1. Abstract

This project aim to explore the dataset to answer the needed questions by charts, That gives us insights and define a relationship between the columns.

The National Information Center (NIC) is providing the latest technology services and digital solutions for government agencies in the kingdom of Saudi Arabia to enhance the efficiency and safety of the government digital assets according to the best global standards. The NIC is aiming to advance the various Saudi sectors as the first provider of insights and analyses based on artificial intelligence technology and as one of the three affiliated agencies of the Saudi Data and Artificial Intelligence Authority (SDAIA).

2. Design

Which station and month can we locate absher machines?

3. Data

I plan to use the New York subway Metropolitan Transportation Authority (MTA) dataset which contains almost all the useful information needed for this project (http://web.mta.info/developers/turnstile.html).

The scope of this project was the most recent three months from June to August of 2021, to take an appropriate decision from up-to-date datasets.

The dataset contains (2,722610 rows × 11 columns)

4. Algorithms

1. Problem understanding
2. Data collection

3. Data preparation
4. Explore data
5. Findings and insights

5. Tools

- Technologies
    1. SQL
    2. SQLite
    3. Python
    4. Jupyter Notebook
- Libraries
    1. NumPy
    2. Pandas
    3. Matplotlib

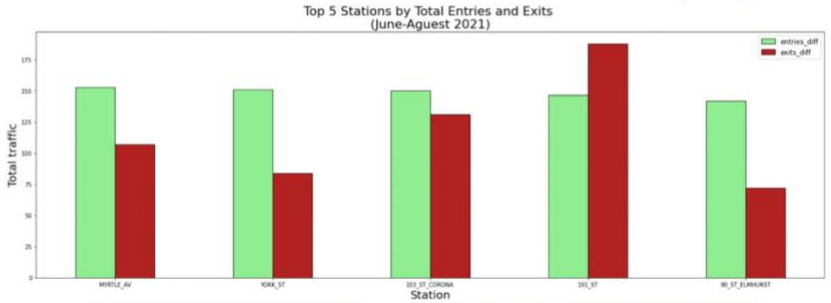
6. Communication

SDAIA
الهيئة السعودية للبيانات والذكاء الاصطناعي
Saudi Data & AI Authority

Findings

Observation

The top five stations has an equal total entries but differ in total exits



SDAIA
الهيئة السعودية للبيانات والذكاء الاصطناعي
Saudi Data & AI Authority

Findings

Observation

York street has the total traffic in july, Zerega avenue has total traffic in june