

# Data Science

*Mega Trend of 21<sup>st</sup> Century*

Sonia Sharma  
Software Engineer-II



# Agenda

- Why To Learn Data Science
- Evolution of Data and Statistics
- What is Data Science
- Purpose and Implementation of Data Science
- Tools and Skills Required
- Categories and Types of ML Models
- Consequences



# Why To Learn Data Science

---

**Glassdoor** has ranked Data Science as its topmost profession.

- A fuel of 21st Century
- Problem of Demand & Supply
- A Lucrative Career



# Evolution of Data and Statistics

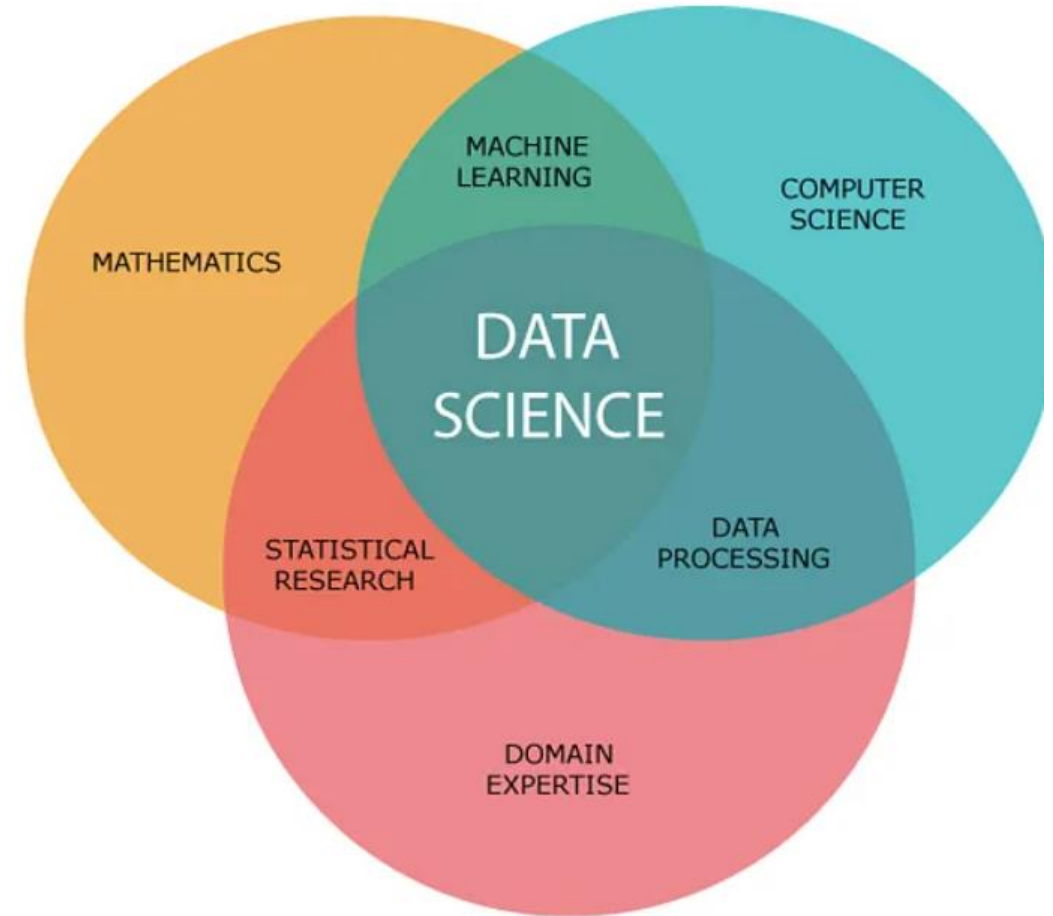
---

- **Data** is raw, unorganized facts that need to be processed. **Data** can be something simple and seemingly random and useless until it is organized.
- **Data** are individual pieces of information, recorded and used for the purpose of analysis. It is the raw information from which **statistics** are created.
- **Statistics** is the discipline that concerns the collection, organization, analysis, interpretation and presentation of data - its interpretation and presentation.

# What is Data Science

---

Statistical and ML tools to draw  
Patterns and Make business Decisions.



# Purpose of Data Science

Better Marketing

Customer Acquisition

Innovation

Enriching Lives





# Implementations of Data Science

# Places of Data Science in Data Universe

---

Artificial Intelligence

---

Machine Learning

---

Deep Learning

---

Big Data

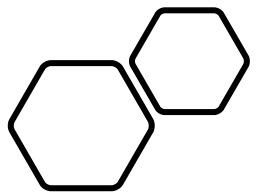
---

Predictive Analysis

---

Business Intelligence

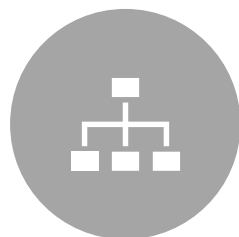




# Role of a Data Scientist



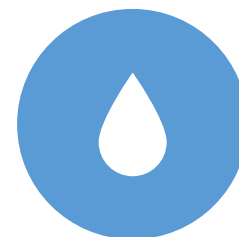
GET DATA



MEANINGFUL  
STRUCTURE



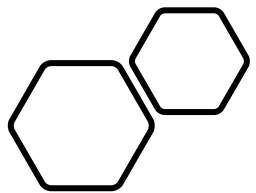
ANALYZE



HYPOTHESIZE



PREDICTION  
(ML)



# Top Skills for Data Science



Business Skills



Practical Skills (Math and Statistical)



Coding Skills



Soft Skills

# Tools



A horizontal bar chart titled 'Tools' comparing the usage of four different tool categories. The categories are listed on the left, and their corresponding usage percentages are shown as horizontal bars on the right. The bars are color-coded: Python / R is orange, SQL is reddish-brown, Hadoop is dark brown, and Excel/PowerBI/Tableau is grey. The bars are arranged in descending order of length from top to bottom.

Tool	Percentage
Python / R	75%
SQL	70%
Hadoop	65%
Excel/PowerBI/Tableau	60%

Python / R

SQL

Hadoop

Excel/PowerBI/Tableau



# Process To Target A Problem

- Identify Data Sources
- Ingest
- Clean Data
- Exploratory Data Analysis
- Train/Test Split
- Build Model on Train Data
- Evaluate
- Deploy, Monitor, Refine



# Machine Learning Categories

Supervised Learning

Unsupervised Learning

# Types Of Models

Regression

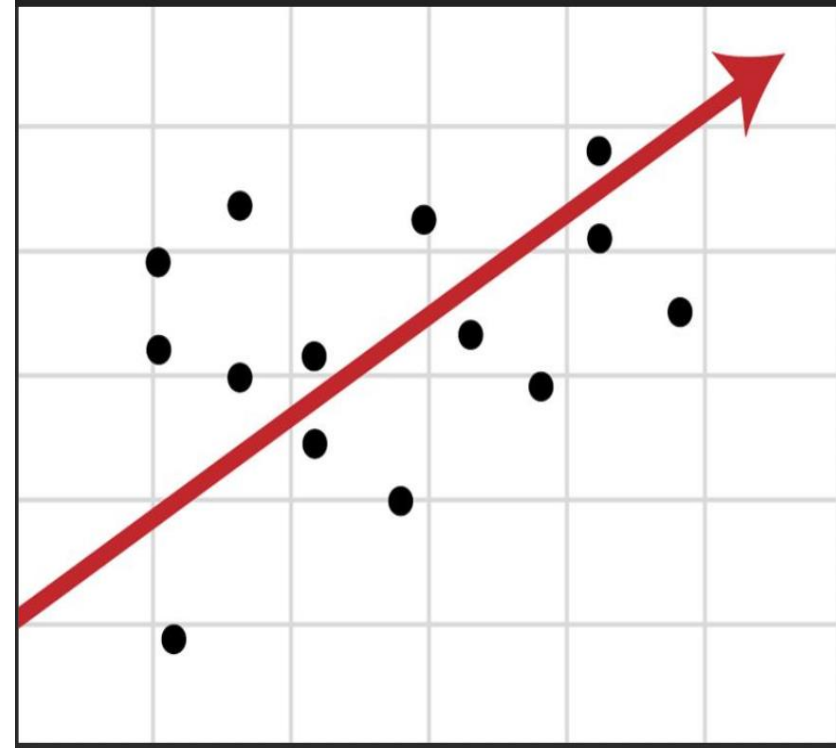
Classifications

Clustering

# Types Of Regression Algorithms

---

- Simple Linear Regression
- Multiple Linear Regression
- Lasso Regression
- Support Vector Machine
- Multiple Variate Regression



# Types of Classification Algorithms

---

- Logistic Regression
- Naïve Bayes Algorithms
- K-Nearest Neighbors
- Decision Trees
- Random Forest
- Support Vector Algorithms

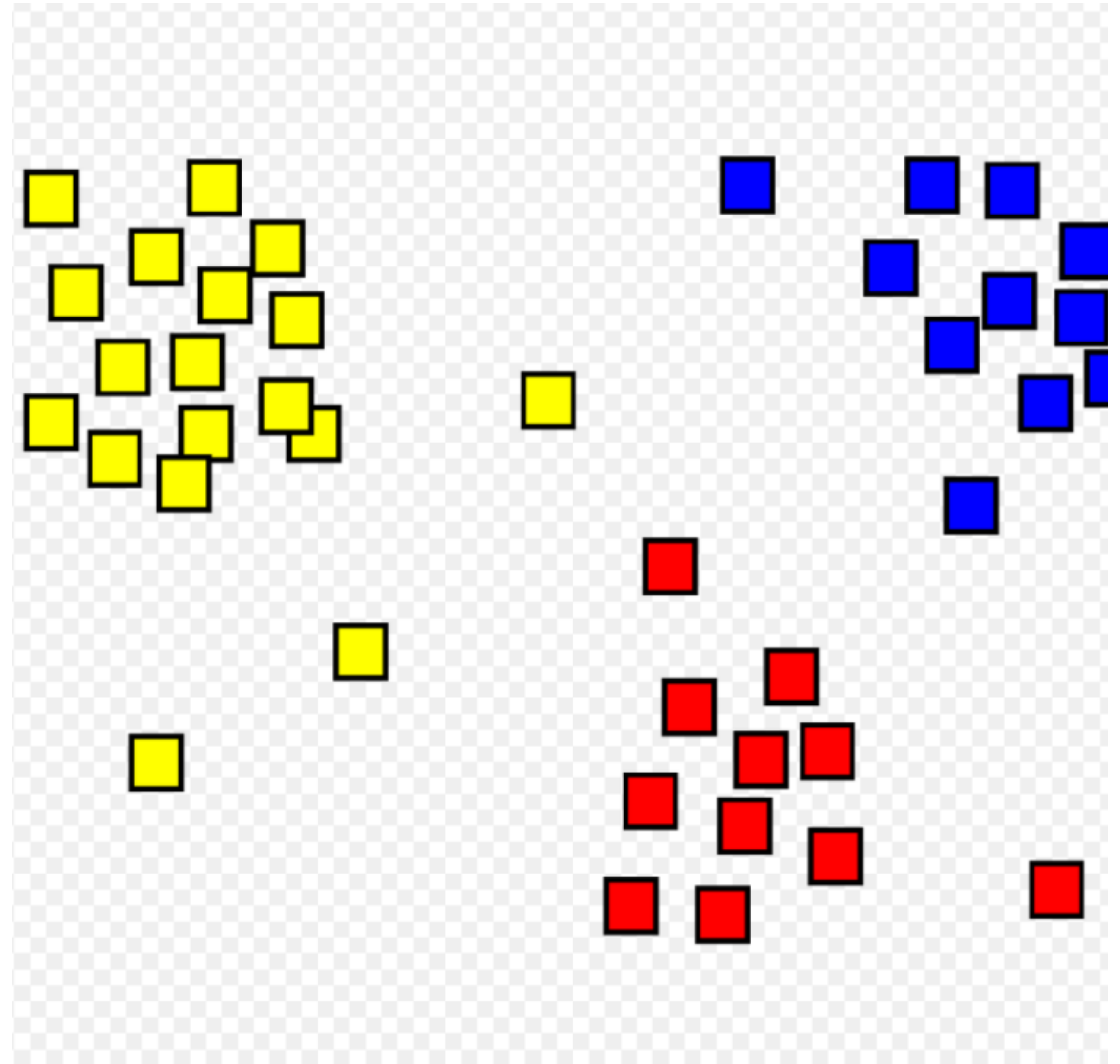




# Types of Clustering Algorithms

---

- K-Means Clustering
- Mean Shift Clustering



**A**

**It is a Blurry Term**

**Mastering Data Science  
is near to impossible**

**B**

**C**

**Large amount of domain  
knowledge required**

**Arbitrary Data May Yield  
Unexpected Results**

**D**

**E**

**Problem of Data Privacy**

Consequences



# Summary

“While Data Science is a vast subject, It is a very robust field that best fits people who have a knack for experimentation and problem-solving. With a large number of applications, Data Science has become the most versatile career.”

# References

A thick yellow horizontal bar spans the width of the slide, with a vertical yellow bar extending downwards from its right end.

- Kaggle
- Analytics Vidya
- LinkedIn Learning
- [Open Source Data Science Masters](#)
- [Elite Data Science](#) (65 Free Data Science Tutorials for Beginners)





# Thank You

**LinkedIn :** [www.linkedin.com/in/sonia~sharma](http://www.linkedin.com/in/sonia~sharma)

**Email:** soniasharma7381@gmail.com