

POWER BI - BECOME AN ANALYST

Path to high-level programming thinking

Present By: Data Engineering Professional - Marcos De Vargas

Temas



Industria 4.0 y la
necesidad de los datos.



Principios básicos de
Power BI.



Power Query.



Structured Query
Language.



Visualizaciones de
datos.

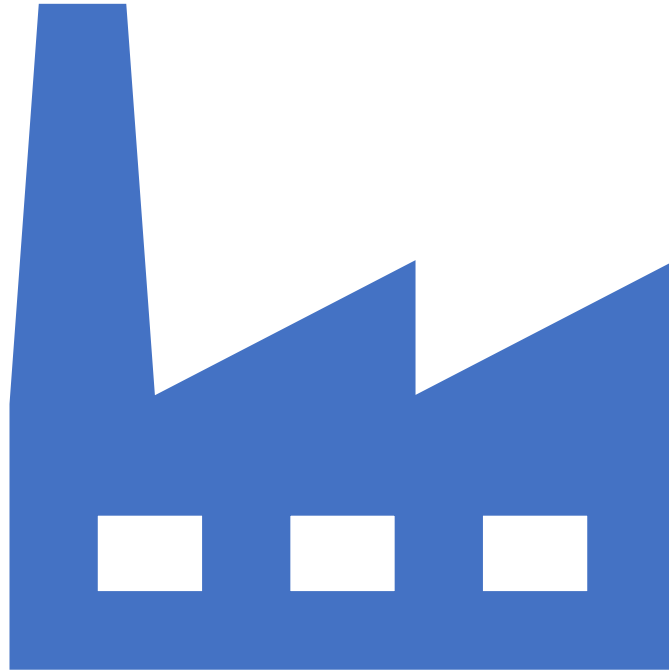


Power BI Service.



La practica hace al
maestro (CI).

Industria 4.0 y la necesidad de los datos





La Industria 4.0 es la realización de la transformación digital, que brinda toma de decisiones en tiempo real, productividad mejorada, flexibilidad y agilidad.



Está revolucionando la forma en que las empresas fabrican, mejoran y distribuyen sus productos/servicios.



Las tecnologías digitales conducen a una mayor automatización, mantenimiento predictivo, optimización automática de mejoras de procesos y, sobre todo, un nuevo nivel de eficiencia y capacidad de respuesta a los clientes que antes no era posible.



Los conceptos y tecnologías de la Industria 4.0 se pueden aplicar en todo tipo de empresas industriales, incluida la manufactura discreta y de procesos, el petróleo y el gas, la minería y otros segmentos industriales.

Revoluciones Industriales

Primera revolución industrial:

Comenzando a fines del siglo XVIII en Gran Bretaña, la primera revolución industrial ayudó a permitir la producción en masa mediante el uso de energía hidráulica y de vapor en lugar de energía puramente humana y animal. Los productos finales se construyeron con máquinas en lugar de producirlos minuciosamente a mano.

Segunda revolución industrial:

Un siglo después, la segunda revolución industrial introdujo las líneas de montaje y el uso de petróleo, gas y energía eléctrica. Estas nuevas fuentes de energía, junto con los avances en las comunicaciones por teléfono y telégrafo, trajeron la producción en masa y cierto grado de automatización a los procesos de manufactura.

Tercera revolución industrial:

La tercera revolución industrial, que comenzó a mediados del siglo XX, agregó computadoras, telecomunicaciones avanzadas y análisis de datos a los procesos de manufactura. La digitalización de las fábricas comenzó con la incorporación de controladores lógicos programables (PLC) en la maquinaria para ayudar a automatizar algunos procesos y recopilar y compartir datos.

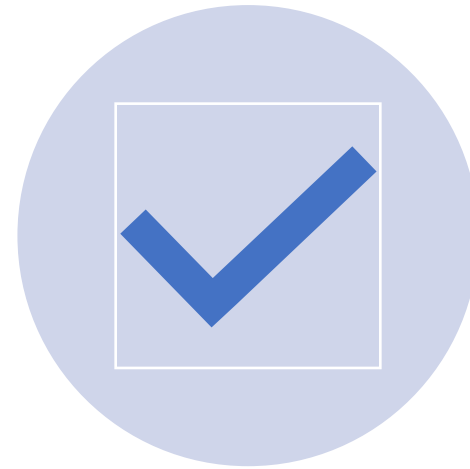
Cuarta revolución industrial:

Ahora estamos en la cuarta revolución industrial, también conocida como la Industria 4.0. Caracterizados por una creciente automatización y el empleo de máquinas y fábricas inteligentes, los datos informados ayudan a producir bienes de manera más eficiente y productiva en toda la cadena de valor. Se mejora la flexibilidad para que los fabricantes puedan satisfacer mejor las demandas de los clientes mediante la personalización masiva. En última instancia, buscan lograr la eficiencia, en muchos casos, con el tamaño de lote 1. Al recopilar más datos de la planta y combinarlos con otros datos operativos de la empresa, una fábrica inteligente puede lograr transparencia de la información y mejores decisiones.

Características de una fábrica inteligente

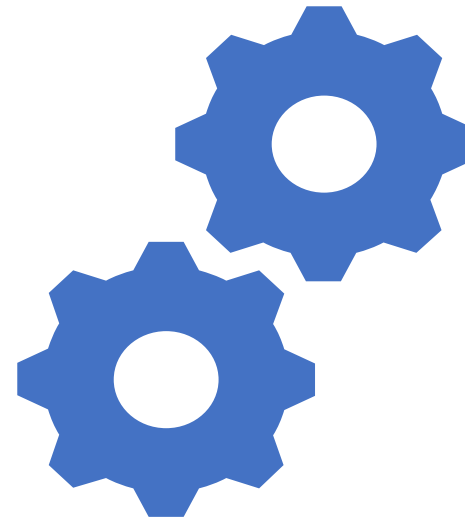


ANÁLISIS DE DATOS PARA UNA
ÓPTIMA TOMA DE DECISIONES



INTEGRACIÓN DE TI

Principios básicos de Power BI.

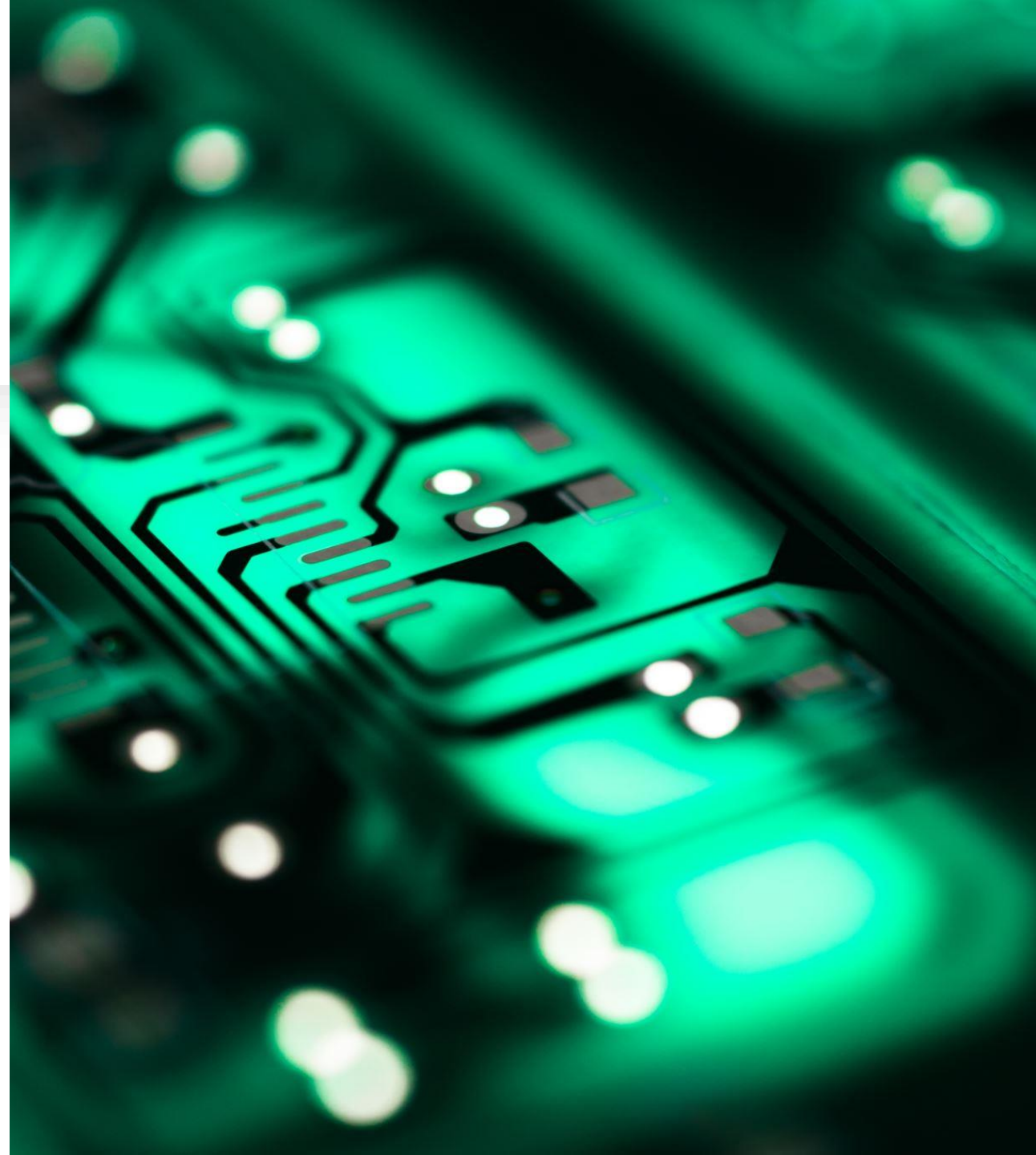


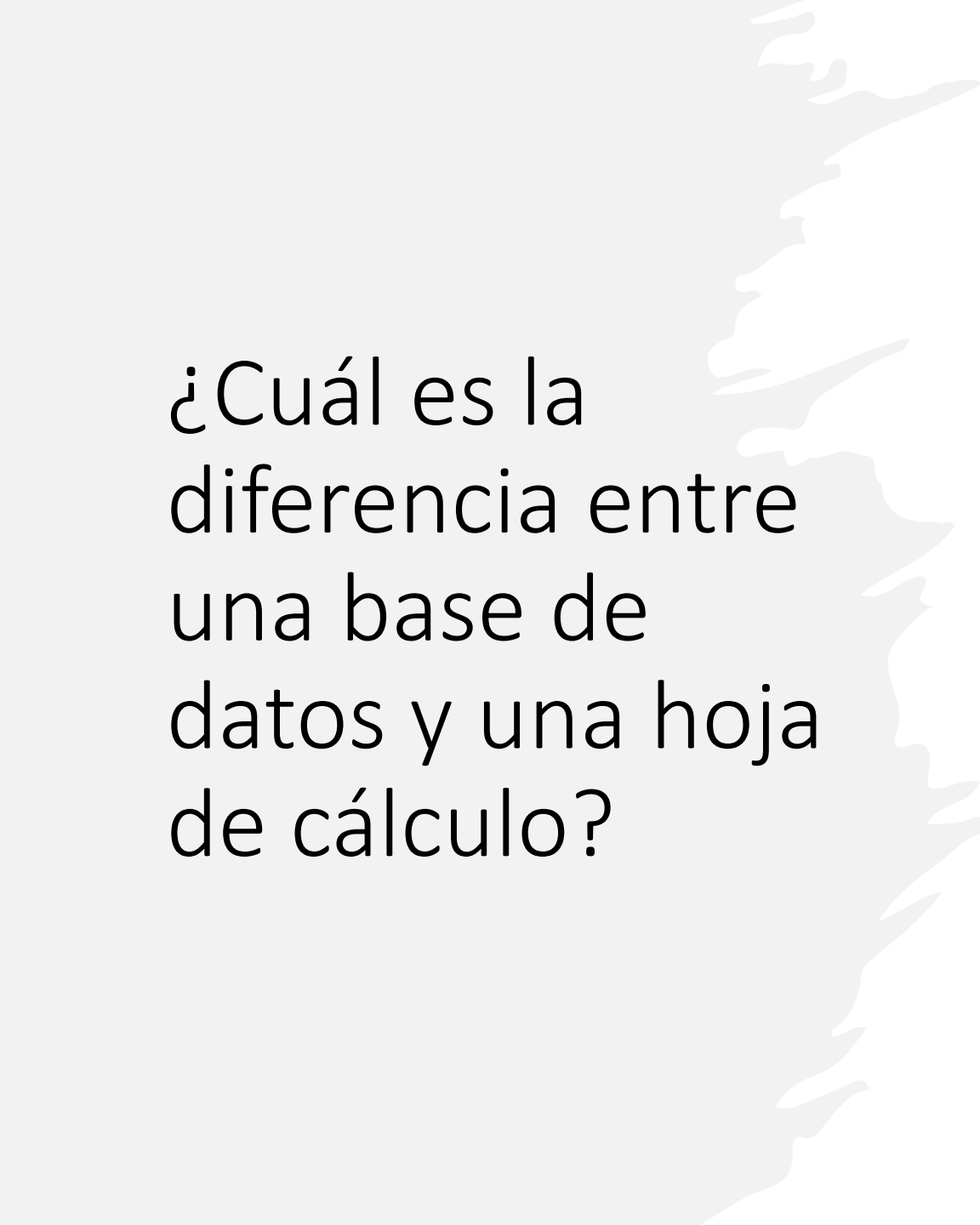
Mundo basado en datos

- Los conjuntos de habilidades están cambiando para adaptarse a un mundo basado en los datos. Para los profesionales es cada vez más valioso poder usar los datos para tomar decisiones y usar elementos visuales para contar historias con los datos para informar quién, qué, cuándo, dónde y cómo. Mientras que la educación tradicional normalmente dibuja una línea distinta entre la narración creativa y el análisis técnico, el mundo profesional moderno también valora a aquellos que pueden cruzar entre las dos: la visualización de datos se encuentra justo en el centro del análisis y la narración visual.

¿Qué es una base de datos?

- Una base de datos es una recopilación organizada de información o datos estructurados, que normalmente se almacena de forma electrónica en un sistema informático. Normalmente, una base de datos está controlada por un sistema de gestión de bases de datos (DBMS). En conjunto, los datos y el DBMS, junto con las aplicaciones asociadas a ellos, reciben el nombre de sistema de bases de datos, abreviado normalmente a simplemente base de datos.
- Los datos de los tipos más comunes de bases de datos en funcionamiento actualmente se suelen utilizar como estructuras de filas y columnas en una serie de tablas para aumentar la eficacia del procesamiento y la consulta de datos. Así, se puede acceder, gestionar, modificar, actualizar, controlar y organizar fácilmente los datos. La mayoría de las bases de datos utilizan un lenguaje de consulta estructurada (SQL) para escribir y consultar datos.





¿Cuál es la
diferencia entre
una base de
datos y una hoja
de cálculo?

- Cómo se almacenan y se manipulan los datos
- Quién puede acceder a los datos
- Cuántos datos pueden almacenarse

Tipos de bases de datos

Existen muchos tipos diferentes de bases de datos. La mejor base de datos para una organización específica depende de cómo pretenda la organización utilizar los datos.

Tipos de bases de datos

- **Bases de datos relacionales:** Las bases de datos relacionales se hicieron predominantes en la década de 1980. Los elementos de una base de datos relacional se organizan como un conjunto de tablas con columnas y filas. La tecnología de bases de datos relacionales proporciona la forma más eficiente y flexible de acceder a información estructurada.
- **Bases de datos orientadas a objetos:** La información de una base de datos orientada a objetos se representa en forma de objetos, como en la programación orientada a objetos.
- **Bases de datos distribuidas:** Una base de datos distribuida consta de dos o más archivos que se encuentran en sitios diferentes. La base de datos puede almacenarse en varios ordenadores, ubicarse en la misma ubicación física o repartirse en diferentes redes.
- **Almacenes de datos:** Un repositorio central de datos, un data warehouse es un tipo de base de datos diseñado específicamente para consultas y análisis rápidos.
- **Bases de datos NoSQL:** Una base de datos NoSQL, o base de datos no relacional, permite almacenar y manipular datos no estructurados y semiestructurados (a diferencia de una base de datos relacional, que define cómo se deben componer todos los datos insertados en la base de datos). Las bases de datos NoSQL se hicieron populares a medida que las aplicaciones web se volvían más comunes y complejas.

Tipos de bases de datos

- **Bases de datos orientadas a grafos:** Una base de datos orientada a grafos almacena datos relacionados con entidades y las relaciones entre entidades. Bases de datos OLTP. Una base de datos OLTP es una base de datos rápida y analítica diseñada para que muchos usuarios realicen un gran número de transacciones.
- **Bases de datos de código abierto:** Un sistema de base de datos de código abierto es aquel cuyo código fuente es de código abierto; tales bases de datos pueden ser bases de datos SQL o NoSQL.
- **Bases de datos en la nube:** Una base de datos en la nube es una recopilación de datos, estructurados o no estructurados, que reside en una plataforma de cloud computing privada, pública o híbrida. Existen dos tipos de modelos de bases de datos en la nube: el modelo tradicional y el de base de datos como servicios (database as a service, DBaaS). Con DBaaS, un proveedor de servicios realiza las tareas administrativas y el mantenimiento.
- **Base de datos multimodelo:** Las bases de datos multimodelo combinan distintos tipos de modelos de bases de datos en un único servidor integrado. Esto significa que pueden incorporar diferentes tipos de datos.
- **Bases de datos de documentos/JSON:** Diseñadas para almacenar, recuperar y gestionar información orientada a los documentos, las bases de datos de documentos son una forma moderna de almacenar los datos en formato JSON en lugar de hacerlo en filas y columnas.
- **Bases de datos de autogestión:** El tipo de base de datos más nuevo e innovador, las bases de datos de autogestión (también conocidas como bases de datos autónomas) están basadas en la nube y utilizan el machine learning para automatizar el ajuste de la base de datos, la seguridad, las copias de seguridad, las actualizaciones y otras tareas de gestión rutinarias que tradicionalmente realizan los administradores de bases de datos.

¿Qué es Power BI?

- Power BI es una colección de servicios de software, aplicaciones y conectores que funcionan conjuntamente para convertir orígenes de datos sin relación entre sí en información coherente, interactiva y atractiva visualmente. Sus datos podrían ser una hoja de cálculo de Excel o una colección de almacenes de datos híbridos locales y basados en la nube. Power BI permite conectarse con facilidad a los orígenes de datos, visualizar y descubrir qué es importante y compartirlo con cualquiera o con todos los usuarios que desee.

Las partes de Power BI



Aplicación de escritorio de Windows llamada Power BI Desktop.



Servicio de software como servicio (SaaS) en línea denominado servicio Power BI.



Aplicaciones para Power BI Mobile para dispositivos Windows, iOS y Android.

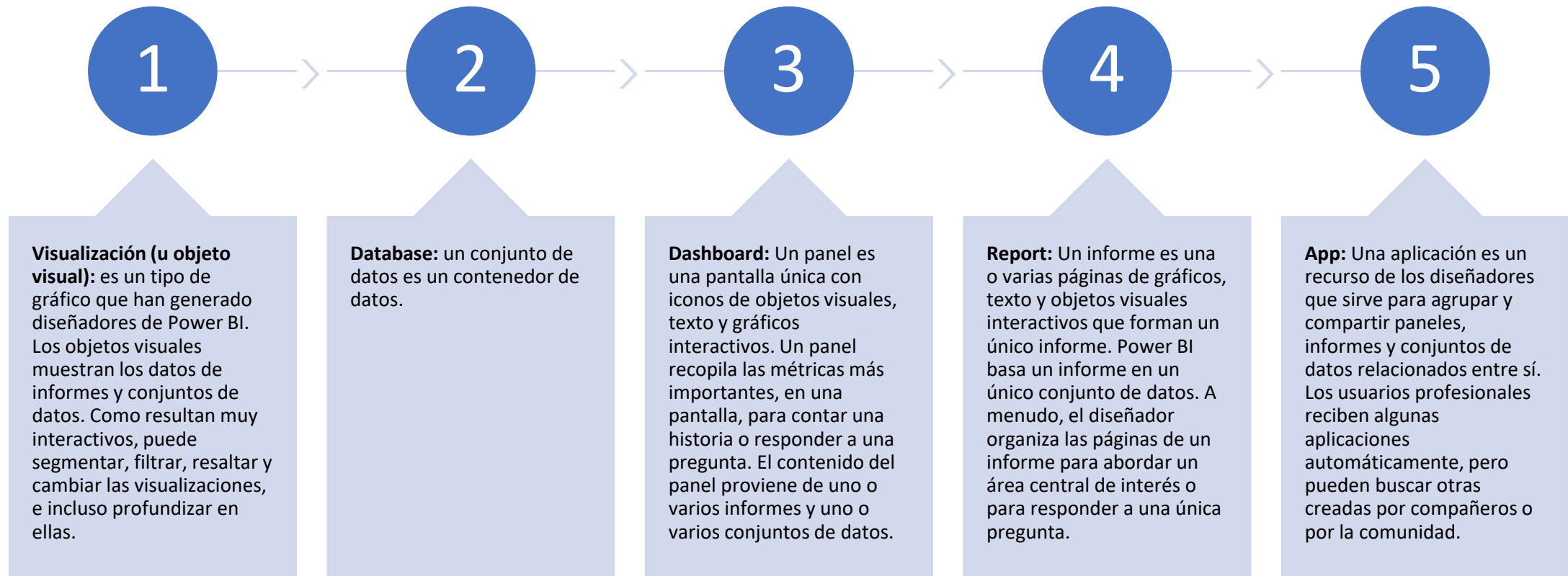


Power BI Report Builder, para crear informes paginados y compartirlos en el servicio Power BI.



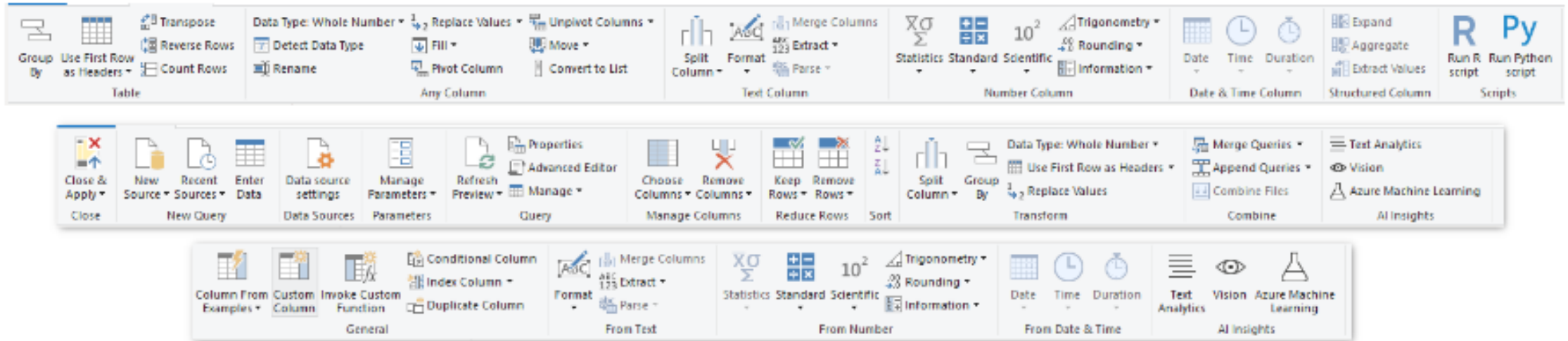
Power BI Report Server, un servidor de informes local en el que puede publicar los informes de Power BI, después de crearlos en Power BI Desktop.

Concepto básico



¿Qué es Power Query?

- Power Query es un motor de preparación de datos y transformación de datos. Este incluye una interfaz gráfica para obtener datos de orígenes y un editor para aplicar transformaciones. Dado que el motor está disponible en muchos productos y servicios, el destino en el que se almacenarán los datos depende de dónde se usó. Con Power Query, puede realizar el procesamiento de extracción, transformación y carga (ETL) de datos.



Funciones

- El motor de transformación de Power Query incluye muchas funciones de transformación precompiladas que se pueden usar a través de la interfaz gráfica del Editor de Power Query. Estas transformaciones pueden ser tan sencillas como quitar una columna o filtrar filas, o tan comunes como usar la primera fila como encabezado de tabla. También hay opciones de transformación avanzadas, como combinar, anexar, agrupar por, dinamizar y anular la dinamización.
- Todas estas transformaciones son posibles eligiendo la opción de transformación en el menú y aplicando las opciones necesarias para esa transformación. En la ilustración siguiente se muestran algunas de las transformaciones disponibles en Editor de Power Query.

Lenguaje M

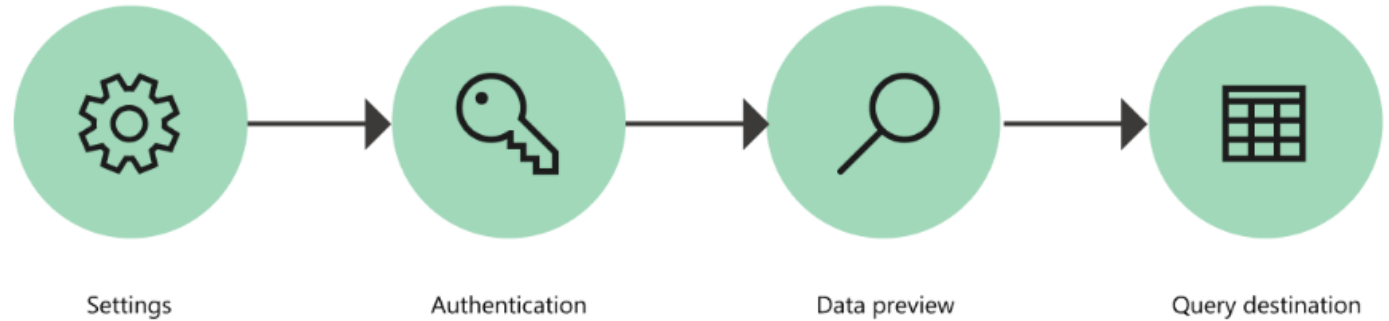
- En cualquier escenario de transformación de datos, hay algunas transformaciones que no se pueden realizar de la mejor manera mediante el editor gráfico. Algunas de estas transformaciones pueden requerir configuraciones y opciones especiales que la interfaz gráfica no admite actualmente. El motor de Power Query usa un lenguaje de scripting en segundo plano para todas las transformaciones de Power Query: el lenguaje de fórmulas M de Power Query, también conocido como M.
- El lenguaje M es el lenguaje de transformación de datos de Power Query. Todo lo que sucede en la consulta se escribe en última instancia en M. Si desea realizar transformaciones avanzadas mediante el motor de Power Query, puede usar el Editor avanzado para acceder al script de la consulta y modificarlo como desee. Si observa que las funciones y transformaciones de la interfaz de usuario no realizarán los cambios exactos que necesita, use el Editor avanzado y el lenguaje M para ajustar las funciones y transformaciones.



```
Power Query M Copiar

let
    Source = Exchange.Contents("xyz@contoso.com"),
    Mail1 = Source{[Name="Mail"]}[Data],
    #"Expanded Sender" = Table.ExpandRecordColumn(Mail1, "Sender", {"Name"}, {"Name"}),
    #"Filtered Rows" = Table.SelectRows(#"Expanded Sender", each ([HasAttachments] = true)),
    #"Filtered Rows1" = Table.SelectRows(#"Filtered Rows", each ([Subject] = "sample file")),
    #"Removed Other Columns" = Table.SelectColumns(#"Filtered Rows1",{"Attachments"}),
    #"Expanded Attachments" = Table.ExpandTableColumn(#"Removed Other Columns", "Attachments", {"Name", "Size", "IsHidden"}),
    #"Filtered Hidden Files" = Table.SelectRows(#"Expanded Attachments", each [IsHidden] = false),
    #"Invoke Custom Function1" = Table.AddColumn(#"Filtered Hidden Files", "Transformed", each Text.Combine([Name], [Size])),
    #"Removed Other Columns1" = Table.SelectColumns(#"Invoke Custom Function1", {"Transformed"}),
    #"Expanded Table Column1" = Table.ExpandTableColumn(#"Removed Other Columns1", "Transformed", {"Name", "Size"}),
    #"Changed Type" = Table.TransformColumnTypes(#"Expanded Table Column1",{{"Column1", type text}, {"Column2", type text}},
in
    #"Changed Type"
```

¿Como
obtener los
datos usando
Power Query?



Disponibilidad de Power Query

Producto	Motor M ¹	Power Query Escritorio ²	Power Query Online ³	Flujos de datos ⁴
Excel para Windows	Sí	Sí	No	No
Excel para Mac	Sí	Sí	No	No
Power BI	Sí	Sí	Sí	Sí
PowerApps	Sí	No	Sí	Sí
Power Automate	Sí	No	Sí	No
Power BI Report Server	Sí	Sí	No	No
Azure Data Factory	Sí	No	Sí	Sí
Data Factory en Microsoft Fabric	Sí	No	Sí	Sí
SQL Server Integration Services	Sí	No	No	No
SQL Server Analysis Services	Sí	Sí	No	No
Dynamics 365 Customer Insights	Sí	No	Sí	Sí

¿Qué es SQL / Structured Query Language?



SQL es un lenguaje de computación para trabajar con conjuntos de datos y las relaciones entre ellos. Los programas de bases de datos relacionales, como Microsoft Office Access, usan SQL para trabajar con datos. A diferencia de muchos lenguajes de computación, SQL no es difícil de leer y entender, incluso para un usuario inexperto. Al igual que muchos lenguajes de computación, SQL es un estándar internacional reconocido por organismos de estándares como ISO y ANSI.



SQL se usa para describir conjuntos de datos que pueden ayudarle a responder preguntas. Al usar SQL, debe usar la sintaxis correcta. La sintaxis es el conjunto de reglas mediante las que se combinan correctamente los elementos de un idioma. La sintaxis SQL se basa en la sintaxis del idioma inglés y usa muchos de los mismos elementos que la sintaxis de Visual Basic para Aplicaciones (VBA).



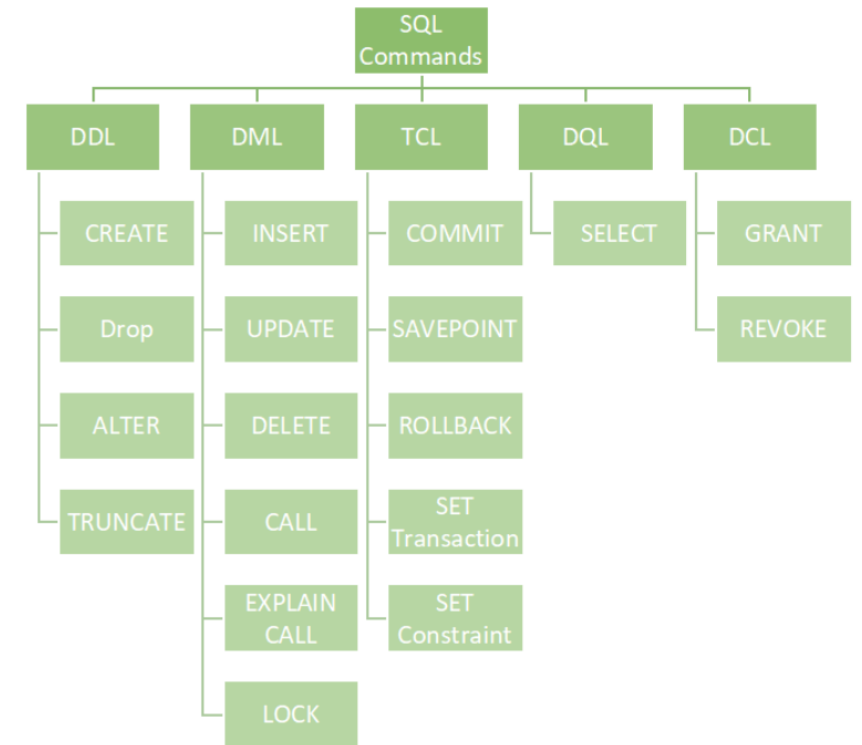
Por ejemplo, una simple instrucción SQL que recupera una lista de los apellidos de contactos cuyo nombre es María podría ser similar a esto:



```
SELECT Last_Name  
FROM Contacts  
WHERE First_Name = 'Mary';
```

SQL Commands

- DDL – Data Definition Language
- DQL – Data Query Language
- DML – Data Manipulation Language
- DCL – Data Control Language
- TCL – Transaction Control Language



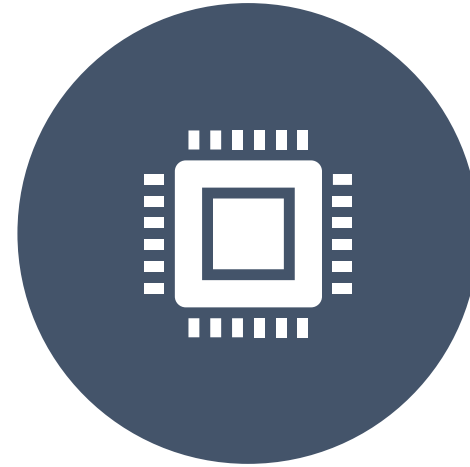
Clausulas y sus funciones

Cláusula SQL	Función	Obligatorio
SELECT	Muestra una lista de los campos que contienen datos de interés.	Sí
FROM	Muestra las tablas que contienen los campos de la cláusula SELECT.	Sí
WHERE	Especifica los criterios de campo que cada registro debe cumplir para poder ser incluido en los resultados.	No
ORDER BY	Especifica la forma de ordenar los resultados.	No
GROUP BY	En una instrucción SQL que contiene funciones de agregado, muestra los campos que no se resumen en la cláusula SELECT.	Solo si están estos campos
HAVING	En una instrucción SQL que contiene funciones de agregado, especifica las condiciones que se aplican a los campos que se resumen en la instrucción SELECT.	No

¿Qué es la visualización de datos?



LA VISUALIZACIÓN DE DATOS ES LA REPRESENTACIÓN GRÁFICA DE INFORMACIÓN Y DATOS. AL UTILIZAR ELEMENTOS VISUALES COMO CUADROS, GRÁFICOS Y MAPAS, LAS HERRAMIENTAS DE VISUALIZACIÓN DE DATOS PROPORCIONAN UNA MANERA ACCESIBLE DE VER Y COMPRENDER TENDENCIAS, VALORES ATÍPICOS Y PATRONES EN LOS DATOS.



EN EL MUNDO DEL BIG DATA, LAS HERRAMIENTAS Y TECNOLOGÍAS DE VISUALIZACIÓN DE DATOS SON ESENCIALES PARA ANALIZAR GRANDES CANTIDADES DE INFORMACIÓN Y TOMAR DECISIONES BASADAS EN LOS DATOS.

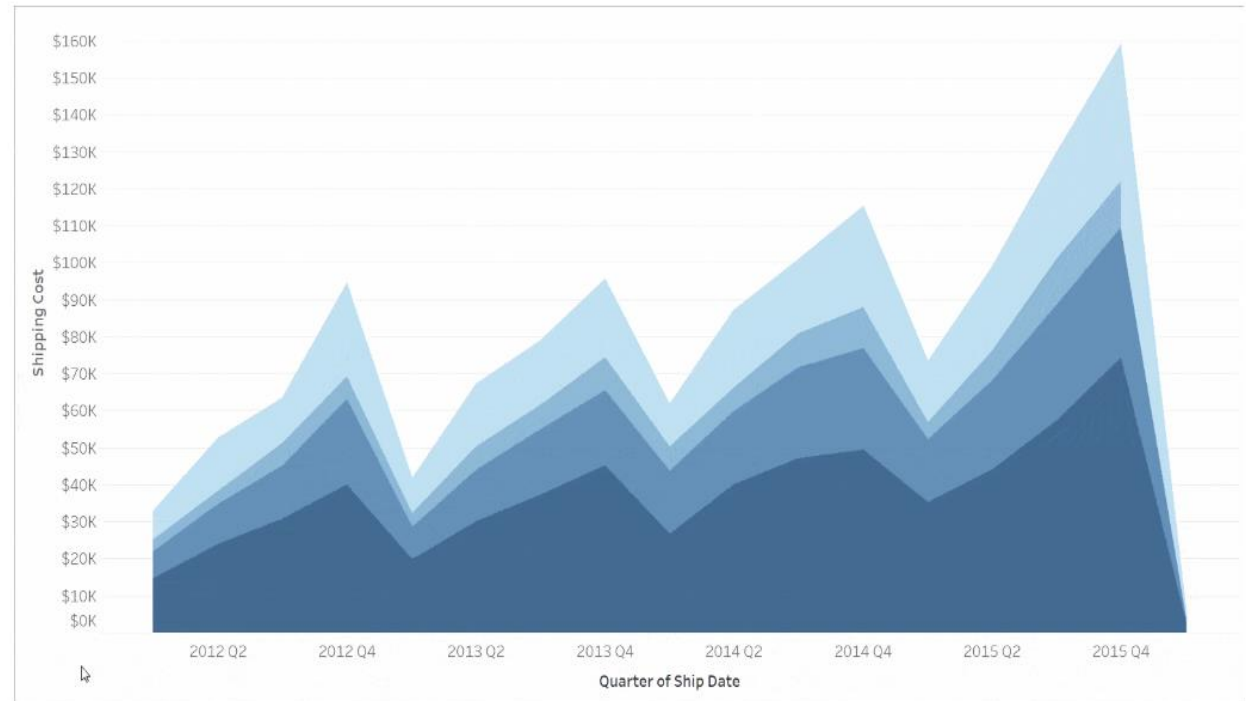
Tipos de visualización de datos

Mas Comunes:

- Cuadros
- Tablas
- Gráficos
- Mapas
- Infografía
- Dashboards

Especificos:

- Gráfico de área
- Gráfico de barras
- Diagramas de caja y bigotes
- Nube de burbujas
- Gráfico de bala
- Cartograma
- Vista circular
- Mapa de distribución de puntos
- Gráfico de Gantt
- Mapa de calor
- Tabla de resaltado
- Histograma
- Matriz



Trabajos en el campo de analisis de datos

Data Worker

Analista de Datos

Ingeniero de Datos

Cientifico de Datos

Data Worker - funciones



Uso continuo de Excel/hoja de calculo.



Exportar data de sistemas



Creación de reportes semanales o mensuales



Recibe la data de otros departamento como IT



Trabajar con funciones

Oportunidades de crecimiento

- Power Query
- Pivot Tables
- Power Point
- Microsoft Word

Analista de Datos - funciones



Oportunidades de crecimiento

- Estadística Básica
- Creación de funciones
- Entender Join

Ingeniero de Datos - funciones



Traduce gran cantidad
de datos en data sets



Conocimiento en
redefinir los data sets a
uno minimizado



Tiene mas acceso a la
data



Entendimiento de
seguridad y privacidad

Oportunidades de crecimiento

- Entendimiento de data estructurada y no estructurada
- Metodología de diseños de base de datos relacional
- Diseño de base de datos

Científico de Datos - funciones



Desarrollo de herramientas que agreguen valor



Interpretar complejos data set



Creación de modelo de Machine Learning



Presentar resultado y hacer sugerencias



Dar soporte y creacion de estrategias para el plan de Data Governance

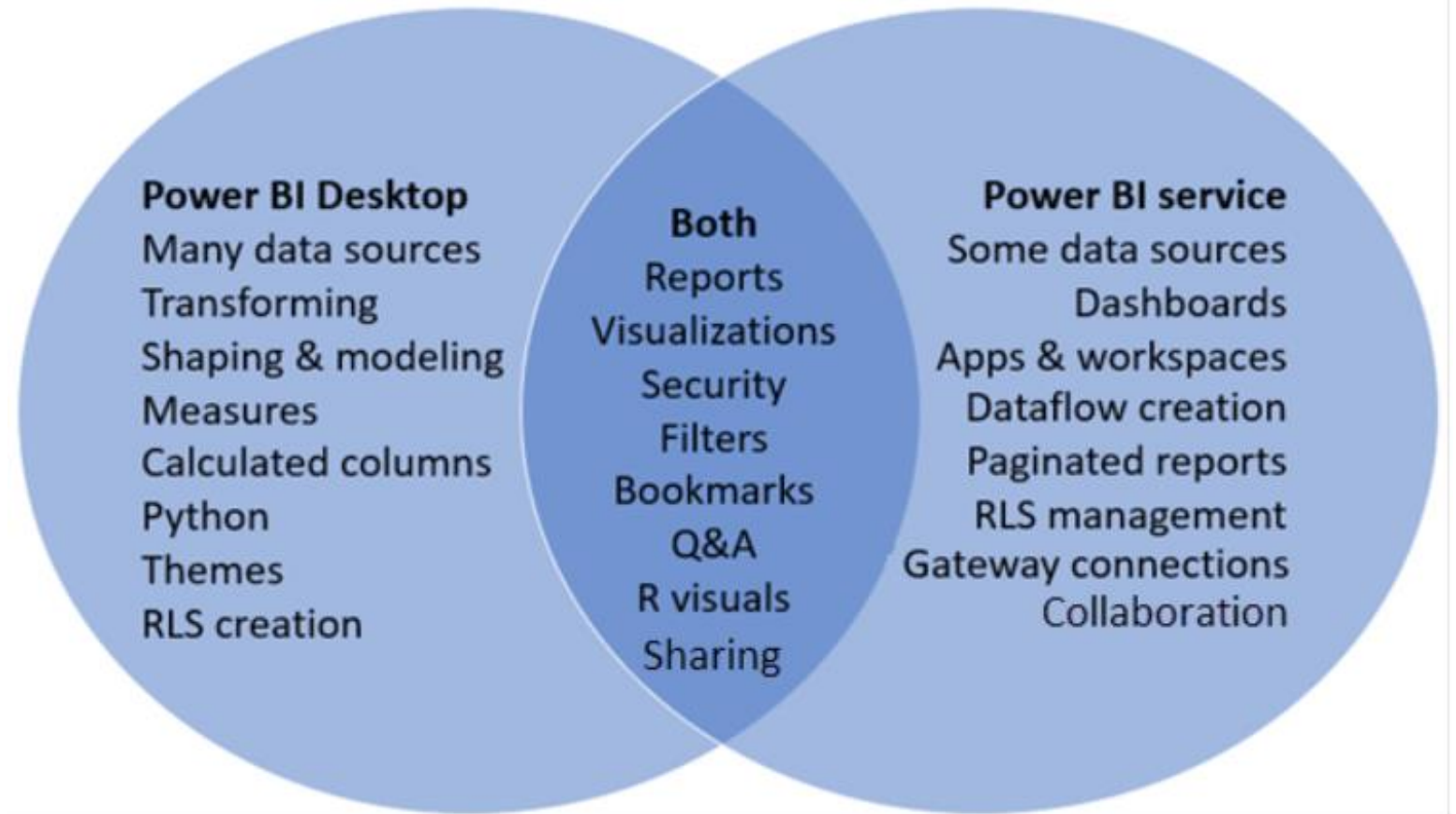
Oportunidades de crecimiento

- Pensamiento críticos
- Programación: Python - coding

¿Qué es el servicio Power BI?

- Power BI es una colección de servicios de software, aplicaciones y conectores que funcionan conjuntamente para ayudarle a crear, compartir y consumir información empresarial de forma que le ayude a usted y a su negocio de la manera más eficaz. El servicio Microsoft Power BI (<https://app.powerbi.com>), que a veces se conoce como Power BI en línea, es la parte de software como servicio (SaaS) de Power BI. Además, los paneles del servicio Power BI le ayudan a tomar el pulso de su empresa. Los paneles muestran iconos, que puede seleccionar para abrir informes y seguir explorando. Los paneles e informes se conectan a conjuntos de datos con el fin de reunir todos los datos relevantes en un solo lugar.

Power BI Desktop vs Service



The image features a dark gray background with three overlapping circles in two shades of blue. A horizontal white band runs across the middle of the image, containing the text "KEEP LEARNING" in a dark blue, sans-serif font.

KEEP LEARNING