Purpose

This folder outlines the procedure I took to answer exam questions; the codes and specific plots are in the readme's of the respective solution folders, updated separately on github.

Question 1

I initially cleaned and filtered the infant names dataset to include pertinent columns before analysing the longevity of naming trends. I then took out the top 25 names for boys and girls for every year starting in 1990. The Spearman rank connection between the names of each year and the names of the next three years was determined by me. This procedure was carried out once for every gender and year. Ultimately, in order to visualize the correlation trends before and after 1990, I converted the data into a lengthy format for plotting and made a time-series graphic.

Question 2

To compare the musical evolution of Coldplay and Metallica, Spotify data was used to analyze their studio recordings, excluding live performances. Key features like tempo, danceability, and energy were extracted for both bands. A boxplot was generated to compare the tempo of Coldplay and Metallica, revealing Coldplay's preference for moderate tempos and Metallica's inclination towards higher, more energetic tempos. Additionally, a scatter plot was created to compare tempo and danceability, showing Metallica's generally higher danceability scores. These visualizations highlight the distinct musical styles and audience engagement strategies of the two bands.

Question 3

I started by determining which EU members were pertinent in order to assess the financial aid payments made to Ukraine by those nations. After that, I filtered the commitment and financial allocation information for these nations. Subsequently, I combined all of the financial pledges and allotments, ordered the nations according to these totals, and classified them according to a cutoff point of 0.7 billion euros to ascertain the level of active support. Lastly, I used a bar chart to display the data, emphasizing the nations who actively support the initiative in green and those that contribute less in red.

Question 4

In order to evaluate India's Olympic performance in comparison to economies of similar size, the GDP dataset was refined to only include countries with a GDP per capita ranging from $1500 to $2000. A scatter plot was generated to depict the GDP per capita of these countries, with each data point annotated. The medal numbers from the Summer Olympics were condensed and restricted to these specific nations. The data was converted into a longitudinal format and utilized to create a bar plot that displays the number of medals by country and type of medal. India's comparative success was highlighted by visualizing both the GDP per capita and medal counts.

Conclusion

I tried to attempt all questions, but I had issues with question 5. The challenges of uploading data prevented me from proceeding and getting results.