

Day 1

Getting Started With Data Science

Amazon → searching Lap

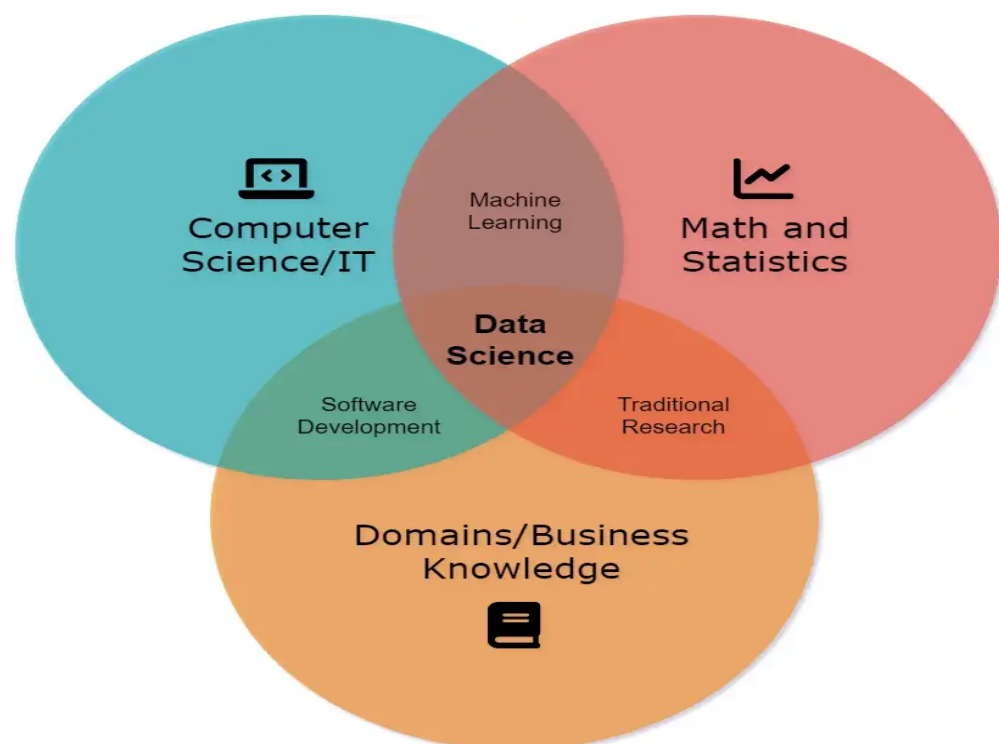
Instagram → Showing an ad of Lap -> Click -> Redirect to amazon

Key → Data

Data Scientist → Plays a role with data in Business

### Introduction:

Data Science -> AI , ML, DL , Mathematics, Statistics, Domain Knowledge (EEE, ECE , Civil , Mech , auto ), Computer Programming skills (Python , R)



### Techniques:

- Classification
- Linear Regression
- Logistic regression
- Neural networks
- Anomaly detection
- Clustering
- Decision Tree
- K - Mean clustering
- Hypothesis testing
- Distribution
- Prob & Statistics

### **Processes:**

PS: Typing issue , Language problem

Sln: Speech to text conversion

- Problem Discovery/ Business understanding  
->Google  
Sln : Speech to text conversion
- Data collection  
India 1k+  
UK  
US  
China  
Japan  
Korea  
German  
Portuguese  
South african etc...

- Data preparation  
Speech → Format which will be easier to convert the text → Verb, Nouns , Gerund
- Exploratory Data Analysis  
Data presentation like Image or graph
- Modelling  
Brain
- Model Evaluation -> Grading
- Model Deployment -> Releasing to real world

#### **Procedures:**

- 1. Algorithm**
- 2. Machine Learning**
- 3. Neural Network – ANN , DNN , RNN**

#### **Applying Data Science to Business and Industry:**

1. Data-Driven Insights  
Data → Applying Data Science (Algo, ML , NN) → O/P
2. Business Processes
  - Opportunities of improvement
  - Reducing cost
  - Working towards Long term goals
3. Analytics and data wrangling  
Data → Analysing → For finding a sln or problem

statment

Data Analyst → Processing data , Visualising,  
Consolidation , Storing and retrieving  
Application -> PowerBI , SQL , Excel etc..

Python -> Pandas , Numpy , Matplotlib etc...

### **Data wrangling: Data Munging**

- The processing of data and using the data for the some business purpose is called data wrangling

Raw I/P (Data) → Collecting -> Modifying -> Analysing(Processing data , Visualising, Consolidation , Storing and retrieving) → Making decision out of it

### **Implementation method of Data Wrangling :**

Python Framework ->Pandas (Sorting, filtering , grouping etc ...)

1. Data Exploration
2. Dealing with missing data
3. Reshaping the data
4. Filtering data
5. Visualisation etc ... (Matplotlib , Seaborn )

## **4. Business Intelligence vs Data Science**

BI , Business Analyst → **Know the DS project req, o/p**

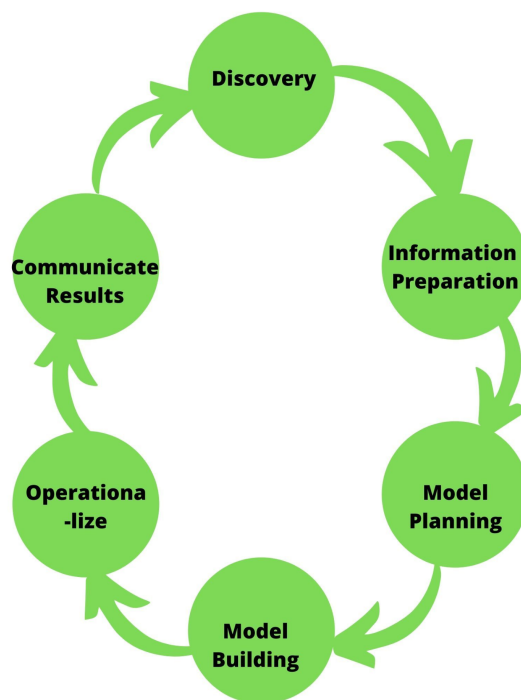
**Ps-> BI -> Cost , Resource planning , S/w , H/W**

**DS -> Req doc -> Yes, No → Complete Data analyst, Data eng ..**

BI	DS
It focus of present and past data	It focus on future planning
Business Analyst , BI Engineer	Data Scientist , Data Analyst , Data Engg

For business growth - Small impact	Greater impact
Power BI , SQL etc..	Prgm Languages , ML , SAAS - AWS , GOOGLE CLOUD

## Processes of Data Science:



## Eg: Diabetic Retinopathy in Hospital

1. Collecting, querying, and consuming data

Scenario 1: Age 40 -> Eye sight prblm

Scenario 2 :Age 35->Getting headache

Scenario 3: Age 45 ->Wearing glass &  
need to be operated

Scenario 4: Age 50 -> Retina replacement process done

**Collecting** -> Patient data

**Querying** -> SQL , MYSQL , MONGODB

Age between 25 to 35 with diabetes & wearing glass

**Consuming** ->. Using the data for problem solving

2. Making use of maths and statistics

Comparing the data with normal users vs diabetic person

3. Applying data science to your subject area

Algo , ML to Health industry

4. Communicating data insights

*Final outcome will identify the patient the patient with diabetic retinopathy*