2.2: Project Planning and Sourcing Data with an API

Katie DallaRosa 9/6/2024

Similar to the example used throughout the Exercise, in this task, you'll plan the concept for your dashboard and enrich it with weather data from NOAA.

Network:ID   GHCND:USW00014732

Directions
Part 1: Make a plan for your dashboard

- Create a strategic dashboard to help the business strategy department create insights for an expansion, using open source data from Citi Bike (2022). Look at the strategic placement of new stations across the city to boost the service supply. Installing stations and maintaining bikes is an expensive endeavor, and nailing the exact spots where the extra bikes are needed will help ensure the investment generates more revenue for the company.

Q1 & Q2. Create a list of the elements you'd like to have on it—it's OK if you want to use the proposed structure in the Exercise, but if you feel strongly about a different format/plot, feel free to approach the research question in your own way. Write down some questions to guide your analysis in a new word-processing document and explain how you intend to visualize the result to answer each of your questions.

Think about why there might be bike shortages in certain places. It's safe to assume that the stations where shortages happen are the most popular. This raises the question:

- What are the most popular stations in the city? A bar chart would best suit our purposes as this variable is categorical.

Let's assume that bike trip numbers aren't the same throughout the year. Perhaps, I can hypothesize that the seasons and months play a role in resulting in more trips in warm weather and fewer trips in colder weather:

- Which are the months with the most trips taken? Is there a weather component at play? It's best to plot time series in a line chart, so I can plot the number of trips for each month as one line and combine it with the average monthly temperature in a second line. This will show whether there's a relationship between the temperature and the number of bikes used.

I can also assume that there are popular routes between stations, e.g., that not only start/end stations are popular but that people also tend to journey between certain stations over others:

- What are the most popular trips between stations? Using a map, I can plot the most common bike trips with aggregation to distinguish one-time trips from recurring ones.

Given the main problem is a lack of demand in certain locations, I could also explore whether the bike stations are evenly distributed. The hypothesis here is that maybe there are locations in the city where some stations have larger gaps than others:

- Are the existing stations evenly distributed? I will use the existing map created from the previous question.

Q9. Push your project into the remote repo and add the link to the word processing document.

https://github.com/katiedallarosa/CitiBike