



# **Introduction to Data Analysis, Visualisation Tools and Techniques**

Bringing your data to life with visualisation

# Contents

- Why the need for Data Analysis ?
- What is Data Analysis?
- Data Analysis STEPS
- Is Data Analysis Mining?
- Data Analysis types
- Data analysis Techniques
- Why Do We Need to Visualisation Data?
- What is Data Visualization?
- Types of Visualization

# WHY IS DATA ANALYSIS IMPORTANT ?

## Five Reasons Why!



### Clear Focus

With data analysis, businesses are able to effectively target customers likely to buy their product or service.



### Solve Problems

With analytics, will have evidence to solve operational costs resulting to efficient internal systems and procedures



### Innovate

An understanding of future trends means businesses are able to design futuristic innovations and solutions.



### Cut Costs

Business with suitable analytics solutions are able to identify and cut avoidable operational costs and other inefficiencies.



### Attract Customers

With analytics, will have evidence to solve operational costs resulting to efficient internal systems and procedures

**Data Analysis**- examining, cleaning, transforming, and modeling data to discover useful information, support decision-making, and identify patterns

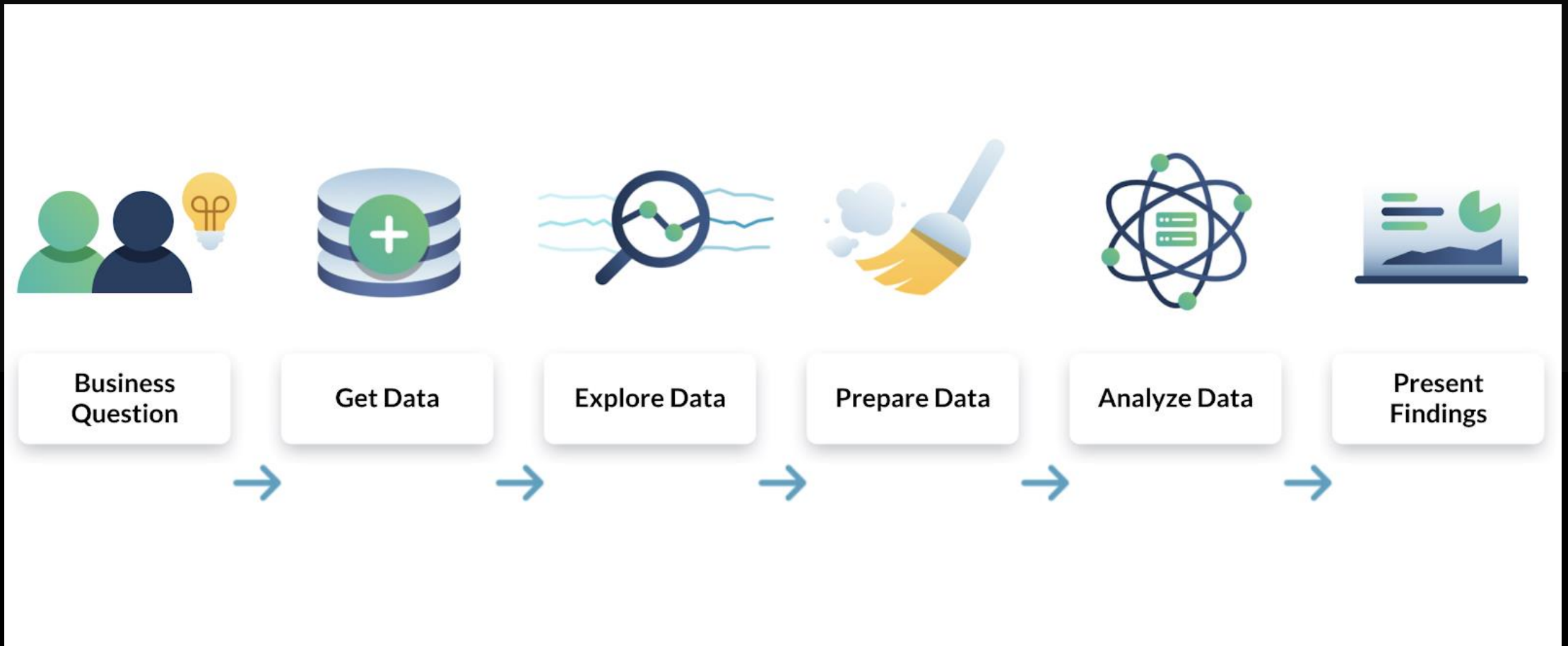


Image: Data CAMP

# Data Analysis-**STEPS**

**Data Collection:** Gathering data from various sources such as databases, sensors, surveys, etc.

**Data Cleaning:** Removing or correcting errors, handling missing data, and ensuring the dataset is consistent.

**Exploratory Data Analysis (EDA):** Summarizing the data using descriptive statistics, detecting outliers, and understanding data patterns.

**Interpretation:** Drawing insights, making decisions, or generating actionable strategies based on analysis.

# IS DATA ANALYSIS MINING ?

ASPECT	DATA ANALYSIS	DATA MINING
PURPOSE	Explain and interpret data	Discover hidden patterns and relationships
APPROACH	Hypothesis-driven , structured	Exploratory, pattern discovery
SCOPE	Specific problems, smaller datasets	Large datasets, complex environments
TECHNIQUES	Statistical methods(regression, correlation, etc.)	Machine learning, clustering, classification
END GOAL	Understanding and decision-making	Prediction and new insights
TOOLS	Excel, R, SPSS, Python (Pandas, Matplotlib)	Weka, RapidMiner, Apache Spark, Scikit-learn

# Data analysis-TYPES



**Descriptive  
Analytics**



**Diagnostic  
Analytics**



**Predictive  
Analytics**



**Prescriptive  
Analytics**



# Data analysis-TYPES...

## Descriptive Statistics:

- Summarizes data using measures like mean, median, mode, standard deviation, and variance.
- It helps understand data distribution and central tendencies.



Descriptive Analysis



# Data analysis-TYPES...

## Prescriptive analysis

- The most advanced type of data analysis.
- It not only predicts future outcomes but also suggests actions to benefit from these predictions.
- It uses sophisticated tools and technologies like machine learning and artificial intelligence to recommend decisions.
- For example, a prescriptive analysis might suggest the best marketing strategies to increase future sales.



Prescriptive Analysis

# Data analysis-**TYPES...**

## Predictive analysis

- It uses statistical models and forecasting techniques to understand the future.
- It involves using data from the past to predict what could happen in the future.
- This type of analysis is often used in risk assessment, marketing, and sales forecasting.
- For example, a company might use predictive analysis to forecast the next quarter's sales based on historical data



# Data analysis-**TYPES...**

## Diagnostic analysis

- It goes a step further than descriptive analysis by determining why something happened.
- It involves more detailed data exploration and comparing different data sets to understand the cause of a particular outcome.
- For instance, if a company's sales dropped in a particular month, diagnostic analysis could be used to find out why.

# Data Analysis - Techniques

According to their unique purpose and application.

- Exploratory analysis,
- Regression analysis,
- Monte Carlo simulation,
- Factor analysis,
- Cohort analysis,
- Cluster analysis,
- Time series analysis, and
- Sentiment analysis.

# Exploratory analysis

- Exploratory analysis is used to understand the main characteristics of a data set.
- It is often used at the beginning of a data analysis process to summarize the main aspects of the data, check for missing data, and test assumptions.
- This technique involves visual methods such as scatter plots, histograms, and box plots.

# Regression analysis

- Regression analysis is a statistical method used to understand the relationship between a dependent variable and one or more independent variables.
- It is commonly used for forecasting, time series modeling, and finding the causal effect relationships between variables.

## Factor analysis

- Factor analysis is a technique used to reduce a large number of variables into fewer factors.
- The factors are constructed in such a way that they capture the maximum possible information from the original variables.
- This technique is often used in market research, customer segmentation, and image recognition.

## Monte Carlo simulation

- Monte Carlo simulation is a technique that uses probability distributions and random sampling to estimate numerical results.
- It is often used in risk analysis and decision-making where there is significant uncertainty.



## Cluster analysis

- Cluster analysis is a technique used to group a set of objects in such a way that objects in the same group (called a cluster) are more similar to each other than to those in other groups.
- It is often used in market segmentation, image segmentation, and recommendation systems.

## Cohort analysis

- Cohort analysis is a subset of behavioral analytics that takes data from a given dataset and groups it into related groups for analysis.
- These related groups, or cohorts, usually share common characteristics within a defined time span.
- This technique is often used in marketing, user engagement, and customer lifecycle analysis.

## Time series analysis

- Time series analysis is a statistical technique that deals with time series data, or trend analysis.
- It is used to analyze the sequence of data points to extract meaningful statistics and other characteristics of the data.
- This technique is often used in sales forecasting, economic forecasting, and weather forecasting.

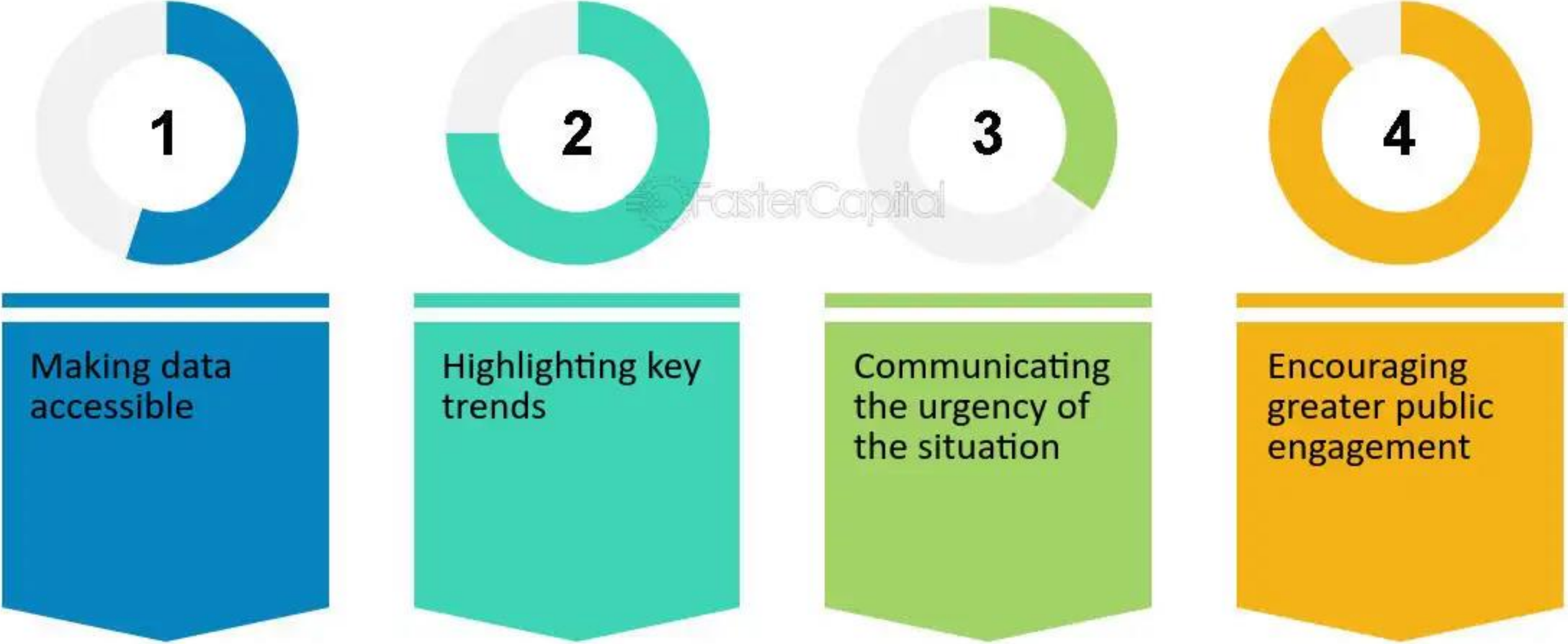
## Sentiment analysis

- Sentiment analysis, also known as opinion mining, uses natural language processing, text analysis, and computational linguistics to identify and extract subjective information from source materials.
- It is often used in social media monitoring, brand monitoring, and understanding customer feedback.

# Why do we need to visualisation data ?



# Importance of Visualizing Climate Change Data



## What is data visualisation?

- The process of representing data in graphical form, using tools like graphs, charts, images and animations.
- It helps illustrate findings from data analysis, and is a useful resource in guiding non-technical parties in decision-making.

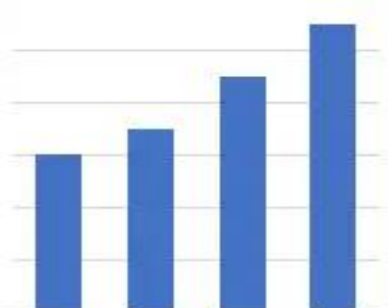
# Data visualisation tools and techniques

- There are a plethora of data visualisation tools and techniques we could use.
- Each of them is best-suited for a particular type of data, and can help illustrate complex ideas and highlight important outcomes of the data analysis process.

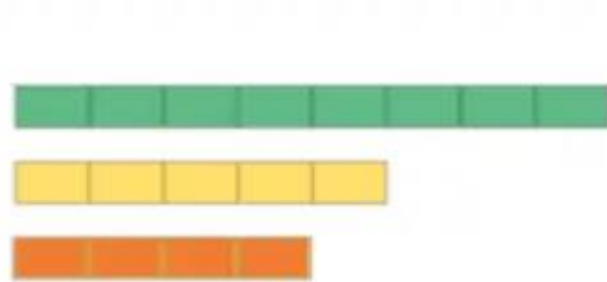




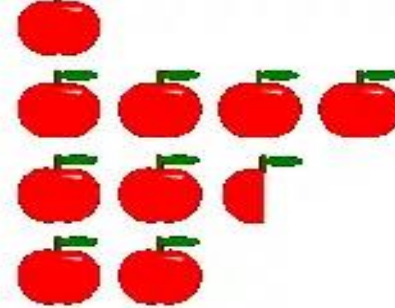
Pie Chart



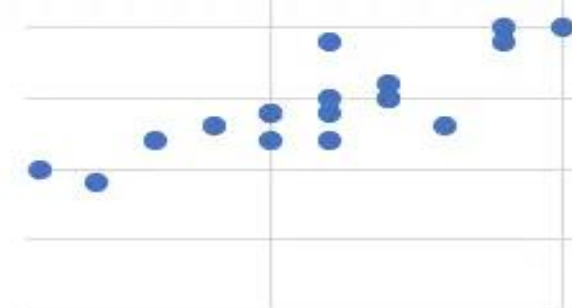
Bar Chart



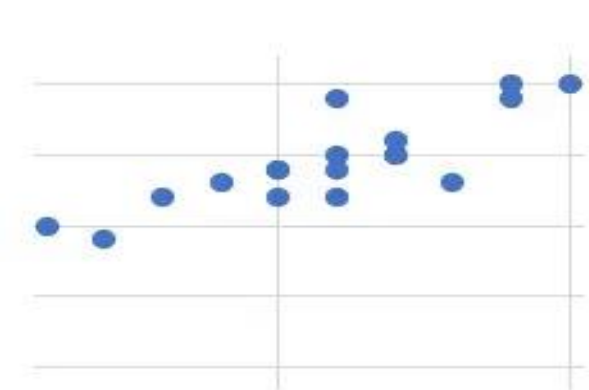
Tape Diagram



Pictograph



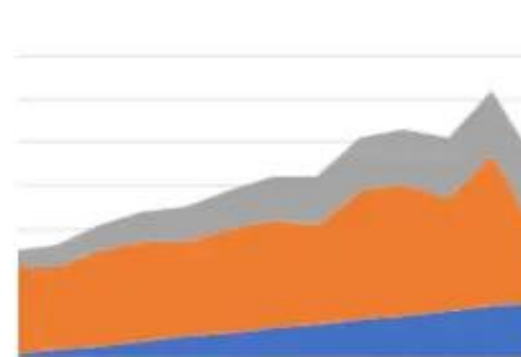
Scatterplot



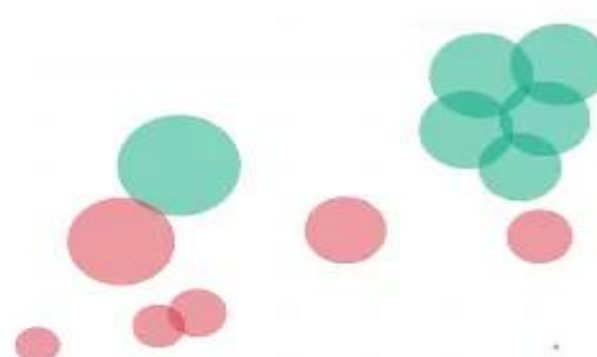
Scatterplot



Time series



Area Chart



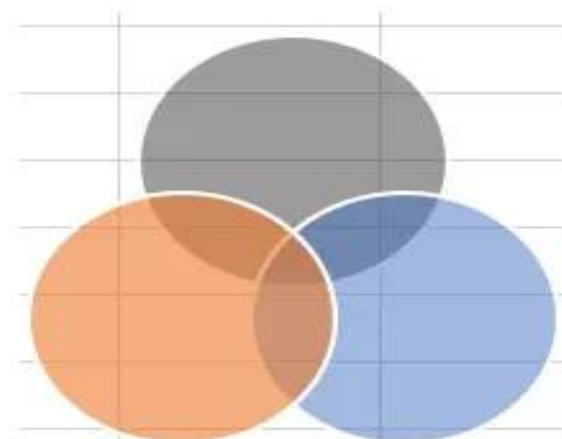
Bubble Graph



Line Chart



Radar Chart



Venn Diagram



Heatmaps

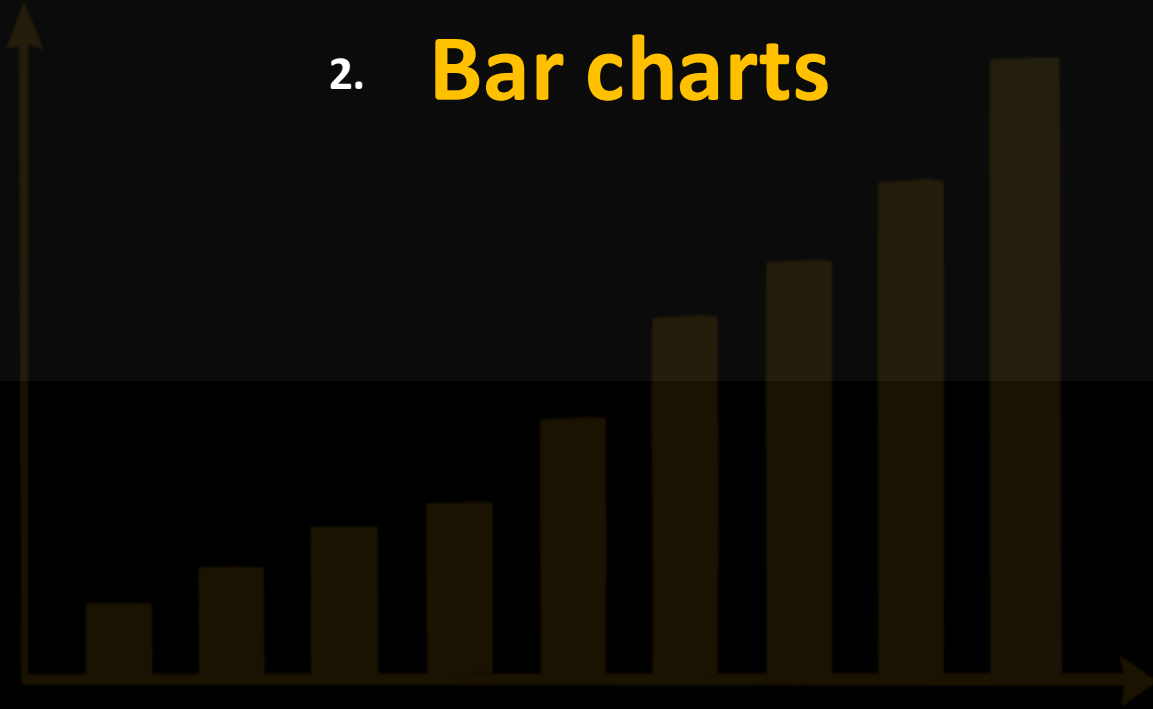
## 1. Time series



**A time series is a set of data points that are indexed in time order. Most often, time series data represents data collected at equally-spaced points in time. Time series data is used to discover how a phenomenon changes over time. An example of time series data is data collected from the stock market showing the prices of a commodity at the end of each trading day, the daily max temperature of a particular state over a year, and the population of a bacterial colony measured each hour.**

**Time series data is often visualised using a line plot or a line graph, where time is the horizontal axis, and the quantity being measured is on the vertical axis.**

## 2. **Bar charts**



- A bar chart is a data visualisation instrument that presents categorical data in the form of rectangular bars with heights or lengths proportional to the values of the quantities they represent.
- It is an easily understandable means of communicating findings from various types of data.

### **Several types of bar charts exist:**

- compound bar charts (which combine multiple data points into a single chart)
- stacked bar charts (where each bar is made up of segments representing different variables).



### 3. **Spatial maps**

# Spatial maps

- When analysing specific types of data, especially data related to specific geographical locations, it can be useful to have the data visualised in form of a spatial map.
- A spatial map is a visualisation of a region in space, which can then be populated with the relevant data.
- Spatial maps are used to illustrate various phenomena such as population distribution, population density, disease outbreak epicentres, weather and climate patterns, and migration trends.



## 4. **Pie charts**

- A pie chart helps organise and show data as fractions of a whole.
- It uses a circle made up of various segments, each representing various parts of the whole, where the area of each segment is proportional to the value of the data it represents.
- Essentially, each segment is a “slice” of the “pie”.
- A pie chart is distinctive and easy to understand.
- As a result, it is a favourite tool for visualising data such as product market share, production defects, time spent on performing various business tasks in a day, cost distributions, and

# Interactive Visualization

Interactive Weather Graph with different data and different visualization tools.

- <https://medium.com/@yuanxie1122/interactive-weather-graph-fca273061cdd>

•THANK YOU