

# MODULE 1

Data: A collection of facts

Data Analysis: The collection, transformation, and organization of data in order to draw conclusions, make predictions, and drive informed decision-making.

Data Analyst: Someone who collects, transforms, and organizes data in order to help make informed decisions

Different processes for Data Analysis: ASK → PREPARE → PROCESS  
ACT ← SHARE ← ANALYZE ←



"Data! Data! Data! ... I can't make bricks without clay!"

— Sherlock Holmes via Sir Arthur Conan Doyle

Data Analytics: The science of data

People Analytics: Aka human resources

analytics or workforce analytics

It is the practice of collecting and analyzing data on the people who make up a company's workforce in order to gain insights to improve how the company operates.

- Video vignettes
- Data journal
- Readings
- Activities
- Discussion prompts

PROGRAM FEATURES

REAL-LIFE EXAMPLE: People analysts answer "How can the organisation improve the retention rate for new employees?"

following are the steps they followed:-

- ① ASK: asked effective questions to define what the project would look like and what qualifies as a successful result.
- ② PREPARE: building a timeline, establishing rules to maintain data security, identification of what data is needed to achieve required result, brainstorming possible project- and data-related issues and how to avoid them.
- ③ PROCESS: made sure employees understood how their data is collected, stored, managed, and protected — ensuring ethical data collection, restricted access to data, raw data uploaded to internal data warehouse.
- ④ ANALYZE: analysed, identified key indicator (employee's experience with processes), proper documentation of findings
- ⑤ SHARE: careful sharing of report with managers providing them with an opportunity to communicate with their team
- ⑥ ACT: implement changes and take actions based on findings

## ARTICLE

Additional Resource

[https://  
online.hbs.edu/  
blog/post/  
business-  
analytics-  
examples](https://online.hbs.edu/blog/post/business-analytics-examples)

## BUSINESS ANALYTICS

- use of math and statistics to collect, analyze and interpret data to make better business decisions
- 4 key types:
  - WHAT HAPPENED? 1) Descriptive Analytics (interpretation of historical data to identify trends)
  - WHAT MIGHT HAPPEN IN FUTURE? 2) Predictive Analytics (used to forecast future outcomes)
  - WHY DID THIS HAPPEN? 3) Diagnostic Analytics (to identify root cause of problem)
  - WHAT SHOULD WE DO NEXT? 4) Prescriptive Analytics (testing techniques to determine best process)

## Business Analytics



DESCRIPTIVE



DIAGNOSTIC



PREDICTIVE



PREScriptive

## Diagnostic Analytics



→ a testing

Aspect	Business Analytics	Data Science
Main Goal	Extract insights to guide business decisions	Turn raw data into conclusions using algorithms and models
Focus Area	Decision-making and strategy	Data-driven discovery and predictive modeling
Typical Tasks	Budgeting, forecasting, product development	Data wrangling, programming, statistical modeling
Tools Used	Excel, Tableau, Power BI, SQL	Python, R, Jupyter, TensorFlow, Spark
Approach	Descriptive and diagnostic analytics	Predictive and prescriptive analytics
Background Needed	Business, economics, management	Computer science, statistics, mathematics
Data Handling	Works with structured business data	Handles structured, semi-structured, and unstructured data
Output	Business reports, dashboards, performance metrics	Machine learning models, data products, advanced analytics

④ 3 disciplines encompassing data science:

- Statistics (<sup>imp</sup> decisions under uncertainty)
- Machine learning (automation: decisions under a lot of uncertainty)
- Analytics (unknown no. of decisions in an unknown situation)

1. **Ask:** business challenge, objective, or question
2. **Prepare:** data generation, collection, storage, and data management
3. **Process:** data cleaning and data integrity
4. **Analyze:** data exploration, visualization, and analysis
5. **Share:** communicating and interpreting results
6. **Act:** putting insights to work to solve the problem

} fundamental data analytic process

## DATA ANALYSIS PROCESSES

1. Discovery
2. Pre-processing data
3. Model planning
4. Model building
5. Communicate results
6. Operationalize



1. Identifying the problem
2. Designing data requirements
3. Pre-processing data
4. Performing data analysis
5. Visualizing data

1. Ask
2. Prepare
3. Explore
4. Model
5. Implement
6. Act
7. Evaluate

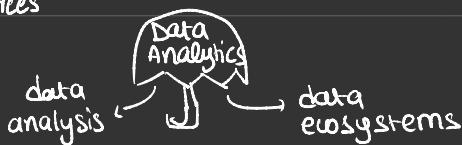
## DATA ECOSYSTEM

The various **elements** that interact with one another in order to produce, manage, store, organise, analyse and share data  
Applications: retail stores, HR departments, agricultural companies, etc.

## DATA ANALYST VS DATA SCIENTIST

finds answers to existing questions by creating insights from data sources

creates new questions using data



# Data Science

Data science is creating new ways of modeling and understanding the unknown by using raw data

## Subject matter Experts

have ability to look at results of data analysis to identify inconsistencies, make sense of gray areas & eventually validate choices made

people familiar with the business problem

## ANALYTICAL SKILLS

qualities & characteristics associated with solving problems using facts

① Curiosity

② Understanding context

③ Having a technical mindset

breaking things down into smaller steps

④ Data Design  
how info is organized

⑤ Data Strategy  
management of people, processes, & tools used in data analysis

ANALYTICAL THINKING  
identifying and defining a problem & then solving it by using data in an organized, step-by-step manner

## Key Aspects

① Visualisation  
graphic representation of information

② Strategy

③ Problem-orientation

⑤ Big-picture & detail-oriented thinking

④ Correlation

"correlation DOES NOT equal causation"

"FIVE WHYS" process to identify root cause of a problem

Common Data Analyst Questions:

① What is the root cause of a problem?

② Where are the gaps in our process?

③ What did we not consider before?

"Gap Analysis" is a method for examining & evaluating how a process works currently in order to get where you want to be in the future

## Quartiles

→ dividing data points into 4 equal parts or quarters