

# Predicting Subway Use in a World with CV-19

By: Patrick Brown

## Abstract

Changing trends in subway use make it hard to plan resources and budgeting. The goal of this project was to improve transit planning by providing insights on what to expect. Using Exploratory Data Analysis (EDA) and statistical modeling (Linear regression, Fast Fourier Transformation), we find there is a positive trend in subway use but it is not predicted to return to pre-Covid levels in the next 12 months. Further, we find that expected growth is not uniform across stations and lines. These discrepancies can be utilized for planning, saving money and improving passenger experience.

## Design

New York City transit use dropped by [90 percent in March of 2020](#). The Covid-19 pandemic and the lockdowns it brought rocked the MTA. All bets were off on what to expect for transit use. This change makes it extremely hard to plan resourcing and budgeting. We ask the questions: How is subway use changing? Will it return to post pandemic levels? How can we better plan for these expected changes? Leveraging the MTA's data and public Covid-19 data along with methods of Exploratory Data Analysis (EDA) and statistical modeling, we will not only answer these questions but give actionable feedback.

## Data

MTA turnstile data was obtained for 2020-Present. This included 22,000,000 unique observations. After aggregating by day and station, 300,000 observations were left for analysis and modeling. Public Covid-19 Data was used for the corresponding time period. A dataset for linking stations with zip codes was used as well.

## Algorithms

Linear regression with ordinary least squares was used for the predictive model. Fast fourier transform was used on residuals. A new regression was fit including the Fourier transformation.

## Tools

- SQLAlchemy (bring data into Python)
- Pandas (Cleaning and EDA)
- Numpy (Preprocessing)
- Sklearn (Modeling)
- Plotly (Visualization)

## Communication

The slides presented can be accessed on my personal site <https://sites.google.com/view/patrick-brown/home>. Code for the project can be found on on my Github <https://github.com/PatrickTyBrown?tab=repositories>.