NY Citi Bike Share: Analyzing User Behavior in Jersey City

Project Overview

Project Context

You're the lead analyst for a NYC-based bike-sharing service that also operates in Jersey City. With increased demand in JC, there have been ensuing issues with bike distribution and logistics. The task is to conduct a descriptive analysis of existing data to help the business strategy team with these issues by providing actionable insights and recommendations.

Project Objectives

- Analyze User Behavior in Jersey City
- Identify Cause(s) of Bike Distribution Problems
- Assess Current Logistics Model of Bike Distribution in JC
- Identify Expansion Opportunities
- Provide Recommendations for Further Research
- Create a Dashboard (Streamlit) and Present Findings
  - *Click here to View Final Dashboard

Project Overview

Preliminary Business Questions

1. Which are the months with the most trips taken? Is there a weather component at play?
2. What are the most popular stations in the city?
3. What are the most popular trips between stations?
4. Are the existing stations evenly distributed?
5. Are certain stations more/less popular depending on time of day, week or year?

Analysis Steps

- **Step 1: Source and Prepare Data**
  - Import Bike-Share Data
  - Use API Key to Source Weather Data
  - Clean and Merge Two Dataframes
- **Step 2: Create Initial Visualizations & Record First Insights**
- **Step 3: Initiate Map**
- **Step 4: Recreate Charts for Dashboard (Plotly)**
- **Step 5: Execute Python Script to Create Dashboard**
- **Step 6: Final Presentation**

Project Overview

Skills & Tools Used

- Python (pandas, numpy, streamlit, pillow, kepler.gl, seaborn, matplotlib, pyplot, numerize)
- Data Cleaning, Prep and Wrangling
- Sourcing Data with API Key
- Dashboard Creating
- Geospatial Plotting
- Virtual Environments

Links: Datasets, Presentation & Repository

- **Links to Datasets:**
  - Dataset 1: Open Source Data from Citi Bike's Database:
  - Dataset 2: Weather Data using NOAA's API Service:
- Link to Final Dashboard
- Link to Final Presentation (YouTube)
- Link to GitHub Repository

Insights: Weather/Month and Bike Usage

Peak Bike Usage Time: May-September

- As temperatures plunge, so does bike use
- Three Noticeable drops in Bike Usage:
  - Early-October
  - Mid-November
  - Mid-December
- Individual Days with “outlier” low usage:
  - 1/29, 3/12, 5/7, 9/6, 10/3 & 12/25

Insights: Most Popular Starting Stations

- **Most Popular Station:**
  - 1st. Grove Street
- **Most Popular Stations: "Second Tier"**
  - 2nd. South Waterfront Walkway
  - 3rd. Hoboken Terminal (River St.)
  - 4th. Hoboken Terminal (Hudson St.)
  - 5th. City Hall
- **Most Popular Stations: "Third Tier"**
  - 6th. Newport Pkwy
  - 7th. Hamilton Park
  - 8th. Newport Park

Insights: Most Popular Routes

There are two big "hubs" or centers for bike usage:

- **Hub 1 & Second Tier Stations**
  - Hub 1 sees the highest concentration of bike trips
  - Most popular trips center around the four "Second Tier" stations
- **Hub 2 & Most Popular Station**
  - Hub 2 sees the 2nd highest concentration of bike trips
  - Most popular trips start/end at Grove Street Station
- **Connecting Hubs 1 & 2 with Third Tier Stations**
  - Popular routes also occur between Hubs 1 and 2
  - The three "Third Tier" stations are all located between Hubs 1 and 2

Final Recommendations for Actions and Further Research

- **Scale back number of bikes after October in phases, according to the three points of drops in bike usage. Further statistical analysis needed to determine how much to scale back**
  - Drop 1: Early October
  - Drop 2: Mid-November (More Significant Drop)
  - Drop 3: Mid-December
- **Increase the number of bikes in three phases according to the noticeable points of rises in bike usage. Further statistical analysis recommended**
  - Rise 1: Mid-February
  - Rise 2: Mid-March
  - Rise 3: Mid-April

Final Recommendations for Actions and Further Research

- **Recommendations for Hub 1:**
  - The four most popular stations in this hub ("Second Tier") are all by the waterfront.
  - Additional analysis is required to assess:
    - If these four stations can manage being the central stations of this hub.
    - If these stations should be expanded
    - If additional stations should be established in the area to alleviate congestion at these four locations.
- **Recommendations for Hub 2:**
  - Further analysis is recommended to assess:
    - Can Grove Street can manage being the single "central station" of this hub?
    - Should Grove Street be expanded?
    - Should additional stations be introduced in the area to reduce congestion?
    - Should nearby stations be expanded to relieve potential congestion?

Final Recommendations for Actions and Further Research

- **Further analysis is necessary to ensure that bikes are consistently available at the most popular locations:**
  - **For popular origin stations and times:**
    - Ensure that an adequate supply of bikes is maintained
    - Replenish the stock frequently.
  - **For popular destination stations/times:**
    - Focus on clearing returned bikes quickly to free up docking space
    - Redistribute them to areas with higher demand for departures