Parul® University

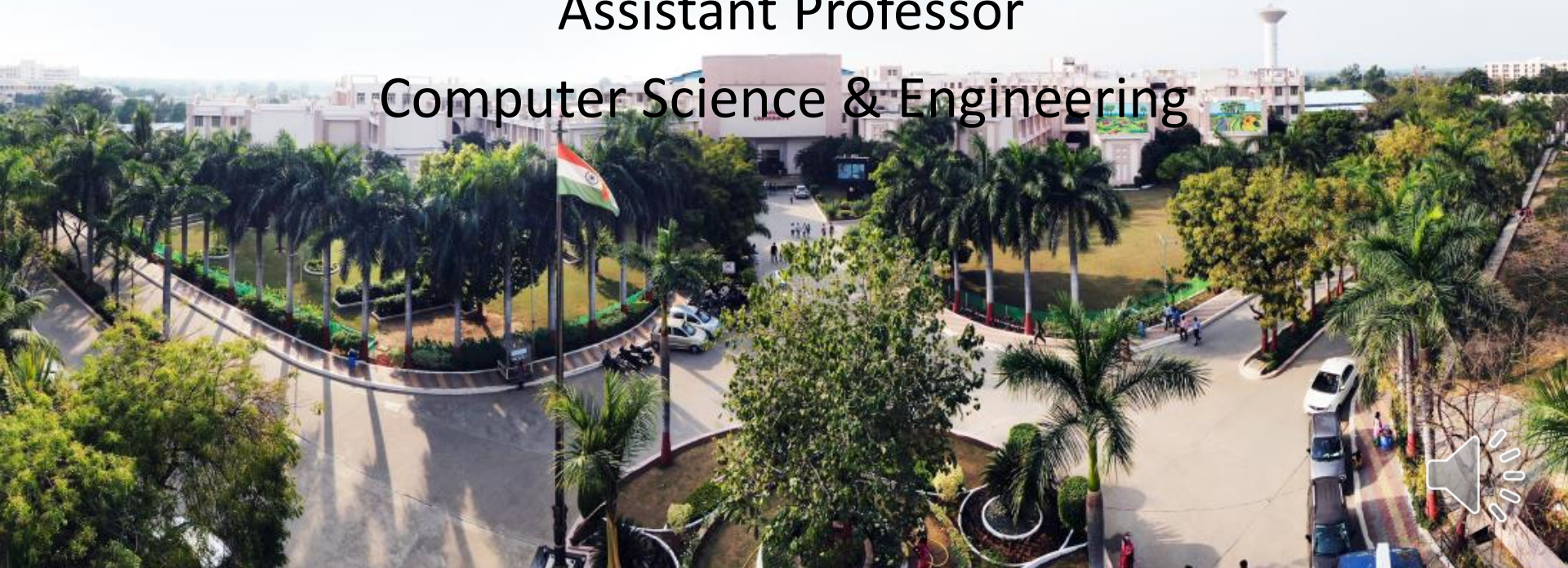
Data Analytics and Data Visualization (30310214)

Unit 1 : Introduction to Data Analytics

Prof. Manisha Chandramaully

Assistant Professor

Computer Science & Engineering



Parul®
University

DIGITAL LEARNING CONTENT

Outline

- Introduction
- Data and its importance
- Data analytics and its types
- Why data analytics is important
- Data analysis Vs. Data analytics
- Classification of data analytics
- Elements of Data analytics
- Data analyst Vs. Data scientist

Introduction

- This is to introduce conceptual understanding using simple and practical examples.
- Focus on:
  - Why to use a particular technique of procedure
  - how to choose the right technique
  - how to use it correctly
  - how to interpret the result.

Parul®
University

DIGITAL LEARNING CONTENT

Data and its importance

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

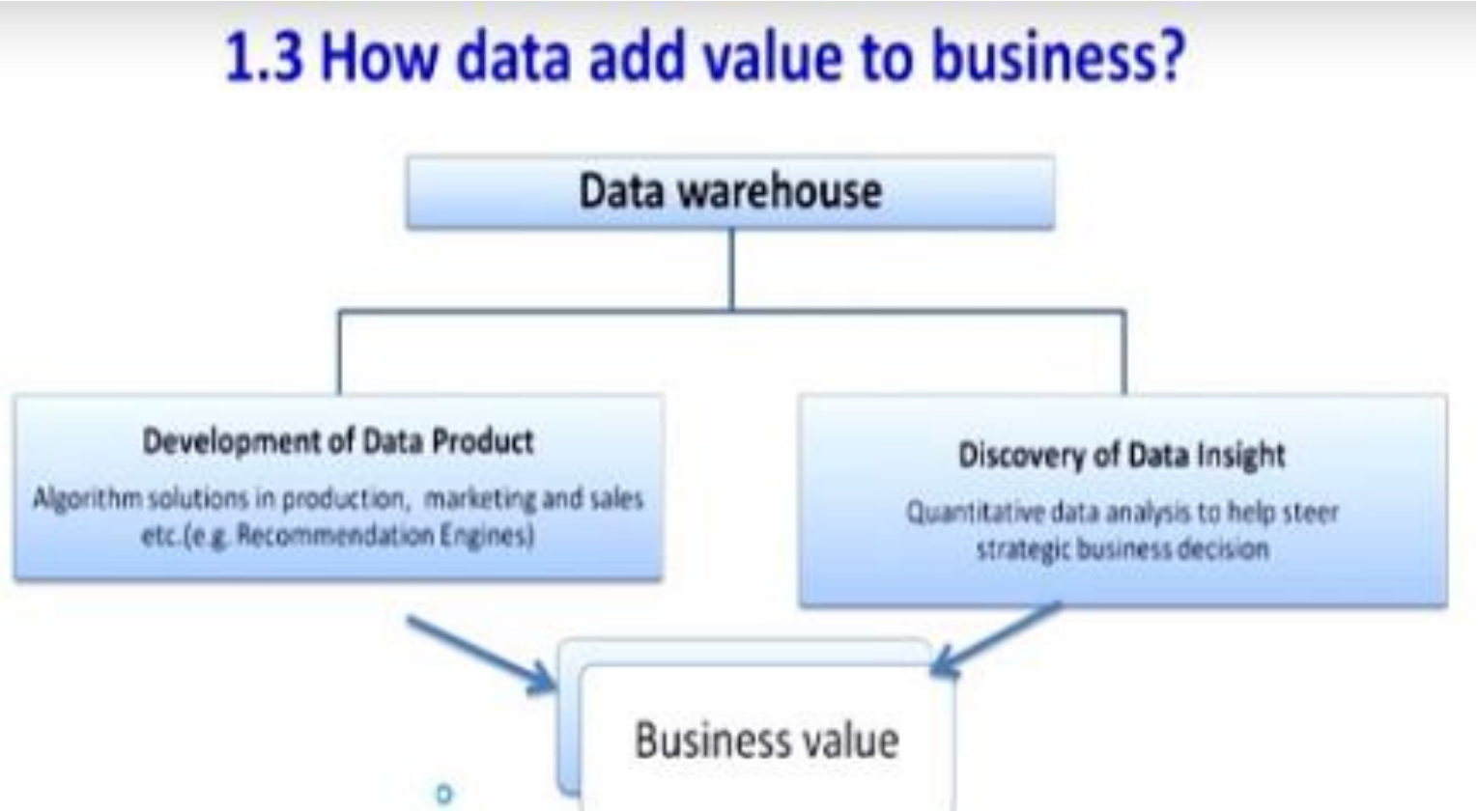
Variable, Measurement and Data

- Variable- is a characteristic of any entity being studied that is capable of taking on different values.
- Measurements- is when a standard process is used to assign numbers to particular attributes or characteristics 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?



Example of 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

Define Data analytics and its types

- Data analytics is defined as “the scientific process of transforming data into insights for making better decisions.”
- Professor James Evans has defined the data analytics that “it is the use of the 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.”

Why data analytics is important?

- Opportunity abounds for the use of analytics and big data such as:
  1. Determining credit risk
  2. Developing new medicines
  3. Finding more efficient ways to deliver products and services
  4. Preventing fraud
  5. Uncovering cyber threats
  6. Retaining the most valuable customers

Data analysis

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

Data analytics Vs. Data Analysis

Analysis

Explain

How?

Why?

Past

Data analysis it is something about -

- what has happened in the past.
- why that has happened?
- how it has happened?
- Example: Post mortem analysis

Data analytics Vs. Data Analysis

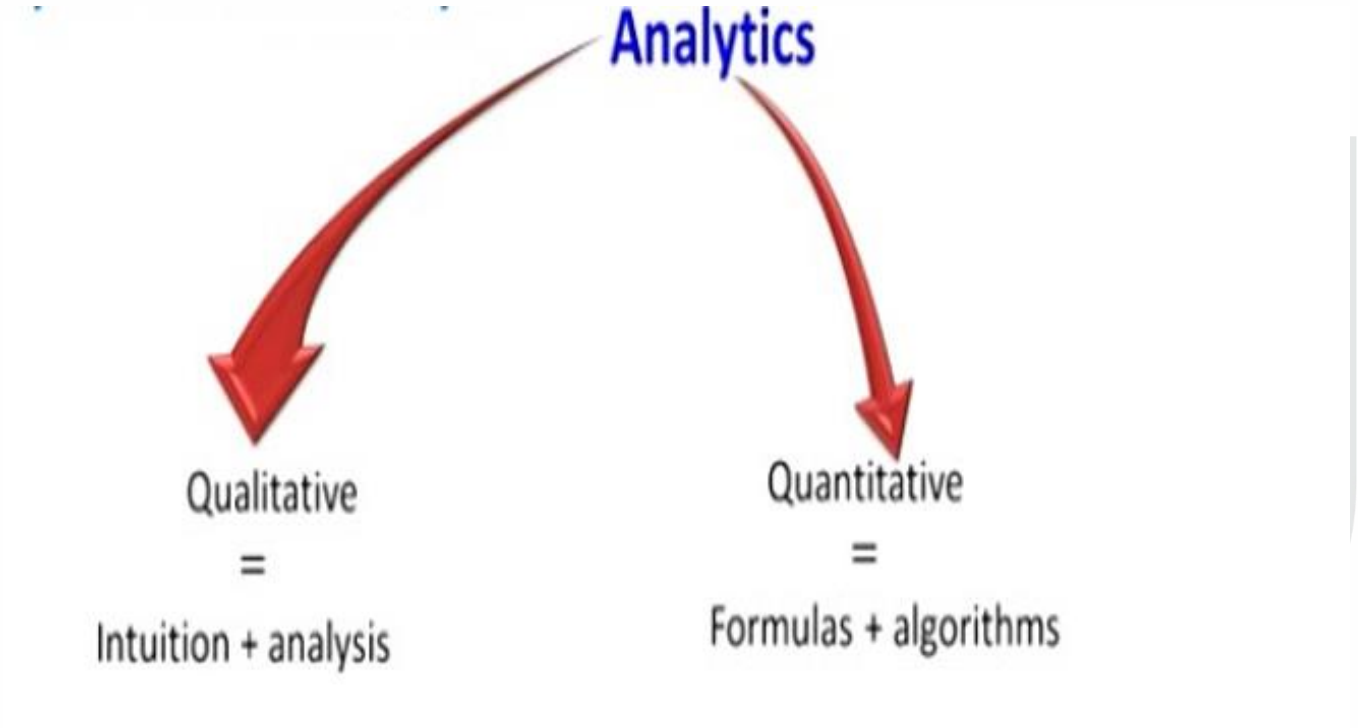
Analytics

Future

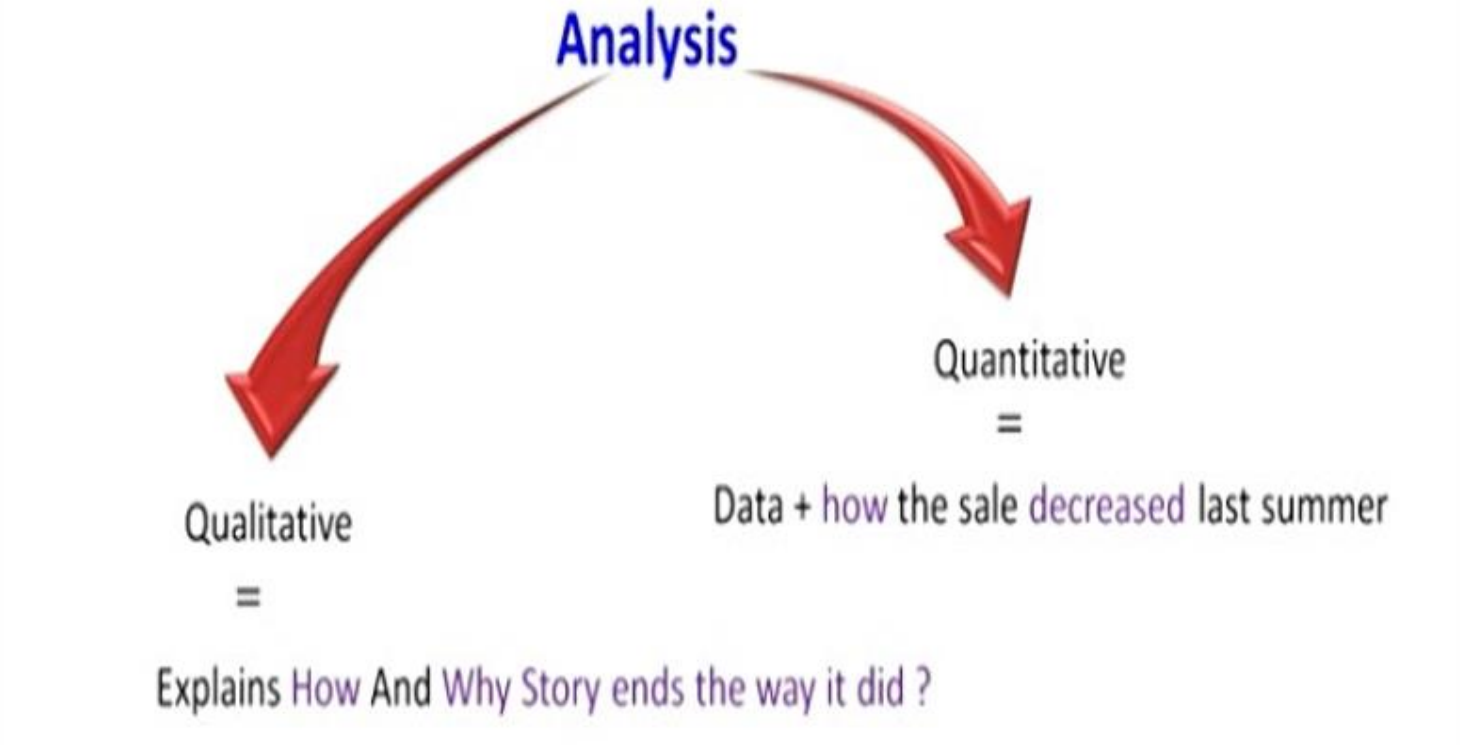
Analytics is studying about –

- what will happen in future.
- predict explore possible potential future events.

Data analytics Vs. Data Analysis



Data analytics Vs. Data Analysis

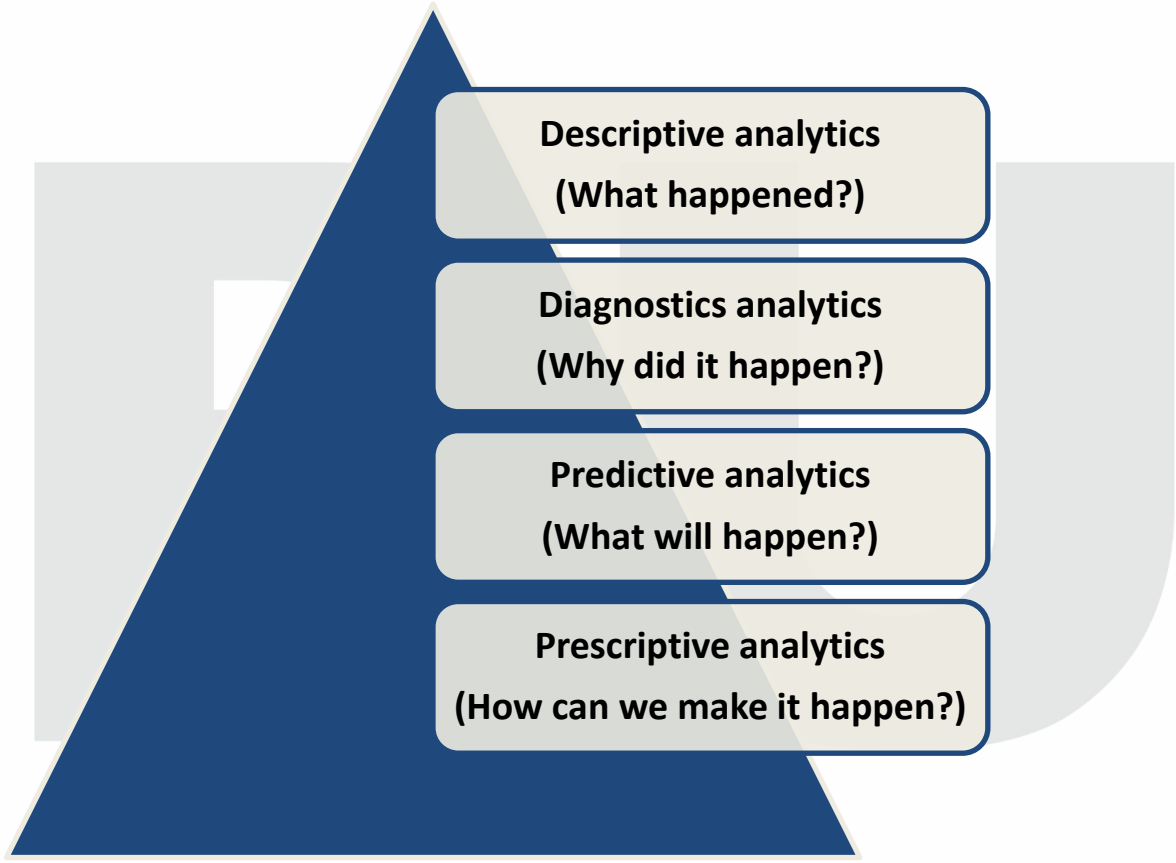


Classification of Data analytics

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

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

Classification of Data analytics



Classification of Data analytics: Descriptive analytics

- Descriptive analytics is the conventional form of business intelligence or data analysis.
- It seeks to provide the depiction or summary view of facts and figures in an understandable format.
- These either inform or prepare data for further analysis. so descriptive analysis or we can say another way in statistics can summarize raw data and convert it into your form that can be easily understood by humans.
- They can describe in detail about an even that has occurred in the past.

Classification of Data analytics: Descriptive analytics

- Descriptive analytics is the conventional form of business intelligence or data analysis.
- It seeks to provide the depiction or summary view of facts and figures in an understandable format.
- These either inform or prepare data for further analysis. so descriptive analysis or we can say another way in statistics can summarize raw data and convert it into your form that can be easily understood by humans.
- They can describe in detail about an even that has occurred in the past.
- Example: company reports that simply provide the historic review like: data queries, reports, descriptive statistics, data visualization and data dashboard.

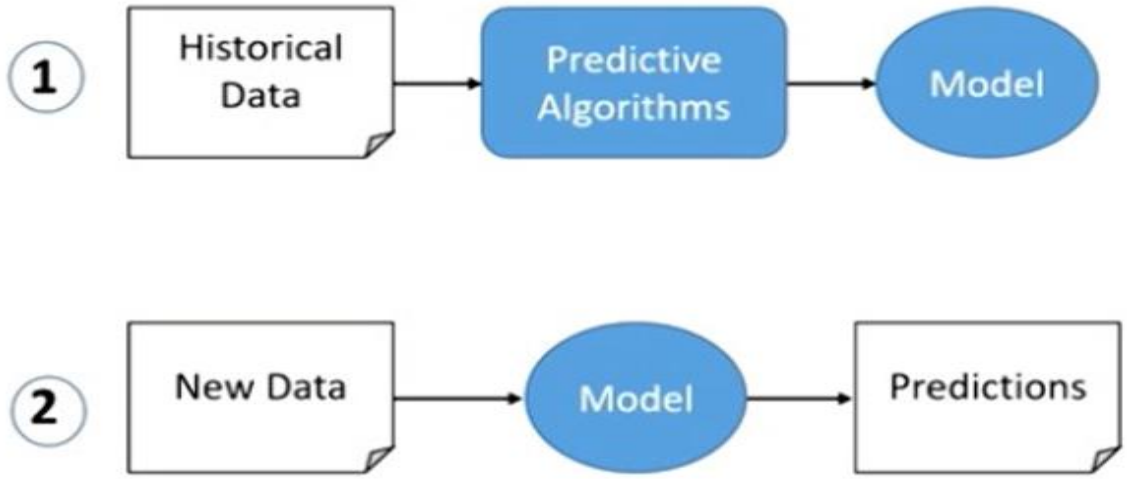
Classification of Data analytics: Diagnostic analytics

- 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 the problem.
- In a structured business environment tools for both descriptive and diagnostic analytics go parallel.
- Example: data discovery, data mining, and correlations.

Classification of Data analytics: Predictive analytics

- Predictive analytics helps to forecast trends based on the current events.
- Predicting the probability of an event happening in future are estimating 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.

Classification of Data analytics: Predictive analytics



- With the help of historical data by using different algorithm, predictive algorithms you can come with a model.
- Once the model is developed a new data can be fit into this model so we can get some predictions about the past events.

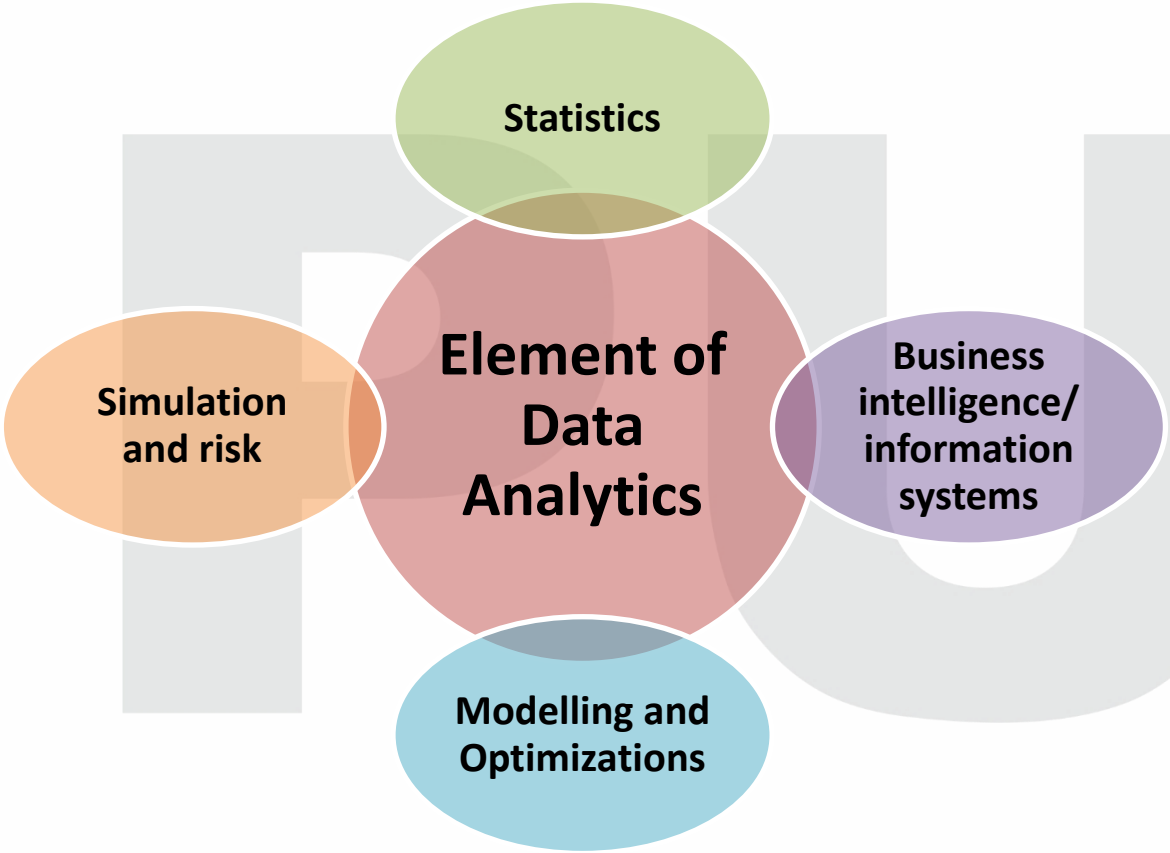
Classification of Data analytics: Predictive analytics

- Example: linear regression, time series analysis and forecasting and data mining.
- These are the techniques for predictive analytics.

Classification of Data analytics: Prescriptive analytics

- A 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: quality improvements, service enhancements, cost reductions and increasing productivity.
- Examples: Some of the tools like optimization models, simulation model, and decision analysis.

Element of Data analytics



Parul® University

Data analyst Vs. Data scientist

Aspect	Data Analyst	Data Scientist
Primary Focus	Analyzing data to provide insights for business decisions.	Using advanced statistical and computational methods to solve complex problems.
Skills Required	– Data Cleaning and Preparation – Statistical Analysis – Data Visualization – SQL – Excel Skills – Problem-Solving – Domain Knowledge – Communication Skills – Attention to Detail – Time Management – Continuous Learning	– Machine Learning – Statistical Analysis – Data Cleaning and Preprocessing – Data Visualization – Big Data Technologies (e.g., Hadoop, Spark) – Deep Learning – Natural Language Processing (NLP) – SQL Database Management – Experiment Design and A/B Testing – Cloud Computing Platforms (e.g., AWS, Azure, Google Cloud) – Communication and Presentation Skills – Domain Knowledge – Time Series Analysis – Feature Engineering

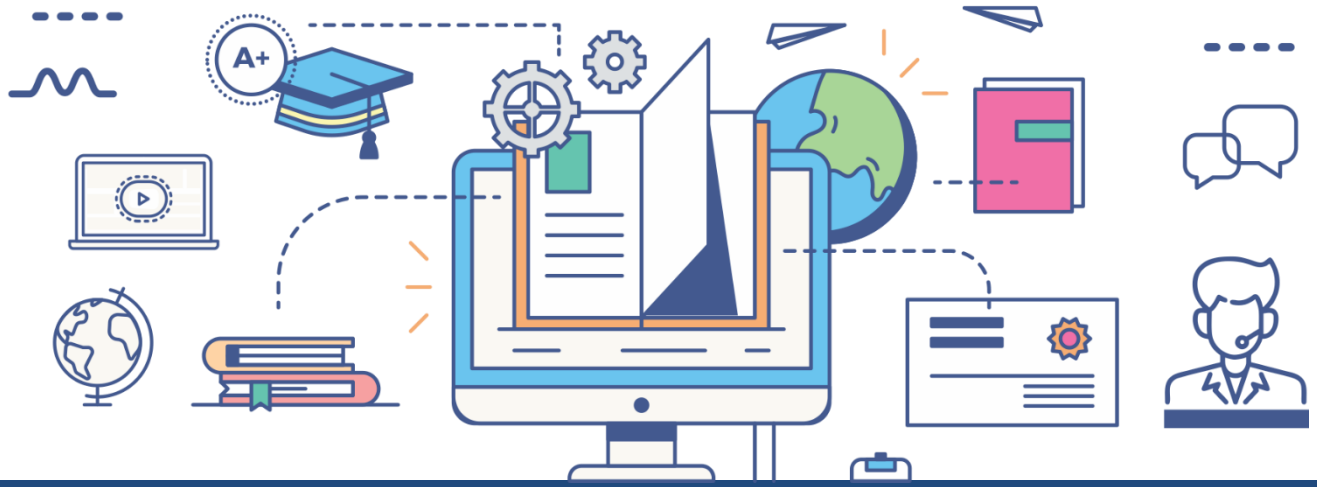
Parul® University

DIGITAL LEARNING CONTENT

Data analyst Vs. Data scientist

Aspect	Data Analyst	Data Scientist
Typical Tasks	Cleaning and organizing data, creating reports, generating dashboards, and performing descriptive analytics.	Building predictive models, conducting A/B testing, developing algorithms, and performing exploratory data analysis.
Example Scenario	Analyzing sales data to identify trends and optimize marketing strategies.	Developing a recommendation system for an e-commerce platform based on customer behavior.
Decision Making	Helps businesses make data-driven decisions by providing insights from existing data.	Involves both providing insights and developing solutions to complex problems using data.
Tools Used	Excel, SQL, Tableau, Power BI, Google Analytics.	Python, R, SQL, TensorFlow, PyTorch, Jupyter Notebooks, Big Data Technologies (e.g., Hadoop, Spark).

DIGITAL LEARNING CONTENT



Parul® University

www.paruluniversity.ac.in