

Department of Artificial Intelligence and Data Science

Fundamentals of Data Science (AD45)

Academic Year : 2024-24, Batch 2023

Credits: 3:0:0

Text Books:

Joel Grus, “Data Science from Scratch”, 2nd Edition, O’Reilly Publications/Shroff Publishers and Distributors Pvt. Ltd., 2019.
ISBN-13: 978- 9352138326

UNIT 1

What is Data Science? Types of Data & Data Sources, Visualizing Data, matplotlib, Bar Charts, Line Charts, Scatterplots, Linear Algebra, Vectors, Matrices, Statistics, Describing a Single Set of Data, Correlation, Simpson's Paradox, Some Other Correlational Caveats, Correlation and Causation. Probability, Dependence and Independence, Random Variables, Continuous Distributions, The Normal Distribution.

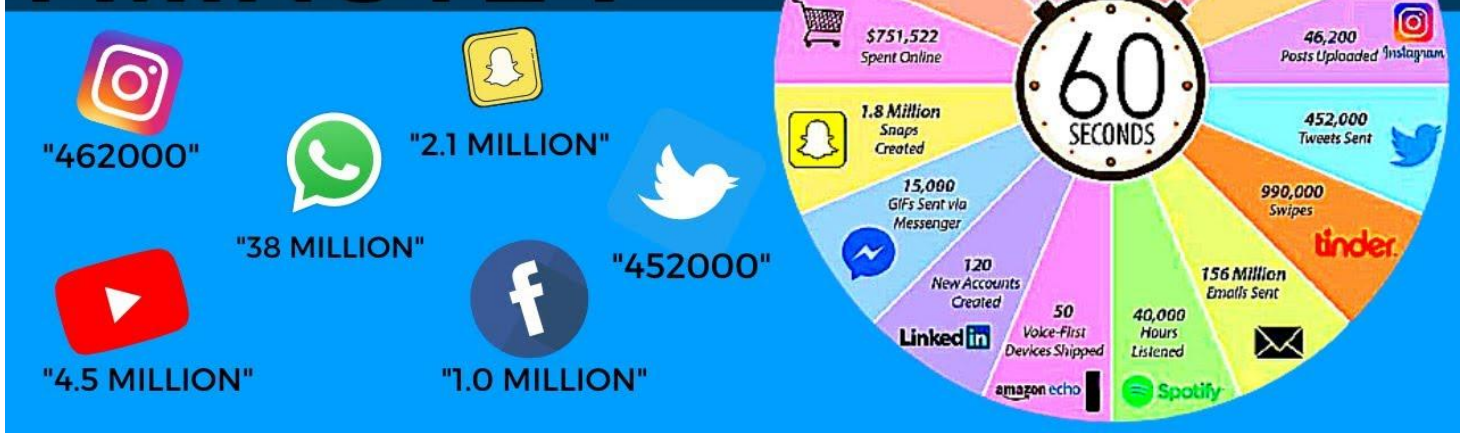
The Ascendancy of Data

- We live in a world that's drowning in data. Websites track every user's every click.
- Your **smartphone** is building up a record of your location and speed every second of every day. "Quantified selfers" wear pedometers-on-steroids that are always recording their heart rates, movement habits, diet, and sleep patterns.
- **Smart cars** collect driving habits, smart homes collect living habits, and smart marketers collect purchasing habits.
- The **internet** itself represents a huge graph of knowledge that contains (among other things) an enormous cross-referenced encyclopedia; domain-specific databases about movies, music, sports results, pinball machines, memes, and cocktails; and too many government statistics (some of them nearly true!) from too many governments to wrap your head around.

What Is Data Science?

- **What is Data?**
- **Data** is a collection of facts, numbers, measurements, observations. It can be raw or processed and is used to gain insights, make decisions, and drive technologies like AI and machine learning.

HOW MUCH DATA GENERATE IN 1 MINUTE ?



What Is Data Science?: Types of Data

- **Structured Data**

- ❖ Organized and stored in a fixed format (e.g., databases, spreadsheets).

- ❖ Examples:

- Customer records (Name, Age, Email)

- Sales data (Price, Quantity, Date)

- SQL Databases

id	name	age
1	Jim	28
2	Pam	26
3	Michael	42

id	subject	Teacher
1	Languages	John Jones
2	Track	Wally West
3	Swimming	Arthur Curry
4	Computers	Victor Stone

student_id	subject_id	grade
2	1	98
1	2	100
1	4	75
3	3	60
2	4	76
3	2	88

