

**Topic :** Project Proposal**Overview:**

The advertising sales department of public transport companies is considered one of the most important sources of income, It contributes to achieving profits through advertisements inside stations and transportation, trains and buses. In this project, We will use data science and take advantage of the available dataset to determine the number of passengers and what are the peak times for the entry and exit of passengers from the stations, which helps the advertising sales department in setting standards for pricing advertisements according to the number of passengers and peak times, therefore the time of advertisement is in peak times. It cost more.

**Question/need:**

- Top 5 most visited stations throughout the week
- How many views are expected for the advertisement in the most visited stations?
- What is the best time to advertise?

**Data Description:**

Field Name	Description
C/A	Control Area (A002)
UNIT	Remote Unit for a station (R051)
SCP	Subunit Channel Position represents an specific address for a device (02-00-00)
STATION	Represents the station name the device is located at
LINENAME	Represents all train lines that can be boarded at this station
DIVISION	Represents the Line originally the station belonged to BMT, IRT, or IND
DATE	Represents the date (MM-DD-YY)
TIME	Represents the time (hh:mm:ss) for a scheduled audit event
DESC	Represent the "REGULAR" scheduled audit event (Normally occurs every 4 hours)
ENTRIES	The cumulative entry register value for a device
EXITS	The cumulative exit register value for a device

By analyzing the dataset from MTA, to observe and determine peak times, which helps the advertising department to know the best time to display ads and set the highest prices accordingly.

**Tools:**

Python | Pandas | SQL