Mr.Gangadhar Immadi,

Immadi.gangadhar@gmail.com

9986789040

Agenda

- Introduction to Data Science

- Need for Data Science

- Data Products

- Machine Learning and its Classification

Background and Motivation

➤ Big data and its Challenges

➤ What is Big Data ? – Volume, Velocity and Variety, Veracity

➤ Benefits of Bigdata and Data Science

➤ Availabity of large datasets

– http://ml-india.org/datasets/

– https://www.kaggle.com/sudalairajkumar/covid19-in-india

– https://www.isical.ac.in/~library/data.php

➤ Availabity of Modern Tools – Python, R, SaS,SPSS, Matlab etc
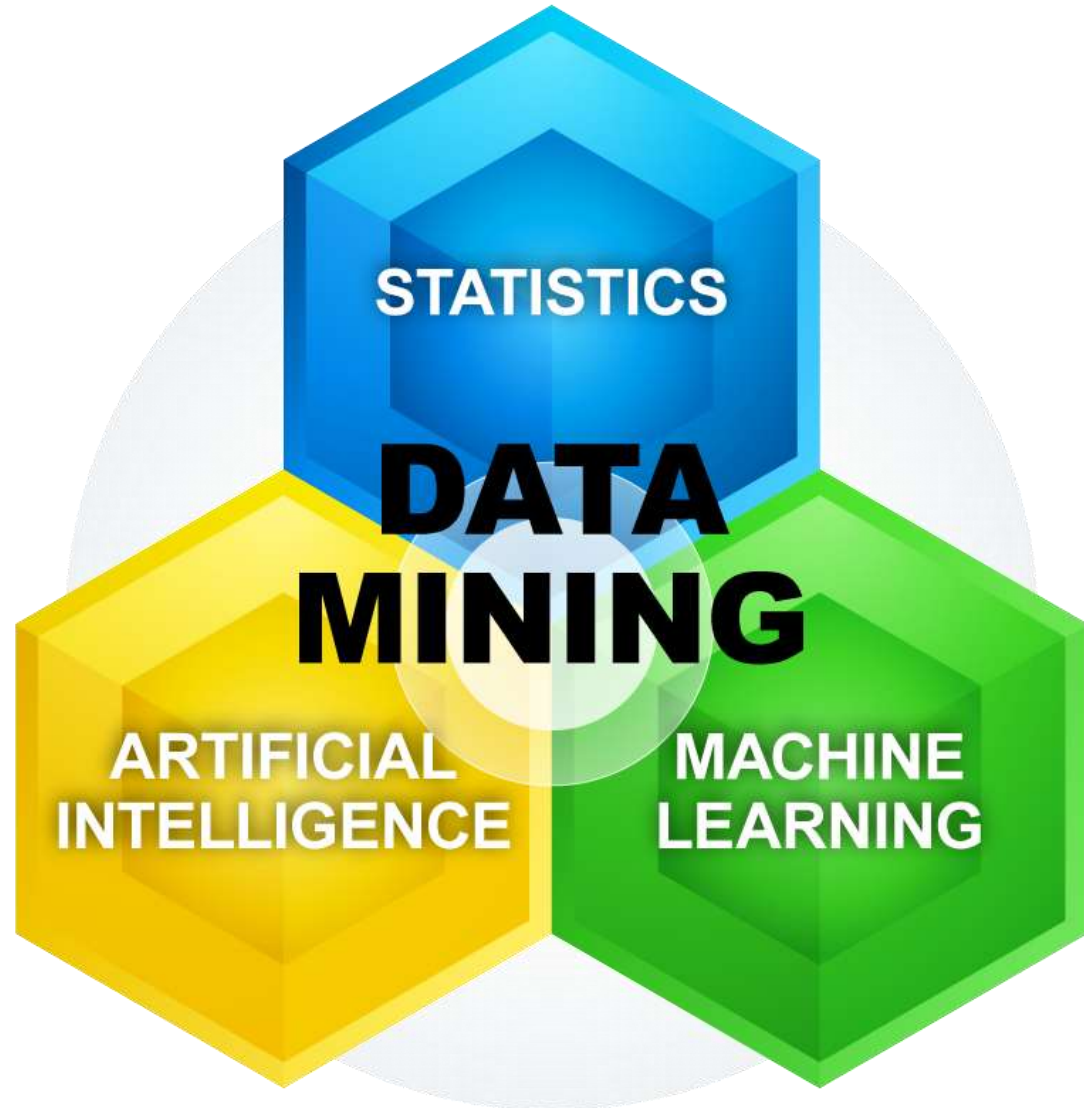
➤ Data driven decission

Data Science

Data science is an inter-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from many structural and unstructured data
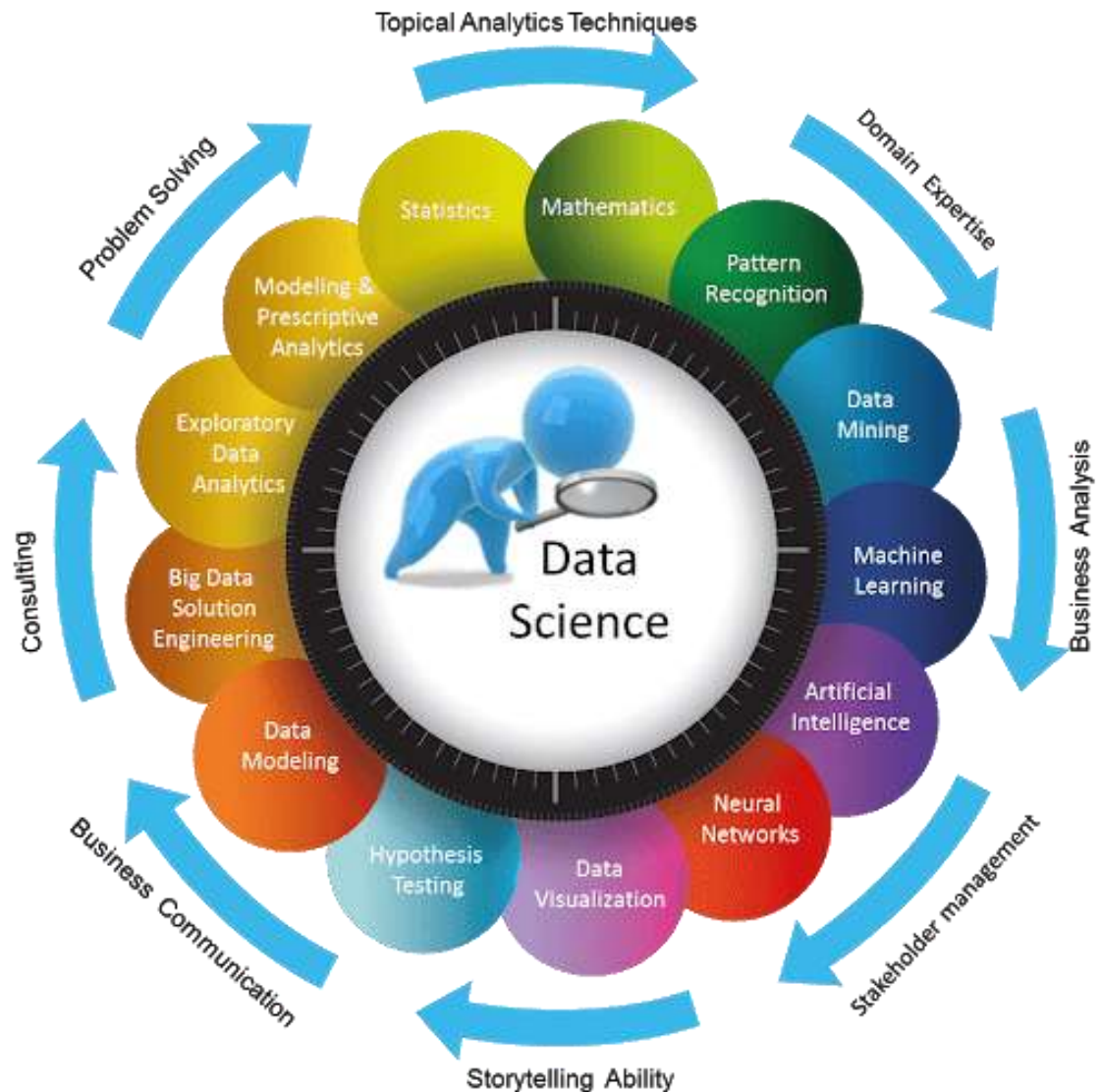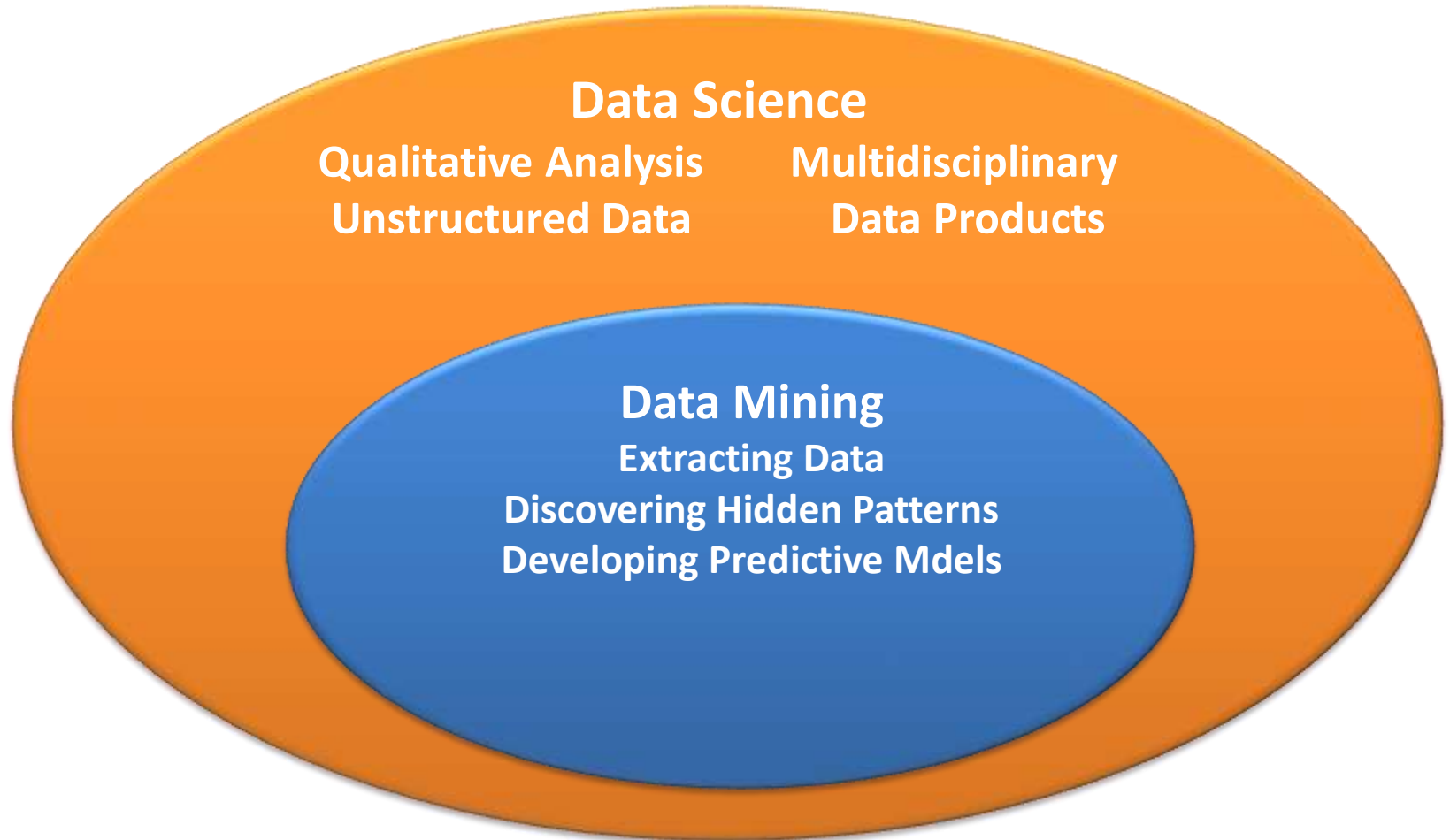
Data Mining

Data mining is a process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems

Data Science & Data Mining

Data Science
Qualitative Analysis   **Multidisciplinary**
Unstructured Data   **Data Products**

Data Mining
Extracting Data
Discovering Hidden Patterns
Developing Predictive Mdels

Mr.Gangadhar Immadi

Data Analysis

➢ Process of examining, transforming, and arranging raw data in a specific way to generate useful information from it

➢ Evaluation of data through analytical and logical reasoning to lead to some sort of outcome or conclusion in some context

➢ Multi-faceted process that involves anumber of steps, approaches, and diverse techniques

➢ Explain how and why about past events

➢ Includes Qualitative(Explains how and why story ends the way it did) and Quantitative(Data + How the sale decreased last summer)

Data Analysis process and Tools

1. Identify the problem or opportunity for value creation
2. Identify the Sources of the Data
3. Preprocess the data for issues such as missing and incorrect data
4. Devide the daasets into training and validation datasets
5. Build Analytical models and identify the model which gives best performance
6. Impleent solution / Decission / Develop Product

Data Analytics

➢ Data analytics is the science of analyzing raw data in order to make conclusions about that information.

➢ Focuses on Future and explore potential future events

➢ Includes Qualitative(Intuition and Analysis) and Quantitative(Formulas and Algorithms)

Business Analytics - Definition

Business analytics (BA) refers to the tools, techniques and processes for continuous exploration and investigation of past data to gain insights and help in decision making and problem solving.

Business Analytics is an integration between **business/problem context**, **technology** and **data science** that assist data driven decision making/problem solving.
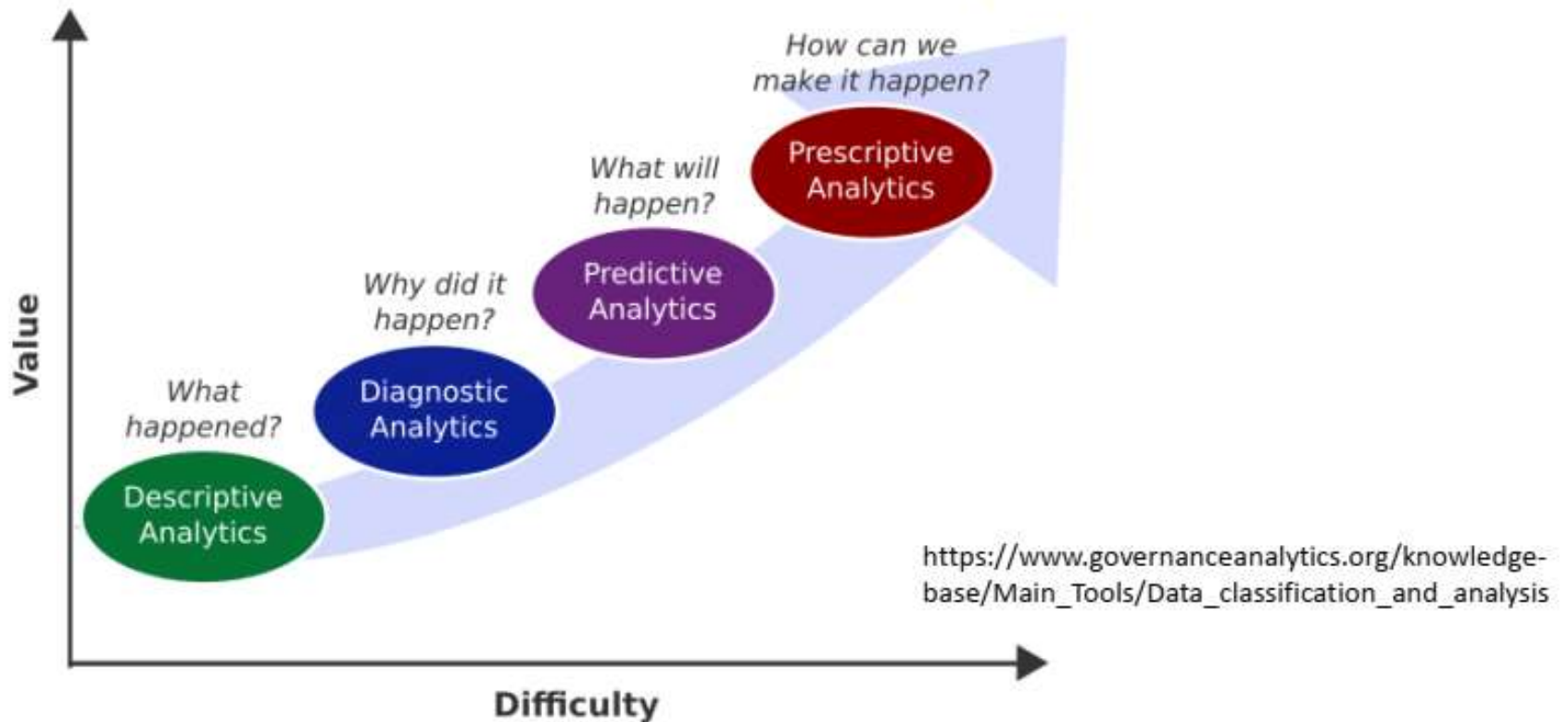
Analysis != Analytics
Data Analysis != Data Analytics
Business Analysis != Business Analytics

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**

Examples

- 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

Source:
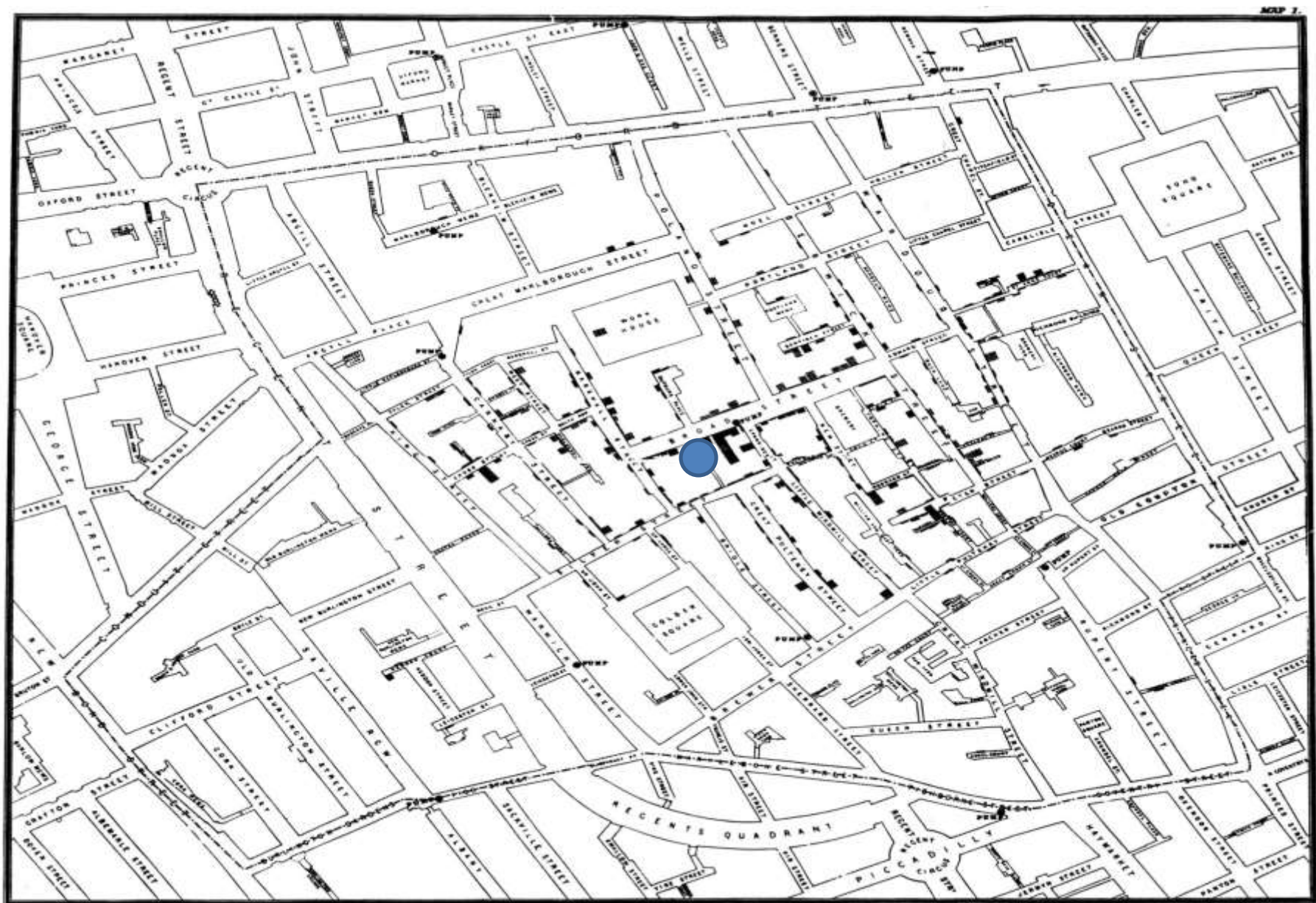https://www.linkedin.com/learning/478e9692-d13d-338f-907e-d76f0724d773

London Cholera Outbreak - 1854

Severe outbreak of cholera that occurred near Broad Street (now Broadwick street) in Soho district of London in 1854.

More than 500 people died within 10 days of the outbreak, the mortality rate in some parts of the city was as high as 12.8%.
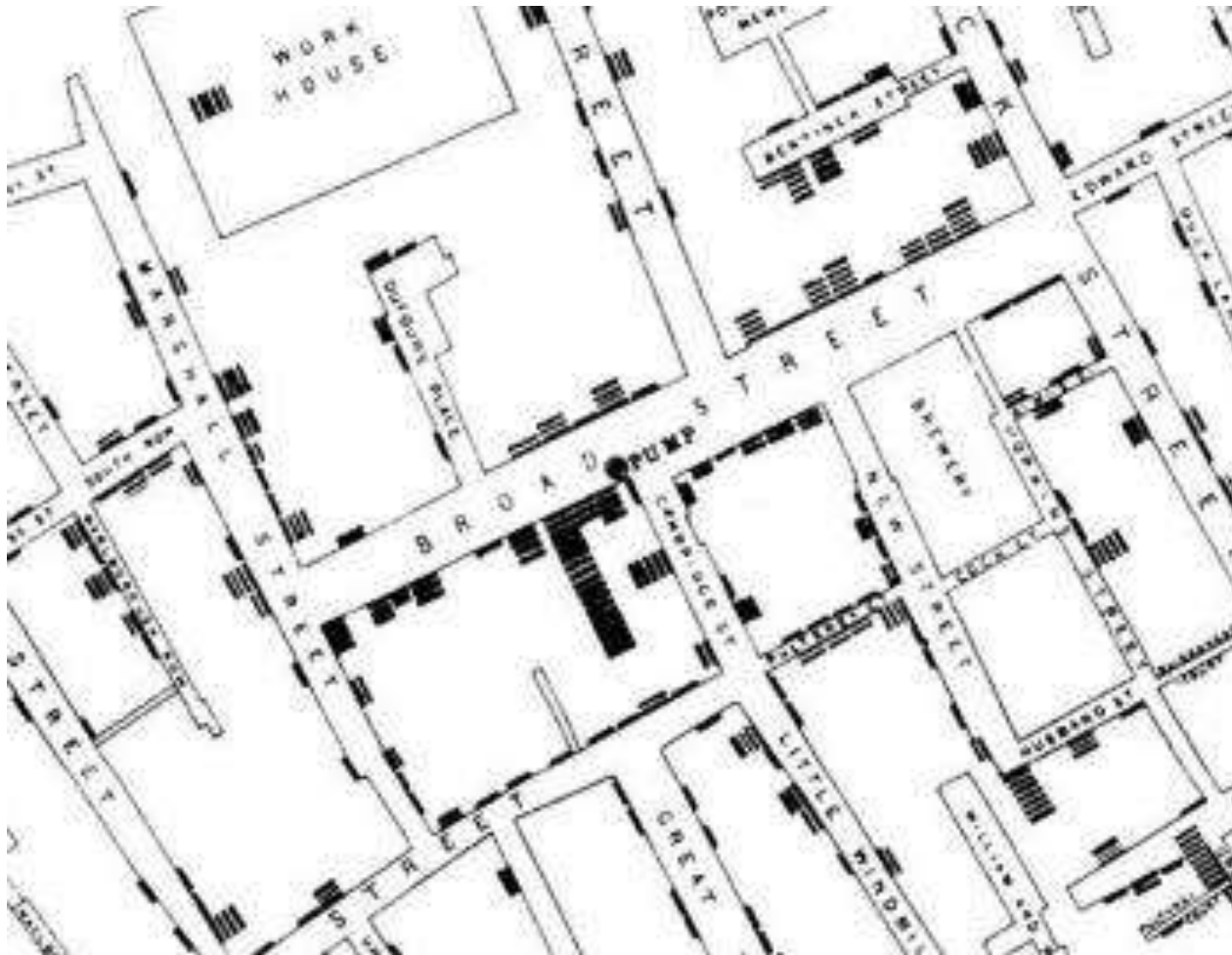
Mr.Gangadhar Immadi

To understand God's thoughts, we must study statistics, for these are the measures of his purpose.

- Florence Nightingale

Descriptive Analytics Applications

➢ Most shoppers turn towards right when they enter the a retail store.

➢ Conversion rate of women shoppers is higher than male shoppers among electronic gadgets purchasers (Radio Shack).

➢ Strawberry pop-tarts sell 7 times more during hurricane compared to regular period (Wal Mart).

➢ Women car buyers prefer women sales person.

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

➢ Examples :Data Discovery,Data Mining,**Correlations**
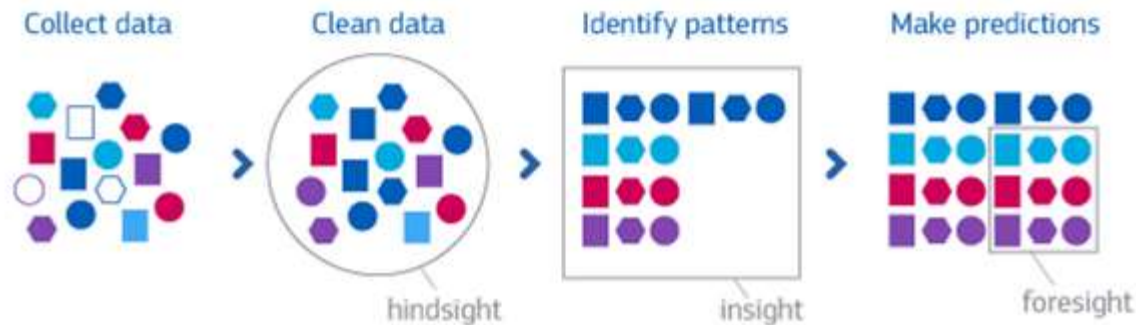
Predictive analytics

- 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

Source: https://www.logianalytics.com/wp-content/uploads/2017/11/predictive-1.png
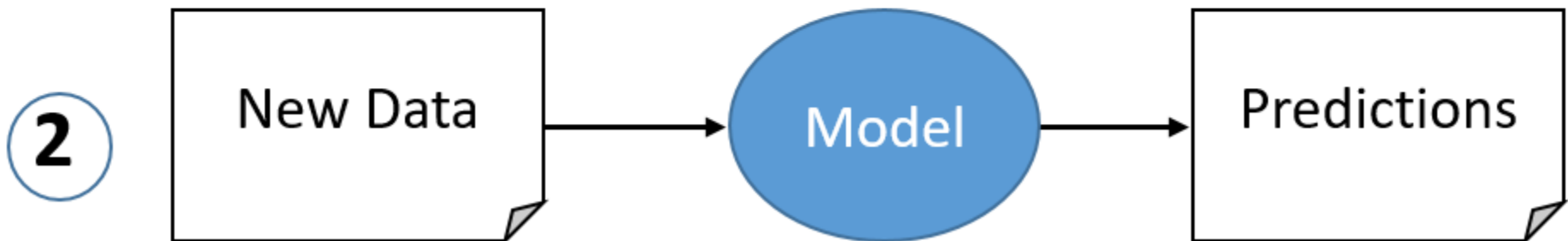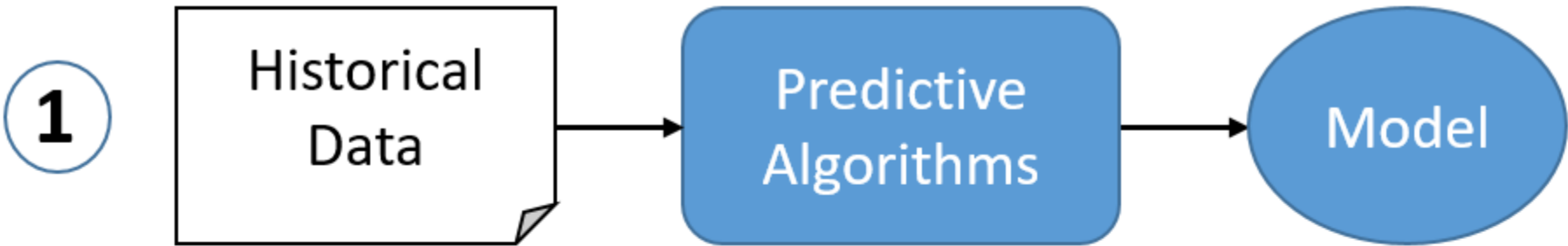
Example

- 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

Source: https://bigdata-madesimple.com/5-examples-predictive-analytics-travel-industry/

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:
  1. **Quality** improvements
  2. Service **enhance**ments
  3. Cost **reduction**s and
  4. Increasing **productivity**

➢ Example - **Optimization** Model, Simulation, Decision Analysis

Prescriptive Analytics Problem

➢ What is the optimal product mix?

➢ What is the optimal route for a delivery truck.

➢ Best markdown pricing for fashion products.

➢ Optimal assignment of aircraft to flight.

➢ How to manage the fleet of vehicles owned by a company for employee drop and pick up?
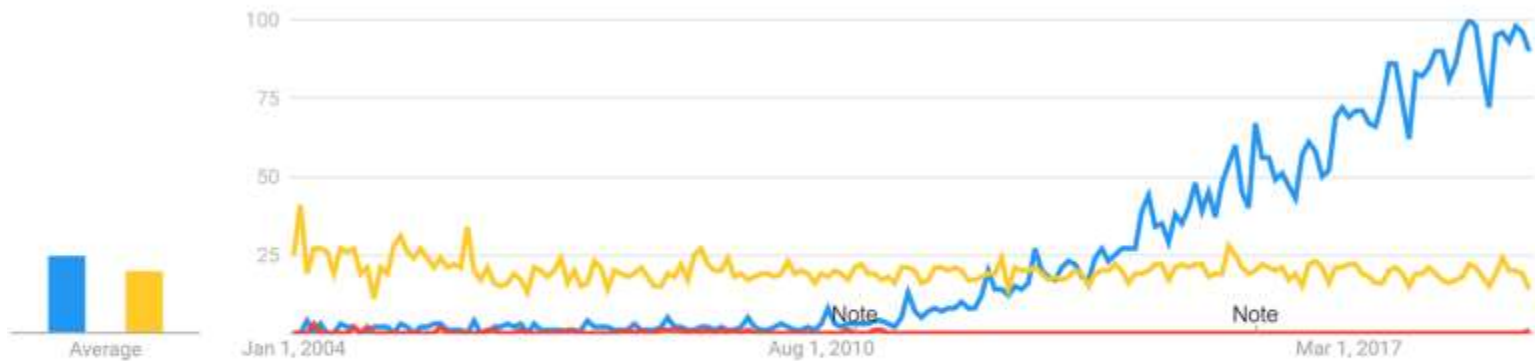
Need for Data Science
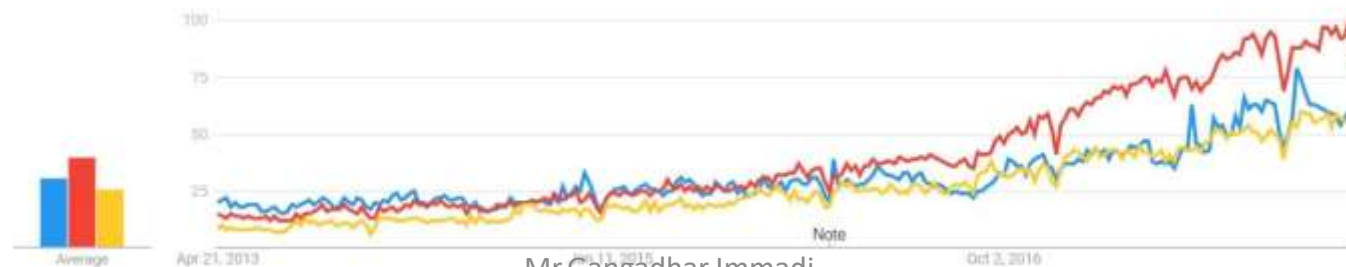
How data add value to business?

Source:https://datajobs.com/
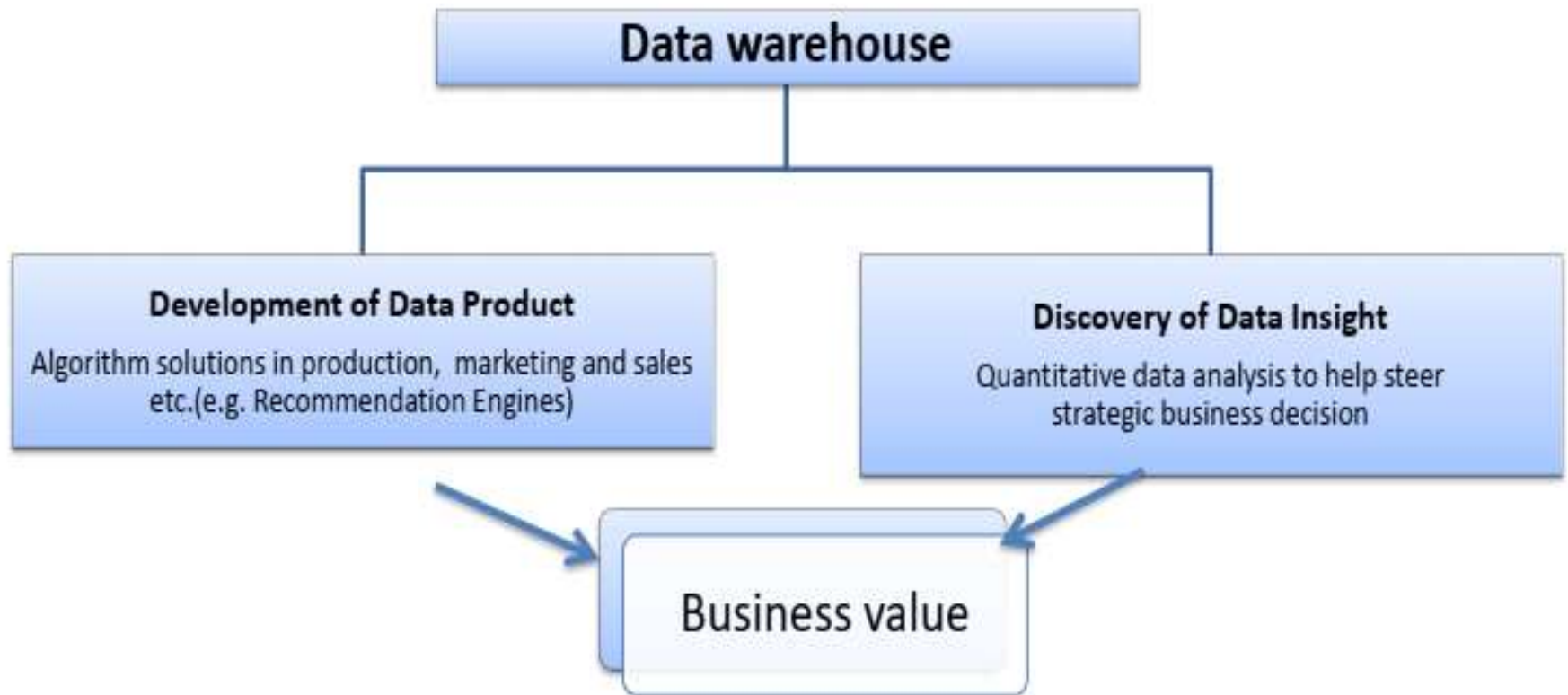
Data Products

RANI CHENNAMMA(16589)

KSR BENGALURU (SBC) → HUBBALLI JN (UBL)
AC 3 Tier (3A)

Current availability as on 3-1-2020:			GNWL31/WL11

Probability of Confirmation: 89% *
Last Year Confirmation Trends for Same Period

Journey Date	Booking Confirmed Upto	Journey Date	Booking Confir
27 Dec 2018	WL/80	28 Dec 2018	WL/10
29 Dec 2018	WL/50	30 Dec 2018	WL/76
31 Dec 2018	WL/16	01 Jan 2019	WL/36
02 Jan 2019	WL/43	03 Jan 2019	WL/51
04 Jan 2019	WL/46	05 Jan 2019	WL/31
06 Jan 2019	WL/31	07 Jan 2019	WL/49
08 Jan 2019	WL/40	09 Jan 2019	WL/31
10 Jan 2019	WL/44		

Click Here to Check CNF Availability on alternate Train

*The information displayed here is purely indicative and is based on analysis of current and past trends. At times, the predication may d
status. Indian Railways / IRCTC will not be liable in case of any such deviation as confirmation of reserved accommodation is based on v.

amazon pay

Dear Customer,

Your postpaid bill of Rs.241.90 for 9986789040
is due on 09 January 2020, Please ignore if
already paid.

Book tickets now!

book my show

Avane Srimannarayana
A story with a quirky fantasy and a touch of historic takes you to the land of mysteries and magic. Read more.

Vertica Components   Data Mining | Cours...

Gmail

It seems like you forgot to attach a file.

You wrote "I have attached" in your message, but there are no files attached. Send anyway?

1 of 11

OK   Cancel

iman

<no-reply@updates.bookmyshow.com> Unsubscribe

Book tickets now!

able

Gangadhar Immadi

Easily approachable

I have attached the file
-Regards,
Gangadhar Immadi
Assistant Professor,
Department of Information Science & Engineering
New Horizon College of Engineering, Bangalore
Mob:9986789040

Sans Serif   B I U A

Send

Mr.Gangadhar Immadi
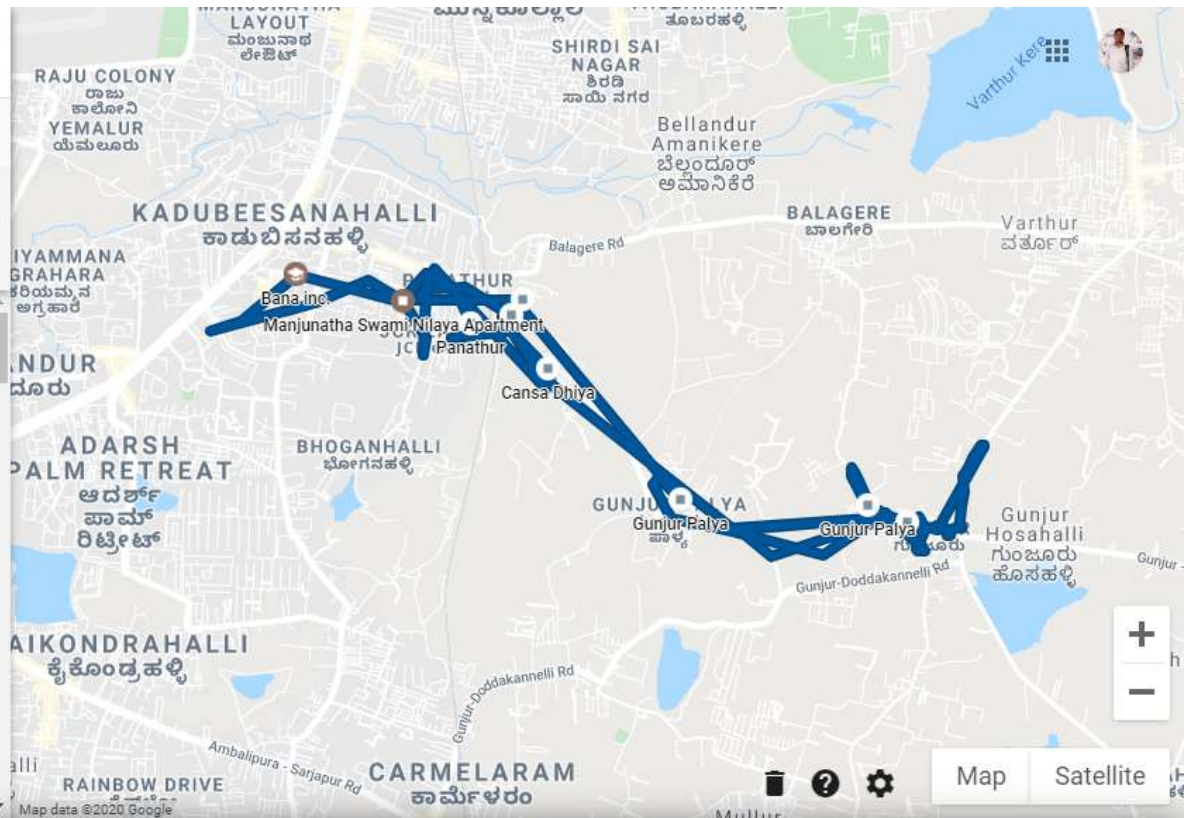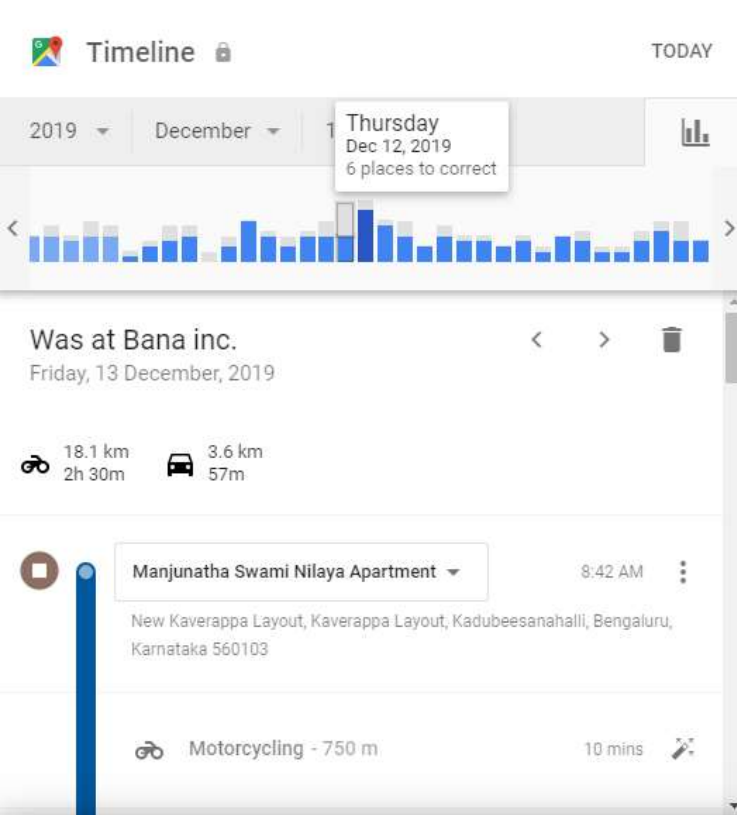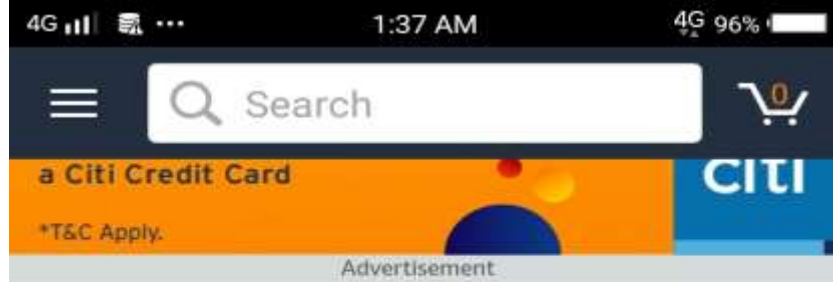
Search

a Citi Credit Card
CITI
*T&C Apply.

Advertisement

Related to items you've viewed

15% off

15% off
BAL'SON
ULTRA 1X
BATHROOM FILTER

Eco Crystal Fresh a...
₹ 2,600 ₹ 3,050

BALSON Ultra 1X B...
₹ 2,796 ₹ 3,290

25% off

10% off

D'Cal Hard Water S...
₹ 2,700 ₹ 3,600

Eco Crystal Fresh A...
₹ 2,199 ₹ 2,450

See more & manage ⟩

< **QP-MEDPLS** Edit

Jan 1, 12:01 PM

AMLOKIND ... are due for purchase.
Pls visit near by Medplus Store or call
040-67006700 or re-order at
https://mdpls.in/oNrCN1

Jan 1, 12:17 PM

AMLOKIND ... are due for purchase.
Pls visit near by Medplus Store or call
040-67006700 or re-order at
https://mdpls.in/oNrCN1

+ Cannot send message to special number Send

Mr.Gangadhar Immadi

Light aircraft

Plantation region

Field layout

Field Overview

Mr.Gangadhar Immadi

An NDVI image of iceberg lettuce field

Data Analyst 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

Source:https://datajobs.com/
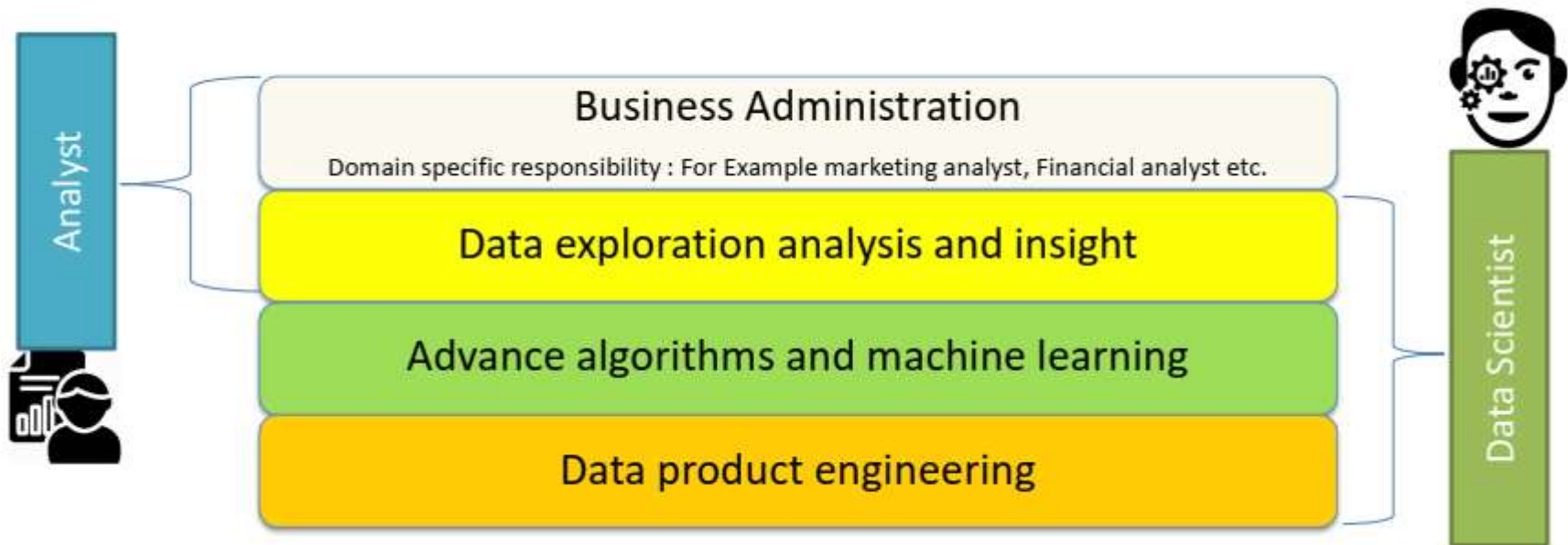
Extracting value from the data

Data Science

Statistical models & Machine Learning Algorithms

Problems, Opportunities, Decision Scenarios

Data collection, storage, retrieval Software tools

Business Context

Technology

Mr.Gangadhar Immadi

ANALYTICS

Competitive
Strategy
Data is everything

Decision Making
(What promotion Strategy to use)

Problem Solver
(Optimal Product Mix)

Process Improvement
(Reduce procurement cycle time)

Analytics for Process Improvement

- ➤ Banking – Cheque clearance time

- ➤ Healthcare – Patient discharge time

- ➤ Manufacturing – Waste minimization

- ➤ Retail – Waiting time at check out counters

- ➤ E-commerce – Time to deliver the customer order

Analytics **for** Problem **Solving**

➢ Banking – Reduce non-performing assets, Predict Fraud

➢ Healthcare – Improve net promoter's score (NPS)

➢ Manufacturing – Reduce inventory management cost

➢ Retail – Assortment planning and shelf space allocation

➢ E-commerce – Predict customer cancellations and Fraud

Analytics for Decision Making

➢ Banking – Loan approval and the interest rate

➢ Healthcare – Introducing new specialties

➢ Manufacturing – Whether to introduce a new product

➢ Retail – Markdown Pricing

➢ E-commerce – Promotions

Machine Learning

- Task

- Experience

- Performance

Source : Machine Learning : Tom Mitchell

Mr.Gangadhar Immadi

Lidar (light detection and ranging)

GPS (global positioning system)

Rear Camera

Video cameras

Ultrasonic sensors

Odometry sensors

Central computer

Radar sensors

Mr.Gangadhar Immadi

More Examples

➢ Comparison of Curriculam

➢ Comparison of Results

➢ Students Result Prediction

➢ Promotion/Increment by HRMS

➢ CAPA generation with minimal intervention

➢ Calender of Events Reminders

➢ Hiring Process Prediction

➢ Matrimony Match

➢ Job Recommendation

➢ Personalization

➢ Food Recommendation

Industry wide applications of analytics

Industry Sector	Sample Analytical problems	Data Sources
Manufacturing	• Supply Chain Analytics • Quality and process improvement • Revenue and cost management • Warranty Analytics	• Procurement , sales and production data • Warranty and after sales service • Commodity Price Data • Manufacturing Data • Macroeconomic Data
Retail	• Assortment Planning • Promotion Planning • Demand forecasting • Market Basket Analysis • Customer Segmentation	• Price data • Demand data at SKU and at category level • SKU level sales data with and without promotions • Planogram • Customer demographics data • Point of sales data • Loyalty program data
Healthcare	• Clinical care • Hospitality related data	• All patient care related data • Hospitality related data • Patient feedback data

Industry Sector	Sample Analytical problems	Data Sources
Service	• Demand forecasting • NPS Optimization • Service Quality Analysis • Customer Segmentation • Promotion	• Transactional and feedback data • Pricing and demand data • Promotional data
Banking & Finance	• Assortment Planning • Promotion Planning • Demand forecasting • Market Basket Analysis • Customer Segmentation	• Customer transactional data • Loan originating data • Credit scoring data
IT and ITES(IT enabling Services)	• Demand for Analytics Services • Software Development Cycle time	• Customer interaction and market research data • Internal product development data
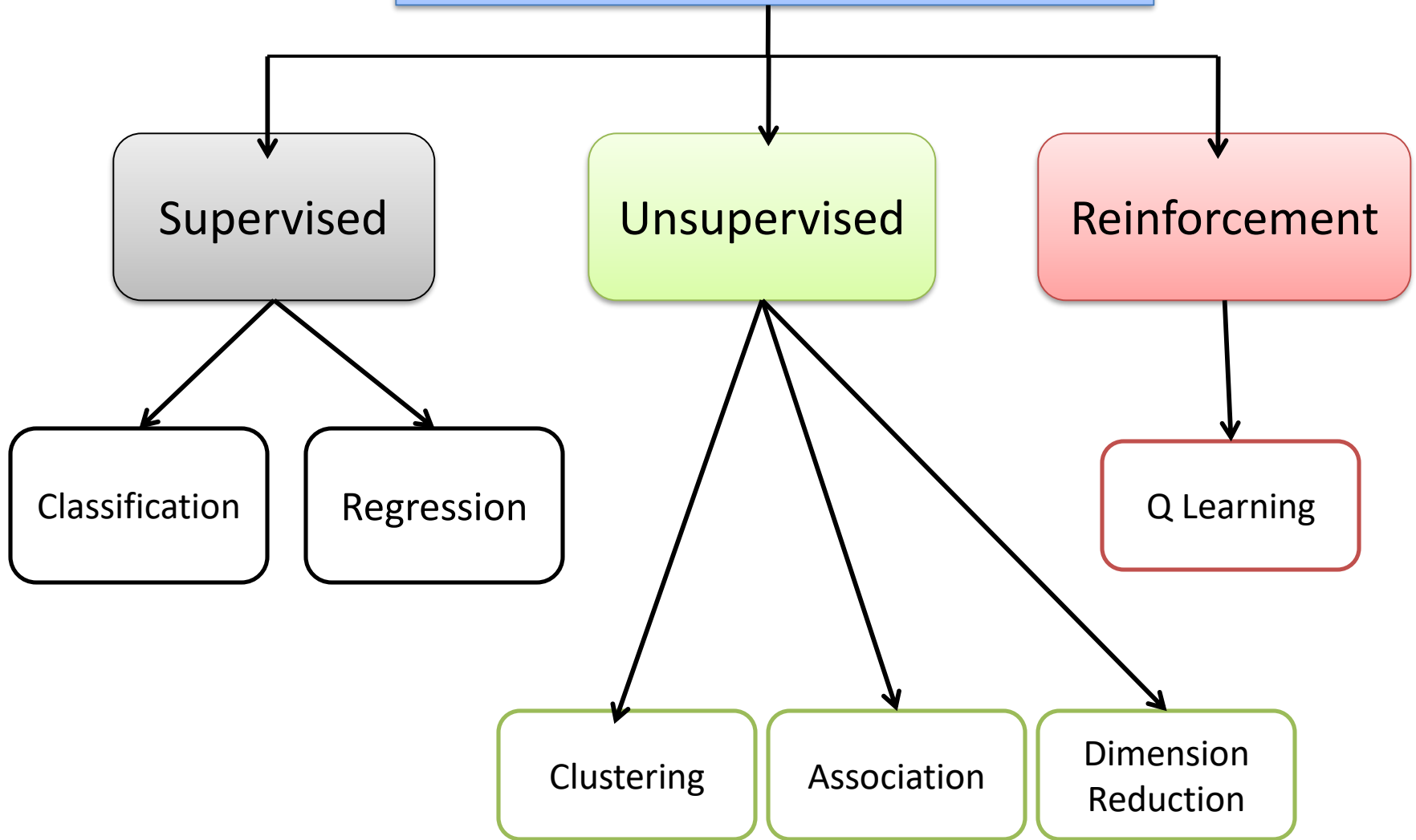
Types of Learning (1)

Mr.Gangadhar Immadi

Picture	Result
	Daddy
	Mommy

Types of Learning (2)

Mr.Gangadhar Immadi

Types of Learning (3)