**Comparing peak ridership trends in the NYC Subway using MTA turnstile data**

By Sean Flanagan

**I. Abstract** – The goal of this project was to identify differences in ridership trends between 2019 and 2022. This was primarily motivated by the assumption that the ongoing Covid-19 pandemic had significantly changed how customers use the MTA. In particular, I was interested in finding out if the pandemic had changed “rush hour” or peak ridership times, and subsequently characterizing that change. In order to explore this question, I used data provided by the MTA in their [turnstile data set](http://web.mta.info/developers/turnstile.html). Using this data, I was able to create several helpful visualizations that begin help us understand how the pandemic has changed peak ridership and ridership in General

**II. Design** – This project is secondarily motivated by the desire to improve MTA’s subway service, which can often prove frustrating. Hopefully by identifying useful ridership trends with these visuals, the MTA can gain insight into its customers’ behavior adjust its service to better meet the needs of its riders.

**III. Data** – The data consists of ~500,000 timestamped entries, consolidated by station for easier data processing. Also included are columns recording the datetime and, separately, the month, year, day, week, and time of day. The most in-depth analysis was performed using these time categories to better understand ridership trends over time.

**IV. Algorithms**

**-** Grouping and consolidating entries by station, time, day, week, and month for better analysis

- Data cleaning by eliminating irregular timestamps (e.g. 01:23:46 instead of easier timestamps like 01:00:00)

- Datetime manipulation for easier processing and analysis

**V. Tools**

**-** SQLite and DB Browser for SQLite for data exploration

- Python Pandas and NumPy for data cleaning and manipulation

- Matplotlib for plotting

**VI. Communication**

In addition to the visuals and slides, this project will expand on my [personal GitHub page.](https://github.com/SeanMartinFlanagan/EDA-Project)