**Group 1: Regular Data Science Questions**

1. Calculate the mean and median Price of artworks.

Mean: $7,261.11; Median: $6,800,000

1. What is the most common Medium used across all artworks?

Oil on Canvas

1. Determine the total number of Visitors for artworks exhibited in the Cubism genre.

13,500 visitors

**Group 2: Multiple Step Hard Data Science Questions**

1. For each Genre, calculate the average Rating. Which Genre has the highest average Rating?

Renaissance: 4.85

1. Identify artworks with a Year of creation after 1950. What is the average Price and total number of Visitors for these artworks?

Average Price is 0 and Total Number of Visitors is 0 because none of the artworks were created after 1950

1. Calculate the ratio of exhibited to non-exhibited artworks. Which Medium has the highest ratio?

Dataset does not contain enough data to answer this question.

1. Identify the top 3 Artists based on the average Price of their artworks. What is the total Price for artworks by these Artists?

Leonardo da Vinci: Average Price $11,000; Sandro Botticelli: Average Price $9,800; Unknown: Average Price $9,200; The total price for artworks by these artists is $83,200.

1. For artworks with a Rating of 4.0 and above, what is the average Year of creation?

1850.86

1. Classify the Artworks based on their Price into three categories: Low (below 3000), Medium (3000-7000), and High (above 7000). What is the average Rating for each category?

Can not calculate average for Low. Medium: 4.12; High: 4.65

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

1. Predict the Price of an artwork based on its Year of creation and Visitors count. What is the predicted Price for an artwork created in 1980 with 3000 Visitors?

$8,463.53

1. Predict the number of Visitors based on the Year of creation. What is the predicted number of Visitors for an artwork created in the year 2025?

1,980

1. Using the Artwork Name, predict the length of the Artwork Name for the next artwork. What is the predicted length for an artwork named "Mystical Sunrise"?

16 characters

1. Use Support Vector Regression to predict the Visitors count based on the Price and Rating of the artwork. What is the predicted number of Visitors for an artwork priced at $7000 with a Rating of 4.5?

2030 visitors

1. Train a neural network to predict the Year of creation of an artwork based on its Price and Visitors count. What is the predicted Year for an artwork priced at $5000 with 2000 Visitors?

1311

1. Perform PCA on the dataset to reduce it to 2 principal components. What is the variance explained by the first principal component?

68.14%