	1990	2000	2010	2020	
Adjuntas	18075	22556	22424	19658	18691	18786	19451	19143	19483	18020	
Aguada	14670	17923	20743	23234	25658	31567	35911	42042	41959	38136	
Aguadilla	28319	34956	44357	47864	51355	54606	59335	64685	60949	55101	
Aguas Buenas	12885	14671	15565	17034	18600	22429	25424	29032	28659	24223	
Aibonito	16361	16819	18191	18360	20044	22167	24971	26493	25900	24637	
Añasco	14276	15701	17235	17200	19416	23274	25234	28348	29261	25596	

- Fíjese en el caso del **Municipio de Adjuntas**. Verá que el punto mínimo está en el dato de población de 2020. Relaciónelo con las otras décadas. Describa qué está pasando en esa serie de datos: _____

¿**Cuáles otros municipios** están pasando por la misma situación? ¿En qué **zona(s)** están ubicados? _____

Con estas líneas, es mucho más fácil poder identificar municipios que han tenido pérdidas constantes de población o crecimiento bien modesto. Esto se puede notar en los municipios como Utuado y Vieques.

Utuado	37434	42531	46625	40449	35494	34505	34980	35336	33149	28287	
Vega Alta	12333	14329	16521	17603	22810	28696	34559	37910	39951	35395	
Vega Baja	20406	23105	28925	30189	35327	47115	55997	61929	59662	54414	
Vieques	10582	10362	9228	7210	7767	7662	8602	9106	9301	8249	
Villalba	11847	12871	14972	16239	18733	20734	23559	27913	26073	22093	

Aparecerán además otros municipios con puntos mínimos posteriores a la década de 1930; estos bajaron su población luego de 1930 en vez de aumentarla, contrario a la tendencia generalizada en

el resto de los municipios. El caso de Vieques es atípico porque muestra mayor población en 1930 que en el censo de 2020.

MÁS PRÁCTICAS CON GRÁFICAS

ESCENARIO

Imagínese que a usted le piden **mostrar gráficas sobre datos de escuelas públicas del Departamento de Educación, a nivel descriptivo**. Se trata de un directorio de escuelas públicas para 2017, el cual fue descargado en 2018 desde el portal data.pr.gov, (ya no está disponible). El interés inicial es **saber si los datos muestran algún tipo de correlación** (positiva o negativa) **entre el nivel de pobreza registrado y el nivel “competente a avanzado” en matemáticas**.

OBJETIVOS

Con las destrezas mostradas a lo largo de este módulo usted practicará:

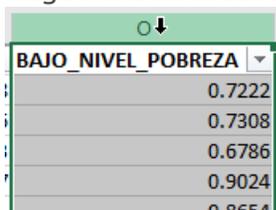
- Hacer gráficas de histograma y diagramas de dispersión para mostrar distribuciones de datos
- Cambiar formatos para hacerlos más legibles
- Filtrar datos

GRÁFICAS Y EXPLORACIÓN DE DATOS

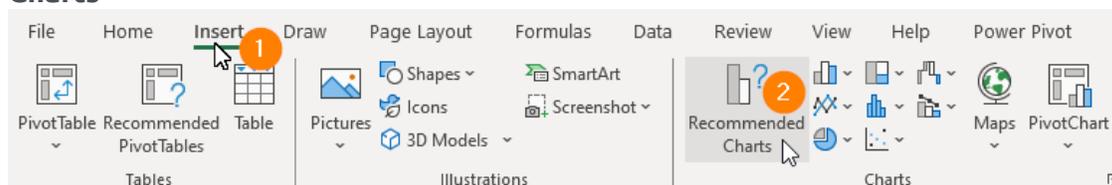
- Abra el workbook llamado **DirectorioEscuelasPublicas2017.xlsx**, localizado en el folder **C:\Tutorial_Excel\DataAnalysis\Modulo_1**
Este workbook tiene una hoja con el directorio de escuelas públicas de 2017. Se eliminaron algunas filas que no tenían datos en las columnas de puntuaciones en las competencias español-inglés-matemáticas.

HACER UN HISTOGRAMA

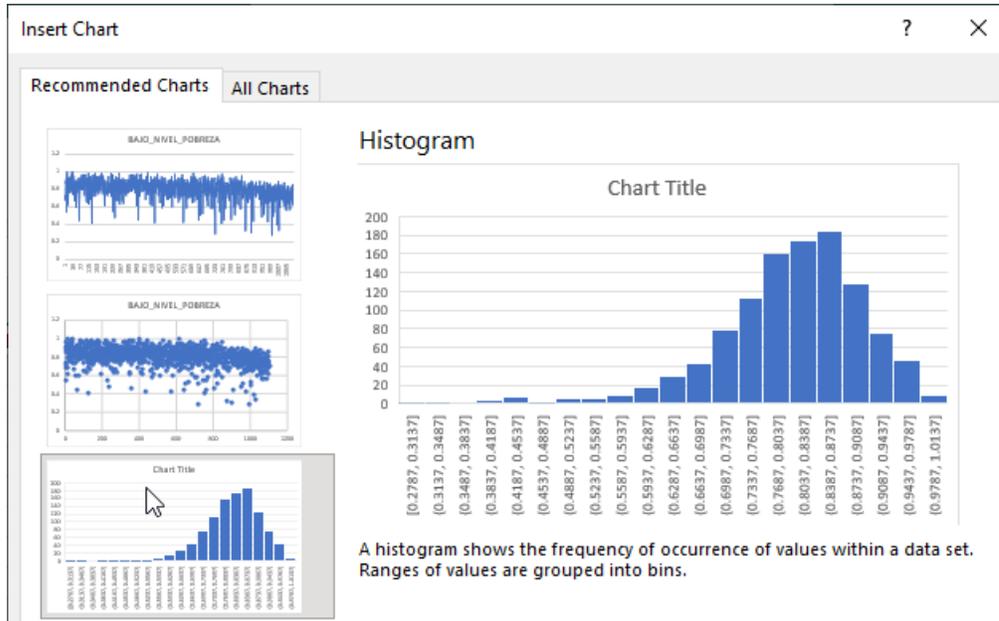
- Haga **click** en la cabecera de la columna **O (BAJO_NIVEL_POBREZA)**



- Vaya al **menú principal** y haga **click** en **Insert**, luego **click** en **Recommended Charts**

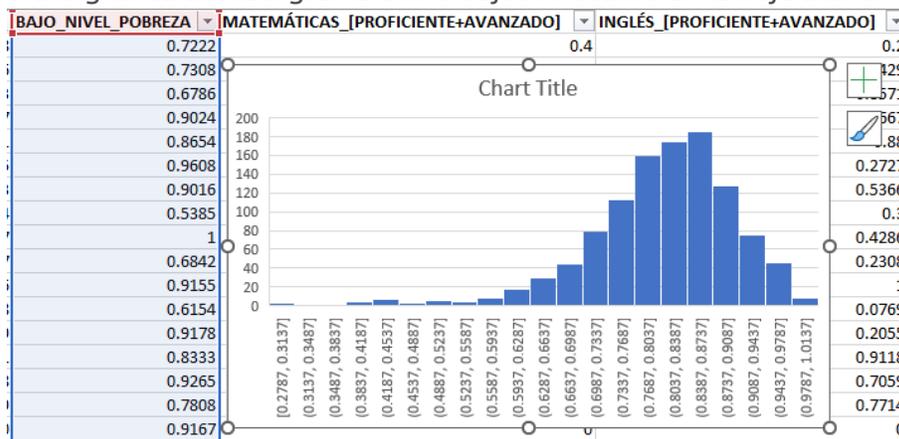


- En la forma **Insert Chart**, escoja la **tercera opción** en el panel izquierdo (*Histogram*).



- Haga **click** en el botón **OK**.

Aparecerá la gráfica de histograma como objeto flotante en la hoja activa:



Eje Y: Altura de las barras: cantidad de casos (escuelas) bajo distintos niveles. Mientras más alta más casos hay en ese bin y viceversa.

Eje X: Contiene los "bins" que vienen a ser "clases o niveles de valores (0.27 a 0.31, etc.).

En este histograma podemos ver que **la mayoría de los datos se concentra en los valores alrededor de 0.8**. Este a su vez, (80%) es el valor promedio (*mean*) de las escuelas. La distribución muestra una "cola" al lado izquierdo. Estas son las escuelas con menos estudiantes bajo el nivel de pobreza.

- **Borre esta gráfica**. Haga **click** en algún lugar del **borde** y presione la tecla **delete**.

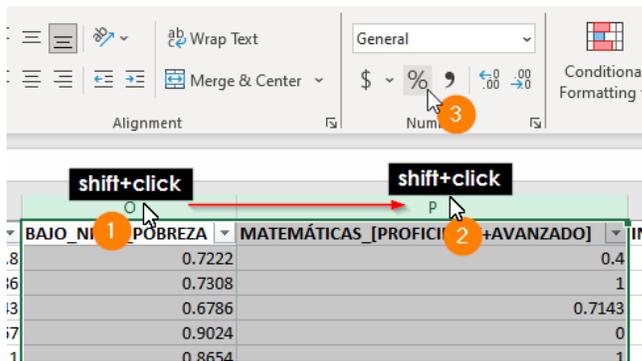
Al lado derecho del data range, hay un pequeño data range que contiene estadísticas descriptivas de los datos de las columnas **Bajo Nivel Pobreza** (BNP) y la columna **Matemáticas** (MAT).

	mean	median	mode	std dev
BNP	80%	82%	80%	10%
MAT	38%	35%	0%	29%

EJERCICIO 3: FORMATEAR DATOS

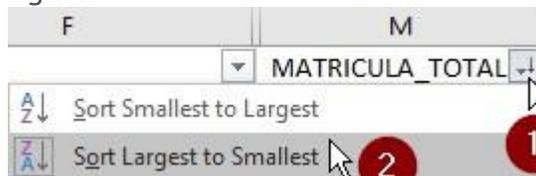
En esta parte, vamos a cambiar el formato de los números para que indiquen que tipo de cifras son. Por ejemplo, números por cantidades mostrando comas, decimales, porcentajes, moneda, y otros. La idea aquí es presentar los números en sus formatos apropiados y hacerlos más comprensibles.

- Vamos a cambiar el formato a **porcentaje** a las columnas **O** y **P**.



- Presione la tecla **SHIFT** y haga **click** en las cabeceras de las columnas **O** y **P**.
- Una vez estén seleccionadas, haga **click** en el botón  **Percent Style** para añadirle los símbolos de porcentaje.
- Haga **click** el botón  **Increase Decimal**, para añadirle un lugar decimal.

- Cambie el formato de la columna **N** a **porcentaje** y use el botón  **Increase Decimal**, para añadirle un lugar decimal.
- En la columna **M** (**MATRÍCULA_TOTAL**), cambie el formato **numérico con comas sin números decimales**.  
- **Ordene** la columna **M** **MATRÍCULA_TOTAL** de forma **descendente**, como aparece en esta figura.



Luego de ordenar la columna M, así deben verse los números.

M	N	O	P	IN
MATRÍCULA_TOTAL	ESPAÑOL [PROFICIENTE+AVANZADO]	BAJO_NIVEL_POBREZA	MATEMÁTICAS [PROFICIENTE+AVANZADO]	IN
1,309	63.6%	69.7%	17.8%	
1,191	41.2%	73.0%	18.9%	
1,185	46.3%	68.8%	2.8%	
1,182	44.0%	71.9%	3.8%	
1,148	36.4%	85.3%	0.3%	
1,094	75.1%	67.1%	65.7%	

EJERCICIO 4: GRÁFICA DE DISPERSIÓN Y FILTRAR DATOS

En esta parte, vamos a **hacer** una **gráfica de dispersión**. Luego de haber preparado la gráfica, **aplicaremos** algunos **filtros para explorar** los **datos** y ver si podemos **saber si hay algún patrón o relación** entre las variables (columnas) **nivel de pobreza** y **competencia avanzada en matemáticas**, así como **matrícula total** y **competencia avanzada en matemáticas**.

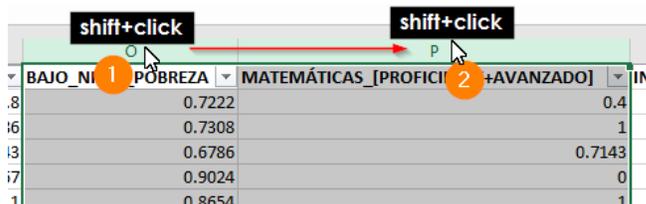
OBJETIVOS:

- Primero haremos la gráfica de dispersión (*scattergram*) con todos los datos
- Luego, cambiaremos de columnas y aplicaremos los filtros.

MODIFICAR POSICIÓN DE COLUMNAS PARA HACER DIAGRAMA DE DISPERSIÓN

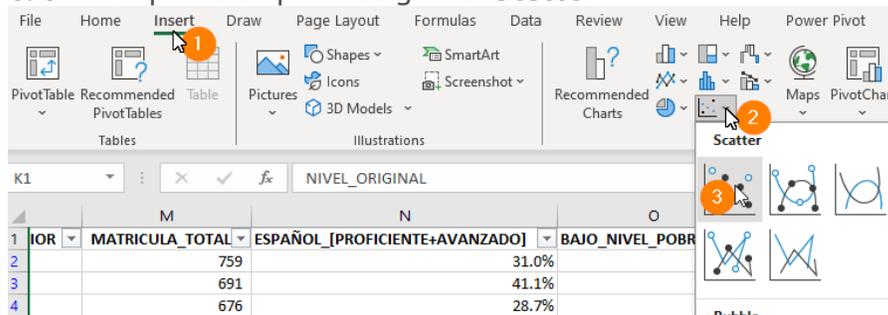
Comencemos por preguntarnos **¿El porcentaje bajo nivel de pobreza tiene alguna correlación positiva o negativa con la competencia en matemáticas?**

- Haga **click** en la cabecera de la columna **O** y **shift+click** en la cabecera de la columna **P**.



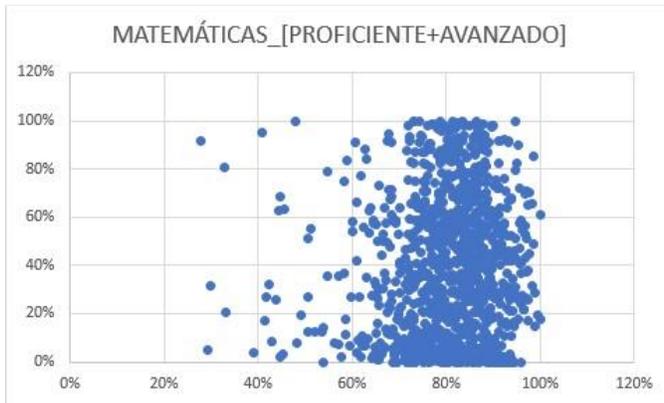
	O	P
8	0.7222	0.4
16	0.7308	1
13	0.6786	0.7143
17	0.9024	0
1	0.8654	1

- Vaya al **menú principal** y haga **click** en **Insert**, luego **click** en la **flecha** del botón **Scatter** y **click** en la primera opción de gráfica **Scatter**



	M	N	O	
1	IOR	MATRÍCULA_TOTAL	ESPAÑOL [PROFICIENTE+AVANZADO]	BAJO_NIVEL_POBREZA
2		759		31.0%
3		691		41.1%
4		676		28.7%

Así aparecerá la gráfica.

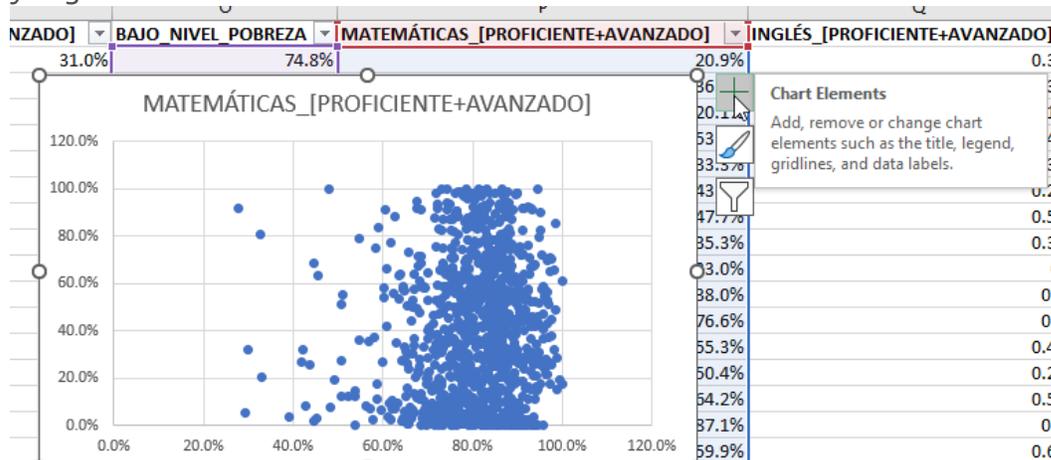


El eje horizontal (X) representa **porcentaje de escuelas bajo el nivel de pobreza**, el eje vertical (Y) representa **porcentajes de nivel de competencia en matemáticas avanzadas**

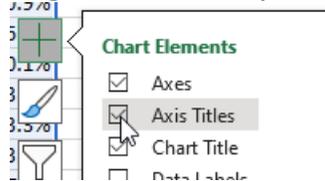
- Los porcentajes bajo el nivel de pobreza fluctúan entre 30-a-100%. Los niveles de competencia en Matemáticas van desde cero a 100%
- Los valores en el eje de X (%bajo nivel de pobreza) se concentran alrededor de 80%.
- Los porcentajes de competencia en matemática van desde 0 a 100
- La distribución de puntos no muestra que haya correlación que sugiera que, al aumentar los niveles de pobreza, baje la competencia en matemáticas

AÑADIR ETIQUETAS EN LOS EJES:

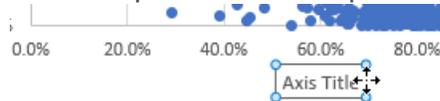
- Para añadir etiquetas a los ejes horizontal X y vertical Y, **mantenga la gráfica activa** y haga **click** en el botón **Add Chart Elements**



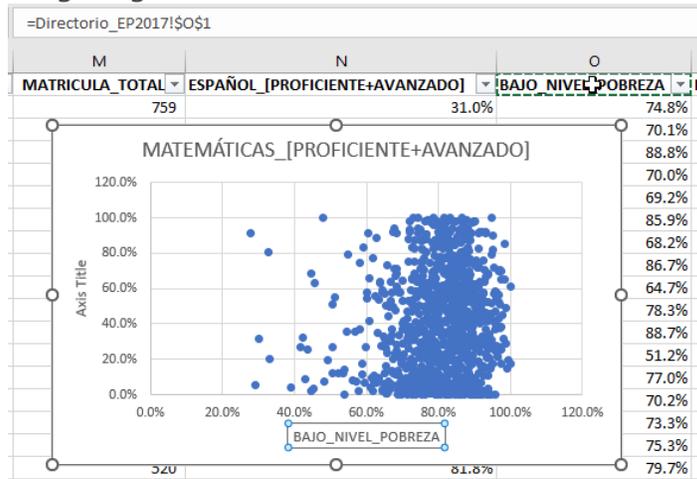
- Haga **check** en la opción **Axis Titles**.



- Cuando aparezcan las etiquetas en los ejes, haga **click** en la etiqueta del eje X

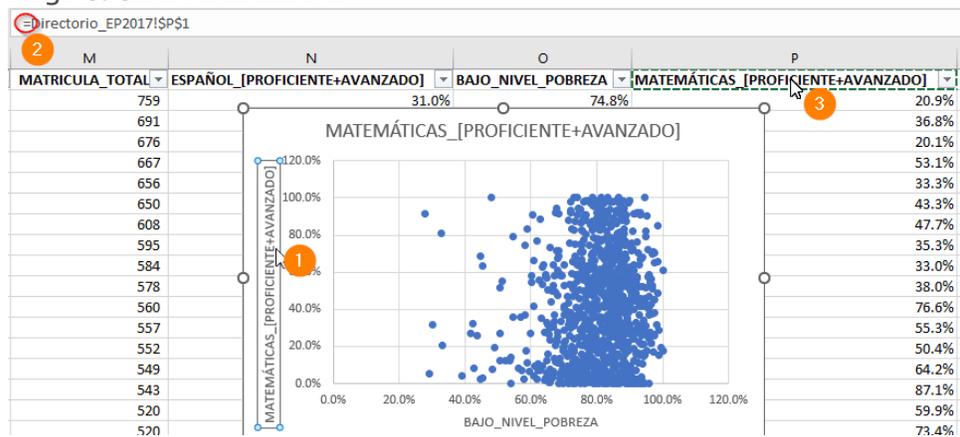


- Vaya a la caja de texto **Formula Bar**, y escriba =
Luego haga **click** en la **celda O1**.



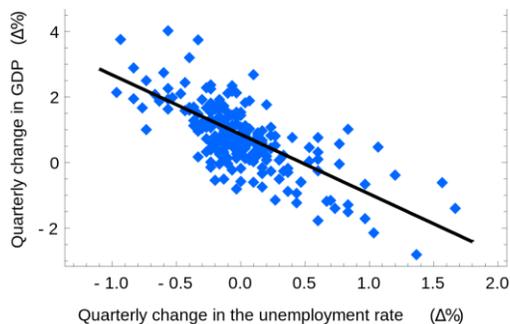
Notará que la etiqueta adquiere el texto que aparece en la celda O1.
Repitamos el proceso para la etiqueta del eje Y.

- Haga **click** en la etiqueta del eje Y.
- Escriba = en la caja de texto **Formula Bar**.
- Haga **click** en la celda **P1**.



AÑADIR LÍNEA DE TENDENCIA

Visualmente podemos constatar en la gráfica anterior que no hay correlación significativa entre el nivel de competencia en matemáticas (en el conjunto de todas las escuelas públicas) y el porcentaje bajo el nivel de pobreza.



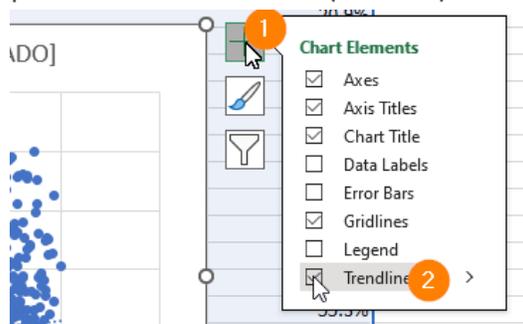
Si hubiese algún tipo de correlación, podríamos ver que los puntos se concentrarían en forma diagonal hacia arriba o hacia abajo (correlación positiva o negativa).

En esta gráfica, la correlación es negativa. Mayor desempleo se correlaciona con menor producto interno bruto.

Fuente: Wikipedia, [Coeficiente de determinación](#), accedido en octubre 20, 2023.

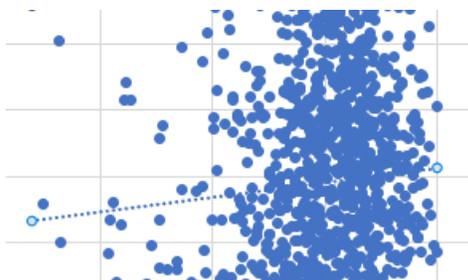
Podemos mostrar una línea de tendencia para sustentar esto y además añadir la estadística [coeficiente de determinación \$R^2\$](#) . Los posibles valores de R^2 van de 0 a 1. Mientras más se aproxime R^2 a 1, mayor será la correlación entre ambas variables.

- Para añadir una línea de tendencia a esta gráfica, haga **click** en la gráfica para activarla. Le aparecerán tres botones (cuadros)



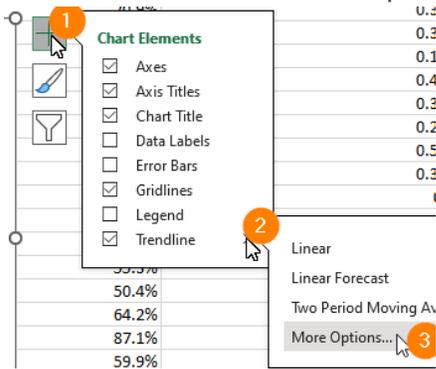
- Haga **click** en el botón **Chart Elements**.
- Haga **check** en la opción **Trendline**.

Aparecerá la línea de tendencia, pero se confunde con el color de los puntos porque son del mismo color.



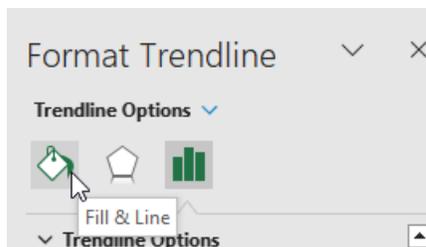
CAMBIAR EL COLOR DE LA LÍNEA DE TENDENCIA

Cambiaremos el color de esta línea para que sobresalga un poco en relación con la nube de puntos.

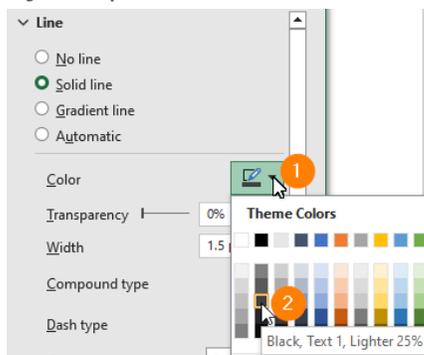


- Haga **click** en el botón **Chart Elements**
- Al final de la lista de opciones haga **click** en el **triángulo** de la opción **Trendline**.
- Haga **click** en la opción **More Options...**

- En la forma **Format Trendline** que aparecerá, bajo la sección **Trendline Options**, haga **click** en el botón **Fill & Line**

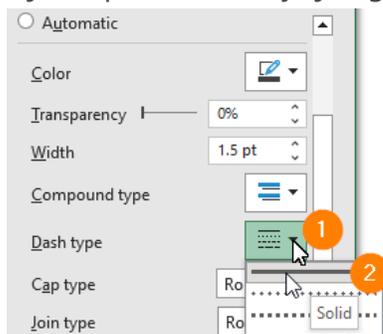


- Vaya al apartado **Color**



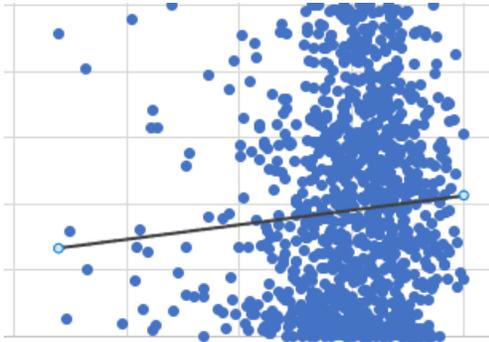
- Haga **click** en el botón **Outline Color**
- Haga **click** en el color **Black, Text 1, Lighter 25%**.

- Vaya un poco más abajo y haga **click** en el botón **Dash Type**



- Escoja la primera opción **"Solid"**

La línea deberá aparecer como esta:

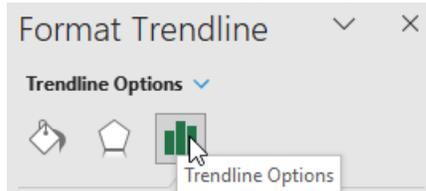


Ahora que la línea llama la atención, pasemos a la próxima sección.

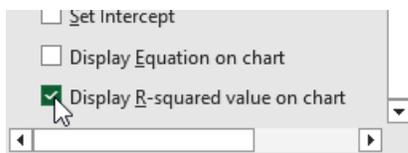
AÑADIR ETIQUETA CON EL COEFICIENTE DE DETERMINACIÓN R^2

Este valor R^2 se usará para conocer cuán fuerte o débil es la correlación entre el nivel de pobreza y el desempeño en matemáticas.

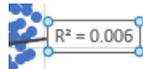
- En la forma **Format Trendline** haga **click** en el botón **Trendline Options**.



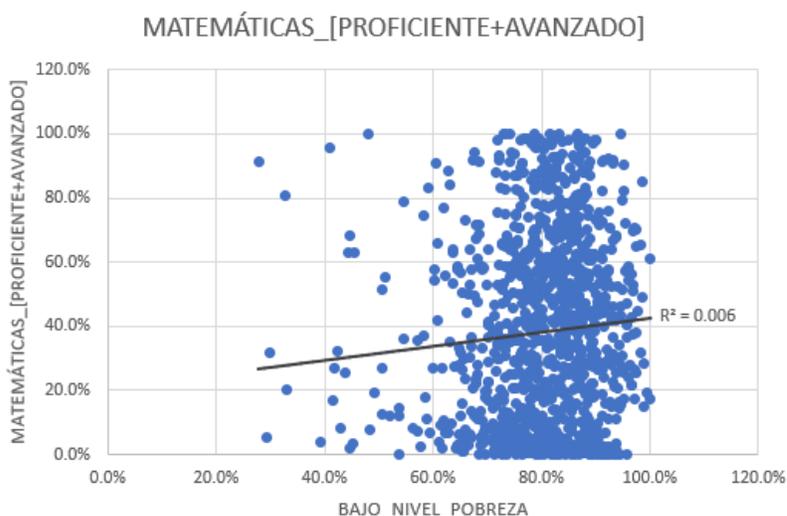
- Navegue hasta el final de la lista y haga **check** en la opción **Display R -squared value on chart**.



Aparecerá el valor R^2 en la gráfica. Muévelo de ser necesario.



El valor $R^2 = 0.0062$ muestra que **no hay correlación** significativa entre los datos de las dos columnas.

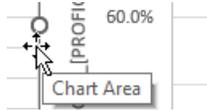


FIJAR LA GRÁFICA

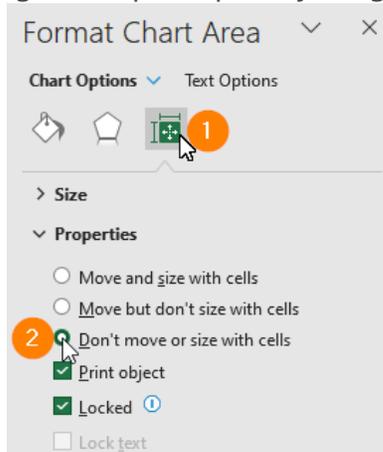
En ocasiones es buena idea fijar la gráfica para que no se mueva o “desaparezca”. Más adelante vamos a segmentar las escuelas por nivel, ya sea superior, intermedio, elemental, para ver si hay

correlación positiva o negativa entre el nivel de pobreza y el desempeño en matemáticas en estos distintos segmentos.

- Para fijar la gráfica, manténgala activada (**click** en el borde).



- Siga estos pasos para fijar la gráfica:



- En la forma **Format Chart Area**, haga **click** en el botón **Size & Properties**.
- Luego **escoja** la opción **Don't move or size with cells**.

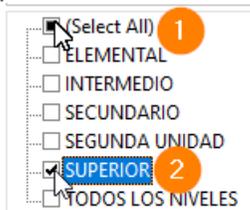
APLICAR FILTROS

Aplicaremos filtros para segmentar/distinguir/separar/discriminar/seleccionar las escuelas por nivel educativo (NIVEL_ORIGINAL). Veremos los resultados en la hoja y en la gráfica.

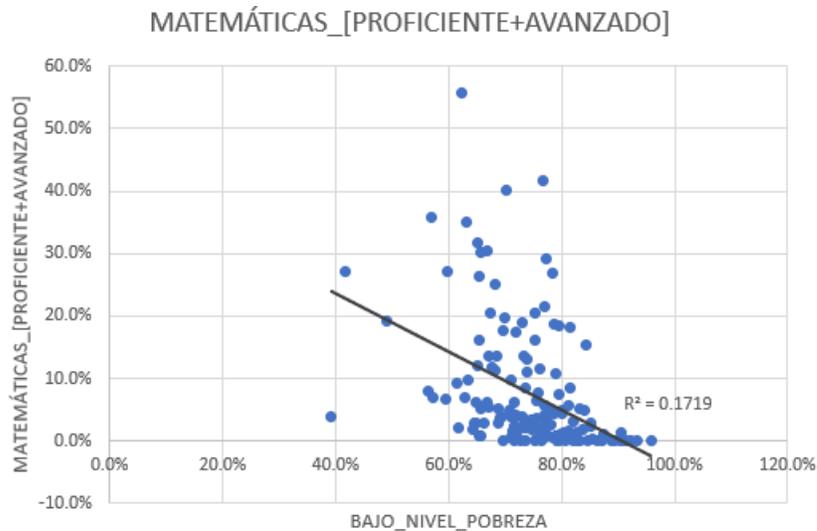
- Bajo la columna **K (NIVEL_ORIGINAL)**, haga **click** en el **botón-flecha de filtros y Sorting**. Verá que este lee **(Showing All)**.

K	L
NIVEL_ORIGINAL	GRADOS_POS
ELEMENTAL	
ELEMENTAL	NIVEL_ORIGINAL: (Showing All)
ELEMENTAL	

- Haga **uncheck** primero en la opción **Select All**. Luego haga **check** en la opción **SUPERIOR**, para seleccionar solamente las escuelas de nivel superior.



Notará que la línea de tendencia y el valor R^2 cambian cuando los datos son solamente para escuelas de nivel superior:



En el caso de escuelas a nivel superior, el valor R^2 aumenta, pero la correlación es débil; es menor que 0.5.

- Repita el proceso** escogiendo solo **escuelas intermedias**, luego de **segunda unidad**, y finalmente las **elementales**.
- Recuerde** que debe quitar el filtro anterior.

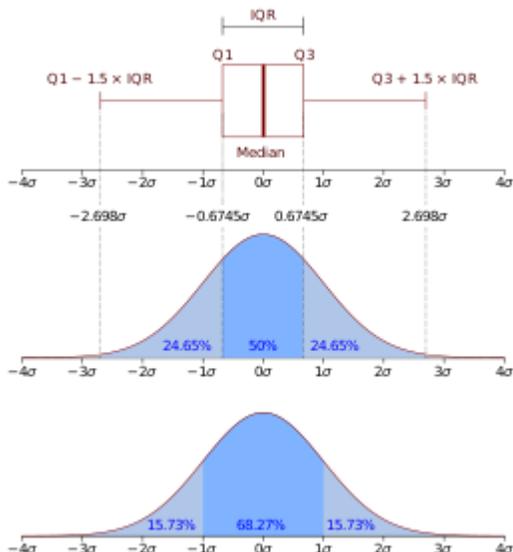
¿Ve alguna diferencia?, ¿Cree que los resultados pueden ser diferentes usando otros filtros?, ¿Por municipio, zona?

- Seleccione la gráfica por el borde y bórrela
- No cierre este workbook porque lo vamos a usar para la próxima gráfica

EJERCICIO 5: GRÁFICA BOXPLOT

Los gráficos de caja (**boxplot and whiskers**) sirven para **resumir mucha información** sobre la distribución de una o múltiples series de datos. Son pequeños resúmenes de estas distribuciones numéricas, dispuestas de manera conveniente y ocupan poco espacio.

En un solo diagrama se muestra:



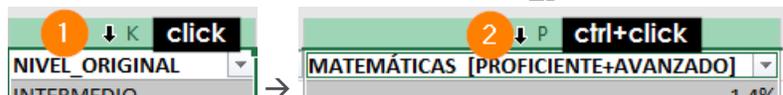
- **Q1: el primer cuartil** (25% de los datos)
- La **mediana**: 50% de los datos (**2do cuartil**)
- **Q3: el tercer cuartil** (75%)
- Los **whiskers** representan las distancias de las líneas fuera de la caja, $Q1 - 1.5 \times IQR$ y $Q3 + 1.5 \times IQR$.
- Los **datos atípicos (outliers)**. Se grafican más allá de las líneas whiskers que están fuera de la caja.
- Si las **cajas** son **alargadas** significa que hay **dispersión** en la distribución de datos.
- En Excel se muestra el promedio (**media**) usando una **x**.

La figura muestra cómo corresponden las longitudes de los whiskers y la caja con respecto a una distribución normal ideal.

Fuente: [Wikipedia](#)

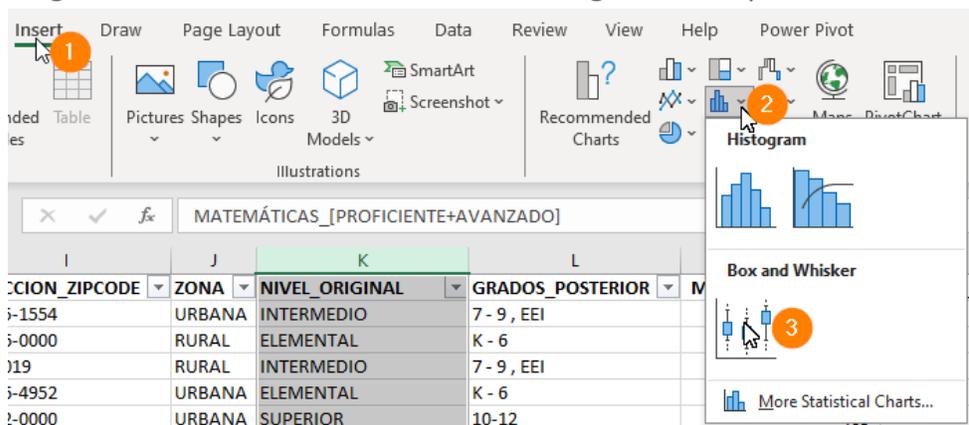
En esta práctica **vamos a usar los diagramas de caja** para mostrar **cómo varía el desempeño en matemáticas** por niveles escolares tales como elemental, intermedio y superior. Hay otros niveles, pero son menos escuelas en esos niveles denotados como segunda unidad o secundario. Vamos a mostrar separadamente en cajas los datos de desempeño para escuela elemental, intermedia y secundaria/superior.

- Utilice el workbook **DirectorioEscuelasPublicas2017.xlsx**, localizado en el folder **C:\Tutorial_Excel\DataAnalysis\Modulo_1**
- Haga **click** en la cabecera de la columna **K: NIVEL_ORIGINAL**, luego haga **ctrl+click** en la cabecera de la columna **P Matemáticas_[Proficiente+Avanzado]**.

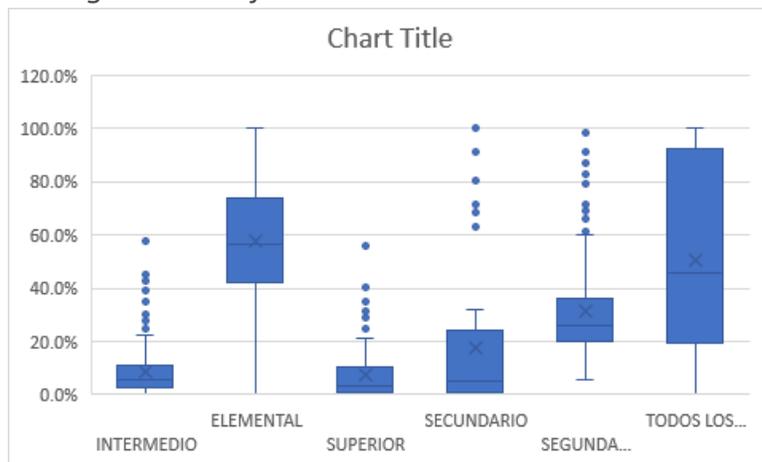


La **columna K** contiene los **datos nominales** (categorías, labels: **eje horizontal X**) mientras que la **columna P** contiene los **porcentajes (datos numéricos: eje vertical Y)** de desempeño en matemáticas.

- Una vez seleccionadas estas dos columnas, vaya al menú principal, **1:** haga **click** en **Insert**, luego **2:** en el botón **Statistical Charts**. Luego **3:** en la opción **Box and Whisker**.



Aparecerá la gráfica de cajas.



- Vamos a cambiar el título de la gráfica. Haga **doble click** encima del label **Chart Title** y cuando esté disponible para modificarlo, escriba **Desempeño en Matemáticas**.



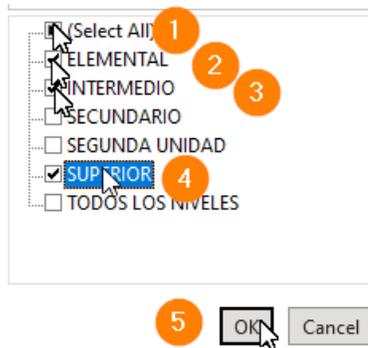
APLICAR FILTRO PARA SELECCIONAR NIVELES DE INTERÉS

Como podemos apreciar en la gráfica de cajas, en el eje horizontal están todos los niveles. Los que nos interesa comparar son: elemental, intermedio y superior. Como mencionamos antes, los otros niveles tienen menos puntos, y pueden tener mucha variabilidad.

- Vaya a la **columna K: NIVEL_ORIGINAL** y haga **click** en el botón de **filtro**.



- En la lista de niveles que aparecerá:



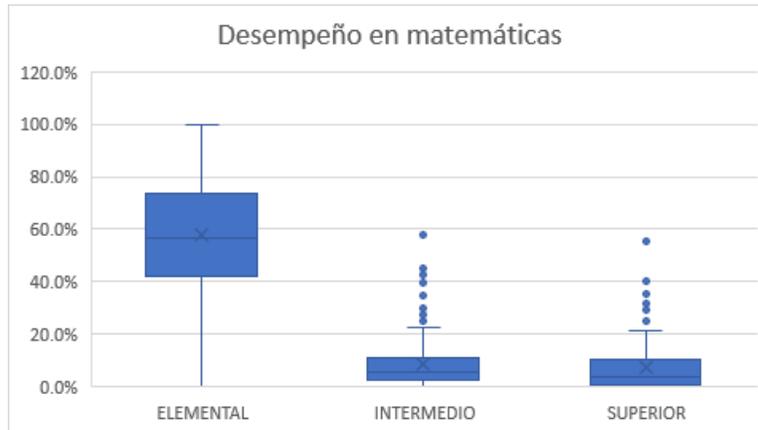
- Haga uncheck primero en la opción **Select All**
- Luego escoja los niveles **elemental, intermedio y superior**;
- Termine haciendo **click** en el botón **OK**

- Ordene las categorías alfabéticamente. Por suerte, el orden alfabético, coincide con la secuencia de niveles.



- Haga **click** en el **botón de filtro** de la **columna K: NIVEL_ORIGINAL** .
- Escoja** la opción **Sort A to Z** (ascendente).

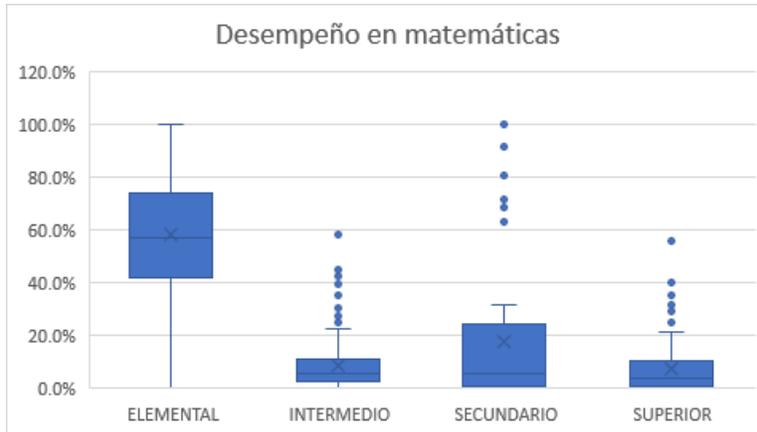
- Dedique un momento a observar esta gráfica:



- ¿Qué está mostrando la gráfica? ¿Cómo va cambiando el desempeño en matemáticas a medida que va avanzando el nivel escolar?

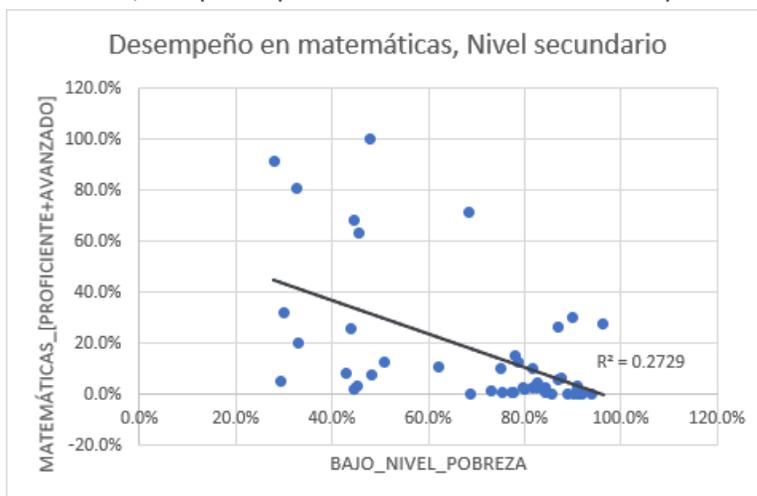
Se debe recordar que **las destrezas matemáticas** se constituyen sobre **fundamentos** elementales que van haciéndose más complejos a medida que se avanza de nivel escolar. Si falta alguna destreza fundamental o no es bien comprendida, el desempeño se verá afectado negativamente en el futuro. Parece haber una brecha desde el nivel elemental al intermedio, que afecta negativamente al próximo nivel. Lo que vemos como outliers en las distribuciones intermedio y superior en realidad no son alentadores para la población, sino para algunas escuelas “exitosas”.

Si añadimos el nivel secundario, la gráfica se ve así:



En esta "caja" que solo contiene **44 escuelas públicas** están **algunas de mejor desempeño en matemáticas** como CROEM, University Gardens. A CROEM solo van estudiantes avanzados; el caso de la escuela University Gardens es parecido, con estudiantes avanzados, algunos con mejores recursos económicos. Estas dos son escuelas especializadas en ciencias y matemáticas, pero también hay otras que tienen cero en desempeño en matemáticas.

Estas dos escuelas con porcentajes bajos en nivel de pobreza coinciden con las de mejor desempeño. Pero no se debe generalizar porque si retomamos la gráfica de dispersión, hay escuelas con porcentajes más bajos en nivel de pobreza, pero con porcentajes bajos en desempeño en matemáticas, así que el problema no es solamente la pobreza.



EJERCICIO 6: GRÁFICA TIPO RADAR

En esta parte vamos a realizar una **gráfica radar** usando datos **sobre fármacos y productos naturales que causan dependencia**. Tenemos una lista de **compuestos de síntesis química y de productos naturales** de los cuales **se estimaron niveles de daño en aspectos** tales como **daño físico, dependencia y daño social**. Esta tabla proviene de un **estudio publicado** en la revista *The Lancet*, vol. 369, marzo 24, 2007, por David Nutt, Leslie A King, William Saulsbury, Colin Blakemore de la Universidad de Bristol, Inglaterra.

Una **gráfica radar** nos da la ventaja de ver muchas categorías y series de datos simultáneamente. Sin embargo, **las series de datos deben estar en la misma escala o normalizadas** para que estas se puedan comparar y visualizar.

- Abra el workbook llamado **rational_scale_drugs.xlsx**, localizado en el folder **C:\Tutorial_Excel\DataAnalysis\Modulo_1**

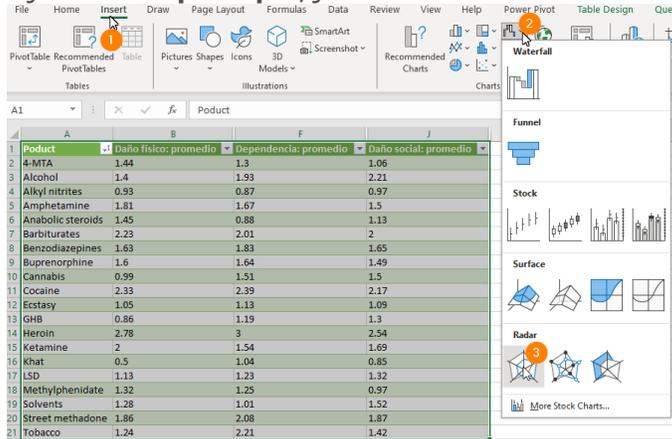
En este workbook tenemos una Tabla Excel; no es un data range. Por tal razón, notamos los colores y los botones de filtro en cada columna, distintivos para las Excel Tables. Discutiremos las Excel Tables en el próximo módulo.

	A	B	C	
1	Product	Daño físico: promedio	Physical Harm: Acute	Physical H
2	4-MTA	1.44	2.2	2.1
3	Alcohol	1.4	1.9	2.4
4	Alkyl nitrites	0.93	1.6	0.9
5	Amphetamine	1.81	1.3	1.8
6	Anabolic steroids	1.45	0.8	2
7	Barbiturates	2.23	2.3	1.9
8	Benzodiazepines	1.63	1.5	1.7
9	Buprenorphine	1.6	1.2	1.3
10	Cannabis	0.99	0.9	2.1
11	Cocaine	2.33	2	2
12	Ecstasy	1.05	1.6	1.6
13	GHB	0.86	1.4	1.2
14	Heroin	2.78	2.8	2.5
15	Ketamine	2	2.1	1.7
16	Khat	0.5	0.3	1.2
17	LSD	1.13	1.7	1.4
18	Methylphenidate	1.32	1.2	1.3
19	Solvents	1.28	2.1	1.7
20	Street methadone	1.86	2.5	1.7
21	Tobacco	1.24	0.9	2.9

- Escoja todas las celdas de las columnas Product, Daño físico: promedio, Daño Social: promedio:**

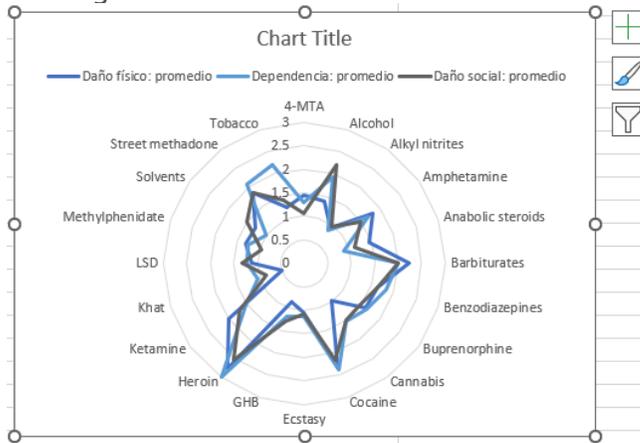
	A	B	F	J
1	Product	Daño físico: promedio	Dependencia: promedio	Daño social: promedio
2	4-MTA	1.44	1.3	1.06
3	Alcohol	1.4	1.93	2.21
4	Alkyl nitrites	0.93	0.87	0.97
5	Amphetamine	1.81	1.67	1.5
6	Anabolic steroids	1.45	0.88	1.13
7	Barbiturates	2.23	2.01	2
8	Benzodiazepines	1.63	1.83	1.65
9	Buprenorphine	1.6	1.64	1.49
10	Cannabis	0.99	1.51	1.5
11	Cocaine	2.33	2.39	2.17
12	Ecstasy	1.05	1.13	1.09
13	GHB	0.86	1.19	1.3
14	Heroin	2.78	3	2.54
15	Ketamine	2	1.54	1.69
16	Khat	0.5	1.04	0.85
17	LSD	1.13	1.23	1.32
18	Methylphenidate	1.32	1.25	0.97
19	Solvents	1.28	1.01	1.52
20	Street methadone	1.86	2.08	1.87
21	Tobacco	1.24	2.21	1.42

- Vaya al **menú principal**, y:



- **1:** Escoja **Insert**,
- **2:** Luego **click** en el botón **Insert waterfall...** radar chart,
- **3:** **Click** en la primera opción de izquierda a derecha: **Radar**

Se mostrará la gráfica:



De primera intención esta gráfica puede ser confusa porque estamos más acostumbrados a ver otros tipos de gráficas como las de barras y pie charts. Sin embargo, podemos ver que hay compuestos que resaltan.

Vamos a hacer algunos cambios y añadirle algunos elementos.

FORMATEAR TEXTO DEL EJE Y; AÑADIR LÍNEAS AL EJE Y

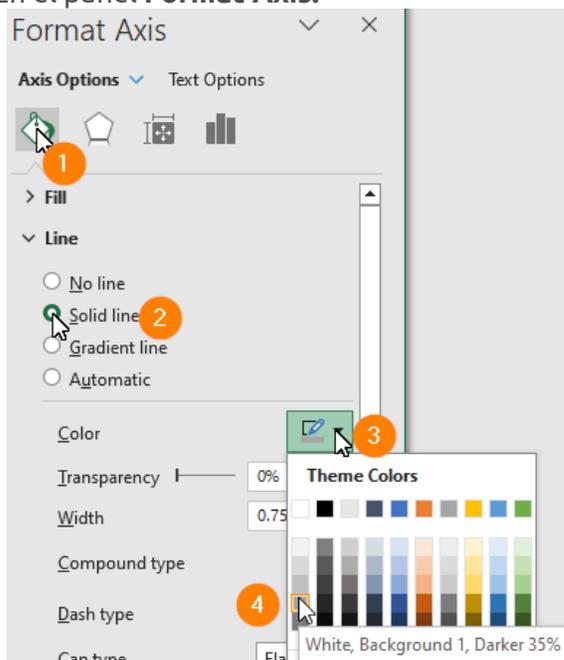
En esta sección haremos resaltar los números del eje Y, los cuales van del 0 al 3 y posteriormente añadiremos las líneas que corresponden a cada tipo de "droga" en la gráfica.

- Haga **doble click** encima del eje Y (encima del número 2.5):



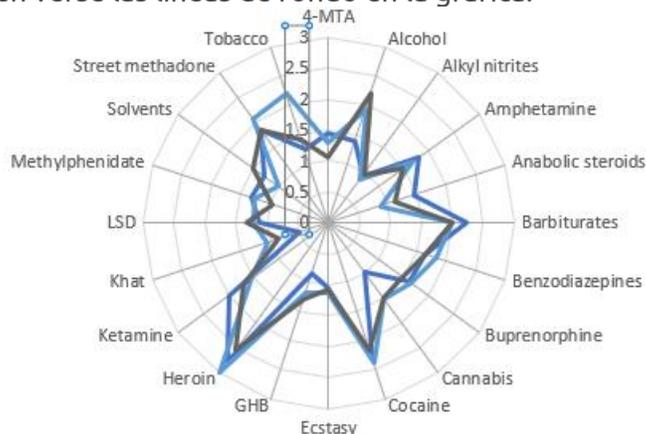
Excel traerá en pantalla el panel **Format Axis**. Vamos a hacer cambios en el eje de Y.

- En el panel **Format Axis**:



- **1:** Haga **click** en el botón **Fill & Line**
- **2:** Bajo el apartado **Line**, escoja la opción **Solid line**
- **3:** **Expanda** el botón **Outline Color**
- **4:** Escoja el **segundo color gris de abajo hacia arriba** de la lista (**White, Background 1, Darker 35%**) para que estas líneas sean visibles, pero no distraigan la atención sobre las líneas principales, que son las de las series de datos.

Así deben verse las líneas de fondo en la gráfica:



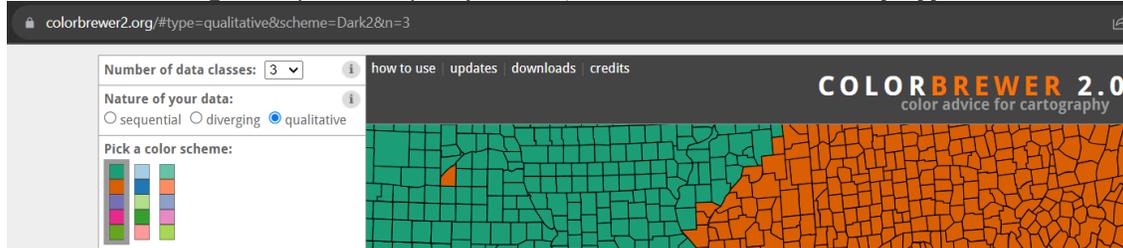
- Mantenga el eje activado. Vamos a cambiar el texto de la escala a negrillas (bold).
- Vaya al **menú principal** y haga **click** en **Home**. Haga **click** en el botón **Bold**.



USAR HERRAMIENTA COLORBREWER® PARA DIFERENCIAR LAS SERIES

Los colores que aparecen por defecto en Excel no necesariamente ayudan a diferenciar las series de datos. En ocasiones se debe recurrir a herramientas para hacerlo o desarrollar esta destreza con la práctica. La herramienta [ColorBrewer](https://colorbrewer2.org) nos puede ayudar a elegir secuencias de color efectivas. Entre las ventajas de esta herramienta está la opción de escoger secuencias que sean seguras para [discromatopsia](#) (color blind). Además, podemos escoger secuencias que se ajusten mejor a impresoras (printer friendly). La herramienta fue diseñada para ayudar a cartógrafos, pero puede usarse en gráficas también.

- Usando el navegador (browser) disponible, escriba la dirección <https://colorbrewer2.org>



- Mantenga el número de **clases** o categorías en **3**.



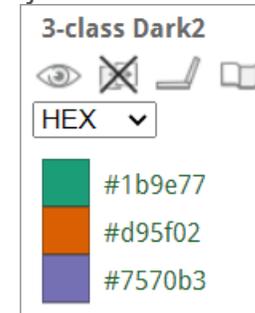
- En la sección **Nature of your data**, escoja la opción **qualitative**.



- Vaya a la sección **Only show**, y escoja las opciones **colorblind safe** y **print friendly**.



- Fijese en la sección **3-class Dark2**.

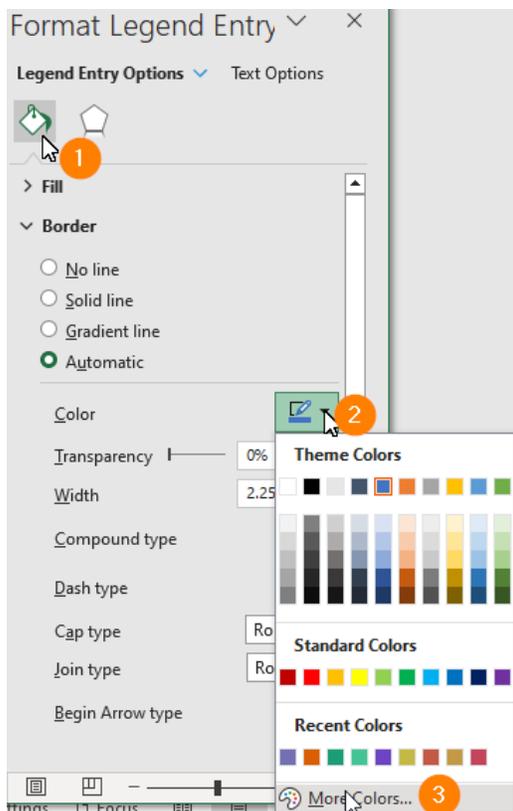


Estos son los colores que vamos a usar en la gráfica radar en Excel para distinguir las series de datos.

- Regresemos a Excel. Haga doble **click** en la **raya azul** del ítem **Daño físico: promedio** de la leyenda. Su selección tiene que verse así para poder cambiar solamente un color a la vez

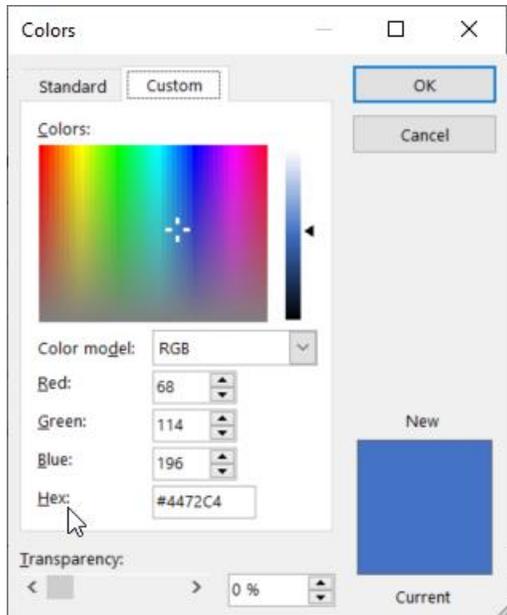


- En el panel **Format Legend Entry** que aparecerá al lado derecho de la interfaz de Excel, **siga esta secuencia de pasos**.



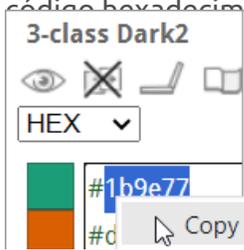
- 1: Haga **click** en el botón **Fill & Line**
- 2: En la sección **Color**, haga **click** en el botón **Outline color**.
- 3: Haga **click** en la opción **More Colors...** para escribir el color que queremos.

Aparecerá la forma **Colors** para modificar el color de este ítem de la leyenda y el color de la serie datos.



En esta parte, vamos a **transferir** (copy-paste) los **códigos de color hexadecimales** desde la interfaz de **ColorBrewer** a los ítems de la leyenda en **Excel**.

- Regrese a la interfaz ColorBrewer en su navegador y copie **solo los números** del primer código hexadecimal. **Click, sombrear, right click, Copy (ctrl+c).**



NO incluya el símbolo #.

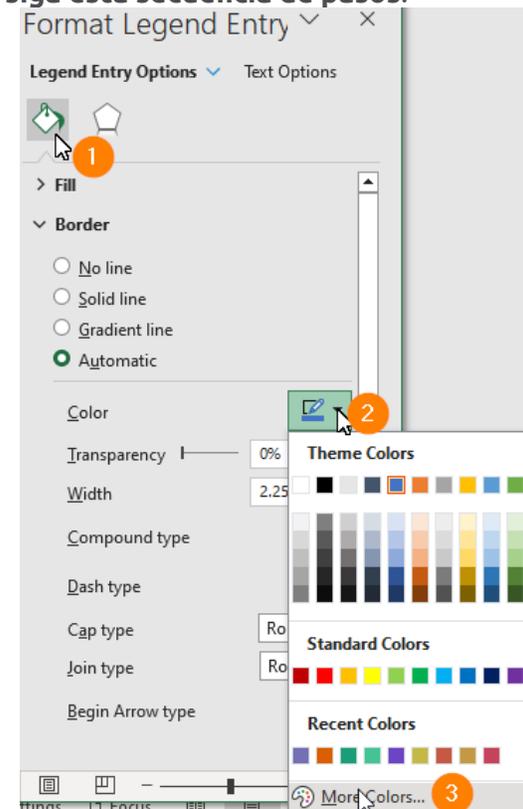
- Regrese a Excel y copie este código en la caja de texto **Hex** de la forma **Colors**



- Presione el botón **OK** para aceptar el cambio y cerrar la forma.
- Haga **click** en el segundo ítem de la leyenda, **Dependencia: promedio**

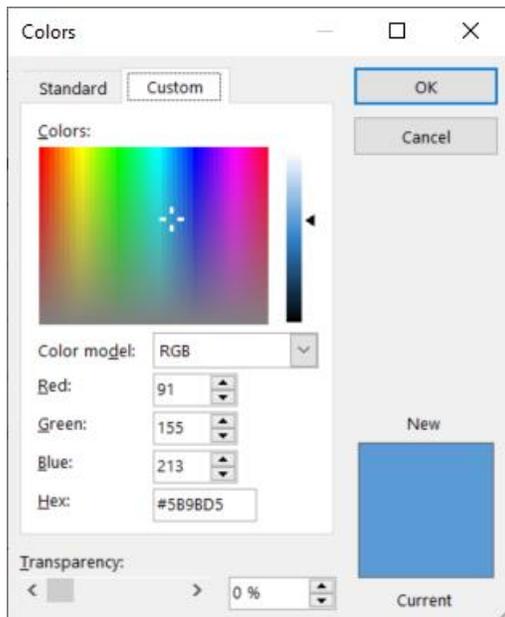


- En el panel **Format Legend Entry** que aparecerá al lado derecho de la interfaz de Excel, **siga esta secuencia de pasos.**

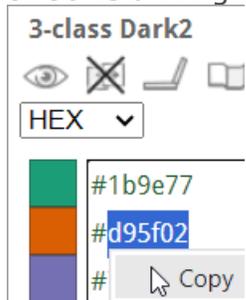


- **1:** Haga **click** en el botón **Fill & Line**
- **2:** En la sección **Color**, haga **click** en el botón **Outline color**.
- **3:** Haga **click** en la opción **More Colors...** para escribir el color que queremos.

Aparecerá la forma **Colors** para modificar el color de este ítem de la leyenda y el color de la serie datos.



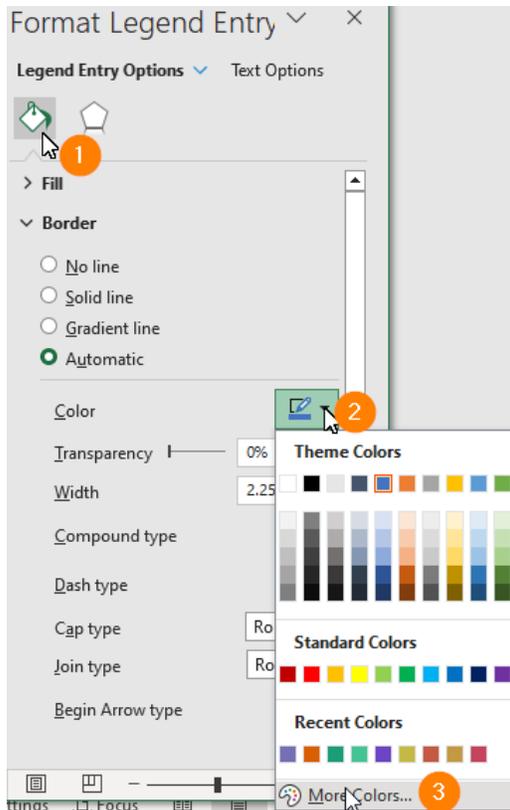
- Regrese a la interfaz **ColorBrewer** en su navegador y copie **solo los números** del **SEGUNDO** código hexadecimal. **Click, sombrear, right click, Copy (ctrl+c).**



- NO** incluya el símbolo #.
- Regrese a **Excel** y copie este código **d95f02** en la caja de texto **Hex** de la forma **Colors**
- Presione el botón **OK** para aceptar el cambio y cerrar la forma.

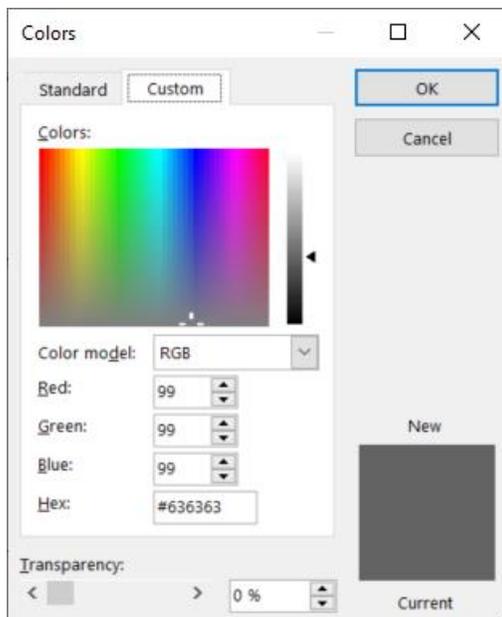
- Haga **click** en el **tercer** ítem de la leyenda, **Daño social: promedio**
- En el panel **Format Legend Entry** que aparecerá al lado derecho de la interfaz de Excel, **siga esta secuencia de pasos.**



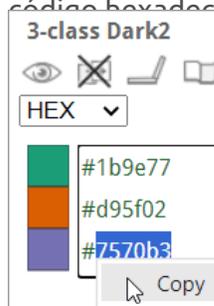


- 1:** Haga **click** en el botón **Fill & Line**
- 2:** En la sección **Color**, haga **click** en el botón **Outline color**.
- 3:** Haga **click** en la opción **More Colors...** para escribir el color que queremos.

Aparecerá la forma **Colors** para modificar el color de este ítem de la leyenda y el color de la serie datos.



- Regrese a la interfaz ColorBrewer en su navegador y copie solo los números del TERCER código hexadecimal. Click, sombrear, right click, Copy (ctrl+c).



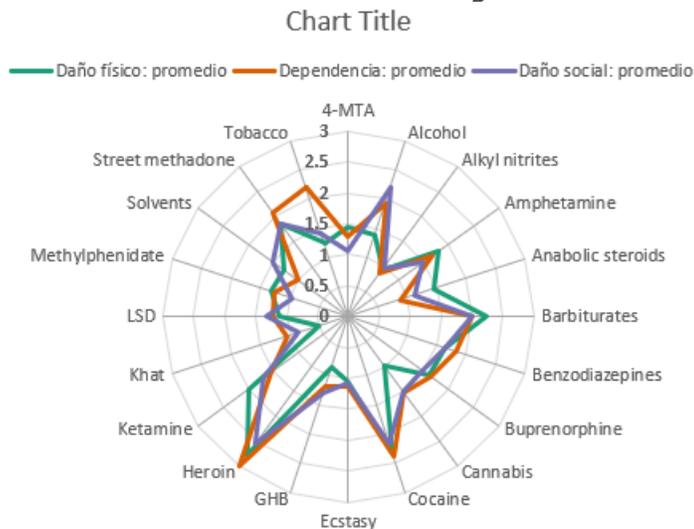
NO incluya el símbolo #.

- Regrese a Excel y copie este código 7570b3 en la caja de texto Hex de la forma Colors



- Presione el botón OK para aceptar el cambio y cerrar la forma.

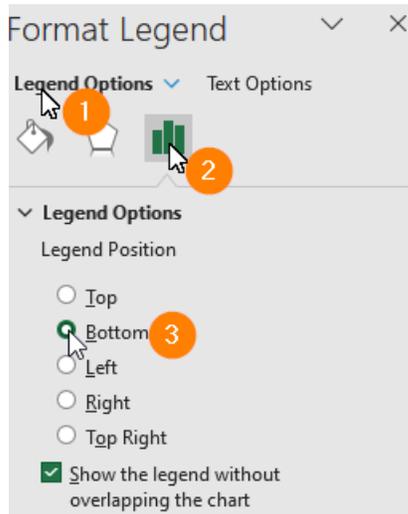
Así aparecerán los colores de las series en la gráfica:



MOVER LA LEYENDA AL FONDO DE LA GRÁFICA

Vamos a mover la leyenda al fondo para que no compita visualmente con el título de la gráfica.

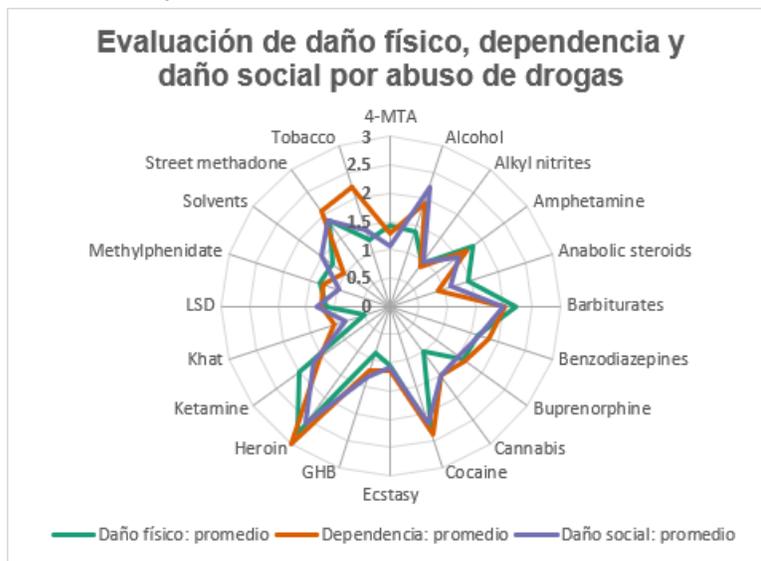
- Para mover la leyenda, siga estos pasos:



- 1: Haga **click** en la opción **Legend Options**
- 2: Haga **click** en el botón **Legend Options**.
- 3: En la sección **Legend Position**, haga **click** en la opción **Bottom**

CAMBIAR EL TÍTULO DE LA GRÁFICA

- Para cambiar el título de la gráfica, haga **doble click** en el título:
- Cambie el texto al siguiente:
Evaluación de daño físico, dependencia y daño social por abuso de drogas
- Cambie el tipo de letra a Arial, bold, tamaño 14.



ORDENAR LAS SERIES DE DATOS PARA HACER MÁS LEGIBLE LA GRÁFICA MEDIANTE CUSTOM SORT

Vamos a ordenar de menor a mayor todas las series de datos. Por defecto, el conjunto de datos está ordenado alfabéticamente por el nombre del producto o compuesto. Ordenar las series numéricas puede ayudar a evaluar más rápidamente las series.

- Primero haga **click** en *cualquiera* de las celdas **dentro de esta tabla**.

	A	
1	Product	Daño físico
2	4-MTA	1.44
3	Alcohol	1.4

- Vaya al **menú principal** y **click** en **Home**, **click** en **Sort and Filter**, luego **click** en **Custom Sort**.



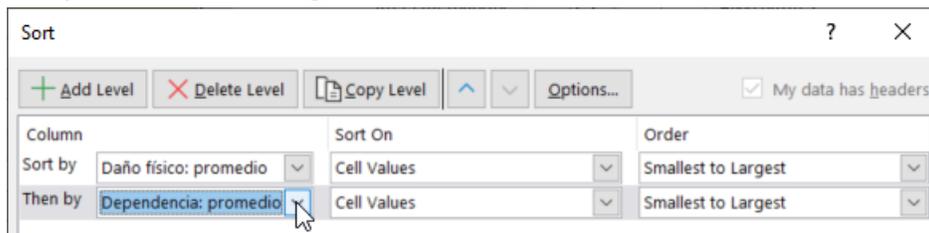
- En la forma **Sort** que aparecerá, vaya a la sección **Sort by**, y escoja la columna **Daño físico: promedio**. En la sección **Order**, escoja **Smallest to Largest**.



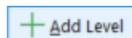
- Añada otro nivel. Haga **click** en el botón **Add Level**.



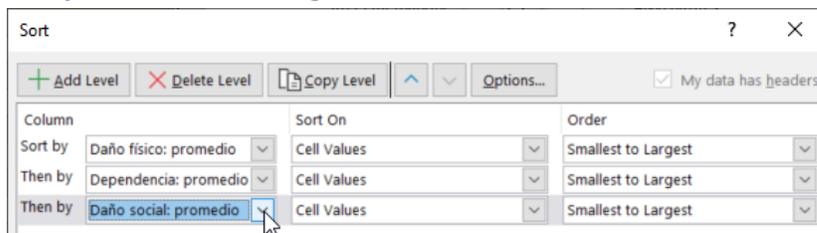
- En la sección **Then by**, escoja la columna **Dependencia: promedio**. En la sección **Order**, escoja **Smallest to Largest**.



- Añada el próximo nivel. Haga **click** en el botón **Add Level**.



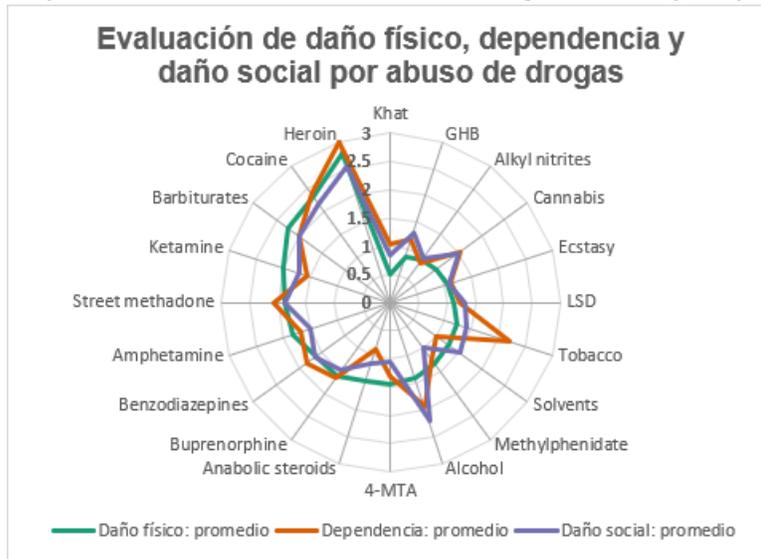
- En la sección **Then by**, escoja la columna **Daño social: promedio**. En la sección **Order**, escoja **Smallest to Largest**.



- Presione el botón **OK** para aceptar los cambios y **cerrar** la forma **Sort**.

De esta manera las series de datos quedan ordenadas y podemos ver mejor, cuáles son los compuestos más **dañosos** en las **categorías física, dependencia y social**. De esta lista, las **tres más perjudiciales** son la **heroína, cocaína** y los **barbitúricos**. No quiere esto decir que el **alcohol** y

el **tabaco** tengan también sus consecuencias. El **alcohol** tiene picos en **dependencia** y **daño social**, mientras que el **tabaco** tiene su máximo en **dependencia** y un poco menos en **daño social** y **físico**.



EJERCICIO 7: GRÁFICAS CON SERIES TEMPORALES

En esta parte vamos a hacer un par de gráficas sobre los temas: **brecha salarial** y **evolución de salarios** a varios niveles educativos desde los años 2010 hasta 2021 en EEUU y Puerto Rico. Los datos provienen de la interfaz del Censo data.census.gov y corresponden al Censo de 2010 y los datos que recopila el American Community Survey para cada año hasta 2021.

Los datos fueron extraídos y procesados previamente, así que lo que resta por hacer es producir las gráficas. Estas serán series temporales representadas con líneas donde los años estarán en el eje horizontal X y el salario en dólares estará en el eje Y.

- Para comenzar, abra el workbook llamado **gender_gap_us_pr2010_201.xlsx**, localizado en el folder **C:\Tutorial_Excel\DataAnalysis\Modulo_1**

Encontrará dos hojas con datos salariales para distintos niveles educativos (educational attainment) para los años entre 2010 a 2021. Una hoja con datos de EEUU y otra hoja con datos de Puerto Rico, con las mismas columnas y años. Estos datos fueron extraídos de la tabla censal: **B20004 MEDIAN EARNINGS IN THE PAST 12 MONTHS (IN 2021 INFLATION-ADJUSTED DOLLARS) BY SEX BY EDUCATIONAL ATTAINMENT FOR THE POPULATION 25 YEARS AND OVER**, para los años 2010 a 2021. Así que se trata de medianas de “salarios” en personas mayores de 25 años.

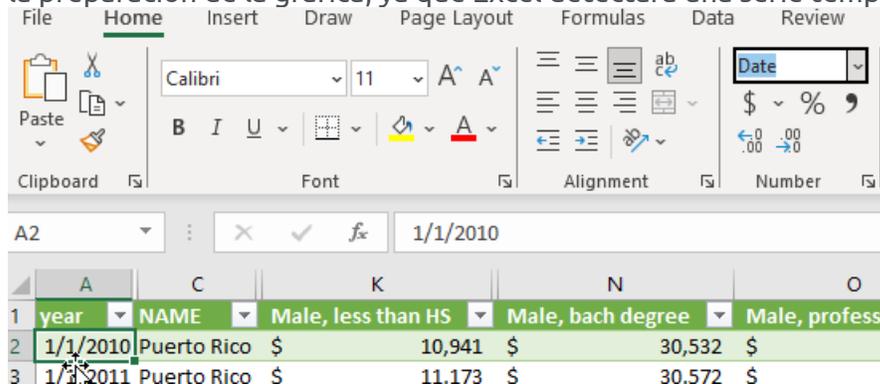
Estados Unidos:

year	NAME	Male, less than HS	Male, bach degree	Male, professional	Female, less than HS	Female, bach degree	Female, professional
1/1/2010	United States	\$ 22,792	\$ 59,922	\$ 80,111	\$ 14,645	\$ 40,313	\$ 53,023
1/1/2011	United States	\$ 23,002	\$ 60,951	\$ 82,216	\$ 14,966	\$ 41,046	\$ 54,539
1/1/2012	United States	\$ 22,728	\$ 61,224	\$ 83,141	\$ 15,052	\$ 41,602	\$ 55,344
1/1/2013	United States	\$ 22,676	\$ 61,250	\$ 83,609	\$ 15,235	\$ 41,814	\$ 55,768
1/1/2014	United States	\$ 23,104	\$ 61,619	\$ 84,137	\$ 15,369	\$ 41,917	\$ 56,185
1/1/2015	United States	\$ 23,668	\$ 61,589	\$ 84,006	\$ 15,510	\$ 41,763	\$ 56,181
1/1/2016	United States	\$ 24,644	\$ 62,242	\$ 85,127	\$ 15,831	\$ 42,154	\$ 56,859
1/1/2017	United States	\$ 25,707	\$ 63,911	\$ 87,504	\$ 16,431	\$ 43,390	\$ 58,868
1/1/2018	United States	\$ 26,856	\$ 65,803	\$ 90,475	\$ 17,146	\$ 45,039	\$ 60,912
1/1/2019	United States	\$ 28,189	\$ 67,947	\$ 92,576	\$ 18,032	\$ 46,624	\$ 62,631
1/1/2020	United States	\$ 30,083	\$ 69,326	\$ 93,923	\$ 19,011	\$ 47,709	\$ 63,781
1/1/2021	United States	\$ 31,722	\$ 73,128	\$ 98,840	\$ 20,448	\$ 50,554	\$ 67,202

Puerto Rico:

year	NAME	Male, less than HS	Male, bach degree	Male, professional	Female, less than HS	Female, bach degree	Female, professional
1/1/2010	Puerto Rico	\$ 10,941	\$ 30,532	\$ 46,575	\$ 9,001	\$ 23,507	\$ 32,602
1/1/2011	Puerto Rico	\$ 11,173	\$ 30,572	\$ 45,464	\$ 9,211	\$ 23,985	\$ 33,460
1/1/2012	Puerto Rico	\$ 11,250	\$ 30,687	\$ 45,697	\$ 9,451	\$ 24,026	\$ 33,133
1/1/2013	Puerto Rico	\$ 11,327	\$ 30,669	\$ 47,299	\$ 9,601	\$ 24,077	\$ 33,599
1/1/2014	Puerto Rico	\$ 11,493	\$ 30,771	\$ 45,994	\$ 9,596	\$ 24,189	\$ 33,458
1/1/2015	Puerto Rico	\$ 11,204	\$ 30,551	\$ 44,897	\$ 9,523	\$ 23,619	\$ 32,855
1/1/2016	Puerto Rico	\$ 11,174	\$ 30,165	\$ 45,529	\$ 9,581	\$ 23,598	\$ 32,628
1/1/2017	Puerto Rico	\$ 11,383	\$ 29,707	\$ 45,826	\$ 9,456	\$ 23,753	\$ 33,017
1/1/2018	Puerto Rico	\$ 11,534	\$ 30,079	\$ 44,335	\$ 9,416	\$ 23,805	\$ 32,477
1/1/2019	Puerto Rico	\$ 11,658	\$ 30,122	\$ 45,406	\$ 9,292	\$ 23,622	\$ 32,365
1/1/2020	Puerto Rico	\$ 11,713	\$ 30,000	\$ 44,156	\$ 9,068	\$ 23,706	\$ 32,203
1/1/2021	Puerto Rico	\$ 12,160	\$ 30,281	\$ 43,751	\$ 8,770	\$ 24,124	\$ 32,527

- Note que las columnas/campos **year**, de ambas hojas están en formato **date**. Esto optimiza la preparación de la gráfica, ya que Excel detectará una serie temporal.



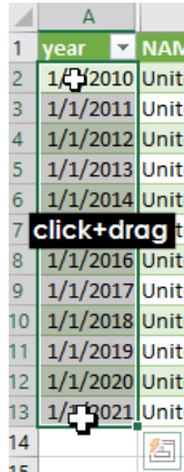
The screenshot shows the Excel Home ribbon with the 'Number' group expanded to show the 'Date' dropdown menu. Below the ribbon, the formula bar shows the date '1/1/2010' and the spreadsheet grid shows the corresponding data for Puerto Rico in 2010.

year	NAME	Male, less than HS	Male, bach degree	Male, profess
1/1/2010	Puerto Rico	\$ 10,941	\$ 30,532	\$
1/1/2011	Puerto Rico	\$ 11,173	\$ 30,572	\$

AJUSTAR EL FORMATO DATE PARA MOSTRAR AÑO

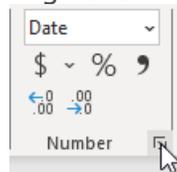
Vamos a ajustar los datos tipo fecha, a una **forma reducida que muestre solo el año**. Esto no cambiará el tipo de datos. **Al hacer esto, no tendremos que hacer cambios en el eje de X cuando hagamos la gráfica.**

- Vaya a la hoja **gender_pay_gap_us** y **sombree (click+drag)** las celdas **A2** hasta la celda **A13**.

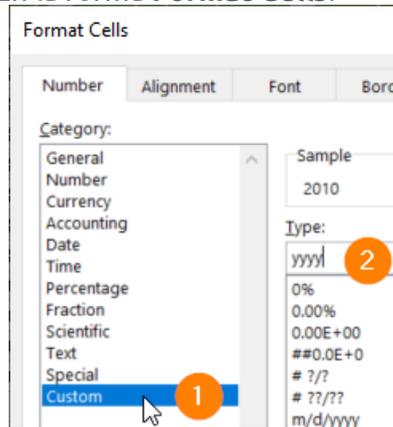


	A	
1	year	NAM
2	1/1/2010	Unite
3	1/1/2011	Unite
4	1/1/2012	Unite
5	1/1/2013	Unite
6	1/1/2014	Unite
7	1/1/2016	Unite
8	1/1/2017	Unite
9	1/1/2018	Unite
10	1/1/2019	Unite
11	1/1/2020	Unite
12	1/1/2021	Unite
13		
14		
15		

- Una vez estén sombreadas las celdas, vaya al **menú principal**, haga **click** en **Home**, vaya a la **sección Number** y haga **click** en el botón **Number Format**

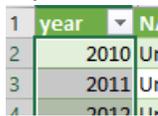


- En la forma **Format Cells**:



- 1: Haga **click** en la opción **Custom**.
- 2: En la caja de texto bajo **Type**: escriba **yyyy**
- Haga **click** en el botón **OK** para aceptar los cambios y cerrar la forma Format Cells.

La fecha aparecerá solo con el año:



	year	N
2	2010	Ur
3	2011	Ur
4	2013	Ur

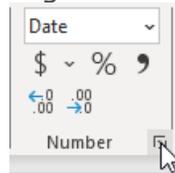
Esto se debe repetir en la hoja **gender_pay_gap_pr**.

- Vaya a la hoja **gender_pay_gap_pr** y **sombree (click+drag)** las celdas **A2** hasta la celda **A13**.

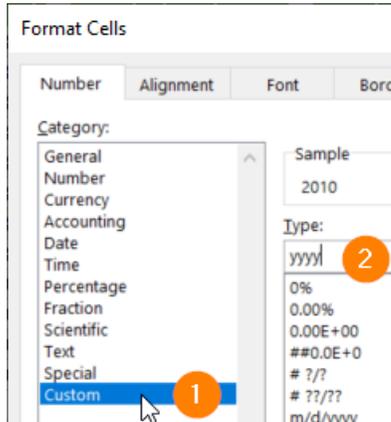
	A	
1	year	NA
2	1/1/2010	Pue
3	1/1/2011	Pue
4	1/1/2012	Pue
5	1/1/2013	Pue
6	1/1/2014	Pue
8	1/1/2016	Pue
9	1/1/2017	Pue
10	1/1/2018	Pue
11	1/1/2019	Pue
12	1/1/2020	Pue
13	1/1/2021	Pue

click+drag

- Una vez estén sombreadas las celdas, vaya al **menú principal**, haga **click** en **Home**, vaya a la **sección Number** y haga **click** en el botón **Number Format**



- En la forma **Format Cells**:



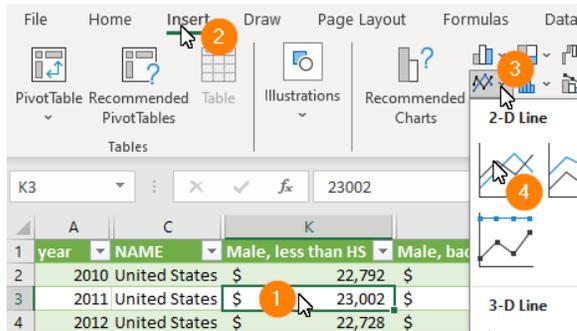
- 1: Haga **click** en la opción **Custom**.
- 2: En la caja de texto bajo **Type**: escriba **yyyy**
- Haga **click** en el botón **OK** para aceptar los cambios y cerrar la forma **Format Cells**.

La fecha aparecerá solo con el año:

	A	
1	year	NA
2	2010	Pue
3	2011	Pue
4	2012	Pue

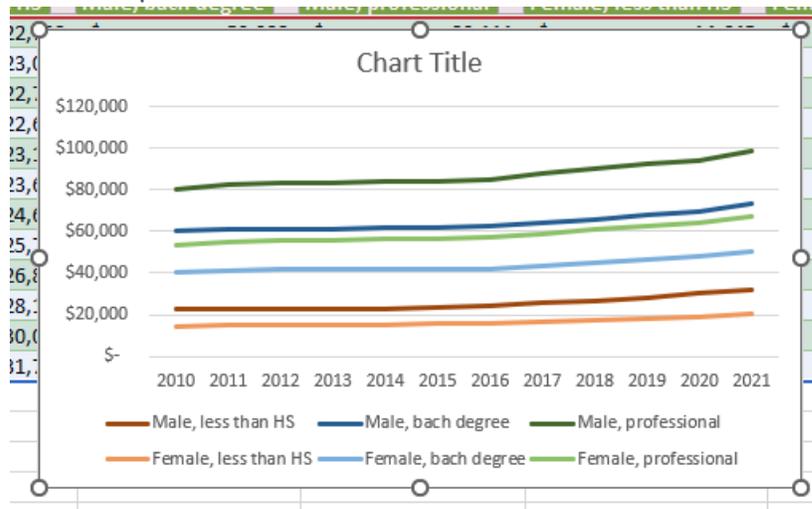
- Regrese a la hoja **gender_pay_gap_us**.

- Haga **click** en **cualquiera** de las celdas **dentro** de la **tabla**



- Vaya al menú principal y escoja **Insert**.
- Vaya a la sección **Charts** y haga **click** en el botón **Line Charts**.
- Luego **click** en la **primera opción** de **2-D Line**.

La gráfica debe aparecer más o menos de esta manera:

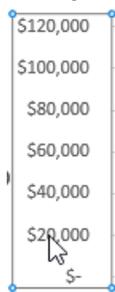


Además de esta gráfica haremos otra con los datos de Puerto Rico en los mismos años. Por ahora, sigamos trabajando esta gráfica con datos de EEUU.

AJUSTAR LA ESCALA DEL EJE Y: MEDIANA DE SALARIO EN DÓLARES

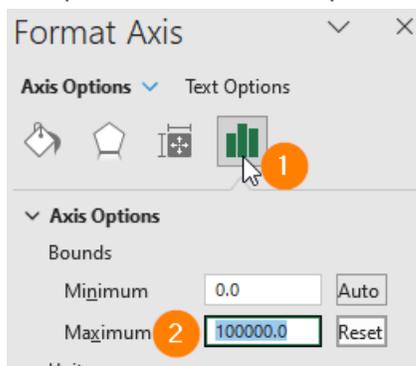
Ya que los salarios no sobrepasan \$100,000, podemos ajustar la escala del eje Y. Más adelante haremos lo mismo con la otra gráfica para Puerto Rico.

- Para ajustar la escala del eje Y;



- Haga **click** encima de cualquiera de los números del eje Y.

- En el panel **Format Axis** que verá a mano derecha de la interfaz de Excel:

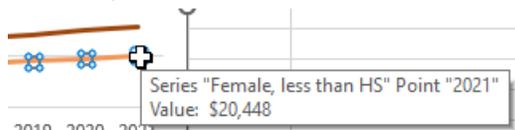


- 1: Haga **click** en el botón **Axis Options**.
- 2: Escriba **100000** (cien mil) Puede escribirlo con una coma, pero Excel la borrará.

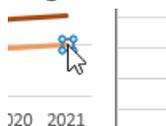
AÑADIR EL NOMBRE DE LA SERIE AL ÚLTIMO PUNTO (AÑO)

Para evitar tener que estar identificando cada serie con la leyenda, podemos incluir el nombre de la serie en uno o el último punto de cada serie. Esto se hace de manera indirecta y se debe hacer con cautela para no incluir el nombre en cada punto de la línea. Cuando tengamos todas las series identificadas, podemos descartar la leyenda.

- Para identificar la serie de una línea, haga **click en el último punto de la línea** de la serie **Female, less than HS**

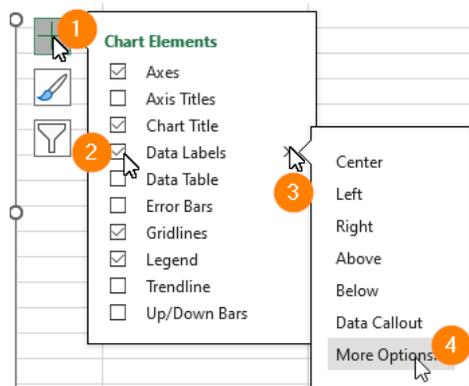


- Haga otro click encima del mismo último punto.



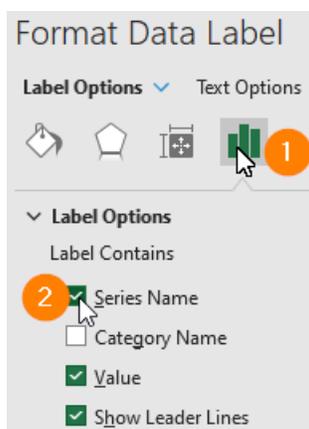
NO debemos dejar que todos los puntos de la serie estén activos, ya que solo necesitamos identificar el último punto.

- Con ese último punto seleccionado, siga estos pasos:



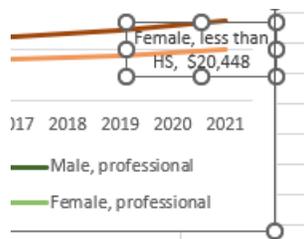
- 1: Haga **click** en el botón **Chart Elements**.
- 2: Haga **check** en la opción **Data Labels**.
- 3: Haga **click** en el **triángulo** para expandir y traer más opciones.
- 4: Haga **click** en la opción **More Options...** para traer el panel **Format Data Label**.

- En el panel **Format Data Label**:



- 1: Haga **click** en el botón **Label Options**.
- 2: Haga **check** en la opción **Series Name**.

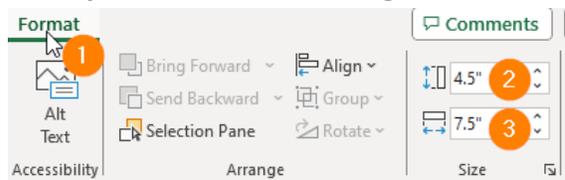
Notará que esta etiqueta no se ve bien, porque, en resumen, no cabe en ese lugar. Para que la etiqueta de la serie se vea mejor, será necesario aumentar el tamaño de la gráfica y reducir el espacio del elemento central.



CAMBIAR NUMÉRICAMENTE EL TAMAÑO DE LA GRÁFICA

Para que todos podamos producir una gráfica más o menos igual, vamos a ajustar el tamaño de la gráfica de manera numérica.

- Para ajustar el tamaño de la gráfica, active la gráfica (click en el borde)



- 1: Vaya al **menú principal** y escoja **Format**.
- 2: En la sección **Size**, en **altura**, escriba **4.5**
- 3: Escriba **7.5** en la caja de texto para **ancho**.

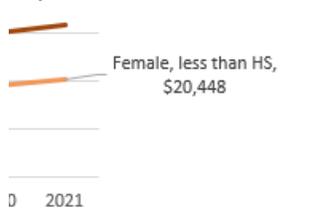
Para que la etiqueta de la serie y último valor no ocupen parte del “plot area”, debemos ajustar el tamaño de esta área, reduciéndola hacia la izquierda.

- Active el área central de la gráfica



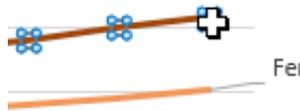
- Haga **click** en el **círculo del borde derecho** del área central (**plot area**) de la gráfica
- Arrastre** ese borde **hacia la izquierda, hasta el año 2019**

Los elementos de la gráfica se ajustarán, dejando espacio en blanco a la derecha. Este espacio lo usaremos para identificar las series.



IDENTIFICAR LA SERIE MALE, LESS THAN HS

- Para identificar la serie **Male, less than HS**, haga **click** en el **último punto de la línea** de la serie

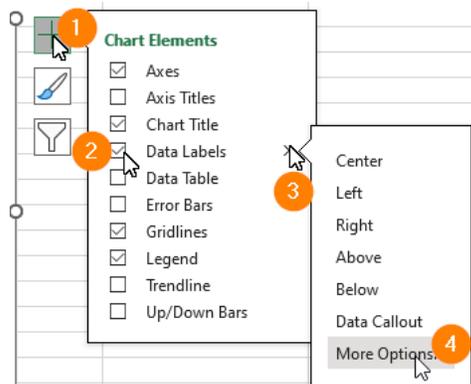


- Haga otro **click** encima del mismo último punto.



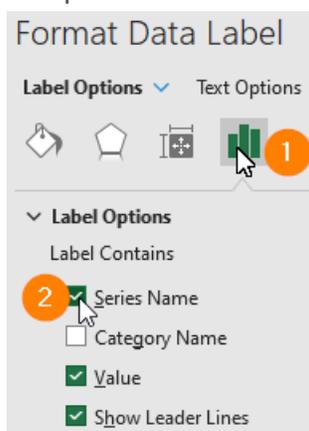
NO debemos dejar que todos los puntos de la serie estén activos, ya que solo necesitamos identificar el último punto.

- Con ese último punto seleccionado, siga estos pasos:



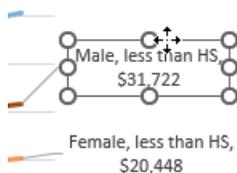
- 1: Haga **click** en el botón **Chart Elements**.
- 2: Haga **check** en la opción **Data Labels**.
- 3: Haga **click** en el **triángulo** para expandir y traer más opciones.
- 4: Haga **click** en la opción **More Options...** para traer el panel **Format Data Label**.

- En el panel **Format Data Label**:



- 1: Haga **click** en el botón **Label Options**.
- 2: Haga **check** en la opción **Series Name**.

- Arrastre** esta etiqueta **al espacio blanco** al lado derecho y **arriba de la etiqueta anterior**.



Repitamos el proceso para la próxima serie.

IDENTIFICAR LA SERIE FEMALE, BACHELOR'S DEGREE

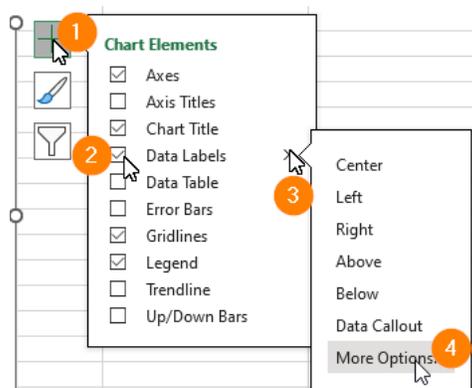
- Para identificar la serie **Female, bach degree**, haga **click** en el **último punto de la línea azul claro** de la serie



- Haga otro **click** encima del mismo último punto.

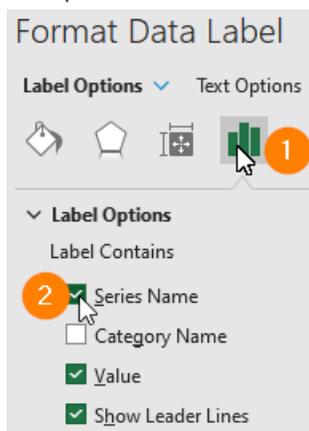


- Con ese último punto seleccionado, siga estos pasos:



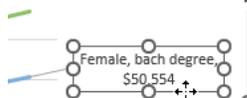
- 1: Haga **click** en el botón **Chart Elements**.
- 2: Haga **check** en la opción **Data Labels**.
- 3: Haga **click** en el **triángulo** para expandir y traer más opciones.
- 4: Haga **click** en la opción **More Options...** para traer el panel **Format Data Label**.

- En el panel **Format Data Label**:



- 1: Haga **click** en el botón **Label Options**.
- 2: Haga **check** en la opción **Series Name**.

- Arrastre** esta etiqueta al espacio blanco al lado derecho y arriba de la etiqueta anterior.



Repitamos el proceso para la próxima serie.

IDENTIFICAR LA SERIE FEMALE PROFESSIONAL

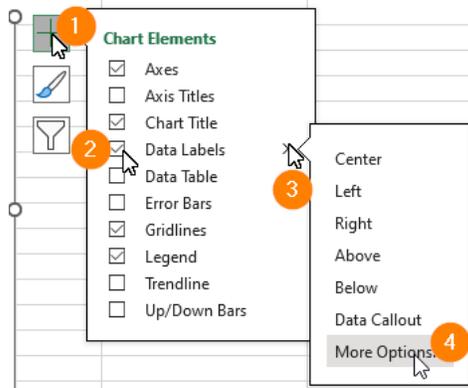
- Para identificar la serie **Female, professional**, haga **click** en el **último punto de la línea verde claro** de la serie



- Haga otro **click** encima del mismo último punto.

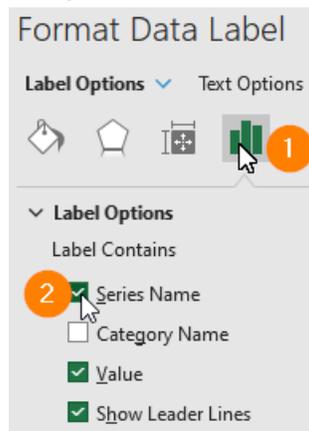


- Con ese último punto seleccionado, siga estos pasos:



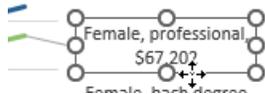
- 1: Haga **click** en el botón **Chart Elements**.
- 2: Haga **check** en la opción **Data Labels**.
- 3: Haga **click** en el **triángulo** para expandir y traer más opciones.
- 4: Haga **click** en la opción **More Options...** para traer el panel **Format Data Label**.

- En el panel **Format Data Label**:



- 1: Haga **click** en el botón **Label Options**.
- 2: Haga **check** en la opción **Series Name**.

- Arrastre** esta etiqueta **al espacio blanco** al lado derecho y **arriba de la etiqueta anterior**.



Repitamos el proceso para la próxima serie.

IDENTIFICAR LA SERIE MALE, BACHELOR'S DEGREE

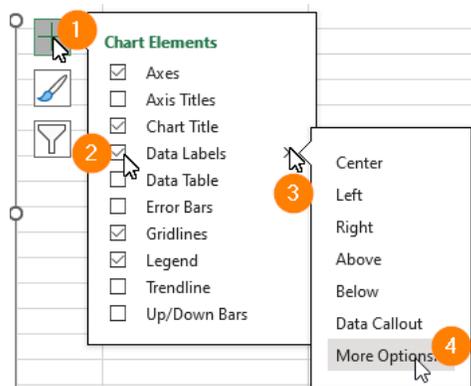
- Para identificar la serie **Male, bach degree**, haga **click** en el **último punto de la línea azul oscuro** de la serie



- Haga otro **click** encima del mismo último punto.

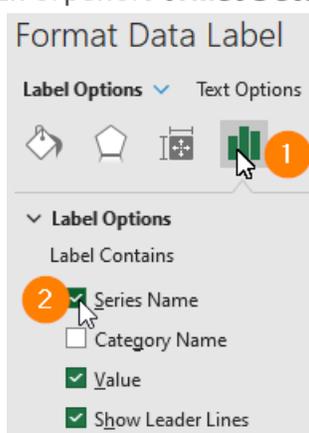


- Con ese último punto seleccionado, siga estos pasos:



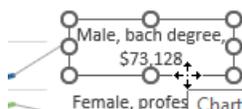
- 1: Haga **click** en el botón **Chart Elements**.
- 2: Haga **check** en la opción **Data Labels**.
- 3: Haga **click** en el **triángulo** para expandir y traer más opciones.
- 4: Haga **click** en la opción **More Options...** para traer el panel **Format Data Label**.

- En el panel **Format Data Label**:



- 1: Haga **click** en el botón **Label Options**.
- 2: Haga **check** en la opción **Series Name**.

- **Arrastre** esta etiqueta **al espacio blanco** al lado derecho y **arriba de la etiqueta anterior**.



Repitamos el proceso para la próxima serie.

IDENTIFICAR LA SERIE MALE, PROFESSIONAL

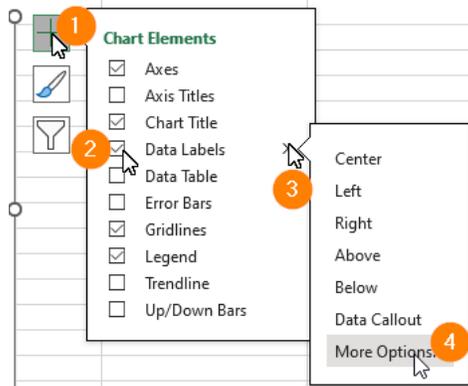
- Para identificar la serie **Male, professional**, haga **click** en el **último punto de la línea verde oscuro** de la serie



- Haga otro **click** encima del mismo último punto.

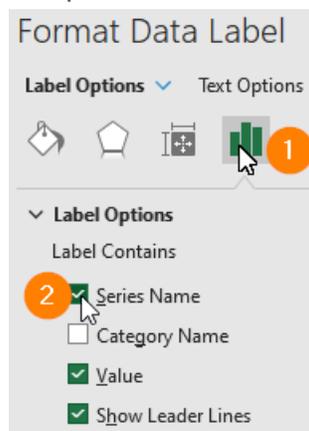


- Con ese último punto seleccionado, siga estos pasos:



- 1: Haga **click** en el botón **Chart Elements**.
- 2: Haga **check** en la opción **Data Labels**.
- 3: Haga **click** en el **triángulo** para expandir y traer más opciones.
- 4: Haga **click** en la opción **More Options...** para traer el panel **Format Data Label**.

- En el panel **Format Data Label**:

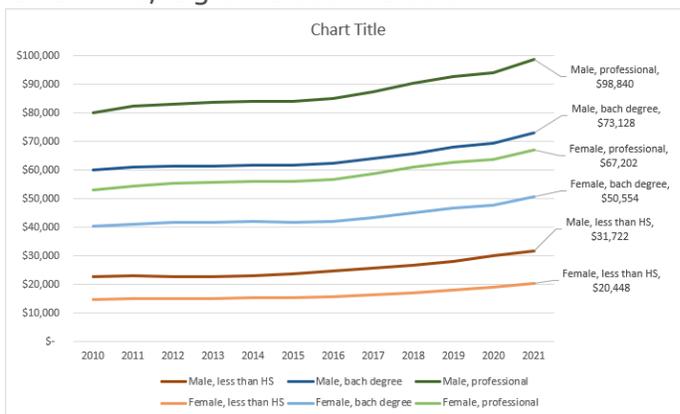


- 1: Haga **click** en el botón **Label Options**.
- 2: Haga **check** en la opción **Series Name**.

- **Arrastre** esta etiqueta **al espacio blanco** al lado derecho y **arriba de la etiqueta anterior**.



Hasta ahora, la gráfica debe verse así:



Queda por hacer:

- 1: **Eliminar** la **leyenda** por ser redundante,
- 2: **Escribir** el **título** de la gráfica y
- 3: **Añadir** **etiquetas** en el **punto de inicio** de cada **serie**.

AÑADIR ETIQUETAS EN PUNTOS DE INICIO DE CADA SERIE

Ya que estamos añadiendo etiquetas, pasemos a añadir a cada punto de inicio de cada serie, una etiqueta. Las etiquetas ayudan a comparar el salario en el primer año contra el último año. Para añadirlas, el proceso es casi el mismo, excepto que no necesitamos identificar la serie.

- Para añadir una etiqueta a la serie **"Male, professional"** haga **click** en la línea verde oscuro.

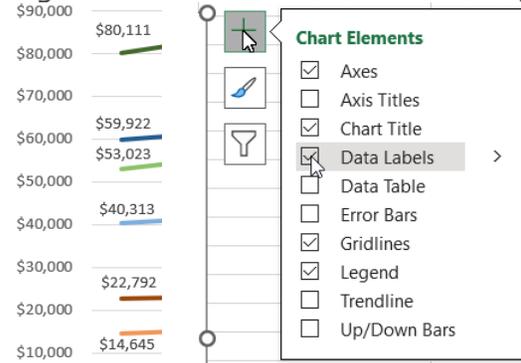


- Haga un **click** en el **punto extremo izquierdo** de esta serie para seleccionarlo.



Observe que no haya seleccionado los demás puntos. Solo éste.

- Haga **click** en el botón **Chart Elements**, luego **check** en la opción **Data Labels**.



- Repita el proceso** para cada una de las series: **Male, bach degree, Female, professional, Female, bach degree, Male, less than HS, Female, less than HS**.
- Debe alejar etiquetas que estén demasiado próximas**. Localice las etiquetas para que estén **cerca de la serie que representan**.

AÑADIR EL TÍTULO

Al añadir el título, vamos a cambiar el tipo de letra y tamaño para diferenciar este elemento de los demás.

Gender pay gap, US 2010-2021.

Arial Black, 14, #595959

- Para cambiar el título, haga **doble click** encima del título existente.

- Borre su contenido original y escriba: Gender pay gap, US 2010-2021

- Haga **click en el borde** de esta caja de texto del título.

Si no activa la caja, el próximo paso no saldrá completo.

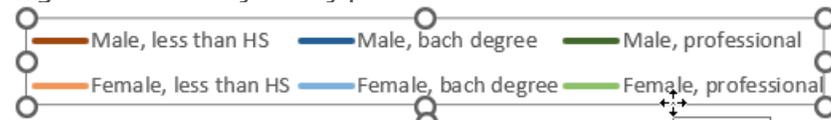
- Vaya al **menú principal**, **click** en **Home**. Vaya a la sección **Font** y **escoja** o escriba el tipo de letra **Arial Black**. El tamaño será **14** puntos.

Gender pay gap, US 2010-2021

DESCARTAR LA LEYENDA

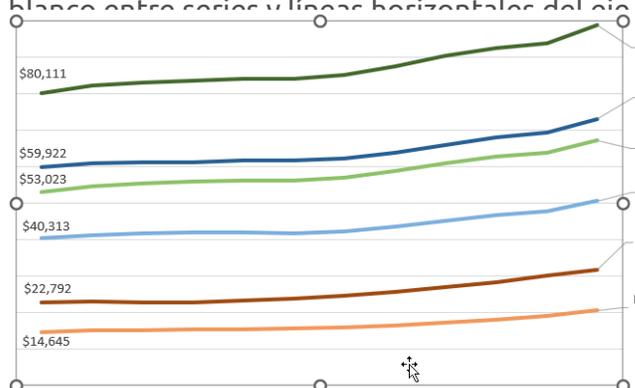
Como se mencionó antes, ya que las series están identificadas, podemos disponer de la leyenda.

- Haga **click** en la leyenda y presione el botón **Delete**.



El espacio que deja la leyenda puede ser ocupado por el "plot area".

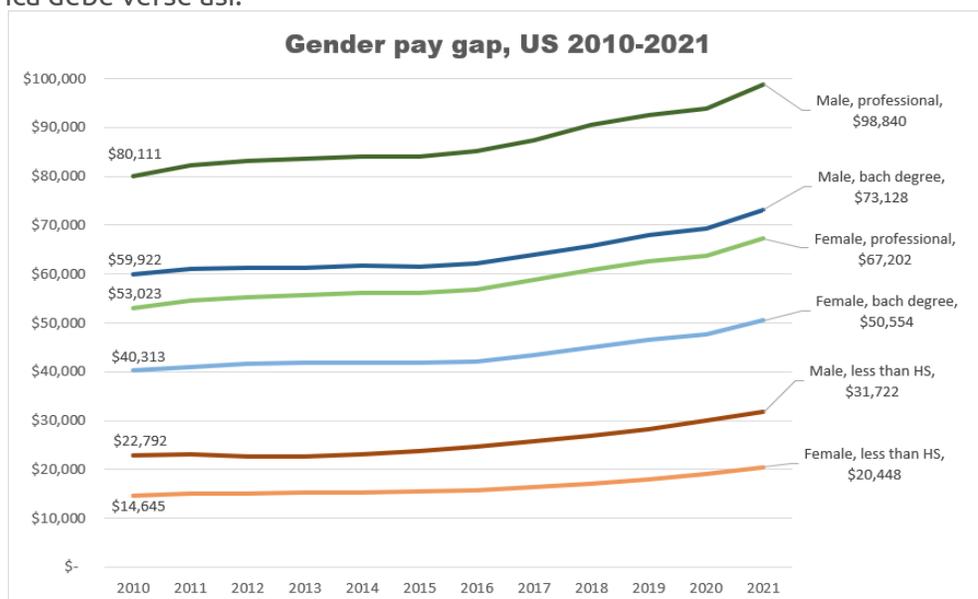
- Seleccione (**click**) la parte central (plot area) de la gráfica. Preferiblemente en espacio en blanco entre series y líneas horizontales del eje Y.



- Estire esta parte central hacia abajo, hasta llegar al extremo inferior de la gráfica.



La gráfica debe verse así:



¿Qué puede observar entre las series **Female, professional** y **Male, Professional**? ¿Le parece que eso es justo?

HACER LA GRÁFICA LINEAL CON DATOS DE PUERTO RICO

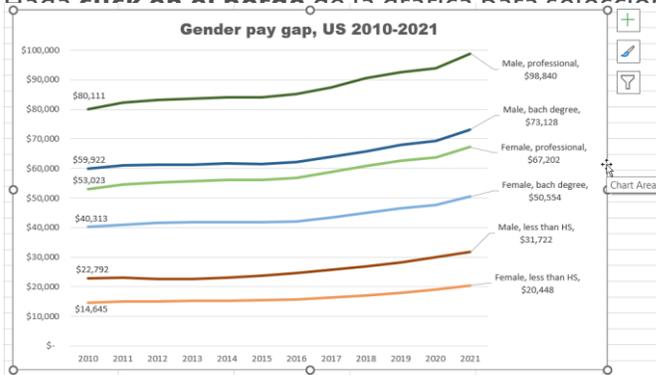
Vamos a hacer la misma gráfica lineal usando los datos para Puerto Rico. Es la misma serie temporal que proviene de la interfaz de datos censales. Estos datos están en la hoja `gender_pay_gap_pr`.

Debido a que **la estructura de los datos es igual en ambas tablas, podemos copiar la gráfica** que hicimos en la hoja para datos de Puerto Rico.

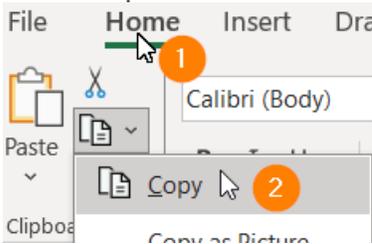
COPIAR Y PEGAR LA GRÁFICA EN OTRA HOJA

En esta parte vamos a copiar la gráfica **dos veces** en otra hoja para comparar ambas, una a cada lado.

- Haga **click** en el borde de la gráfica para seleccionarla.



- Vaya al **menú principal** y haga **click** en **Home**, luego en la primera opción del botón **Copy**. También puede hacer solo **ctrl+c**.



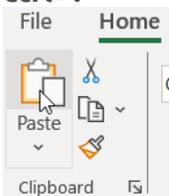
- Ahora ubíquese en la hoja **gender_pay_gap_pr**, haciendo **click** en el tab de esta hoja.

gender_pay_gap_pr

- Antes de copiar la gráfica, haga **click** en la celda K15. Así ubicaremos la gráfica en un lugar análogo a donde está ubicada la gráfica en la otra hoja.

year	NAME	Male, less than HS
2010	Puerto Rico	\$ 10,941
2011	Puerto Rico	\$ 11,173
2012	Puerto Rico	\$ 11,250
2013	Puerto Rico	\$ 11,327
2014	Puerto Rico	\$ 11,493
2015	Puerto Rico	\$ 11,204
2016	Puerto Rico	\$ 11,174
2017	Puerto Rico	\$ 11,383
2018	Puerto Rico	\$ 11,534
2019	Puerto Rico	\$ 11,658
2020	Puerto Rico	\$ 11,713
2021	Puerto Rico	\$ 12,160

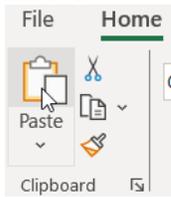
- En el **menú principal**, haga **click** en **Home**, luego en el botón **Paste**. También puede usar **ctrl+v**



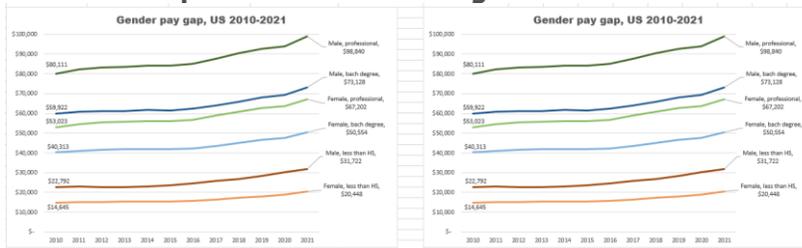
- Para copiar la otra gráfica, haga **click** en la celda U15.

	N	O	Q	T	U
13	\$ 30,281	\$ 43,751	\$ 8,770	\$ 24,124	\$ 32,527
14					
15					

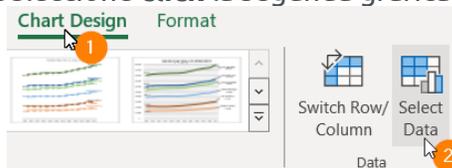
- En el **menú principal**, haga **click** en **Home**, luego en el botón **Paste**. También puede usar **ctrl+v**



La gráficas van a copiarse exactamente igual.



- Una vez copiadas la gráficas, solo tenemos que **cambiar la referencia de los datos** a la gráfica de la derecha
- Seleccione **click** la segunda gráfica a la derecha.



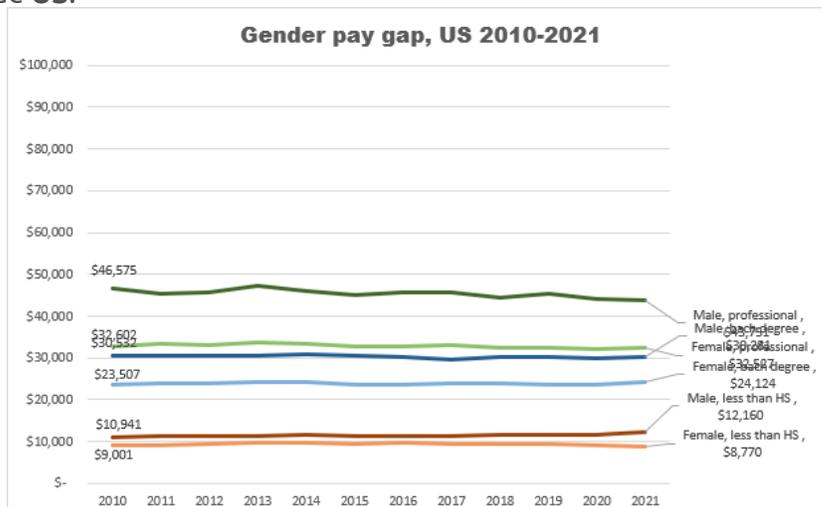
- Vaya al **menú principal** y escoja **Chart Design**
- Luego haga **click** en el botón **Select Data**.

- En la forma **Data Source**, vaya a la caja de texto **Chart data range**, y cambie la parte del nombre de la hoja **us** por **pr**

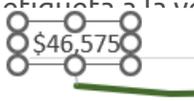


- No hay que cambiar nada más aquí. Presione **Enter**.

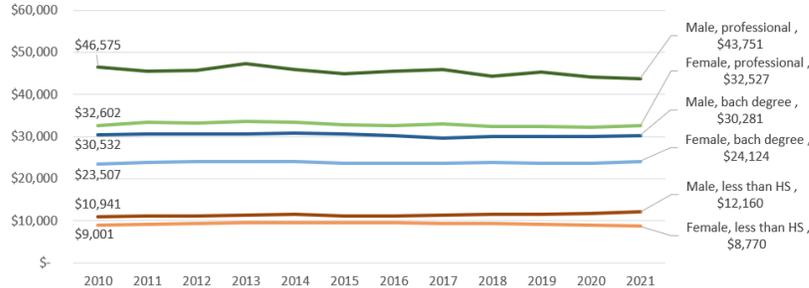
La gráfica reflejará el cambio de hoja de datos. Solo será necesario mover las etiquetas que estén demasiado cerca o que estén solapadas. Además, tenemos que cambiar el título para que lea **PR** en vez de **US**.



- Practique mover las etiquetas como hizo en la parte anterior. Recuerde que para mover una etiqueta a la vez, esta debe estar seleccionada de esta manera:

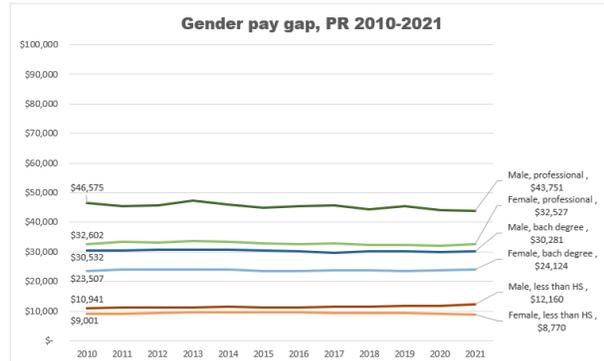
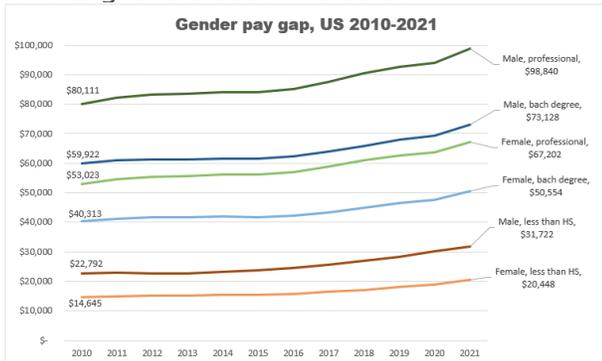


- Las etiquetas deben quedar más o menos así:



- Modifique el título. Intercambie US por PR.
Gender pay gap, PR 2010-2021

Ambas gráficas deben verse así:



La escala de las gráficas es la misma.

Note las diferencias salariales en conjunto entre las dos gráficas.

¿Qué está pasando con los salarios en Puerto Rico? ¿Cómo cree que esto afecta la economía?

¿Quiénes tienen mayor brecha salarial entre las dos gráficas?

- Con esto terminamos este módulo sobre gráficas.
- No es necesario guardar el archivo/ workbook.

Revisión de conceptos

Data range: _____

Filtro: _____

Sparkline: _____

MÓDULO 2: EXCEL TABLES Y FUNCIONES

RESUMEN

Según los datos van acumulándose y haciéndose más detallados, aumenta la necesidad de poder manipular eficientemente los datos. Más allá de los *data ranges* y los *named ranges*, las nuevas “**tablas**” en Excel ayudan a el manejo, análisis y resumen de los datos mediante ayudas visuales y funcionalidades diferentes aplicadas a estas **tablas**.

OBJETIVOS

Luego de completar este módulo, usted podrá:

- Describir qué es una tabla y su utilidad
- Crear una tabla Excel
- Usar tablas para resumir datos
- Describir algunas funciones para aplicar fórmulas y agregar datos

BENEFICIOS DE LAS TABLAS EXCEL

Una tabla es una extensión de celdas (*data range*) formateadas para ayudar a visualizar y analizar datos

- Todas las celdas de una tabla son tratadas en conjunto
- Se le aplican colores alternados o esquemas de color para hacer los datos más legibles
- Aparecerán botones de filtros a las columnas
- Se puede añadir una fila de totales o resúmenes numéricos
- Provee una manera rápida para comenzar el análisis de datos

QUÉ ES UNA TABLA

Una tabla de Excel es un grupo de celdas formateadas de forma particular para ayudar en la visualización y el análisis de datos.

FORMATEO

Al crear una tabla, las celdas se formatean utilizando colores por fila y columna, filtros en cada columna. El formato visual se utiliza para ayudar en el proceso de análisis de los datos; no solamente para que las celdas luzcan bien. Esto ayuda en el proceso de análisis, ya que además de proveer botones para hacer filtros, las celdas aparecerán mejor organizadas.

UTILIDAD DE LAS TABLAS

Cuando se genera una tabla, todas las celdas que la componen son tratadas como un conjunto, de manera similar a un **data range con nombre** (*named data range*) pero con mayor funcionalidad. Por ejemplo, al tratar de cambiar formatos, se cambiarán los formatos a las filas o a las columnas. Además, se aplicarán cálculos en ristra a las celdas de una columna hacia abajo automáticamente.

CREAR TABLAS

Para crear tablas:

- **Sombrear un data range** y hacer **click** en **Insert | Table**



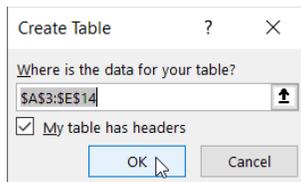
- De *otro modo*: hacer **click** en **Tell me what to do** y **click** en **Insert a Table**



- Aparecerá el tab **Design**, el cual le permitirá las funcionalidades de:
 - Color y estilos
 - Color en filas y/o columnas
 - Ennegrecer (*bold*) los nombres de las columnas
 - Mostrar cabeceras de columnas y fila de totales
 - Remover duplicados
- Para **convertir una tabla a data range**, haga **click** en **Design | Convert to Range**.



- Puede **crear una tabla a partir de un data range** usando **CTRL+A** *dentro* del data range y luego **CTRL+T**. Si su data range ya tiene nombres de columnas, mantenga **check** en **My table has headers**. Finalmente, **OK** en la forma **Create Table**.



FORMATEO DE LA TABLA

Luego de haber creado la tabla, aparecerá el tab **Table Design** con sus botones y múltiples opciones



Properties	Para nombrar la tabla
Tools	Manipulación de tablas
External Table Data	Conexión a datos en tablas externas
Table Style Options y Table Styles	Maneras para cambiar la apariencia de las tablas

Hablemos un poco del grupo **Tools**; especialmente los botones **Convert to Range** y **Remove Duplicates**.

Remove Duplicates: como dice el nombre, este botón se usa para eliminar récords duplicados. En ocasiones, pueden ocurrir duplicados, ya sea por error o de manera intencional. En ocasiones, se debe analizar usando una sola ocurrencia o un récord para no sobreestimar estos datos.

Convert to Range: se usa para **convertir** la tabla **en un data range**.

ORDENAR Y FILTRAR

Al convertir a tabla, se añaden automáticamente los botones de filtros a cada columna.

Zona	POP2010	POP2020
Sureste	89,018	72,843
Este	109,936	95,630
Centro Oeste	141,972	127,386
Noreste	189,073	162,995
Noroeste	229,767	212,465
Oeste	273,053	239,427
Centro	262,070	240,332
Norte	392,939	358,596
Centro Este	433,025	383,596
Sur	464,633	387,523
Metro	1,144,785	1,005,081

TIPOS DE FILTROS

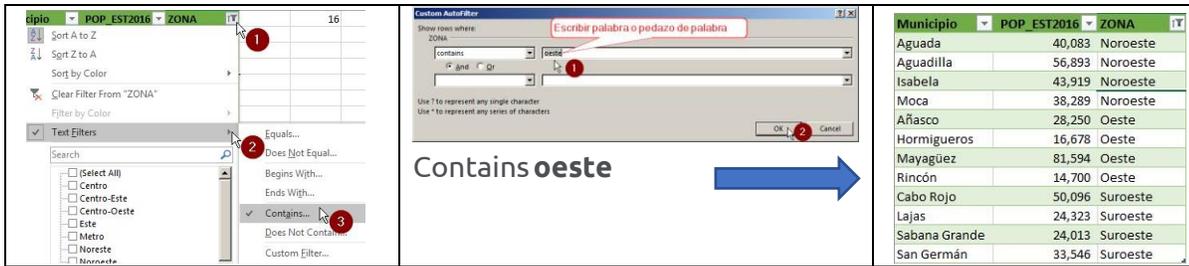
Criterios	Selección de datos de acuerdo con ciertos criterios que se escojan, ya sea por un número, nombre o código, filas antes o después de un valor, celdas entre dos valores, primeros 10 , o por porcentaje
Formatos	Selección de acuerdo con formatos de las celdas : color, etc.
Lista	Seleccionar mediante una lista de valores existentes , ya sean nombres o números

Asegúrese que cada columna sea de un solo tipo de dato: texto o numérico. No mezclar números y texto en una columna numérica porque afectará los cálculos.

ORDENAR (SORTING)

- Ascendente o descendente
- Por fechas
- Las opciones varían según el tipo de dato.
- Ordenamiento personalizado (custom sorting)

Ejemplo de tabla con un filtro aplicado y ordenado ascendente en la columna ZONA



1

2

3

1

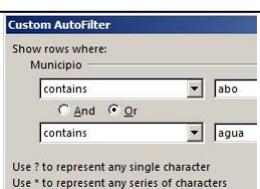
2

3

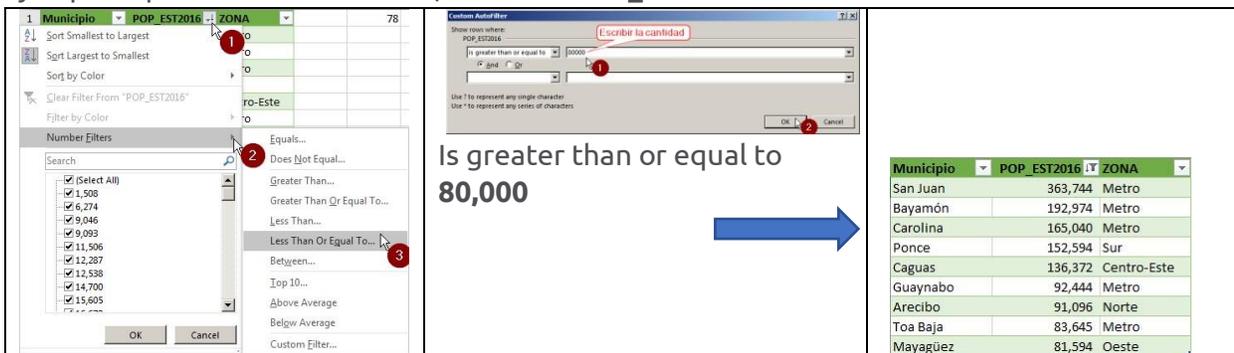
Contains oeste

Municipio	POP_EST2016	ZONA
Aguada	40,083	Noroeste
Aguadilla	56,893	Noroeste
Isabela	43,919	Noroeste
Moca	38,289	Noroeste
Añasco	28,250	Oeste
Hormigueros	16,678	Oeste
Mayagüez	81,594	Oeste
Rincón	14,700	Oeste
Cabo Rojo	50,096	Suroeste
Lajas	24,323	Suroeste
Sabana Grande	24,013	Suroeste
San Germán	33,546	Suroeste

Comparación de cadenas alfanuméricas (strings): texto, códigos

	Equals...	Pareo exacto del texto
	Does not Equal...	El texto no es igual a
	Begins with...	Comienza con texto...
	Ends With...	Termina con texto...
	Contains...	Celda(s) que contengan la cadena de texto...
	Does not contain...	Celda(s) que NO contengan la cadena de texto...
		Custom filter... Hasta dos condiciones para filtrar. En este caso, municipios que contengan abo y agua en las letras de su nombre . Por ejemplo, Guaynabo, Gurabo, Caguas, Aguada, Aguadilla, Aguas Buenas

Ejemplo: Aplicar filtro numérico, columna POP_EST2016:



1

2

3

1

2

3

Is greater than or equal to 80,000

Municipio	POP_EST2016	ZONA
San Juan	363,744	Metro
Bayamón	192,974	Metro
Carolina	165,040	Metro
Ponce	152,594	Sur
Caguas	136,372	Centro-Este
Guaynabo	92,444	Metro
Arecibo	91,096	Norte
Toa Baja	83,645	Metro
Mayagüez	81,594	Oeste

Equals...	Es igual a
Does Not Equal...	No es igual a
Greater Than...	Mayor que
Greater Than or Equal To...	Mayor o igual que
Less Than...	Menor que
Less Than Or Equal To...	Igual o menor que
Between...	Entre dos valores
Top 10...	Los 10 más altos
Above Average	Sobre el promedio
Below Average	Bajo el promedio
Custom Filter...	Similar al Custom de texto, pero con las opciones numéricas

ALGUNOS CONCEPTOS ANTES DE DISCUTIR SOBRE FUNCIONES

En esta sección, discutiremos algunas funciones que nos pueden ayudar en las labores diarias o para la preparación de informes que requieran expresar resúmenes por selecciones y categorías.

OBJETIVOS

Luego de terminar esta sección, usted conocerá algunos conceptos importantes sobre **nombrar grupos de celdas** (*named ranges*, y *named columns*), **referencias relativas y absolutas a celdas, filas y columnas**, además de una serie de **funciones** tales como **AGGREGATE**, **VLOOKUP**, **INDEX(MATCH)**, **COUNT**, **COUNTA**, **COUNTIFS**, **SUMIFS**, y funciones con fechas.

NAMED RANGES: NOMBRAR GRUPOS DE CELDAS Y COLUMNAS

En ocasiones es práctico seleccionar un subgrupo de celdas y ponerles un nombre. A esto se le conoce como *named range* y puede usarse para evitar estar sombreando frecuentemente un grupo o zona de celdas (data range) para diferentes tareas.

En ocasiones, tenemos un data range con varias columnas y solo queremos las primeras o últimas columnas, sin tener que esconderlas. Por ejemplo, la función **VLOOKUP** requiere que la primera columna tenga unos códigos para identificar y parear, pero en nuestro data range no está en la primera columna. En este caso, podemos asignarle nombre a un subconjunto dentro del data range de referencia que incluya como primera columna a esta columna que tiene los códigos.

CÓMO HACER UN NAMED RANGE

Seleccionar el **data range** usando **click** y **arrastrar** hacia la **celda final** hasta cubrir toda la zona de interés.

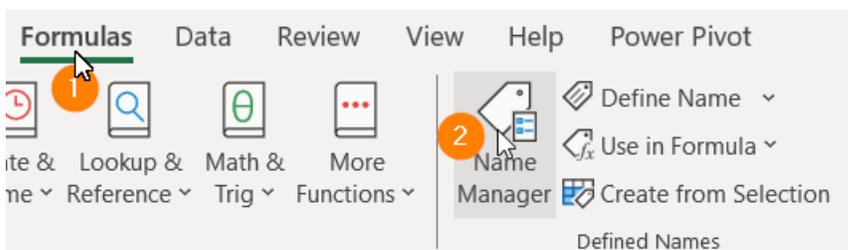


Luego **escribir el nombre** del nuevo **data range** en la **caja de texto** donde aparece el nombre de la celda. En este caso era A1.

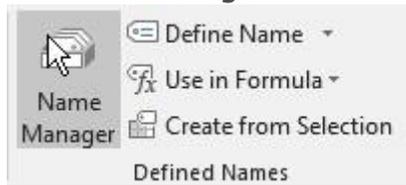


BORRAR O REDEFINIR NAMED RANGES

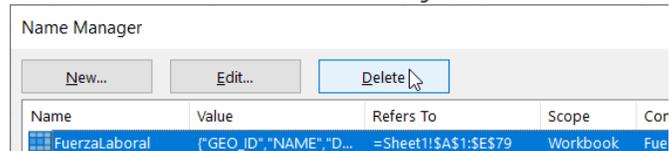
Para **borrar un named range**, vaya al **menú principal** en el tab **Formulas** y haga **click** en el botón **Name Manager**, localizado en el grupo **Defined Names**.



1: Click en Name Manager



2: Click en el ítem con nombre y click en botón Delete



REFERENCIAS RELATIVAS

En Excel nos podemos referir a celdas de manera relativa o absoluta. Por defecto, las referencias son relativas. Las referencias relativas a celdas cambian cuando una misma fórmula se aplica a otra celda. Esto facilita el cómputo en ristra, tomando como partida a la celda siguiente para aplicar la fórmula. A1+B1, SUM(A1:R1), etc. son ejemplos de referencias relativas.

REFERENCIAS ABSOLUTAS:

En ocasiones, solo queremos mantener la celda de origen sin cambiar su referencia, y aplicarle una fórmula a las demás que deseamos calcular. Podemos mantener referencia fija (absoluta) a una celda, a una columna o a una fila. Estas las podemos identificar mediante el signo de dólar \$.

Ejemplos de referencias absolutas:

Solo a una celda	\$A\$1	La referencia a la celda (fila y columna) se mantiene fija.
Solo a una fila	A\$1	La referencia a la fila (row) no cambia al copiar fórmulas.
Solo a una columna	\$A1	La referencia a la columna no cambia al copiar fórmulas.

CÓDIGOS DE ERROR EN EXCEL Y QUÉ SIGNIFICAN

#DIV/0!	División por cero
#VALUE!	Inclusión de valor incorrecto en un argumento en una fórmula. Por ejemplo, entrar un texto en un argumento de una función cuando Excel espera un número.
#NAME?	Nombre incorrecto de función, data_range o referencia a una celda.
#REF!	Referencia inexistente. Por ejemplo, una fórmula escrita D2+E2 y luego borrar la columna D.
#NUM!	Hay un problema con un número en una fórmula o una función. Aparece cuando hay algún número no manejable por Excel (out of range). Puede ocurrir cuando Excel espera un número en una función y escribimos texto.
#N/A	Generalmente aparece cuando Excel no encuentra algún valor en una lista. Puede incluir otros errores.
#####	No es error. Significa que el tamaño (visual) de la celda no es suficiente para acomodar el valor que contiene. Estire la celda o haga doble click en el borde derecho de la columna para poder ver el valor completo.

FUNCIONES

En esta parte, discutiremos algunas funciones de interés de Excel. La mayoría tienen que ver con resumir datos y manejo de datos temporales (fechas y horas).

AGGREGATE

Una función útil para resumir datos que expresen valores tomados de listas (arrays). Estas pueden devolver valores de toda la lista o solamente de récords filtrados, ignorando celdas escondidas, celdas con errores y celdas que contengan subtotales

SINTAXIS:

AGGREGATE(function_num, options, array, [k])

function_num	Número asignado a la función a utilizarse. La lista de funciones aparece a continuación en la tabla de funciones disponibles en AGGREGATE
options	Número asignado a la opción a usarse para contar o no ciertas celdas según su condición. La lista de opciones aparecerá a continuación.
array	Un array es una lista. Esta puede ser numérica o alfanumérica, el array puede ser compuesto por una columna, una referencia a un grupo de celdas (range) o a una función que devuelva un array.
k	Opcional, dependiendo del tipo de función escogida. En algunos casos, las funciones requieren un argumento adicional. Estas aparecerán en una tabla aparte más adelante.

Lista de funciones disponibles en AGGREGATE(function_num, options, array, [k])

1	AVERAGE	Promedio, <i>media</i> . Medida de tendencia central.
2	COUNT	Contar celdas que tengan algún valor o dato.
3	COUNTA	Contar las celdas que no estén vacías. Esto no incluye celdas que tengan una cadena alfanumérica vacía "". COUNTA evalúa (<i>cuenta</i>) celdas con error
4	MAX	Devuelve el valor más alto en un conjunto de datos.
5	MIN	Devuelve el valor menor en un conjunto de datos.
6	PRODUCT	Usa el contenido de celdas aisladas o en grupos como argumentos y los multiplica
7	STDEV.S	Desviación (error) estándar de la muestra
8	STDEV.P	Desviación (error) estándar de la población
9	SUM	Sumar valores individuales, referencias a celdas, data ranges o todas las anteriores.
10	VAR.S	Varianza de la muestra (error estándar al cuadrado)
11	VAR.P	Varianza de la población (error estándar al cuadrado)
12	MEDIAN	Valor central de una distribución ordenada de números
13	MODE.SNGL	Moda. Valor que más se repite
14	LARGE	Devuelve el valor n-ésimo más alto. Primer valor más alto, tercer, etc.
15	SMALL	Devuelve el valor n-ésimo más bajo. Primer valor más alto, tercer, etc.
16	PERCENTILE.INC	Valor que divide un conjunto ordenado de datos estadísticos de forma que un porcentaje de tales datos sea inferior a dicho valor. Un individuo en el percentil 80 está por encima del 80 % del grupo a que pertenece. (según la RAE)

17	QUARTILE.INC	Devuelve el cuartil, que es cualquiera de los valores de percentil 25, 50 o 75. Los parámetros son 0,1,2,3,4. MIN, MAX, MEDIAN, devuelven el resultado de los parámetros <i>quartile</i> 0, 4 y 2 respectivamente.
18	PERCENTILE.EXC	Igual que percentile.inc excepto que no incluye valores que esté en el límite de la clase.
19	QUARTILE.EXC	Igual que quartile.inc pero excluyente.

OPCIONES DE AGGREGATE(FUNCTION_NUM, OPTIONS, ARRAY,[K])

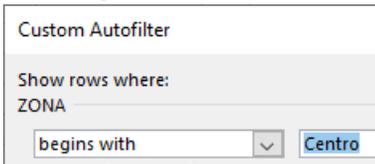
0 or omitted	Ignorar subtotales anidados y funciones AGGREGATE
1	Ignorar celdas escondidas (<i>hidden</i>), subtotales anidados y funciones AGGREGATE
2	Ignorar celdas con valor ERROR, subtotales anidados y funciones AGGREGATE
3	Ignorar celdas escondidas, celdas con ERROR, subtotales anidados y funciones AGGREGATE
4	No ignorar nada
5	Ignorar solo celdas escondidas
6	Ignorar celdas con valor ERROR
7	Ignorar celdas escondidas y celdas con valor ERROR

FUNCIONES QUE USAN EL PARÁMETRO K

LARGE(array, k)
SMALL(array, k)
PERCENTILE.INC y PERCENTILE.EXC
QUARTILE.INC y QUARTILE.EXC

Veamos un ejemplo de AGGREGATE(). En este ejemplo, hay dos *medidas* (números) *dinámicas*. Es decir: **cambian según se apliquen filtros**. Una celda devuelve el número de municipios usando **AGGREGATE con la opción COUNT** y la otra celda usa **AGGREGATE con la opción SUM** para la suma de población de los municipios que estén en la lista. En ambas celdas se usa la opción "1" para que la función AGGREGATE no evalúe celdas escondidas.

En estos casos, podemos ver que la columna **Zona** tiene un filtro aplicado con todas las celdas que contengan la secuencia alfanumérica "Centro".



Ambas funciones AGGREGATE usan la **opción 5** para **ignorar celdas escondidas**.

	A	B	C	D	E	F	G	H	I	J	K	L
1	GEO_ID	Municipio	POP2020	ZONA		Municipios: 22	población:		694,041			
2	72001	Adjuntas	18020	Centro		FormulaText: =AGGREGATE(3,1,B2:B79)			=AGGREGATE(9,1,C2:C79)			
3	72007	Aguas Buenas	24223	Centro-Este								
4	72009	Albionito	24637	Centro-Este								
5	72019	Barranquitas	28983	Centro-Este								
6	72025	Caguas	127244	Centro-Este								
7	72035	Cayey	41652	Centro-Este								
8	72039	Ciales	16984	Centro								
9	72041	Cidra	39970	Centro-Este								
10	72045	Comerio	18883	Centro-Este								
11	72047	Corozal	34571	Centro								
12	72063	Gurabo	40622	Centro-Este								
13	72073	Jayuya	14779	Centro								
14	72077	Juncos	37012	Centro-Este								
15	72081	Lares	28105	Centro-Oeste								
16	72083	Las Marías	8874	Centro-Oeste								
17	72093	Maricao	4755	Centro-Oeste								
18	72101	Morovis	28727	Centro								
19	72105	Naranjito	29241	Centro								
20	72107	Orocovis	21434	Centro								
21	72129	San Lorenzo	37693	Centro-Este								
22	72131	San Sebastián	39345	Centro-Oeste								
23	72141	Utuado	28287	Centro								

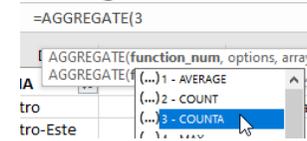
	A	B	C	D	E	F	G	H	I	J	K	L
1	GEO_ID	Municipio	POP2020	ZONA		Municipios: 10	población:		1,064,653			
2	72021	Bayamón	185187	Metro								
3	72031	Carolina	154815	Metro								
4	72033	Cataño	23155	Metro								
5	72051	Dorado	35879	Metro								
6	72061	Guaynabo	89780	Metro								
7	72087	Loíza	23693	Metro								
8	72127	San Juan	342259	Metro								
9	72135	Toa Alta	66852	Metro								
10	72137	Toa Baja	75293	Metro								
11	72139	Trujillo Alto	67740	Metro								

El filtro aplicado devuelve las zonas *Centro*, *Centro-Este* y *Centro-Oeste*. Hay 22 municipios en estas zonas y suman 694,041 habitantes (Censo 2020).

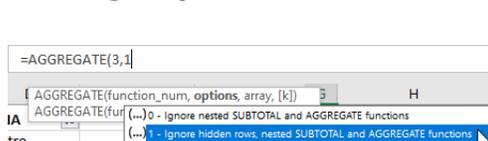
Otro ejemplo: Hay 10 municipios y 1,064,653 habitantes en la *Zona Metro* (Censo 2020)

- Conteo de municipios: =AGGREGATE(3,1,B2:B79)
- Alternativa: =AGGREGATE(3,5,B2:B79) usando función **COUNTA** con *nombres* de municipios
- Sumatoria de Población: =AGGREGATE(9,1,C2:C79)

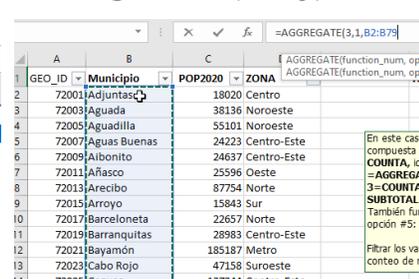
1: Escoger función



2: Escoger opción



3: Escoger lista (array)



LOOKUP

Esta función es útil para clasificar una lista de valores. A diferencia de la función VLOOKUP, la **función LOOKUP trabaja con una sola columna**. Por ejemplo, tener una columna con valores desde 0 a 1,000. Para propósitos de agrupar, tomamos los valores y los convertimos en clases y creamos diez grupos de cien: 0-100; 100-200; 200-300; 300-400; 400-500; 500-600; 600-700; 700-800; 800-900; 900-1,000.

LOOKUP va celda por celda y evalúa/compara el valor de la celda de inicio, con el valor del lookup_vector a asignarse en la otra celda o celdas en una columna. Puede usarse una lista existente en celdas o crear un arreglo (array).

SINTAXIS

LOOKUP(lookup_value, lookup_vector, [result_vector])

lookup_value	Celda que contiene el valor a buscar y parrear.
lookup_vector	Grupo de celdas o una lista contenida en una sola columna o fila. Los valores de la lista deben estar en orden ascendente . De lo contrario, los resultados no serán confiables.

[result_vector]	Una sola columna o fila que contiene los valores a ser devueltos. Para cada valor de lookup_vector debe corresponder un valor en el result_vector.
-----------------	--

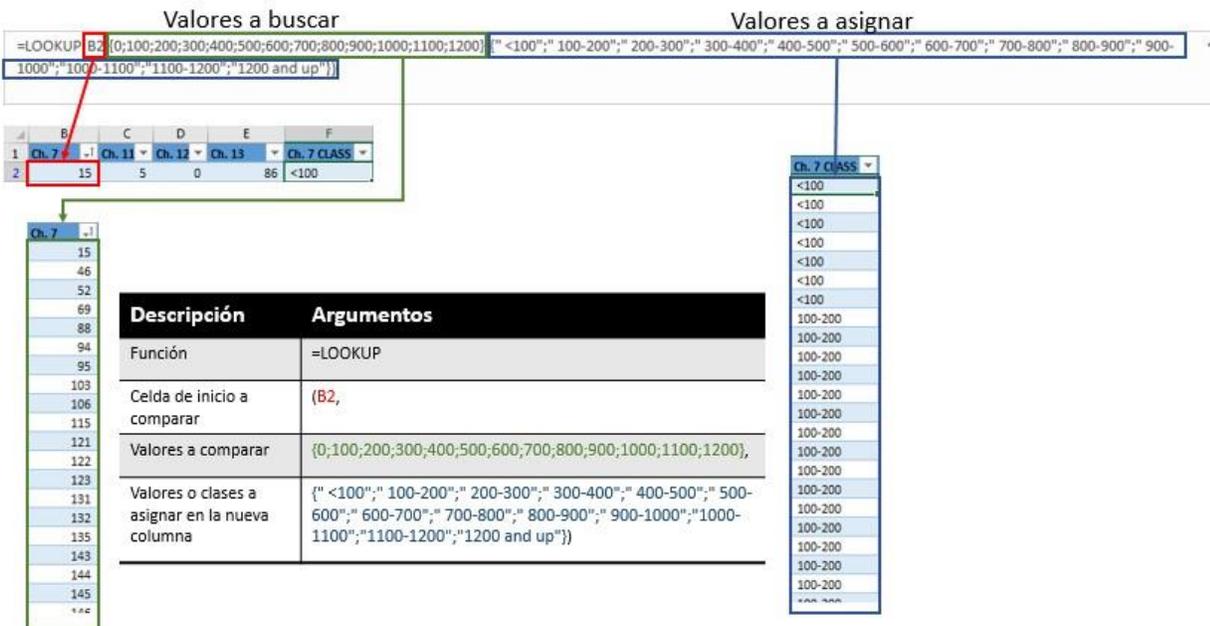
Por ejemplo:

=LOOKUP(B2,

{0;100;200;300;400;500;600;700;800;900;1000;1100;1200}, (ESTE ES EL LOOKUP VECTOR)

{" <100";" 100-200";" 200-300";" 300-400";" 400-500";" 500-600";" 600-700";" 700-800";" 800-900";" 900-1000";"1000-1100";"1100-1200";"1200 and up"}) (ESTE ES EL "RESULT" VECTOR)

¿Qué es un **vector**? En álgebra lineal, un vector es una **lista de valores**. En ciencias de cómputos puede ser llamado un arreglo (array) de una dimensión. **En Excel, un vector es una lista de una sola columna o de una fila.**



Descripción	Argumentos
Función	=LOOKUP
Celda de inicio a comparar	(B2,
Valores a comparar	{0;100;200;300;400;500;600;700;800;900;1000;1100;1200},
Valores o clases a asignar en la nueva columna	{" <100";" 100-200";" 200-300";" 300-400";" 400-500";" 500-600";" 600-700";" 700-800";" 800-900";" 900-1000";"1000-1100";"1100-1200";"1200 and up"}

VLOOKUP

Esta función se usa para **parear** (*lookup* = buscar) los **valores de la primera columna de una lista** (array) moviéndose hacia abajo (*vertical*) a través de las filas y **devolver el valor encontrado en otra celda o celdas**. VLOOKUP, HLOOKUP y XLOOKUP son parte del grupo de funciones de búsqueda y referencia ([lookup and reference](#)).

SINTAXIS:

VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])

Según la documentación de Microsoft sobre [VLOOKUP](#), la forma más simple de esta función es:

- =**VLOOKUP**(**lookup value**: *valor de referencia que desea buscar*,
- table_array**: *grupo de celdas de donde se busca este valor*,
- col_index_number**: *número de la columna en el grupo de celdas que contienen el valor a devolver*,
- range_lookup**: *pareo exacto o aproximado indicado como 0/FALSE o 1/TRUE*).

Recordar:

- Un requisito de **VLOOKUP** es que el *lookup_value* debe estar en la primera celda de la primera columna.
- No puede estar en otra columna, a menos que se designe de antemano, un *named_range* que incluya la columna de referencia como columna #1.
- Los valores en el *table_array* deberán estar ordenados (*sorted*).

Por ejemplo:

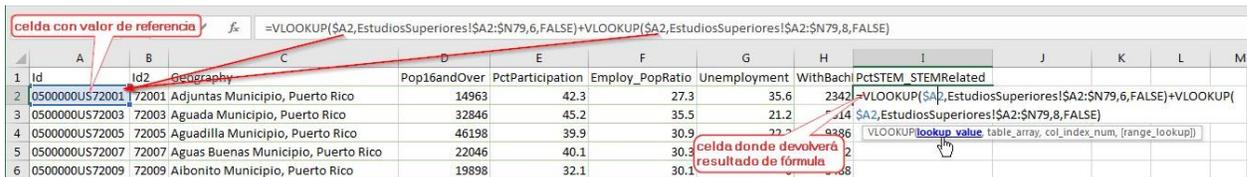
```
=VLOOKUP($A2,EstudiosSuperiores!$A2:$N79,6,FALSE)
```

lookup_value: \$A2	Celda que contiene el valor a buscar y parear. Note que esta <i>referencia a celdas es fija en la columna</i> , como se discutió anteriormente.
table_array: <i>EstudiosSuperiores!\$A2:\$N79</i>	En la hoja <i>EstudiosSuperiores</i> , el grupo de celdas va desde la \$A2:\$N79. Note que esta referencia a celdas es fija en las columnas A a la N.
col_index: 6	Número de la columna cuyos valores queremos transferir
range_lookup: FALSE	Requerimos pareo exacto.

En este caso, queremos sumar dos columnas usando dos LOOKUPS:

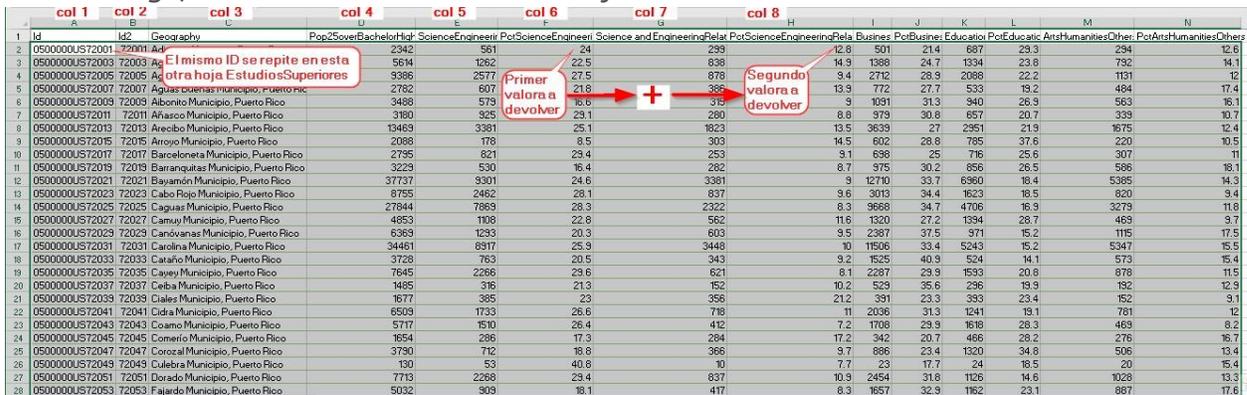
```
=VLOOKUP($A2,EstudiosSuperiores!$A2:$N79,6,FALSE)+VLOOKUP($A2,EstudiosSuperiores!$A2:$N79,8,FALSE)
```

Celda con valor de referencia.



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Id	Id2	Geography	Pop16andOver	PctParticipation	Employ_PopRatio	Unemployment	WithBach	PctSTEM_STEMRelated				
2	0500000US72001	72001	Adjuntas Municipio, Puerto Rico	14963	42.3	27.3	35.6	2342	=VLOOKUP(\$A2,EstudiosSuperiores!\$A2:\$N79,6,FALSE)+VLOOKUP(\$A2,EstudiosSuperiores!\$A2:\$N79,8,FALSE)				
3	0500000US72003	72003	Aguada Municipio, Puerto Rico	32846	45.2	35.5	21.2	114					
4	0500000US72005	72005	Agua Dulce Municipio, Puerto Rico	46198	39.9	30.9	22.2	2386					
5	0500000US72007	72007	Aguas Buenas Municipio, Puerto Rico	22046	40.1	30.3	30.3	2					
6	0500000US72009	72009	Aibonito Municipio, Puerto Rico	19898	32.1	30.1	30.1	28					

De este data range, toma el valor de la columna 6 y suma el valor de la columna 8



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Id	Id2	Geography	Pop16andOver	PctParticipation	Employ_PopRatio	Unemployment	WithBach	PctSTEM_STEMRelated					
2	0500000US72001	72001	Adjuntas Municipio, Puerto Rico	14963	42.3	27.3	35.6	2342						
3	0500000US72003	72003	Aguada Municipio, Puerto Rico	32846	45.2	35.5	21.2	114						
4	0500000US72005	72005	Agua Dulce Municipio, Puerto Rico	46198	39.9	30.9	22.2	2386						
5	0500000US72007	72007	Aguas Buenas Municipio, Puerto Rico	22046	40.1	30.3	30.3	2						
6	0500000US72009	72009	Aibonito Municipio, Puerto Rico	19898	32.1	30.1	30.1	28						

Resultado devuelto:

Formula Bar: `=VLOOKUP($A2,EstudiosSuperiores!$A2:$N79,6,FALSE)+VLOOKUP($A2,EstudiosSuperiores!$A2:$N79,8,FALSE)`

B	C	D	E	F	G	H	I
Id2	Geography	Pop16andOver	PctParticipation	Employ_PopRatio	Unemployment	WithBach	PctSTEM STEMRelated
72001	Adjuntas Municipio, Puerto Rico	14963	42.3	27.3	35.6	2342	36.8
72003	Aguada Municipio, Puerto Rico	32846	45.2	35.5	21.2	5614	37.4
72005	Aguadilla Municipio, Puerto Rico	46198	39.9	30.9	22.2	9386	36.9
72007	Aguas Buenas Municipio, Puerto Rico	22046	40.1	30.3	24.6	2782	35.7
72009	Aibonito Municipio, Puerto Rico	19898	32.1	30.1	6	3488	25.6
72011	Añasco Municipio, Puerto Rico	23190	43.8	37	15.5	3180	37.9
72013	Arecibo Municipio, Puerto Rico	74784	37.6	30.8	18.1	13469	38.6
72015	Arroyo Municipio, Puerto Rico	14869	33.6	26.1	22.3	2088	23
72017	Barceloneta Municipio, Puerto Rico	19656	33	27.2	17.7	2795	38.5

Annotations: "Resultado devuelto" points to the value 36.8 in cell I2. "Fórmula copiada a celdas subsiguientes" points to the range I3:I17.

XLOOKUP

Es una nueva [función de búsqueda](#). No está disponible en las versiones 2016 y 2019 de Excel. Hace una búsqueda en un data range o un array y devuelve el ítem que corresponde al valor de la primera celda que encuentre. Si no hay pareo, entonces devuelve el valor más parecido, aunque se puede especificar para que solo devuelva solo celdas con pareo exacto.

Sintaxis

<code>=XLOOKUP(lookup_value,</code>	Celda con el valor que deseamos buscar
<code>lookup_array,</code>	Arreglo de celdas de referencia
<code>return_array,</code>	Arreglo de celdas que tienen los datos que queremos copiar
<code>[if_not_found],</code>	Valor asignado si no hay pareo
<code>[match_mode],</code>	Tipo de pareo, 0: Pareo exacto, -1: pareo exacto o el próximo más pequeño, 1: pareo exacto o el próximo más grande, 2: pareo usando un comodín
<code>[search_mode])</code>	Tipo de búsqueda, 1: Del primero al último, -1 del último al primero, 2: búsqueda binaria en orden ascendente, -2: búsqueda binaria en orden descendente

Por ejemplo:

```
=XLOOKUP(A2,'pop2020'!A2:A79,'pop2020'!C2:C79,,0)
```

A2 (valor=72001) tiene el dato para comparar con el primer grupo de celdas

	A	B	C	D	E	F	G	H	I	J	K	L
1	GEO_ID	Municipio	1930	1940	1950	1960	1970	1980	1990	2000	2010	2020
2	72001	Adjuntas	18075	22556	22424	19658	18691	18786	19451	19143	19483	18020
3	72003	Aguada	14670	17923	20743	23234	25658	31567	35911	42042	41959	38136

POP_HISTORIC

Este dato se compara con el dato del grupo de datos de la columna A en otra hoja llamada pop2020

	A	B	C
1	GEO_ID	NAME	POP2020
2	72001	Adjuntas	18020
3	72003	Aguada	38136
4	72005	Aguadilla	55101
5	72007	Aguas Buenas	24223
6	72009	Aibonito	24637

pop2020

Esta es la hoja identificada en la fórmula XLOOKUP como 'pop2020'. Contiene los códigos **geo_id**, el nombre de los municipios y la **población en 2020**.

La lista de **valores a comparar** está en la **columna A** desde **A2** hasta **A79**.

La lista de valores que queremos transferir a la columna L de la hoja **POP_HISTORIC** está en la columna **C** desde **C2** hasta **C79**.

Es como decir celda L2= compara lo que hay en la celda A2 con la lista que hay en la columna A de la hoja POP 2020 y si porean los códigos, asigna el valor de la misma fila que está en la lista de población en la columna C en la hoja POP2020 a la celda L2 de la hoja POP_HISTORIC

Si los datos son iguales 72001=72001, va a seleccionar el valor correspondiente en la misma fila #2: 18020. Lo demás es repetir la fórmula modificada a las demás celdas de la fila L.

Esta función es mucho menos restrictiva que las versiones VLOOKUP y HLOOKUP, además de ser más simple.

COUNT, COUNTA, COUNTIF, COUNTIFS

Estas funciones son para conteo de celdas.

COUNT	Contar celdas que contengan números en un grupo de celdas.
Sintaxis:	COUNT(value1, [value2], ...)
	value1 puede ser un data range, celda, valor value2 es opcional e igualmente puede ser data range, celda, valor ... hasta 255 argumentos (<i>value</i>)
COUNTA	Cuenta celdas que no estén vacías en un grupo de celdas.
Sintaxis:	COUNTA(value1, [value2], ...)
	value1 puede ser un data range, celda, valor value2 es opcional e igualmente puede ser data range, celda, valor ... hasta 255 argumentos (<i>value</i>) COUNTA cuenta celdas que contengan <i>cualquier tipo de dato</i> incluyendo <i>error_values</i> y texto vacío ("")
COUNTIF	Conteo de celdas que cumplan una condición específica. Por ejemplo, cuántas escuelas hay en el Municipio de San Germán.
Sintaxis:	COUNTIF(range, criteria)

	<p>Range: grupo de celdas donde se quiere buscar</p> <p>Criteria: qué es lo que se va a buscar. Añadir comillas (") si se busca en cadenas alfanuméricas (texto). Puede añadir <i>wildcards</i> (*) asterisco en el argumento de búsqueda. Por ejemplo: =COUNTIF(F2:F1102,"*MUNOZ RIVERA*") busca en el data range F2:F1102, la cadena alfanumérica *MUNOZ RIVERA*. Esto incluye cualquier celda que contenga la secuencia MUNOZ RIVERA en cualquier parte de la celda.</p>
COUNTIFS	Similar a COUNTIF. Incorpora más condiciones para seleccionar.
Sintaxis:	COUNTIFS(criteria_range1, criteria1, [criteria_range2, criteria2]...)
	A diferencia de COUNTIF, COUNTIFS se usa para añadir múltiples criterios de selección . El resultado debe ser cierto para todas las condiciones incluidas en la fórmula.

SUMIFS

Sumar cantidades basando una o más condiciones.

SUMIFS	Similar a SUMIF. Incorpora más condiciones para seleccionar.	
Sintaxis:	SUMIFS(sum_range, criteria_range1, criteria1, [criteria_range2, criteria2], ...)	
	sum_range:	Grupo de celdas a sumar.
	criteria_range1	Grupo de celdas a comparar con <i>la condición (criteria1)</i> .
	criteria1	Condición para comparar con en el grupo de celdas <i>criteria_range1</i> .
	criteria_range2	Otro grupo de celdas a comparar con <i>la condición (criteria2)</i> .
	criteria2	Condición para comparar con en el grupo de celdas <i>criteria_range2</i> .

Ejemplos usando COUNTIFS y SUMIFS

A	B	C	D	E	F	G
Municipio	POP_EST2016	ZONA		Formulas: COUNTIFS & SUMIFS	Resultado	Explicación
Adjuntas	18,760	Centro		=COUNTIFS(C2:C79,"Metro",A2:A79,"Ca**")	2	el data range C2:C79 = "Metro" & A2:A79 comienza con Ca
Aguada	40,083	Noroeste		=SUMIFS(B2:B79,C2:C79,"Metro",A2:A79,"Ca**")	191,177	Sumar range B2:B79 si C2:C79 lee Metro y A2:A79 comienza con Ca
Aguadilla	56,893	Noroeste				
Aguas Buenas	27,323	Centro-Este				
Aibonito	24,438	Centro-Este				
Añasco	28,250	Oeste				
Arecibo	91,096	Norte				
Arroyo	18,778	Sur				
Barceloneta	24,699	Norte				
Barranquitas	29,548	Centro-Este				
Bayamón	192,974	Metro				
Cabo Rojo	50,096	Suroeste				
Caguas	136,372	Centro-Este				
Camuy	33,481	Norte				
Canóvanas	47,174	Noreste				
Carolina	165,040	Metro				
Cataño	26,137	Metro				

Primera fórmula: COUNTIFS(C2:C79,"Metro",A2:A79,"Ca*")

El data range C2:C79 contiene las zonas. El data range A2:A79 contiene los nombres de los municipios.

Condición	Explicación
C2:C79,"Metro"	En el data range C2:C79 de zonas , seleccionar solamente los récords que pertenezcan a la zona " Metro " (hay 10 seleccionados)
A2:A79,"Ca*"	Que <i>además</i> entre esos 10 seleccionados, el data range A2:A79 , de nombres de los municipios deben comenzar en "Ca*" (Carolina y Cataño)

Segunda fórmula: SUMIFS(B2:B79,C2:C79,"Metro",A2:A79,"Ca*")

sum_range:	Explicación
B2:B79	Data range que contiene números a sumar
Condiciones:	
C2:C79,"Metro"	En el data range C2:C79 de zonas , seleccionar solamente los récords que pertenezcan a la zona " Metro "
A2:A79,"Ca*"	Que <i>además</i> el data range A2:A79 , de nombres de los municipios deben comenzar en "Ca*" (Carolina y Cataño)

Note que los datos alfanuméricos (texto) llevan comillas. Excel no devolverá error; puede devolver cero (0).

Estas fórmulas sirven para resumir datos. Recuerde que puede usar filtros y usar la función AGGREGATE para validar resultados de las funciones SUMIFS y COUNTIFS.

DATE FUNCTIONS: FUNCIONES TEMPORALES, FECHAS Y HORAS

LOS DATOS TEMPORALES SE REGISTRAN COMO NÚMEROS

Las [fechas en Excel](#) no deben escribirse como texto. En realidad Excel las convierte y se registran internamente como **números secuenciales**. Esto quiere decir que **Excel usa un origen arbitrario: día # 1** y continúa **sumando días consecutivamente** a partir de este origen.

Por ejemplo, en Excel, **el día # 1 es el 1ro de enero de 1900**. Por otro lado, las **horas, minutos y segundos** se expresan como **fracciones**. De esta manera se pueden realizar en Excel las **operaciones aritméticas (+-*/) usando fechas y horas**.

DOS SISTEMAS DE FECHAS

Excel puede usar **dos sistemas de fechas**: uno de ellos tiene como fecha de origen el 1ro de enero de **1900**. Este sistema es el que se utiliza con el sistema operativo **Windows**. Por su parte las computadoras **Mac**, usan otro sistema con fecha de origen el 1ro de enero de **1904**. Ambos sistemas tienen números secuenciales diferentes para las fechas posteriores al 1º de enero de 1904. La diferencia es 1,462 días. Excel tiene una **opción** para **intercambiar** el uso de estos **sistemas**. En **Excel 2016: File | Options | Advanced | When calculating this workbook | Use 1904 date system**.

DATE FUNCTION

Esta es una función fundamental para el manejo de fechas. Se recomienda usar la función DATE para registrar fechas, ya que no todas las funciones de Excel reconocen las fechas escritas en texto.

SINTAXIS

DATE(year,month,day)

year	<p>Representa el año de la fecha. Excel interpreta este a partir del sistema de fecha de la computadora. Excel usa como partida el año 1900 por defecto.</p> <ul style="list-style-type: none"> Si escribimos años entre 1900 y 9999 inclusive, ese será el año registrado por Excel. <ul style="list-style-type: none"> Por ejemplo, =DATE(2018,02,28) devolverá February 28, 2018. Si escribimos 0 a 1899 inclusive en el argumento <i>year</i>, Excel le suma ese número a 1900 <ul style="list-style-type: none"> Por ejemplo, =DATE(68,2,29) devolverá 2/29/1968 =DATE(1899,2,29) devolverá 3/1/3799 Si entramos un año negativo o mayor de 9999, Excel devolverá el código de error #NUM!. <p>Recomendación: siempre escriba el año usando cuatro cifras: 2000, 2018, etcétera.</p>
month	<p>Representa el mes de la fecha: de 1 a 12.</p> <ul style="list-style-type: none"> Si escribimos un número mayor de 12, Excel le añade el sobrante de 12 al primer mes del próximo año. <ul style="list-style-type: none"> Por ejemplo, =DATE(2018,18,28) devolverá Friday, June 28, 2019. Añade 6 meses comenzando en enero, 2019 Mes 0: =DATE(2018,0,28) devolverá 12/28/2017 Mes negativo: =DATE(2018,-1,28) devolverá 11/28/2017. Note que la fecha devuelta es noviembre 2017. Por lo tanto, no es restar un mes. El resultado será -1+(-n), donde -n es la cifra del mes negativo.
day	<p>Representa el día de la fecha, de 1 a 31.</p> <ul style="list-style-type: none"> Aplican las mismas reglas del argumento <i>month</i> para números cero y negativos <ul style="list-style-type: none"> Por ejemplo, =DATE(2018,1,31) devolverá 1/31/2018. =DATE(2018,1,32) devolverá 2/1/2018. Devuelve 1ro de febrero, aun cuando el mes se escribió 1. =DATE(2018,1,0) devolverá 12/31/2017. Devuelve 31 de diciembre de 2017 =DATE(2018,1,-1) devolverá 12/30/2017. Note que devuelve 30 de diciembre de 2017

DATEDIF

Calcula el número de días, meses, años entre dos fechas. Esta función está escondida y existe todavía para tener compatibilidad con versiones pasadas. La documentación Microsoft advierte que esta función puede dar errores en algunas circunstancias, especialmente calculando con la opción MD: diferencia de días entre dos meses comparados.

SINTAXIS

DATEDIF(start_date,end_date,unit)

DATEDIF(start_date,end_date,unit)	
start_date	Fecha inicial
end_date	Fecha final
unit	Unidad. El tipo de dato que deseamos encontrar.

unit:

Y	Número de años en el periodo
M	Número completo de meses en el periodo
D	Número de días en el periodo
YM	Diferencia en meses entre dos fechas. Con esta opción, los días y años son ignorados.
YD	Diferencia en días entre dos fechas. Con esta opción, los días entre años son ignorados.

Ejemplos:

A	B	C	D	E	F
	start_date	end_date	Fórmula	Resultado	Descripción
	1/2/2015	1/1/2018	=DATEDIF(B3,C3,"Y")	2	Dos años completados entre la fecha inicial y la final (Falta un día para que sean 3 años)
	1/1/2008	1/1/2009	=DATEDIF(B4,C4,"D")	366	366 días entre la fecha inicial y final. 2008 es año bisiesto.
	6/1/2010	8/15/2012	=DATEDIF(B5,C5,"YD")	75	75 días entre junio 1 y agosto 15. Note que está ignorando los años de estas fechas.

DAYS

Devuelve el número de días entre dos fechas: final e inicio.

SINTAXIS

DAYS(end_date,start_date)

Ejemplo

DAYS(end_date,start_date)				
	A	B	C	D
1	Fechas	Formula	Resultado	Explicación
2	12/31/2011	=DAYS("3/15/2011","2/1/2011")	42	Hay 42 días entre el 15 de marzo y el 1ro de febrero de 2011. Recuerde escribir las fechas con comillas para que Excel las reconozca como tales en las fórmulas.
3	1/1/2011	=DAYS(A2,A3)	364	Hay 364 días entre la fecha final A2 y la de inicio A3.

Notas:

Si las fechas no son escritas correctamente, Excel no las convertirá al formato interno secuencial numérico de fechas. Por lo tanto, los resultados serán erróneos.

NETWORKDAYS

Devuelve el número de días entre dos fechas, descartando los días sábado y domingo. Opcional: Añadir lista (array) de celdas con fechas de días feriados.

SINTAXIS

NETWORKDAYS(start_date, end_date, [holidays])

start_date	Fecha de inicio
end_date	Fecha final
holidays	Feriatos: días libres que no sean sábado o domingo

Ejemplo: Número de días de trabajo y lista de días feriados oficiales para el año 2018

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Fecha	Descripción	Fórmula	Días de trabajo	Días libres	% libre														
2	1/1/2023	Inicio de año 2023	=NETWORKDAYS(A2,A3)	261	104	28.5%	+ NETWORKDAYS(A2,A3) No incluye sábados y domingos													
3	1/1/2024	Inicio de año 2024	=NETWORKDAYS(A2,A3,A9:A28)	246	119	32.6%	+ NETWORKDAYS(A2,A3,A9:A24) No incluye sábados y domingos, ni los días feriados que están en la lista													
4																				
5		Días feriados:	=COUNTA(B9:B28)		20		+ COUNTA() Conteo de celdas no vacías en la lista A9:A25													
6		Días en el año 2023:	=DAYS(A3,A2)	365			+ DAYS() Número de días en el año 2023, sin incluir el 1o de enero 2024													
7																				
8	Feriatos 2018	Descripción																		
9	1/1/2023	Día de Año Nuevo																		
10	1/6/2023	Día de Reyes																		
11	1/16/2023	Natalicio del Dr. Martin Luther King, Jr.																		
12	2/20/2023	Día de los Presidentes y el Día del Prócer y la Mujer Ilustre de Puerto Rico en honor a la vida y obra de: Eugenio María de Hostos, Luis Muñoz Rivera, Ramón Emeterio Betances, Román Baldorioty de Castro, Luis Muñoz Marín.																		
13	3/2/2023	Día de la Ciudadanía Estadounidense																		
14	3/22/2023	Día de la Abolición de la Esclavitud																		
15	4/7/2023	Viernes Santo																		
16	4/17/2023	Natalicio de José de Diego																		
17	5/29/2023	Día de la Conmemoración de los Muertos en la Guerra (Memorial Day)																		
18	6/17/2023	Juneteenth																		
19	7/4/2023	Día de la Independencia de los Estados Unidos																		
20	7/25/2023	Constitución ELA																		
21	7/27/2023	Natalicio de José Celso Barbosa																		
22	9/4/2023	Día del Trabajo																		
23	10/9/2023	Día de la Raza																		
24	11/11/2023	Día del Veterano																		
25	11/19/2023	Día de la Cultura Puertorriqueña y el Descubrimiento de Puerto Rico																		
26	11/23/2023	Día de Acción de Gracias																		
27	12/24/2023	Medio día por motivo de la Noche Buena																		
28	12/25/2023	Día de Navidad																		

En este ejemplo estamos usando la función NETWORKDAYS, a la cual le añadimos una lista de días feriados para 2023.

=NETWORKDAYS(A2,A3). Suma los lunes a viernes del año 2018. La fecha final es 1ro de enero de 2024 para que se pueda incluir el 31 de diciembre de 2023.

=NETWORKDAYS(A2,A3,A9:A24). Suma los lunes a viernes del año 2023. La lista A9:A24 contiene las fechas de los días feriados en Puerto Rico para el año 2023. (16 días). En realidad, son 15.5 días, considerando el medio día por motivo de Noche Buena.

TABLAS EXCEL: RESUMIR DATOS

A medida que va pasando el tiempo, se acumulan cantidades de datos; a eso se le añade la creciente capacidad de guardar más datos que antes en las computadoras. Cuando la información es mucha, es preferible resumirla. El resumir nos da una visión general (*overview*) de las características del conjunto de datos en poco tiempo.

OBJETIVOS

- Explicar la importancia de resumir datos
- Explicar maneras apropiadas de resumir distintos tipos de datos
- Resumir *tablas* Excel usando Total Row

PROPÓSITO DE RESUMIR DATOS

Resumir datos nos ayuda a contestar varias preguntas sobre un conjunto de datos. Por ejemplo:

- Conteo total

- Sumas totales o parciales,
- Medidas de tendencia central: media, mediana, moda,
- Medidas de dispersión: varianza, desviación estándar, coeficiente de variación
- Valores mínimo y máximo

En general, **los resúmenes nos ayudan a entender los datos y comunicar el significado** de ellos a otras personas.

Otras ventajas de resumir datos son:

- Validar trabajo hecho
- Probar hipótesis
- Conocer mejor los datos
- Hacer comparaciones

TOTAL ROW

Excel provee la opción de **añadir** una fila **Total Row** al final de una **tabla**. Con esta opción, se usa la función compuesta **SUBTOTAL()** para añadir conteos, sumatorias, medidas de tendencia central, etcétera a cada columna de la tabla. Estas son las opciones disponibles para Total Row (**SUBTOTAL()**):

No resumir	MIN
AVERAGE	SUM
COUNT	STDEV
COUNTA	VAR y otras...
MAX	

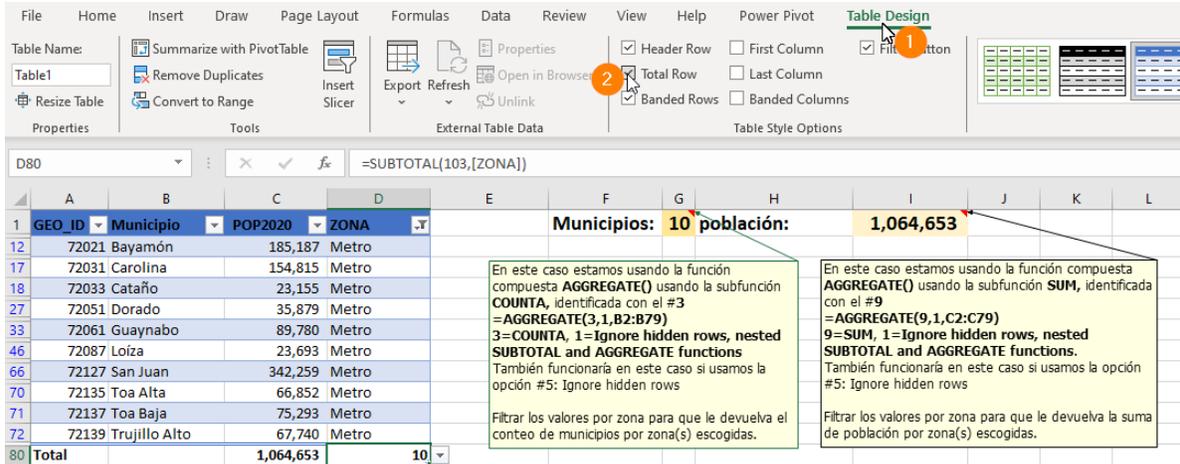
FUNCIONES INCLUIDAS EN COMANDO SUBTOTAL:

Sintaxis

SUBTOTAL(function_num, ref1,[ref2],...)

Esta función tiene los siguientes argumentos:

Función	Código de la función Incluye celdas escondidas	Código de la función Ignora celdas escondidas
AVERAGE	1	101
COUNT	2	102
COUNTA	3	103
MAX	4	104
MIN	5	105
PRODUCT	6	106
STDEV	7	107
STDEVP	8	108
SUM	9	109
VAR	10	110
VARP	11	111



The screenshot shows an Excel spreadsheet with a table containing data for various municipalities. The formula bar displays `=SUBTOTAL(103,[ZONA])`. The table design ribbon is active, showing options for filtering and formatting. Two callout boxes provide detailed explanations of the functions used:

- Left Callout:** Explains the use of the `AGGREGATE()` function with the subfunction `COUNTA` (identified by #3) as `=AGGREGATE(3,1,B2:B79)`. It notes that `3=COUNTA, 1=Ignore hidden rows, nested SUBTOTAL and AGGREGATE functions`. It also mentions that the option #5 (Ignore hidden rows) would also work.
- Right Callout:** Explains the use of the `AGGREGATE()` function with the subfunction `SUM` (identified by #9) as `=AGGREGATE(9,1,C2:C79)`. It notes that `9=SUM, 1=Ignore hidden rows, nested SUBTOTAL and AGGREGATE functions`. It also mentions that the option #5 (Ignore hidden rows) would also work.

Vea que el resultado de las funciones AGGREGATE son los mismos que SUBTOTAL al final de la lista.

Ejemplos usados en este caso:

Fórmula	Función
<code>=SUBTOTAL(103,[Municipio])</code>	<code>103=COUNTA</code>
<code>=SUBTOTAL(109,[POP_EST2016])</code>	<code>109=SUM</code>

PRÁCTICAS CON TABLAS:

Objetivo: Derivar una tabla a partir de datos existentes en un data range.

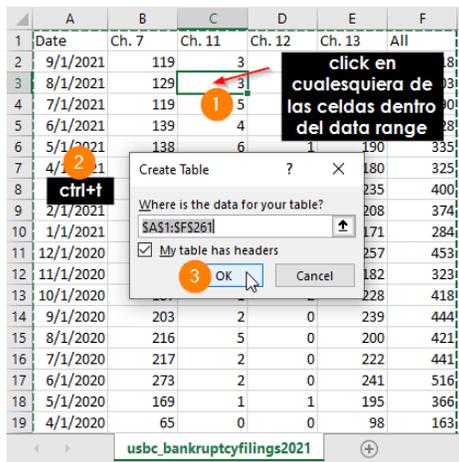
Las tablas pueden generarse a partir de hojas de cálculo que ya tienen datos. En estas usted escoge el grupo de celdas *data range* y luego las convierte a una tabla Excel. La otra manera es mediante la importación de datos externos utilizando un grupo de instrucciones conocido como Query. Para esto Excel provee la interfaz Query Editor.

EJERCICIO 1: CREAR TABLAS

1: Para esta práctica, abra el archivo **usbc_bankruptcyfilings2021.xlsx**, localizado en **C:\Tutorial_Excel\DataAnalysis\Modulo_2**. El archivo contiene los datos solicitudes de quiebras por capítulo presentadas al tribunal federal desde 2000 a octubre de 2021. [Los datos fueron descargados](#) del portal de datos abiertos del Instituto de Estadísticas en noviembre 6 de 2023.

2: Derive una tabla Excel a partir de los datos que están en la hoja de cálculo. Estos son los pasos:

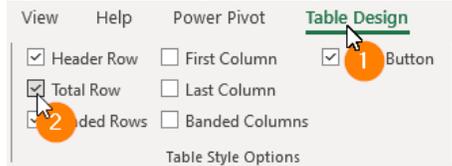
- Siga estos pasos:



Date	Ch. 7	Ch. 11	Ch. 12	Ch. 13	All
9/1/2021	119	3			
8/1/2021	129	3			
7/1/2021	119	5			
6/1/2021	139	4			
5/1/2021	138	6			
4/1/2021					
2/1/2021					
1/1/2021					
12/1/2020					
11/1/2020					
10/1/2020					
9/1/2020	203	2	0	239	444
8/1/2020	216	5	0	200	421
7/1/2020	217	2	0	222	441
6/1/2020	273	2	0	241	516
5/1/2020	169	1	1	195	366
4/1/2020	65	0	0	98	163

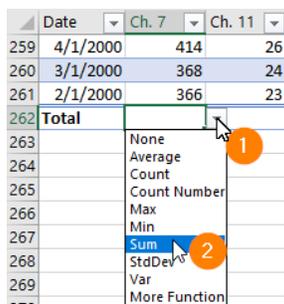
- Click en *cualquiera* de las celdas dentro del data range.
- Pulse las teclas **ctrl+t** simultáneamente.
- Haga click en el botón **OK**

- **Añada** una fila **Total Row** (al final de la tabla) que contenga las sumatorias de las columnas y el promedio de porcentajes en la última columna generada.



- Vaya al **menú principal** y haga click en **Table Design**
- Haga **check** en la opción **Total Row**.

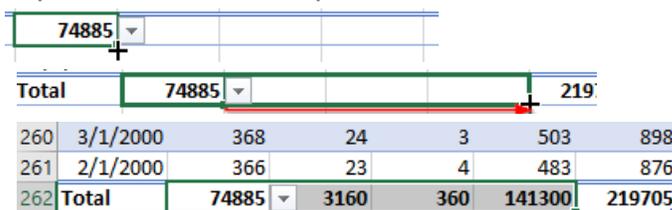
- Añada la sumatoria a las otras columnas. Siga estos pasos:



Date	Ch. 7	Ch. 11	
259	4/1/2000	414	26
260	3/1/2000	368	24
261	2/1/2000	366	23
262	Total		
263			
264			
265			
266			
267			
268			
269			
270			

- Haga **click** en el botón **drop down** (triángulo) al final de la columna Ch. 7.
- Escoja con **click** la opción **Sum**.

- Para añadir la sumatoria a las demás filas hacia la derecha de la columna Ch. 7, haga **click** en el punto verde de la esquina inferior derecha de la celda B262



Total	74885					
260	3/1/2000	368	24	3	503	898
261	2/1/2000	366	23	4	483	876
262	Total	74885	3160	360	141300	219705

- Guarde el archivo (**Save**) para la próxima práctica. Vaya al **menú principal** y escoja **File > Save**.

EJERCICIO 2: TRABAJAR CON TABLAS

Objetivos: En esta parte:

- **Añadiremos una columna**

- **Calcularemos valores** de porcentajes usando una fórmula simple
- **Cambiaremos el formato** de los números para representar cantidades y porcentajes.

- Añada una columna adyacente a la última columna derecha (All).**
Llámela **pctCH7&13**.

	A	B	C	D	E	F	G
1	Date	Ch. 7	Ch. 11	Ch. 12	Ch. 13	All	PctCh7&Ch13
2	9/1/2021	119	3	0	196	318	
3	8/1/2021	129	3	1	170	303	

Esta nueva columna debe contener el **porcentaje** de **quiebras** por **capítulo 7** y **Capítulo 13** (fórmula $= (Ch\ 7 + Ch\ 13) / All$) Ingrese los paréntesis para que no se altere el orden de las operaciones. Recuerde que con tablas Excel, las referencias no son por celdas sino por columnas...

	A	B	C	D	E	F	G	H
1	Date	Ch. 7	Ch. 11	Ch. 12	Ch. 13	All	PctCh7&Ch13	
2	9/1/2021	119	3	0	196	318	$=([Ch. 7]+[Ch. 13])/[All]$	
3	8/1/2021	129	3	1	170	303		
4	7/1/2021	119	5	1	165	290		
5	6/1/2021	139	4	0	185	328		
6	5/1/2021	138	6	1	190	335		
7	4/1/2021	141	4	0	180	325		
8	3/1/2021	161	4	0	235	400		
9	2/1/2021	162	4	0	208	374		
10	1/1/2021	109	2	2	171	284		
11	12/1/2020	194	2	0	257	453		
12	11/1/2020	139	2	0	182	323		

Click en la celda G2.

Escribir =

Abra un paréntesis (

- Haga **click** en la celda **G2**.
Escriba = y **abra** un **paréntesis** (
Haga **click** en la celda **B2**

	A	B	C	D	E	F	G	H
1	Date	Ch. 7	Ch. 11	Ch. 12	Ch. 13	All	PctCh7&Ch13	
2	9/1/2021	119	3	0	196	318	$=([@Ch. 7])$	
3	8/1/2021	129	3	1	170	303		

Verá que Excel registra la columna **Ch. 7** con corchetes [](brackets). Notará también que se añade una arroba @. Esto significa que el cómputo hace referencia solo a la fila.

- Continúe y **añada** un signo de suma +
Haga **click** en la celda **E2**
Cierre con paréntesis) Esto es para que se sume primero las dos celdas y luego haremos la división, fuera de paréntesis.

	A	B	C	D	E	F	G	H
1	Date	Ch. 7	Ch. 11	Ch. 12	Ch. 13	All	PctCh7&Ch13	
2	9/1/2021	119	3	0	196	318	$=([@Ch. 7]+[@Ch. 13])$	

- Continuando, **añada** el signo de división /
Haga **click** en la celda **F2**. Se verá como $[@All]$

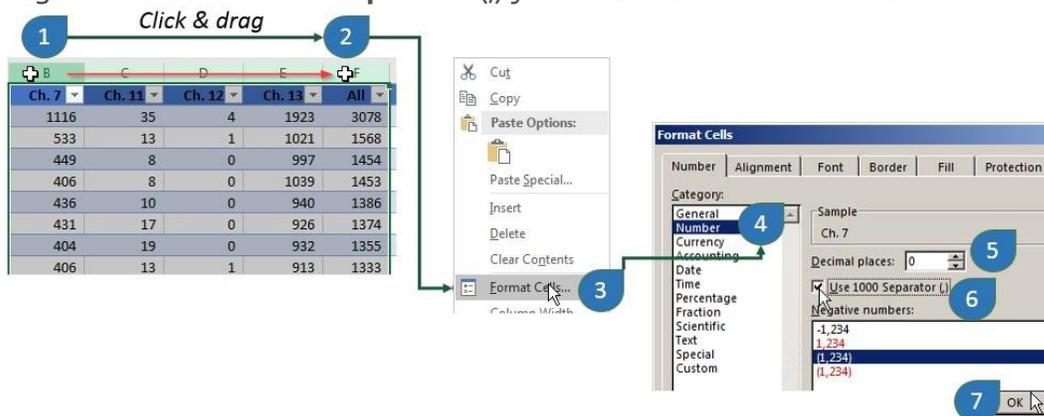
	A	B	C	D	E	F	G	H	I
1	Date	Ch. 7	Ch. 11	Ch. 12	Ch. 13	All	PctCh7&Ch13		
2	9/1/2021	119	3	0	196	318	$=([@Ch. 7]+[@Ch. 13])/[@All]$		

- Vea que el resto de las celdas hacia abajo en esta columna adquieren el valor por la fórmula que se escribió en el paso anterior. Esa es otra ventaja de usar tablas.

	A	B	C	D	E	F	G
1	Date	Ch. 7	Ch. 11	Ch. 12	Ch. 13	All	PctCh7&Ch13
2	9/1/2021	119	3	0	196	318	0.990566038
3	8/1/2021	129	3	1	170	303	0.98679868
4	7/1/2021	119	5	1	165	290	0.979310345
5	6/1/2021	128	4	0	195	279	0.987904970

Cambie el formato a las columnas.

- **Seleccione las columnas B hasta la columna F.**
Haga **right click | Format Cells**.
Haga **check en Use 1000 Separator (,)** y cambie el **número de decimales a cero (0)**.

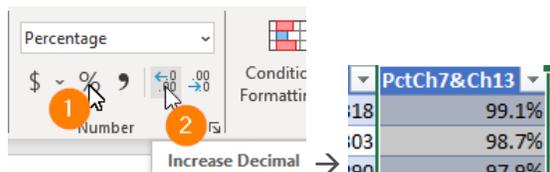


Seleccione la columna **pctCH7&13** y repita el proceso. Esta vez cambie el formato a porcentaje **Percentage** con **un lugar decimal**.

- Haga **click** en la cabecera de la columna G.



- Vaya a la sección **Number**, haga **click** en el **botón de porcentaje** y luego **click** en el **botón Increase Decimal**.

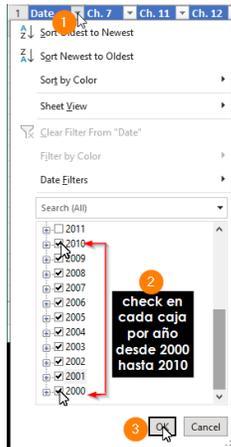


EJERCICIO 3: ANÁLISIS DE DATOS

Objetivo: Con el workbook **usbc_bankruptcyfilings2021.xlsx**, usar herramientas **Filter** y **Total Row** para contestar preguntas sobre casos de quiebras entre los años 2000 y 2021.

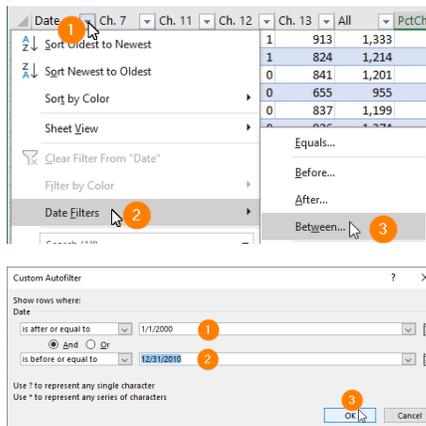
Casos de quiebras: Usando la tabla Excel que acaba de generar: **¿Cuántos casos de quiebras hay entre los años 2000 y 2010?**

- Averigüe usando herramientas **Filter**, y **Total Row**. Informe solamente las **quiebras** por **capítulo 7**, **capítulo 13**, el **promedio de pctCh7&Ch13** y de la columna **All** (todas)
- Primero, **filtre datos por fecha**, incluyendo todos los meses de los años entre **1/1/2000 al 12/31/2010** inclusive. Hay distintas maneras de hacerlo.



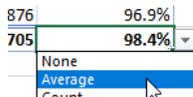
- Haga **click** en el botón de filtro del campo **Date**, columna A.
- Haga **check en cada año, desde 2000 hasta 2010**. NO haga doble click porque hará expandir las opciones dentro del año.
- Haga **click** en el botón **OK**, una vez tenga todos los años desde 2000 hasta 2010 en check.

- Otra manera de filtrar es:



- Hacer click en el botón de filtro.
- Escoger **Date Filters** y luego escoger la opción **Between...**
- En la caja **Custom Autofilter**, escribir **1/1/2010** en la primera caja de texto
- Escribir 12/31/2010** en la segunda caja de texto.

- No olvide asignar el cómputo de promedio (average) al campo PctCh7&Ch13.



Los resultados aparecerán **al final** en la fila **Total**.

- Escriba las cantidades para los años 2000 al 2010.**
- En la columna 2011-2021** escriba las cantidades para esos años. Deberá repetir el proceso, cambiando las fechas a 1/1/2011 a 12/31/2021

Capítulo	2000-2010	2011-2021
Capítulo 7		
Capítulo 13		
Todas (All)		
Porcentaje (Caps 7&13)		

Hagamos una gráfica compuesta con estos datos.

- Guarde** el archivo (Save). Lo vamos a utilizar en otro ejercicio.

EJERCICIO 4: USO DE FUNCIONES SUMIF Y COUNTIF

Abra el workbook **Municipios_2020.xlsx**, localizado en el folder **C:\Tutorial_Excel\DataAnalysis\Modulo_2**.

Complete la siguiente tabla. Utilice las funciones **SUMIF** para sumar habitantes por zona y **COUNTIF** para contar los municipios que pertenecen a cada región. Si va a copiar la fórmula a otras celdas, recuerde usar **referencia absoluta (\$) a la columna o a la fila**, según sea el caso.

ZONAS	POBLACIÓN	cantidad municipios
Metro		
Centro		
Centro-Este		
Centro-Oeste		
Este		
Sureste		
Noreste		
Norte		
Noroeste		
Oeste		
Suroeste		
Sur		

- Para comenzar, haga **click** en la **celda H3**.
- Presione la tecla =** y **escriba sumif**, seguido de pulsar la tecla **tab** para añadir la función SUMIF

G	H	
ZONAS	POBLACIÓN	cantidad
Metro	=SUMIF(
Centro	SUMIF(range, criteria, [sum_range])	

- Luego del paréntesis abierto **escriba** el nombre de la tabla Excel **PopMun** y haga **doble click** en el icono de la tabla **PopMun2020** para registrar esta tabla en la fórmula

G	H	
ZONAS	POBLACIÓN	cantidad
Metro	=SUMIF(PopMun2020	
Centro	SUMIF(range, criteria, [sum_range])	
Centro-Este		

- Pulse la tecla [corchete (bracket) para que aparezca la lista de campos disponibles. Escoja con **doble click** el campo llamado ZONA

ZONAS		
Metro	=SUMIF(PopMun2020[
Centro	SUMIF(range, criteria, [sum_range])	
Centro-Este		
Centro-Oeste		
Este		
Sureste		
Noreste		

@ - This Row

(...) GEO_ID

(...) Municipio

(...) POP2020

(...) ZONA

All

Data

Headers

Totals

- Cierre con el corchete]

ZONAS	POBLACIÓN	cantidad mun
Metro	=SUMIF(PopMun2020[ZONA]	
Centro	SUMIF(range, criteria, [sum_range])	

- Escriba una coma , para el próximo argumento de la fórmula **criteria**:

ZONAS	POBLACIÓN	cantidad mun
Metro	=SUMIF(PopMun2020[ZONA],	
Centro	SUMIF(range, criteria, [sum_range])	

- Escriba** el nombre de la celda, en este caso es la celda **G3** que contiene el nombre **Metro**.

ZONAS	
Metro	=SUMIF(PopMun2020[ZONA],G3
Centro	SUMIF(range, criteria, {sum_range})

- Escriba una coma , para el próximo argumento de la fórmula: **sum_range** o las celdas que vamos a escoger para sumar población en este caso.

ZONAS	
Metro	=SUMIF(PopMun2020[ZONA],G3,
Centro	SUMIF(range, criteria, {sum_range})

- Luego de la coma, **escriba** el nombre de la tabla **PopMun2020** seguido de un [corchete y **escoja** con **doble click** el campo **POP2020**

Metro	=SUMIF(PopMun2020[ZONA],G3,PopMun2020[
Centro	SUMIF(range, criteria, {sum_range})
Centro-Este	
Centro-Oeste	

- Cierre con] corchete de cierre y) paréntesis de cierre

ZONAS	
Metro	=SUMIF(PopMun2020[ZONA],G3,PopMun2020[POP2020])

- Presione **Enter** para terminar.

La celda debe tener la siguiente *sumatoria por criterio* para la zona metro:

ZONAS	POBLACIÓN
Metro	1,064,653

- Repita las fórmulas para el resto de las celdas hacia abajo. Ya sabe cómo hacerlo con un par de clicks. ¿Cierto?

	1,064,653
--	-----------

Añada la cantidad de municipios que componen la zona Metro. La función en este caso no es SUMIF. Deberá usar la función =COUNTIF. Esta tiene menos argumentos. =COUNTIF(range, criteria). Para esta práctica, el *range* es la columna ZONA de la tabla PopMun2020 y el criterio es lo que contiene la celda G3 Metro.

- Escriba la fórmula como aparece en la figura:

ZONA			
Centro	ZONAS	POBLACIÓN	
Noroeste	Metro	1,064,653	=COUNTIF(PopMun2020[ZONA],G3)
Morona	Centro		

- Complete el resto de la tabla.

Preguntas: ¿Se debe/puede o no usar referencias fijas (\$G3 o \$G\$3) en estas fórmulas?
 ¿Por qué sí o por qué no o no es necesario?

¿Qué pasaría si el área que está usando para registrar los datos por zona fuera una tabla Excel?

EJERCICIO 5: FUNCIONES XLOOKUP Y LOOKUP

XLOOKUP

Como se explicó anteriormente, la función XLOOKUP se usa para traer valores de un sitio a otro usando un identificador común para parear y asignar valores a una o más celdas. Trabajaremos un ejemplo con datos censales de población por décadas y municipios.

El workbook tiene dos hojas: una contiene los datos históricos de población y la otra hoja tiene el conteo de población para el Censo de 2020.

Tareas:

El objetivo es practicar **cómo pasar los valores del Censo 2020 a otra hoja que tiene los datos de población histórica sin hacer copy/paste**.

Insertar columna y darle nombre:

Escribir y ejecutar la función XLOOKUP

Como vamos a trabajar con una Tabla Excel, no tendremos que copiar la fórmula al resto de las celdas de la columna.

Abra el archivo **EjercicioLookup.xlsx**, localizado en el folder **C:\Tutorial_Excel\DataAnalysis\Modulo_2**.

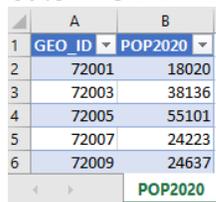
El workbook tiene dos hojas. 

- Deberá abrir primero la hoja **PobSerieHist** así:



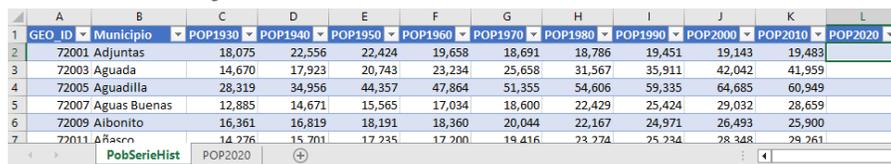
	A	B	C	D	E	F	G	H	I	J	K	L
1	GEO_ID	Municipio	POP1930	POP1940	POP1950	POP1960	POP1970	POP1980	POP1990	POP2000	POP2010	
2	72001	Adjuntas	18,075	22,556	22,424	19,658	18,691	18,786	19,451	19,143	19,483	
3	72003	Aguada	14,670	17,923	20,743	23,234	25,658	31,567	35,911	42,042	41,959	
4	72005	Aguadilla	28,319	34,956	44,357	47,864	51,355	54,606	59,335	64,685	60,949	
5	72007	Aguas Buenas	12,885	14,671	15,565	17,034	18,600	22,429	25,424	29,032	28,659	
6	72009	Aibonito	16,361	16,819	18,191	18,360	20,044	22,167	24,971	26,493	25,900	
7	72011	Añasco	14,276	15,701	17,235	17,200	19,416	23,274	25,234	28,348	29,261	
8	72013	Arecibo	56,525	69,192	75,361	69,879	73,468	86,766	93,385	100,131	96,440	
9	72015	Arroyo	8,199	10,746	12,936	13,315	13,033	17,014	18,910	19,117	19,575	
10	72017	Barceloneta	15,751	18,545	19,897	19,334	20,792	18,942	20,947	22,322	24,816	
11	72019	Barranquitas	14,901	17,096	17,605	18,978	20,118	21,639	25,605	28,909	30,318	
12	72021	Bayamón	29,524	37,190	48,000	72,221	156,192	196,206	220,262	224,044	208,116	

- Vamos a **trasladar** los **conteos de población para 2020** presentes en la hoja **POP2020** a la **columna L** de la hoja **PobSerieHist**.



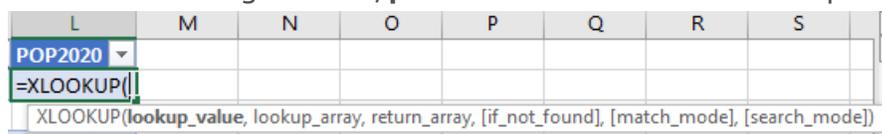
	A	B
1	GEO_ID	POP2020
2	72001	18020
3	72003	38136
4	72005	55101
5	72007	24223
6	72009	24637

- Regrese a la hoja PobSerieHist** y haga **click** en la **primera celda** de la **columna L** y escriba **POP2020**, seguido de **Enter**. Este será el nombre de la columna. Verá que la columna se añadió a la tabla Excel. Se añadirá una columna a la tabla siempre y cuando la nueva columna sea adyacente a la última columna de la derecha.



	A	B	C	D	E	F	G	H	I	J	K	L
1	GEO_ID	Municipio	POP1930	POP1940	POP1950	POP1960	POP1970	POP1980	POP1990	POP2000	POP2010	POP2020
2	72001	Adjuntas	18,075	22,556	22,424	19,658	18,691	18,786	19,451	19,143	19,483	
3	72003	Aguada	14,670	17,923	20,743	23,234	25,658	31,567	35,911	42,042	41,959	
4	72005	Aguadilla	28,319	34,956	44,357	47,864	51,355	54,606	59,335	64,685	60,949	
5	72007	Aguas Buenas	12,885	14,671	15,565	17,034	18,600	22,429	25,424	29,032	28,659	
6	72009	Aibonito	16,361	16,819	18,191	18,360	20,044	22,167	24,971	26,493	25,900	
7	72011	Añasco	14,276	15,701	17,235	17,200	19,416	23,274	25,234	28,348	29,261	

- Haga **click** en la celda **L2** y escriba **=XLOOKUP**. Cuando Excel traiga el nombre de la función con los argumentos, **presione** la tecla **tab**. Esto hará que Excel abra el paréntesis.



	L	M	N	O	P	Q	R	S
1	POP2020							
2	=XLOOKUP(
3	XLOOKUP(lookup_value, lookup_array, return_array, [if_not_found], [match_mode], [search_mode])							

- Para llenar el argumento **lookup_value**, haga **click** en la celda **A2**. La celda A2 tiene el valor de búsqueda que vamos a comparar con las celdas de la tabla que está en otra hoja.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	GEO_ID	Municipio	POP1930	POP1940	POP1950	POP1960	POP1970	POP1980	POP1990	POP2000	POP2010	POP2020		
2	72001	Adjuntas	18,075	22,556	22,424	19,658	18,691	18,786	19,451	19,143	19,483	=XLOOKUP([@[GEO_ID]]		
3	72003	Aguada	14,670	17,923	20,743	23,234	25,658	31,567	35,911	42,042	41,959	XLOOKUP(lookup_value, look		

- **Escriba una coma ,** para continuar con el próximo argumento **lookup_array**.
- Haga **click** en el tab de la hoja **POP2020** y luego **click** en la cabecera de la **columna A**

	A	B	C	D	E	F	G
1	GEO_ID	POP2020					
2	72001	18020					
3	72003	38136					
4	72005	55101					
5	72007	24223					
6	72009	24637					
7	72011	25596					

La columna A de esta tabla contiene los valores que vamos a comparar con los que están en la columna A de la tabla PobSerieHist.

- **Escriba una coma ,** para continuar con el próximo argumento **result_vector**.
- Haga **click** en la cabecera de la **columna B**.

	A	B	C	D	E
1	GEO_ID	POP2020			
2	72001	18020			
3	72003	38136			
4	72005	55101			
5	72007	24223			
6	72009	24637			
7	72011	25596			

La columna B contiene el vector o lista de valores (de población 2020) que deseamos copiar con la condición de que correspondan con los valores que están en la columna A de la tabla PobSerieHist.

- Presione la tecla **Enter** para completar los argumentos de la función XLOOKUP. Los demás argumentos no son necesarios aquí.

Podrá ver que los valores fueron pareados mediante el código GEO_ID y copiados en las celdas correspondientes.

	A	B	C	D	E	F	G	H	I	J	K	L
1	GEO_ID	Municipio	POP1930	POP1940	POP1950	POP1960	POP1970	POP1980	POP1990	POP2000	POP2010	POP2020
2	72001	Adjuntas	18,075	22,556	22,424	19,658	18,691	18,786	19,451	19,143	19,483	18,020
3	72003	Aguada	14,670	17,923	20,743	23,234	25,658	31,567	35,911	42,042	41,959	38,136
4	72005	Aguadilla	28,319	34,956	44,357	47,864	51,355	54,606	59,335	64,685	60,949	55,101
5	72007	Aguas Buenas	12,885	14,671	15,565	17,034	18,600	22,429	25,424	29,032	28,659	24,223
6	72009	Aibonito	16,361	16,819	18,191	18,360	20,044	22,167	24,971	26,493	25,900	24,637
7	72011	Añasco	14,276	15,701	17,235	17,200	19,416	23,274	25,234	28,348	29,261	25,596
8	72013	Arecibo	56,525	69,192	75,361	69,879	73,468	86,766	93,385	100,131	96,440	87,754
9	72015	Arroyo	8,199	10,746	12,936	13,315	13,033	17,014	18,910	19,117	1	#N/A
10	72017	Barceloneta	15,751	18,545	19,897	19,334	20,792					22,657
11	72019	Barranquitas	14,901	17,096	17,605	18,978	20,118					28,983
12	72021	Bayamón	20,534	27,100	48,000	73,231	155,103	105,205	120,253	124,044	108,116	185,187

Preguntas:

Note que hay unas celdas con valores **#N/A**. **¿Cuántas celdas tienen ese aviso de error?, ¿Por qué ocurrió esto?**

LOOKUP

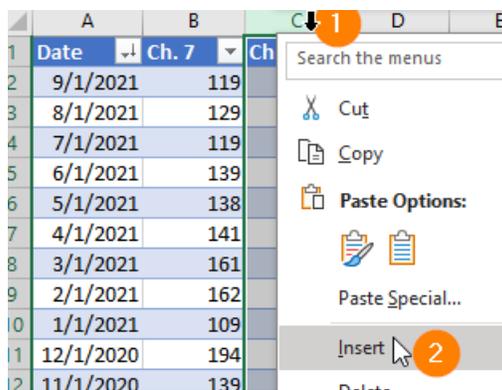
Volvamos a abrir el workbook **usbc_bankruptcyfilings2021.xlsx**, para clasificar datos numéricos, **agrupando por categorías**: 0 a 100, 101 a 200, etcétera.

Tareas:

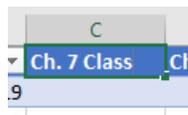
- **Insertar (field) columna nueva** para guardar los grupos de valores (clases) para las quiebras por capítulo 7. Estas “clases” estarán guardadas en formato alfanumérico.
- **Escribir la fórmula** que incluya la función **LOOKUP**. En este ejemplo, escribiremos (*hard coding*) los valores de inicio y final de cada clase, en lugar de tomarlos desde celdas en un data range.

INSERTAR COLUMNA

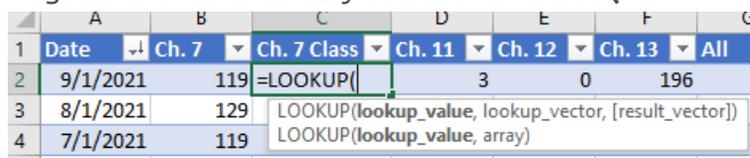
- Abra el archivo **usbc_bankruptcyfilings2021.xlsx**, localizado en el folder **C:\Tutorial_Excel\DataAnalysis\Modulo_2**.
- Haga **right click** en la cabecera de la columna **C** y escoja la opción **Insert** para insertar una columna.



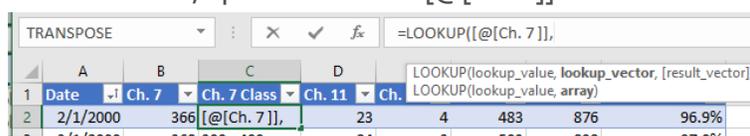
- Cambie el nombre de la columna a **Ch. 7 Class** y presione **Enter**



- Haga **click** en la celda C2 y **escriba =LOOKUP(**



- Haga **click** en la celda **B2**. Ese es el valor que vamos a usar para comparar. Como estamos usando tablas, aparecerá como **[@[Ch. 7]]**



- Escriba una coma ,** para continuar con el próximo argumento **lookup_vector**
- Haga **click** en el tab de la hoja **Classes**.

- En la hoja **Classes**, haga **click** en la celda **B1** y **arrastre el cursor hasta la celda N1**. Ese es el vector de búsqueda de valores.

B1		=LOOKUP([@[Ch. 7]],Classes!B1:N1)																	
	A	B	C	D	E	LOOKUP(lookup_value, lookup_vector, [result_vector])										K	L	M	N
1	lookup_vector	0	100	200	300	LOOKUP(lookup_value, array)										900	1000	1100	1200
2	Result_vector	<100	100 - 200	200 - 300	300 - 400	400 - 500	500 - 600	600 - 700	700 - 800	800 - 900	900 - 1,000	1,000 - 1,100	1,100 - 1,200	1,200 and up					

Este vector debe ser fijo.

Para fijar el vector, debemos **insertar el símbolo \$ al lado izquierdo de la letra de la columna y al lado izquierdo del número de las celdas que componen la lista (el vector)**.

- Para **añadir referencia fija**, presione la tecla **F4**. Notará que aparece como \$B para fijar la columna de inicio y \$1 para fijar la fila de inicio, de manera análoga se fija también en la celda final del vector \$N\$1.

```
=LOOKUP([@[Ch. 7]],Classes!$B$1:$N$1)
```

- **Escriba una coma ,** para continuar con el próximo argumento **result_vector**
- Haga **click** en la celda **B2** y **arrastre hasta la celda N2**. Luego **presione** la tecla **F4** para fijar las celdas de este vector como aparece aquí.

B2		=LOOKUP([@[Ch. 7]],Classes!\$B\$1:\$N\$1,Classes!\$B\$2:\$N\$2)																	
	A	B	C	D	E	LOOKUP(lookup_value, lookup_vector, [result_vector])										K	L	M	N
1	lookup_vector	0	100	200	300	LOOKUP(lookup_value, array)										900	1000	1100	1200
2	Result_vector	<100	100 - 200	200 - 300	300 - 400	400 - 500	500 - 600	600 - 700	700 - 800	800 - 900	900 - 1,000	1,000 - 1,100	1,100 - 1,200	1,200 and up					

- Presione **Enter** para cerrar los argumentos de esta función.

Aparecerán las clases en las celdas que componen esta columna.

C3		=LOOKUP([@[Ch. 7]],Classes!\$B\$1:\$N\$1,Classes!\$B\$2:\$N\$2)																
	A	B	C	D	E	F	G	H	I									
1	Date	Ch. 7	Ch. 7 Class	Ch. 11	Ch. 12	Ch. 13	All	PctCh7&Ch13										
2	2/1/2000	366	300 - 400	23	4	483	876	96.9%										
3	3/1/2000	368	300 - 400	24	3	503	898	97.0%										
4	4/1/2000	414	400 - 500	26	0	536	976	97.3%										
5	5/1/2000	358	300 - 400	25	1	511	895	97.1%										

Note que **las clases resultantes tienen un espacio al inicio**. Esto se usa para que Excel ordene (Sort) correctamente las clases. Al usar alfanuméricos, el 1 y 10, 11 hasta el 19 van primero que el 2. Por lo tanto, es recomendable poner espacios en los dígitos según la cantidad de espacios que tengan las cifras. En este caso, solo falta un espacio.

Así se verá la tabla Excel luego de aplicar la función. Automáticamente calculará los valores en las demás celdas inferiores de la columna **Ch. 7 CLASS**.

	B	C	D
h	Ch. 7	Ch. 7 CLASS	Ch. 11
2005	1,116	1,100-1,200	35
2005	533	500-600	13
2003	449	400-500	8
2001	406	400-500	8
2004	436	400-500	10
2001	421	400-500	17

- **Esto concluye el ejercicio. Guarde este workbook Excel (Save).**

MÓDULO 3: PIVOTTABLES Y PIVOTCHARTS

RESUMEN

Una **tabla Pivot** es una herramienta para resumir y analizar datos, especialmente si son voluminosos. Es una de las herramientas más útiles de Excel para el análisis de datos desde diferentes perspectivas. Una **gráfica Pivot** es igual a cualquier otra gráfica Excel. La diferencia es que **las PivotCharts están ligadas a una tabla Pivot**.

DIMENSIONES Y MÉTRICAS EN PIVOTTABLES

Para entender un poco mejor, imagínese un conjunto de datos, un data range, una Tabla Excel. A menudo los datos contienen descripciones y mediciones. Algunos llaman **dimensiones** o **perspectivas** a estas **descripciones**. Las **métricas** son mediciones: conteos, volúmenes, sumatorias, proporciones **numéricas** que componen la segunda parte de estos datos. Algunos conjuntos de datos tienen mayor número de dimensiones que métricas y viceversa.

DIMENSIONES						MÉTRICAS						
ESCUELA	DIRECCION	MUNICIPIO	DIRECCION	ZIPCODE	ZONA	NIVEL ORIGINAL	GRADOS POSTERIOR	MATRICULA	ESPAÑOL	MATEMATICAS	INGLES	BAJO NIVEL
COTTO (ANEXO)	ARECIBO		00613-9582		URBANA	TODOS LOS NIVELES	EEE - EES	26	0.8	40%	20%	72%
JARDINES DE CAGUAS	CAGUAS		7267396		URBANA	ELEMENTAL	EEE	51	0.4286	100%	14%	73%
SU REXFORD GUY TUGWELL	CAYEY		00737-0000		RURAL	SEGUNDA UNIDAD	K - 9	52	0.7143	71%	86%	68%
PRE-VOCACIONAL CASIANO CEPEDA	RIO GRANDE		7450000		RURAL	SECUNDARIO	EEL - EES	53	0.1667	0	67%	90%
JOSE GUALBERTO PADILLA	CAYEY		7370000		RURAL	ELEMENTAL	K - 6	54	1	1	88%	87%
TOMAS VERA AYALA	PATILLAS		00723-0000		RURAL	SEGUNDA UNIDAD	K - 9 , EEE	60	0.4545	0.5455	27%	96%
JOAQUIN FERRAN LLUIS	PONCE		00732-7105		URBANA	ELEMENTAL	K - 6 , EEE	61	0.7143	0.7317	54%	90%
MANUEL CRUZ MACEIRA	COMERIO		00782-0000		URBANA	ELEMENTAL	1-6	65	0.4	0%	30%	54%
LENA M FRANCESCH - RUBIAS KM 2.	YAUCO		-698		RURAL	ELEMENTAL	K - 6	71	0.6327	61%	43%	100%
AVELINO PENNA REYES	HUMACAO		00791-0000		URBANA	SECUNDARIO	EEL - EES	72	0.2857	71%	23%	68%
PASO PALMA I	UTUADO		00641-0000		RURAL	ELEMENTAL	K - 6	75	0.8056	0.9167	100%	92%
VICTOR PARES COLLAZO	SAN JUAN		00918-0000		URBANA	TODOS LOS NIVELES	PK - EEE - EES	77	0.1538	27%	8%	62%
LAMBOGLIA	PATILLAS		00723-0000		RURAL	ELEMENTAL	K - 6	77	0.2329	0.0685	21%	92%
DAMIAN ABAJO	OROCOVIS		00720-0000		RURAL	ELEMENTAL	K - 6	81	1	1	91%	83%
INDIERA FRIA	MARICAO		00606-0000		RURAL	ELEMENTAL	PK - 6 , EEE	81	0.9118	0.7059	71%	93%
CARMEN APONTE	UTUADO		-611		RURAL	ELEMENTAL	K - 6	81	0.9429	97%	77%	78%
JUANITA RIVERA ALBERT	VIEQUES		7659712		RURAL	ELEMENTAL	PK - 1 , EEE	82				92%
MARIA C HUERTAS	COMERIO		00782-0000		RURAL	ELEMENTAL	K - 6 , EEE	84	0.9444	1	100%	81%
BRYAN	LAS MARIAS		00670-9069		RURAL	ELEMENTAL	K - 6	84	0.4545	0.6818	61%	94%
CRISTOBAL VICENS	CIALES		00638-0000		RURAL	ELEMENTAL	K - 6	85	0.2927	41%	22%	76%
LUDOVICO COSTOSO	BAYAMON		00959-5259		RURAL	TODOS LOS NIVELES	EEE - EES	86	0.2593	0.5185	21%	81%

Dimensiones y mediciones (*métricas* también) en una tabla o un data range.

Las dimensiones pueden verse como diferentes perspectivas para organizar los datos. Las PivotTables utilizan estas dimensiones para agrupar y resumir las mediciones mediante conteos, sumatorias, promedios, etcétera.

NOTA DE PRECAUCIÓN: CELDAS VACÍAS EN UN DATA RANGE O UNA TABLA

Tenga cuidado cuando haya celdas vacías dentro del data range o una tabla. Cuando hay espacios vacíos, Excel presume que los datos no son numéricos.

UTILIDAD DE LAS TABLAS PIVOT

Además de resumir datos, las PivotTables pueden usarse para organizar, filtrar, resumir y analizar datos voluminosos. Por ejemplo, resumir datos de una tabla de inventario de escuelas para saber cuántas escuelas hay por región, distrito, municipio, zona, nivel, ... Además, Excel permite hacer columnas de datos nuevos en las PivotTables, basadas en operaciones matemáticas.

REGION	DISTRITO	MUNICIPIO_ESCOLAR	CODIGO	ESCUELA	DIRECCION_MUNICIPIO	DIRECCION_ESCUELA
ARECIBO	ARECIBO	ARECIBO II	10173	COTTO (ANEXO)	ARECIBO	00782-0000
CAGUAS	GURABO	CAGUAS I	23150	JARDINES DE CAGUAS	CAGUAS	00723-0000
CAGUAS	CIDRA	CAYEY	21345	SU REXFORD GUY TUGWELL	CAYEY	00724-0000
HUMACAO	CANOVANAS	RIO GRANDE	35543	PRE-VOCACIONAL CASAS DE CEPEDA/RIO GRANDE	RIO GRANDE	00725-0000
CAGUAS	CIDRA	CAYEY	21303	JOSE GUALBERTO PARRALES	CAYEY	00724-0000
HUMACAO	YABUCOA	PATILLAS	25189	TOMAS VERA AYALA	PATILLAS	00723-0000
PONCE	PONCE	PONCE II	52175	JOAQUIN FERRANDEZ	PONCE	00726-0000
CAGUAS	BARRANQUITAS	COMERIO			COMERIO	00782-0000
PONCE	YAUCO	YAUCO			YAUCO	00782-0000
HUMACAO	LAS PIEDRAS	HUMACAO			HUMACAO	00791-0000
PONCE	UTUADO	UTUADO			UTUADO	00641-0000
SAN JUAN	SAN JUAN I	SAN JUAN II	65953	VICTOR PARES COLLAZO	SAN JUAN	00918-0000
HUMACAO	YABUCOA	PATILLAS	25080	LAMBOGLIA	PATILLAS	00723-0000
BAYAMON	OROCOVIS	OROCOVIS	12567	DAMIAN ABAJO	OROCOVIS	00720-0000
MAYAGUEZ	MAYAGUEZ	MARICAO	41913	INDIERA FRIA	MARICAO	00606-0000
PONCE	UTUADO	UTUADO	13094	CARMEN APONTE	UTUADO	00611-0000
HUMACAO	FAJARDO	VIEQUES	32300	JUANITA RIVERA ALBERT	VIEQUES	00659-0000
CAGUAS	BARRANQUITAS	COMERIO	21782	MARIA C HUERTAS	COMERIO	00782-0000
MAYAGUEZ	MAYAGUEZ	LAS MARIAS	41632	BRYAN	LAS MARIAS	00670-9069

DATOS CRUDOS

PROCESADOR:
Interfaz PivotTable Fields

REGION	DISTRITO	MUNICIPIO_ESCOLAR	Sum of # escuelas	Sum of MATRICULA_TOTAL	Avg BAJO_NIVEL_POBREZA	Avg Competencia INGLÉS
SAN JUAN			168	54,645	82.4%	35.4%
	CAROLINA		30	12,911	82.0%	32.7%
		CAROLINA II	17	7,476	81.1%	34.5%
		CAROLINA I	13	5,435	83.2%	30.4%
	GUAYNABO		32	10,937	80.0%	43.0%
		GUAYNABO	17	5,964	78.3%	43.9%
		TRUJILLO ALTO	15	4,973		
	SAN JUAN I		44	12,847		
		SAN JUAN II	21	6,779		
		SAN JUAN I	23	6,068		
	SAN JUAN II		62	17,950		
		SAN JUAN IV	26	7,020		
		SAN JUAN III	22	6,725	87.3%	29.7%
		SAN JUAN V	14	4,205	87.0%	24.3%
ARECIBO			134	50,011	81.3%	49.4%

PivotTable:
Resultados devueltos por el procesador

Con esta herramienta podemos contestar preguntas que no pueden ser contestadas fácilmente usando solamente filtros y fórmulas en Excel.

INTERFAZ PIVOTTABLE FIELDS

Antes de comenzar a trabajar con PivotTables, debemos entender la interfaz **PivotTable Fields**. A través de esta interfaz escogemos cuáles serán las **dimensiones** que queremos usar para resumir datos, las **mediciones** que vamos a calcular y qué **filtros** vamos a aplicar para separar unos datos de otros, si es necesario.

- Tools** para cambiar el look-and-feel de la interfaz PivotTable Fields
- Columnas originales** de la tabla o data range
- Filtros** aplicables al PivotTable
- Columnas nuevas** que serán parte del PivotTable
- Filas** que contienen las dimensiones o perspectivas para resumir/agrupar datos
- Valores de las celdas** en el PivotTable en una o más columnas

DISPOSICIÓN DE LOS DATOS PARA QUE FUNCIONEN LAS TABLAS PIVOT

A continuación presentamos un resumen de cómo deben estar dispuestos los datos para el mejor funcionamiento del mecanismo de tablas Pivot.

Apropiado	Inapropiado
<ul style="list-style-type: none"> • Rectangular: variables en columnas observaciones en filas • Sin formateo extra • Contiene solo dimensiones y mediciones • Headers (cabeceras de columnas) claras y precisas • No debe tener headers extra, footers (calce), subtotales, ni campos calculados 	<ul style="list-style-type: none"> • Datos traspuestos (TRANSPOSE) variables como filas, observaciones como columnas • Formateo innecesario • Contiene campos calculados • Nombres de columnas/headers ambiguos • Filas extra en las cabeceras de columnas

PIVOTTABLE TOOLS GROUP TABS

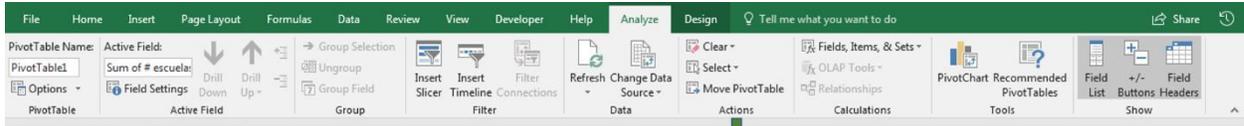
HERRAMIENTAS DEL TAB ANALYZE:



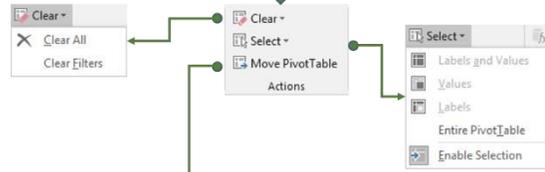
Grupos:

PivotTable	Opciones de configuración y nombre del PivotTable
Active Field	Opciones de configuración de campos o columnas (fields)
Group	Para agrupar o desagrupar datos
Filter	Seleccionar/separar datos
Data	Opciones para cambiar la fuente de datos y traer a memoria los datos (refresh)
Actions	Opciones para borrar datos en memoria, seleccionar medidas y labels y mover la tabla Pivot
Calculations	Cálculo de nuevas columnas (fields, items and sets) y relaciones entre tablas
Tools	Generar PivotCharts y recomendaciones para hacer otros PivotTables
Show	Opciones para configurar listas de campos, botones y cabeceras de columnas

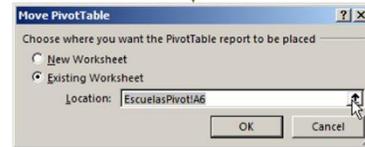
Botones del grupo **Actions**: Mover tabla Pivot, limpiar contenido, seleccionar valores, filas o la tabla completa.



Limpiar filtros y borrar el contenido de la tabla Pivot



Seleccionar valores y filas o la tabla Pivot completa



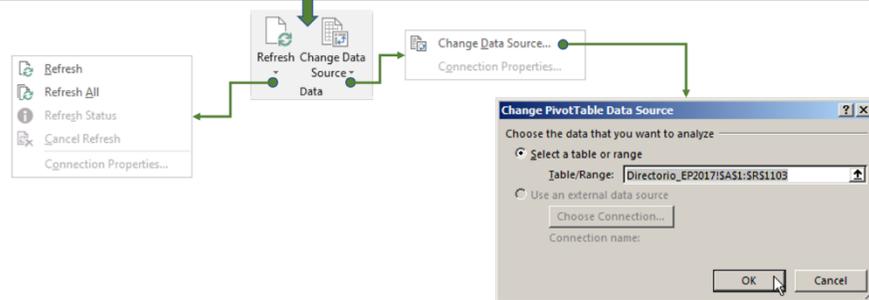
HERRAMIENTAS DEL TAB DESIGN



GRUPOS:

Layout	Opciones para cambiar la apariencia de la tabla Pivot en subtotales, totales, filas vacías.
PivotTable Style Options	Opciones para incluir cabeceras de filas y columnas y sombreado de filas y columnas

Botones del grupo **Data | Refresh** de los datos cuando haya cambios en los datos crudos, cambiar la fuente de datos.



Refresh actualiza trayendo los datos del grupo de celdas (data range) a la tabla Pivot.

CÓMO FUNCIONA UNA TABLA PIVOT

Para generar una tabla Pivot, Excel selecciona y agrupa los datos necesarios para devolver un resultado, basado en las dimensiones que escoja el usuario. En este caso, se está resumiendo (mediante sumatoria) la población 2016 por zonas que incluyen cierto número de municipios.

1 Detectar/evaluar coordenadas

- Zona: **Centro**
- Métrica: **POP_EST2016**
- Filtro: POP_EST2016 **todos los valores** (All)

POP_EST2016 (All)	
Row Labels	Sum of POP_EST2016
Centro	202,174
Centro-Este	451,902
Centro-Oeste	83,174
Este	163,138
Metro	1,135,058
Noreste	133,565
Noroeste	179,184
Norte	364,146
Oeste	141,222
Sur	478,533
Sureste	65,311
Suroeste	131,978
Grand total	3,529,385

Municipio	POP_EST2016	ZONA
Corozal	35,484	Centro
Morovis	32,087	Centro
Utua	30,869	Centro
Naranjito	29,467	Centro
Orocovis	22,258	Centro
Adjuntas	18,760	Centro
Ciales	17,644	Centro
Jayuya	15,605	Centro
Caguas	136,372	Centro-Este
Gurabo	46,990	Centro-Este
Cayey	46,071	Centro-Este
Cidra	41,782	Centro-Este

Excel va a la tabla o data range y aísla los datos relevantes

2 Operación aritmética

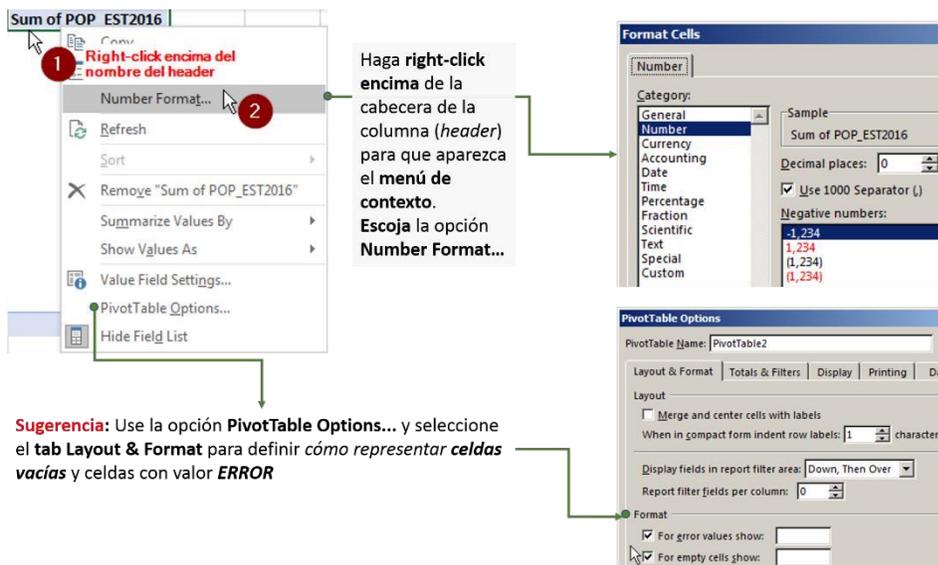
- Summarize values by: **SUM**
- Puede usarse COUNT, MAX, MIN, y otras

3 Mostrar resultados

- $35,484 + 32,087 + 30,869 + 29,467 +$
- $22,258 + 18,760 + 17,644 + 15,605 =$ **202,174**

FORMATEAR NÚMEROS

Como se mencionó anteriormente, es preferible darle un formato claro a los números para que se entienda qué representan, especialmente si son cantidades, moneda, porcentajes o proporciones. Escoja la opción apropiada para sus números, así como la cantidad de decimales que usted especifique.



1 Right-click encima del nombre del header

2 Haga right-click encima de la cabecera de la columna (header) para que aparezca el menú de contexto. Escoja la opción **Number Format...**

Sugerencia: Use la opción **PivotTable Options...** y seleccione el tab **Layout & Format** para definir cómo representar celdas vacías y celdas con valor **ERROR**

Format Cells

Category: **Number**

Sample: Sum of POP_EST2016

Decimal places: 0

Use 1000 Separator (,)

Negative numbers: -1,234; 1,234; (1,234); (1,234)

PivotTable Options

PivotTable Name: PivotTable2

Layout & Format | Totals & Filters | Display | Printing | Data

Layout

Merge and center cells with labels

When in compact form indent row labels: 1 character(s)

Display fields in report filter area: Down, Then Over

Report filter fields per column: 0

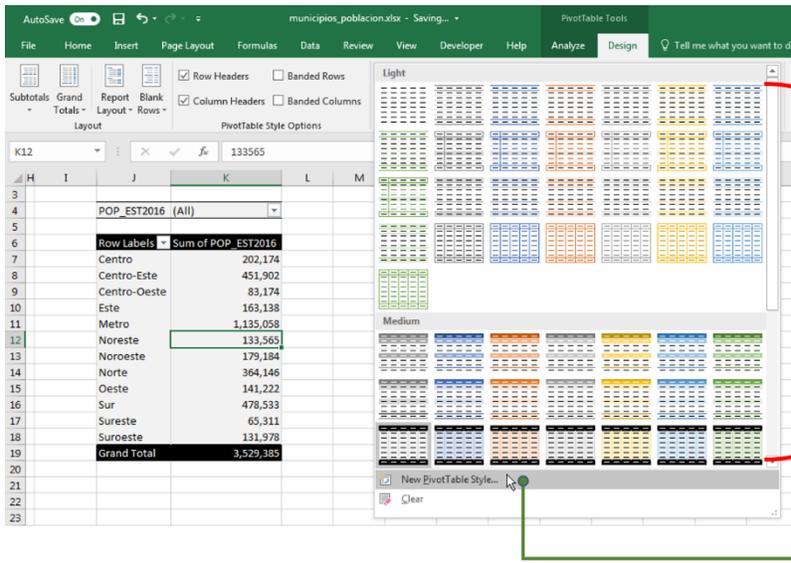
Format

For error values show:

For empty cells show:

ESTILOS PARA TABLAS PIVOT

Excel le provee múltiples opciones para cambiar los colores y las líneas de las tablas. Usted puede escoger de la lista o desarrollar un estilo particular y guardarlo.



The screenshot shows the Excel PivotTable Design tab with various style options. A red bracket highlights the 'Light' and 'Medium' style categories. A green arrow points from the 'New PivotTable Style...' button to the 'New PivotTable Style' dialog box.

Row Labels	Sum of POP_EST2016
Centro	202,174
Centro-Este	451,902
Centro-Oeste	83,174
Este	163,138
Metro	1,135,058
Noreste	133,565
Noroeste	179,184
Norte	364,146
Oeste	141,222
Sur	478,533
Sureste	65,311
Suroeste	131,978
Grand Total	3,529,383

Seleccione entre la variedad de estilos o desarrolle uno y guárdelo para aplicarlo a otras tablas

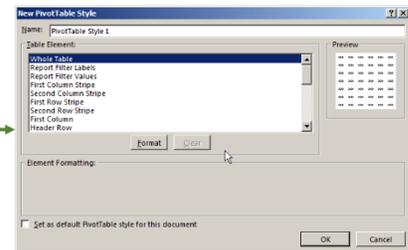
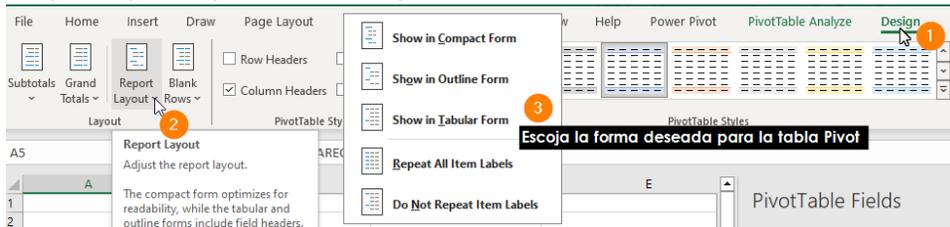


TABLE LAYOUTS: ORGANIZACIÓN DE LA TABLA

Podemos cambiar el diseño de la tabla Pivot para hacerla más legible o prepararla para usarla como tabla de datos crudos para otros procesos de análisis o compartir los datos. Para el análisis de datos, es preferible usar la forma extendida o de bosquejo (*outline*).

Para cambiar el diseño de la tabla Pivot, puede hacerlo desde el **menú principal | Design | Report Layout** y escoja el tipo de diseño que desee.



Row Labels	Sum of # escuelas	Sum of MATRICULA_TOTAL	Avg BAJO_NIVEL_POBREZA	Avg Competencia INGLÉS	Densidad Estudiantil
SAN JUAN	168	54,645	82.4%	35.4%	325.27
CAROLINA	30	12,911	82.0%	32.7%	430.37
CANOVANAS	2	750	80.7%	26.4%	375.00
CAROLINA	28	12,161	82.1%	33.2%	434.32
GUAYNABO	32	10,937	80.0%	43.0%	341.78
CAGUAS	2	441	86.6%	66.3%	220.50
CATANO	2	509	60.1%	48.9%	254.50
GUAYNABO	12	4,547	79.5%	40.8%	378.92
SAN JUAN	2	733	80.2%	25.8%	366.50
TRUJILLO ALTO	14	4,707	82.2%	43.3%	336.21
SAN JUAN I	44	12,847	76.4%	38.8%	291.98
SAN JUAN	44	12,847	76.4%	38.8%	291.98
SAN JUAN II	62	17,950	88.0%	30.3%	289.52
GUAYNABO	1	413	94.8%	23.9%	413.00
SAN JUAN	60	17,008	88.2%	30.2%	283.47
TRUJILLO ALTO	1	529	69.8%	43.7%	529.00
ARECIBO	134	50,011	81.3%	49.4%	373.22

Forma compacta: (default)

Los campos están anidados en la primera columna. Permite aplicar solo un filtro personalizado (custom filter)

REGION	DISTRITO	DIRECCION_MUNICIPIO	Sum of # escuelas	Sum of MATRICULA_TOTAL	Avg BAJO_NIVEL_POBREZA	Avg Competencia INGLÉS	Densidad Estudiantil
SAN JUAN			168	54,645	82.4%	35.4%	325.27
CAROLINA			30	12,911	82.0%	32.7%	430.37
CANOVANAS			2	750	80.7%	26.4%	375.00
CAROLINA			28	12,161	82.1%	33.2%	434.32
GUAYNABO			32	10,937	80.0%	43.0%	341.78
CAGUAS			2	441	86.6%	66.3%	220.50
CATANO			2	509	60.1%	48.9%	254.50
GUAYNABO			12	4,547	79.5%	40.8%	378.92
SAN JUAN			2	733	80.2%	25.8%	366.50
TRUJILLO ALTO			14	4,707	82.2%	43.3%	336.21
SAN JUAN I			44	12,847	76.4%	38.8%	291.98
SAN JUAN			44	12,847	76.4%	38.8%	291.98
SAN JUAN II			62	17,950	88.0%	30.3%	289.52
GUAYNABO			1	413	94.8%	23.9%	413.00
SAN JUAN			60	17,008	88.2%	30.2%	283.47
TRUJILLO ALTO			1	529	69.8%	43.7%	529.00
ARECIBO			134	50,011	81.3%	49.4%	373.22

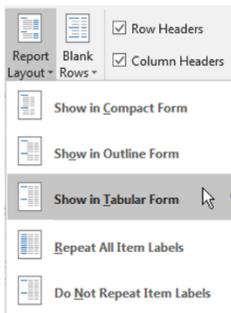
Forma Outline: Bosquejo (recomendada)

Los campos aparecen cada uno en su columna. Permite aplicar filtros a cada columna tipo label: (las que tienen los botones de filtrado)

Las columnas que permiten las opciones para ordenar son las que aparecen con los botones de filtrado.

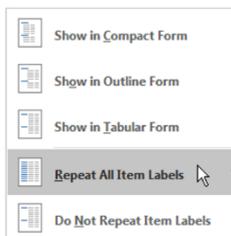
FORMA TABULAR

Estas opciones son útiles para convertir una tabla Pivot en una tabla “rectangular”; primero estirando (*flattening*) la tabla, expandiendo las columnas anidadas en nuevas columnas.



Forma tabular: (sin repeticiones)

REGION	DISTRITO	Sum of # escuelas	Sum of MATRICULA_TOTAL	Avg BAJA_NIVEL_POBREZA
SAN JUAN	CAROLINA	30	12,911	82.0%
	GUAYNABO	32	10,937	80.0%
	SAN JUAN I	44	12,847	76.4%
	SAN JUAN II	62	17,950	88.0%
SAN JUAN Total		168	54,645	82.4%
ARECIBO	CAMUY	27	11,303	79.1%
	VEGA ALTA	37	14,049	82.7%
	ARECIBO	36	13,321	80.1%
	MANATI	34	11,338	82.9%
ARECIBO Total		134	50,011	81.3%
MAYAGUEZ	SAN SEBASTIAN	40	14,054	79.0%
	AGUADILLA	46	14,876	81.9%
	CABO ROJO	40	12,731	78.9%
	MAYAGUEZ	43	11,429	83.4%
MAYAGUEZ Total		169	53,090	80.9%



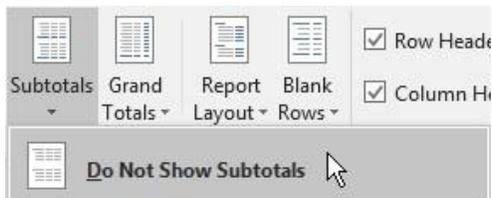
Forma tabular: (con repeticiones)

REGION	DISTRITO	Sum of # escuelas	Sum of MATRICULA_TOTAL	Avg BAJA_NIVEL_POBREZA
SAN JUAN	CAROLINA	30	12,911	82.0%
SAN JUAN	GUAYNABO	32	10,937	80.0%
SAN JUAN	SAN JUAN I	44	12,847	76.4%
SAN JUAN	SAN JUAN II	62	17,950	88.0%
SAN JUAN Total		168	54,645	82.4%
ARECIBO	CAMUY	27	11,303	79.1%
ARECIBO	VEGA ALTA	37	14,049	82.7%
ARECIBO	ARECIBO	36	13,321	80.1%
ARECIBO	MANATI	34	11,338	82.9%
ARECIBO Total		134	50,011	81.3%
MAYAGUEZ	SAN SEBASTIAN	40	14,054	79.0%
MAYAGUEZ	AGUADILLA	46	14,876	81.9%
MAYAGUEZ	CABO ROJO	40	12,731	78.9%
MAYAGUEZ	MAYAGUEZ	43	11,429	83.4%
MAYAGUEZ Total		169	53,090	80.9%

Sugerencia: Use la opción **Tabular con repeticiones** para crear una tabla nueva como “datos crudos”.

Más aún, podemos rellenar los espacios vacíos repitiendo los valores “de cabecera”. Esto es útil para *hacer la transición de un PivotTable a una tabla de una base de datos*. Deberá eliminar también las filas Subtotales, si le interesa convertir la tabla en una tabla de datos crudos *raw dataset*. Esto aparece en el menú **PivotTable Tools** en **Layout | SubTotals | Do not Show Subtotals**.

Eliminar subtotales:



REGION	DISTRITO	Sum of # escuelas	Sum of MATRICULA_TOTAL	Avg BAJA_NIVEL_POBREZA
SAN JUAN	CAROLINA	30	12,911	82.0%
SAN JUAN	GUAYNABO	32	10,937	80.0%
SAN JUAN	SAN JUAN I	44	12,847	76.4%
SAN JUAN	SAN JUAN II	62	17,950	88.0%
ARECIBO	CAMUY	27	11,303	79.1%
ARECIBO	VEGA ALTA	37	14,049	82.7%
ARECIBO	ARECIBO	36	13,321	80.1%
ARECIBO	MANATI	34	11,338	82.9%
MAYAGUEZ	SAN SEBASTIAN	40	14,054	79.0%
MAYAGUEZ	AGUADILLA	46	14,876	81.9%
MAYAGUEZ	CABO ROJO	40	12,731	78.9%
MAYAGUEZ	MAYAGUEZ	43	11,429	83.4%

FORMATEO CONDICIONAL (CONDITIONAL FORMATTING)

Las tablas Pivot pueden contener celdas con formateo condicional (gráfico y colores) de la misma manera que las tablas Excel y data ranges. Para representar una columna numérica usando barras, por ejemplo, se debe duplicar la columna para producir el gráfico. Estos son los pasos:

1 Seleccionar celdas

2 Escoger una opción: **Conditional Formatting**

3 Esconder números; Right click **Number Format...**

4 ..En **Category:** | **Custom**, escriba tres punto-y-comas **;;;**

5 Resultado

Opciones para dar formato gráfico condicional

Gráfica	Densidad Estudiantil	Gráfica
3	7%	430.37
4	0%	341.78
3	8%	291.98
3	3%	289.52
5	5%	418.63
3	9%	379.70
5	2%	370.03
5	0%	333.47
4	6%	351.35
5	7%	323.39
3	1%	318.28
4	7%	265.79
4	6%	358.22
3	3%	354.43
3	4%	294.61
3	4%	285.29

Note que en el paso #4 hacemos que los números sean transparentes, dejando ver en el paso #5, la gráfica limpia sin números.

ORDENAR Y FILTRAR

Al igual que en tablas y grupos de celdas, Excel provee herramientas para ordenar y filtrar datos en tablas Pivot. Hay opciones simples y otras más complejas para escoger (filtrar) valores, ya sean textuales (*Label Filters*) o numéricos (*Value Filters*).

1: Click en botón para ordenar y filtrar

Orden de valores o labels: ascendente o descendente

More Sort Options...

Sort (REGION)

Sort options

Manual (you can drag items to rearrange them)

Ascending (A to Z) by:

Descending (Z to A) by:

Densidad Estudiantil

Summary

Sort REGION by Densidad Estudiantil in descending order

Label Filters (text)

Clear Filter

Equals...

Does Not Equal...

Begins With...

Does Not Begin With...

Ends With...

Does Not End With...

Contains...

Does Not Contain...

Greater Than...

Greater Than Or Equal To...

Less Than...

Less Than Or Equal To...

Between...

Not Between...

Value Filters (numbers)

Clear Filter

Equals...

Does Not Equal...

Greater Than...

Greater Than Or Equal To...

Less Than...

Less Than Or Equal To...

Between...

Not Between...

Top 10...

Selección manual: check/uncheck o escribir el valor

REGION	DISTRITO	Sum of # escuelas	Sum of MATRICULA_TOTAL	Avg BAJO_NIVEL_POBREZA
4	SAN JUAN	CAROLINA	30	12.9
			32	10.9
			44	12.4
			62	17.5
			27	11.3
			37	14.4
			36	13.3
			34	11.3
			40	14.4
			46	14.4
			40	12.7
			12	11.1

AGRUPAR DATOS

Excel permite agrupar datos para hacer cálculos a modo de subtotales con los mismos.

1: Escoger las celdas a agrupar

POP_EST2016 (All)	
Row Labels	Sum of POP_EST2016
Centro	202,174
Centro-Este	451,902
Centro-Oeste	83,174
Este	163,138
Metro	1,135,058
Noreste	133,565
Noroeste	179,184
Norte	364,146
Oeste	141,222
Sur	478,533
Sureste	65,311
Suroeste	131,978
Grand Total	3,529,385

2: Hacer click en botón Group Selection o ...



2: ... right click y escoger la opción Group...

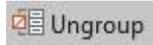
3: Aparecerá Group1 en la cabecera de filas (row)

POP_EST2016 (All)	
Row Labels	Sum of POP_EST2016
Group1	737,250
Centro	202,174
Centro-Este	451,902
Centro-Oeste	83,174
Este	163,138
Este	163,138
Metro	1,135,058

4: Puede cambiar el nombre por defecto a uno de mayor significado

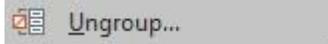
POP_EST2016 (All)	
Row Labels	Sum of POP_EST2016
Interior montañoso	737,250
Centro	202,174
Centro-Este	451,902
Centro-Oeste	83,174
Este	163,138
Este	163,138
Metro	1,135,058

De manera análoga, podemos **desagrupar** (*Ungroup*) el grupo que hayamos creado usando el botón



Ungroup... o hacer **right click** en la celda con agrupación y

escoger

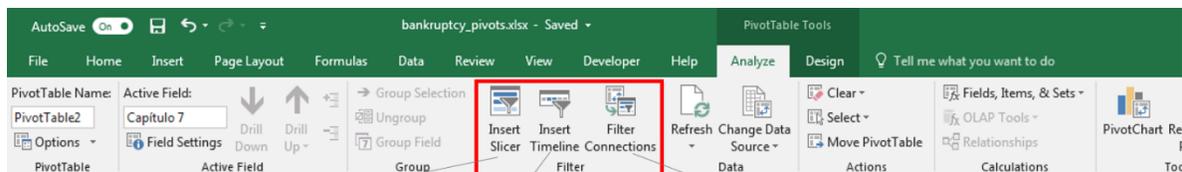


la opción

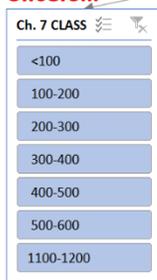
Ungroup...

SLICERS Y TIMELINES

Los slicers y timelines son filtros gráficos interactivos. Como aparece en la figura siguiente, los slicers tienen utilidad cuando los datos a filtrar son discretos, son alfanuméricos o están agrupados.



Slicers...

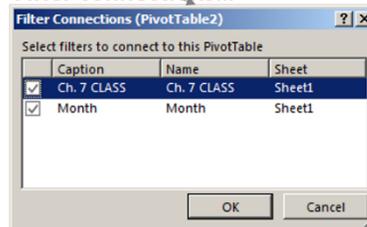


... son útiles cuando existe una columna compuesta por clases o grupos de números, labels o datos nominales con más de una ocurrencia

Timelines...



Filter connections...



... para asegurar la conexión entre slicers/timelines y la tabla Pivot

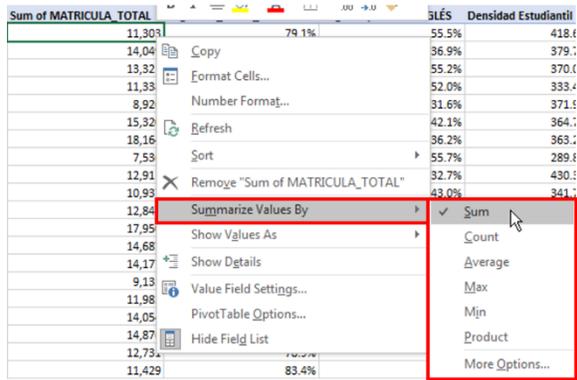
Para este ejemplo se calculó una columna adicional en la tabla Excel para guardar valores de clases que agruparan la cantidad de presentaciones de quiebras por Capítulo 7 al tribunal federal.

CÁLCULO DE VALORES Y DE COLUMNAS

En esta sección discutiremos varias de las opciones para calcular valores de las columnas, así como generar columnas nuevas en una tabla Pivot.

SUMMARIZE VALUES BY...

Resume los datos por la o las dimensiones escogidas. Los resúmenes pueden ser sumas, promedios, conteos y otros.



Sum of MATRICULA_TOTAL	SILÉS	Densidad Estudiantil
11,303	79.1%	55.5%
14,04		36.9%
13,32		55.2%
11,33		52.0%
8,92		31.6%
15,32		42.1%
18,16		36.2%
7,53		55.7%
12,91		32.7%
10,93		43.0%
12,84		43.0%
17,95		43.0%
14,68		43.0%
14,17		43.0%
9,13		43.0%
11,98		43.0%
14,05		43.0%
14,87		43.0%
12,75		43.0%
11,429	83.4%	43.0%

Summarize Values By: determina cómo se van a tratar los números y con cuál función agregarlos, ya sea mediante SUM, COUNT, AVERAGE, MAX, MIN, etc.

Advertencia: Cuando hayan celdas vacías o celdas con texto, Excel calcula por defecto el conteo de celdas (COUNT) en vez de sumar.

Esto se puede arreglar,

- 1) volviendo a los datos y cotejar si hay alguna falta o error.
- 2) También es posible “forzar” la suma mediante el uso de

Summarize Values By | Sum

Como dice la advertencia, **tenga cuidado con las sumas o conteos que Excel haga por defecto.** En ocasiones, hay celdas con alguna falta o error. Puede haber una celda vacía o una celda con valores alfanuméricos donde debería haber un valor numérico. En ese caso, debe cotejar para arreglarlo si es posible. De lo contrario, estar consciente de que deberá forzar la suma cuando aplique.

SHOW VALUES AS

Esta serie de funciones contiene funciones analíticas de gran importancia porque pueden mostrar los datos según su contribución a un gran total, al total de un grupo, diferencia entre fechas, running total (total movable), rank (posición), índices.

Se trata de una manera cómoda y rápida de analizar los datos desde diferentes ángulos sin tener que utilizar funciones ni calcular datos en el data range (datos crudos) o en la tabla Excel.

PERCENT OF GRAND TOTAL:

Muestra la participación de matrícula de cada distrito con relación al total. Para este ejemplo, los valores **% of Grand Total** coinciden con **% of Column Total**

	A	B	C	D	E	F	G	H	I	J
3			Sum of # escuelas	MATRICULA_TOTAL	Sum of MATRICULA_TOTAL	Ave BAJO NIVEL POBREZA	Ave Competencia INGLÉS	Densidad Estudiantil	Gráf	
4	ARECIBO	CAMUY	27	11,303	3.12%			55.5%	418.63	
5	ARECIBO	VEGA ALTA	37	14,049	3.87%			36.9%	379.70	
6	ARECIBO	ARECIBO	36	13,321	3.67%			55.2%	370.03	
7	ARECIBO	MANATI	34	11,338	3.13%			52.0%	333.47	
8	BAYAMON	COROZAL	24	8,926	2.46%			31.6%	371.92	
9	BAYAMON	TOA BAJA	42	15,320	4.23%			42.1%	364.76	
10	BAYAMON	BAYAMON	50	18,164	5.01%			36.2%	363.28	
11	BAYAMON	OROCOVIS	26	7,536	2.08%			55.7%	289.85	
12	SAN JUAN	CAROLINA	30	12,911	3.56%					
13	SAN JUAN	GUAYNABO	32	10,937	3.02%					
14	SAN JUAN	SAN JUAN I	44	12,847	3.54%					
15	SAN JUAN	SAN JUAN II	62	17,950	4.95%					
16	HUMACAO	LAS PIEDRAS	41	14,687	4.05%					
17	HUMACAO	CANOVANAS	40	14,177	3.91%					
18	HUMACAO	FAJARDO	31	9,133	2.52%					
19	HUMACAO	YABUCOA	42	11,982	3.30%					
20	MAYAGUEZ	SAN SEBASTIAN	40	14,054	3.88%					
21	MAYAGUEZ	AGUADILLA	46	14,876	4.10%					
22	MAYAGUEZ	CABO ROJO	40	12,731	3.51%	78.9%				
23	MAYAGUEZ	MAYAGUEZ	43	11,429	3.15%	83.4%				
24	PONCE	SANTA ISABEL	44	15,447	4.26%	75.0%				
25	PONCE	PONCE	50	16,924	4.67%	84.2%				
26	PONCE	UTUADO	32	9,106	2.51%	82.7%				
27	PONCE	YAUCO	48	13,080	3.61%	79.7%				
28	CAGUAS	CIDRA	38	13,783	3.80%	76.5%				
29	CAGUAS	GURABO	48	15,109	4.17%	79.1%				
30	CAGUAS	BARRANQUITAS	36	10,575	2.92%	75.8%				
31	CAGUAS	GUAYAMA	39	10,874	3.00%	78.7%				

Show Values As: En este caso, estamos mostrando cuál es la participación de cada escuela dentro del total de matrícula de las más de 1,000 escuelas en este directorio.

PERCENT OF PARENT ROW TOTAL

Con esta opción, Excel devuelve el cómputo de la participación de cada DISTRITO dentro de la REGIÓN educativa o *Parent Row* en este caso:

	A	B	C	D	E	F	G	H	I	J
REGION	DISTRITO	Sum of # escuelas	MATRICULA_TOTAL	Sum of MATRICULA_TOTAL	Ave BAJO NIVEL POBREZA	Ave Competencia INGLÉS	Densidad Estudiantil	Gráf		
ARECIBO	VEGA ALTA	37	14,049	28.09%			36.9%	379.70		
ARECIBO	ARECIBO	36	13,321	26.64%			55.2%	370.03		
ARECIBO	MANATI	34	11,338	22.67%			52.0%	333.47		
ARECIBO	CAMUY	27	11,303	22.60%			55.5%	418.63		
BAYAMON	BAYAMON	50	18,164	36.37%			36.2%	363.28		
BAYAMON	TOA BAJA	42	15,320	30.67%			42.1%	364.76		
BAYAMON	COROZAL	24	8,926	17.87%						
BAYAMON	OROCOVIS	26	7,536	15.09%						
SAN JUAN	SAN JUAN II	62	17,950	32.85%						
SAN JUAN	CAROLINA	30	12,911	23.63%						
SAN JUAN	SAN JUAN I	44	12,847	23.51%						
SAN JUAN	GUAYNABO	32	10,937	20.01%						
HUMACAO	LAS PIEDRAS	41	14,687	29.39%						
HUMACAO	CANOVANAS	40	14,177	28.37%						
HUMACAO	YABUCOA	42	11,982	23.97%						
HUMACAO	FAJARDO	31	9,133	18.27%						
MAYAGUEZ	AGUADILLA	46	14,876	28.02%						
MAYAGUEZ	SAN SEBASTIAN	40	14,054	26.47%						
MAYAGUEZ	CABO ROJO	40	12,731	23.98%	78.9%					
MAYAGUEZ	MAYAGUEZ	43	11,429	21.53%	83.4%					
PONCE	PONCE	50	16,924	31.02%	84.2%					
PONCE	SANTA ISABEL	44	15,447	28.31%	75.0%					
PONCE	YAUCO	48	13,080	23.97%	79.7%					
PONCE	UTUADO	32	9,106	16.69%	82.7%					
CAGUAS	GURABO	48	15,109	30.01%	79.1%					
CAGUAS	CIDRA	38	13,783	27.38%	76.5%					

Show Values As: % of Parent Row Total

En este caso, estamos mostrando cuál es la participación de cada DISTRITO dentro del total de matrícula de cada REGIÓN.

OTROS EJEMPLOS DE SHOW VALUES AS

Aquí mostramos otros ejemplos dentro de las opciones incluidas en la lista **Show Values As**. Estamos usando datos de casos de solicitudes de permisos de construcción en varios años sometidos a la Oficina de Gerencia de Permisos (OGPe), mostrando el costo estimado del proyecto.

			No calculation	% of Column Total	% of Parent Row Total	% of Percent Difference from... (Years)	Running Total In... (Years)	Rank Largest to Smallest
5	Región	Years	Costo Estimado Proyectos	Costo Estimado2	Costo Estimado3	Costo Estimado4	Costo Estimado5	Costo Estimado6
6	Aguadilla		\$273,009,983	18.03%	18.03%			3
7		2013	\$103,951,243	6.87%	38.08%		\$103,951,243	4
8		2014	\$128,683,130	8.50%	47.13%	23.79%	\$232,634,373	2
9		2015	\$40,375,609	2.67%	14.79%	-68.62%	\$273,009,983	3
10	Arecibo		\$250,253,371	16.53%	16.53%			4
11		2013	\$119,794,314	7.91%	47.87%		\$119,794,314	3
12		2014	\$95,348,841	6.30%	38.10%	-20.41%	\$215,143,155	3
13		2015	\$35,110,215	2.32%	14.03%	-63.18%	\$250,253,371	4
14	Central		\$448,992,237	29.66%	29.66%			1
15		2013	\$200,448,853	13.24%	44.64%		\$200,448,853	1
16		2014	\$131,917,467	8.71%	29.38%	-34.19%	\$332,366,320	1
17		2015	\$116,625,916	7.70%	25.98%	-11.59%	\$448,992,237	1
18	Humacao		\$339,860,954	22.45%	22.45%			2
19		2013	\$173,524,073	11.46%	51.06%		\$173,524,073	2
20		2014	\$92,884,783	6.14%	27.33%	-46.47%	\$266,408,856	4
21		2015	\$73,452,098	4.85%	21.61%	-20.92%	\$339,860,954	2
22	Ponce		\$201,757,384	13.33%	13.33%			5
23		2013	\$96,313,551	6.36%	47.74%		\$96,313,551	5
24		2014	\$82,301,193	5.44%	40.79%	-14.55%	\$178,614,744	5
25		2015	\$23,142,640	1.53%	11.47%	-71.88%	\$201,757,384	5

Las diferentes opciones nos muestran:

No calculation	La suma del Costo Estimado en dólares por región OGPe
% of Column Total	Participación de cada región por año dentro del total de la columna Costo Estimado.
% of Parent Row Total	Participación de cada región por año dentro del total de la fila de origen (parent) en este caso, los años dentro de la Región OGPe.
% Difference from...	Diferencia porcentual dentro de cada región entre años, usando la opción <i>previous</i> .
Running Total In...	Totales, según se van acumulando por años.
Rank Largest to Smallest	Escala ordinal de costo estimado entre regiones, ordenado de mayor a menor . Si se usa la opción por años, los resultados serán diferentes.

SHOW VALUES AS... INDEX

Esta opción la presentamos aparte para entender cómo trabajaría con el conjunto de datos previamente usado de Solicitudes de permisos de construcción. Le aplicamos **Conditional formatting** a las celdas para hacer resaltar las celdas con valores extremos más altos o bajos.

Costo Estimado Proyectos		Región				
Years		Aguadilla	Arecibo	Central	Humacao	Ponce
2013		0.83	1.04	0.97	1.11	1.04
2014		1.34	1.09	0.84	0.78	1.16
2015		0.78	0.74	1.36	1.13	0.60

En este caso, los costos estimados se están viendo como índices por año y por región de la OGP. Como podemos ver, los índices más altos están en la región Central (San Juan) para el año 2015 y en Aguadilla para el año 2014. Los índices más bajos están en Ponce (0.60) y en Arecibo (0.74) para el año 2015.

Los índices son una especie de promedios ponderados. Si tomáramos en cuenta solamente los costos estimados y le aplicamos **Conditional Formatting**, veríamos algo diferente, que puede o no coincidir. Fíjese en los valores crudos de 2013 en la región Central. En cuanto al índice, es uno bajo, pero al ponderar entre filas y columnas, su efecto es mucho menor.

Costo Estimado Proyectos		Región					
Years		Aguadilla	Arecibo	Central	Humacao	Ponce	Grand Total
2013		\$103,951,243	\$119,794,314	\$200,448,853	\$173,524,073	\$96,313,551	\$694,032,034
2014		\$128,683,130	\$95,348,841	\$131,917,467	\$92,884,783	\$82,301,193	\$531,135,415
2015		\$40,375,609	\$35,110,215	\$116,625,916	\$73,452,098	\$23,142,640	\$288,706,479
Grand Total		\$273,009,983	\$250,253,371	\$448,992,237	\$339,860,954	\$201,757,384	\$1,513,873,928

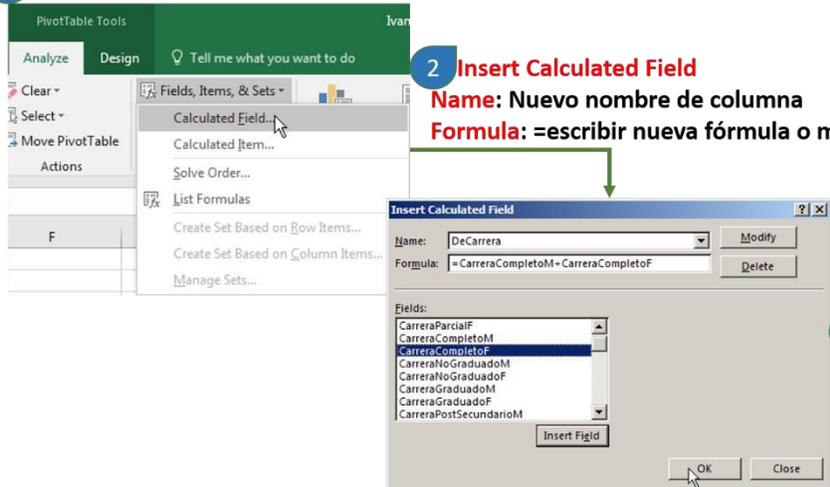
Los **índices** son calculados de la siguiente manera:

$$\text{Valor de la celda} * \text{Gran Total} / (\text{Total de la fila} * \text{Total de la columna})$$

CAMPOS CALCULADOS

A menudo tenemos columnas con conteos divididos en las cuales queremos hacer ciertas operaciones aritméticas, tales como sumar, restar, multiplicar o dividir. En una tabla Pivot podemos generar nuevas columnas que contengan cálculos basados en otras columnas. En este ejemplo, tenemos el **Registro Anual de Puestos por Demografía**. En este registro se dividen los conteos por género. Podemos sumar conteos sin afectar los datos originales. A continuación, mostramos cómo producir una columna nueva para una tabla Pivot:

1 Pivot Table Tools | Analyze | Fields, Items, & Sets | Calculated Field... Para crear una nueva columna (field)



2 Insert Calculated Field
Name: Nuevo nombre de columna
Formula: =escribir nueva fórmula o modificar existente

3 Resultados

1		
2	AnoFiscal	2012-2013
3		
4	Row Labels	Sum of DeCarrera
5	Corporaciones públicas	35,200
6	Departamentos y Agencias	100,643
7	Municipios	26,600
8	Oficina del Gobernador	1,069
9	Otras entidades	344
10	Rama Judicial	3,568
11	Rama Legislativa	529
12	Grand Total	167,953

CÁLCULOS BASADOS EN CONTEOS

En PivotTables, las operaciones numéricas normales están restringidas a campos numéricos basados en sumatorias; no en conteos. Siguiendo el tema de campos calculados, ¿qué pasaría si queremos utilizar un campo basado en conteos en lugar de estar basado en sumatorias para hacer operaciones aritméticas?

Al hacer resúmenes por dimensión, (región, clase, tipo, etc.) lo hacemos generalmente mediante sumas y cálculos agregados como promedios. En ocasiones, queremos utilizar algún campo basado en conteos para calcular otros campos.

El siguiente diagrama muestra primero un ejemplo de una falta al no considerar que un campo es basado en conteos.

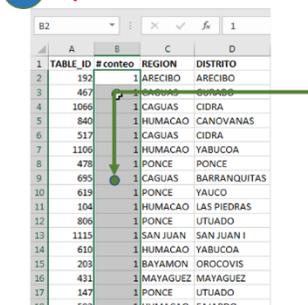
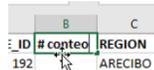
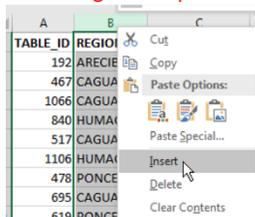
Falta al calcular columna:


Debe existir una columna que registre los conteos (basada en una medición) para que funcione correctamente.

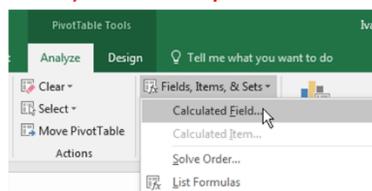
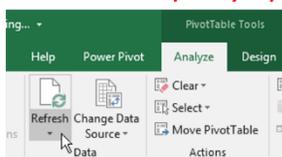
Este es el **procedimiento recomendado**: crear columna nueva en la tabla de datos. ¿Cómo hacerlo?

Primero, añadir una columna de conteos en la tabla fuente...

- 1 **Crear columna nueva** usando **right click | Insert**
- 2 **Ponerle nombre** a la columna
- 3 **Copiar el número 1 ...** en todas las celdas de esta columna



- 4 **Regresar a la tabla pivot...**
Pivot Table Tools | Analyze | Refresh
- 5 **Pivot Table Tools |**
Fields, Items & Sets | Calculate Field...
- 6 **Calcular la columna**



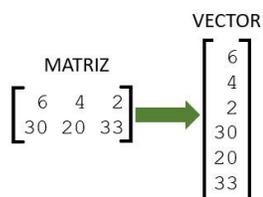
En este ejemplo se genera una columna con valores 1. Excel generará una suma y esta nueva columna podrá ser usada para cálculos aritméticos. En este caso: promedio de matrícula por escuela.

Resultados:

5	REGION	Count of CODIGO	Sum of MATRICULA_TOTAL	Sum of Densidad Estudiantil
6	ARECIBO	134	50,011	373
7	BAYAMON	142	49,946	352
8	CAGUAS	161	50,341	313
9	HUMACAO	154	49,979	325
10	MAYAGUEZ	169	53,090	314
11	PONCE	174	54,557	314
12	SAN JUAN	168	54,645	325
13	Grand Total	1102	362,569	329

CONSOLIDAR COLUMNAS (VECTORIZACIÓN)

En ocasiones, **una misma métrica está dividida por años o alguna otra subclase**. Por lo tanto, es preferible **consolidar varias columnas que tratan sobre el mismo tema en una sola columna**. Este concepto es llamado **vectorización**, en términos de **Data Science**. En este proceso, como en álgebra lineal, se convierte una matriz en un vector.



Un **vector** es una **secuencia de datos de un mismo tipo**. Los miembros en un vector son llamados **componentes**.

Para **programas de manejo de datos y visualización** como **Excel, Power BI, Tableau**, etcétera, puede trabajarse de manera más eficiente con datos vectorizados.

Podemos trabajar un ejemplo que puede ser una tabla presupuestaria, la cual tiene múltiples dimensiones y una sola métrica dividida en años u otras características de ese y otros años.

Esta es una hoja derivada de una consulta (*SQL Query*) a la base de datos presupuestarios de la Oficina de Gerencia y Presupuesto. El resultado del query vuelve el presupuesto aprobado para los años 2001 al 2016 de cada agencia, sus características, los conceptos de gasto y origen de fondos. La métrica viene siendo la misma: **fondos**, divididos por año.

A	B	C	D	E	F	G	H	I	J	K	L	N	Q	T	V	Z	AC	AF	AI	AL	AO	
RAMA	TIPO	AGENCIA	NOMBRE	PROGRAMA	NOMBSF	PROGRAMA	ORIGEN	CONCEPTO_DE_GASTO	CUENTA	DESCRIPCION	Área	Sector	AF 2001 Apr	AF 2002 Apr	AF 2003 Apr	AF 2004 Apr	AF 2005 Apr	AF 2006 Apr	AF 2007 Apr	AF 2008 Apr	AF 2009 Apr	AF 2010 Apr
1	E	A	132	Administración 1331	Administración y Conserva	AE	Nómina y Costos Relaciona	1100	Sueldos; puestos	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
2	E	A	132	Administración 1331	Administración y Conserva	AE	Donativos, Subsidios y Distri	6170	Donativos y Apor	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	E	A	132	Administración 1331	Administración y Conserva	AE	Donativos, Subsidios y Distri	6190	Donativos y Apor	Desarr	Energía	290.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	E	A	132	Administración 1331	Administración y Conserva	AR	Asignaciones Empleadas	0001	Asignaciones em	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	93.00
5	E	A	132	Administración 1331	Administración y Conserva	AR	Nómina y Costos Relaciona	1100	Sueldos; puestos	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	691.00
6	E	A	132	Administración 1331	Administración y Conserva	AR	Nómina y Costos Relaciona	1100	Sueldos; puestos	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	E	A	132	Administración 1331	Administración y Conserva	AR	Nómina y Costos Relaciona	1100	Sueldos; Puestos	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	E	A	132	Administración 1331	Administración y Conserva	AR	Servicios Profesionales	1340	Servicios de Cont	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	180.00
9	E	A	132	Administración 1331	Administración y Conserva	AR	Servicios Profesionales	1360	Servicios Profesi	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.00
10	E	A	132	Administración 1331	Administración y Conserva	AR	Anuncios y Pautas en Medio	2020	Anuncios de Des	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	E	A	132	Administración 1331	Administración y Conserva	AR	Servicios Profesionales	1360	Servicios profes	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2,222.00
12	E	A	132	Administración 1331	Administración y Conserva	AR	Servicios Comprados	1034	Servicios Privats	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	24.00
13	E	A	132	Administración 1331	Administración y Conserva	AR	Nómina y Costos Relaciona	1100	Compensaciones	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	E	A	132	Administración 1331	Administración y Conserva	AR	Nómina y Costos Relaciona	1100	Bono Navidad - P	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	E	A	132	Administración 1331	Administración y Conserva	AR	Nómina y Costos Relaciona	1100	Liquidación Licen	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	E	A	132	Administración 1331	Administración y Conserva	AR	Nómina y Costos Relaciona	1100	Liquidación Licen	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	E	A	132	Administración 1331	Administración y Conserva	AR	Anuncios y Pautas en Medio	2000	Anuncios y Aviso	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	E	A	132	Administración 1331	Administración y Conserva	AR	Anuncios y Pautas en Medio	2020	Anuncios Teatros	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19	E	A	132	Administración 1331	Administración y Conserva	AR	Servicios Comprados	2030	Impresos y envas	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	32.00
20	E	A	132	Administración 1331	Administración y Conserva	AR	Servicios Comprados	212	Franqueo	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.00
21	E	A	132	Administración 1331	Administración y Conserva	AR	Gastos de Transportación	2220	Servicios de Tran	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.00
22	E	A	132	Administración 1331	Administración y Conserva	AR	Gastos de Transportación	2300	Pasajes de Viajes	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
23	E	A	132	Administración 1331	Administración y Conserva	AR	Gastos de Transportación	2320	Gastos de subst	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.00
24	E	A	132	Administración 1331	Administración y Conserva	AR	Gastos de Transportación	2330	Pasajes de Viajes	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	9.00
25	E	A	132	Administración 1331	Administración y Conserva	AR	Gastos de Transportación	2340	Subsistencia a per	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.00
26	E	A	132	Administración 1331	Administración y Conserva	AR	Gastos de Transportación	2350	Bonificaciones pc	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.00
27	E	A	132	Administración 1331	Administración y Conserva	AR	Servicios Comprados	2362_3	Transportación Si	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
28	E	A	132	Administración 1331	Administración y Conserva	AR	Gastos de Transportación	2380	Gastos de Viaje fu	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
29	E	A	132	Administración 1331	Administración y Conserva	AR	Gastos de Transportación	2390	Gastos de viaje fu	Desarr	Energía	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.00

Para un mejor uso de las funciones PivotTable, vamos a combinar las **columnas presupuestarias separadas** (que **son una sola métrica dividida por años**), consolidándolas en dos columnas: una contendrá los presupuestos (*Value*) y la otra contendrá el año fiscal (*Attribute*). En Excel prepararemos el archivo recibido y traspondremos dichas columnas.

En lugar de **abrir directamente el archivo** de hoja de cálculo presupuestario, **usaremos la interfaz Query Editor de Excel** para **preparar los datos antes de traer la hoja Excel** con los datos. No haremos cambios al archivo original.

TRASPONER USANDO UNPIVOT COLUMNS Query Editor: Unpivot columns

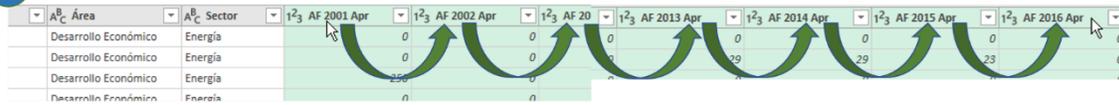
- Combinar múltiples columnas en una.
 - Ejemplo: columnas 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017.
 - Combinarlas en una sola
- CUIDADO: Esto multiplicará el número de filas (rows) multiplicándolas por cada columna
- **Importar datos** a Excel usando
[Data](#) | [Get Data](#) | [From File](#) | [From Workbook](#)

Los pasos para traer los datos y trasponer (*unpivot*) las columnas son los siguientes:
Hay dos procesos principales:

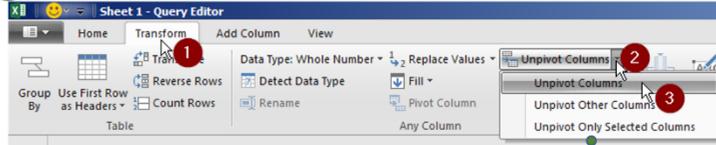
- 1 **Importar el archivo Excel** original usando **Data | Get Data... From Workbook**
- 2 Usar la interfaz **Query Editor** de Excel para **escoger las columnas a consolidar y aplicar la herramienta tab Transform | Unpivot Column**.

TRANSFORM | UNPIVOT COLUMNS

A Ctrl+Click en cada cabecera de columna AF 2001 al 2016 Apr



B Click en Transform tab | Unpivot Columns | Unpivot Columns



Nota: Aunque vamos a *trasponer* columnas, **NO** vamos a usar el botón **Transpose**.

Sólo vamos a trasponer 16 de las columnas presupuestarias (métricas) y dejaremos las dimensiones (las otras columnas descriptivas) como están. Por esto usamos el botón **Unpivot Columns**.

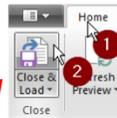
D Cambie los nombres de las columnas:

Attribute → **Año**
Value → **Presupuesto**

C Resultados...
Las 16 columnas fueron reducidas a dos: **Attribute** y **Value**

Área	Sector	Attribute	Value
Desarrollo Económico	Energía	AF 2010 Apr	691
Desarrollo Económico	Energía	AF 2011 Apr	183
Desarrollo Económico	Energía	AF 2012 Apr	55
Desarrollo Económico	Energía	AF 2013 Apr	0
Desarrollo Económico	Energía	AF 2014 Apr	0
Desarrollo Económico	Energía	AF 2015 Apr	0
Desarrollo Económico	Energía	AF 2016 Apr	0

D Para ejecutar cambios, Haga click en: **Home** tab | **Close & Load**

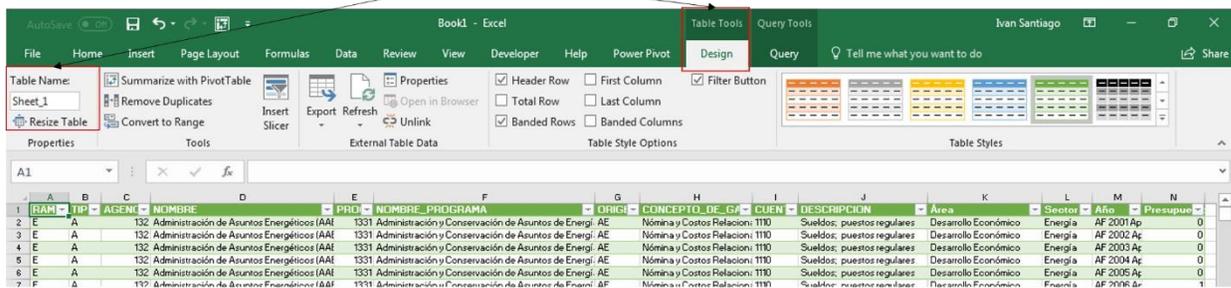


Al cerrar la interfaz **Query Editor**, el **resultado** será una **tabla Excel**. Esta tendrá sus filas (rows) multiplicadas por cada columna que consolidamos.



Espere el procesamiento de los datos

La **Tabla** resultante consiste de **753,536** filas y **14** columnas



Al hacer la **trasposición** usando **Unpivot Columns**, consolidamos los años 2001 al 2016 en una sola columna, **repetiendo** filas por cada año. El archivo original contiene **47,096** filas y **58** columnas...

Esta tabla ya está lista para usar las funciones de PivotTable. Usaremos este archivo para uno de los ejercicios.

PIVOTCHARTS

Una gráfica Pivot es igual a cualquier otra gráfica Excel, salvo que está directamente atada a una tabla Pivot. La gráfica Pivot se irá ajustando conforme se le apliquen filtros y campos a la tabla Pivot a la cual esté enlazada.



En este ejemplo arriba, mostramos en una gráfica Pivot las **solicitudes de quiebras ante el tribunal federal, distrito de San Juan** por los **capítulos 7 y 13**. Se aplicaron filtros seleccionando solamente los grupos <100, 100-200 y 200-300, y para los años 2000 al 2006.



La misma gráfica puede mostrar ahora los casos sometidos para todas las clases desde <100 a 1,200, además de incluir los años 2000 a 2010.

CÓMO AÑADIR UNA GRÁFICA PIVOT

Excel provee dos opciones para añadir una gráfica Pivot. Luego de haber activado la tabla Pivot, puede añadir una gráfica Pivot desde el menú principal:

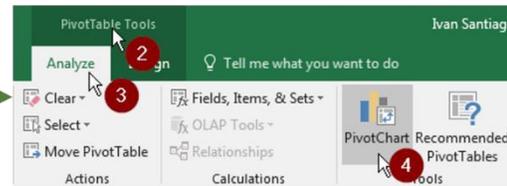
A: PivotTable Tools | Analyze | PivotChart

B: Insert | PivotChart

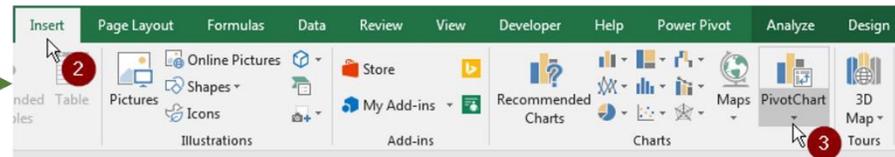
1 Activar la tabla Pivot

Row Labels	Capítulo 7	Capítulo 13
2000	4,116	5,462
2001	4,276	9,973
2002	4,084	9,549
2003	4,221	9,874
2004	4,247	8,853
2005	4,027	9,032
2006	1,284	4,066
2007	1,756	5,877
2008	2,311	6,589
2009	3,148	7,974
2010	4,253	7,942
Grand Total	37,723	85,191

Opción A: Pivot Table Tools | Analyze | PivotChart



Opción B: Insert | PivotChart



OPCIONES DISPONIBLES PARA FORMATEAR GRÁFICAS PIVOT

Cuando una gráfica Pivot es activada, se activan a su vez una serie de pestañas (tabs) y botones con diferentes funciones. Estas funciones van desde las relacionadas a:

- **Analysis tab:** aplicar filtros, slicers, timelines, conexiones a tablas Pivot, y eliminar de la gráfica los botones de filtros, entre otras.
- **Design tab:** Cambiar cómo luce la gráfica Pivot. Tal como se hace con las gráficas regulares en Excel.
- **Format tab:** Ir más al detalle de los elementos de la gráfica para hacerle cambios individuales a diferentes elementos que la componen.

Analyze tab



Design tab



Format tab

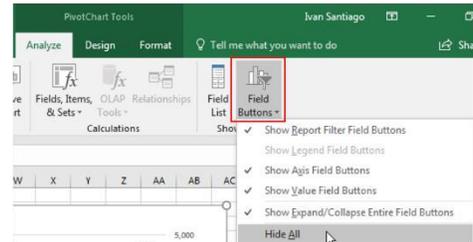


FIELD BUTTONS

Por defecto, aparecerán botones correspondientes a las dimensiones que aparecen en la gráfica Pivot. Estos botones son para filtrar datos o quitar un campo para moverlo al área de filtros.



Field Buttons: Pueden ser apagados o encendidos desde el grupo **Analyze | Botón Field Buttons**

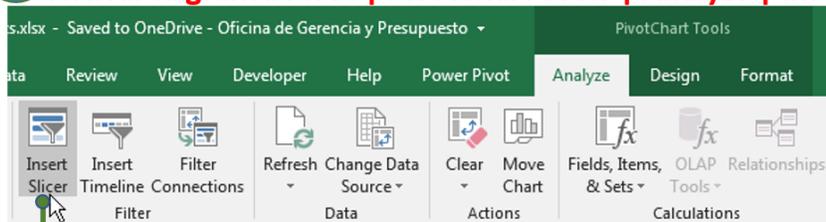


Los botones se pueden apagar o mostrar en la gráfica desde el menú **Analyze | Field Buttons**.

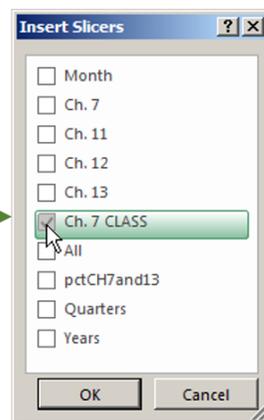
SLICERS

Los slicers son filtros interactivos. Contienen botones que corresponden a las dimensiones que deseamos aplicar como filtros. Los slicers se añaden desde **PivotChart Tools | tab Analyze | botón Insert Slicer**.

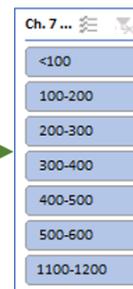
1 Activar la gráfica Pivot | PivotChart Tools | Analyze | Insert Slicer



2 Escoger columna(s)



3 Resultado

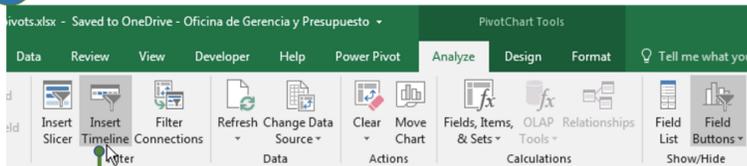


Es más práctico crear slicers para datos **dimensionales** (categorías). Si va a usar mediciones, recomendamos dividirlos en clases (usar funciones LOOKUP) para evitar tener slicers con listas demasiado largas haciéndolos inmanejables.

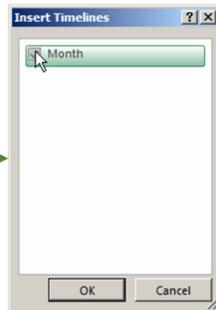
TIMELINES

Los timelines son slicers para usarlos con datos temporales (fechas, tiempo). Para añadir un timeline, active la gráfica Pivot, y haga **click** en el **botón Insert Timeline**, localizado en el **grupo Filter**, dentro del **tab Analyze** del menú **PivotChart Tools**.

1 Activar la gráfica Pivot | PivotChart Tools | Analyze | Insert Timeline



2 Escoger columna(s)



3 Resultado



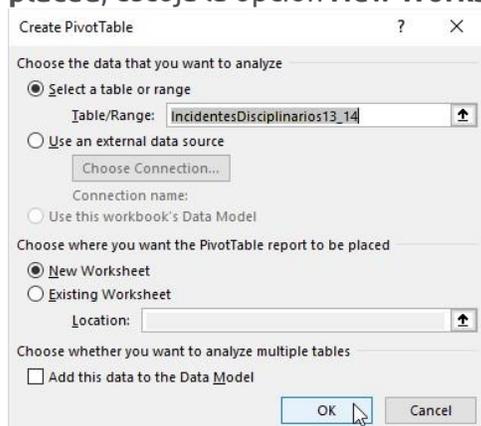
PRÁCTICAS: CREAR PIVOTTABLES Y PIVOTCHARTS

En esta parte vamos a practicar con algunos conjuntos de datos que se han presentado anteriormente como ejemplos de PivotTables y PivotCharts.

EJERCICIO 1: CREAR UNA TABLA PIVOT

Objetivo: generar tabla Pivot utilizando datos crudos en formato Excel, previamente descargados desde el portal del [Instituto de Estadísticas](#). El Departamento de Educación publicó gráficas para años hasta 2021. A pesar de eso, los datos no se pueden descargar o la interfaz para descarga no se ve o no está habilitada.

- Abra el workbook **IncidentesDisciplinariosEscuelasDE2013-14.xlsx**, localizado en **C:\Tutorial_Excel\DataAnalysis\Modulo_3**.
- Active la tabla** haciendo **click** y vaya al **menú principal, Insert | PivotTable**. En la forma **Create PivotTable**, en el apartado **Choose where you want the PivotTable report to be placed**, escoja la opción **New Worksheet**.



A la **izquierda** aparecerá la tabla vacía **PivotTable1**. A la **derecha** aparecerá la forma/interfaz **PivotTable Fields**.

- En la forma **PivotTable Fields**, arrastre el campo **Región** en el cuadro **Rows**. Arrastre luego el campo **Incidentes** al cuadro **Values**.



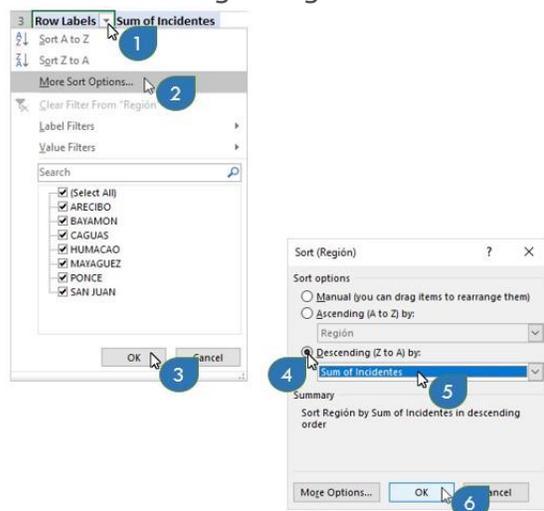
- Regrese un momento a la hoja **Incidentes**. Incidentes Inspeccione la columna C: Incidentes. A base de lo que se discutió anteriormente en la sección CÁLCULOS BASADOS EN CONTEOS, ¿Por qué todas las celdas de esta columna tienen el valor 1?

- Volvamos a la hoja que tiene la nueva tabla pivot. La tabla Pivot deberá aparecer así. Estas son las sumas de distintos incidentes disciplinarios sumados por región.

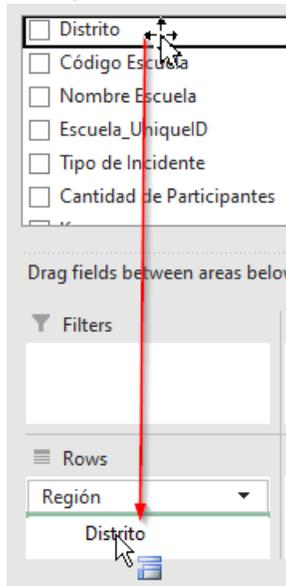
Row Labels	Sum of Incidentes
ARECIBO	15
BAYAMON	18
CAGUAS	132
HUMACAO	87
MAYAGUEZ	12
PONCE	70
SAN JUAN	100
Grand Total	434

- ¿Cuáles son las (3) regiones educativas con más incidentes disciplinarios registrados?

- Puede **ordenar** los datos en la columna **Sum of Incidentes** para ordenarlas de manera descendente. Siga la siguiente secuencia:



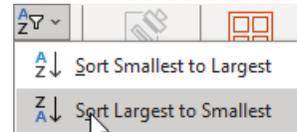
- Ahora, **arrastre** la columna **Distrito** *debajo de* la columna **Región** en la caja **Rows**.



La tabla pivot deberá verse así. La tabla muestra ahora más detalle. Esta vez desglosa los incidentes por **región educativa y distrito**. Un **distrito escolar** puede estar compuesto de uno o más municipios.

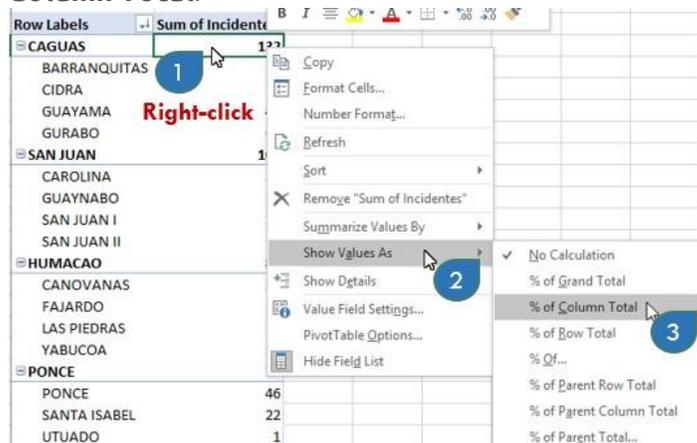
Row Labels	Sum of Incidentes
CAGUAS	132
GURABO	62
GUAYAMA	42
BARRANQUITAS	15
CIDRA	13
SAN JUAN	100
GUAYNABO	32
CAROLINA	27
SAN JUAN I	26
SAN JUAN II	15
HUMACAO	87
LAS PIEDRAS	29
YABUCOA	25
CANOVANAS	25
FAJARDO	8
PONCE	70
PONCE	46
SANTA ISABEL	22
YAUACO	1
UTUADO	1
BAYAMON	18
BAYAMON	12

Haga **click** en la celda **B5** y ordene los valores de distrito.



MOSTRAR VALORES COMO PORCENTAJES (COLUMN TOTAL)

- Haga **right click** en **cualquier celda** debajo de la columna **Sum of Incidentes**.
- En el menú de contexto, escoja la opción **Show Values As** y luego **click** en la opción **% of Column Total**.



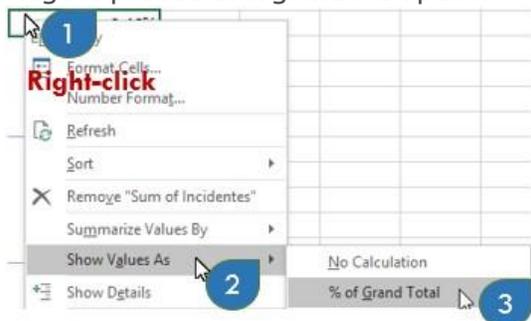
La tabla Pivot debe verse así:

Row Labels	Sum of Incidentes
CAGUAS	30.41%
GURABO	14.29%
GUAYAMA	9.68%
BARRANQUITAS	3.46%
CIDRA	3.00%
SAN JUAN	23.04%
GUAYNABO	7.37%
CAROLINA	6.22%
SAN JUAN I	5.99%
SAN JUAN II	3.46%
HUMACAO	20.05%
LAS PIEDRAS	6.68%
YABUCOA	5.76%
CANOVANAS	5.76%
FAJARDO	1.84%
PONCE	16.13%

Fíjese en lo que muestra la tabla. Aquí aparecen las **proporciones en por cientos por el total en cada columna**. Cada **región educativa** y **distrito** muestran su **proporción con relación al valor total de esta columna (Cantidad de Incidentes)**.

En este conjunto de datos **no habrá diferencia si usamos** las opciones **% of Grand Total** y **% of Column Total**.

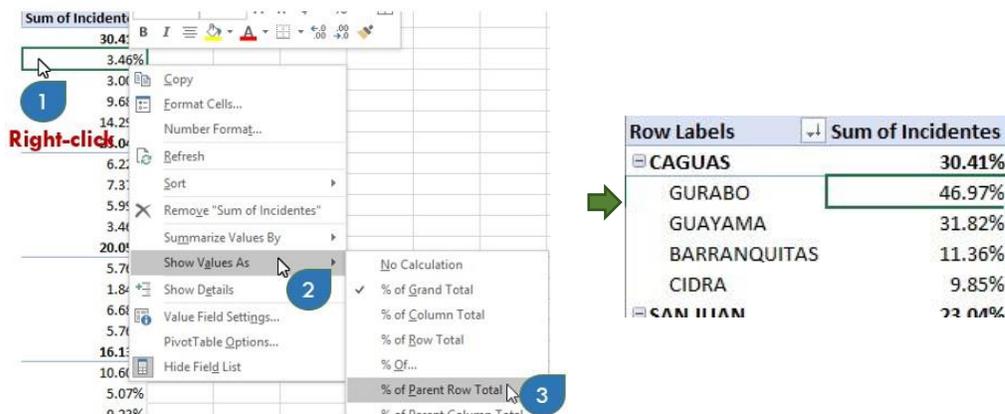
- Haga la prueba escogiendo la opción **% of Grand Total** en la columna **Sum of Incidentes**.



VER PARTICIPACIÓN DE CADA DISTRITO DENTRO DE LA REGIÓN

Hasta el momento, hemos visto la participación de cada distrito dentro del total de la columna. Ahora, vamos a **mostrar la participación de cada distrito dentro de cada región**.

- Escoja ahora la opción **% of Parent Row Total** para que pueda ver la participación de cada distrito.

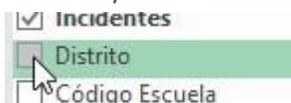


Note que dentro de la región educativa de Caguas, casi 80% de los incidentes están en los distritos de Gurabo y Guayama.

RANKS (POSICIÓN)

Se puede preguntar, **cuál será el orden** de cantidad de incidentes por regiones. Esta pregunta puede contestarse usando la opción **Rank**

- Primero, **remueva** los **distritos** de la caja Rows haciendo **uncheck** en el campo **Distrito**



- Haga **right click** encima de cualquier celda de la columna **Sum of Incidentes**.

- Escoja la opción **Show Values | Rank Largest to Smallest...**



- Escoja el campo **Región** en la forma **Show values as (Sum of Incidentes)**



- Como era de esperarse, así aparecen las posiciones.

Row Labels	Sum of Incidentes
CAGUAS	1
SAN JUAN	2
HUMACAO	3
PONCE	4
BAYAMON	5
ARECIBO	6
MAYAGUEZ	7

- Guarde los cambios usando el mismo nombre del archivo original.**

Opcional: Contestar:

¿Cuál es el distrito escolar con mayor cantidad de incidentes disciplinarios? _____

¿Cuáles son las 10 escuelas con mayor cantidad de incidentes disciplinarios?

Ayuda: use el campo **Escuela_UniqueID** **Escuela_UniqueID** y aplique filtro **Value Filter | Top 10**

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

¿Cuáles son los 10 tipos de incidentes más frecuentes?

En el próximo ejercicio, haremos el proceso de integrar varias columnas y compactarlas en dos columnas mediante vectorización (unpivot), usando datos presupuestarios.

EJERCICIO 2: VECTORIZACIÓN O CONSOLIDACIÓN (UNPIVOT) DE MÚLTIPLES COLUMNAS

Como mencionamos anteriormente, cuando los datos están divididos en múltiples columnas, la preparación de cómputos para las opciones *Show Values As* son más tediosas o puedan dar resultados inconsistentes.

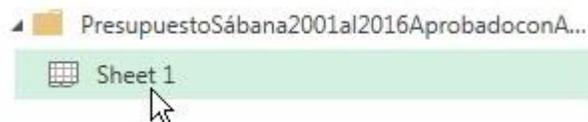
Para esta parte, usaremos un workbook con **datos presupuestarios de la Oficina de Gerencia y Presupuesto** para los años **2001 al 2016**.

- Abra un nuevo archivo workbook de Excel. **New | Blank workbook.**

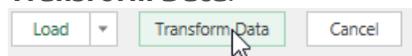
Vamos a usar la interfaz Power Query de Excel para extraer, transformar y cargar los datos de un workbook Excel a otra sesión de Excel. Durante el proceso de importación, vamos a realizar la operación "Unpivot". Con esta función vamos a convertir todas las columnas presupuestarias de los años 2001 a 2016 a una sola columna. En realidad se trata de la misma métrica: presupuesto; y la podemos integrar en una sola columna, mientras que los años se recogerán en otra columna. La desventaja es que la cantidad de récords se multiplicará.

USAR POWER QUERY PARA TRANSFORMAR DATOS

- Si está usando Office 365, vaya al **menú principal** y escoja **Data | Get data | From File | From Excel Workbook**
- Si está usando Office 2016, deberá ir al **menú principal** y escoja **Data | New Query | Get data | From File | From Workbook**
- Traiga (**Import**) el archivo **PresupuestoSábana2001al2016AprobadoconAreaySector.xlsx**, localizado en **C:\Tutorial_Excel\DataAnalysis\Modulo_3**.
- En la forma **Navigator de Power Query**, escoja (click) en la hoja **Sheet1**



- Ya que vamos a hacer modificaciones al archivo resultante, haga **click** en el botón **Transform Data**.



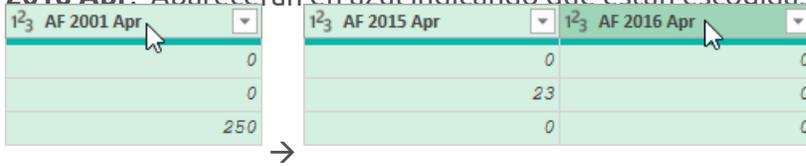
Aparecerá la interfaz **Query Editor** que mostrará el contenido del workbook a importar.

- Navegue hacia la derecha para que pueda ver las columnas del **presupuesto aprobado (AF Apr)** para los años 2001 al 2016.

VECTORIZAR (UNPIVOT) MÚLTIPLES COLUMNAS

Como hemos mencionado antes, este proceso trata sobre integrar los valores guardados en varias columnas (16 en estos datos) y guardándolas en dos columnas. Una columna tiene la dimensión de los años 2001 al 2016 y la otra columna resultante tendrá los presupuestos aprobados que corresponden a cada año fiscal.

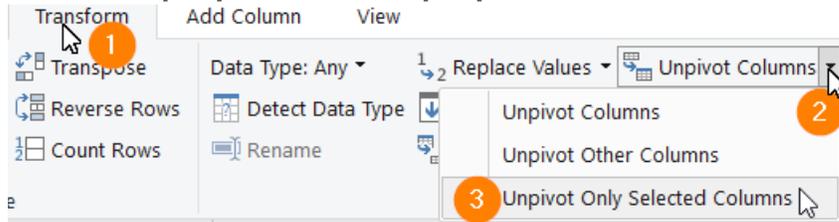
- Haga **shift+click** en las **cabeceras headers** de las **columnas AF 2001 Apr** a la **columna AF 2016 Apr**. Aparecerán en azul indicando que están escogidas.



AF 2001 Apr	AF 2015 Apr	AF 2016 Apr
0	0	0
0	23	0
250	0	0

Debe escoger todas las columnas desde 2001 a 2016.

- Vaya al **menú principal** de la Interfaz **Query Editor** y escoja **Transform | Unpivot Columns | Unpivot Columns**



CAMBIAR NOMBRES DE LAS COLUMNAS

Es mejor renombrar las columnas para que tengan nombres reconocibles luego.

- Haga **doble click** en el nombre de la columna **Attribute** y escriba **Año**.
- Haga **doble click** en el nombre de la columna **Value** y escriba **Presupuesto**.

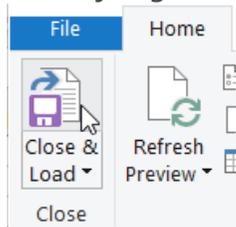


Año	Presupuesto
AF 2001 Apr	0

ACEPTAR CAMBIOS Y CERRAR LA INTERFAZ QUERY EDITOR

No vamos a hacer más cambios. Recuerde que estos cambios no afectan al archivo original. Los cambios se hacen en la memoria RAM y se guardarán en otro workbook Excel.

- Para aceptar los cambios y cerrar la interfaz **Query Editor**, vaya al **menú principal** y escoja **Home** y haga **click** en el botón **Close and Load**



Esto iniciará el proceso de conversión de los datos para traerlos a una hoja de Excel.



A	B	C
ExternalData_1: Getting Data ...		

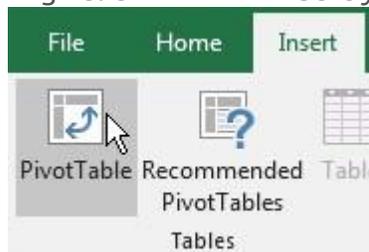
El proceso tardará debido a que se están multiplicando las filas (rows) para cada año fiscal. Inspeccione la tabla resultante:

Sector	Año	Presupuesto
Energía	AF 2002 Apr	0
Energía	AF 2003 Apr	0
Energía	AF 2004 Apr	0
Energía	AF 2005 Apr	0
Energía	AF 2006 Apr	0
Energía	AF 2007 Apr	0
Energía	AF 2008 Apr	0
Energía	AF 2009 Apr	0
Energía	AF 2010 Apr	0
Energía	AF 2011 Apr	0
Energía	AF 2012 Apr	0
Energía	AF 2013 Apr	29
Energía	AF 2014 Apr	29
Energía	AF 2015 Apr	23
Energía	AF 2016 Apr	0
Energía	AF 2001 Apr	250

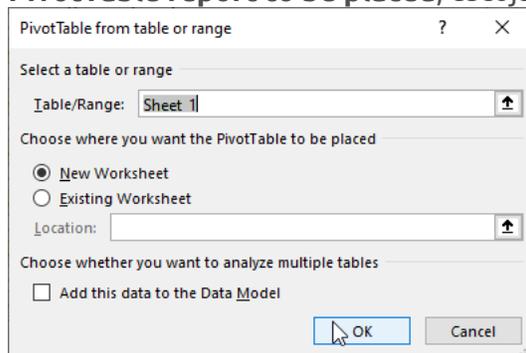
Notará además que **el resultado es una *tabla Excel***.

GENERAR LA TABLA PIVOT

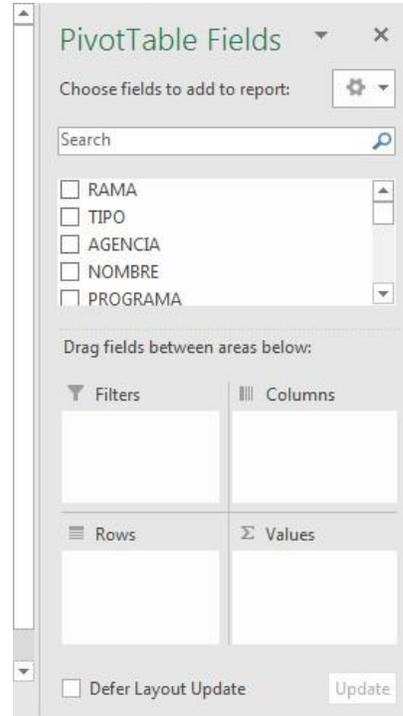
- Haga **click** en **cualquier celda dentro de la tabla Excel**. Vaya al **menú principal** (ribbon) y haga **click** en el tab **Insert** y **click** en el botón **Pivot Table**, del **grupo Tables**.



- En la forma **PivotTable from table or range**, bajo el apartado **Choose where you want the PivotTable report to be placed**, escoja la opción **New Worksheet**.

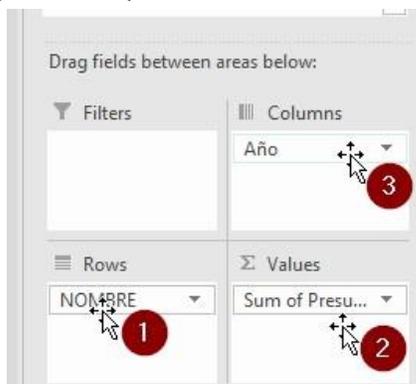


Aparecerán en la nueva hoja la **tabla pivot vacía** a la izquierda y la interfaz **PivotTable Fields**



GENERAR INFORME PRESUPUESTARIO POR AGENCIA

- Para generar el primer informe simple por agencia gubernamental, arrastre los campos (columnas)...



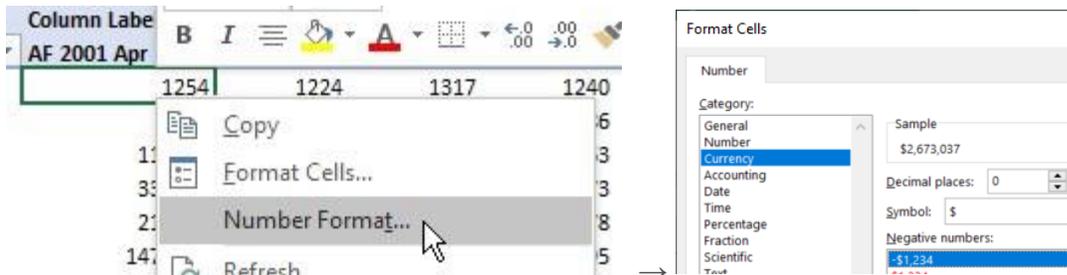
- 1 **Nombre:** a la caja **Rows**
- 2 **Presupuesto:** ... a la caja **Values**
- 3 **Año:** a la caja **Columns**

La tabla Pivot deberá verse así:

Sum of Presupuesto	Column Labels									
Row Labels	AF 2001 Apr	AF 2002 Apr	AF 2003 Apr	AF 2004 Apr	AF 2005 Apr	AF 2006 Apr	AF 2007 Apr	AF 2008 Apr	AF 2009 Apr	AF 2010 Apr
Administración de Asuntos Energéticos (AAE)	1254	1224	1317	1240	1696	1628	1627	1063	1405	
Administración de Asuntos Federales de Puerto Rico	9350	9232	9245	9986	10257	7296	6255	6255	6044	
Administración de Compensaciones por Accidentes de Automóviles	114629	85663	92890	103463	102780	102002	103586	102616	97758	
Administración de Corrección	332724	315500	343525	383773	380794	478804	439558	403090	406069	
Administración de Desarrollo Laboral	218099	262087	224984	94878	107560	95347	81476	107998	145707	
Administración de Desarrollo Socioeconómico de la Familia	1475056	1505138	1575573	1597895	1686526	1704348	1743746	1813684	2049123	
Administración de Desarrollo y Mejoras de Viviendas	15469	18541	19604	16717	13051	11568	0	0	0	
Administración de Familias y Niños	317613	327545	361462	173327	234404	206573	217316	223591	225079	
Administración de Fomento Comercial	13559	13201	12882	9676	0	0	0	0	0	
Administración de Instituciones Juveniles	89972	82551	101580	83947	82358	91507	90578	80630	79175	
Administración de la Industria y el Deporte Hípico	4009	4362	4701	4716	5073	3820	4656	5078	5483	
Administración de Recursos Naturales	72858	74056	98524	68182	80008	121192	89339	69641	77629	
Administración de Reclamaciones y Permisos	71800	77588	77157	10514	10007	71367	71087	77030	73065	

CAMBIAR FORMATO A CURRENCY

- Haga **right click** en cualquiera de las celdas bajo la columna **AF 2001 Apr**.
- Escoja la opción **Number Format....** En la forma **Format Cells**, bajo la lista **Category**, escoja **Currency**.
- Escriba **0** en **Decimal places** y haga **click** en el botón **OK** para aceptar los cambios.



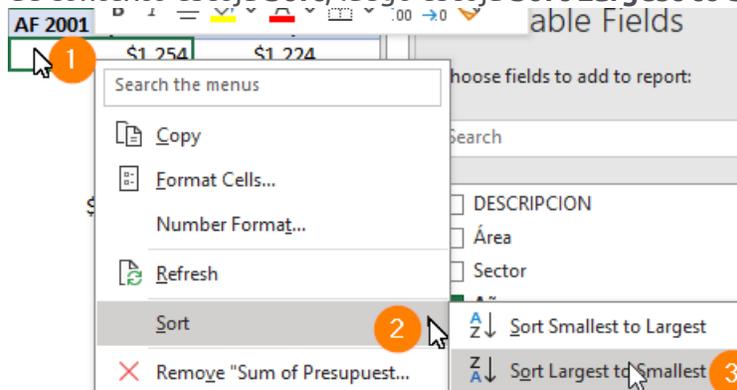
Notará que todas las columnas de años fiscales en la tabla Pivot, cambiarán a formato **Currency**:

2001 Apr	AF 2002 Apr	AF 2003 Apr	AF 2004 Apr	AF 2005 Apr	AF 2006 Apr	AF 2007 Apr	AF 2008 Apr	AF 2009 Apr	AF 2010 Apr	AF 2011 Apr	AF 2012 Apr	AF 2013 Apr	AF 2014 Apr	AF 2015 Apr	AF 2016 Apr
\$1,254	\$1,224	\$1,317	\$1,240	\$1,696	\$1,628	\$1,627	\$1,063	\$1,405	\$7,561	\$27,843	\$104,194	\$39,384	\$36,115	\$24,575	\$9,350
\$9,350	\$9,232	\$9,245	\$9,986	\$10,257	\$7,296	\$6,255	\$6,044	\$4,443	\$4,568	\$4,641	\$4,535	\$4,468	\$4,288	\$114,629	\$85,663
\$332,724	\$315,500	\$343,525	\$383,773	\$380,794	\$478,804	\$439,558	\$403,090	\$406,069	\$0	\$0	\$0	\$0	\$0	\$0	\$0

ORDENAR DE MANERA DESCENDENTE

Vamos a ordenar el presupuesto del año 2001 para ver las agencias con mayor presupuesto.

- Haga **right click** en la **celda B5** o el primer valor presupuestario del año 2001. En el menú de contexto escoja **Sort**, luego escoja **Sort Largest to Smallest**.



- ¿Cuáles son las primeras 5 entidades públicas con mayor presupuesto en 2001? ¿Cómo esto cambió en 2016?

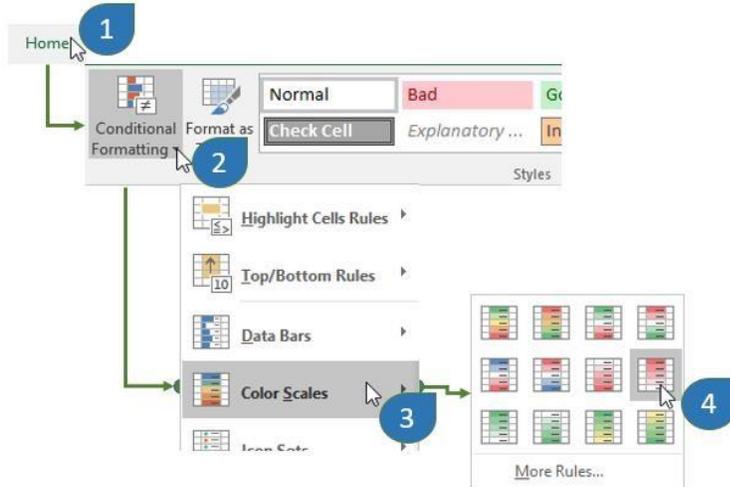
REMOVER TOTALES PARA LAS FILAS

Para propósitos de este ejercicio, no usaremos los valores de Grand Total de las filas.

Apr	AF 2016 Apr	Grand Total
\$1,575	\$0	\$252,126
\$4,288	\$3,840	\$104,703
\$4,266	\$113,266	\$1,602,223
\$0	\$0	\$3,483,837
\$1,254	\$0	\$1,830,097

- Estas se pueden **remover** yendo al **menú principal** en **Design**

- Vaya al **menú principal**, click en el tab **Home**. Haga **click** en el botón **Conditional Formatting**. En el menú de contexto escoja **Color Scales** y luego **escoja** la opción **Red - White** en la **segunda fila, cuarta columna**.



Si ordenó los datos en las secciones anteriores, el resultado deberá verse como este. Note cómo resaltan los valores más altos en rojo. Así podrá identificar más fácilmente los valores extremos.

Column Labels	AF 2001 Apr	AF 2002 Apr	AF 2003 Apr	AF 2004 Apr	AF 2005 Apr	AF 2006 Apr	AF 2007 Apr	AF 2008 Apr	AF 2009 Apr	AF 2010 Apr	AF 2011 Apr	AF 2012 Apr	AF 2013 Apr	AF 2014 Apr	AF 2015 Apr	AF 2016 Apr
	\$2,673,037	\$2,642,173	\$2,705,900	\$3,094,059	\$3,630,328	\$4,272,704	\$4,304,661	\$5,043,958	\$4,496,273	\$4,569,730	\$4,832,449	\$5,259,977	\$5,172,201	\$4,977,514	\$4,884,877	\$4,901,749
	\$2,303,022	\$2,269,323	\$2,697,452	\$2,550,989	\$2,688,655	\$3,624,215	\$3,438,995	\$3,546,651	\$3,548,156	\$3,822,139	\$3,655,058	\$3,714,259	\$3,374,077	\$3,653,007	\$3,294,743	\$2,904,300
	\$1,475,056	\$1,505,138	\$1,575,573	\$1,597,895	\$1,686,526	\$1,704,348	\$1,743,746	\$1,813,684	\$2,049,123	\$2,374,104	\$2,188,334	\$2,187,447	\$2,179,405	\$2,084,746	\$2,087,143	\$2,083,516
	\$1,296,428	\$1,271,244	\$1,247,901	\$1,362,671	\$1,359,410	\$1,592,551	\$1,641,700	\$1,723,432	\$1,495,025	\$1,628,908	\$2,017,294	\$1,994,698	\$2,250,274	\$2,226,236	\$2,223,908	\$2,823,427
	\$1,185,783	\$1,033,122	\$1,021,329	\$1,302,056	\$1,060,791	\$1,501,870	\$1,523,456	\$1,591,807	\$1,084,843	\$1,012,907	\$1,035,496	\$2,331,484	\$2,049,915	\$1,298,227	\$1,242,800	\$1,242,800
	\$1,068,373	\$1,172,542	\$1,195,733	\$1,146,912	\$1,264,191	\$1,445,712	\$1,455,604	\$1,479,598	\$1,831,820	\$1,585,362	\$1,527,242	\$1,513,297	\$1,590,302	\$1,604,691	\$1,472,789	\$1,479,805
	\$757,304	\$807,687	\$758,902	\$845,284	\$812,752	\$879,845	\$1,253,913	\$1,569,105	\$1,296,903	\$1,257,022	\$1,191,904	\$1,435,426	\$1,242,096	\$1,464,049	\$1,506,801	\$1,436,112
	\$626,339	\$744,210	\$682,795	\$646,642	\$480,159	\$593,359	\$489,357	\$400,451	\$244,816	\$409,933	\$518,063	\$403,834	\$392,589	\$523,372	\$463,351	\$479,101
	\$590,821	\$591,544	\$669,011	\$797,981	\$759,188	\$867,493	\$862,780	\$886,995	\$933,897	\$792,250	\$750,568	\$786,841	\$814,741	\$847,953	\$853,957	\$752,494
	\$562,320	\$385,763	\$314,771	\$343,743	\$366,859	\$237,462	\$248,037	\$372,710	\$350,486	\$363,841	\$175,392	\$127,996	\$121,683	\$219,922	\$125,244	\$124,171
	\$528,992	\$593,482	\$650,691	\$792,157	\$742,088	\$829,838	\$824,483	\$823,728	\$951,347	\$955,216	\$896,910	\$876,186	\$830,169	\$816,986	\$784,900	\$747,358
	\$363,047	\$331,499	\$399,730	\$418,764	\$704,289	\$383,312	\$525,467	\$410,621	\$389,147	\$979,990	\$1,031,392	\$1,096,095	\$937,841	\$955,002	\$905,415	\$915,839
	\$356,747	\$406,100	\$648,444	\$439,431	\$508,365	\$537,652	\$422,374	\$457,344	\$374,189	\$485,184	\$373,302	\$412,558	\$401,586	\$462,849	\$418,535	\$394,580
	\$347,354	\$432,457	\$439,480	\$451,569	\$490,362	\$235,340	\$632,322	\$571,081	\$402,594	\$404,406	\$311,988	\$295,680	\$242,176	\$275,763	\$857,718	\$1,126,210
	\$332,724	\$315,500	\$343,525	\$383,773	\$380,794	\$478,804	\$439,558	\$403,090	\$406,069	\$0	\$0	\$0	\$0	\$0	\$0	\$0
	\$326,210	\$375,231	\$311,980	\$366,927	\$488,795	\$480,165	\$449,572	\$433,595	\$552,776	\$489,786	\$469,309	\$464,520	\$405,898	\$624,147	\$390,196	\$397,752
	\$317,613	\$327,545	\$361,462	\$173,327	\$234,404	\$206,573	\$217,316	\$223,591	\$225,079	\$265,870	\$250,971	\$234,627	\$234,556	\$230,900	\$228,623	\$231,407
	\$297,061	\$320,362	\$319,705	\$357,292	\$394,807	\$405,319	\$422,983	\$421,373	\$468,501	\$468,586	\$436,850	\$528,518	\$467,684	\$468,929	\$461,909	\$461,909

- Navegue hacia abajo... podrá ver otros valores altos en rojo.

MOSTRAR VALORES COMO PORCENTAJE DEL TOTAL

Ya que hemos *vectorizado* los valores por año fiscal, tendremos acceso a las opciones que provee el menú de contexto **Show Values as**.

No Calculation	Presenta los valores, solo con el cómputo agregado.
% of Grand Total	Porcentaje del total.
% of Column Total	Porcentaje del total de una columna.
% of Row Total	Porcentaje del total de una fila.
% of ...	Porcentaje comparando con otra columna.
% of Parent Row Total	Porcentaje de una fila en relación con el total de la fila madre. Mostrará el porcentaje correcto de la fila madre con relación al total global de la columna.
% of Parent Column Total	Porcentaje en relación con el total de la columna
% of Parent Total	Porcentaje en relación con el total de la fila madre. La diferencia es que la fila madre presentará valor de 100%.
Difference From	Resta entre una columna y otra.
% Difference From	Porcentaje de diferencia entre dos columnas.
Running Total In	Total, acumulado, usualmente entre fechas.
% Running Total In	Porcentaje acumulado, usualmente entre fechas.
Rank Smallest to Largest	Posicionamiento ordenado de menor a mayor.
Rank Largest to Smallest	Posicionamiento ordenado de mayor a menor.
Index	Convierte el valor a un índice ponderado como se explicó anteriormente.

No vamos a practicarlos todos aquí, pero haremos algunos.

Experimente presentar los valores como:

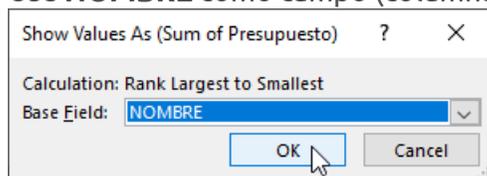
Percent of Grand Total (¿Qué representan los valores?) _____

Percent of Column Total (¿Qué representan los valores?) _____

Percent of Row Total (¿Qué representan los valores?) _____

Rank: Largest to Smallest & Smallest to Largest.

Use **NOMBRE** como campo (columna base) para la comparación

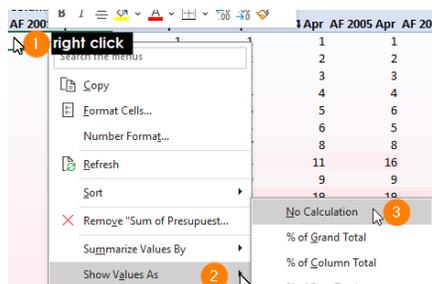


Note cómo algunas agencias mantienen su posición a lo largo de los años. Otras pueden cambiar drásticamente.

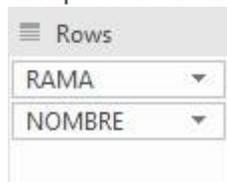
OTRAS OPCIONES: AGRUPAR POR JERARQUÍA

En esta parte usaremos la **jerarquía existente** para agrupar agencias y presentaremos los valores por grupo o Parent Row (Ramas Gubernamentales) y Agencias como filas.

- Haga **right click** en cualquiera de las **celdas** con números y escoja **Show Values As | No calculation**



- Para agrupar por jerarquía, **arrastre el campo RAMA a la caja Rows**, justo **encima** del campo NOMBRE



Podrá ver ahora la tabla jerarquizada por rama gubernamental, mostrando además las agencias (o Aportaciones presupuestarias) que las componen.

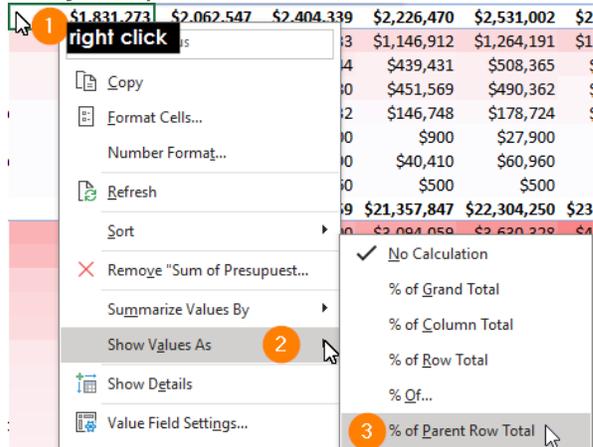
Sum of Presupuesto		Column Labels				
Row Labels	AF 2001 Apr	AF 2002 Apr	AF 2003 Apr	AF 2004 Apr	AF 2005 Apr	
A	\$1,831,273	\$2,062,547	\$2,404,339	\$2,226,470	\$2,531,002	
Aportaciones a los Municipios	\$356,747	\$406,100	\$648,444	\$439,431	\$508,365	
Aportaciones a los Partidos Políticos	\$12,029	\$900	\$900	\$900	\$27,900	
Aportaciones al Quehacer Cultural	\$360	\$360	\$360	\$500	\$500	
Aportaciones para Pensiones y Seguridad Social - Sistema de Retiro	\$43,126	\$43,624	\$96,732	\$146,748	\$178,724	
Aportaciones para Pensiones y Seguridad Social - Sistema de Retiro	\$3,284	\$6,564	\$22,690	\$40,410	\$60,960	
Deuda Pública	\$347,354	\$432,457	\$439,480	\$451,569	\$490,362	
Universidad de Puerto Rico	\$1,068,373	\$1,172,542	\$1,195,733	\$1,146,912	\$1,264,191	
E	\$17,579,732	\$17,945,070	\$19,431,459	\$21,357,847	\$22,304,250	
Administración de Asuntos Energéticos (AAE)	\$1,254	\$1,224	\$1,317	\$1,240	\$1,696	
Administración de Asuntos Federales de Puerto Rico	\$9,350	\$9,232	\$9,245	\$9,986	\$10,257	
Administración de Compensaciones por Accidentes de Automóviles	\$114,629	\$85,663	\$92,890	\$103,463	\$102,780	
Administración de Corrección	\$332,724	\$315,500	\$343,525	\$383,773	\$380,794	
Administración de Desarrollo Laboral	\$218,099	\$262,087	\$224,984	\$94,878	\$107,560	
Administración de Desarrollo Socioeconómico de la Familia	\$1,475,056	\$1,505,138	\$1,575,573	\$1,597,895	\$1,686,526	
Administración de Desarrollo y Mejoras de Viviendas	\$15,469	\$18,541	\$19,604	\$16,717	\$13,051	

Al organizar por jerarquía, podemos notar otros patrones: qué aportaciones aumentaron o se redujeron en la "Rama A" de Aportaciones Presupuestarias.

PERCENT OF PARENT ROW TOTAL

Use esta opción para mostrar los valores presupuestarios como porcentajes *dentro* de una jerarquía. Además, mostrará la participación en porcentaje de cada rama con respecto al total de la columna.

- Haga **right click** en la celda **B5**. En el menú de contexto, escoja **Show Values As**, luego escoja la opción **% of Parent Row Total**.

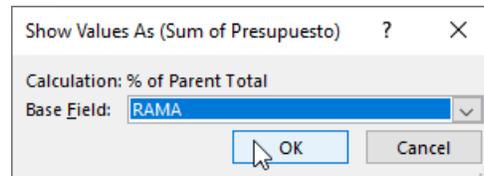
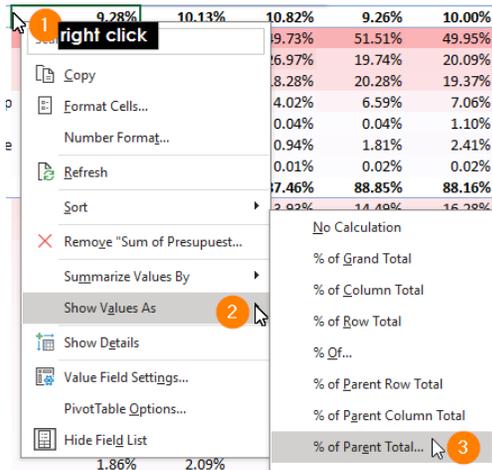


Fíjese cuáles son las proporciones por rama de gobierno y cómo han aumentado o disminuido en **A=Aportaciones del gobierno central**, por ejemplo.

Sum of Presupuesto	Column Labels				
Row Labels	AF 2001 Apr	AF 2002 Apr	AF 2003 Apr	AF 2004 Apr	AF 2005 Apr
A	9.28%	10.13%	10.82%	9.26%	10.00%
Aportaciones a los Municipios	19.48%	19.69%	26.97%	19.74%	20.09%
Aportaciones a los Partidos Políticos	0.66%	0.04%	0.04%	0.04%	1.10%
Aportaciones al Quehacer Cultural	0.02%	0.02%	0.01%	0.02%	0.02%
Aportaciones para Pensiones y Seguridad Social - Sistema de Retiro	2.35%	2.12%	4.02%	6.59%	7.06%
Aportaciones para Pensiones y Seguridad Social - Sistema de Retiro	0.18%	0.32%	0.94%	1.81%	2.41%
Deuda Pública	18.97%	20.97%	18.28%	20.28%	19.37%
Universidad de Puerto Rico	58.34%	56.85%	49.73%	51.51%	49.95%
E	89.04%	88.09%	87.46%	88.85%	88.16%
Administración de Asuntos Energéticos (AAE)	0.01%	0.01%	0.01%	0.01%	0.01%

PERCENT OF PARENT TOTAL

- Utilice la opción **% of Parent Total**. Use el campo **RAMA** como **Base Field**.



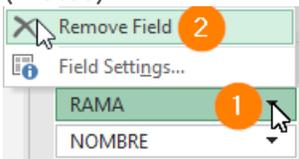
Mostrará las jerarquías como 100% y las proporciones dentro del grupo.

Sum of Presupuesto	Column Labels	AF 2001 Apr	AF 2002 Apr	AF 2003 Apr	AF 2004 Apr	AF 2005 Apr
A		100.00%	100.00%	100.00%	100.00%	100.00%
Aportaciones a los Municipios		19.48%	19.69%	26.97%	19.74%	20.09%
Aportaciones a los Partidos Políticos		0.66%	0.04%	0.04%	0.04%	1.10%
Aportaciones al Quehacer Cultural		0.02%	0.02%	0.01%	0.02%	0.02%
Aportaciones para Pensiones y Seguridad Social - Sistema de Retiro de Empleados		2.35%	2.12%	4.02%	6.59%	7.06%
Aportaciones para Pensiones y Seguridad Social - Sistema de Retiro de Maestros		0.18%	0.32%	0.94%	1.81%	2.41%
Deuda Pública		18.97%	20.97%	18.28%	20.28%	19.37%
Universidad de Puerto Rico		58.34%	56.85%	49.73%	51.51%	49.95%
E		100.00%	100.00%	100.00%	100.00%	100.00%
Administración de Asuntos Energéticos (AAE)		0.01%	0.01%	0.01%	0.01%	0.01%

INDEX

Como se [explicó anteriormente](#), esta opción mostrará cada celda con métricas (measures) ponderando el valor relativo en términos del total de filas y total de columnas.

- Primero, **remueva** el campo **RAMA** de la **caja Rows**, **desplazándolo** al área de las columnas (**Fields**).



- Haga **right click** en cualquier celda con números y escoja **Show Values as | Index**

Asignaciones bajo la Custodia de la Oficina de Gerencia y Presupuesto	0.092418256	0.802529807	0.656599339	0.603930918	0.475722881
Asignaciones para Propósitos Especiales	2.512128199	3.283449478	4.305437566	4.100114758	4.362199123
Asignaciones para Propósitos Generales	10.34221877	10.64664276	0	0	0
Autoridad de Acueductos y Alcantarillados	0.872323744	0.901704681	0.776788017	0.79971335	0.730546541
Autoridad de Carreteras y Transportación	1.175230608	0.99239204	0.899482907	1.059916838	0.820408402
Autoridad de Conservación y Desarrollo de Culebra	0.976072734	0.91384617	1.042547919	0.854601479	1.447165435
Autoridad de Desperdicios Sólidos de Puerto Rico	0.60239344	2.502755702	7.573704977	0.818309511	0.493080531
Autoridad de Edificios Públicos	1.64947375	1.899526005	1.597848735	1.398698525	0.986741574
Autoridad de Energía Eléctrica	0.845049249	0.809565679	0.760148021	0.803396425	0.895581913
Autoridad de las Navieras	4.495401424	4.356586718	3.996623313	3.694095294	0.764427943
Autoridad de los Puertos	0.857819601	0.858018411	1.333443635	1.312461011	1.172973631
Autoridad de Tierras de Puerto Rico	1.191661421	0.974134725	0.720057545	0.621711858	0.823364156
Autoridad de Transporte Integrado	0	0	0	0	0
Autoridad de Transporte Marítimo de Puerto Rico y las Islas Municipio	1.026234509	1.035452028	1.044025787	0.943034663	0.783523364
Autoridad del Centro de Convenciones de Puerto Rico	21.32718924	0	0	0	0
Autoridad del Distrito del Centro de Convenciones	0.380514182	0	0.9231733	1.545443178	2.173123039
Autoridad del Puerto de las Américas	0	0	0.253538762	0.234346916	0.824396672

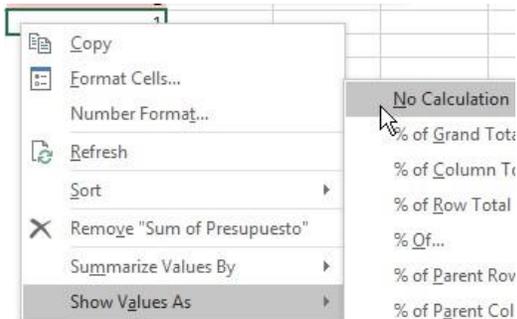
Recuerde que este índice puede verse influido *por grandes diferencias entre columnas*. Por eso algunos índices pueden aparecer muy altos. Pueden tener una asignación significativa en un año, pero en los demás años no aparecerán aportaciones.

DIFFERENCE FROM, % DIFFERENCE FROM

¿Cuáles son las diferencias entre años, para cada agencia, por ejemplo? Para esta y otras opciones de **Show Values As**, haremos algunos cambios.

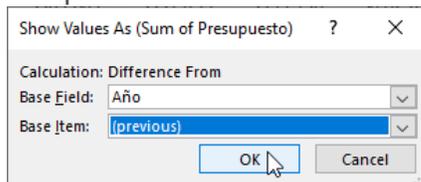
Para ver estas diferencias:

- Quite los valores como índices. Haga **right click** en cualquiera de los valores de años y escoja **Show Values as | No calculation**

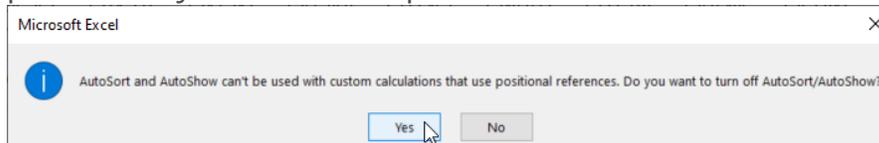


De esta manera, podrá ver los valores originales agregados antes de mostrar las diferencias por año.

- Haga **right click encima** de cualquiera de las celdas con números.
- En el menú de contexto que aparecerá, escoja **Show Values as | Difference from**.
- En **Base Field**, escoja el campo **Año**. En **Base Item**, escoja **(previous)**... lo cual representa comparar con el año anterior.



- Si le aparece esta forma de aviso, haga **click** en el botón **Yes** para quitar las referencias posicionales y ordenamientos previos.

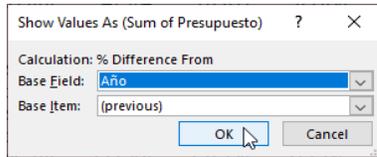


Notará que los valores negativos aparecerán en colores blancos, mientras que los aumentos sustanciales a las asignaciones presupuestarias aparecerán en colores rojos más oscurecidos.

Column Labels	AF 2002 Apr	AF 2003 Apr	AF 2004 Apr	AF 2005 Apr	AF 2006 Apr	AF 2007 Apr	AF 2008 Apr	AF 2009 Apr	AF 2010 Apr	
AF 2001 Apr										
		-\$30	\$93	-\$77	\$456	-\$68	-\$1	-\$564	\$342	\$6,156
		-\$118	\$13	\$741	\$271	-\$2,961	-\$1,041	\$0	-\$211	-\$1,601
		-\$28,966	\$7,227	\$10,573	-\$683	-\$778	\$1,584	-\$970	-\$4,858	-\$2,962
		-\$17,224	\$28,025	\$40,248	-\$2,979	\$98,010	-\$39,246	-\$36,468	\$2,979	-\$406,069
		\$43,988	-\$37,103	-\$130,106	\$12,682	-\$12,213	-\$13,871	\$26,522	\$37,709	\$11,784
		\$30,082	\$70,435	\$22,322	\$88,631	\$17,822	\$39,398	\$69,938	\$235,439	\$324,981
		\$2,077	\$1,063	-\$2,887	-\$3,666	-\$1,483	-\$11,568	\$0	\$0	\$0

- Repita el proceso**; esta vez, utilice **Show Values as | % Difference From**.

Recuerde usar: **Año** como **Base Field** y **(previous)** como **Base Item**



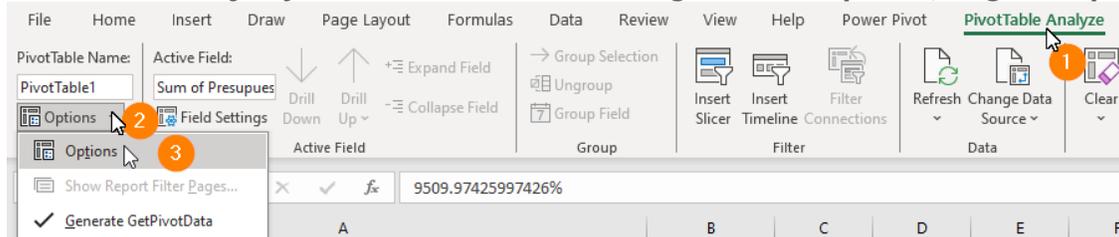
Note que aparecerán celdas con valor **#DIV/0!**. Esto se debe a que al comparar hay valores cero en el denominador.

Corporación del Centro Regional del ELA de PR	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
Corporación del Conservatorio de Música de Puerto Rico	-43.71%	-3.24%	55.04%	57.69%	59.98%	-18.10%	-26.59%	-37.47%	9.11%
Corporación del Fondo del Seguro del Estado	7.84%	-0.21%	11.76%	10.50%	2.66%	4.36%	-0.38%	11.18%	0.02%
Corporación del Proyecto ENLACE del Caño Martín Peña	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	-7.60%
Corporación para el Desarrollo de las Artes, Ciencias e Industria Cinematogr	28.00%	228.19%	-9.88%	0.37%	-15.33%	19.71%	-17.16%	5.82%	-8.61%
Corporación para el Desarrollo de las Exportaciones de Puerto Rico	-2.62%	-9.29%	56.82%	-100.00%	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
Corporación para el Desarrollo Rural	-65.95%	13.74%	143.29%	1.09%	-59.29%	35.10%	-47.70%	-51.16%	-100.00%
Corporación Pública para la Supervisión y Seguro de Cooperativas de Puerto R	-1.33%	17.80%	8.01%	-6.52%	-4.79%	10.67%	5.70%	20.67%	24.72%
Cuerpo de Bomberos de Puerto Rico	3.45%	9.06%	21.78%	8.69%	12.75%	1.47%	0.72%	5.98%	-11.96%
Cuerpo de Emergencias Médicas de Puerto Rico	13.97%	-0.04%	32.59%	-19.83%	48.14%	-6.45%	9.01%	2.08%	-3.29%
Departamento de Agricultura	78.75%	-28.44%	-30.75%	24.83%	58.68%	-42.00%	-4.69%	-3.08%	10.68%
Departamento de Asuntos del Consumidor	-3.68%	14.03%	3.24%	13.28%	4.92%	4.18%	-1.06%	7.01%	-23.88%
Departamento de Corrección y Rehabilitación	-17.71%	90.43%	9.65%	15.08%	17.55%	-20.83%	2.13%	-0.81%	9509.97%
Departamento de Desarrollo Económico y Comercio	91.87%	61.71%	2.16%	70.18%	7.50%	10.00%	107.10%	48.80%	601.83%

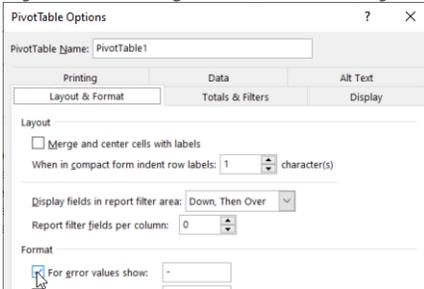
Hay una manera de “esconder” este valor en PivotTables

ESCONDER VALORES DE ERROR EN PIVOTTABLE

- Para **esconder** el valor de error **#DIV/0!**, vaya al **menú principal**, haga click en el tab **PivotTable Analyze** y en la sección **PivotTable** haga click en **Options**, luego en **Options**



- Vaya al tab **Layout & Format** y bajo **Format**,



- Haga **check** en la opción **For error values show**:
- En la caja de texto, **escriba** un cero (0) o una raya (-).

En la fila de totales podrá ver los cambios porcentuales globales.

3.18%	9.07%	8.19%	5.25%	4.99%	0.65%	3.64%	-3.49%	9.35%	-4.38%	6.88%	-4.03%	1.02%	-2.99%	2.50%
-------	-------	-------	-------	-------	-------	-------	--------	-------	--------	-------	--------	-------	--------	-------

RUNNING TOTAL IN Y % RUNNING TOTAL IN

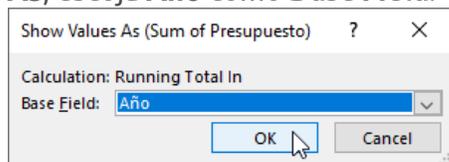
Esta opción le mostrará la acumulación de valores (presupuestos), en este caso, a medida que pasan los años.

NOTA: En este caso (presupuesto) no es que haya necesariamente acumulaciones; por lo tanto, no hace mucho sentido usarlo porque el presupuesto gubernamental no es para acumular dinero; se asigna y se gasta, por lo general completamente, con pequeños ahorros o con deudas...

Si quiere ver los números de todos modos, este sería el proceso:

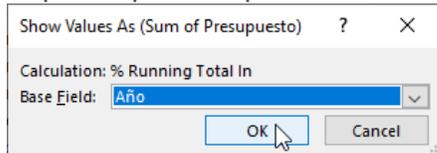
- Haga **right click** encima de cualquier celda que contenga valores numéricos.

- En el menú de contexto escoja **Show Values as | Running Total In**. En la caja **Show Values As**, escoja **Año** como **Base Field**.



\$0	\$0	\$0	\$0	\$0	\$0	\$275,000
9,977	\$3,272,077	\$3,922,108	\$4,384,058	\$5,263,079	\$5,894,904	\$6,932,275
6,744	\$1,136,744	\$1,136,744	\$1,136,744	\$1,136,744	\$1,136,744	\$1,136,744
2,567	\$72,567	\$72,567	\$72,567	\$72,567	\$72,567	\$72,567
8,717	\$11,430,621	\$12,866,047	\$14,108,143	\$15,572,192	\$17,078,993	\$18,515,105
7,964	\$13,353,460	\$15,684,944	\$17,734,859	\$19,033,086	\$20,275,886	\$21,518,686
7,274	\$7,821	\$8,340	\$8,880	\$9,542	\$10,272	\$10,925
0,232	\$220,629	\$230,042	\$237,993	\$248,498	\$254,222	\$258,202
8,061	\$5,836,124	\$6,239,958	\$6,632,547	\$7,155,919	\$7,619,270	\$8,098,371
2,823	\$42,265,272	\$47,525,249	\$52,697,450	\$57,674,964	\$62,559,841	\$67,461,590
4,777	\$114,777	\$114,777	\$114,777	\$114,777	\$114,777	\$114,777

- Repita el proceso para **Show Values as | % Running Total in**, usando **Año** como **Base Field**



Notará que los porcentajes van aumentando hacia 100% hacia el AF 2016 en las filas. Recuerde que la aplicación de esta parte es solo demostración. El presupuesto no se acumula.

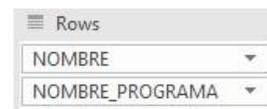
OTROS DESGLOSES

Usted podrá agregar y resumir datos usando las dimensiones que estén disponibles para analizar los datos. Para este conjunto de datos presupuestarios, podemos **agregar datos por origen de fondos, concepto de gasto, cuenta, área, sector**, etcétera.

- Este es un ejemplo que puede aplicar, desplazando el campo (field, columna)

NOMBRE_PROGRAMA a la caja Rows

Sum of Presupuesto	Column Labels	
Row Labels	AF 2001 Apr	AF 2002 Apr
Autoridad de Energía Eléctrica	\$2,673,037	\$2,642,173
Operación del Sistema Eléctrico e Ingeniería	\$1,516,638	\$1,514,820
Deuda Pública y Gerencial	\$386,853	\$409,541
Transmisión y Distribución de Energía Eléctrica	\$291,136	\$254,332
Apoyo Técnico y Administrativo	\$210,587	\$215,572
Contribución a los Municipios	\$121,890	\$105,696
Facturación y Cobro del Servicio Eléctrico y Atención al Cliente	\$101,100	\$102,889
Contribución al Gobierno de Puerto Rico	\$30,253	\$21,993
Dirección y Administración General	\$7,579	\$8,201
Administración y Conservación de los Sistemas de Riego Público	\$7,001	\$9,129
Departamento de Educación	\$2,303,022	\$2,269,323
Escuelas de la Comunidad	\$1,655,157	\$1,596,449
Servicios de Alimentos al Estudiante	\$27,473	\$16,308



- Este otro ejemplo muestra el detalle añadiendo el campo **CONCEPTO_DE_GASTO**

Sum of Presupuesto	Column Labels			
Row Labels	AF 2001 Apr	AF 2002 Apr	AF 2003 Apr	AF 2004 Apr
Autoridad de Energía Eléctrica	\$2,673,037	\$2,642,173	\$2,705,900	\$3,094,059
Operación del Sistema Eléctrico e Ingeniería	\$1,516,638	\$1,514,820	\$1,500,946	\$1,780,187
Servicios Comprados	\$886,748	\$834,770	\$176,423	\$436,762
Otros Gastos Operacionales	\$290,686	\$273,991	\$922,028	\$888,714
Inversión en Mejoras Permanentes	\$204,854	\$264,068	\$259,580	\$267,459
Nómina y Costos Relacionados	\$127,809	\$135,256	\$135,712	\$152,273
Gastos de Transportación	\$4,898	\$4,265	\$4,351	\$5,345
Materiales y Suministros	\$1,643	\$2,470	\$2,852	\$29,634
Facilidades y Pagos por Servicios Públicos	\$0	\$0	\$0	\$0
Servicios Profesionales	\$0	\$0	\$0	\$0
Deuda Pública y Gerencial	\$386,853	\$409,541	\$433,924	\$481,085
Pago de la Deuda	\$386,853	\$409,541	\$433,924	\$481,085
Transmisión y Distribución de Energía Eléctrica	\$291,136	\$254,332	\$265,053	\$348,096
Inversión en Mejoras Permanentes	\$135,290	\$104,045	\$110,797	\$157,450
Nómina y Costos Relacionados	\$130,547	\$124,047	\$124,729	\$157,777
Otros Gastos Operacionales	\$12,672	\$9,877	\$12,332	\$15,832
Materiales y Suministros	\$6,097	\$9,183	\$10,263	\$11,864
Gastos de Transportación	\$4,610	\$5,474	\$5,584	\$5,173



Los ejemplos muestran los valores presupuestarios en sumas agregadas. Usted puede experimentar usando las opciones contenidas en **Show Values as**.

EJERCICIO 3: CREAR PIVOTCHART

Antes de terminar, haremos un ejemplo de una gráfica pivot usando el mismo conjunto de datos presupuestarios. Para hacer la gráfica Pivot primero debemos preparar los datos un poco:

- Para **simplificar** la gráfica, vamos a **aplicar** un **filtro (Top 10)** para seleccionar un grupo de agencias con los presupuestos más grandes.
- Presentaremos los **datos** presupuestarios **agregados** solamente (**No Calculation**)
- Cambiaremos los **Años** a las filas (**Rows**) y las 10 a 15 **agencias** seleccionadas a la caja **Columns**.

Regrese los valores de las celdas a datos agregados (No Calculation).

- Haga **right click** en cualquier celda numérica. En el menú de contexto escoja **Show Values as | No Calculation**.
- En el panel **Pivot Table Fields**, **remueva** las columnas **NOMBRE_PROGRAMA** y **CONCEPTO_DE_GASTO** de la caja Rows

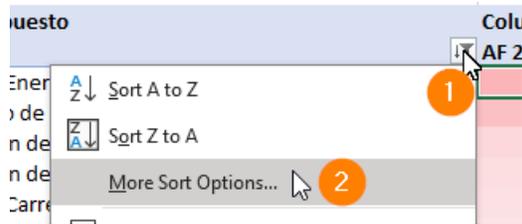
APLICAR FILTRO TOP 10

- Haga **click** en el **botón** al lado derecho de la celda ancha **Row Labels**, **click** en **Value Filters | Top 10**. Aparecerá la forma **Top 10 Filter**. Aplique (escriba) filtro para **11 Items** by **Sum of Presupuesto**

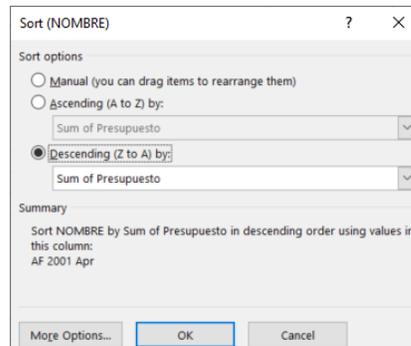


- Ordene los datos de forma descendente para que la gráfica aparezca más organizada.

- Haga **click** en el **botón de filtro** en la fila **Row Labels**. Escoja la opción **More Sort Options**.



- En la forma **Sort**, escoja la opción **Descending (Z to A) by**, luego escoja el ítem **Sum of Presupuesto** de la lista en el combo box. Luego **click** en botón **OK** para aceptar las opciones y ejecutar el ordenamiento.

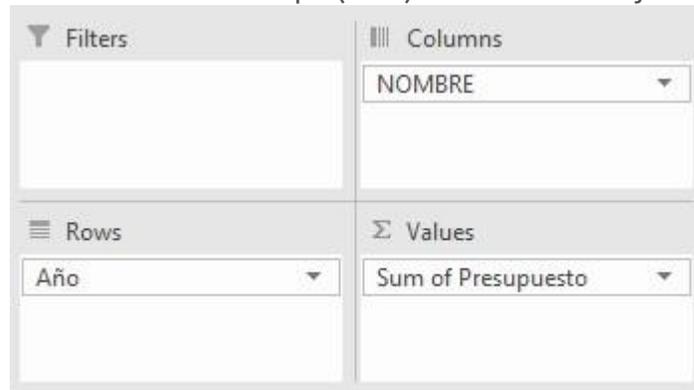


La tabla deberá aparecer así:

Sum of Presupuesto	Column Labels							
Row Labels	AF 2001 Apr	AF 2002 Apr	AF 2003 Apr	AF 2004 Apr	AF 2005 Apr	AF 2006 Apr	AF 2007 Apr	AF 2008 Apr A
Autoridad de Energía Eléctrica	\$2,673,037	\$2,642,173	\$2,705,900	\$3,094,059	\$3,630,328	\$4,272,704	\$4,304,661	\$5,043,958
Departamento de Educación	\$2,303,022	\$2,269,323	\$2,697,452	\$2,550,989	\$2,688,655	\$3,624,215	\$3,438,995	\$3,546,651
Administración de Desarrollo Socioeconómico de la Familia	\$1,475,056	\$1,505,138	\$1,575,573	\$1,597,895	\$1,686,526	\$1,704,348	\$1,743,746	\$1,813,684
Administración de Seguros de Salud de Puerto Rico	\$1,296,428	\$1,271,244	\$1,247,901	\$1,362,671	\$1,359,410	\$1,592,551	\$1,641,700	\$1,723,432
Autoridad de Carreteras y Transportación	\$1,185,783	\$1,033,122	\$1,021,329	\$1,302,056	\$1,060,791	\$1,501,870	\$1,523,456	\$1,591,807
Universidad de Puerto Rico	\$1,068,373	\$1,172,542	\$1,195,733	\$1,146,912	\$1,264,191	\$1,445,712	\$1,455,604	\$1,479,598
Autoridad de Acueductos y Alcantarillados	\$757,304	\$807,687	\$758,902	\$845,284	\$812,752	\$879,845	\$1,253,913	\$1,569,105
Autoridad de Edificios Públicos	\$626,339	\$744,210	\$682,795	\$646,642	\$480,159	\$593,359	\$489,357	\$400,451
Policía de Puerto Rico	\$590,821	\$591,544	\$669,011	\$797,981	\$759,188	\$867,493	\$862,780	\$886,995
Departamento de Salud	\$528,992	\$593,482	\$650,691	\$792,157	\$742,088	\$829,838	\$824,483	\$823,728
Banco Gubernamental de Fomento para Puerto Rico	\$363,047	\$331,499	\$399,730	\$418,764	\$704,289	\$383,312	\$525,467	\$410,621
Grand Total	\$12,868,202	\$12,961,964	\$13,605,017	\$14,555,410	\$15,188,377	\$17,695,247	\$18,064,162	\$19,290,030

CAMBIAR AGENCIAS A COLUMNAS Y AÑOS A ROWS

- Ahora arrastre el campo (field) **NOMBRE** a la caja **Columns** y el campo **Años** a la caja **Rows**



- **Quite** cualquier resumen que parezca a Running Totals. Haga **right click** encima de cualquier celda numérica y escoja **Show Values As | No Calculation**.
- **Para poner las columnas en orden descendente**, haga **right click** en la celda **B21** y escoja **Sort | Largest to Smallest**.

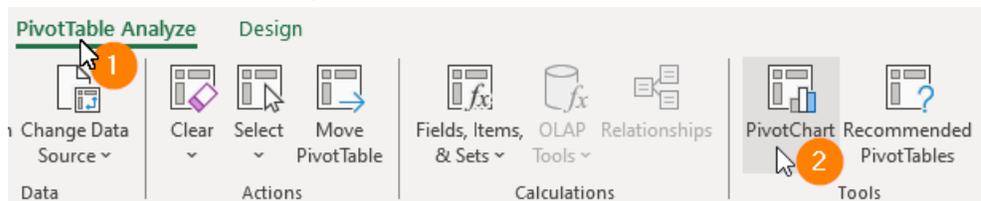
- La tabla debe verse así. Si los años no están ordenados de 2001 a 2016, aplique ordenamiento ascendente en la columna Row Labels.

Sum of Presupuesto	Column Labels		
Row Labels	Autoridad de Energía Eléctrica	Departamento de Educación	Administración de Desarrollo Socioeconómico de la Familia
AF 2001 Apr	\$2,673,037	\$2,303,022	\$1,475,056
AF 2002 Apr	\$2,642,173	\$2,269,323	\$1,505,138
AF 2003 Apr	\$2,705,900	\$2,697,452	\$1,575,573
AF 2004 Apr	\$3,094,059	\$2,550,989	\$1,597,895
AF 2005 Apr	\$3,630,328	\$2,688,655	\$1,686,526
AF 2006 Apr	\$4,272,704	\$3,624,215	\$1,704,348
AF 2007 Apr	\$4,304,661	\$3,438,995	\$1,743,746
AF 2008 Apr	\$5,043,958	\$3,546,651	\$1,813,684
AF 2009 Apr	\$4,496,273	\$3,548,156	\$2,049,123
AF 2010 Apr	\$4,569,730	\$3,822,139	\$2,374,104
AF 2011 Apr	\$4,832,449	\$3,655,058	\$2,188,334
AF 2012 Apr	\$5,259,977	\$3,714,259	\$2,187,447
AF 2013 Apr	\$5,172,201	\$3,374,077	\$2,179,405
AF 2014 Apr	\$4,977,514	\$3,653,007	\$2,084,746
AF 2015 Apr	\$4,884,877	\$3,294,743	\$2,087,143
AF 2016 Apr	\$4,901,749	\$2,904,300	\$2,083,516
Grand Total	\$67,461,590	\$51,085,041	\$30,335,784

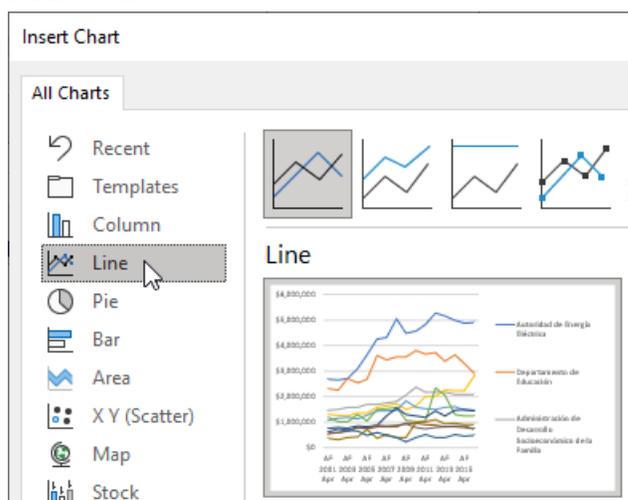
GENERAR GRÁFICA PIVOT

Ahora que esta tabla está lista, pasemos a generar una gráfica Pivot.

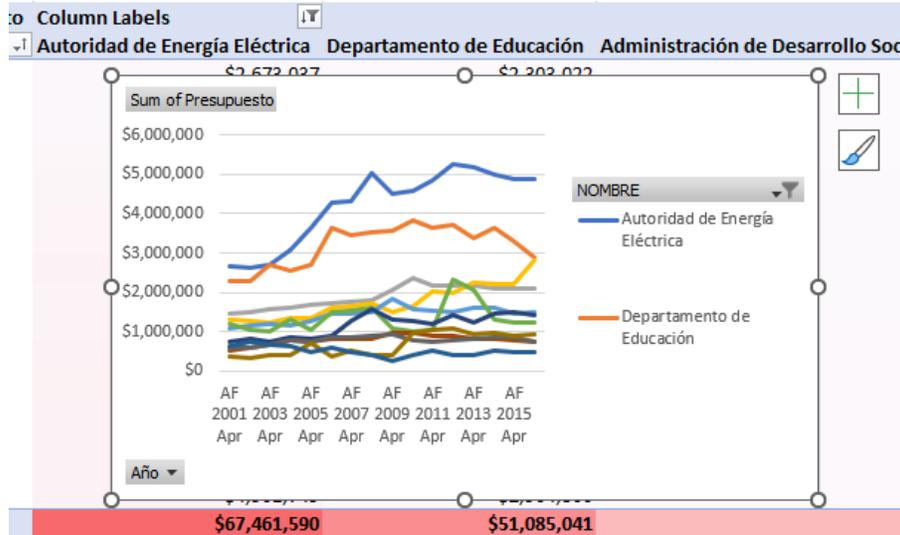
- Primero, asegúrese que la tabla pivot esté activada (click en cualquier celda *dentro* de la tabla pivot).
- Vaya al **menú principal**. Asegúrese de ver en el menú **PivotTable Analyze**.
- Haga **click** en **PivotTable Analyze**
- En la sección **Tools**, haga **click** en el botón **PivotChart**.



- En la forma **Insert Chart** escoja la opción **Line**. Observe la previsualización. Haga **click** en botón **OK**

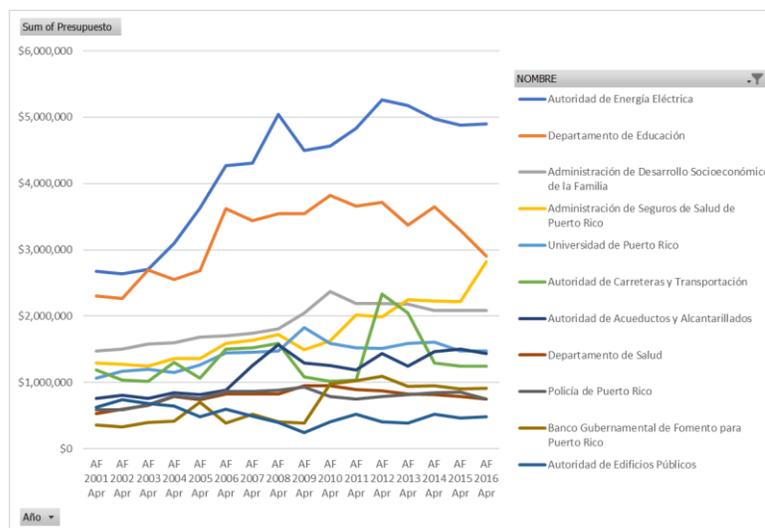


La gráfica Pivot aparecerá en algún lugar dentro o fuera de la tabla Pivot:



- Mueva la gráfica debajo de la tabla. Expanda la gráfica, estirando desde los círculos en las esquinas. Estire lo suficiente para que pueda ver en la leyenda los once ítems

La gráfica debe verse así:

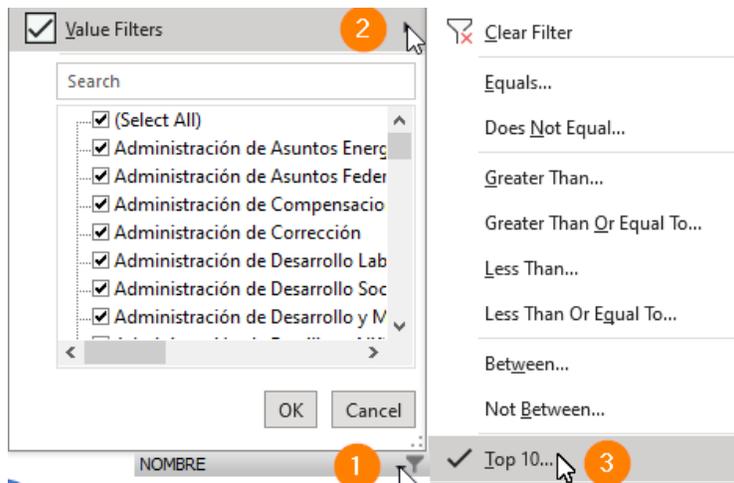


APLICAR FILTRO EN LA GRÁFICA PIVOT (TOP 10, OPCIÓN PERCENT)

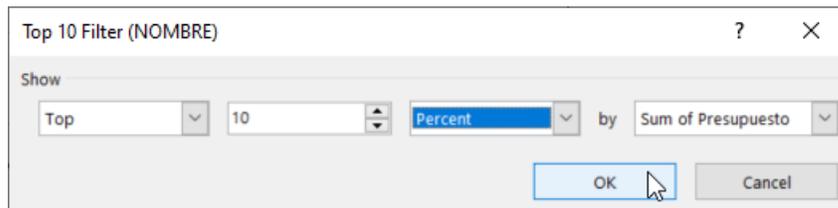
Note que hay varios botones en la gráfica. Estos son para seleccionar datos y presentar subconjuntos de estos.

Vamos a experimentar seleccionar usando el filtro **Top 10** con la opción *Percent*. Aplicaremos filtro para presentar agencias que estén en el rango "Top 10%" ...

- Haga **click** en el botón de filtro **NOMBRE | Value Filters | Top 10 ...**



- En la caja **Top 10 Filter** escoja **Top**, escriba **10** y escoja de la lista a la derecha el ítem **Percent**. **Click** en **OK** para aceptar cambios y ejecutar el filtro.



- Para filtro 10%: ¿Cuántas agencias (entidades gubernamentales) aparecen en la gráfica?
-
- Para filtro 20%: ¿Cuántas agencias (entidades gubernamentales) aparecen en la gráfica?
-
- Para filtro 30%: ¿Cuántas agencias (entidades gubernamentales) aparecen en la gráfica?
-
- Para filtro 50%: ¿Cuántas agencias (entidades gubernamentales) aparecen en la gráfica?
-
- ¿Qué tipo de filtro propone utilizar para considerar solo las agencias de gobierno central?
-
- ¿Qué filtro propone utilizar para obtener agencias que dependen del Fondo General, Ingresos Propios, Fondos Federales, ¿etc.?
-

Esto concluye este ejercicio.

- Guarde** su workbook Excel y nómbrelo **PrácticaPresupuestoPivot.xlsx** en el folder **C:\Tutorial_Excel\DataAnalysis\Modulo_3**

EJERCICIO 4: CAMPOS CALCULADOS EN PIVOTTABLES

En esta parte vamos a practicar generar campos adicionales en PivotTables. En ocasiones, los datos originales no proveen campos calculados, especialmente si vienen de bases de datos. La derivación de datos: [resumen y análisis](#), viene a ser trabajo del *analista de datos*. Las tablas en la base de datos deben ser eficientes y ocupar el menor espacio posible.

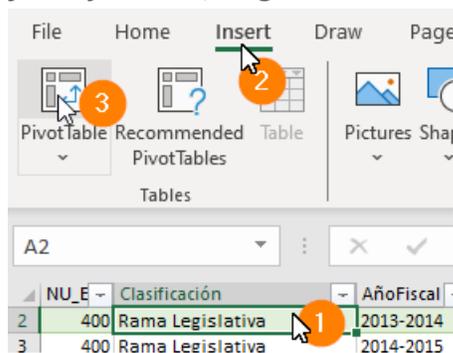
- En este ejercicio usaremos el workbook Excel: **Registro Anual de Puestos por Demografía_26abr2017.xlsx**, localizado en **C:\Tutorial_Excel\DataAnalysis\Modulo_3**

Al abrir este workbook, esta es una muestra de tabla Excel:

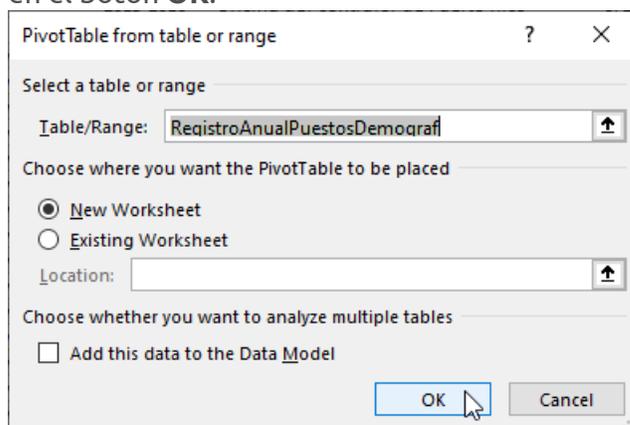
1	A	B	C	D	E	F	G	H	I	J	K	L
2	NU_E	Clasificación	AñoFiscal	NB_ENTIDAD	FechaTran	FechaCierre	Estatus	CarreraParcialM	CarreraParcialF	CarreraCompletoM	CarreraCompletoF	CarreraNoGraduadoM
3	400	Rama Legislativa	2013-2014	Oficina del Contralor de Puerto Rico	1/20/2015 0:00	2/5/2015 0:00	1	0	2	214	227	0
4	400	Rama Legislativa	2014-2015	Oficina del Contralor de Puerto Rico	2/8/2016 0:00	2/18/2016 0:00	1	0	4	211	224	0
5	400	Rama Legislativa	2015-2016	Oficina del Contralor de Puerto Rico	3/1/2017 0:00	3/2/2017 0:00	1	0	0	203	220	0
6	400	Rama Legislativa	2012-2013	Oficina del Contralor de Puerto Rico	2/6/2014 0:00	2/7/2014 0:00	1	0	2	224	232	0
7	1000	Departamentos y Agencias 2013-2014		Departamento de Agricultura	2/6/2015 0:00	2/6/2015 0:00	1	0	0	152	125	2
8	1000	Departamentos y Agencias 2014-2015		Departamento de Agricultura	3/1/2016 0:00	3/1/2016 0:00	1	0	0	158	127	2
9	1000	Departamentos y Agencias 2015-2016		Departamento de Agricultura	2/1/2017 0:00	2/1/2017 0:00	1	0	0	159	132	2
10	1000	Departamentos y Agencias 2012-2013		Departamento de Agricultura	1/22/2014 0:00	1/27/2014 0:00	1	1	0	161	137	2
11	1000	Departamentos y Agencias 2014-2015		Departamento de Estado	1/14/2016 0:00	1/14/2016 0:00	1	0	0	36	69	4

AÑADIR PIVOTTABLE

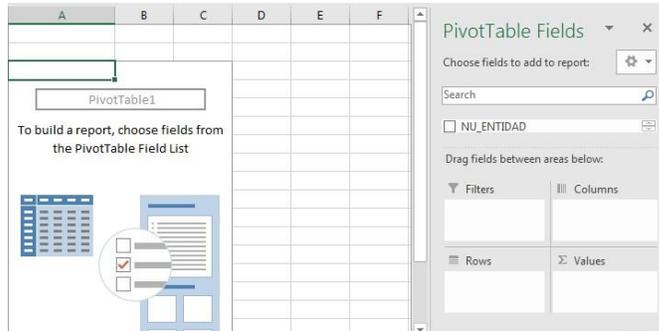
- Para añadir una tabla Pivot, haga **click** en cualquier celda de la tabla, vaya al **menú principal** y escoja **Insert**, luego **click** en el botón **PivotTable**.



- En la forma **PivotTable from table or range**, acepte las **opciones** por defecto y haga **click** en el botón **OK**.



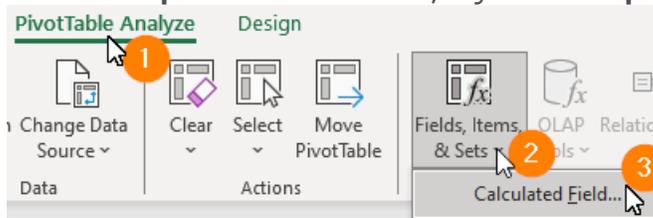
Aparecerá la tabla Pivot vacía, junto con la interfaz PivotTable Fields



AÑADIR CAMPO CALCULADO A LA TABLA PIVOT

Según aparece en la tabla original de datos, los empleados de carrera (métrica) aparecen divididos en múltiples categorías: por género, jornada completa, jornada parcial, niveles educativos, y otras características. A diferencia de la tabla presupuestaria, estos datos son un poco más complejos por la cantidad de columnas. Por lo tanto, si queremos tener un conteo de empleados públicos en la tabla Pivot, podemos generar un campo calculado.

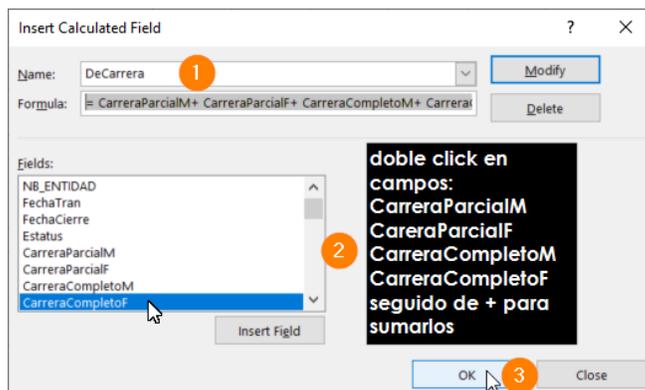
- Con la **tabla pivot vacía activada**, vaya al **menú principal**:



- Click en **Pivot Table Analyze**
- En la sección **Calculations**, haga click en el botón **Fields, Items, & Sets**
- Click en la opción **Calculated Field...**

- En la forma **Insert Calculated Field**:

- En **Name**: escriba **DeCarrera**
- En **Formula**: después del signo de igualdad = borre el cero 0.

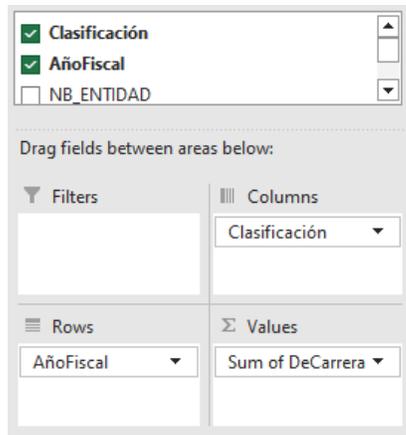


- Use la **lista** de campos en la **sección Fields** para **insertar** los campos que vamos a sumar. Haga **doble click** en los campos: **CarreraParcialM**, **CarreraParcialF**, **CarreraCompletoM**, **CarreraCompletoF** seguidos del **signo de suma +** :
- Click en botón **Add** para añadir el nuevo campo calculado (en la tabla Pivot)
- Click en botón **OK** para aceptar los cambios y salir de esta forma

Su tabla Pivot deberá aparecer así. Aquí aparecerá solamente la sumatoria del nuevo campo calculado **DeCarrera**

2	
3	Sum of DeCarrera
4	402436
5	

- Añada las columnas** para hacer el desglose por **clasificación** de la entidad pública y por **año fiscal**, campo con los años que van desde 2012-13 al 2015-16.



- Arrastre el campo **Clasificación** a la caja **Columns**
- Arrastre el campo **AñoFiscal** a la caja **Rows**

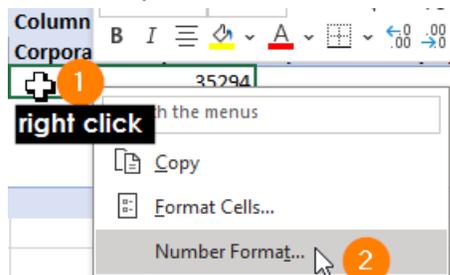
Su tabla Pivot aparecerá así:

Sum of DeCarrera	Column Labels								
Row Labels	Corporaciones públicas	Departamentos y Agencias	Municipios	Oficina del Gobernador	Otras entidades	Rama Judicial	Rama Legislativa	Grand Total	
2012-2013	35294	100668	27872	1069	429	3568	531	169431	
2013-2014	18433	40501	14937	1087	225	3663	812	79658	
2014-2015	18472	44792	19287	818	149	3581	507	87606	
2015-2016	16044	31820	16876	364	147	0	490	65741	
Grand Total	88243	217781	78972	3338	950	10812	2340	402436	

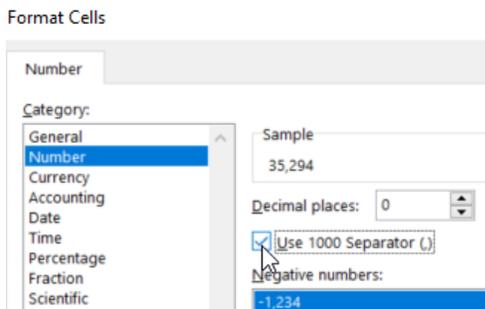
CAMBIAR FORMATO NUMÉRICO

Cambie el formato numérico para añadir una coma para separar millares.

- Haga **right click** encima de cualquiera de las celdas numéricas.
- Click** en opción **Number Format...**



Aparecerá la forma **Format Cells**.



- Click en **Number**. En **Decimal Places**, escriba **0**.
- Check** en opción **Use 1000 separator**.
- Click en **OK** para aceptar cambios y cerrar esta forma.

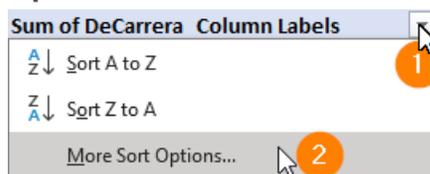
La tabla Pivot debe verse así:

Sum of DeCarrera	Column Labels	Departamentos y Agencias	Municipios	Oficina del Gobernador	Otras entidades	Rama Judicial	Rama Legislativa	Grand Total
Row Labels	Corporaciones públicas							
2012-2013	35,294	100,668	27,872	1,069	429	3,568	531	169,431
2013-2014	18,433	40,501	14,937	1,087	225	3,663	812	79,658
2014-2015	18,472	44,792	19,287	818	149	3,581	507	87,606
2015-2016	16,044	31,820	16,876	364	147	0	490	65,741
Grand Total	88,243	217,781	78,972	3,338	950	10,812	2,340	402,436

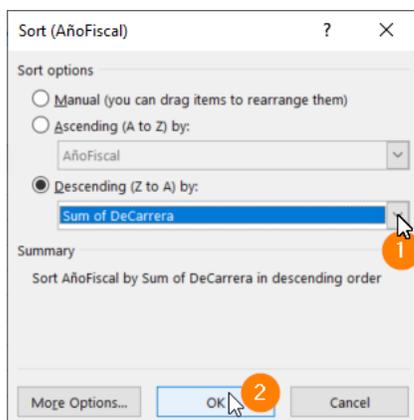
ORDENAR DATOS DE FORMA DESCENDENTE

Ordene los datos de mayor a menor.

- Haga **click** en el **botón** al lado derecho de la celda **Row Labels** y **escoja** la opción **More Sort Options**.



Aparecerá la forma Sort (Año Fiscal)



- En la forma **Sort**, escoja la opción **Descending (Z to A) by**. **Escoja** de la lista el campo **Sum of DeCarrera**.
- Haga **click** en el botón **OK** para aceptar cambios y cerrar esta forma.

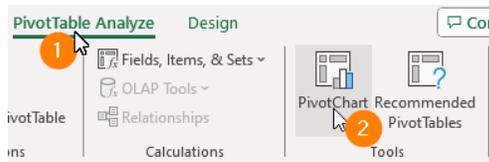
La tabla pivot debe verse así:

Sum of DeCarrera	Column Labels	Departamentos y Agencias	Municipios	Oficina del Gobernador	Otras entidades	Rama Judicial	Rama Legislativa	Grand Total
Row Labels	Corporaciones públicas							
2012-2013	35,294	100,668	27,872	1,069	429	3,568	531	169,431
2014-2015	18,472	44,792	19,287	818	149	3,581	507	87,606
2013-2014	18,433	40,501	14,937	1,087	225	3,663	812	79,658
2015-2016	16,044	31,820	16,876	364	147	0	490	65,741
Grand Total	88,243	217,781	78,972	3,338	950	10,812	2,340	402,436

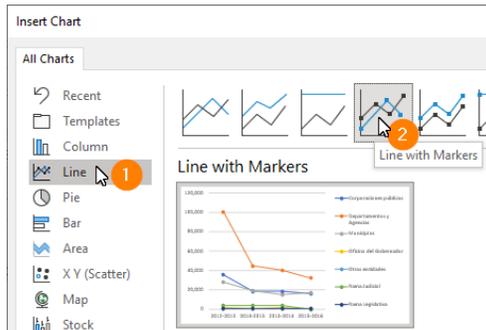
PREPARAR PIVOTCHART

Para terminar, prepare una gráfica Pivot.

- Vaya al menú principal y haga click en **PivotTable Analyze** y en la sección **Tools** haga click en el botón **PivotChart**.

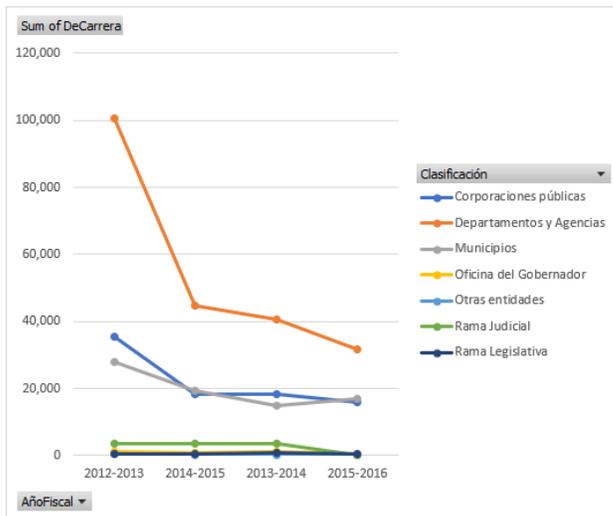


- En la forma **Insert Chart** que aparecerá,



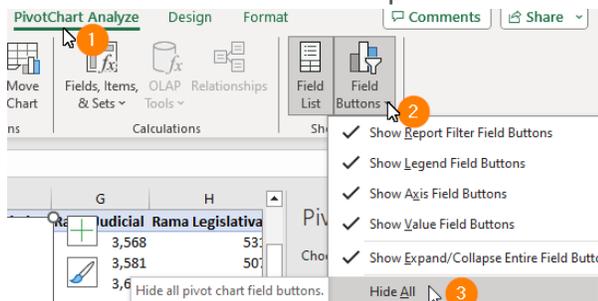
- Haga **click** en la opción **Line** y
- A la derecha, escoja la opción **Line with Markers**.
- Haga **click** en el botón **OK** para aceptar los cambios y cerrar la forma.

Luego de mover y estirar un poco, la gráfica Pivot debe verse así:



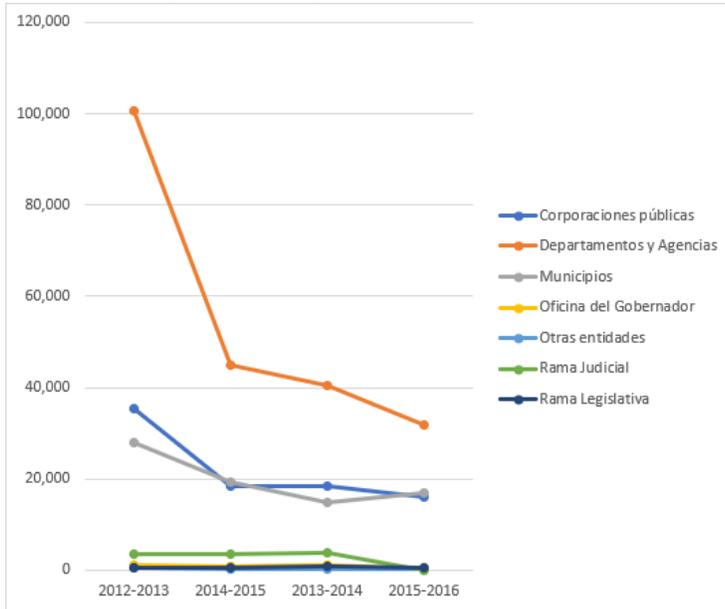
QUITAR BOTONES A LA GRÁFICA PIVOT

Si desea remover los botones para filtrado de la gráfica Pivot. Estos son los pasos:



- Mantenga la gráfica activada.**
- Haga **click** en el tab **PivotChart Analyze**
- En la Sección **Show/Hide**, haga click en el triángulo del botón **Field Buttons**.
- Haga **click** en la opción **Hide All** para remover los botones en la gráfica Pivot.

Así deberá aparecer la gráfica:



¿Cuál fue la clasificación de mayor reducción de personal? _____

Esta reducción de personal en el gobierno central, corporaciones y municipios, fue parte de las múltiples disposiciones de la [Ley #7 de 2009](#).

Al momento de escribir este tutorial, los datos más recientes no están disponibles en el Instituto de Estadísticas. Los datos más recientes están en el portal de la Oficina del Contralor. Para acceder al inventario, es necesario crear una cuenta y la interfaz no provee información de cómo crear una cuenta nueva.

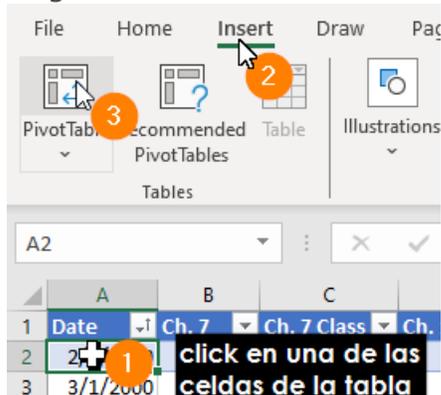
- Guarde este archivo en el folder **C:\Tutorial_Excel\DataAnalysis\Modulo_3**

EJERCICIO 5: AÑADIR SLICERS Y TIMELINE

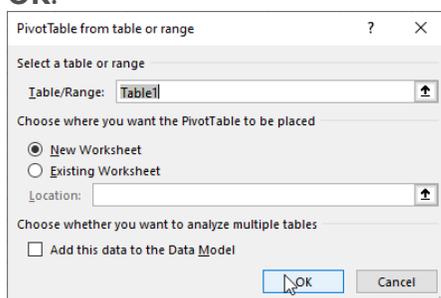
Usaremos el workbook trabajado anteriormente sobre el tema de quiebras. Este mismo archivo será usado para añadir una línea de tiempo (timeline) y luego interactuar con los botones de campos (field buttons). Los slicers, al igual que los timelines sirven para hacer sub-selección de datos de manera gráfica sin tener que escribir fórmulas. Son muy utilizados en tableros de datos (dashboards).

- Para comenzar, **localice** y **abra** el workbook **usbc_bankruptcyfilings2021.xls** localizado en el folder **C:\Tutorial_Excel\DataAnalysis\Modulo_2**

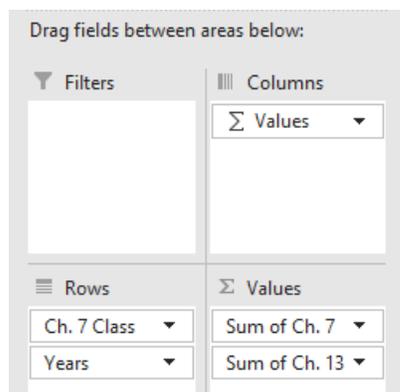
- Haga **click** en **cualquiera de las celdas de la tabla**. Vaya al **menú principal** y escoja **Insert**, luego **click** en el botón **PivotTable**.



- En la forma **Pivot Table from table or range**, acepte las opciones y haga **click** en el botón **OK**.



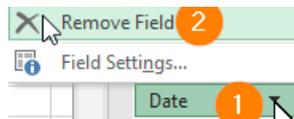
- Aparecerá la tabla pivot vacía. Será necesario añadir los campos en la tabla pivot de la siguiente manera.



- **Desplace** el campo **Date** a la sección **Rows**.
- **Desplace** los campos **Ch. 7** y **Ch. 13** en la sección **Values**.

Aparecerá el botón **Σ Values** en la sección **Columns**. Eso hará que aparezcan las columnas Ch. 7 y Ch. 13 con sus sumatorias por año.

- **Elimine** las opciones **Date** y **Quarters** dejando solamente **Years**



La tabla pivot debe verse de esta manera.

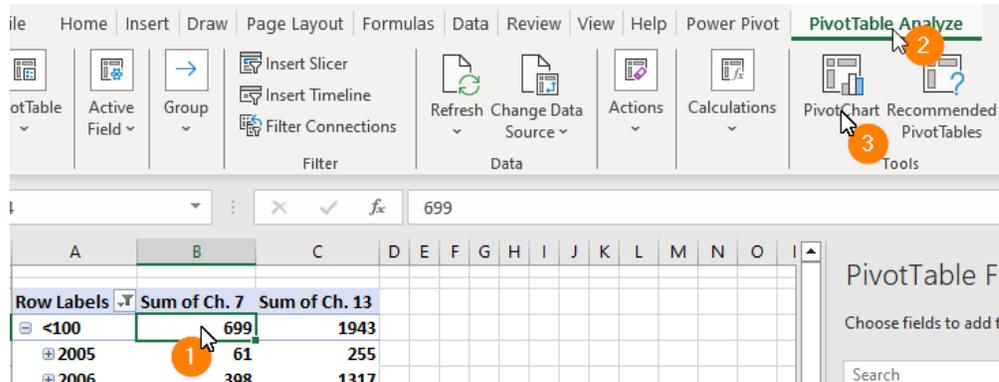
Row Labels	Sum of Ch. 7	Sum of Ch. 13
<100	240	371
2017	175	273
2020	65	98
100 - 200	2442	3318
2017	181	211
2020	1044	1407
2021	1217	1700
200 - 300	12318	19401
300 - 400	500	1071

Para cada clase del campo Ch. 7 Class, aparecerán todos los años que tengan uno o más casos en determinados años. Por ejemplo: dentro de la clase <100 quiebras (quiebras por Capítulo 7) hay 175 quiebras en 2017 y 65 quiebras en 2020.

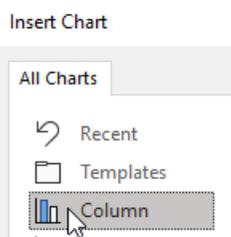
AÑADIR LA GRÁFICA PIVOT

Añadamos la gráfica pivot para visualizar los datos con la tabla pivot.

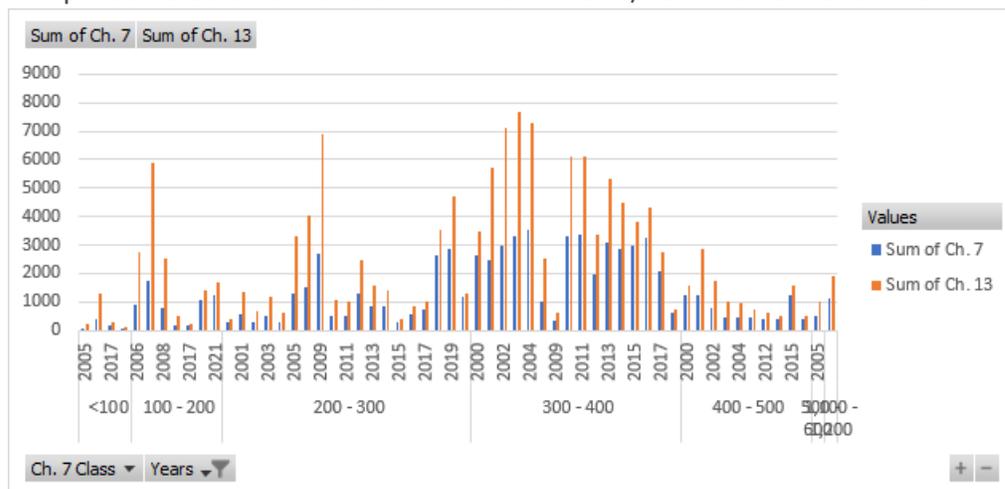
- Haga **click** en una de las celdas de la tabla pivot. Vaya al menú principal y escoja **Pivot Table Analyze**, luego **click** en el botón **PivotChart** que está en la sección **Tools**.



- En la forma **Insert Chart**, escoja la opción **Column** y haga **click** en el botón **OK** para aceptar las opciones y cerrar la forma **Insert Chart**.



La gráfica aparecerá de esta manera. Procure estirla, haciéndola más ancha.



Verá que para cada clase, se pueden repetir los años. También las clases de 500 a 600 en adelante aparecerán apiñadas por falta de espacio. Con el uso de slicers y timelines podrá ver los datos de manera dinámica y en menos espacio.

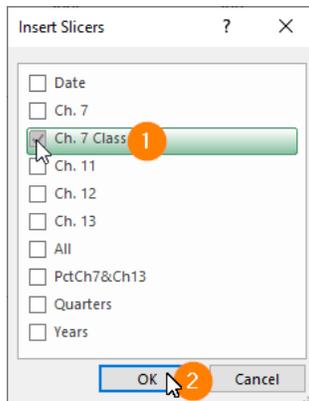
AÑADIR SLICERS

- Para añadir slicers deberá **activar la gráfica pivot o la tabla pivot**.

- Haga **click** en la **gráfica pivot** para activarla. Vaya al **menú principal** y haga **click** en el tab **PivotChart Analyze**, luego **click** en el botón **Insert Slicer**.

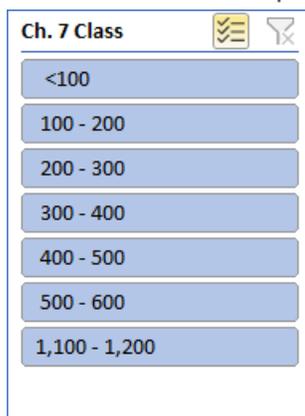


- En la forma **Insert Slicers** que aparecerá,



- Haga **check** en el campo **Ch. 7 Class** y
- Haga **click** en el botón **OK** para cerrar la forma.

- El **slicer Ch. 7 Class** aparecerá en algún lugar de la hoja.



Por defecto, todas las clases están seleccionadas. Verá también que **no aparecen las clases 600 a 700, 700 a 800, etc.**, hasta la clase 1,000 a 1,100. Esto ocurre **porque esas clases no tienen casos**.

- Experimente con el slicer. **1:** Remueva filtros anteriores si fuere necesario y **2:** Haga **click** en el botón **Multi-Select** **3:** Haga **click** en el botón **<100** del Slicer.

Slicer:

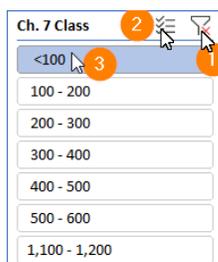
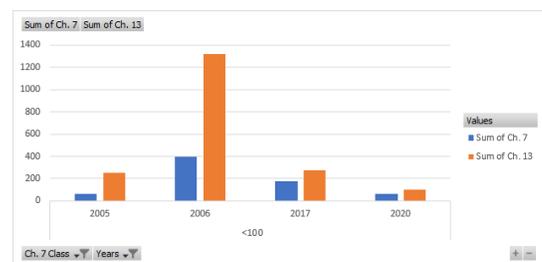


Tabla Pivot:

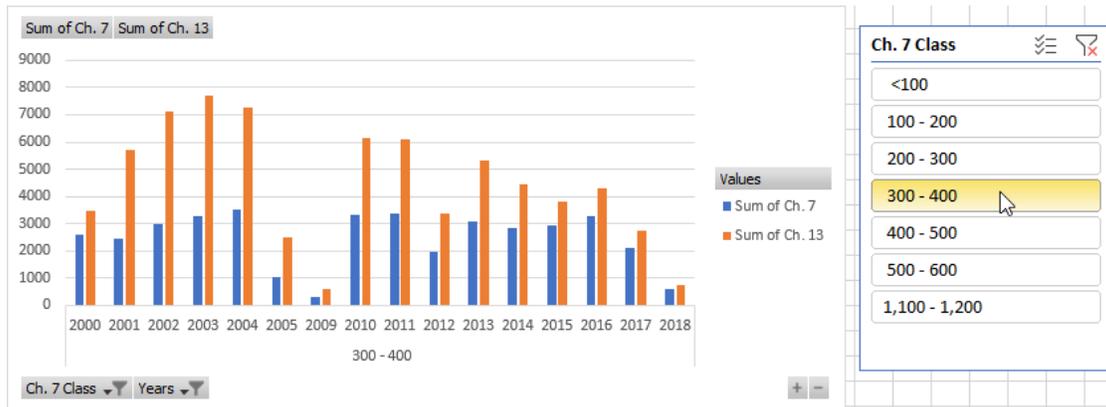
Row Labels	Sum of Ch. 7	Sum of Ch. 13
<100	699	1943
2005	61	255
2006	398	1317
2017	175	273
2020	65	98
Grand Total	699	1943

Gráfica Pivot:



Para los años 2005, 2006, 2017 y 2020 hay casos dentro de la clase <100 quiebras. Esto no quiere decir que para esos años no hubo más casos. Solo que dentro de esa clase <100 hay casos para esos años. En otros años no había casos dentro de esa clase <100.

- Experimente con otras clases como 300 a 400 para que vea las diferencias en años y conteos dentro de la clase

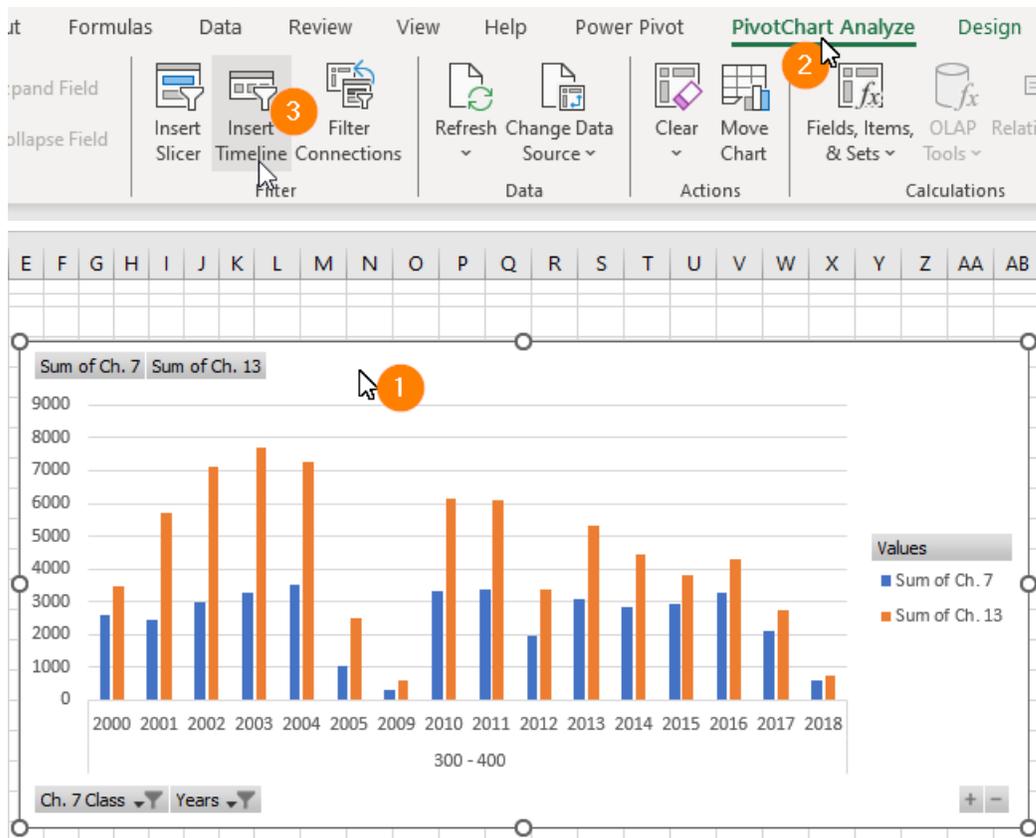


Recuerde que por lo general, las quiebras por capítulo 13 son de corporaciones. El capítulo 7 es de individuos.

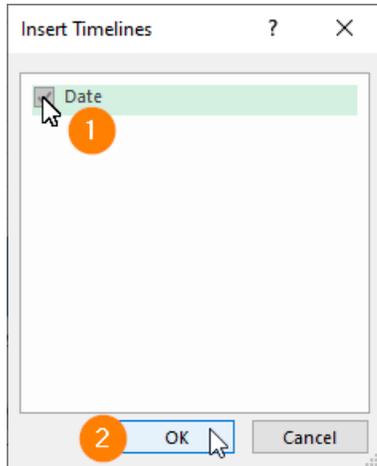
AÑADIR UN TIMELINE

Añadiremos una línea de tiempo asociada a los años registrados en la tabla de informe de quiebras.

- Para añadir un timeline (línea de tiempo):
Haga **click en la gráfica pivot** para activarla. Luego vaya al **menú principal** y escoja el tab **PivotChart Analyze**, luego haga **click** en el botón **Insert Timeline**.



- En la forma **Insert Timelines:**



- Haga **check** en el campo **Date**, el cual contiene las fechas de sometimiento de quiebras en el tribunal federal.
- Luego haga **click** en el botón **OK** para aceptar los cambios y cerrar la forma.

Seguramente tendrá que estirar el timeline para que pueda ver todos los años. (2000 a 2021)



Vamos a usar el timeline para escoger solo los años de 2011 a 2021. Recuerde que el slicer todavía está activo y filtrando datos por clase en el campo capítulo 7.

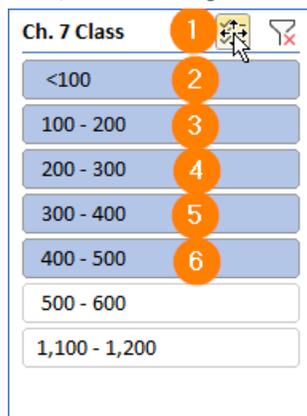
- Para cambiar de años, puede hacer **click y arrastrar** los años desde 2000 hasta el inicio de 2011.



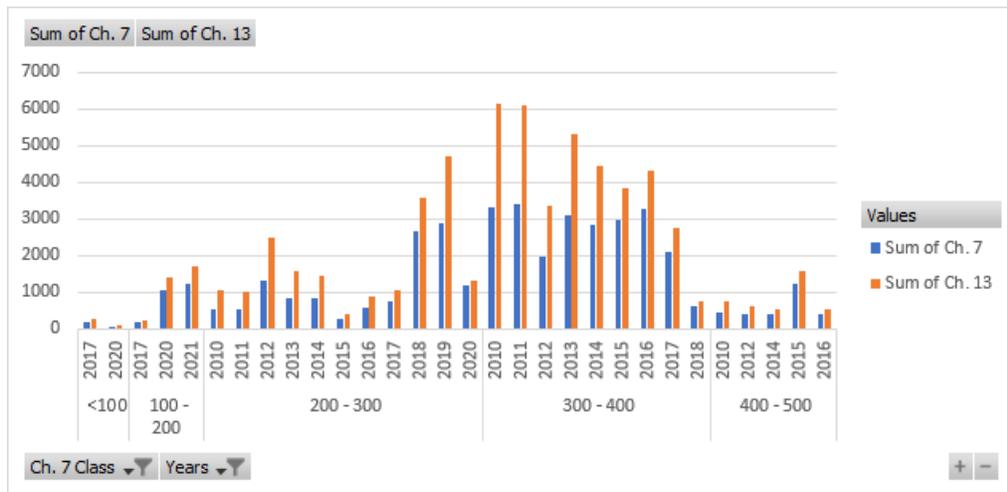
- El timeline debe verse así:



- Vaya al **slicer** y con la opción **Multi-Select activada**, escoja las clases **<100, 100 – 200, 200 – 300, 300 – 400 y 400 – 500**.

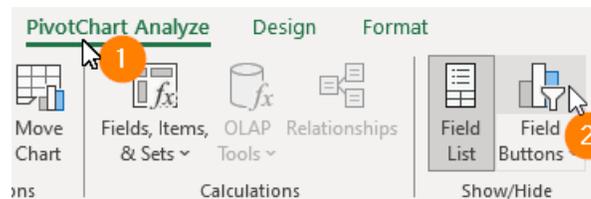


La gráfica con los años seleccionados de 2010 a 2021 y con las clases <100 hasta 400 – 500 casos, debe verse así:

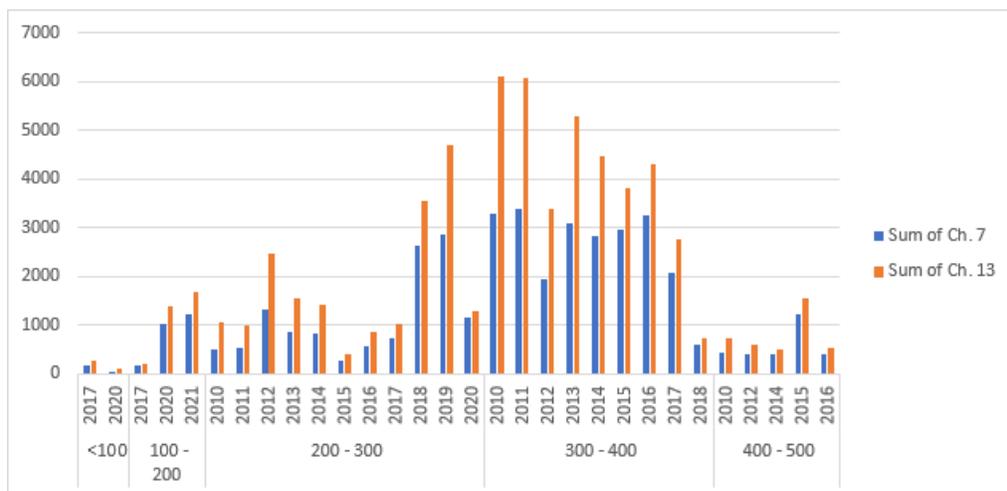


Observe que para cada clase hay diferentes años. En la clase 200-300 hay casos para los años 2010 a 2020 pero no para 2021. Sin embargo, en la clase 100 – 200, solo hay casos para los años 2017, 2020 y para 2021.

- Si desea quitar los field buttons de la gráfica pivot, deberá ir al **menú principal, escoger** el tab **PivotChart Analyze** y en la sección **Show/Hide**, haga **click** en el botón **Field Buttons**



La gráfica deberá aparecer sin los botones.



Con este ejercicio, finalizamos este tutorial. Guarde el archivo si lo desea.

FUENTES

FUNCIONES ÚTILES, EXCEL

AGGREGATE

<https://support.office.com/en-us/article/AGGREGATE-function-43b9278e-6aa7-4f17-92b6e19993fa26df>, recuperado en noviembre 17, 2023.

SUMIFS

<https://exceljet.net/excel-functions/excel-sumifs-function>, recuperado en noviembre 17, 2023.

Keyboard shortcuts <https://www.analyticsvidhya.com/blog/2016/12/cheatsheet-excel-functions-keyboardshortcuts/>, recuperado en noviembre 17, 2023.

10 Excel Functions You Need to Know for Data Analysis, <https://blog.adeptmarketing.com/10excel-functions-you-need-to-know-for-data-analysis>, recuperado en noviembre 17, 2023.

LOOKUP

<https://support.office.com/en-us/article/LOOKUP-function-446d94af-663b-451d-8251369d5e3864cb>, recuperado en noviembre 17, 2023.

VLOOKUP

Excel Vlookup Tutorial Part 3 - Vlookup Example with Exact Match, <http://www.excelfunctions.net/Vlookup-Example-Exact-Match.html>, recuperado en noviembre 17, 2023.

XLOOKUP

XLOOKUP Function, Microsoft Support, <https://support.microsoft.com/en-us/office/xlookup-function-b7fd680e-6d10-43e6-84f9-88eae8bf5929>, recuperado en noviembre 17, 2023.

Referencia a celdas

Relative and Absolute Cell References, <https://www.gcflernfree.org/excel2016/relative-andabsolute-cell-references/1/>, recuperado en noviembre 17, 2023.

Trasponer celdas:

How to transpose in Excel - convert row to column and viceversa <https://www.ablebits.com/office-addins-blog/2014/03/06/transpose-excel-rows-columns/> recuperado en noviembre 17, 2023.

UTILIDADES, PROGRAMAS ADD-IN

The XY Chart Labeler, <http://www.appspro.com/Utilities/ChartLabeler.htm> se usa para tener más opciones para añadir etiquetas en los diagramas de dispersión en Excel. Recuperado en noviembre 17, 2023.

DATOS TEMPORALES

Date and time functions (reference). Documentación Microsoft.

<https://support.office.com/en-us/article/date-and-time-functions-reference-fd1b5961-c1ae-4677-be58-074152f97b81>, recuperado en noviembre 17, 2023.

Using DATE function in Excel - formula examples to calculate dates,

<https://www.ablebits.com/office-addins-blog/2015/06/03/excel-date-formulas/>, recuperado en noviembre 17, 2023.

Año 1900 aparece como bisiesto en Lotus-1-2-3 y por consiguiente en Excel: *Excel incorrectly assumes that the year 1900 is a leap year*, Documentación Microsoft:

<https://learn.microsoft.com/en-US/office/troubleshoot/excel/wrongly-assumes-1900-is-leap-year>, recuperado en noviembre 17, 2023

Why are Excel weekdays wrong for 1900?, <https://superuser.com/questions/481496/why-are-excel-weekdays-wrong-for-1900>, recuperado en noviembre 17, 2023

Method to determine whether a year is a leap year, Documentación Microsoft, <https://support.microsoft.com/en-us/help/214019/method-to-determine-whether-a-year-is-a-leap-year>, recuperado en noviembre 17, 2023.

DATEDIF FUNCTION

DATEDIF function. Documentación Microsoft, <https://support.office.com/en-us/article/datediff-function-25dba1a4-2812-480b-84dd-8b32a451b35c>, recuperado en noviembre 17, 2023.

Secret Excel DATEDIF Function, <https://www.myonlinetraininghub.com/secret-excel-functiondatedif>, recuperado en noviembre 17, 2023. Provee ejemplos más detallados de los errores en esta función y alternativas a esta función.

Cómo representar fechas en inglés a español: Specifying a Language for the TEXT Function https://excelribbon.tips.net/T011782_Specifying_a_Language_for_the_TEXT_Function.html recuperado en noviembre 17, 2023.

Muestra cómo desplegar una fecha en otro idioma cuando Excel está en otro lenguaje. El código que deberá insertar en la función TEXT para español es **0C0A**. Ejemplo:

`=TEXT(DATEVALUE("11/09/1999"),["$-0C0A]dddd, d ""de"" mmmm, yyyy")`

VISUALIZACIÓN, GRÁFICAS

RECOMENDACIONES:

Data Visualization 101: How to Choose the Right Chart or Graph for Your Data, <https://blog.hubspot.com/marketing/data-visualization-choosing-chart>, recuperado en noviembre 17, 2023.

Data Visualization – How to Pick the Right Chart Type? <https://eazybi.com/blog/data-visualization-and-chart-types/>, recuperado en noviembre 17, 2023

Choosing a good chart, http://extremepresentation.typepad.com/blog/2006/09/choosing_a_good.html, recuperado en noviembre 17, 2023.

How to Use the Gestalt Principles for Visual Storytelling #PoDV, <https://www.fusioncharts.com/blog/how-to-use-the-gestalt-principles-for-visual-storytellingpodv/>, recuperado en noviembre 17, 2023.

PIVOT TABLES

Pivot Table Index Calculations

<https://www.youtube.com/watch?v=NhN87Xzsqg>, recuperado en noviembre 17, 2023.

OTROS TEMAS:

QUIEBRAS

[Guía básica sobre quiebras](#), recuperado en octubre 12, 2023.

[Se disparan las quiebras en Puerto Rico](#), recuperado en octubre 12, 2023.