

Project proposal

backstory:

In this project, the purpose is to find out which stations are the busiest peak times as well as the busiest days in order to help Variety the Children's Charity of New York run a fundraiser, the Variety the Children's Charity needs to set up 5 booths across New York City

Data description:

MAT Turnstile Data:

Data from The Metropolitan Transportation Authority (MTA) involving the number of entries and exits at various turnstiles around each subway station in New York City. The data is updated each week and the sample size of this data is 3 months.

Tools:

- 1- SQL lite: Basic relational database management system (RDBMS) software.
- 2- SQLAlchemy : Python SQL toolkit and Object Relational Mapper that gives application developers the full power and flexibility of SQL.
- 3- Pandas library: it's a fast, powerful, flexible and easy to use open source data analysis and manipulation tool
- 4- Matplotlib : library that allows visualization within Python