



1. Big Data

Definición:

Big Data se refiere a conjuntos de datos extremadamente grandes y complejos que no pueden ser gestionados, procesados o analizados por las herramientas de procesamiento de datos tradicionales. El volumen de datos es tan grande que se requiere de tecnologías avanzadas para almacenarlos, procesarlos y analizarlos de manera efectiva.

Características clave de Big Data (las 3 Vs):

- **Volumen:** La cantidad de datos que se genera. Ejemplo: Terabytes de datos por día en redes sociales, transacciones financieras, etc.
- **Velocidad:** La rapidez con la que los datos son generados y procesados. Por ejemplo, los datos de sensores en tiempo real o las transacciones de comercio electrónico que se actualizan constantemente.
- **Variedad:** Los diferentes tipos de datos, que incluyen datos estructurados (tablas de bases de datos), semi-estructurados (archivos JSON, XML), y no estructurados (videos, imágenes, correos electrónicos, etc.).

Importancia en Power BI:

Power BI es una herramienta que permite conectar, visualizar y analizar grandes cantidades de datos, lo cual es clave cuando trabajamos con Big Data. Sin embargo, Power BI no necesariamente maneja Big Data por sí solo (especialmente cuando se trata de grandes volúmenes que requieren procesamiento en tiempo real), pero permite conectar a bases de datos y plataformas que gestionan Big Data, como **Azure Data Lake** o **Google BigQuery**.



2. Business Intelligence (BI)

Definición:

Business Intelligence (BI) se refiere al uso de tecnologías, procesos, y herramientas para transformar datos crudos en información procesable que ayuda a tomar decisiones empresariales. BI incluye la recopilación, integración, análisis, y presentación de datos. El objetivo es proporcionar a las organizaciones información valiosa que les permita optimizar sus operaciones y tomar decisiones basadas en datos.

Componentes de BI:

- **Data Warehousing:** Es el proceso de almacenar datos históricos y actuales en una base de datos centralizada.
- **ETL (Extract, Transform, Load):** Son los procesos que permiten extraer datos de diferentes fuentes, transformarlos (limpieza, integración) y cargarlos en un sistema de almacenamiento.
- **Análisis de Datos:** Uso de herramientas (como Power BI) para transformar los datos en informes y dashboards interactivos.
- **Visualización de Datos:** Presentación de los datos de manera comprensible, utilizando gráficos, tablas, mapas y otras representaciones visuales.

Power BI y BI:

Power BI es una de las herramientas de Business Intelligence más populares. Permite a los usuarios obtener datos de múltiples fuentes, transformarlos mediante Power Query, y crear visualizaciones interactivas en tiempo real. Con Power BI, los usuarios pueden realizar análisis de ventas, proyecciones financieras, análisis de tendencias, y mucho más, todo desde un único dashboard interactivo.



3. Data Warehousing (Almacenamiento de Datos)

Definición:

El **Data Warehousing** es el proceso de almacenar grandes cantidades de datos de manera estructurada en una base de datos centralizada. Los datos en un "data warehouse" provienen de diversas fuentes y se procesan y optimizan para facilitar el análisis.

Relación con BI:

Los sistemas de BI suelen conectarse a un **Data Warehouse** para extraer los datos necesarios para su análisis. Estos almacenes de datos permiten que los usuarios tengan acceso a información histórica y actualizada de manera eficiente.

En Power BI:

Power BI se conecta a diversas fuentes de datos, incluidas bases de datos y **Data Warehouses** como SQL Server, Oracle, y Azure SQL Database, lo que permite obtener información centralizada para su análisis.



4. Análisis Predictivo y Machine Learning

Definición:

Análisis Predictivo es el uso de técnicas estadísticas y algoritmos de machine learning para hacer predicciones sobre eventos futuros basados en datos históricos. El objetivo es identificar patrones y tendencias que ayuden a predecir comportamientos futuros.

En Power BI:

Power BI permite integrar **Modelos Predictivos** utilizando lenguajes como **R** o **Python**. Por ejemplo, un analista podría usar un modelo predictivo para predecir las ventas futuras de un producto basándose en datos históricos. Además, Power BI tiene la capacidad de incorporar resultados de modelos de Machine Learning creados en **Azure Machine Learning**.



5. Datos Estructurados y No Estructurados

Datos Estructurados:

Son aquellos que se almacenan en bases de datos relacionales, como tablas de SQL, con un formato y estructura bien definida (columnas, filas, tipos de datos).

Ejemplo: datos de ventas, inventarios, datos de clientes.

Datos No Estructurados:

Son aquellos que no tienen un formato fijo y suelen ser difíciles de organizar.

Ejemplo: correos electrónicos, imágenes, videos, archivos de texto, etc.

Power BI:

Power BI es muy eficiente para trabajar con **datos estructurados**. Sin embargo, se puede integrar con otras herramientas y servicios que ayudan a procesar **datos no estructurados** (como el análisis de texto con Power BI y Azure Cognitive Services).



6. KPIs (Key Performance Indicators - Indicadores Clave de Desempeño)

Definición:

Los **KPIs** son métricas utilizadas para medir el rendimiento de una organización en relación con sus objetivos estratégicos. Los KPIs pueden ser financieros (por ejemplo, ingresos) o no financieros (por ejemplo, satisfacción del cliente).

En Power BI:

Power BI permite a los usuarios crear y monitorear KPIs de forma interactiva. Por ejemplo, un KPI puede mostrar el porcentaje de ventas alcanzado en relación con la meta mensual de ventas.



7. Data Governance (Gobernanza de Datos)

Definición:

Data Governance se refiere a las políticas, procedimientos y normas establecidas para gestionar la calidad, privacidad y seguridad de los datos dentro de una organización.

En Power BI:

La gobernanza de datos en Power BI implica gestionar quién tiene acceso a los datos, cómo se transforman y cómo se comparten los informes. Power BI permite establecer permisos y roles para controlar el acceso a los datos.



8. Dashboard vs. Reporte

Dashboard:

Es una representación visual interactiva de los datos que proporciona una vista general de los KPIs y métricas clave en tiempo real. Los dashboards son más visuales e interactivos, permitiendo a los usuarios hacer clic y filtrar la información.

Reporte:

Es un documento detallado que muestra los resultados del análisis de datos. Los reportes suelen ser más estáticos y menos interactivos que los dashboards.

En Power BI:

Power BI permite crear tanto **dashboards** como **reportes**. Un **dashboard** proporciona una visión general rápida y visual, mientras que un **reporte** puede ser más detallado y contener múltiples páginas de información.