Assignment 1

1. Data Identification and Collection Assignment

Objective: identify a real-world problem that can be analyzed using data, outline their objectives, and collect the relevant data.

Task:

- Select an area of interest (e.g., sports, health, economics).

- Define a clear problem statement yourself.

- Identify potential data sources (surveys, databases, datasets, APIs).

- Collect initial data, ensuring to handle privacy and security aspects.

Question 1: Definition and Characteristics

- What are the main differences between structured and unstructured data? Provide two examples for each type of data.

Question 2: Data Observation

- For each of the following types of data, determine if it is structured or unstructured and justify your answer:

    1. Email correspondence between employees about project management.

    2. Database of customer addresses and phone numbers.

    3. Tweets related to a recent product launch.

    4. Sales figures compiled in a monthly Excel spreadsheet.

Question 3: Data Conversion

- Is it possible to convert unstructured data into structured data? Explain the process and provide an example where this conversion might be useful in a business context.