

WomenTechWomenYes (WTWY)

The Last Data Benders

Agenda

01

The Team

02

Introduction

03

Methadology

04

Conclusion

05

Future Work

01

Team



Halil Kolatan



Tuğçe Ayçetin



Muhammed Maral



Egemen Doğruel

&

Ebrar Ölçer



Introduction

The Project Overview:

In this EDA project, we explore how socio-economic, coffee shops, pubs, restaurants and subway ridership in New York City by using census, restaurant audit results and subway data. Our project supports the **WomanTechWomanYes** initiative to promote women in technology.

Methodology

Data Collection

Gathering relevant data



NYC OpenData

Data Cleaning

Identifying and fixing data issues

pandas



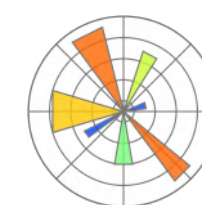
Data Exploration

Understanding data structure and patterns.



Data Visualization

Gathering relevant data



NYC-Census

Data Source: [Kaggle](#)

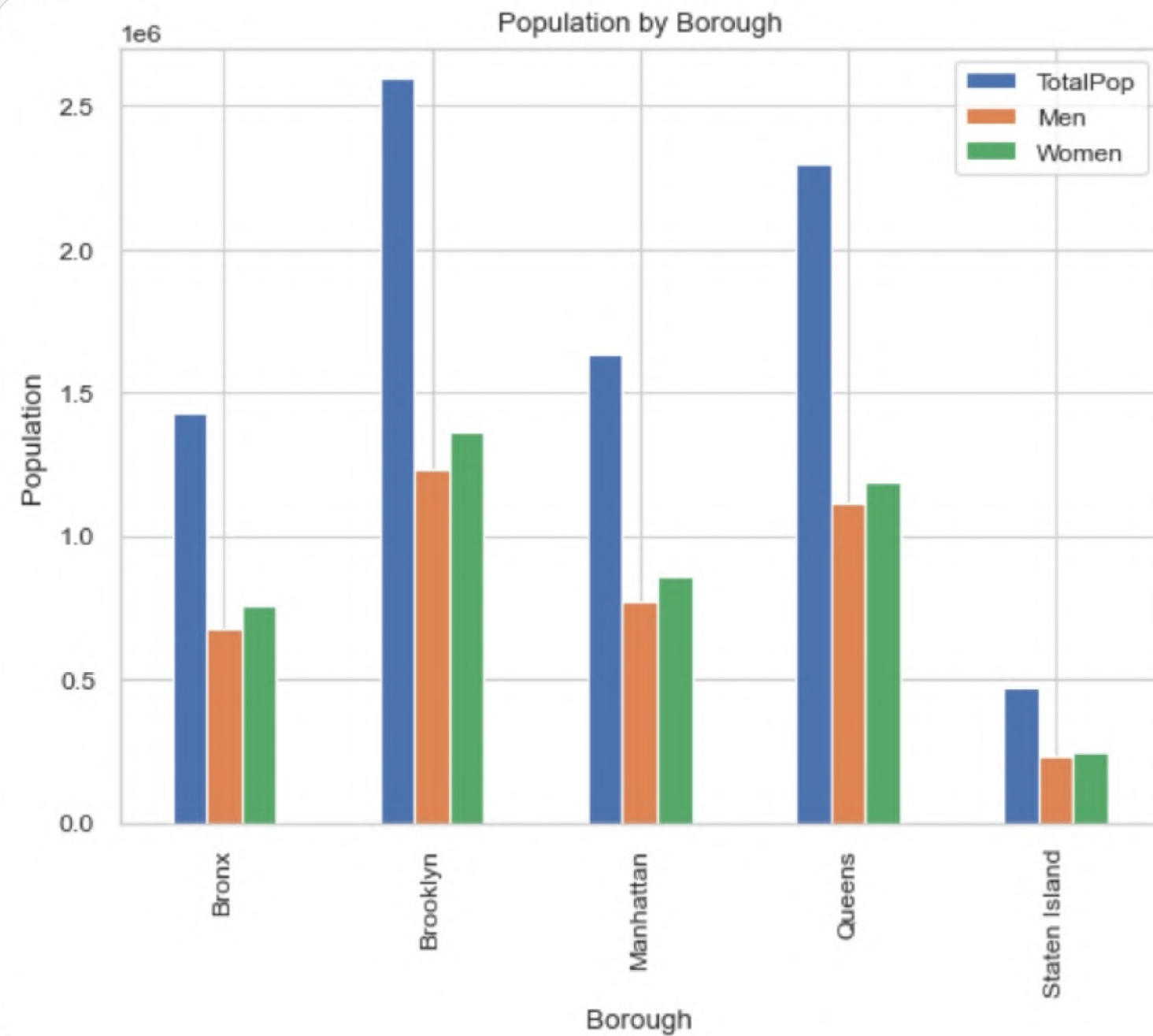
Variables: The important variables among **32 columns** were county, borough, population, race, and income.

Data quality: Race ratios incomplete, income was null **66/ 2167** rows.

Data insights: Visualizations of income, and population distribution based on borough.



Visualization – Population



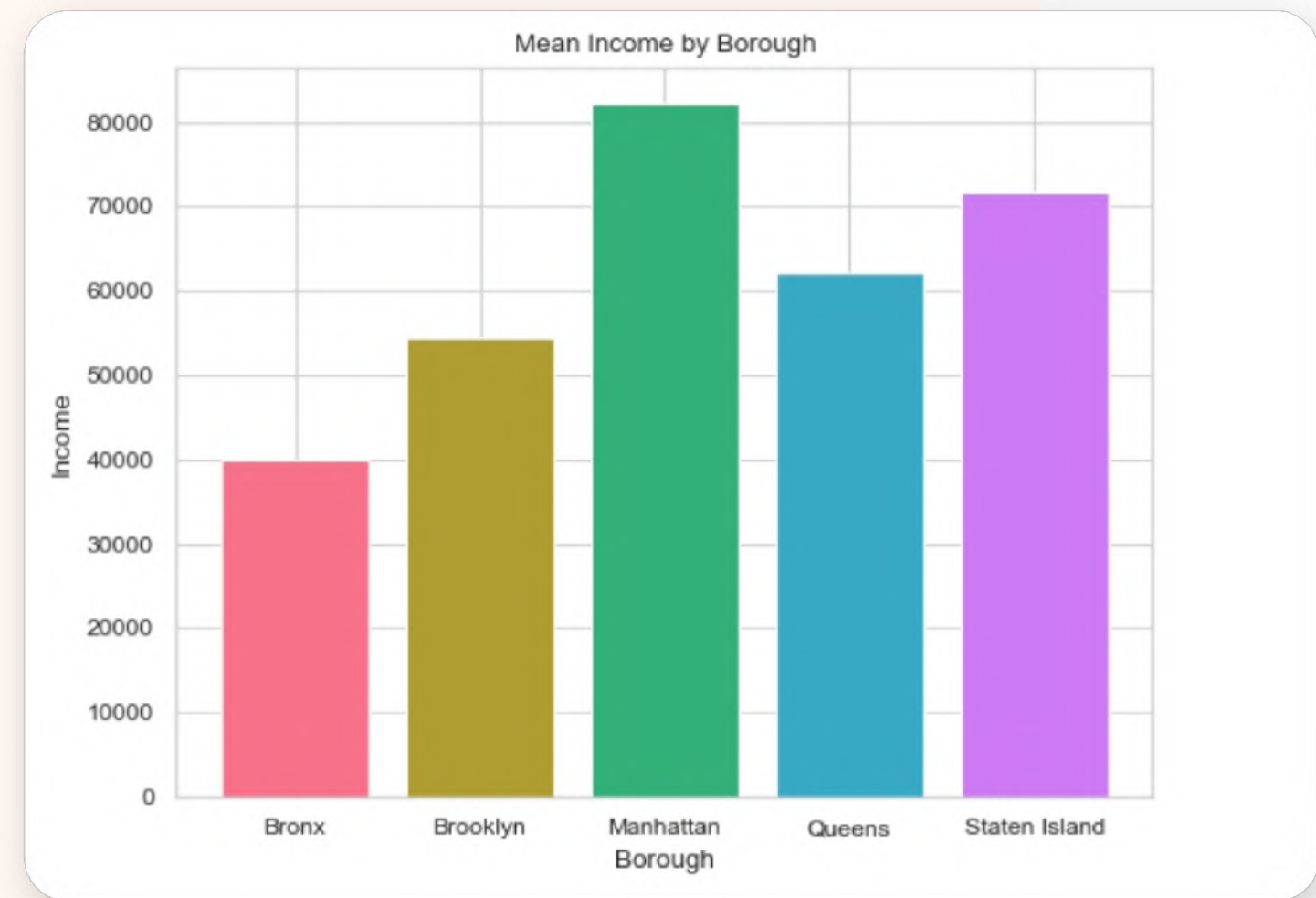
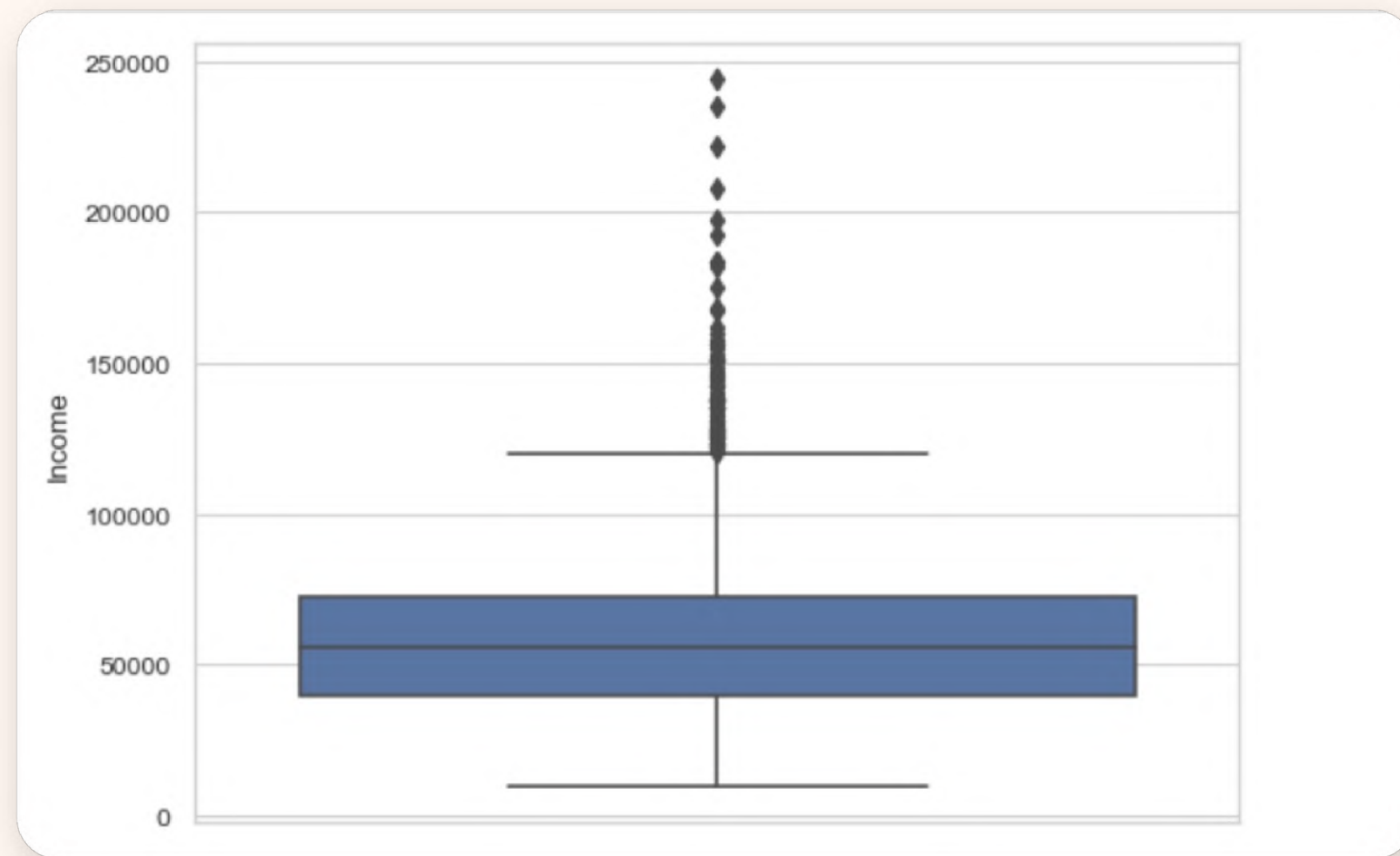
Total Pop

Men

Women

Distributions

Visualization – Income



MTA Turnstile Data

Data Source: [MTA](#)

Variables: Station, time, entries, exits, and **Total Traffic**.

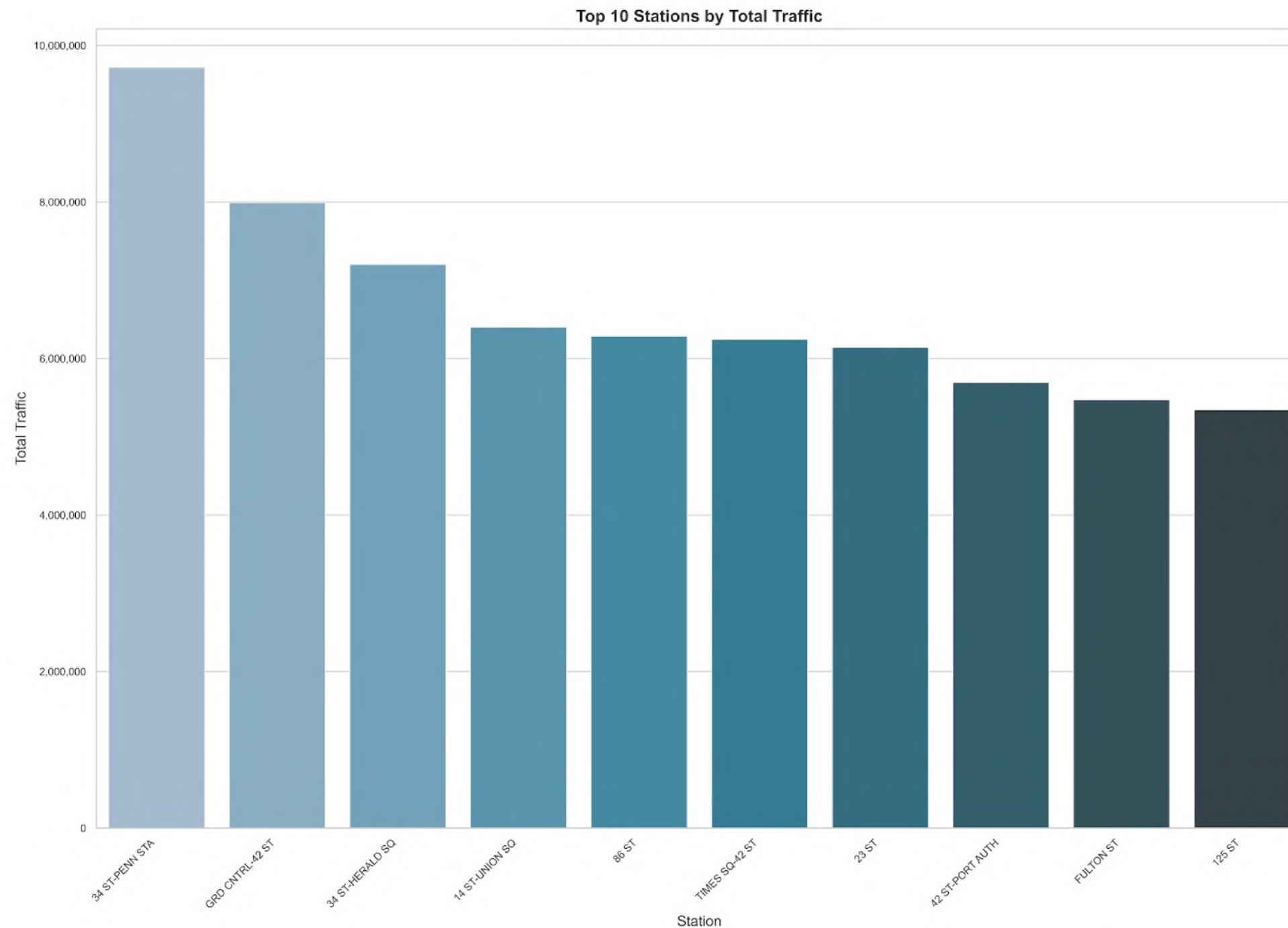
Data quality: Inconsistent time, entry & exit data were summed, **+2M rows of data**.

Data insights:

- A collection of 70 days analyzed
- Top 10 station with Total Traffic
- Busiest Days & Times 34th Pen
- Average number of passengers by time on Wednesdays



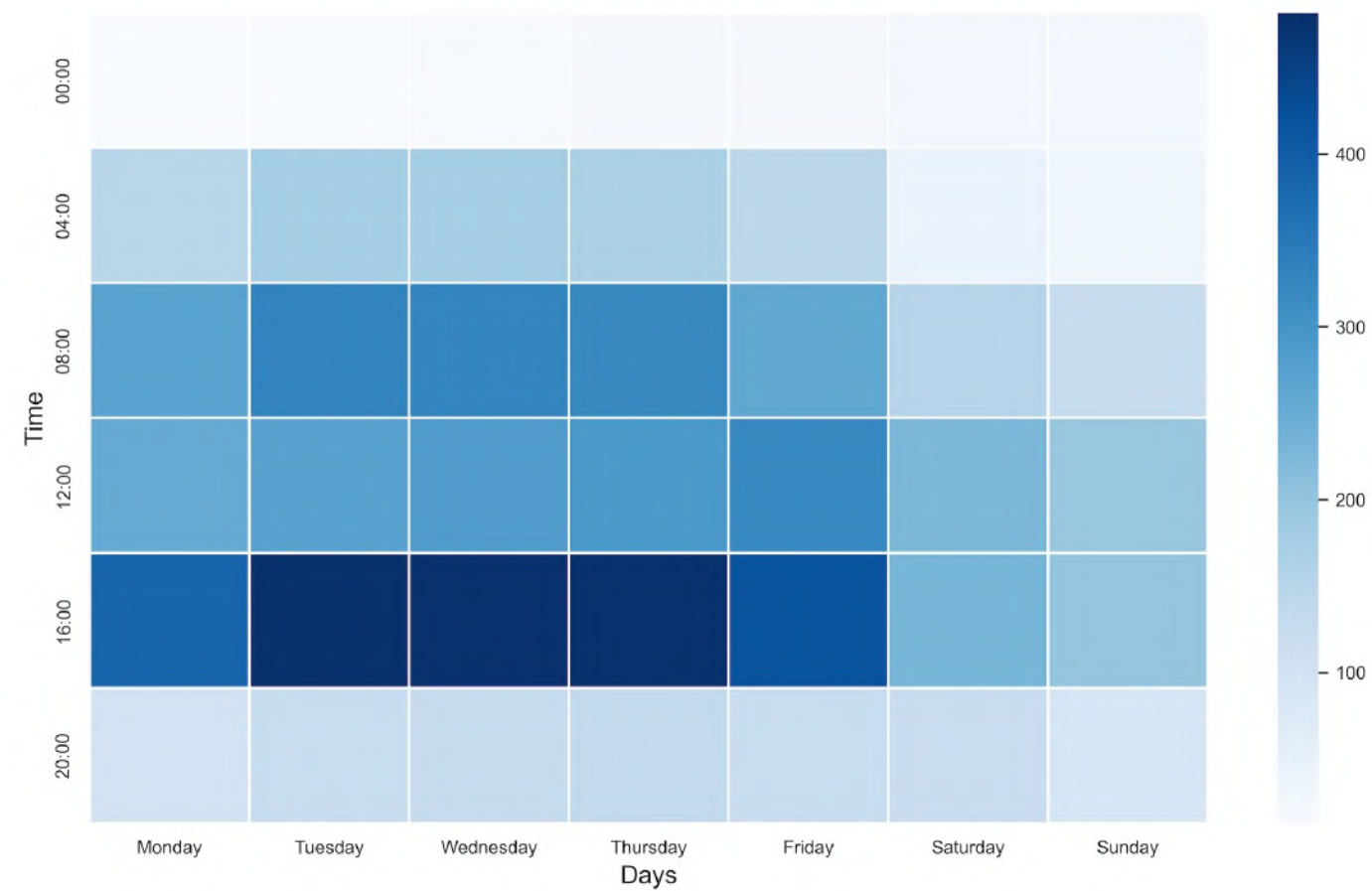
Visualization – Top 10 Stations



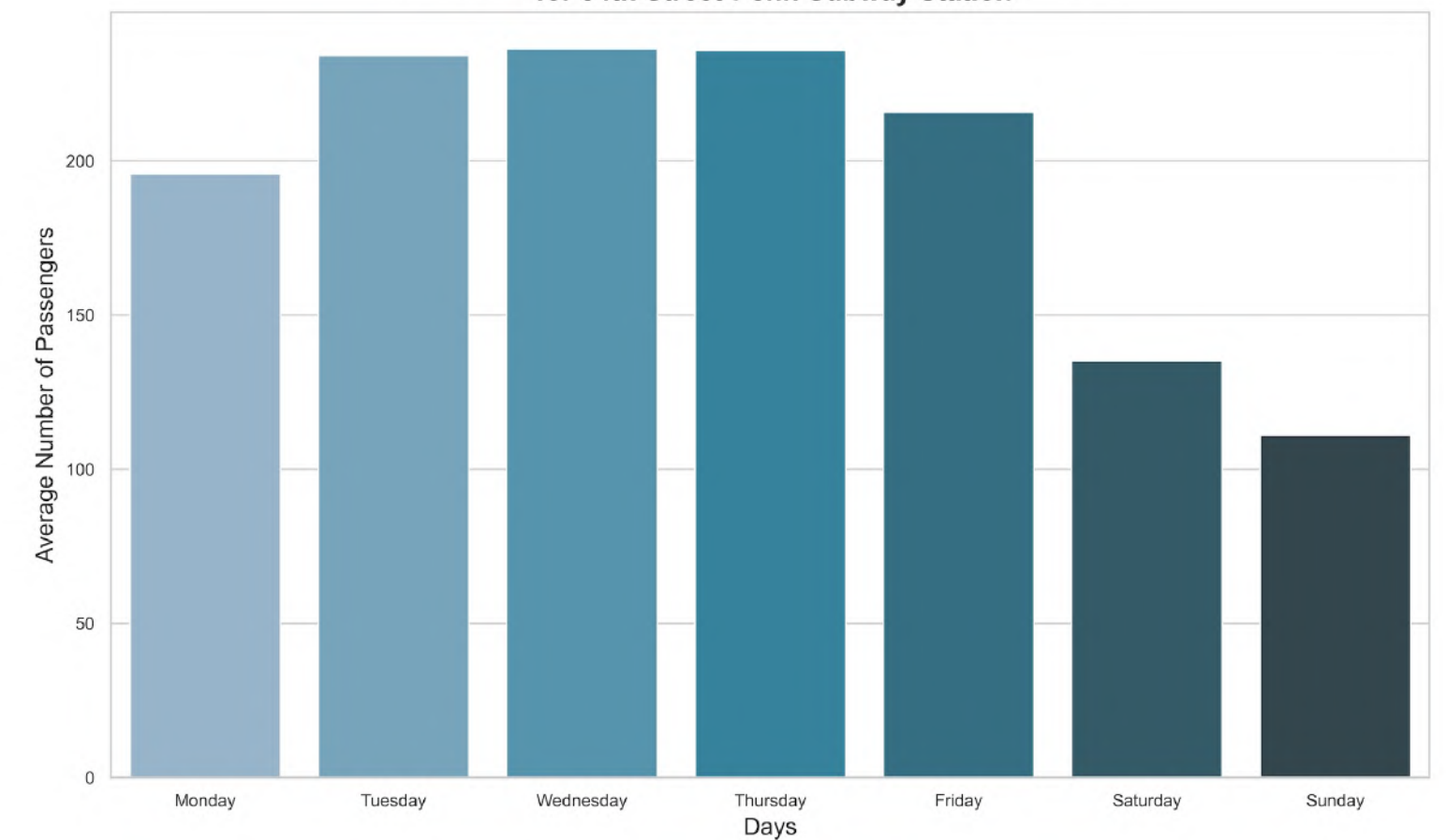
**34-ST-PENN Station
and
GR CNTRL-42 ST**
Were busiest stations

Visualization – Busiest Days and Times

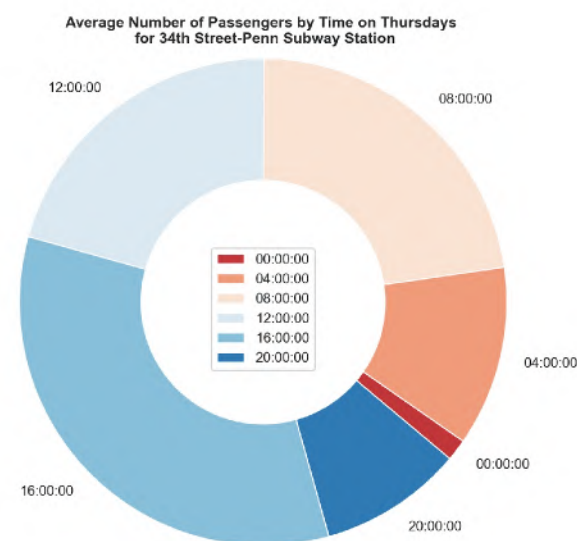
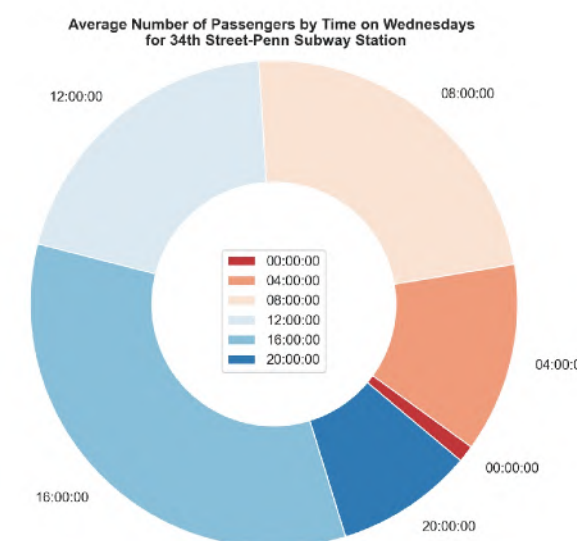
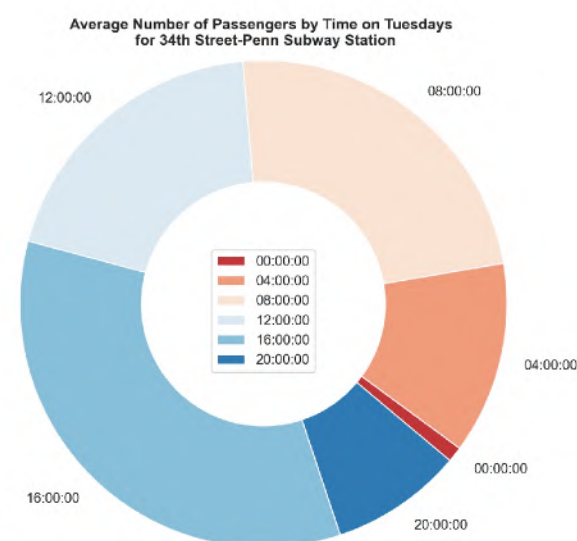
Busiest Days and Times at 34th Street-Penn Station



Average Number of Passengers by Time on Wednesdays for 34th Street-Penn Subway Station

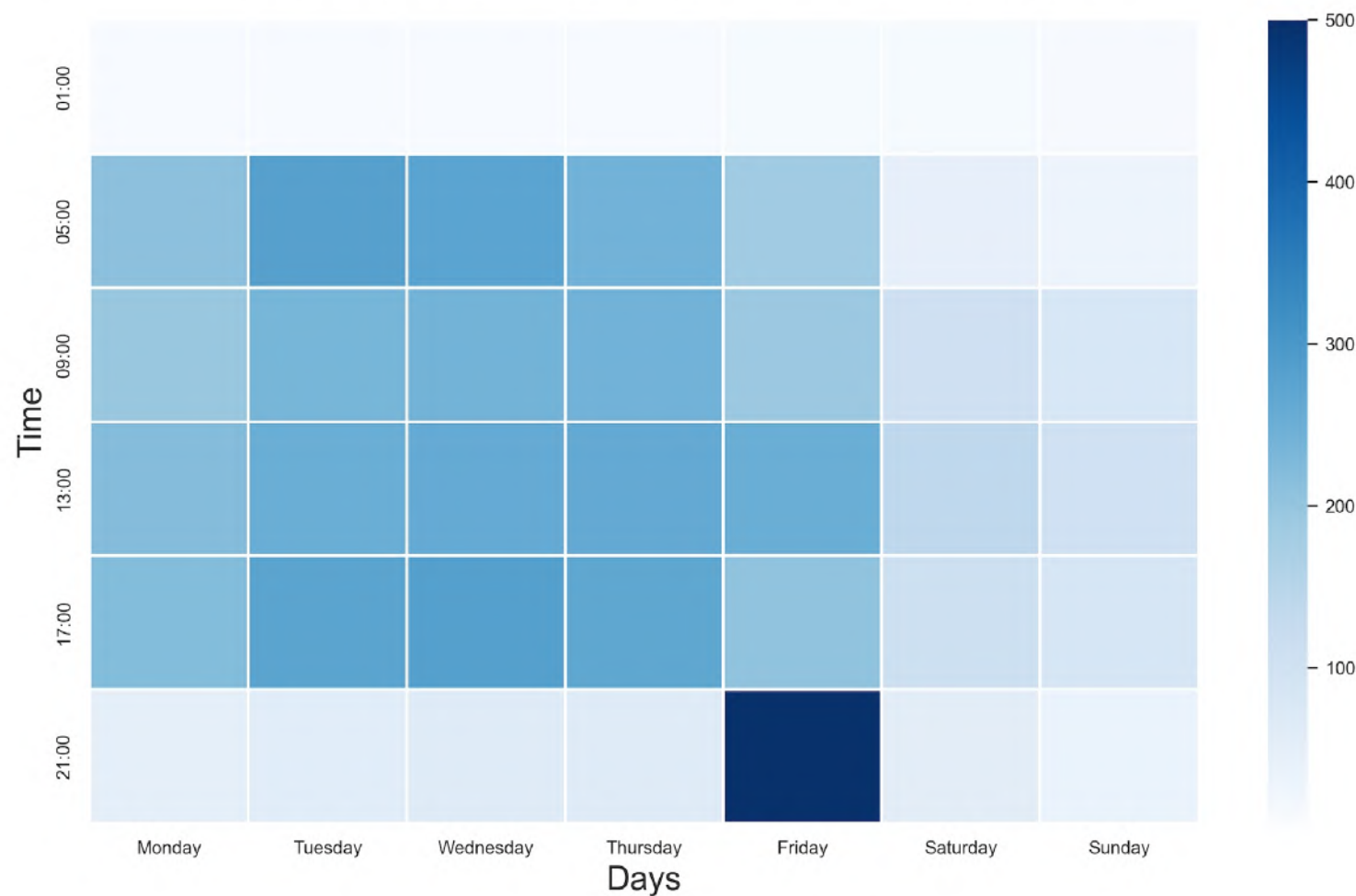


Visualization – Busiest Days and Times

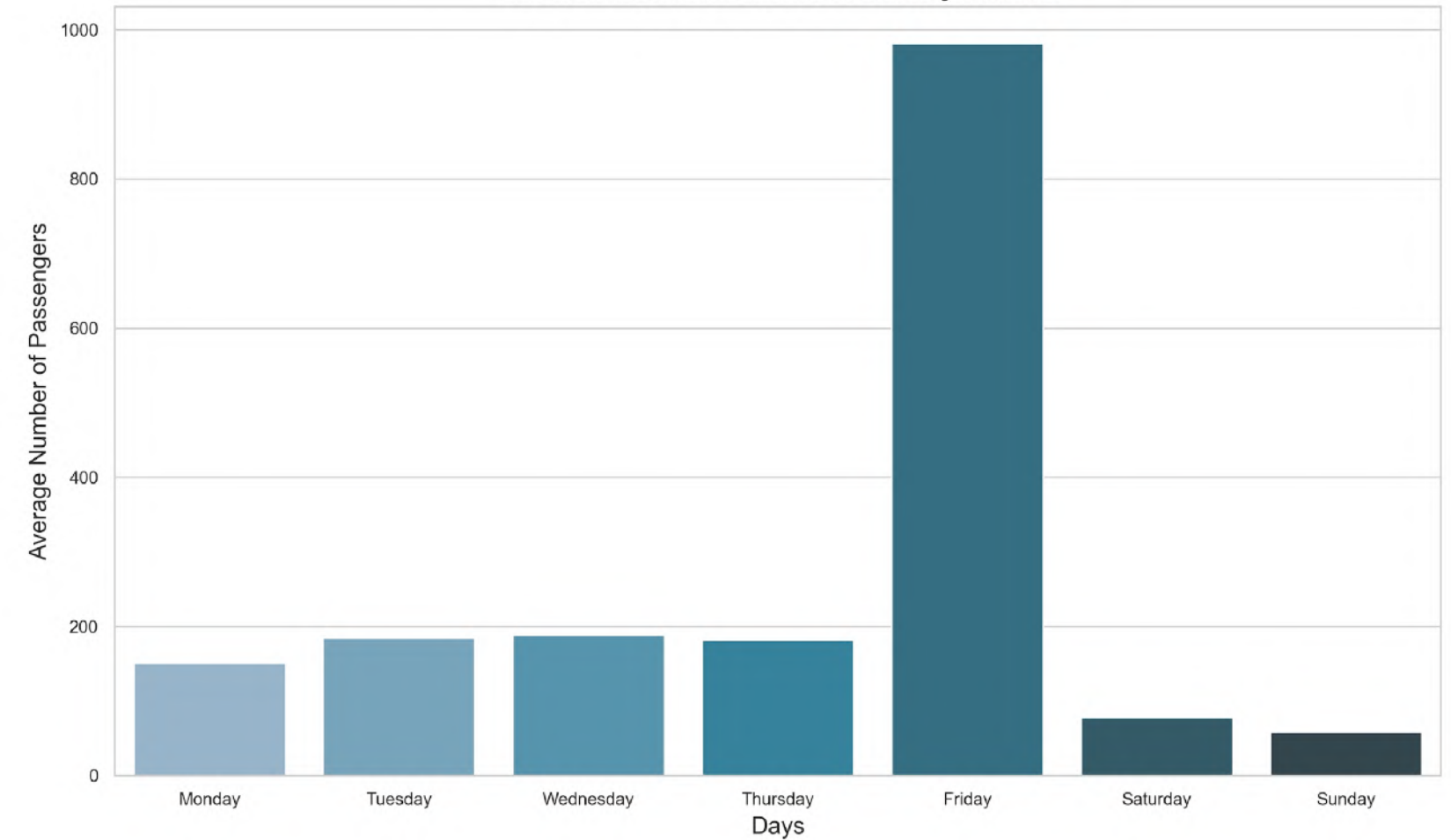


Visualization – Busiest Days and Times Heatmap

Busiest Days and Times at Grand Central 42nd St Subway Station

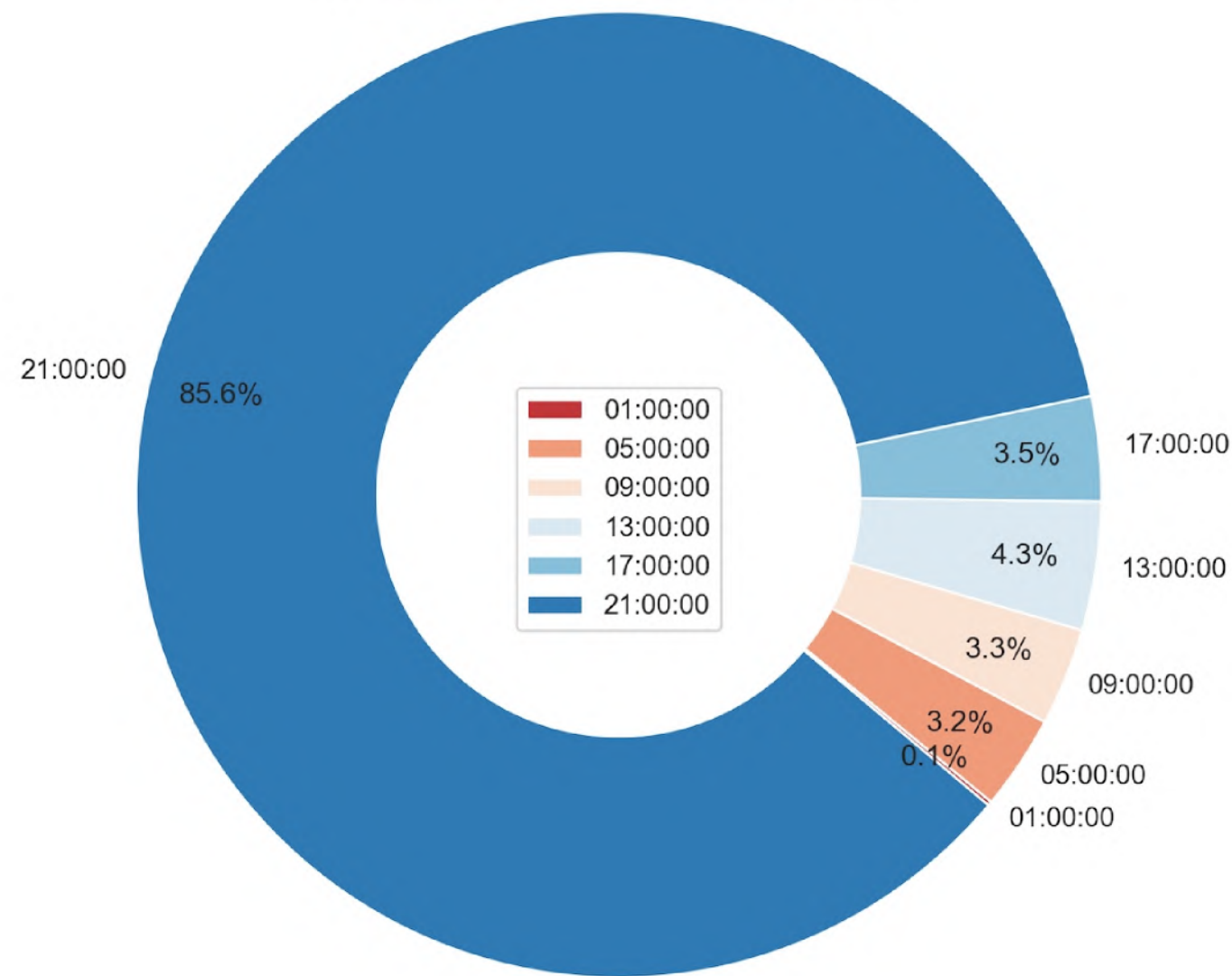


Average Number of Passengers by Time on Fridays for Grand Central 42nd St Subway Station



Visualization – Busiest Day and Time

Average Number of Passengers by Time on Fridays
for Grand Central 42nd St Subway Station



42-GR-CNTR Station
time
on Fridays

21:00:00 Busiest Period

NYC-DOHMH New York City Restaurant Inspection Results

Data Source: [NYC Open Data](#)

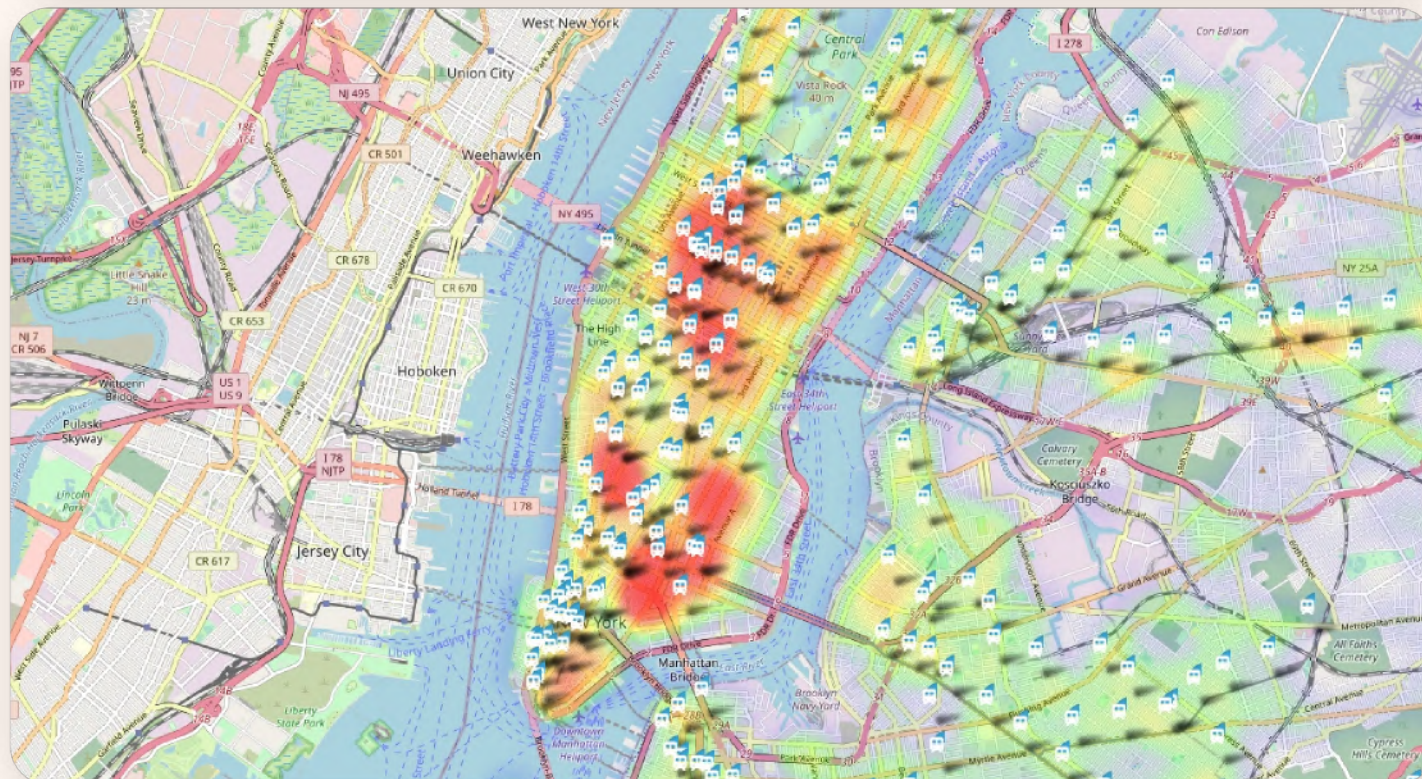
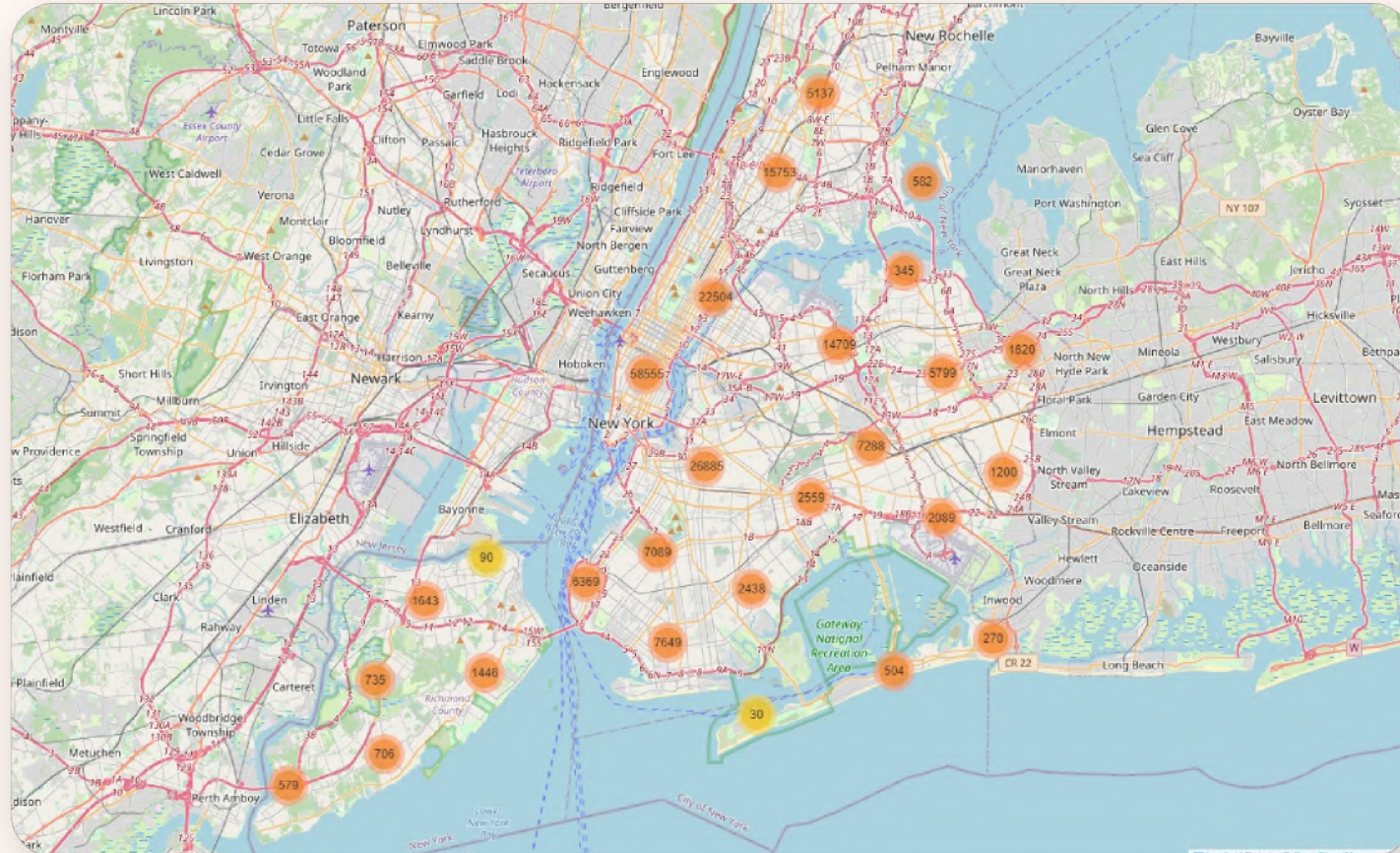
Variables: Restaurant names, locations, types and latitude and longitude, +200k rows of data.

Data quality: Some restaurants have not been audited and therefore are not up to date.

Data insights: Heat maps of restaurant locations.



Mapping with Folium Library





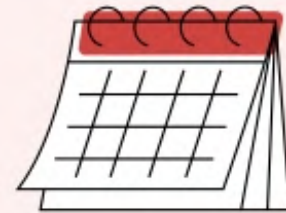
WomenTechWomenYes

CONCLUSION



34th Street-Penn
Subway Station

Wednesday



Busiest Day

16:00 - 20:00

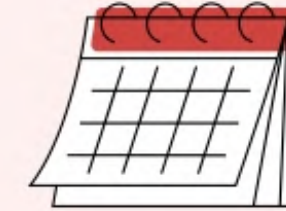


Busiest Time



Grand Central
42nd St Subway
Station

Friday

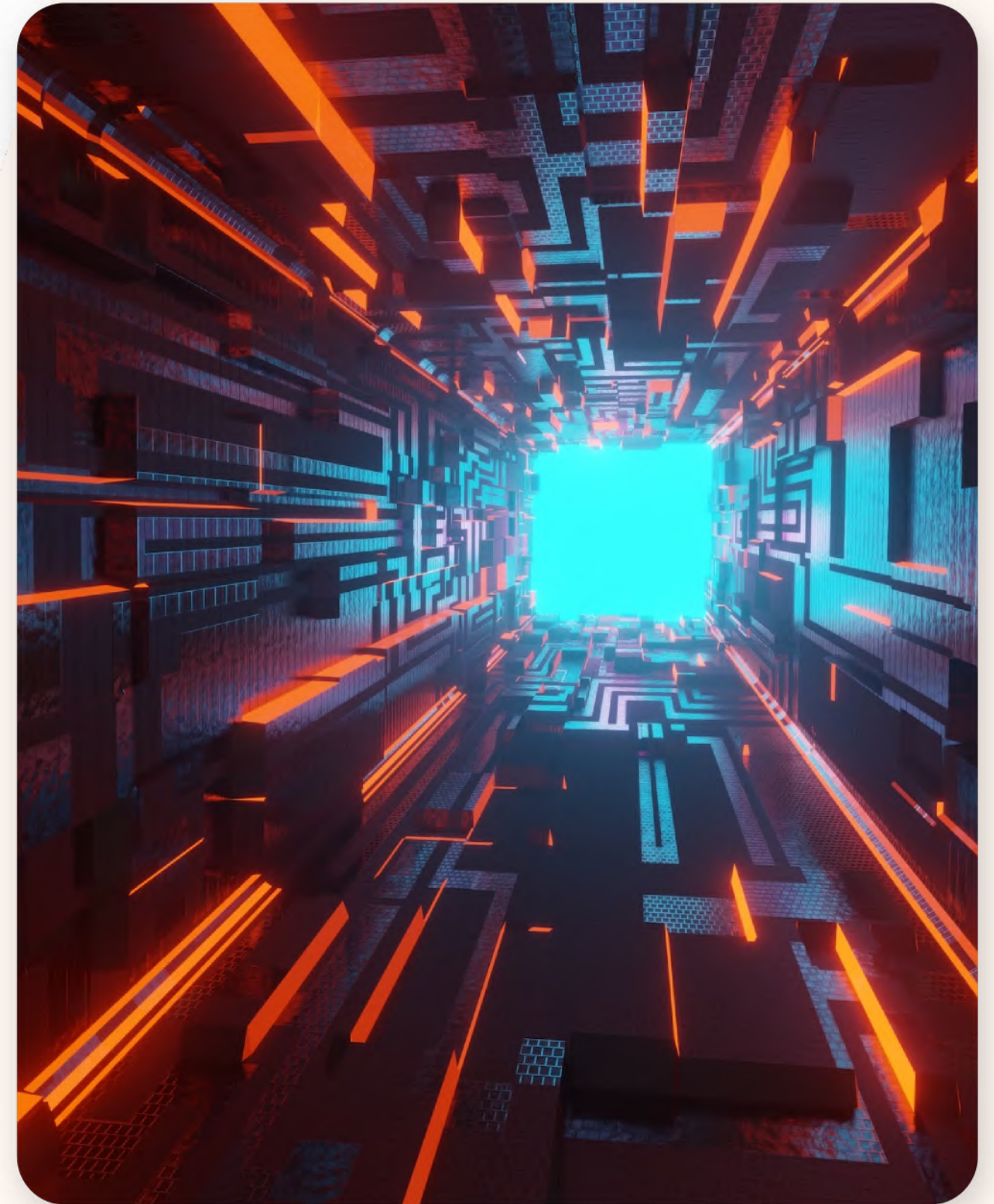


21:00 - 00:00



Future Work

- Explore the relationship between restaurant, coffee shop, and pub data and variables like population density or income.
- Analyze the impact of the COVID-19 pandemic on subway usage patterns, demographics, and economic conditions.





Thank you!

**Feedback
&
Questions?**