**Module 5 Challenge Summary**

Summarizes data displayed in the DataFrame and implications

The pyber\_data\_df\_new dataframe shows the date, time, city and price for each ride. It also shows the type of city the city is located in(rural, urban or suburban). The dataFrame also starts from “urban” cities and it ends with “rural” cities. The problem in with this dataframe is that the dates are not in order and therefore you are not able to see the trends within the month.

Pyber\_data\_new

A screenshot of a cell phone

Description automatically generated

The pyber\_sum dataframe is also showing the date, time, city type and fare. But the difference between the two is that, pyber\_sum dataframe is showing the date in order starting from January 1.

Pyber\_sum

A screenshot of a cell phone

Description automatically generated

Another dataframe I created is the filtered\_resample dataframe. This dataframe differs by the fact that the date has a one week increment between them. Another difference is that it shows the total fare for the whole week. This dataframe also showing the prices of the different city types on the same date.

Filtered\_Resample

A screenshot of a cell phone

Description automatically generated

✓ ​Summarizes the data displayed in the Multiple-line chart

The line chart displays the filtered\_resample dataframe. The blue line represents the rural cities, the orange are the suburban cities and the green line represents the urban cities. From the graph, it shows, Urban cities have the highest t amount of fares. However, it does not show how many rides took place in weeks. It just indidiates the total amount in one week. It also shows that the middle of February, has the highest fares out of all the four months. In addition, the end of March to the first week of April, has the most expensive fares for the Rural rides.

A close up of a map

Description automatically generated

✓ ​Explains the implication of the data summarized in the DataFrame ​**and**​ the Multiple-line chart

The problems with the chart is that you are not able to see how far each drive was. It only gives you the total price for each drive. Therefore, it is unsure if the fare for each ride in different cities is different or if each ride itself was long/far as in distance.

The filtered\_resample dataframe shows that the Urban city type have the most expensive rides/most rides. It is not surprising that the urban cities have the most people live in urban cities. What all three city types have in common are that the third week of February has most amount of total fares. The graph also shows that the first week of April has the highest total fare for the rural cities.