أكاديمية طويق
Tuwaiq Academy

أكاديمية طويق

معسكر علم البيانات وتعلم الآلة

أكاديمية طويق
Tuwaiq Academy

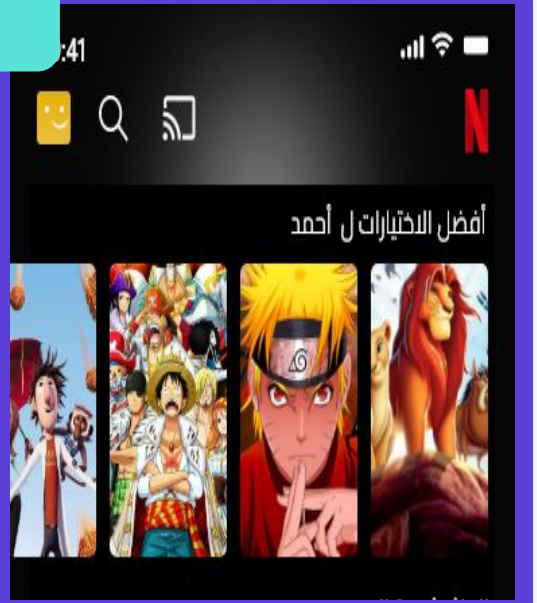
Data Science Introduction

By: eng. Esraa Madhi

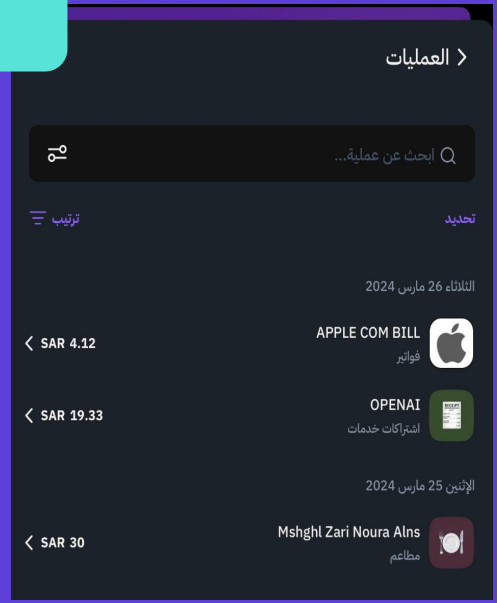
Where We Can See Data Science Products?

01

02

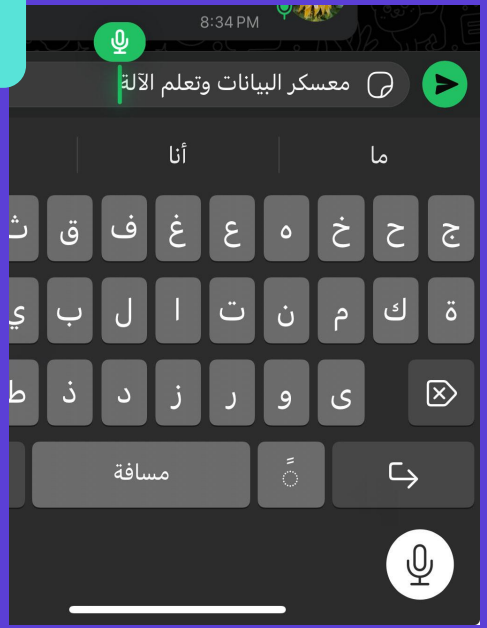


03

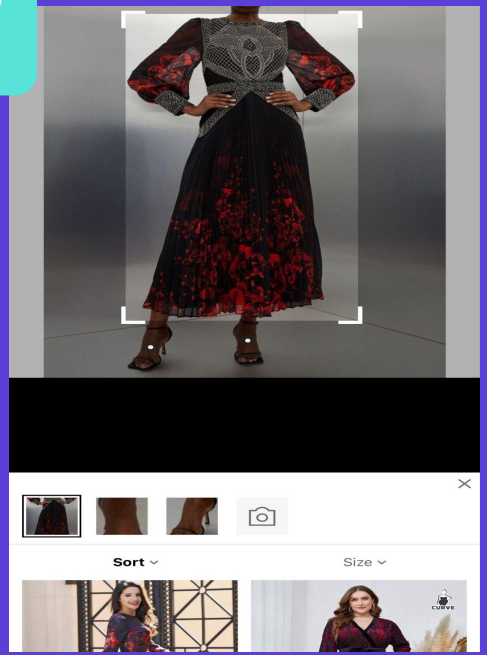


Where We Can See Data Science Products?

04



05



06

What Benefits Arise From Designing These Features/Function?

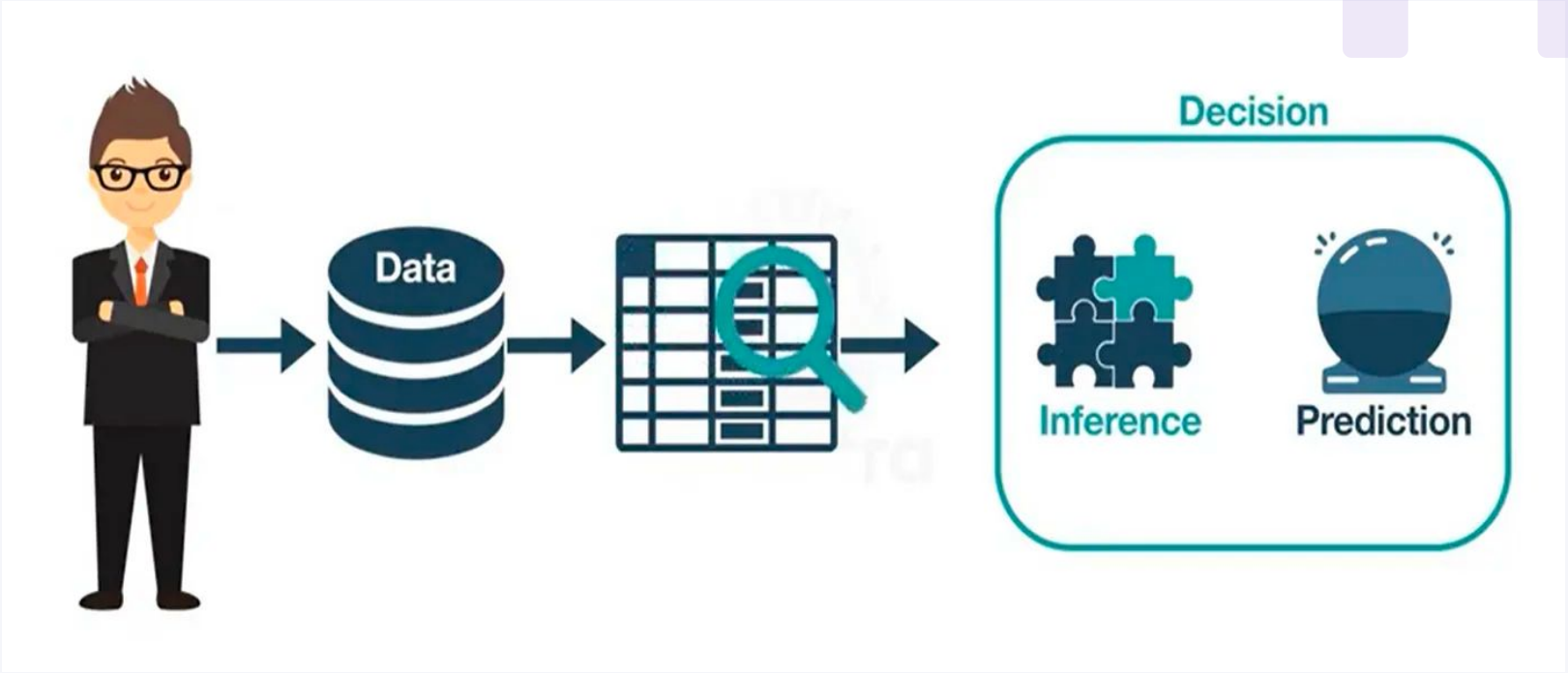
Manage Costs

Increase Efficiencies

Identify New Market Opportunities

Make Informed Decisions

How could they built such an application?



So We Can Define Data Science As:

It's the process of asking interesting questions, and then answering those questions using data

Data science enables businesses to process huge amounts of structured and unstructured big data to detect patterns

Is it a big deal ? to analyze data ? Why Data Science is Important?

- Approximately 328.77 million terabytes of data are created each day
- Around 120 zettabytes of data will be generated this year
- 181 zettabytes of data will be generated in 2025
- The amount of data that exists grows exponentially
- This means there is a huge amount of work in data science—much left to uncover. According to The Guardian, in 2012 only about 0.5 percent of all data was analyzed.

quantra

Introduction to Data Science

What is Data?

Data refers to raw, unorganized facts or figures that are collected and stored. It can be in the form of numbers, text, images, or any other type of input

Data, by itself, lacks context and meaning. It is the most basic form of representation and requires further processing to become useful

Examples of data:

- A database with customer information
- Temperature readings from a weather station

What is Information?

Information is created when data are processed, organized, or structured to provide context and meaning. Information is essentially processed data

It is the result of data being transformed into a more meaningful and useful state

Examples of information:

- A sales report highlighting top-selling products

What is Knowledge?

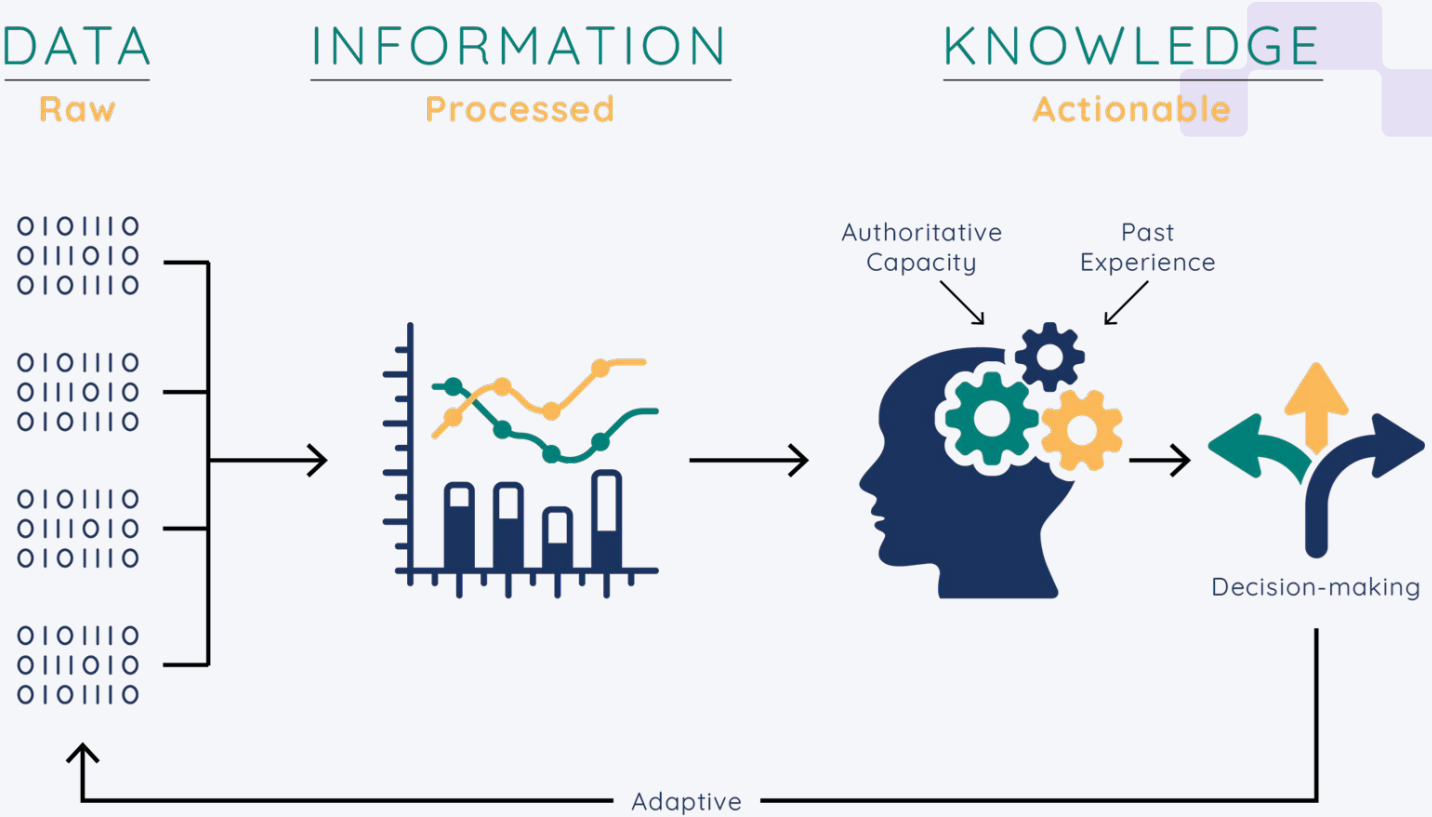
Knowledge goes beyond information in that it involves understanding and expertise. It is the result of gaining insights, experience, and being able to apply information in a meaningful way

Knowledge is the culmination of information and personal understanding, allowing individuals to make informed judgments and take effective action

Examples of knowledge:

- Recommend product for users based on their selling behaviour

Data Information and Knowledge



Data Types

02

Open Data
Vs
Public Data
Vs
Private data

01

Structured data
VS
Unstructured data

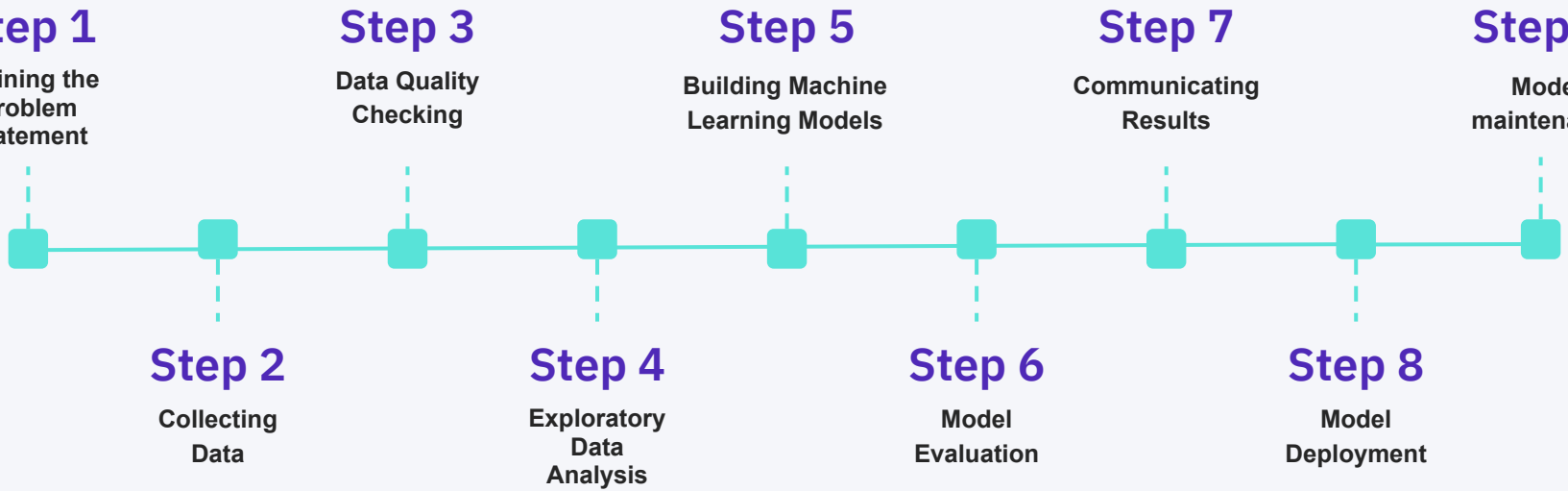
Structured data VS Unstructured data:

- **Structured data** is data that fits neatly into data tables and includes discrete data types such as numbers, short text, and dates. **Unstructured data** doesn't fit neatly into a data table because its size or nature: for example, audio and video files and large text documents.
- **Structured data** is often stored in data warehouses, while **unstructured data** is stored in data lakes.
- **Structured data** is easy to search and analyze, while **unstructured data** requires more work to process and understand.
- **Structured data** exists in predefined formats, while **unstructured data** is in a variety of formats.

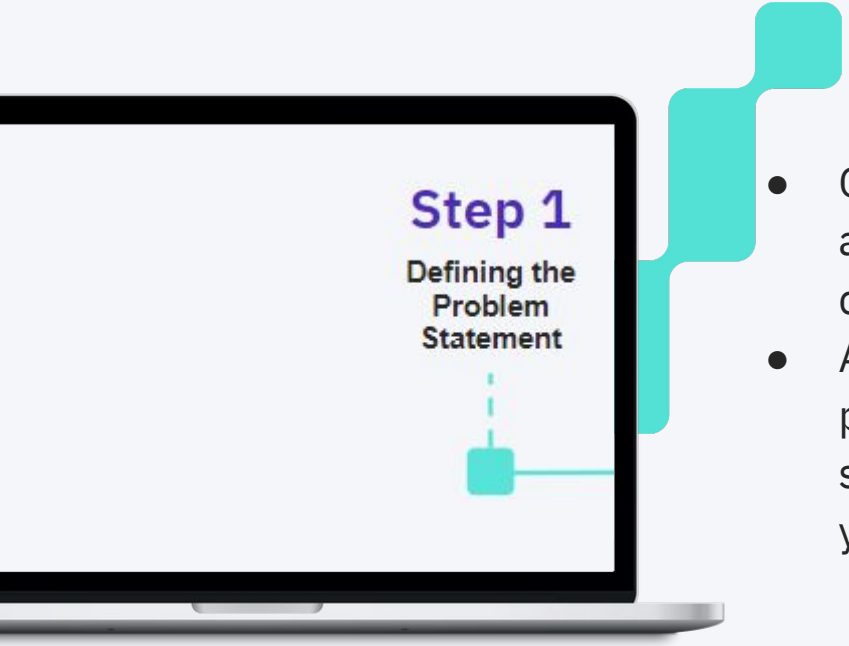
Open Data Vs Public Data Vs Private data:

- **Open data** is data that is available for everyone to access, use and share. It is generally published by governments on freely accessible portals and might include information about local areas, or statistics on topics such as the economy, health, and the environment.
- **Public data** is the data that exists everywhere else. This is information that's freely available (but not really accessible) on the web. It is frequently unstructured and unruly, and its usage requirements are often vague.
- **Private data** refers to any information that is personally identifiable or sensitive and protected

Data Science Life Cycle:



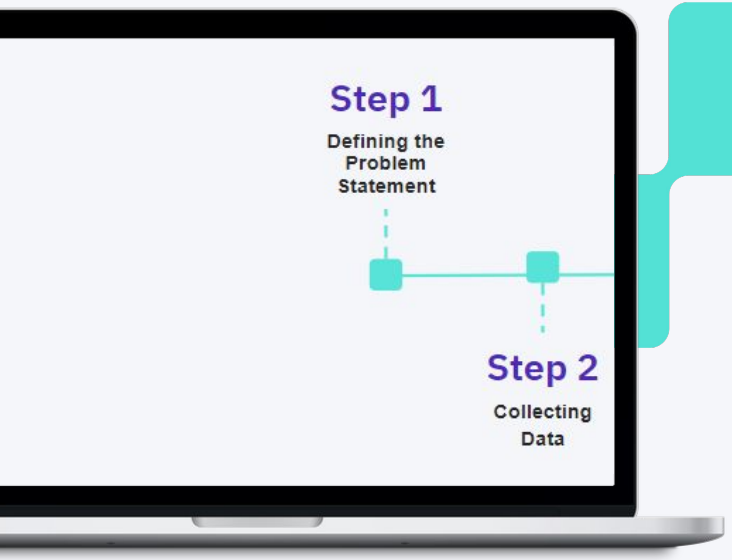
Data Science Life Cycle:



- Creating a well-defined problem statement is a first and critical step in data science. It is a brief description of the problem that you are going to solve.
- Also, all the efforts and work you do after defining the problem statement is to solve it. The problem statement is shared by your client. Your client can be your boss, colleague or it can be your personal project.

Step 1

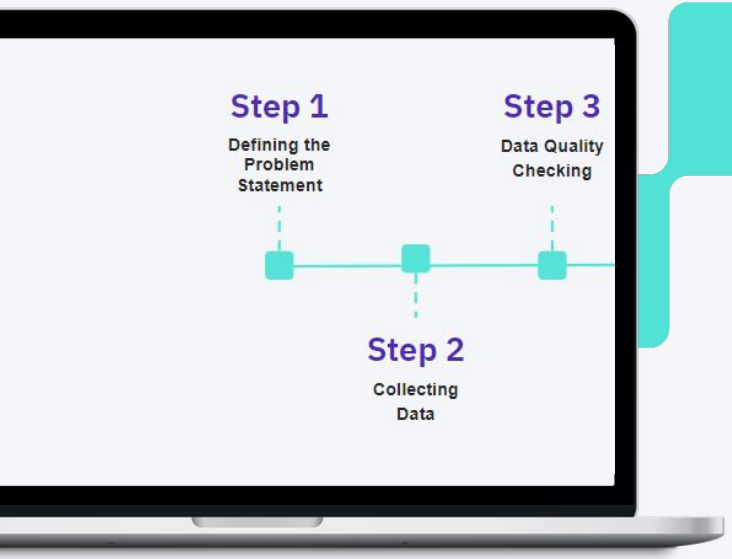
Data Science Life Cycle:



Step 2

- You need to collect the data which can help to solve the problem. Data collection is a systematic approach to gather relevant information from a variety of sources. Depending on the problem statement, the data collection method is broadly classified into two categories.
- First, when you have some unique problem and no related research is done on the subject. Then, you need to collect new data.
- Another method is to use the data which is readily available or collected by someone else. These data can be found on the internet, news articles, government census and so on.

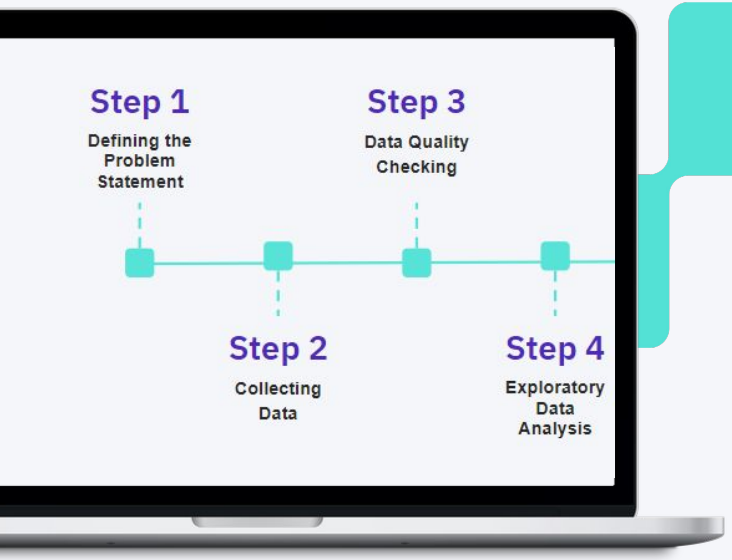
Data Science Life Cycle:



- One of the most important and often ignored aspects by data scientists is ensuring the data that is used for analysis and interpretation is of good quality.
- After collecting the data, most people start the analysis on it. Often, they forgot to do a sanity check on the data. If the data is of bad quality, it can give misleading information. Simply said: “Garbage in, garbage Out”.

Step 3

Data Science Life Cycle:



- Before you model the steps to arrive at a solution, it's important to analyse the data. It is the most exciting step as it helps you to build familiarity with the data and extract useful insights. If you skip this step then you might end up generating inaccurate models and choosing the insignificant variables in your model.

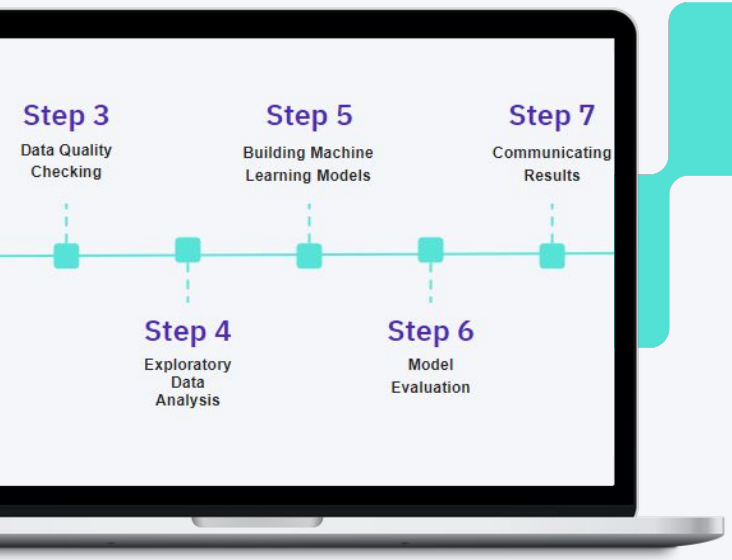
Step 4

Data Science Life Cycle:

- ML Modeling means formulating every step and gather the techniques required to achieve the solution.

Step 5 & 6

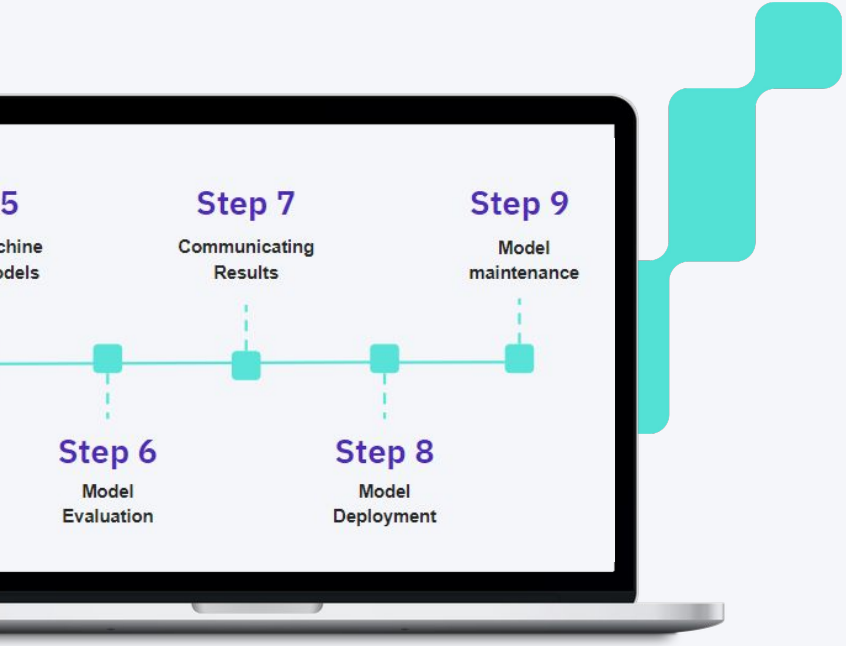
Data Science Life Cycle:



Step 7

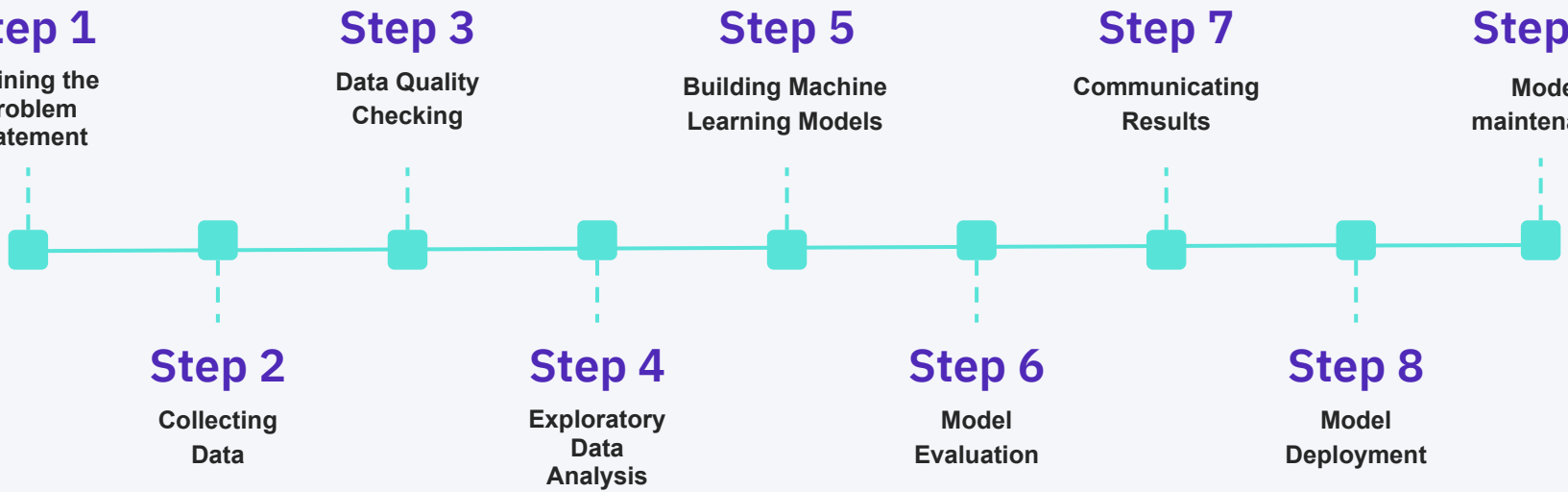
- This is the final step where you present the results from your analysis to the stakeholders. You explain to them how you came to a specific conclusion and your critical findings.
- Most often you need to present your findings to a non-technical audience, such as the marketing team or business executives. You need to communicate the results in a simple to understand manner. And the stakeholders should be able to chalk out an actionable plan from it.

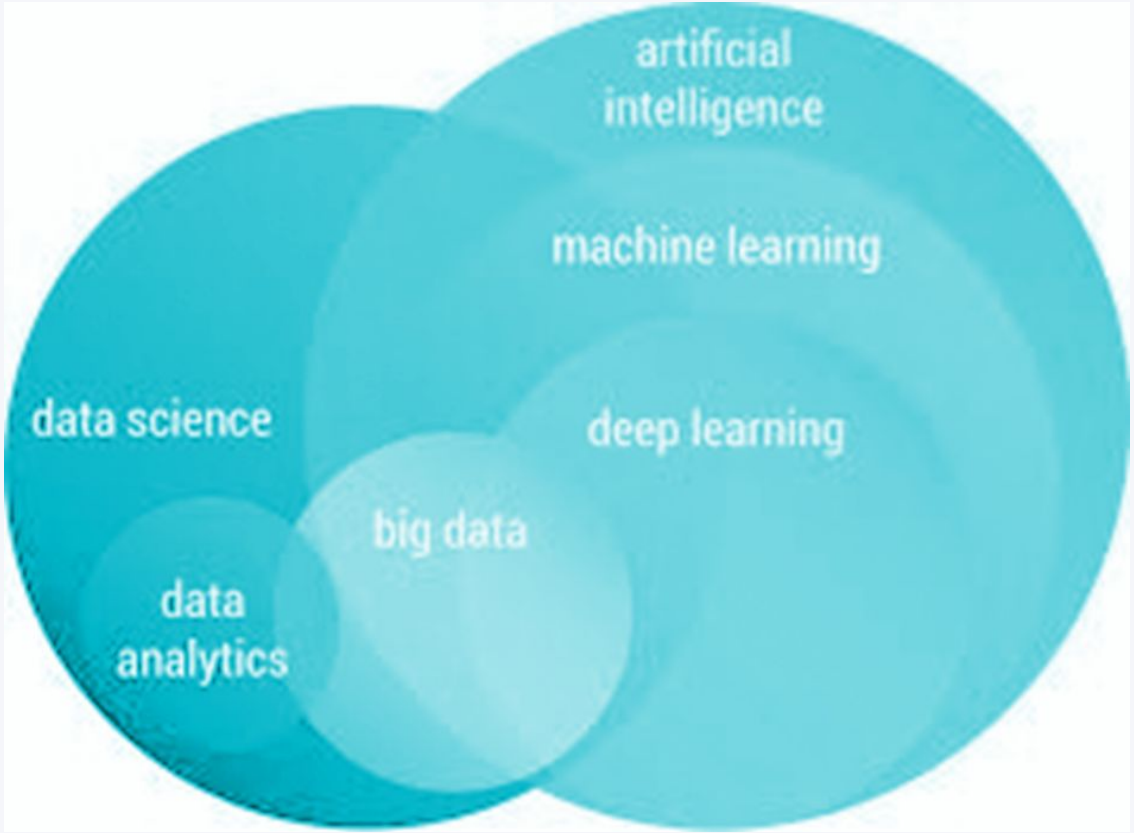
Data Science Life Cycle:



Step 8 & 9

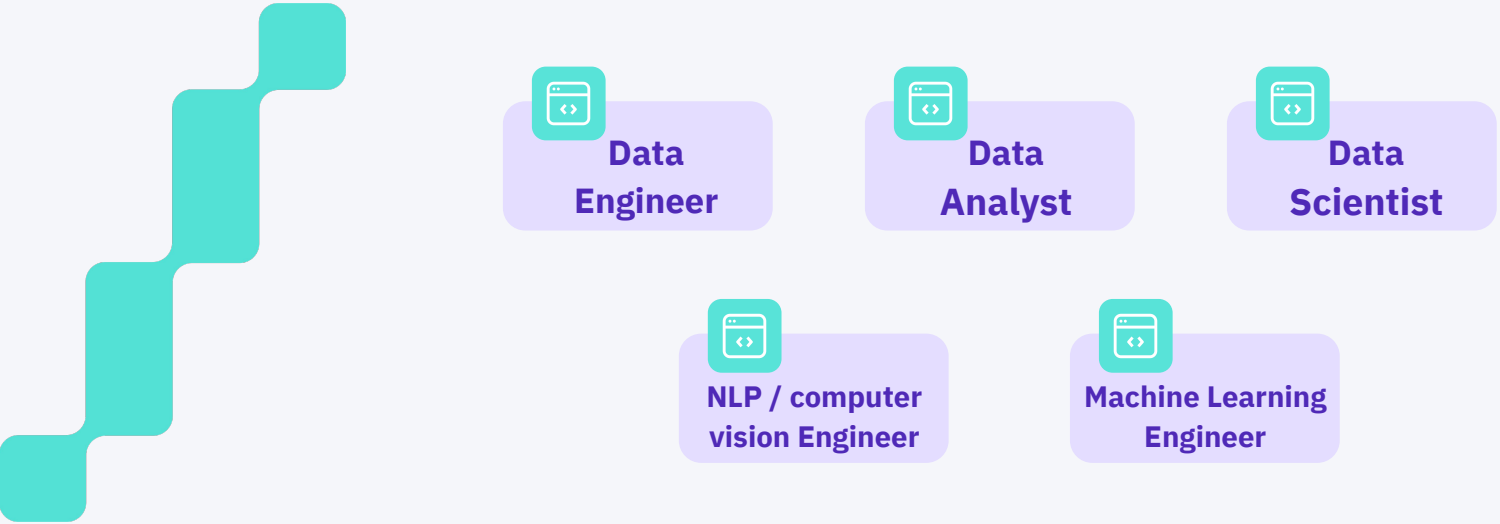
Data Science Life Cycle:




artificial
intelligence
machine learning
data science
deep learning
big data
data
analytics

Job Titles for Data Professionals depend on:

- Required Skill Set
- Type of Problems Solved
- Types of Data Utilized



Who else ?Who Oversees the Data Science Process?

Business Managers: Their primary responsibility is to collaborate with the data science team to characterise the problem and establish an analytical method.

Data Science Managers: They primarily trace and supervise the working procedures of all data science team members.

IT Managers: They are primarily responsible for developing the infrastructure and architecture to enable data science activities.

Data Governance Manager: who give access to data.

Now, it is your turn.

Take a problem and solve it as a data scientist.

How to launch your data science career (with Python)?

Data science isn't about mastering every advanced topic from the get-go.

Esraa's Recipe to launch your data science career and progress

PHASE 1

Discover data science basics and how it enable different applications using data

- **Get comfortable with Python**
- **Understand data handling techniques.**
- **Learn machine learning**

Esraa's Recipe to launch your data science career and progress

PHASE 2

Understand machine learning in more depth

Esraa’s Recipe to launch your data science career and progress

PHASE 3

Find "the thing" that motivates you to practice what you learned and to learn more

Esraa’s Recipe to launch your data science career and progress

PHASE 4

Practice in your thing (Work on personal data science projects to hone your skills)

Esraa's Recipe to launch your data science career and progress

PHASE 5

Keep learning and refining your expertise to differentiate yourself.

Esraa's Recipe to launch your data science career and progress

PHASE 6

Stay up to date

Example of Last bootcamp capstone:

أكاديمية طويق
Tuwaiq Academy

شـــكـــرًا لكـــم

THANK YOU