Dashboard plan document

Overview

Objective

The objective of this project is to analyse the user behaviour of those using the CitiBike bike-sharing platform in New Youk City. The analysis will help the CitiBike business analysis strategy department evaluate the success of the current bike distribution logistics model across the city and identify where expansion opportunities may arise.

Context

As bike-sharing platform take-up has increased since the Covid pandemic, so have distribution issues. Popular stations may have fewer bikes available than required and other stations have too many docked bikes, making the return of a hired bike more difficult.

Our task is to workout the reason for distribution problems and advise CitiBike on how to ensure the correct number of useable bikes are in the best places. The issue may stem from the number of bikes in use, from seasonal demand or another reason entirely.

Data

The data for this dashboard will come from open source data from the 2022 CitiBike database and will be supplemented by weather from NOAA, via its API service. A number of Python libraries will be used, among them Plotly, Matplotlob and Seaborn for the charts, pandas and Kepler.gl for maps and Streamlit for the final dashboard design

Considerations

The dashboard needs to clear and simple to understand, featuring no more than four charts.

We must also be mindful of what data is used on the dashboard. There should be no user-sensitive information that could be used to identify individuals or compromise their security.

Analysis questions

1. Where in the city are the most popular bike stations located?
2. Is demand affected by particular times of the year? Does temperature affect demand?
3. Which bike stations are the most popular?
4. Is demand highest for electric or classic bikes?

Visualisations

1. For a question with a geographic element, a map showing bike stations with graduated symbols depicting where usage was greatest would be an ideal choice.
2. With time series data being used to show demand over the course of the year, the best chart to use would be a line chart. It might be useful to see if temperature affects hire of a particular type of bike, so the visualisation could show two lines, one for each type of bike.
3. A bar chart would be good for a categorical variable such as the 10 most popular stations.
4. A donut chart showing the two choices of bikes for hire (another categorical variable), which would be made even more useful when linked to another element – for example, if the user was viewing the data for a particular station or month.