

Project proposal

backstory:

In this project, the purpose is to identify and evaluate the most popular types of fare cards used by customers, to identify the most requested features of cards that are of interest to customers, so that we can develop new cards that offer new features that can benefit both parties customers and MTA.

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.

Fare Card History for MTA:

Data from NYCT involving the different types of MTA fare cards purchased at different subway station in New York City. The data is updated each week and the sample size of this data is 3 months, it shows the number of MetroCard swipes made each week by customers entering each station of the New York City Subway.

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