**Project Overview 🡪 Olympic Dataset**

The dataset contains information about athletes who participated in the Olympics, including:

- ID: Unique identifier for each athlete.

- Name: Athlete's name.

- Sex: Athlete's gender.

- Age: Athlete's age.

- Height: Athlete's height.

- Weight: Athlete's weight.

- Team: Team the athlete represents.

- NOC: National Olympic Committee code.

- Games: Specific Olympic games.

- Year: Year of the event.

- Season: Season (Summer or Winter).

- City: Host city.

- Sport: Sport category.

- Event: Specific event within the sport.

- Medal: Medal won (if any).

**Basic Questions**

1. Data Exploration: How many rows and columns does the dataset have?

2. Unique Values: How many unique sports are there in the dataset?

3. Gender Distribution: What is the distribution of male and female athletes?

4. Age Statistics: What are the minimum, maximum, and average ages of the athletes?

5. Medal Count: How many medals were won in total? Break it down by Gold, Silver, and Bronze.

6. Top Teams: Which team has the most athletes represented?

7. Missing Data: Which columns have missing values, and how many?

**Intermediate Questions**

8. Height and Weight Distribution: What is the distribution of heights and weights of the athletes? Use histograms for visualization.

9. Age Distribution by Gender: Plot the age distribution for male and female athletes.

10. Medals by Country: Which country has won the most medals? Create a bar chart.

11. Height vs. Weight: Create a scatter plot to show the relationship between height and weight.

12. Age vs. Medals: Is there a correlation between the age of athletes and the number of medals won?

13. Participation Over Time: How has the number of participating athletes changed over the years?

14. Top Sports: Which sports have the most athletes? Create a bar chart.

15. City Hosting: How many times has each city hosted the Olympics?

**Advanced Questions**

16. Medals by Season: Compare the number of medals won in Summer vs. Winter Olympics.

17. Team Performance: Analyse the performance of top teams over different years.

18. Height and Weight Over Time: How have the average height and weight of athletes changed over the years?

19. Medal Predictions: Can you predict the likelihood of winning a medal based on age, height, weight, and gender using a logistic regression model?

20. Clustering Sports: Use clustering techniques (like K-Means) to group sports based on the physical attributes (height, weight, age) of their athletes.