



Data Analytics

Lecture No: 01

Instructor: Dr. Muhammad Abeer Irfan



Objective of the course

- The principal focus of this course is to introduce conceptual understanding using simple and practical examples rather than repetitive and point click mentality.
- This course should make you comfortable using analytics in your career and life.
- You will know how to work with real data, and might have learned many different methodologies but choosing the right methodology is important.



Objective of the course

- The danger in using quantitative method does not generally lie in the inability to perform the calculation.
- The real threat is lack of fundamental understanding of:
 - Why to use a particular technique or procedure.
 - How to use it correctly and,
 - How to correctly interpret the result.





Learning Objectives

1. Define data and its importance.
2. Define data analytics and its types.
3. Explain why analytics is important in today's business environment.
4. Explain how statistics, analytics and data science are interrelated.
5. Why Python?
6. Explain the four different levels of Data:
 - a. Nominal
 - b. Ordinal
 - c. Interval and
 - d. Ratio



Define data and its importance

- Variable, measurement and data.
- What is generating so much data?
- How data add value to the business?
- Why data is important?



Variable, Measurement and Data

- **Variable** – is a characteristic of any entity being studied that is capable of taking on different values.
- **Measurement** – is when a standard process is used to assign numbers to particular attributes or characteristic of a variable.
- **Data** – data are recorded measurements.

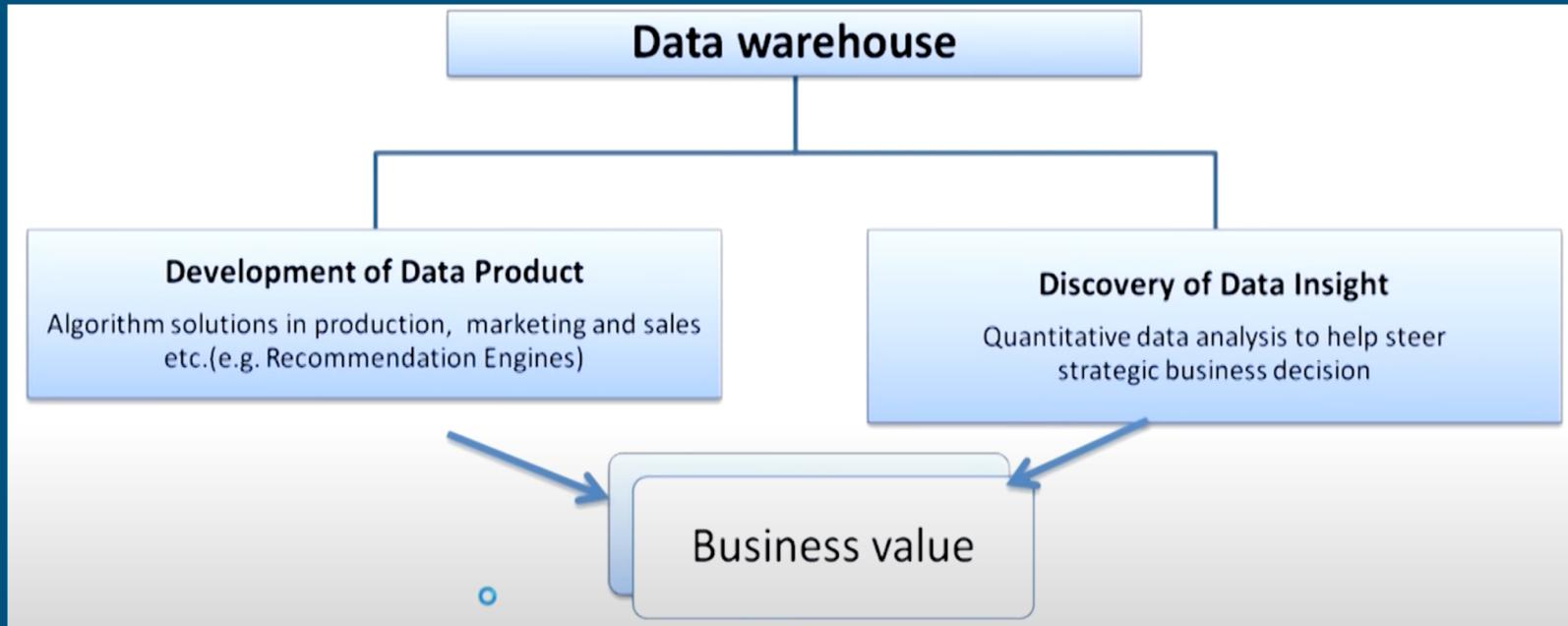


What is generating so much data?

- Data can be generated by:
 - Humans,
 - Machines or
 - Humans-machines combines.
- It can be generated anywhere, where any information is generated and stored in structured or unstructured formats.



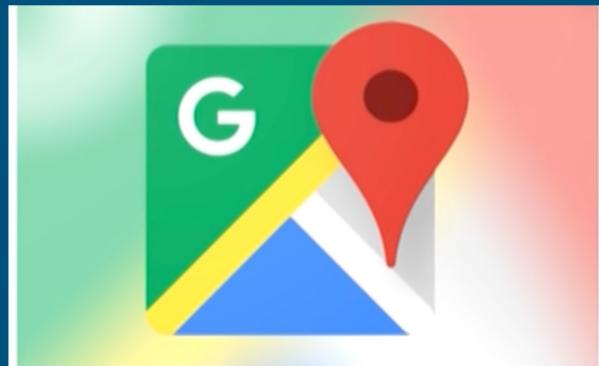
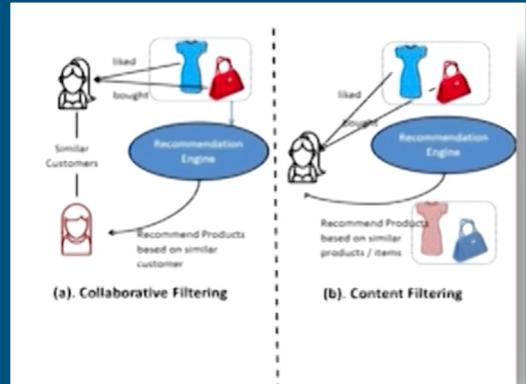
How data add value to business?



Source: <https://datajobs.com>



Data Products





Why data is important?

- Data helps in make better decisions?
- Data helps in solve problems by finding the reason for underperformance.
- Data helps one to evaluate the performance.
- Data helps one improve processes.
- Data helps one understand consumers and the market.



Con'td

- Define data analytics?
- Why analytics is important?
- Data analysis
- Data analytics vs Data analysis
- Types of data analytics



Define data analytics

- Analytics is defined as “ the scientific process of transforming data into insights for making better decisions”.
- Analytics, is the use of data, information technology, statistical analysis, quantitative methods, and mathematical or computer-based models to help managers gain improved insight about their business operations and make better, fact-based decisions (Professor James Evans).
- Analysis = Analytics?



Why analytics is important?

- Opportunity abound for the use of analytics and big data such as:
 - a. Determining credit risk
 - b. Developing new medicines
 - c. Finding more efficient ways to deliver products and services
 - d. Preventing fraud*
 - e. Uncovering cyber threats**
 - f. Retaining the most valuable customers!.

*<https://www.meadenmoore.com/blog/iaq/using-data-analytics-help-prevent-detect-fraud>

**<https://seleritysas.com/blog/2019/10/01/how-big-data-analytics-combats-cybercrime/#:~:text=Whether%20it%20>

!<https://www.arpatech.com/blog/how-big-data-and-business-intelligence-help-to-improve-customer-retention/>



Data analysis

- Data analysis is the process of **examining**, **transforming**, and **arranging** raw **data** in a specific way to generate useful information from it.
- Data analysis allows for the *evaluation* of data through **analytical** and **logical** reasoning to lead to some sort of outcome or conclusion in some context.
- Data analysis is a multi-faceted process that involves a number of steps, approaches, and diverse techniques.



Data analytics vs data analysis

Analysis

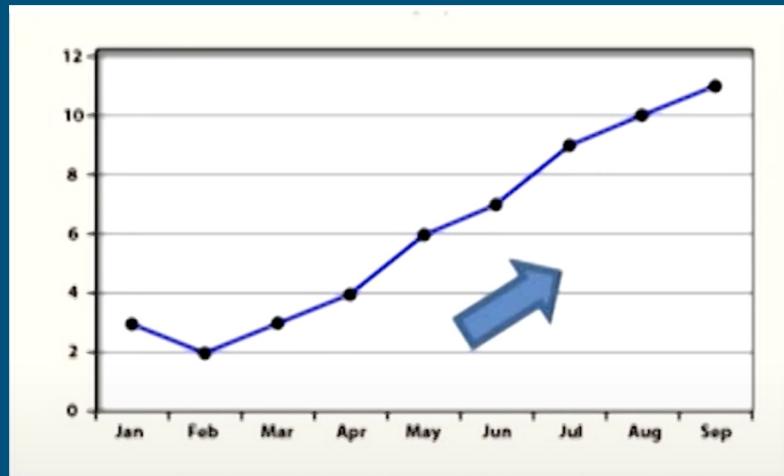


Past

Explain

How?

Why?





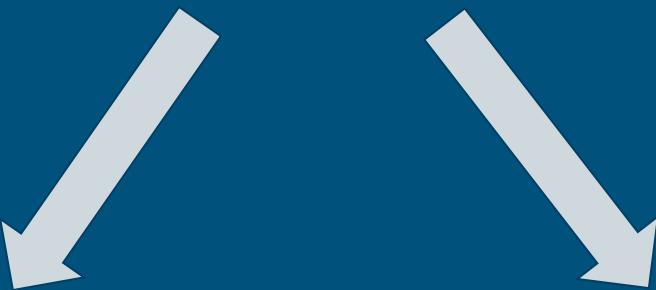
Cont'd





Cont'd

Analysis



Qualitative

Explains **how** and **why** story ends the way it did?

Quantitative

Data+ **how** the sale **decreased** last summer



Analysis	≠	Analytics
Data analysis	≠	data analytics
Business analysis	≠	Business analytics



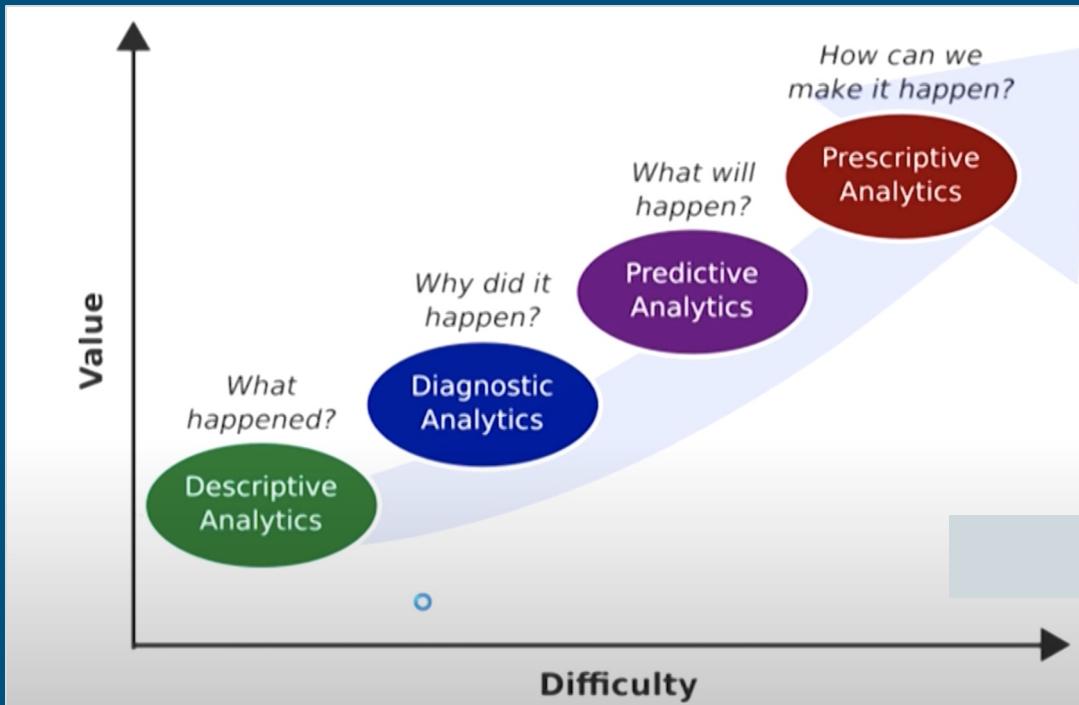
Classification of data analytics

Based on the phase of workflow and the kind of analysis required, there are four major types of data analytics.

- Descriptive analytics.
- Diagnostic analytics.
- Predictive analytics.
- Prescriptive analytics.



Classification of data analytics





Descriptive analytics

- Descriptive analytics, is the conventional form of Business intelligence and data analysis.
- It seeks to provide a depiction or “summary view” of facts and figures in an understandable format.
- This either inform or prepare data for further analysis.
- Descriptive analysis or statistics can summarize raw data and convert it into a form that can be easily understood by humans.
- They can describe in detail about an event that has occurred in the past.



Example

A common example of descriptive analytics are company reports that simply provide a historic review like:

- Data Queries.
- Reports.
- Descriptive statistics
- Data visualization
- Data dashboard.





Diagnostic analytics

- Diagnostic analytics is a form of advanced analytics which examines data or content to answer the question “ Why did it happen? ”.
- Diagnostic analytical tools aid an analyst to dig deeper into an issue so that they can arrive at the source of a problem.
- In a structured business environment, tools for both descriptive and diagnostic analytics go parallel.



Example

1. Data Discovery
2. Data mining
3. Correlations.

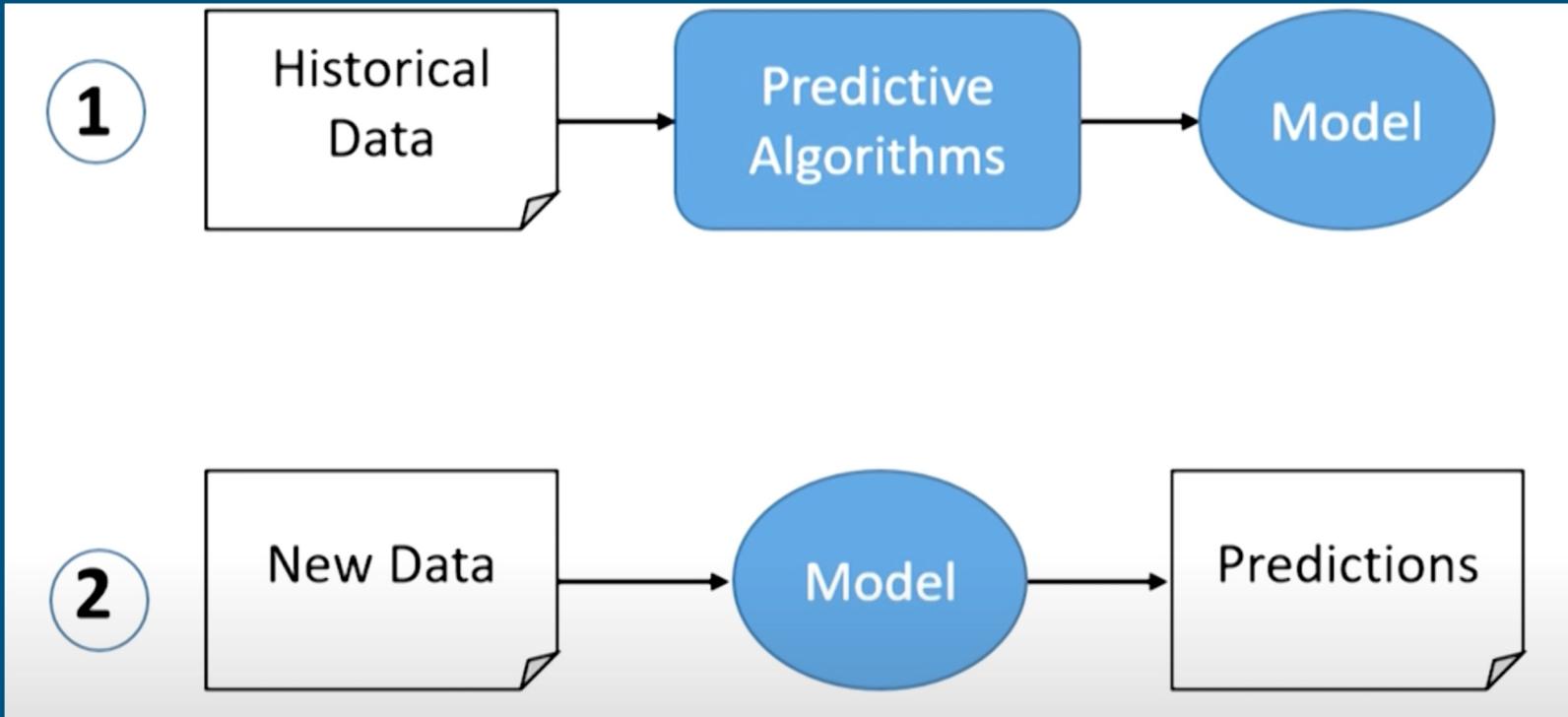


Predictive analytics

- Predictive analytics helps to forecast trends based on the current events.
- Predicting the probability of an event happening in future or estimating the accurate time it will happen can all be determined with the help of predictive analytical models.
- Many different but co-dependent variables are analysed to predict a trend in this type of analysis.



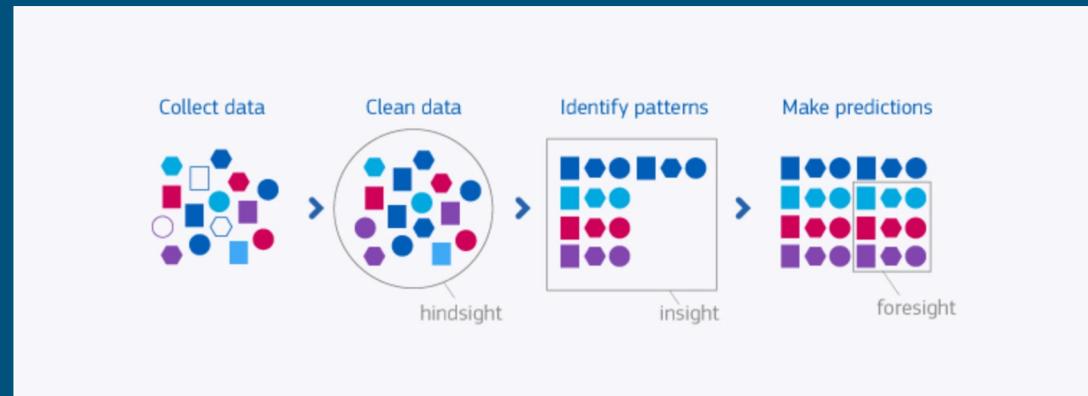
Cont'd





Example

- Set of techniques that use model constructed from past data to predict the future or ascertain impact of one variable on another:
 1. Linear regression
 2. Time series analysis and forecasting
 3. Data mining





Prescriptive analytics

- Set of techniques to indicate the best course of action.
- It tells what decision to make to optimize the outcome
- The goal of prescriptive analytics is to enable:
 - a. Quality improvements
 - b. Service enhancements
 - c. Cost reductions and,
 - d. Increasing productivity



Prescriptive analytics: Example

- Optimization Model
- Simulation
- Decision analysis



Explain why analytics is important

- Demand for data analytics.
- Element of data analytics



Demand for data analysts

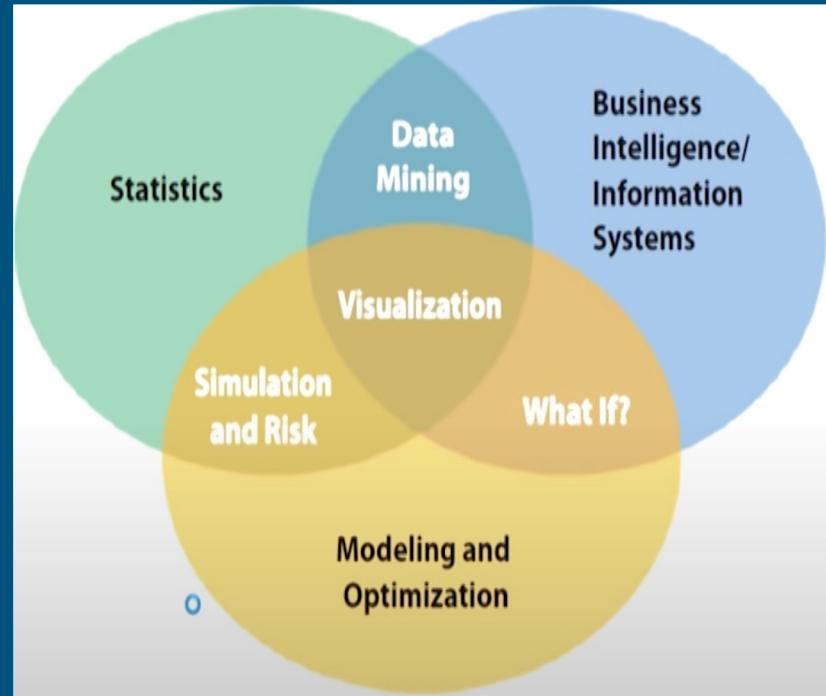
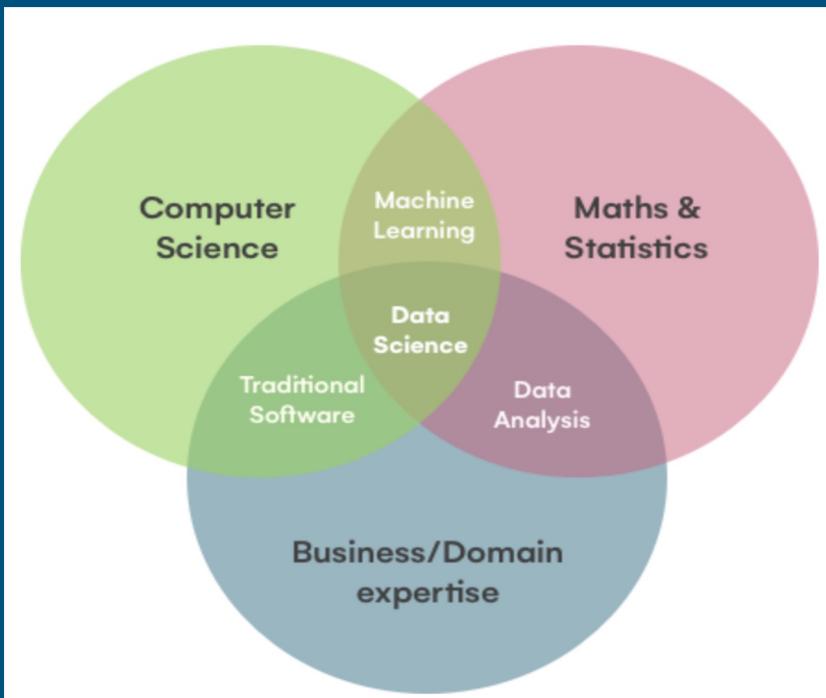
- Become a data analyst/scientist.
- Data marketplaces and exchanges
- Automated decision making with pinch of AI

The Data Analytics Profession And Employment Is Exploding —Three Trends That Matter

<https://www.forbes.com/sites/bernhardschroeder/2021/06/11/the-data-analytics-profession-and-employment-is-exploding-three-trends-that-matter/?sh=625a22493f81>



Element of data analytics



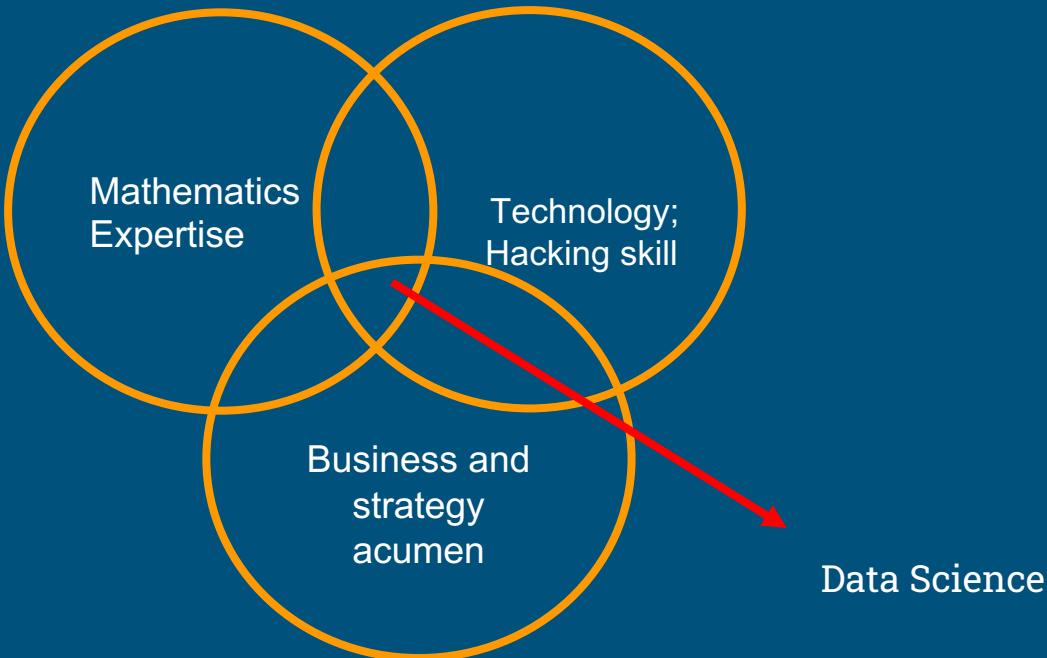


Data analyst and data scientist

- The requisite skill set
- Difference between data analyst and data scientist



The requisite skill set





Difference between data analytics and data scientist

Analyst



Business Administration

Domain specific responsibility : For Example marketing analyst, Financial analyst etc.

Data exploration analysis and insight

Advance algorithms and machine learning

Data product engineering

Data Scientist



Why Python?



Features

- Simple and easy to learn
- Freeware and Open source
- Interpreted
- Dynamically Typed
- Extensible
- Embedded
- Extensive library



Cont'd

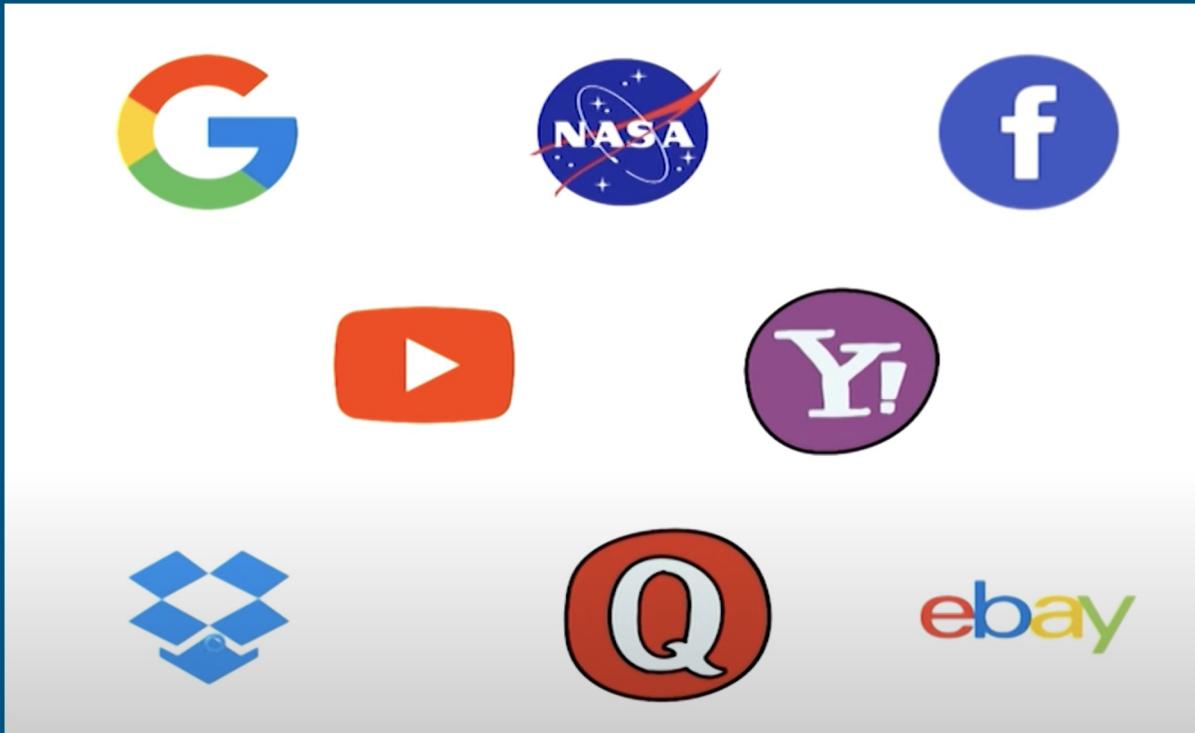


Usability

- Desktop and web applications
- Database applications
- Networking applications
- Data analysis (Data Science)
- Machine learning
- IoT and AI applications
- Games



Companies using Python



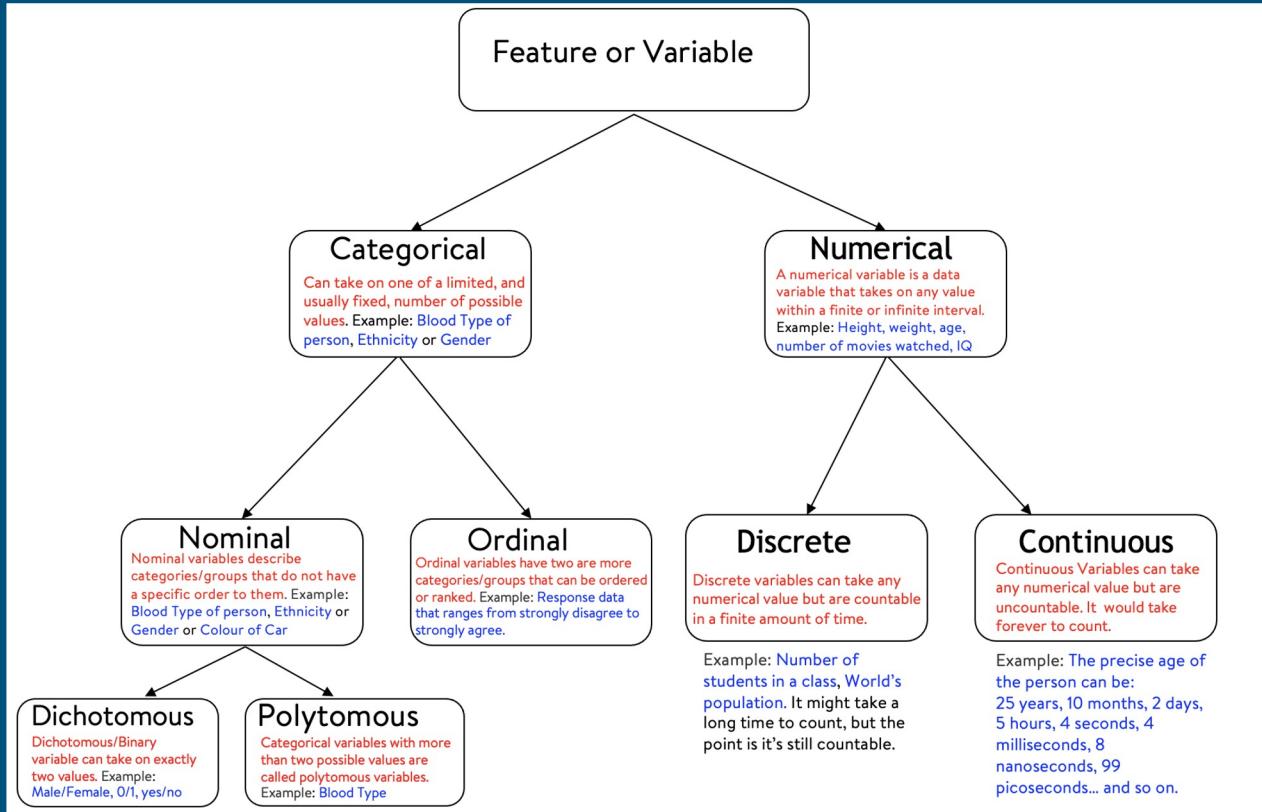


Levels of data

- Types of variables.
- Levels of data measurement
- Compare the four different levels of data:
 - Nominal
 - Ordinal
 - Interval and
 - Ratio
- Usage potential of various levels of data
- Data level, Operations, and statistical methods.



Types of variables





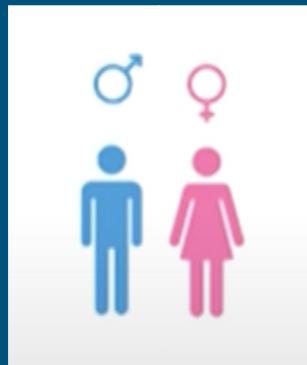
Levels of data measurement

- Nominal – Lowest level of measurement
- Ordinal
- Interval
- Ratio – Highest level of measurement



Nominal

- A **nominal scale** classifies data into distinct categories in which no ranking is implied.
- Example: Gender, marital status.





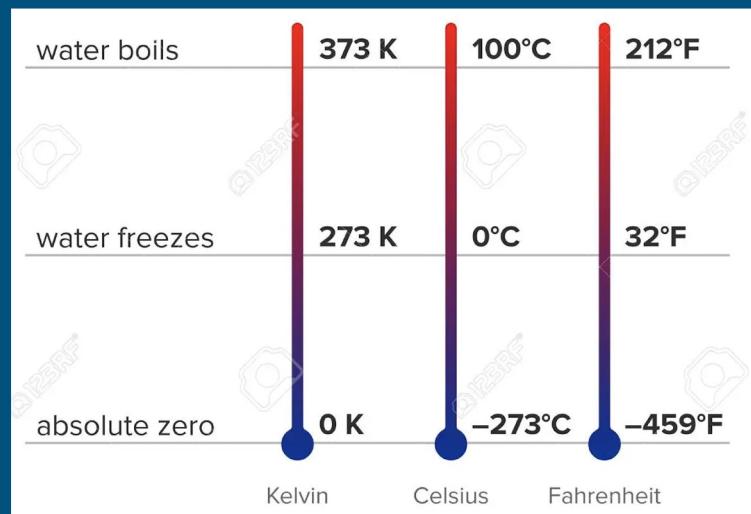
Ordinal Scale

- An **ordinal scale** classifies data into distinct categories in which ranking is implied.
- Example:
 - Product satisfaction → Satisfied, neutral, Unsatisfied.
 - Faculty rank → Professor, Associate professor, Assistant professor, Lecturer.
 - Student grades → A,A-,B+,B,B-,C+,C,C-,D+,D,F



Interval Scale

- An **interval scale** is an ordered scale in which the difference between measurements is a meaningful quantity but the measurements do not have a true zero point .
- Example
 - Temperature in Fahrenheit and celsius
 - Year.



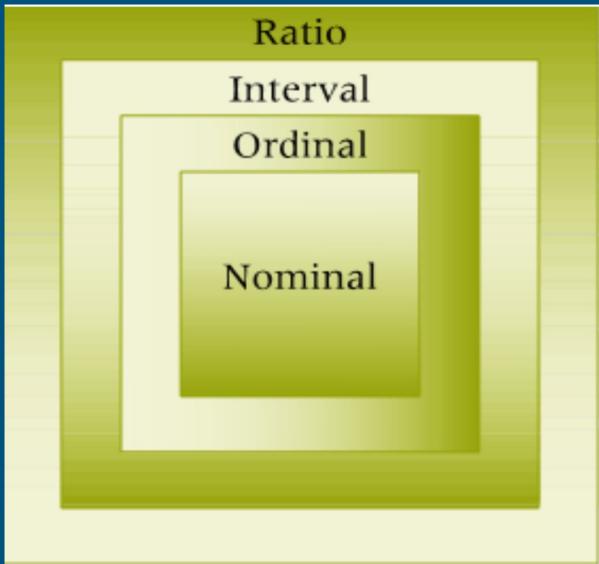


Ratio Scale

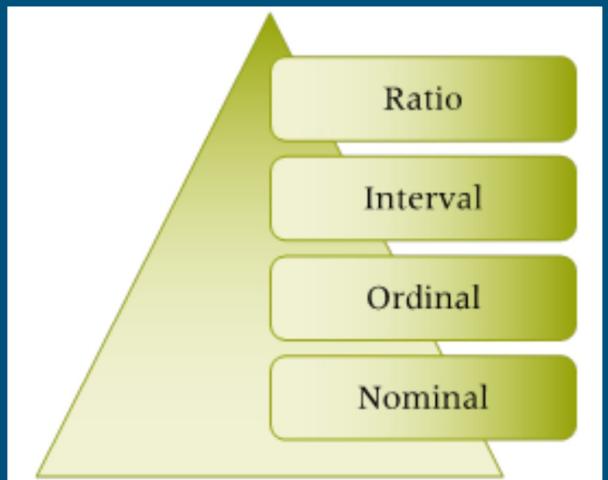
- A ratio scale is an ordered scale in which the difference between the measurements is a meaningful quantity and the measurements have a true zero point.
- Example:
 - Weight.
 - Height.
 - Age.
 - Salary.



Usage potential of various levels of data



Highest level of data measurement



Lowest level of data measurement



Impact of choice of measurement scale

Data Level	Meaningful Operations	Statistical Methods
Nominal	Classifying and counting	Non parametric
Ordinal	All of the above plus ranking	Non parametric
Interval	All of the above plus Addition, Subtraction	Parametric
Ratio	All the above plus Multiplication and division	Parametric