DATA Analytics

PREPARED BY: HEND OSAMA

SCAN ME

Pre-Assessment QR Code

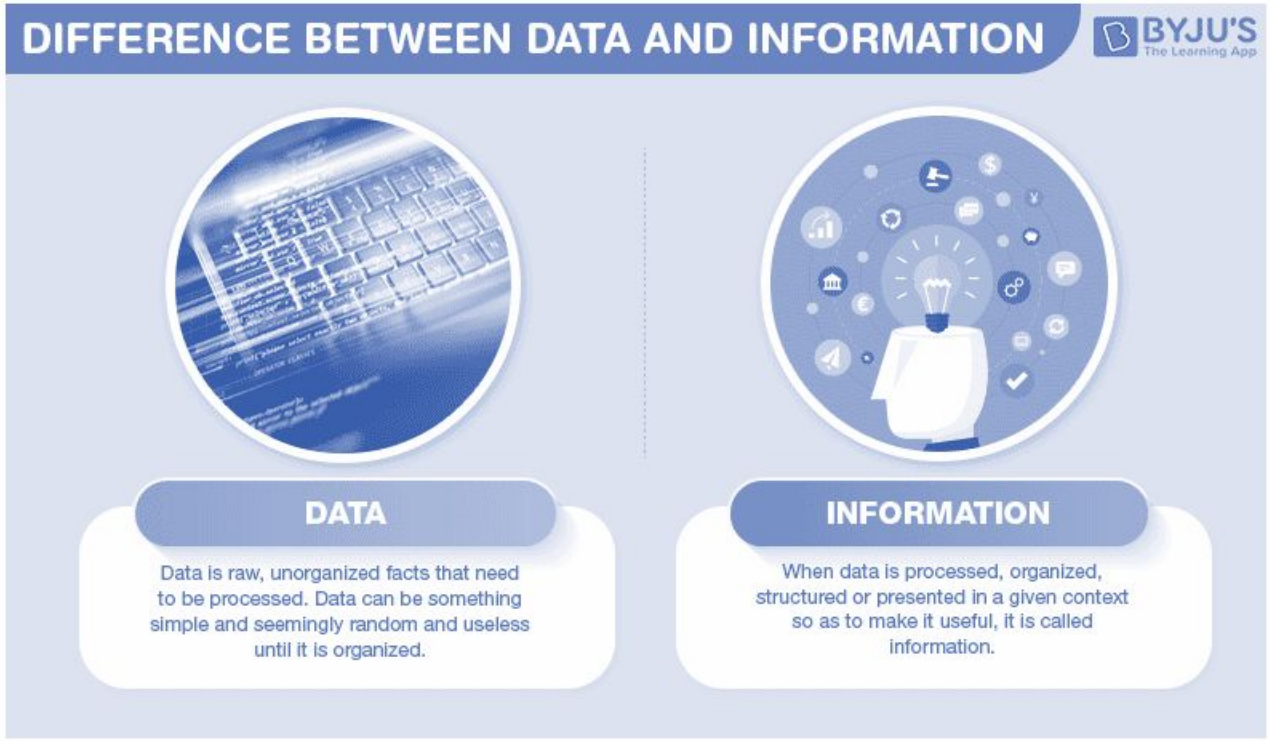
Training Expectation Questions

Agenda

- Importance of Data
- DIKW Pyramid
- Types of Information Processing
- Difference Between OLTP and OLAP
- Data Types
- Data Ecosystem & Lifecycle
- Business Intelligence
- Overview to data analytics
- Types of Data analytics
- Stages in Data Analytics Process

Importance of Data

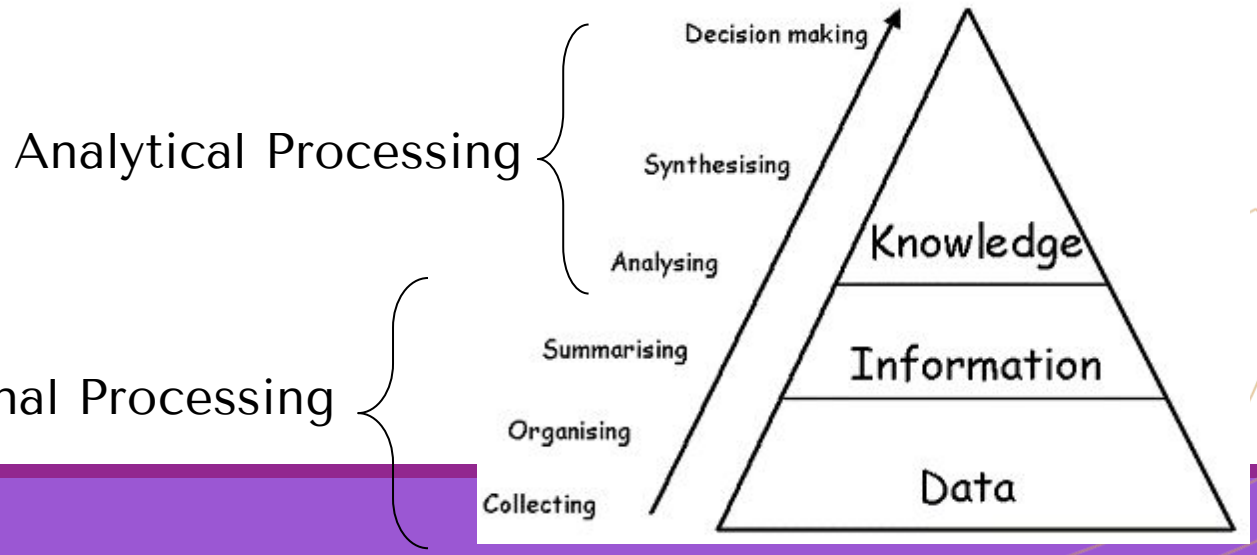
Data is becoming increasingly important in today's digital age. It is the raw material that fuels modern businesses and drives innovation across industries. When analysed effectively, data can reveal patterns and insights that businesses can use to make informed decisions, optimize performance, and gain a competitive advantage.



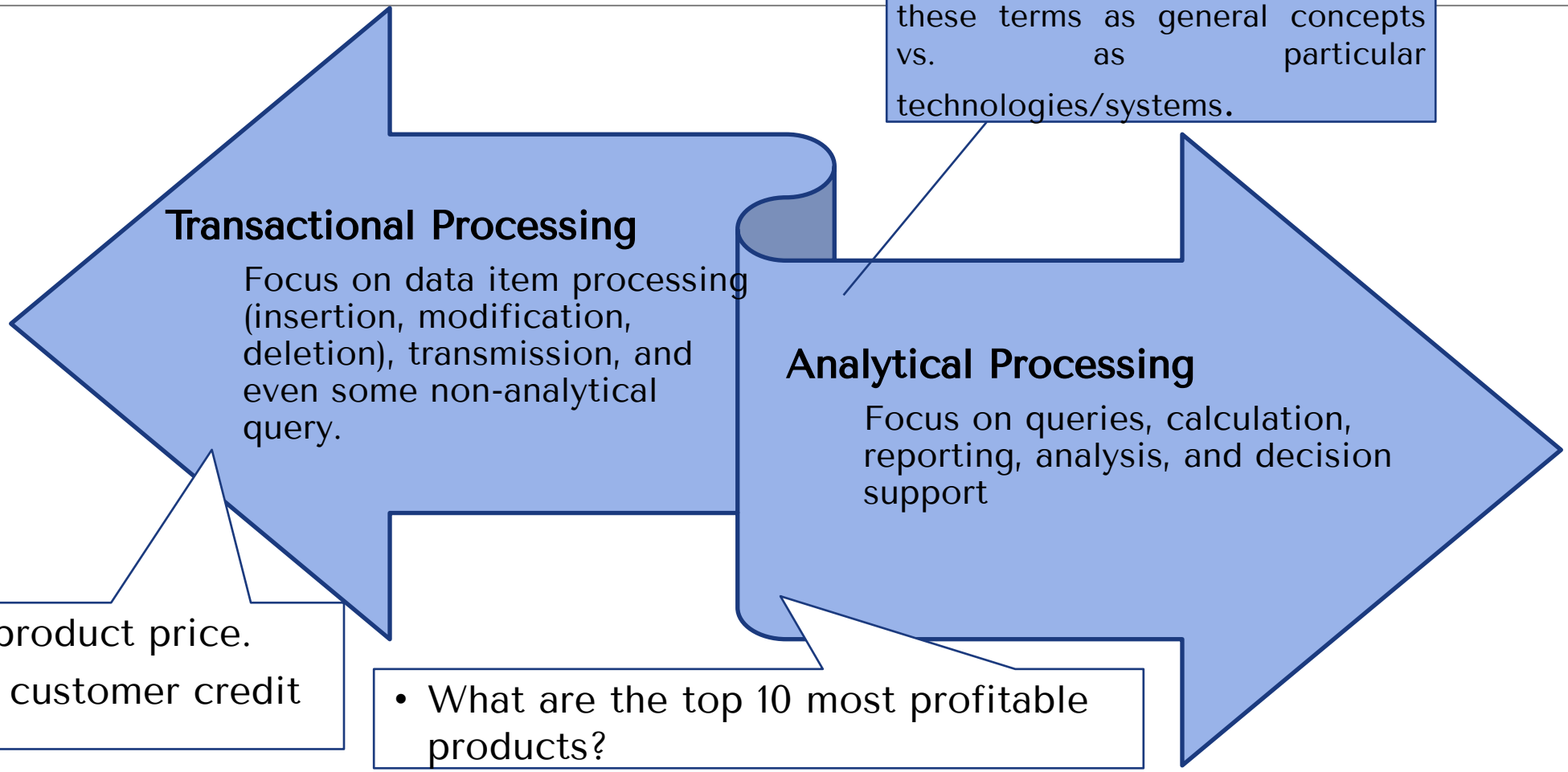
DIKW Pyramid

The DIKW hierarchy depicts relationships between data, information, knowledge (and wisdom).

- Data: raw value elements or facts
- Information: the result of collecting and organizing data that provides context and meaning
- Knowledge: the concept of understanding information. It provides insight to information, so it becomes useful and actionable.
- Wisdom: involves understanding and ability to make use of the data and information to answer questions, solve problems, make decisions.



Types of Information Processing



- Change product price.
- Increase customer credit limit.
- Import data from another source

- What are the top 10 most profitable products?
- Is there a significant increase of operational cost?

Difference Between OLTP and OLAP

OLTP and OLAP both are the online processing systems. OLTP is a transactional processing while OLAP is an analytical processing system. OLTP is a system that manages transaction-oriented applications on the internet for example, ATM. OLAP is an online system that reports to multidimensional analytical queries like financial reporting, forecasting, etc.

Data Types

Data are facts that are collected, recorded, stored, and processed by an information system.

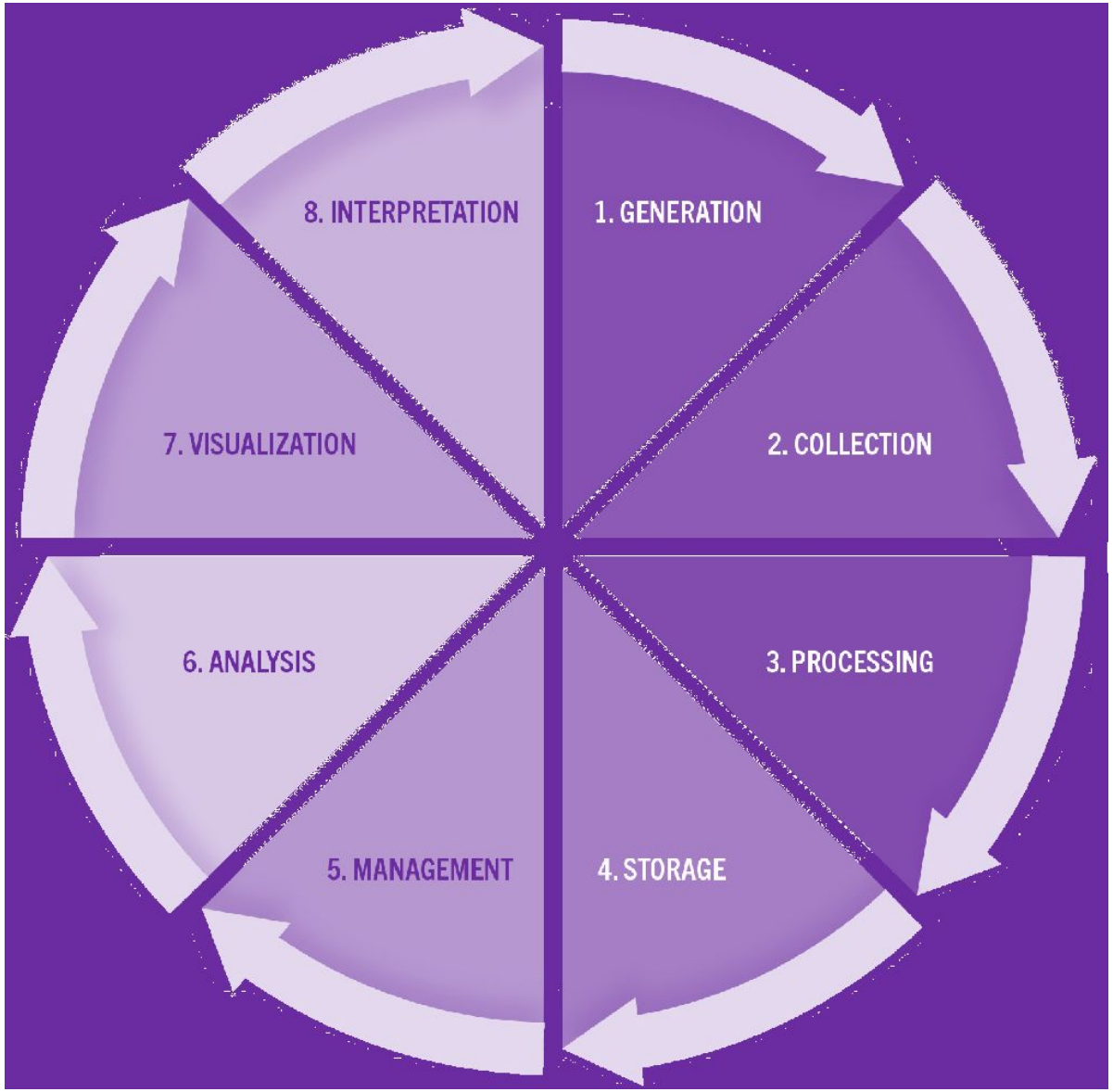
Information is processed data

»» Structured: Data is stored, processed, and manipulated in a traditional relational database management system (RDBMS).

»»Unstructured: Data that is commonly generated from human activities and does not fit into a structured database format. including formats like audio, video, and social media postings.

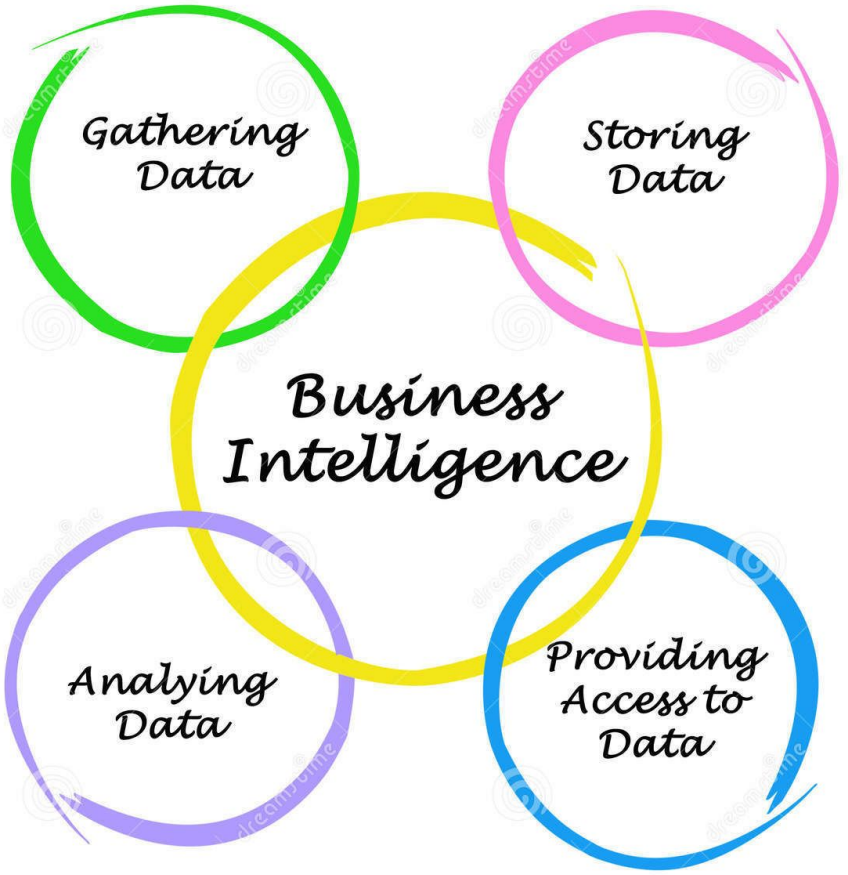
»»Semi-structured: Data does not fit into a structured database system, however its structured by tags that are useful for creating a form of order and hierarchy in the data. . Example of semi-structured data is a data represented in an XML file or excel sheets.

Data Ecosystem & Lifecycle

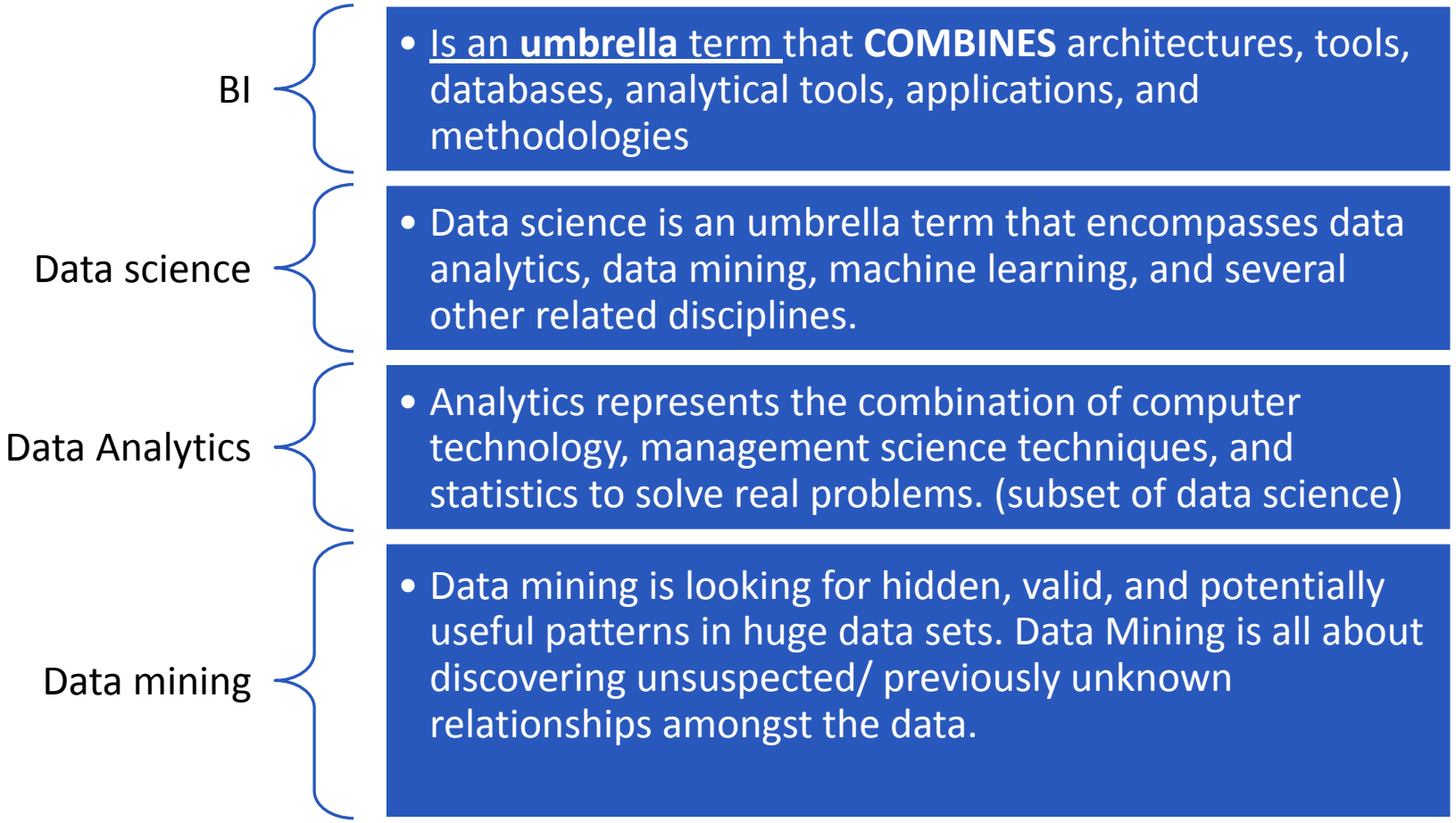


Business Intelligence

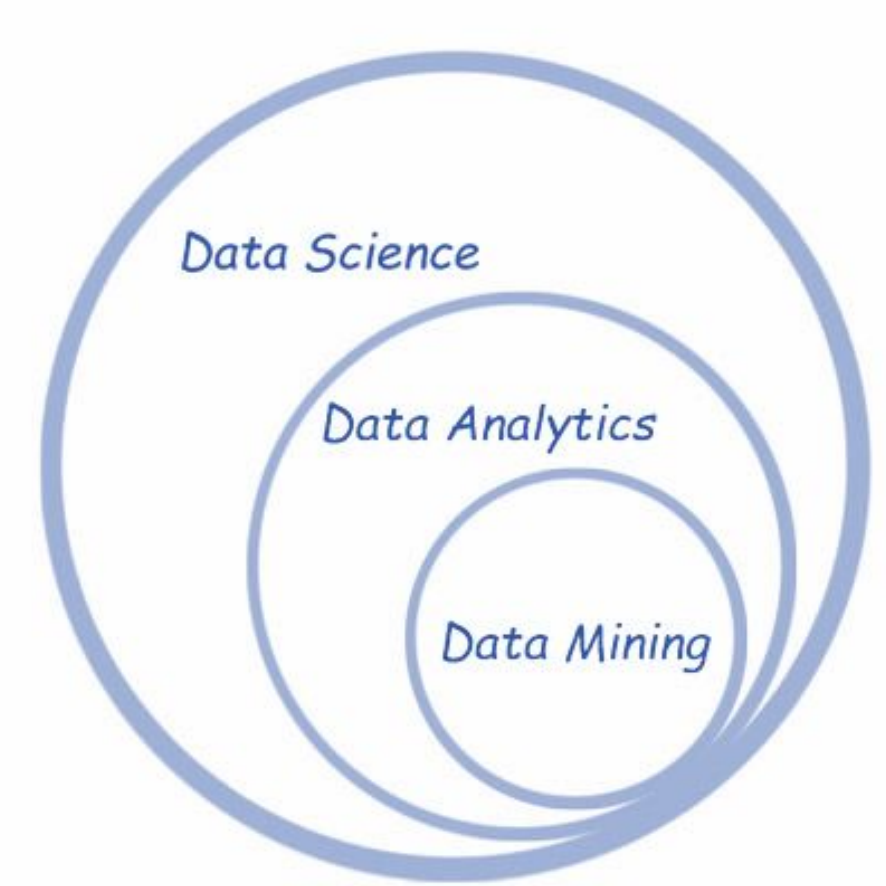
Business intelligence comprises the strategies and technologies used by enterprises for the data analytics of business information. BI technologies provide historical, current, and predictive views of business operations.



An Overview of Business Intelligence, Analytics, and Data Science



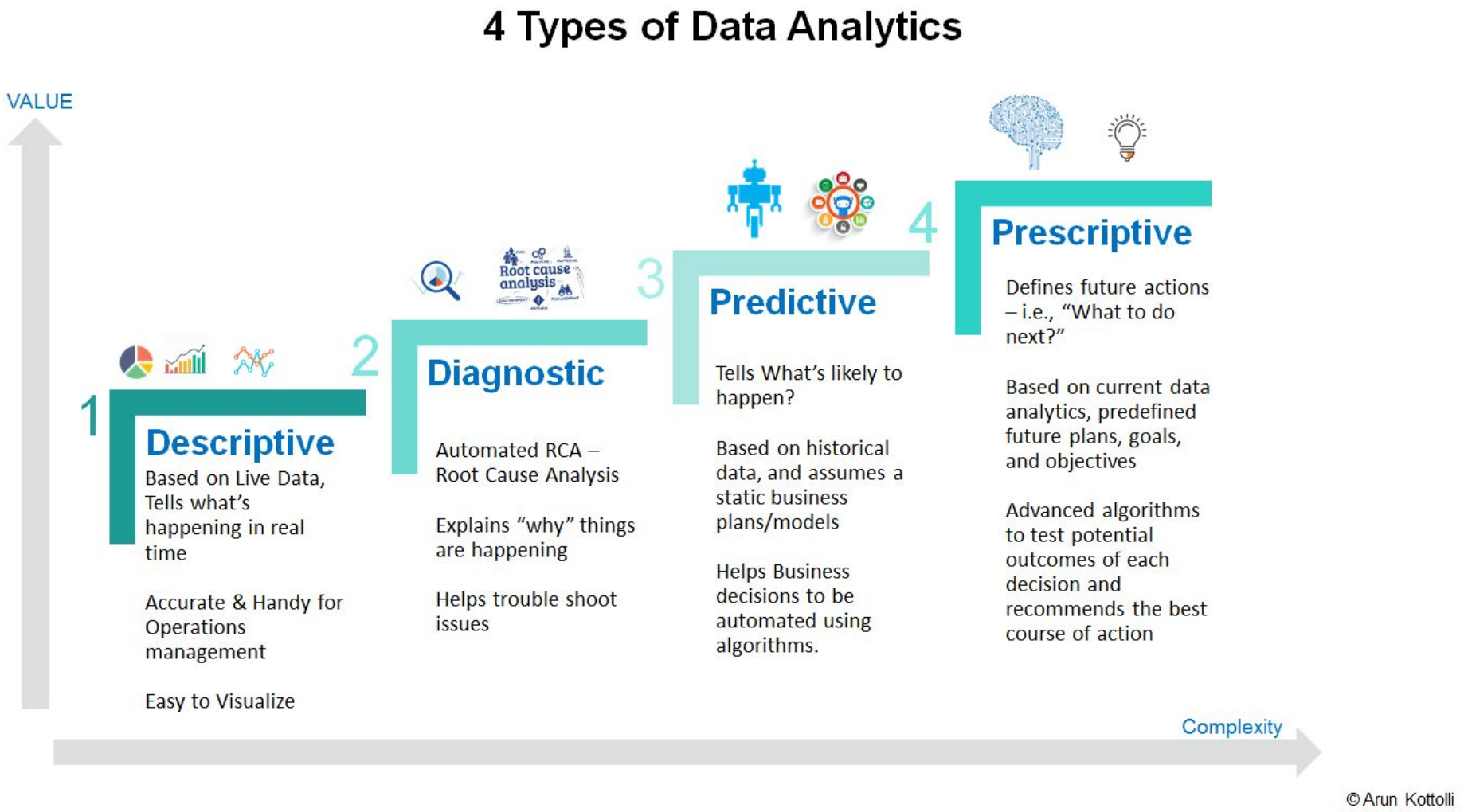
Data Analytics, and Data Science, and Data Mining



Overview to data analytics

- Data analytics is the process of identifying, cleaning, transforming, and modeling data to discover meaningful and useful information. The data is then crafted into a story through reports for analysis to support the critical decision-making process.
- As the world becomes more data-driven, storytelling through data analysis is becoming a vital component and aspect of large and small businesses. It is the reason that organizations continue to hire data analysts.
- Data-driven businesses make decisions based on the story that their data tells, Data analysis is, and should be, a critical aspect of all organizations to help determine the impact to their business, including evaluating customer sentiment, performing market and product research, and identifying trends or other data insights.

Types of Data analytics



Descriptive analytics

Descriptive analytics help answer questions about ***what has happened*** based on historical data. Descriptive analytics techniques summarize large datasets to describe outcomes to stakeholders.

By developing key performance indicators (KPIs), these strategies can help track the success or failure of key objectives. Metrics such as return on investment (ROI) are used in many industries, and specialized metrics are developed to track performance in specific industries.

An example of descriptive analytics is generating reports to provide a view of an organization's sales and financial data.

Diagnostic analytics

Diagnostic analytics help answer questions about ***why events happened***. Diagnostic analytics techniques supplement basic descriptive analytics, and they use the findings from descriptive analytics to discover the cause of these events. Then, performance indicators are further investigated to discover why these events improved or became worse. Generally, this process occurs in three steps:

- Identify anomalies in the data. These anomalies might be unexpected changes in a metric or a particular market.
- Collect data that's related to these anomalies.
- Use statistical techniques to discover relationships and trends that explain these anomalies.

Predictive analytics

Predictive analytics help answer questions about ***what will happen*** in the future. Predictive analytics techniques use historical data to identify trends and determine if they're likely to recur.

Predictive analytical tools provide valuable insight into what might happen in the future. Techniques include a variety of statistical and machine learning techniques such as neural networks, decision trees, and regression.

Prescriptive analytics

Prescriptive analytics help answer questions about ***which actions should be taken*** to achieve a goal or target. By using insights from predictive analytics, organizations can make data-driven decisions. This technique allows businesses to make informed decisions in the face of uncertainty.

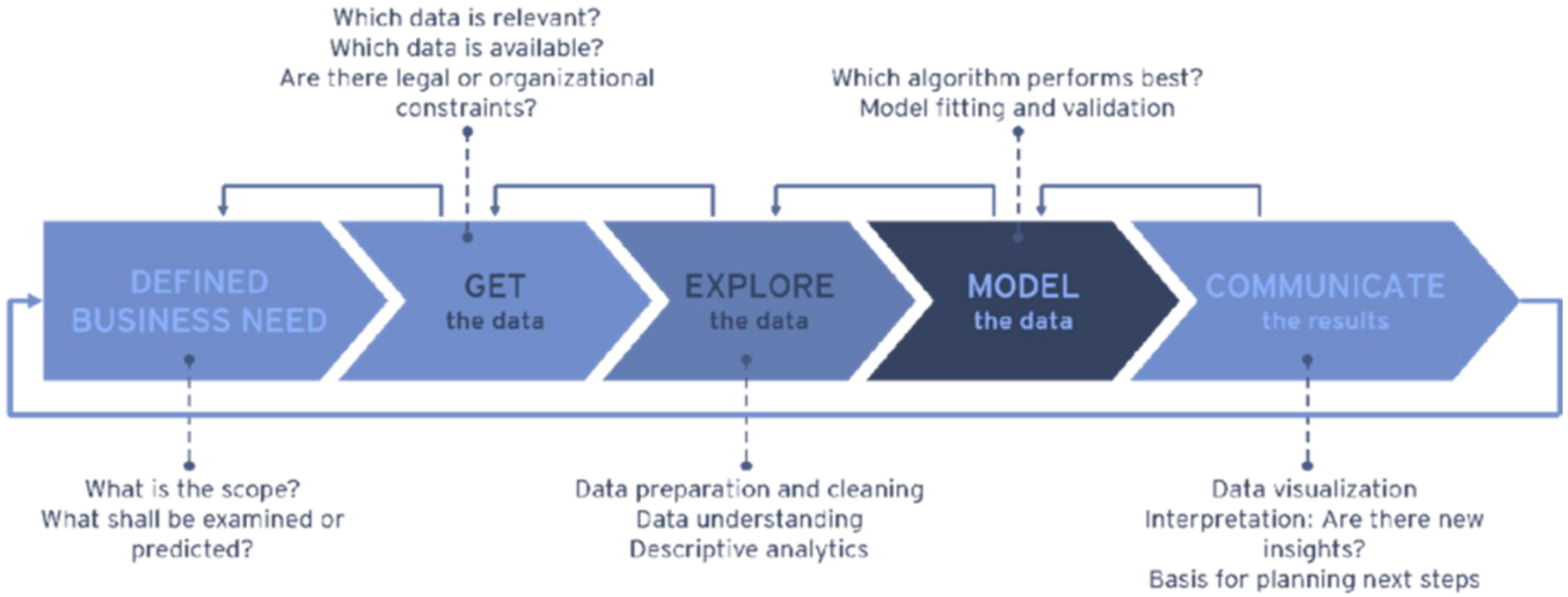
Prescriptive analytics techniques rely on machine learning strategies to find patterns in large datasets. By analyzing past decisions and events, organizations can estimate the probability of different outcomes.

Example

- By enabling reporting and data visualizations, a retail business uses descriptive analytics to look at patterns of purchases from previous years to determine what products might be popular next year. The company might also look at supporting data to understand why a particular product was popular and if that trend is continuing, which will help them determine whether to continue stocking that product.

- A business might determine that a certain product was popular over a specific timeframe. Then, they can use this analysis to determine whether certain marketing efforts or online social activities contributed to the sales increase.

Stages in Data Analytics Process



Stakeholders Mapping

In general, these are the main stakeholder groups for most companies.

Our goal is to study the expectations of stakeholder groups

Remember !

- ✔ Neglecting the expectations of even one stakeholder group can have a massive impact on profitability.
- ✔ The easiest and most effective way to understand the expectations of different stakeholder groups is by asking questions.
  - Ask "who", "what", "when", "where", "why" and "how" questions.
- ✔ The knowledge you will get from these interviews is the backbone on which you will proceed and start building the firm's business intelligence.

ETL Process

- Conducting analyses involve exploring relevant data that could be available at different resources.
- The process of extracting data from the source, transforming, and loading data into the specific target is known as ETL.
- ETL is a fundamental step in preparing data and delivering it in a format that could be used to perform accurate analyses.

ETL Process: Transform data

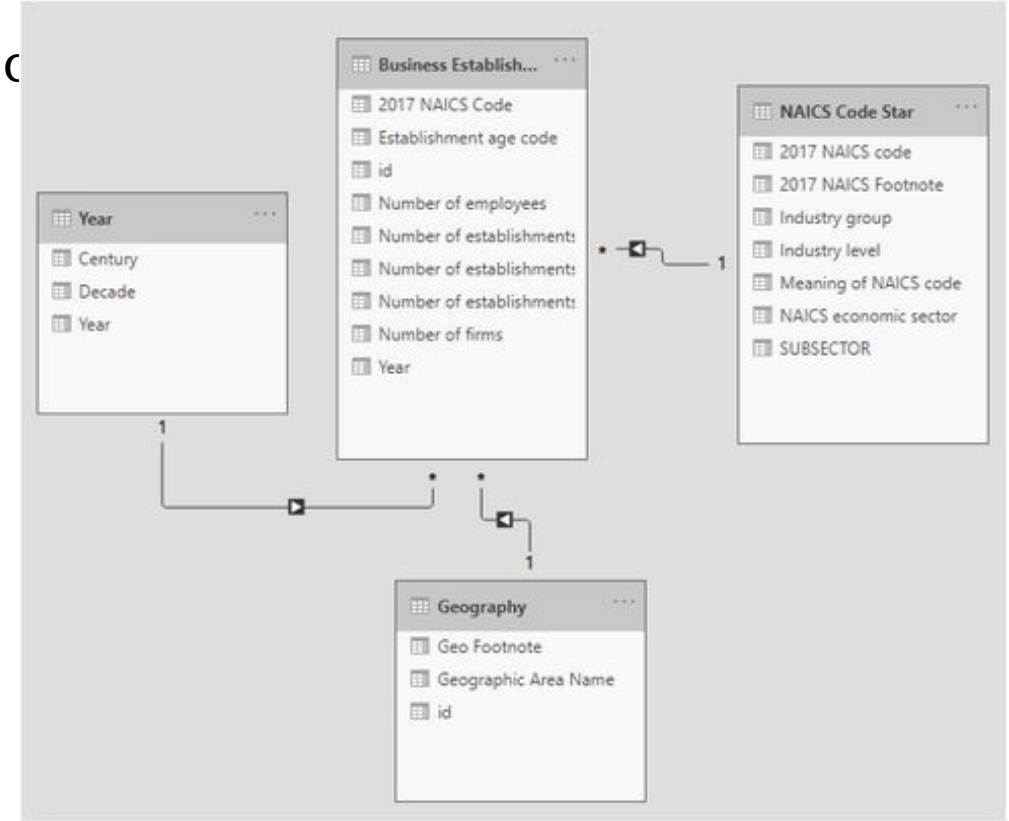
- Transformation is the main step where the ETL process adds value.
- It changes data and provides guidance whether data can be used for its intended purposes

Transforming includes:

- Selecting specific columns
- Recoding value
- Deriving new value through calculation
- Deduplicating, i.e., identifying and removing duplicate, records
- Joining data from multiple sources
- Aggregating multiple rows of data
- Conforming data so that separate sources can be used together
- Cleansing data to ensure data quality and consistency

What is Data Model?

- Conceptual view of data elements.
- Typically, it is a visual representation of connected tables.
- **Data models include:**
  - Tables
  - Columns
  - Relationships between tables
  - Data types
  - Keys



Power BI introduction

Microsoft Power BI is a collection of software services, apps, and connectors that work together to turn your unrelated sources of data into coherent, visually immersive, and interactive insights.

Whether your data is a simple Microsoft Excel workbook, or a collection of cloud-based and on-premises hybrid data warehouses.

Power BI lets you easily connect to your data sources, visualize (or discover) what's important, and share that with anyone or everyone you want.

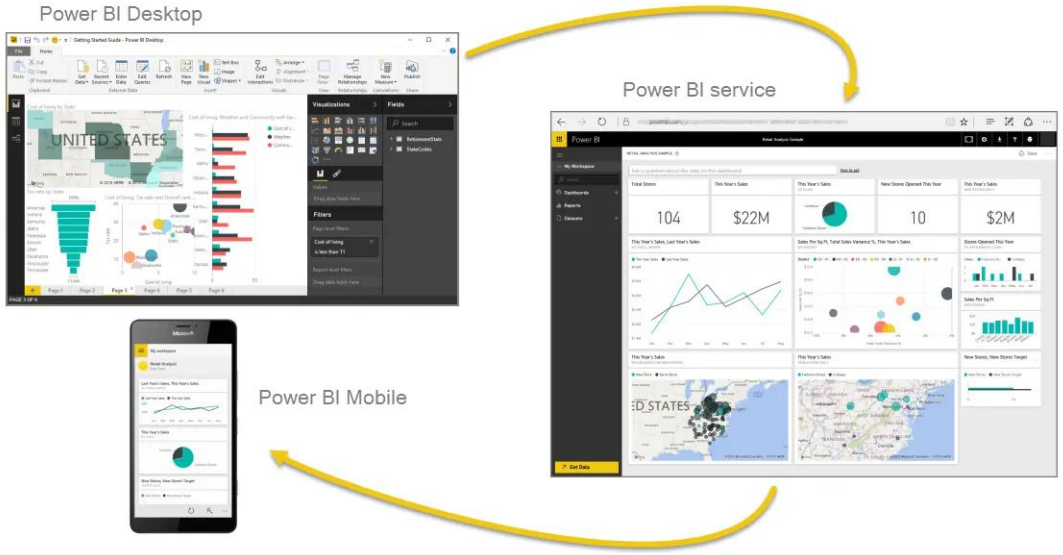
The parts of Power BI

Power BI consists of a Microsoft

a) Windows desktop application called Power BI Desktop,

b) an online SaaS (Software as a Service) service called the Power BI service, and

c) mobile Power BI apps that are available on any device, with native mobile BI apps for Windows, iOS, and Android.



How Power BI matches your role

How you use Power BI might depend on your role on a project or a team. And other people, in other roles, might use Power BI differently, which is just fine.

Download Power BI Desktop

- You can download Power BI Desktop from the web or as an app from the Microsoft Store on the Windows tab.

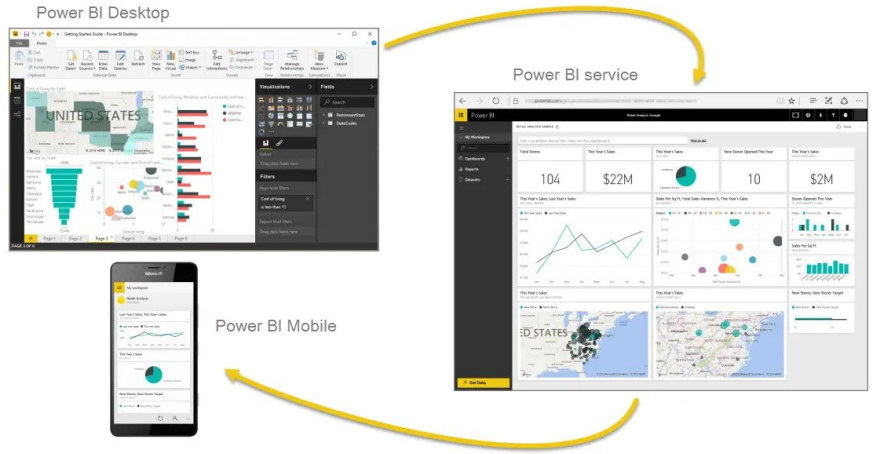
- For Mac users you can use power BI service as unfortunately there is no power BI desktop version for Mac.

Download Strategy	Link	Notes
Windows Store App	Windows Store	Will automatically stay updated
Download from web	Download .msi	Must manually update periodically

Power BI service & Mobile app

- Before you can sign into Power BI, you'll need an account. To get a free trial, go to app.powerbi.com and sign up with your email address.
- For detailed steps on setting up an account, see Sign in to Power BI service

- A common flow of work in Power BI begins in Power BI Desktop, where a report is created. That report is then published to the Power BI service and finally shared, so that users of Power BI Mobile apps can consume the information.



Building blocks of Power BI

Your first order of business is to learn the basic building blocks of Power BI, which will provide a solid basis for turning data into cool reports and visuals.

Everything you do in Microsoft Power BI can be broken down into a few basic building blocks.

Here are the basic building blocks in Power BI:

Datasets

Visualizations

Reports

Dashboards

Tiles

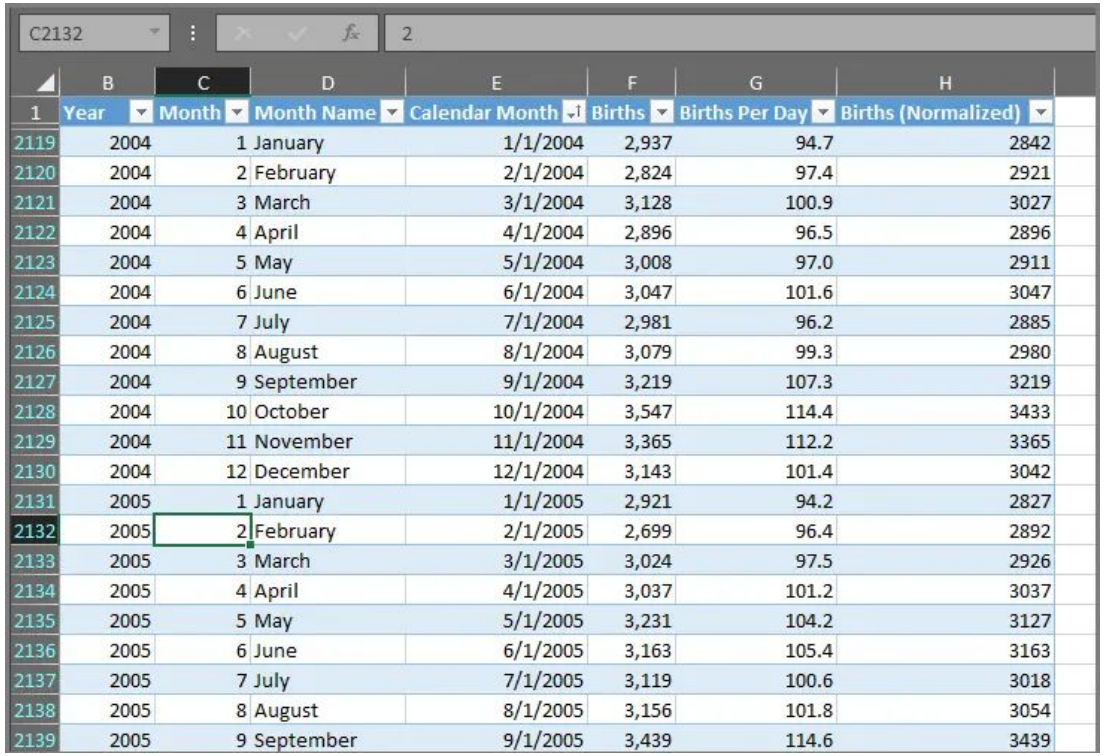
Datasets

A dataset is a collection of data that Power BI uses to create its visualizations.

You can have a simple dataset that's based on a single table from a Microsoft Excel workbook, similar to what's shown in the following image.

Datasets can also be a combination of many different sources, which you can filter and combine to provide a unique collection of data (a dataset) for use in Power BI.

Power BI has built-in data connectors that let you easily connect to that data, filter it if necessary, and bring it into your dataset.

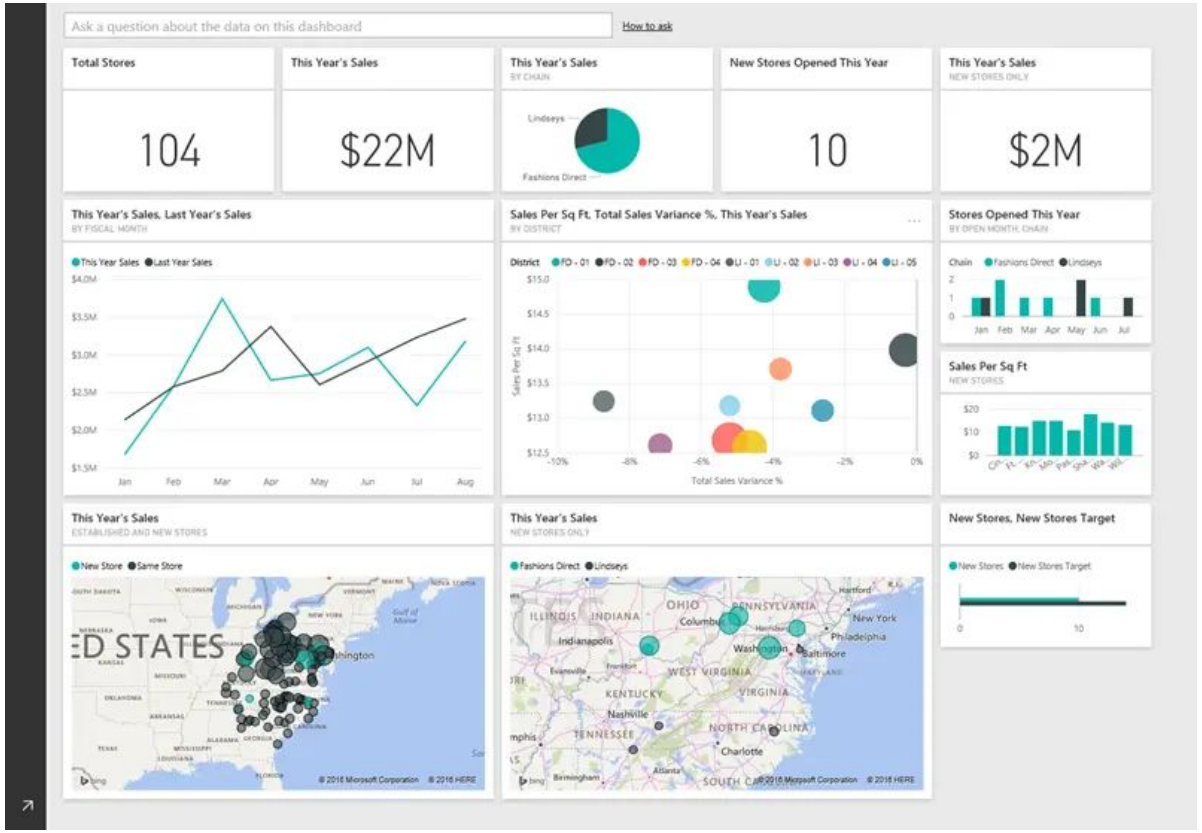


C2132 | 2

	B	C	D	E	F	G	H
1	Year	Month	Month Name	Calendar Month	Births	Births Per Day	Births (Normalized)
2119	2004	1	January	1/1/2004	2,937	94.7	2842
2120	2004	2	February	2/1/2004	2,824	97.4	2921
2121	2004	3	March	3/1/2004	3,128	100.9	3027
2122	2004	4	April	4/1/2004	2,896	96.5	2896
2123	2004	5	May	5/1/2004	3,008	97.0	2911
2124	2004	6	June	6/1/2004	3,047	101.6	3047
2125	2004	7	July	7/1/2004	2,981	96.2	2885
2126	2004	8	August	8/1/2004	3,079	99.3	2980
2127	2004	9	September	9/1/2004	3,219	107.3	3219
2128	2004	10	October	10/1/2004	3,547	114.4	3433
2129	2004	11	November	11/1/2004	3,365	112.2	3365
2130	2004	12	December	12/1/2004	3,143	101.4	3042
2131	2005	1	January	1/1/2005	2,921	94.2	2827
2132	2005	2	February	2/1/2005	2,699	96.4	2892
2133	2005	3	March	3/1/2005	3,024	97.5	2926
2134	2005	4	April	4/1/2005	3,037	101.2	3037
2135	2005	5	May	5/1/2005	3,231	104.2	3127
2136	2005	6	June	6/1/2005	3,163	105.4	3163
2137	2005	7	July	7/1/2005	3,119	100.6	3018
2138	2005	8	August	8/1/2005	3,156	101.8	3054
2139	2005	9	September	9/1/2005	3,439	114.6	3439

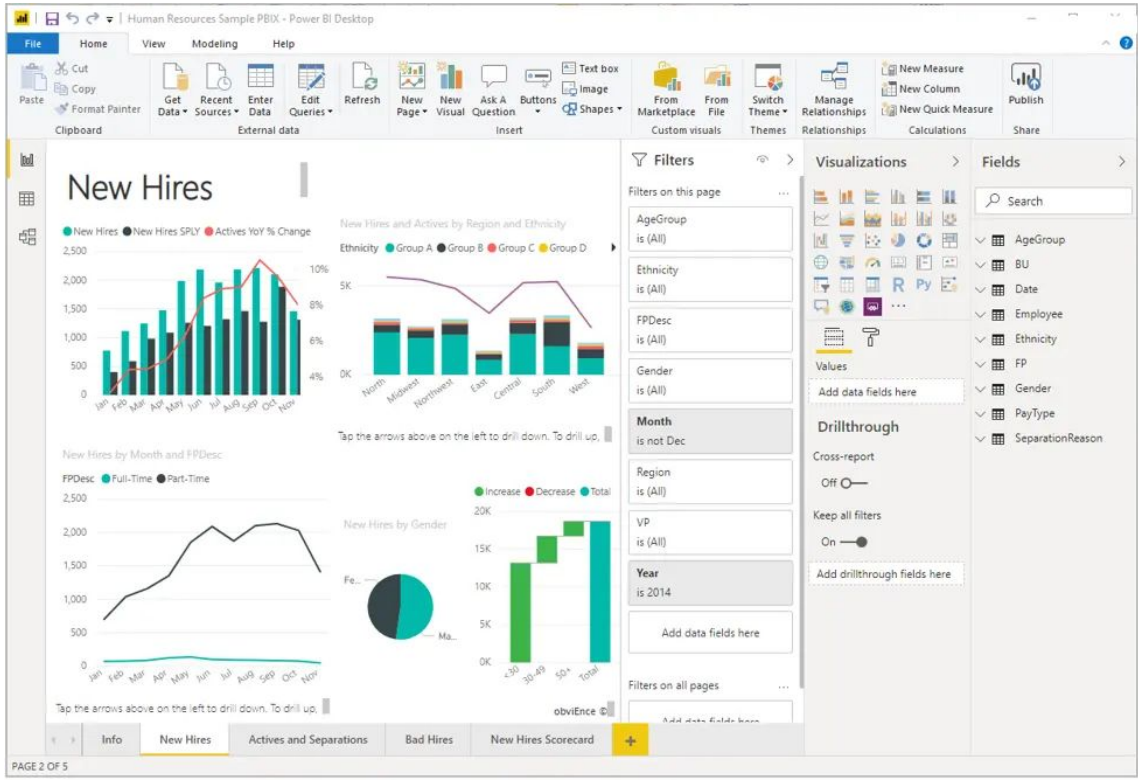
Visualizations

A visualization (sometimes also referred to as a visual) is a visual representation of data, like a chart, a color-coded map, or other interesting things you can create to represent your data visually. Power BI has all sorts of visualization types, and more are coming all the time. The following image shows a collection of different visualizations that were created in Power BI.



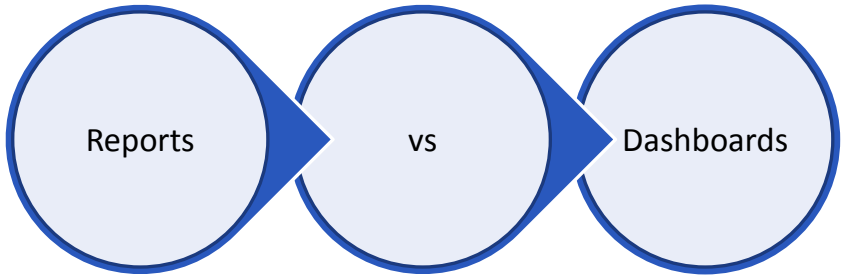
Reports

In Power BI, a report is a collection of visualizations that appear together on one or more pages. Just like any other report you might create for a sales presentation or write for a school assignment, a report in Power BI is a collection of items that are related to each other. The following image shows a report in Power BI Desktop—in this case, it's the second page in a five-page report. You can also create reports in the Power BI service.



Dashboards

When you're ready to share a report, or a collection of visualizations, you create a dashboard. Much like the dashboard in a car, a Power BI dashboard is a collection of visuals from a single page that you can share with others. Often, it's a selected group of visuals that provide quick insight into the data or story you're trying to present.

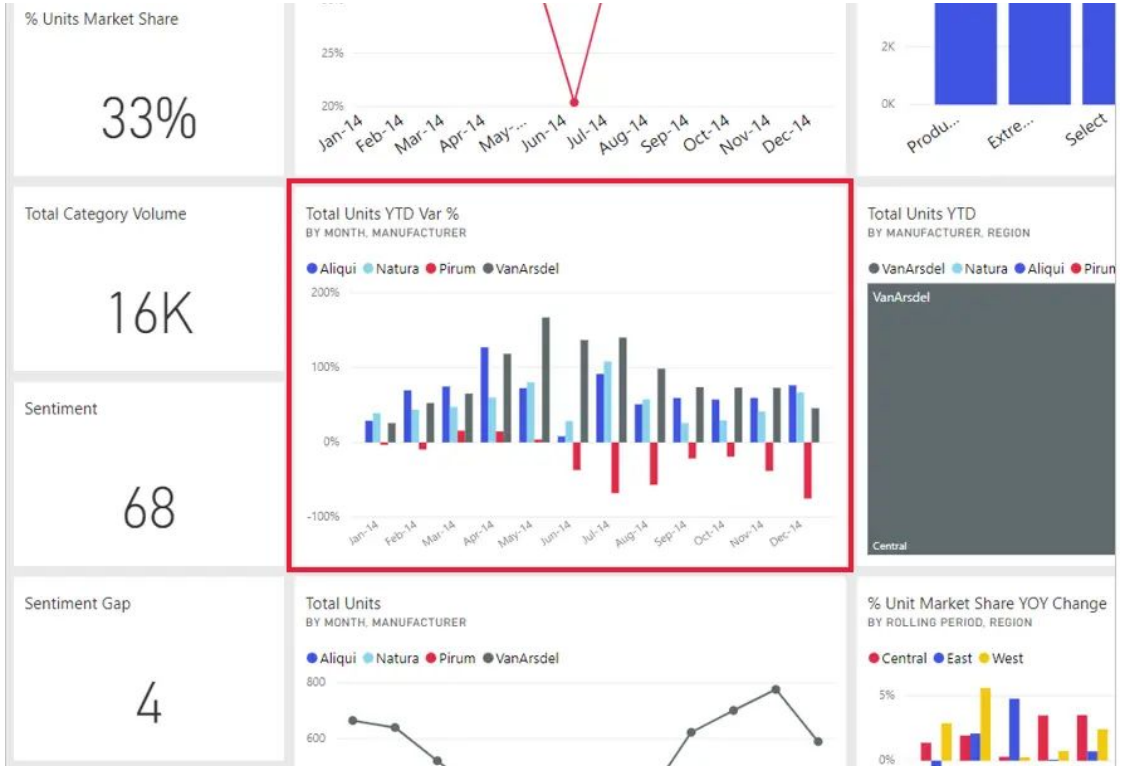


A dashboard must fit on a single page, often called a canvas. Think of it like the canvas that an artist or painter uses—a workspace where you create, combine, and rework interesting and compelling visuals. You can share dashboards with other users or groups, who can then interact with your dashboards when they're in the Power BI service or on their mobile device.

Tiles

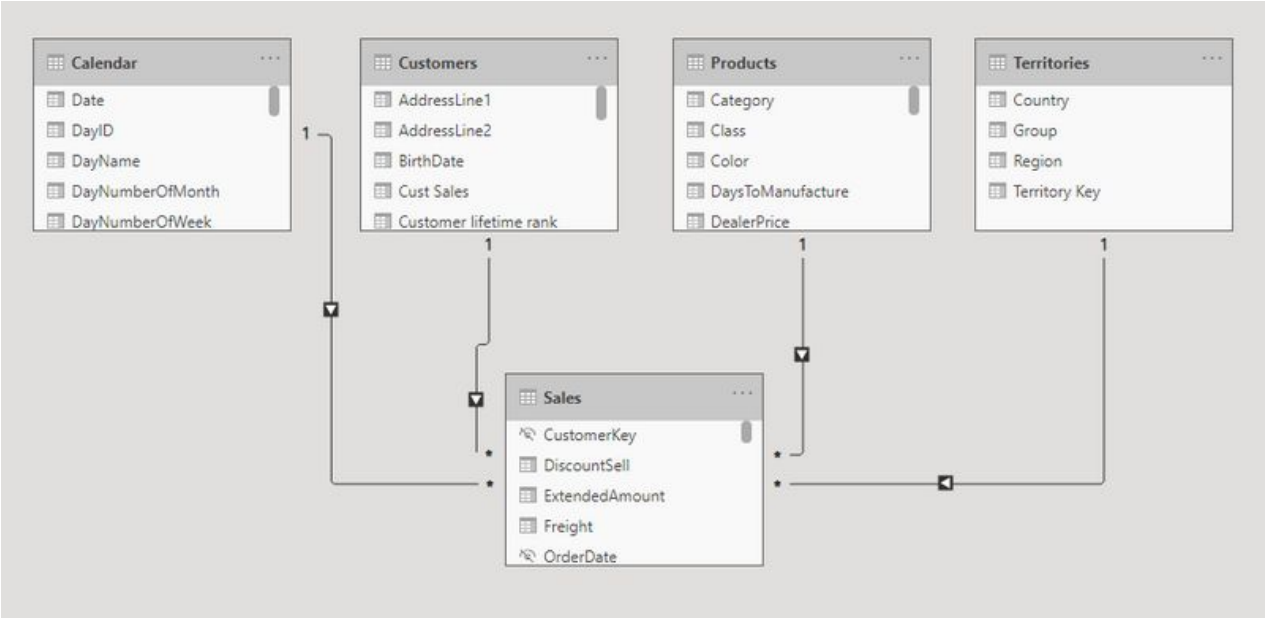
In Power BI, a tile is a single visualization on a dashboard. It's the rectangular box that holds an individual visual. In the following image, you see one tile, which is also surrounded by other tiles.

Summary: using Power BI you create datasets, build compelling visuals, and share them with others.



Data modeling

Data modeling is the process of analyzing and defining all different data your business collect and produces, as well as the relationships between those bits of data.



Practical Example

Finally, we will conduct practical example about how to import data from different data sources (excel, database) to develop a dataset and start visualizing data to create a report and a dashboard

Thank you