

Usecase 1

By: eng. Esraa Madhi

So we defined data science as: It's the process of asking interesting questions, and then answering those questions using data.

For any **Data project** we will go through these steps:

1. Defining the Problem Statement
 2. Collecting Data
 3. Data Quality Checking and Remediation
 4. Exploratory Data Analysis
 5. Building Machine Learning Models
 6. Model Evaluation
 7. Communicating Results
 8. Model Deployment
 9. Model Performance Maintenance in Production
-

Usecase 1

Step 1: Defining the Problem Statement

We are eager to learn about the Riyadh Marathon. We would like to know:

1. In the Kingdom, which gender shows more interest in participating in the marathon, and which is more likely to secure the top positions?
2. What was the fastest time recorded for a participant to reach the finish line?
3. Is the event considered international? Can you provide an estimate of the proportion of participants from other countries?"
4. Please begin your analysis, and don't hesitate to consider additional relevant questions.

Step 2: Collecting Data

Marathon data is not collected for 2023 or [2024](#) (it needs scrapping), the most updated available data is [data at 2022](#)

Step 3: Data Quality Checking and Remediation

To do in the notebook

Step 4: Exploratory Data Analysis

To do in the notebook

Step 5: Building Machine Learning Models

Not applicable

Step 6: Model Evaluation

Not applicable

Step 7: Communicating Results

To do in the notebook

Step 8: Model Deployment

Not applicable

Step 9 : Model Performance Maintenance in Production

Not applicable