

Material Imprimible

Curso Data Analytics

### **Módulo 3: Visualización con PowerBI**

#### **Contenidos:**

- Introducción a Powerbi y carga de datos
- Transformación de datos
- Introducción a Dax: Medidas y columnas calculadas
- Visualizaciones
- Formato de gráficos y Construcción de un tablero

## 1. Introducción a PowerBI

PowerBI es el software de análisis de datos de Microsoft. Permite realizar conexiones a cientos de fuentes de datos y transformar los mismos en tableros interactivos que permiten optimizar la toma de decisiones.

Tiene diversas versiones: la de escritorio (o desktop) que es gratuita y permite utilizar todas las funcionalidades de análisis sin límite alguno, PowerBI Service, que permite analizar los datos en un entorno web o PowerBI Pro, que está destinado a usuarios profesionales que comparten sus análisis en un ámbito organizacional.

Durante este módulo de nuestro curso de Data Analytics, usaremos la versión de Escritorio, que puede instalarse desde el siguiente Link:

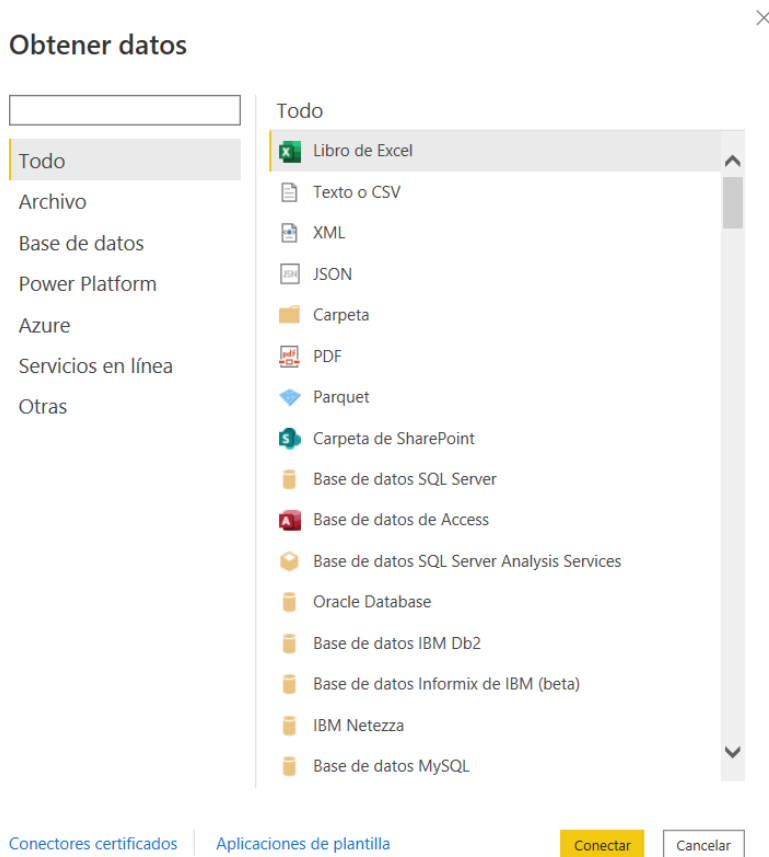
<https://www.microsoft.com/es-ar/download/details.aspx?id=58494>

### **Carga de datos: Conectores**

Esta herramienta tiene la capacidad de conectarse a cientos de orígenes de datos, destacándose que permite hacerlo con otros proveedores tecnológicos y no restringido solamente a Microsoft.

- Orígenes de datos locales: Excel, Archivos de Texto, Archivos separados por Coma (CSV), Access, Pdf, etc.
- Orígenes Cloud: SQL Azure, Google Big Query, Amazon Redshift, etc.
- Orígenes on premise: SQL Server, Oracle DB, IBM, Teradata, etc.

El proceso de carga incluye un análisis automático que ofrece la herramienta, que permite detectar encabezados de las columnas y el tipo de dato.



## 2. Transformación de datos

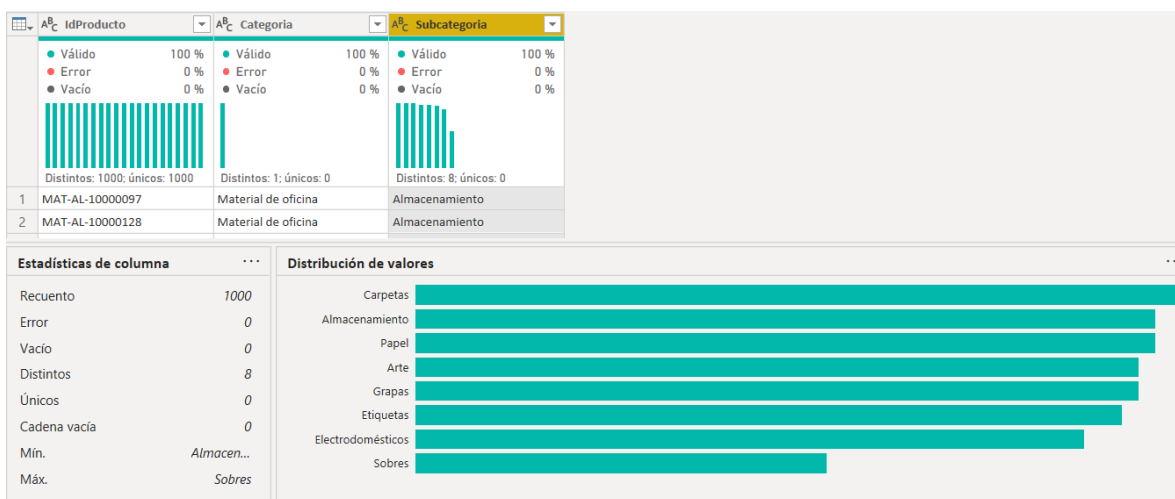
Una vez que se inicia el proceso de carga de datos, en varios momentos se nos permite realizar una transformación sobre los mismos.

Ingresando al módulo de Transformación que ofrece PowerBI, también denominado “Editor de PowerQuery”, vamos a poder realizar acciones de validación, limpieza, optimización y construcción de datos.

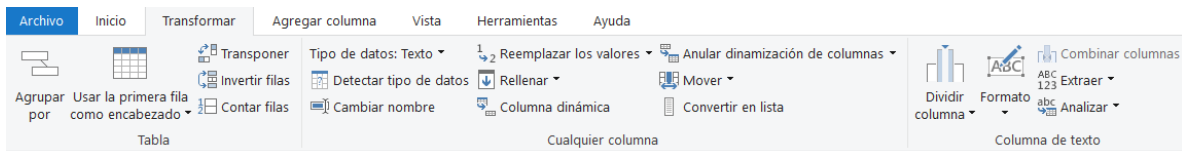
Es una herramienta potente con funcionalidades de ETL (Entrada, Transformación y Carga) que permite realizar todas las modificaciones previamente a utilizar nuestros datos en la capa de análisis.

Sus principales funcionalidades son las siguientes:

- Validación de datos: Ofrece la posibilidad de realizar muestreo y análisis de distribución de todas las columnas, permitiendo detectar valores nulos o erróneos fácilmente. De esta forma, el analista de datos ya sabe con qué desafíos va a encontrarse una vez que inicie su análisis, o qué errores debe subsanar previamente.



- Limpieza de datos: Una vez que hayamos detectado errores o fallas en nuestros datos, es fácilmente corregible mediante diversas funcionalidades de Reemplazo de datos (con o sin errores), filtrado de columnas y filas, extracción y concatenación de datos, entre otros.
- Optimización de datos: Ofrece funcionalidades que transforman nuestros datos de acuerdo a necesidades que tengamos, tales como, pivot de filas en columnas, dinamizar filas, hacer agrupaciones, combinar tablas,anexar tablas, etc.



- Construcción de nuevos datos: Son herramientas de generación de tablas y/o nuevas columnas para complementar los datos existentes. Las funcionalidades de construcción van desde un campo condicional hasta el diseño de campos dinámicos a partir de ejemplos, confiando en la interpretación que hace la herramienta.



### 3. Introducción a DAX: Medidas y columnas calculadas

Cuando ya nuestros datos están transformados y cargados, comienza la etapa de análisis y visualización. Allí, vamos a generar información relevante explotando los datos desde múltiples gráficos y construyendo medidas dinámicas. En esta etapa aparece DAX, que es el conjunto de fórmulas que nos permiten construir métricas dinámicas e interactivas. Sí, son fórmulas similares a Excel en muchos casos, pero que en PowerBI potencian cualquier análisis, permitiendo automatizar múltiples vistas sobre un mismo conjunto de datos y nos ofrecen de manera simple la realización de cálculos avanzados o que en otras herramientas no serían posible de realizar.

Dichas funciones pueden ser de agregación de datos (Contar, Sumar, Promediar), de escalado de datos (construcción de nuevas columnas), de inteligencia de tiempo (variación entre períodos), matemáticas, de filtrado, lógicas, ¡y muchas más!

Antes de continuar, es importante sentar precedente de lo que es una columna y lo que es una métrica, para poder utilizar el máximo poder de DAX

Columna:

- Son los atributos que poseen las distintas tablas y se pueden visualizar en dicho nivel. En el panel de campos, se los ve identificados con un símbolo de tabla y pueden ser de distinto tipo: Tiempo (fecha / hora), Ubicación (Mapa), Texto, Numérico, etc.
- Dax nos permite generar columnas calculadas que necesitemos, mediante el desarrollo de diversas fórmulas. Podemos extraer datos, realizar combinaciones o cálculos, o fórmulas condicionales. Es importante entender que dicha fórmula tendría impacto sobre una tabla puntual e iría fila a fila aplicando la sintaxis.

Veamos un ejemplo:

A partir de la tabla de clientes, genero una nueva columna calculada que sea Rango de Edad según la Edad que tengo en el atributo correspondiente. Si la persona es mayor a 28 años la nueva columna tomará el valor de “Adulto”, de lo contrario el valor será “Joven”



The screenshot shows the Power BI interface with a calculated column formula and a data table. The formula bar at the top displays: `1 Rango de Edad = IF(CLIENTES[Edad] > 28, "Adulto", "Joven")`. Below the formula bar is a table with the following columns: **CodigoCliente**, **Segmento**, **Nombre**, **Edad**, and **Rango de Edad**. The **Rango de Edad** column is highlighted in yellow. The table contains the following data rows:

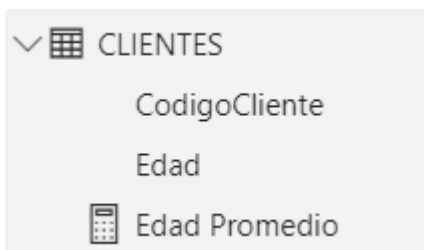
CodigoCliente	Segmento	Nombre	Edad	Rango de Edad
AA-10015	Cliente	Abel Ángel	47	Adulto
AA-10150	Cliente	Alonso Arévalo	44	Adulto
AA-10165	Cliente	Adán Atencio	20	Joven
AA-10390	Cliente	Arturo Anguiano	50	Adulto
AA-11665	Cliente	Alina Anaya	58	Adulto
AB-10645	Cliente	Alexia Bonilla	43	Adulto
AB-10885	Cliente	Ana Borrego	39	Adulto
AC-10105	Cliente	Abraham Cedillo	32	Adulto

- Medidas:

Las medidas resumen la información de un determinado atributo y permiten su uso desde diversos gráficos de manera dinámica. Si bien utilizan datos de las distintas tablas (una o muchas) las mismas no se almacenan directamente en ninguna, por lo cuál es una buena práctica armar una tabla vacía o carpeta y almacenarlas allí.

Cuando un analista crea una medida, está generando en memoria un cálculo que podrá ser utilizado en múltiples visualizaciones en paralelo, segmentado por diversas columnas y funcionará en todos los casos. No vamos a poder encontrar las medidas en ninguna tabla, porque no se calcula campo a campo, es un resultado almacenado en memoria listo para visualizarse.

Una forma rápida de detectar una medida en el panel de campos, es ver si tiene un símbolo de una calculadora en el margen izquierdo



Por ejemplo, una medida sería “La edad promedio de mis clientes”. Cada línea de la tabla es un cliente, y la medida resumiría todos los clientes otorgando ese valor y dejándolo disponible para ser analizado en gráficos dinámicos.





Hay una forma rápida de hacer medidas con PowerBI que es el Quick Measures, donde las funciones simples ya están precargadas en algunos campos y listas para ser utilizadas. Siempre se aconseja que, más allá de la funcionalidad *drag and drop*, el analista de datos interprete la fórmula, los datos y optimice su uso.

Fuente:

<https://dax.guide/>

<https://docs.microsoft.com/es-es/power-bi/transform-model/desktop-quickstart-learn-dax-basics>

<https://docs.microsoft.com/es-es/power-bi/transform-model/desktop-measures>

#### **4. Visualizaciones y Segmentación de datos**

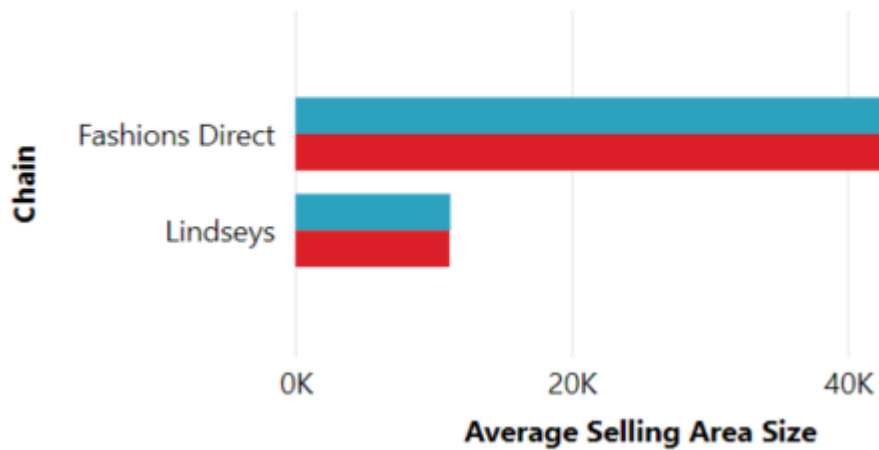
Es el conjunto de gráficos que ofrece PowerBi para explorar nuestros datos (columnas y medidas) y empezar a construir nuestros tableros dinámicos.

Tipos de gráficos

Gráficos de Barra

Average Selling Area Size by Store Type, Chain

Store Type ● New Store ● Same Store



Estos gráficos permiten introducir una medida en un eje y una columna que segmenta dicha medida en el otro. Se recomienda utilizar las barras cuando queremos analizar alguna volumetría (como las Ventas o las cantidades de ítems) por diversos criterios categóricos: Canales, Productos, Tiendas, Países, Segmentos, etc.

En el caso de que tengamos múltiples aperturas se aconseja usar las barras apiladas. También es común utilizar gráficos de barra combinados con líneas.

Gráficos de Distribución

This Year Sales by Chain



Permiten visualizar cómo se distribuye un todo en múltiples valores que lo segmentan. Los gráficos más utilizados en este apartado son el Pie Chart, el Donut Chart y el TreeMap. Es importante mencionar que este tipo de gráficos nos da una visión única y no puede ser utilizado para ver la evolución histórica, salvo que se lo combine con barras.

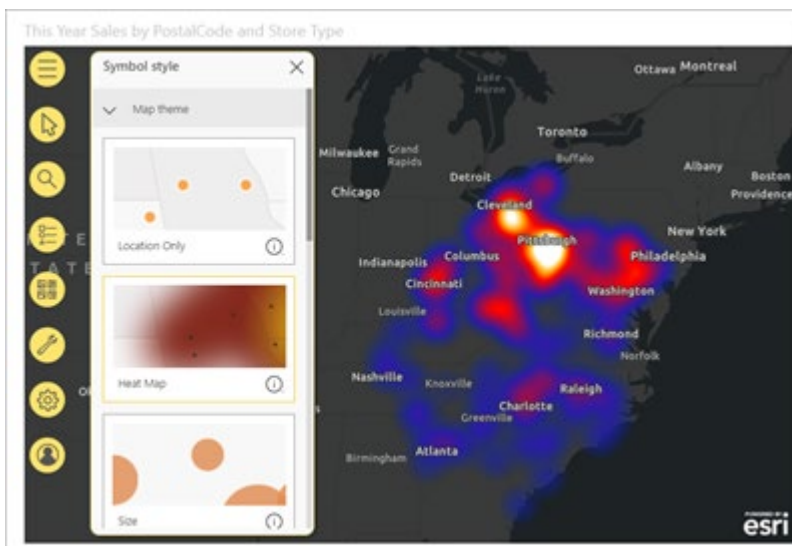
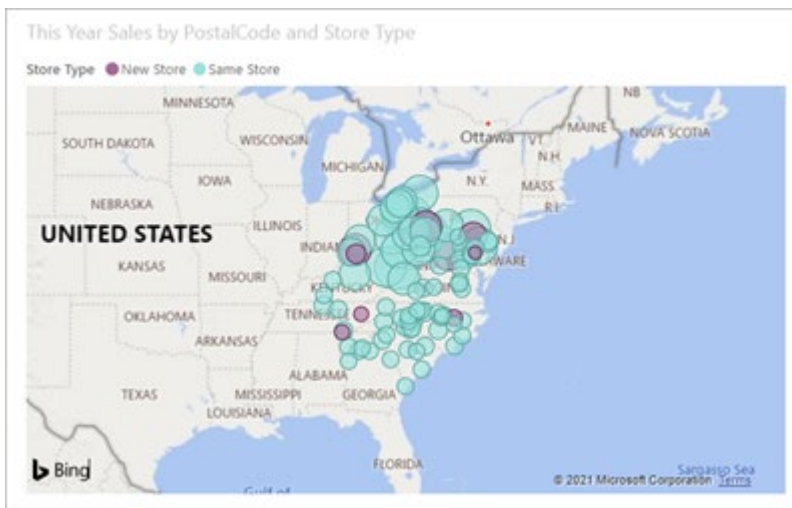
### Gráficos de Línea

This Year Sales and Last Year Sales by FiscalMonth



Es el tipo de gráfico ideal para ver la evolución en el tiempo. Se aconseja siempre utilizar en el eje X una variable temporal (Períodos, meses, años, etc) y un indicador numérico en el eje Y (variaciones, promedios, montos, etc). Pueden combinarse fácilmente con gráficos de barra.

### Gráficos de Mapa



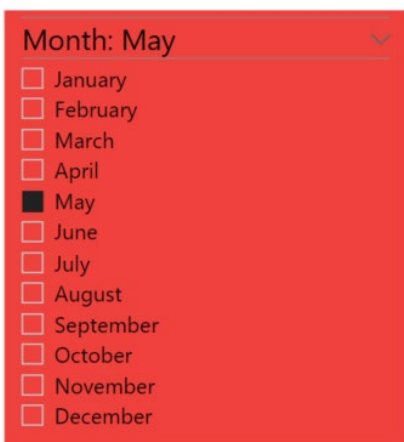
Requieren de datos de tipo geográfico (ubicaciones, direcciones, latitud y longitud, ciudades, países, etc) y permiten visualizar en distintos puntos del mapa los valores relativos y/o absolutos que queramos mostrar, integrándose con diversas capas de colores para profundizar el análisis.

#### Tarjetas y KPI

104  
Total Stores

Son universalmente utilizados en Analytics para exponer resultados rápidamente de manera elegante. Es importante que en estas tarjetas pongamos los datos más relevantes que el usuario debe ver primero, para luego profundizar con otros tipos de gráficos exploratorios.

#### Segmentadores



Son objetos visuales que permiten seleccionar valores de una lista o cuadro para poder filtrar los restantes gráficos del tablero. En Excel se los conoce como Segmentadores o *Slicers*.

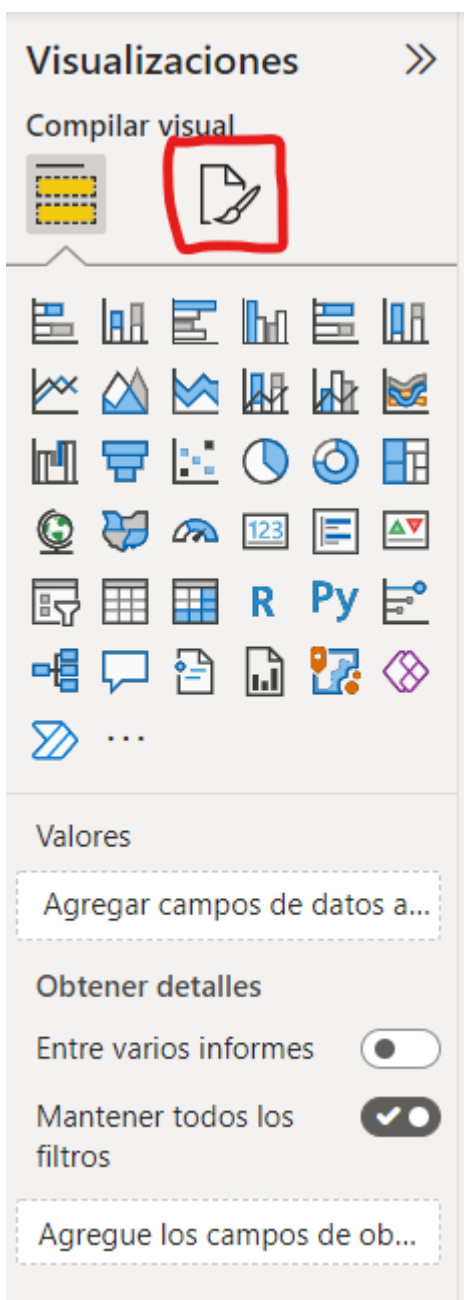
Fuente:

<https://docs.microsoft.com/es-es/power-bi/visuals/power-bi-visualization-types-for-reports-and-q-and-a>

## **5. Formato de gráficos y construcción de un tablero**

Todos los objetos de nuestro espacio de trabajo cuentan con una vista de Diseño, en el cuál podremos editar múltiples propiedades relacionadas con el formato, los colores, y la posición de los mismos.

Estas funcionalidades se ubican en el margen derecho con el siguiente ícono:



Allí, según el objeto que hayamos seleccionado podremos acceder a todas las funciones de diseño que tiene disponible.

Es importante mencionar que cada objeto tiene sus particularidades. Por ejemplo, si estamos dando formato a una página podremos elegir una imagen o color de fondo, el borde, o el tamaño, y si estamos dando formato a un gráfico de barra, podremos

establecer barras de colores según el valor de los datos, tamaño y profundidad de objetos, títulos, diseño de ejes, leyendas y mucho más.

### **Construcción de un tablero**

Un Tablero es un conjunto de gráficos en armonía que están ubicados en una misma hoja de trabajo. Típicamente, se recomiendan entre 4 a 6 gráficos en una misma página de PowerBI y las principales reglas de diseño que hay que considerar son las siguientes:

- Partir de un indicador general que introduzca el tema.
- No utilizar gráficos muy cargados de datos. Ser simples y cumplir el lema “*menos es más*”
- Explorar los datos con gráficos entendibles
- Colocar segmentadores en el margen superior o izquierdo, donde el usuario del tablero los pueda usar fácilmente
- Colocar imágenes, logos, títulos y colores amigables

Un ejemplo de una hoja de tablero es la siguiente:



