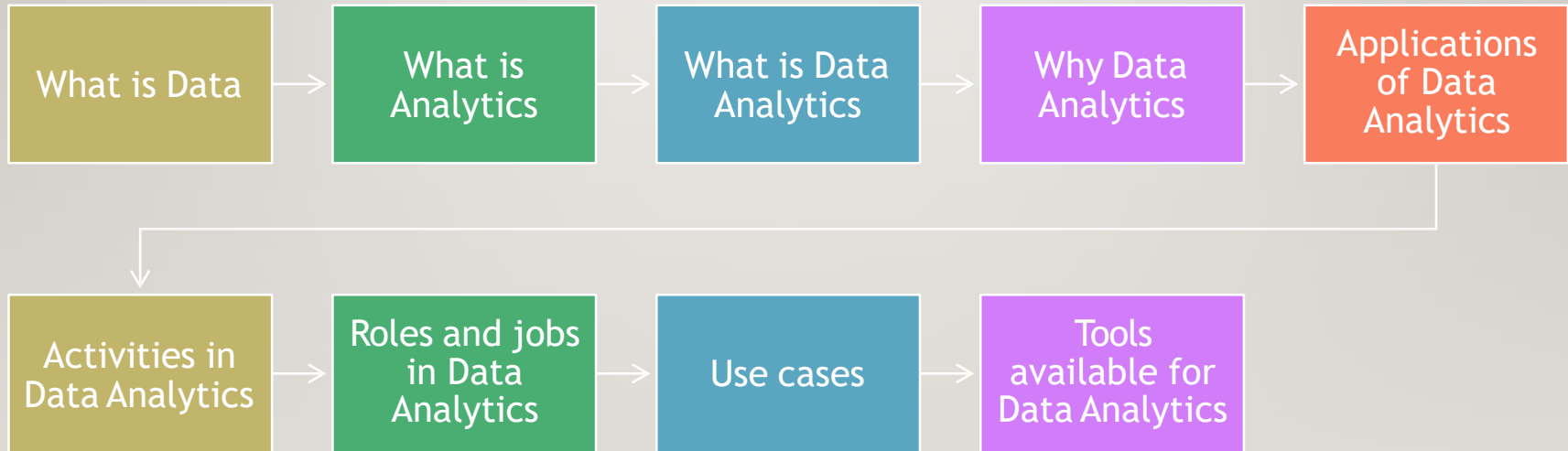


INTRODUCTION TO DATA ANALYTICS



WHAT IS DATA ?

- ✓ Data is simply facts or figures — bits of information.
- ✓ When data is organized, processed, interpreted, structured to make it meaningful or useful, it is called complete information.
- ✓ Data can be images, sounds, videos etc.
- ✓ Data is Measured, Reported, and Analyzed
- ✓ Can be structured or unstructured
- ✓ Can be discrete or continuous

WHAT IS ANALYTICS ?

It is the method of understanding or identifying or interpreting the hidden pattern in the data for effective decision making.



What is Data Analytics ?

It is the process of inspecting, cleaning, transforming and modelling data by using statistics



TYPES OF ANALYTICS

- ✓ Descriptive Analytics
- ✓ Prescriptive Analytics
- ✓ Predictive Analytics

DESCRIPTIVE ANALYTICS

- ✓ We use Descriptive statistics
- ✓ Identifying what has happened
- ✓ Example: What is the average test score of the class

PRESCRIPTIVE ANALYTICS

- ✓ We use inferential statistics
- ✓ Interpreting what should we do
- ✓ Example: Whether to reject or accept a batch of sops manufactured by a machine

PREDICTIVE ANALYTICS

- ✓ We use predictive statistics
- ✓ Predicting what could happen
- ✓ Example: Predicting the sales of Walmart

WHY DATA ANALYTICS ?

- ✓ Effective decision making
- ✓ Enhanced customer service
- ✓ Improves Business/sales
- ✓ Identify anomalies



APPLICATIONS OF DATA ANALYTICS

Self Driving Cars

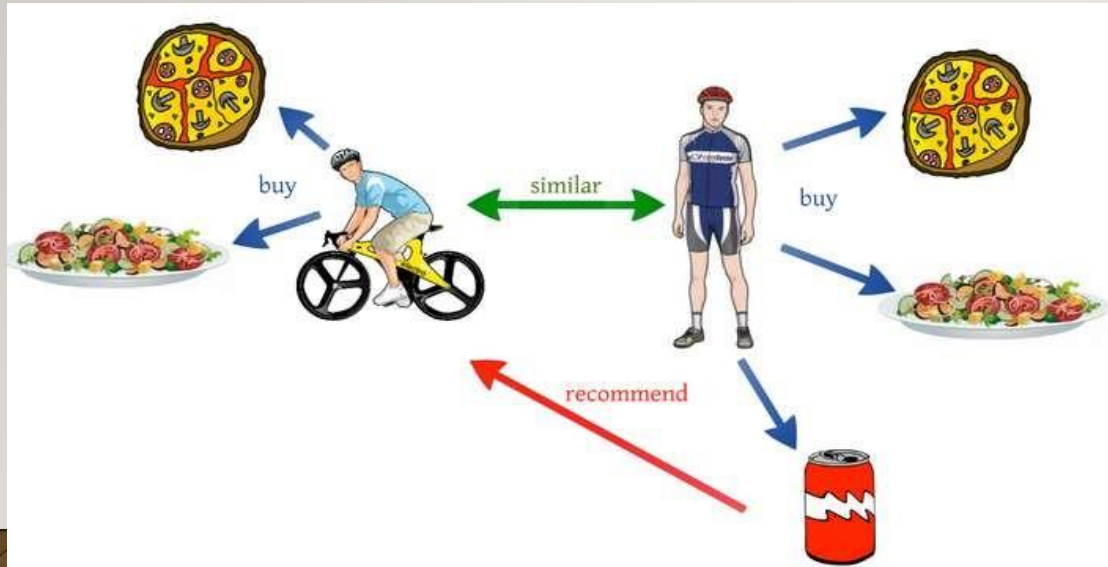


Google Maps

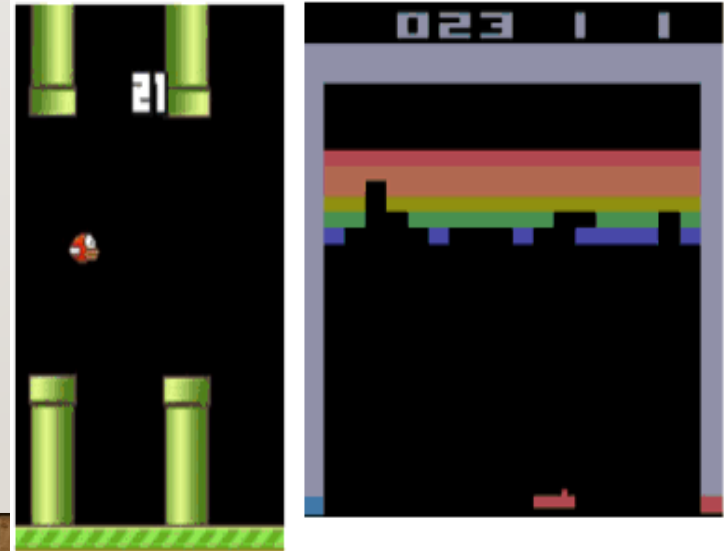


APPLICATIONS OF DATA ANALYTICS

Recommendation Engine



Games



Applications of Data Analytics

Decision Making



ACTIVITIES IN DATA ANALYTICS

- ✓ Data Extraction – Fetch data from database
- ✓ Data Analysis – Observe the data
- ✓ Data Manipulation – Manipulate the data
- ✓ Data Modelling – Fit a model to the data
- ✓ Data Visualization – Visualize the data

CAREER IN DATA ANALYTICS

- ✓ Data Engineer
- ✓ Data Analyst
- ✓ Data Modeler
- ✓ Data Scientist
- ✓ Full stack developer

CAREER IN DATA ANALYTICS

DATA ENGINEER

SOFTWARE ENGINEERS BY TRADE

Role

Develops, constructs, tests and maintains architectures (such as databases and large scale processing systems)

Languages

SQL, Hive, Pig, R, Matlab, SAS, SPSS, Python, Java, Ruby, C++, Perl

Mindset

All-purpose everyman

Skills & Talents

- Database systems (SQL & NO SQL based)
- Data modeling & ETL tools
- Data APIs
- Data warehousing solutions



CAREER IN DATA ANALYTICS

DATA ANALYST

DATA DETECTIVE

Role

Collects, processes and performs statistical data analyses

Languages

R, Python, HTML, Javascript, C/C++, SQL

Mindset


Intuitive data junkie with high "figure-it-out" quotient

Skills & Talents

- Spreadsheet tools (e.g. Excel)
- Database systems (SQL and NO SQL based)
- Communication & visualization
- Math, Stats, Machine Learning



CAREER IN DATA ANALYTICS

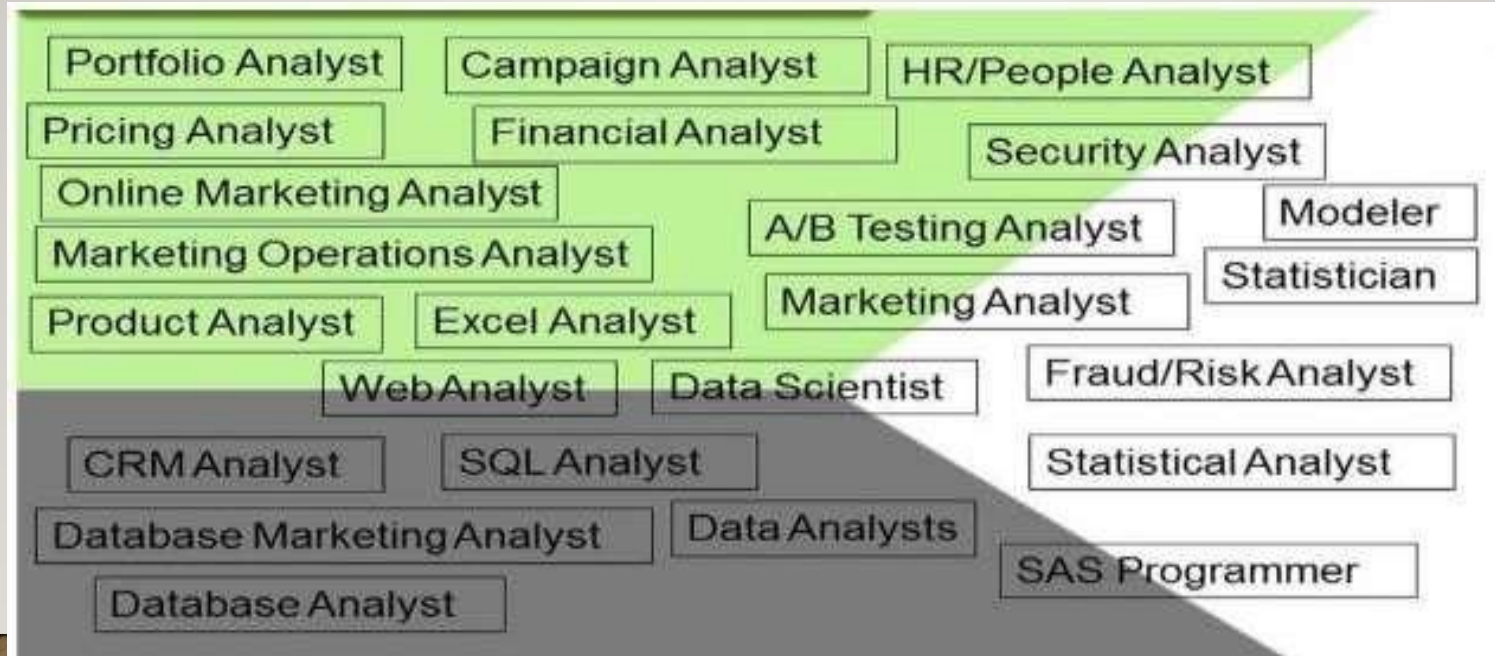


DATA SCIENTIST

AS RARE AS UNICORNS

Role	Mindset
Cleans, massages and organizes (big) data	Curious data wizard
Languages	Skills & Talents
R, SAS, Python, Matlab, SQL, Hive, Pig, Spark	<ul style="list-style-type: none">• Distributed computing• Predictive modeling• Story-telling and visualizing• Math, Stats, Machine Learning

ROLES IN DATA ANALYTICS



USE CASE



USE CASE - YOUTUBE



- ✓ YouTube uses recommendation systems to bring videos to a user.
- ✓ Increases the time user spends in the site
- ✓ Makes money by advertising
- ✓ Entertains the user
- ✓ Ultimately became popular

USE CASE - BMW

- ✓ Sensors are installed in the car to monitor the condition of the vehicle
- ✓ BMW gets the data continuously, analyses it to understand the condition of the car
- ✓ Customer is contacted when the car needs service



USE CASE - OTHER

- ✓ Identifying fraudulent banking transactions
- ✓ Determining a defaulter before issuing a credit card
- ✓ Advertising a brand on T-Shirts – FIFA
- ✓ Customer churn - Telecommunication

TOOLS AVAILABLE IN DATA ANALYTICS

– DATA EXTRACTION

- ✓ Data Stage
- ✓ Informatica
- ✓ R
- ✓ Python
- ✓ SAS

TOOLS AVAILABLE IN DATA ANALYTICS

– DATA ANALYSIS & MODELLING

- ✓ Python
- ✓ R
- ✓ SAS
- ✓ SPSS
- ✓ IBM WATSON
- ✓ MATLAB
- ✓ Statistica
- ✓ WEKA
- ✓ MS Excel

TOOLS AVAILABLE IN DATA ANALYTICS

– DATA VISUALIZATION

- ✓ COGNOS
- ✓ Qlikview
- ✓ Tableau
- ✓ Excel
- ✓ R
- ✓ Python
- ✓ Duen
- ✓ Domo
- ✓ BIRT
- ✓ JaperSoft
- ✓ SpagoBI
- ✓ Pentaho