NEW YORK CITIBIKE

NEW YORK CITIBIKE USAGE ANALYSIS

TOOLS USED: [PYTHON, JUPYTER, PANDAS, MATPLOTLIB, NUMPY, KEPLER.GL, STREAMLIT]

DATASET:

- NYC CITIBIKE TRIP DATA (API SOURCED)
- WEATHER DATA (NOAA INTEGRATION)

Summary

Analyzed NYC Citi Bike trip data to identify usage patterns, station demand, and geographic distribution. Performed data wrangling, aggregation, and geospatial visualization to uncover high traffic routes and operational gaps. The goal was to support data driven decisions for bike redistribution, station planning, and service optimization.

NEW YORK CITIBIKE

NEW YORK CITIBIKE USAGE ANALYSIS

[PROCESS - ANALYSIS]

KEY STEPS TAKEN

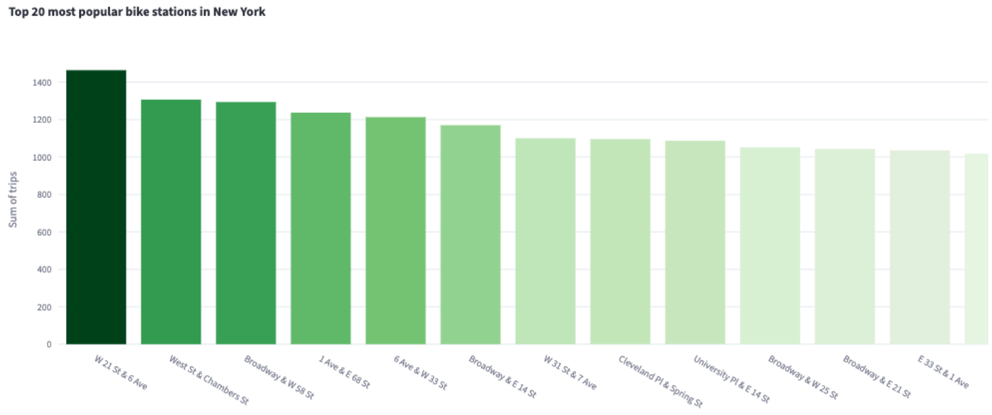
- Retrieved Citi Bike data via API and structured datasets for analysis.
- Cleaned and transformed data (missing values, duplicates, data types).
- Created new variables (trip counts, time features, station pairs).
- Aggregated trips using groupby (start/end stations).
- Integrated geospatial coordinates for mapping and visualization.

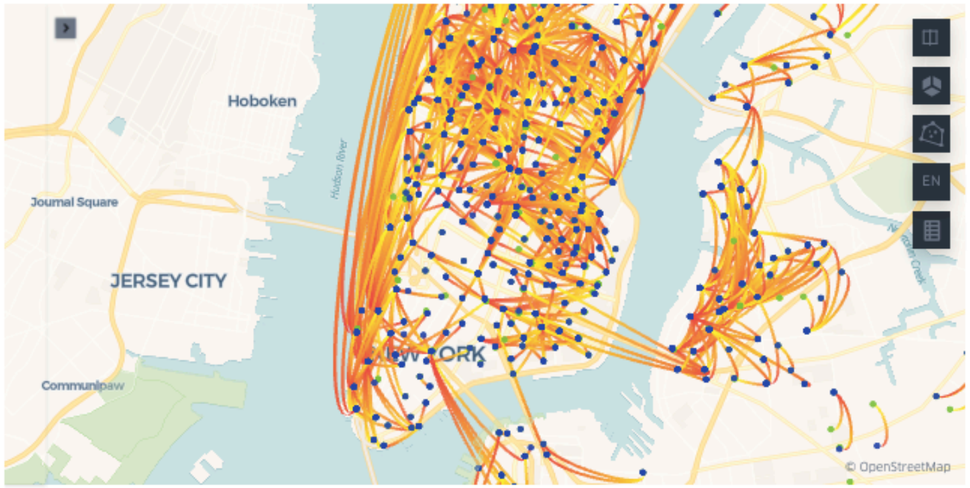
NEW YORK CITIBIKE USAGE ANALYSIS

[RESULTS - INSIGHTS]

Total Bike Rides

339.03K





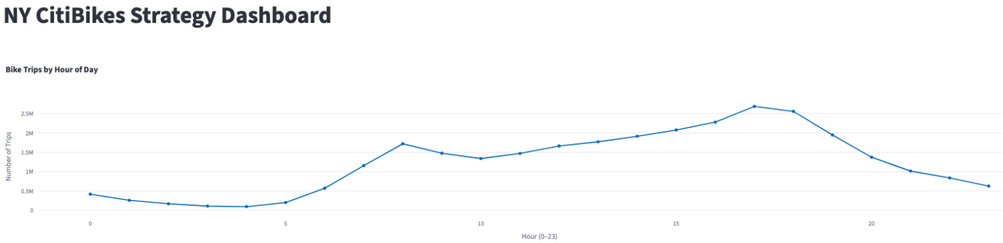
KEY FINDINGS

- High trip density observed in central Manhattan and commuter-heavy zones.
- Peak usage occurs during morning and evening commute hours.
- Popular routes connect business districts and transit hubs.
- Weekday usage exceeds weekends, indicating commuter driven demand.

NEW YORK CITIBIKE USAGE ANALYSIS

[RESULTS - INSIGHTS]

📊 Daily Bike Trips vs Temperature (2022)



⏱ Bike Trips by Hour of Day

KEY FINDINGS

- Certain stations show consistent imbalance (high departures vs arrivals).
- Tourist areas display irregular but high-volume trip patterns.
- Weather conditions impact trip frequency and duration.
- Filtering trips highlights critical high demand routes needing optimization.

NEW YORK CITIBIKE USAGE ANALYSIS

[REFLECTION - LEARNINGS]

This project strengthened my ability to work with real world API data, perform data aggregation, and build geospatial visualizations using Python. I improved my understanding of mapping tools like kepler.gl and how to translate raw data into actionable insights.

If extended, I would develop predictive models for demand forecasting and optimize bike redistribution using real-time and historical patterns.

LINKS

https://elia-nycitibike.streamlit.app

https://citibike.elialanz.com

https://github.com/elialanz/New_York_City_Bike