**Group 1: Regular Data Science Questions**

1. What is the average 'Distance from the Equator' for cities in the Southern Hemisphere?

-24.27 degrees

1. What is the range of the 'Average Temperature throughout the Year' for cities in Europe?

12.2 degrees

1. How many unique languages are spoken in cities located more than 40 degrees from the Equator?

17 unique languages

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

1. Group the dataset by 'Continent' and calculate the median 'Distance from the Equator' for each continent. Which continent has the highest median?

Europe

1. For each language, calculate the mean 'Average Temperature throughout the Year'. What is the difference in temperature between the highest and lowest averages?

22.2 degrees Celsius

1. Identify the city with the maximum 'Distance from the Equator' in each hemisphere. What are these cities and their respective distances?

Northern Hemisphere: Helsinki, Finland, with a distance of 60.17 degrees from the Equator; Southern Hemisphere: Auckland, New Zealand, with a distance of -36.85 degrees from the Equator.

1. For cities in the Northern Hemisphere, what is the correlation coefficient between 'Distance from the Equator' and 'Average Temperature throughout the Year'?

-0.925

1. Rank the continents based on the sum of 'Average Temperature throughout the Year' of their cities. What is the ranking order?

Asia, Europe, North America, Africa, South America, Australia

1. For cities that speak English, calculate the average 'Distance from the Equator'. How does this compare to the overall average 'Distance from the Equator'?

Speak English: 16.44, Overall average: 27.42; English speaking cities are closer to the equator

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

1. Using 'Distance from the Equator' as a predictor, build a linear regression model to predict the 'Average Temperature throughout the Year'. What is the predicted temperature for a city 30 degrees from the Equator?

15.32°C

1. Apply Support Vector Regression to predict 'Average Temperature throughout the Year'. What is the predicted temperature for the city with the highest 'Distance from the Equator' in the dataset?

8.49°C

1. Employ Ridge Regression to predict 'Average Temperature throughout the Year' using both 'Distance from the Equator' and a one-hot encoding of 'Continent'. What is the R-squared value of this model?

0.63

1. Use Lasso Regression to predict 'Average Temperature throughout the Year' based on 'Distance from the Equator' and a one-hot encoding of 'Hemisphere'. What is the coefficient value for the 'Southern Hemisphere' in the model?

0.0

1. Apply Elastic Net Regression to predict 'Distance from the Equator' using 'Average Temperature throughout the Year' and a one-hot encoding of 'Language'. Which language's coefficient has the greatest absolute value in the model?

4.16

1. Implement Principal Component Regression to predict 'Average Temperature throughout the Year' using 'Distance from the Equator' and principal components derived from one-hot encoding of 'Continent' and 'Language'. What is the total explained variance ratio of the principal components used in the model?

95.59%