

# Introduction to Data Science

# What is Data Analytics?

- **Data Science** is a science of analyzing raw data using statistics and machine learning techniques with the purpose of drawing insight from the data
- Data Science is used in many industries to allow them to make better business decisions, and in sciences to test models or theories.
- This require process of inspecting, cleaning, transforming, modeling, analyzing, and interpreting raw data

# What is Data Analytics?

Extraction of knowledge from large volumes of data that are structured or unstructured.

It is a continuation of the fields **data mining** and **predictive analytics**

Height (Cm)
160
165
142
178
134
...
149

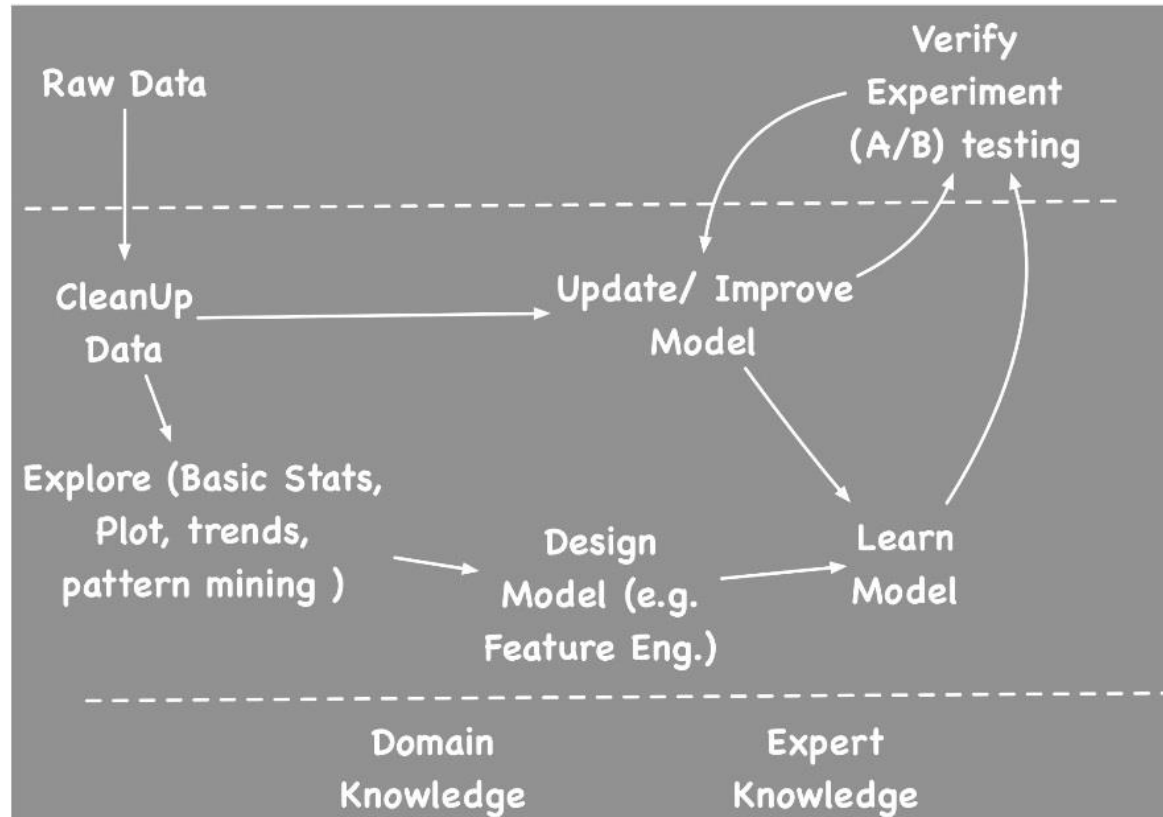


Range = 35 cm  
Avg = 155 cm

# Process of Data Science

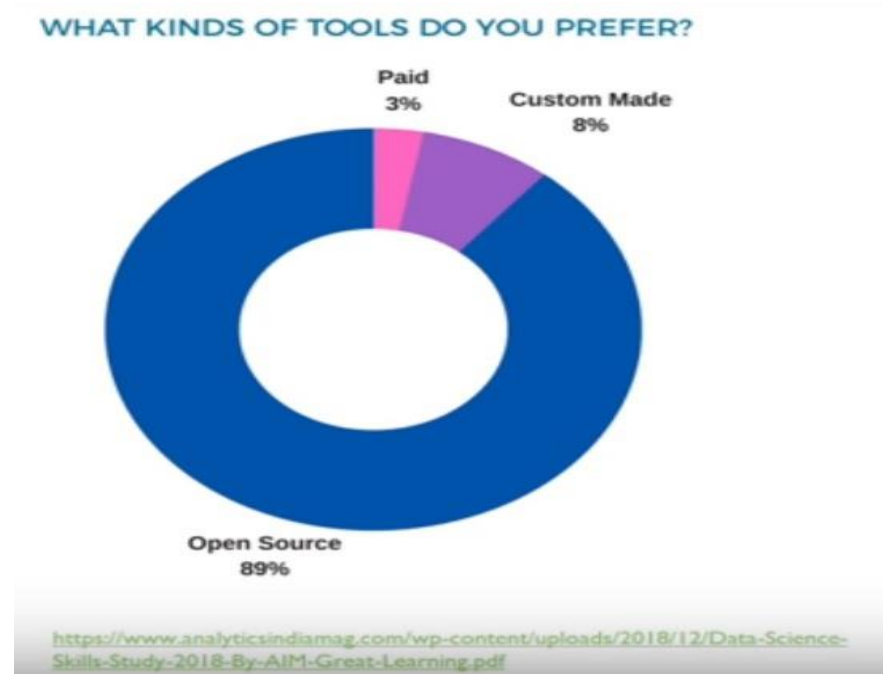


# Data Perspective



# Trends in tools used for Data Science

- As of 2018, most Indian scientists prefer to use open source tools over paid or custom made tools



# Trends across the Globe

## US, Europe, Japan and China

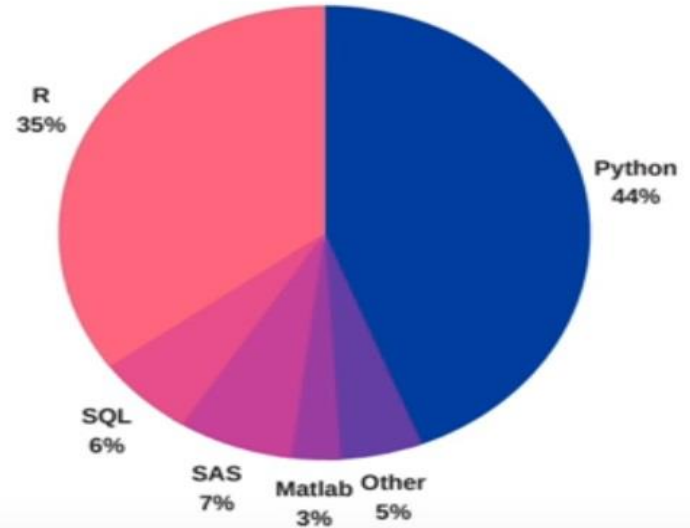
### Programming Languages and Tools

#### Main programming language for data analysis

Number of answers: 1522



## India



<https://www.jetbrains.com/research/data-science-2018/>

<https://www.analyticsindiamag.com/wp-content/uploads/2018/12/Data-Science-Skills-Study-2018-By-AIM-Great-Learning.pdf>

# Session Plan

Sl. No.	Topic Name
1.	Numpy Library
2.	Pandas Library
3.	Static Visualization
4.	Data Cleaning & Feature Extraction
5.	Dynamic & Geographical Visualization
6.	Time Series Analysis
7.	Case Studies and Mini Project



# Data Science using Python

- Python libraries provide key feature sets which are essential for data science
- Data manipulation and pre-processing
  - Python's Numpy & Pandas' library offers a variety of functions for data wrangling and manipulation
- Data Cleaning, Feature Extraction and Summary
- Visualization
  - Plotting libraries like 'matplotlib', 'seaborn', 'plotly', 'cufflinks' and 'folium' aid in corresponding statistical information and help in identifying trends and relationship
- Various case studies exploring Analytical Skill-set

# Installation of Anaconda Software

Thank You