**Group 1: Regular Data Science Questions**

1. What is the average height of monuments built before 1900?

149.66 meters

1. Calculate the median visitor count for monuments made of 'Marble'.

6,000,000 visitors

1. Find the range (difference between the maximum and minimum) of entry fees for 'Religious' type monuments.

$0.00 (no variation in fees)

**Group 2: Multistep Hard Data Science Questions**

1. Group the monuments by 'Material' and calculate the average number of visitors and the standard deviation of height for each group. Which material has the highest average visitors?

Stone

1. For each 'Type' of monument, determine the proportion of total visitors it brings compared to the overall visitors. Which 'Type' attracts the most visitors proportionally?

Wall

1. Identify the correlation between monument height and visitor count. Are taller monuments attracting more visitors?

0.027 (very low correlation)

1. Create a new column 'Age' representing the age of the monument from its 'Year\_Built' to 2020. Then, find the average age of monuments for each 'Type'.

Icon: 131 years., Mausoleum: 367 years., Memorial: 55 years., National Symbol: 134 years., Religious: 89 years.

1. Divide the dataset into two categories based on the median 'Fee'. For each category, find the average 'Height' and 'Visitors'.

Low Fee: Average Height - 70.87 meters, Average Visitors - 3,259,259.; High Fee: Average Height - 117.23 meters, Average Visitors - 7,245,283.

1. Determine the top 3 most common materials used for constructing monuments and calculate the mean fee for these materials.

Copper: $12.00., Steel: $18.00., Stone: $19.00.

**Group 3: Multistep Data Analysis and Machine Learning Questions**

1. Using clustering algorithms like K-Means, cluster the monuments based on 'Height' and 'Visitors'. Determine the optimal number of clusters.

8 clusters

1. Predict the 'Fee' for a monument based on 'Height', 'Visitors', and 'Year\_Built' using a linear regression model. What is the R-squared value of the model?

0.8978

1. Apply a decision tree classifier to predict the 'Type' of monument using 'Height', 'Visitors', 'Fee', and 'Material'. What is the accuracy of the model?

100%

1. Use a random forest classifier to predict whether a monument's visitor count is above or below the median. Determine the feature importance in the model.

Height: 14.97%, Fee: 58.69%, Year Built: 26.34%

1. Employ a time series forecasting model to predict the number of visitors for the next year based on the historical data, assuming a 5% annual increase in visitors.

5,495,327 visitors

1. Implement anomaly detection using Isolation Forest to identify monuments with unusually high fees. What characteristics do these monuments share?

Monument\_ID: Varied (mean: 53)., Height: Average 79.25 meters., Year Built: Average 1835.95., Visitors: Average 5,594,340., Fee: Average $13.72.