



# **The Metropolitan Transportation Authority**

---

**Exploratory Data Analysis (EDA)**

---

Nada alruwaythi

October 10, 2021

# Outline

---

**01. Introduction**

---

**04. Data set**

---

**02. Backstory**

---

**05. Data Cleaning**

---

**03. Design**

---

**06. Tools**

---

**07. Analysis result**



# “ Introduction

The New York City Subway is a rapid transit system that serves four of the five boroughs of New York City, New York: the Bronx, Brooklyn, Manhattan, and Queens. In 2016, an average of 5.66 million passengers used the system daily, making it the busiest rapid transit system in the United States and the seventh busiest in the world .





## Backstory

As a well-known cookies store working in New York, which provides improves the product's quality by taking feedback from clients while also increasing the store's income. Every customer who buys cookies and completes the survey will receive free cookies.

It targets the most crowded station and during rush hours to put the booth cookies. I used these months August, September, October, and November because it's almost the end of summer.



*Coookies*

“

## Design

The data set in this project is taken from the MTA website. This data set presents the complete status of the metro in New York City.



## Data Set

The used data from MTA data set are 4 months of 2019 August, September, October, and November

Dataset contains 3,708,990 with 11 row

- Turnstile
- Stations
- Line names
- Number of entries and exits
- Date and Time

# Tools

---



matplotlib



seaborn



SQLAlchemy



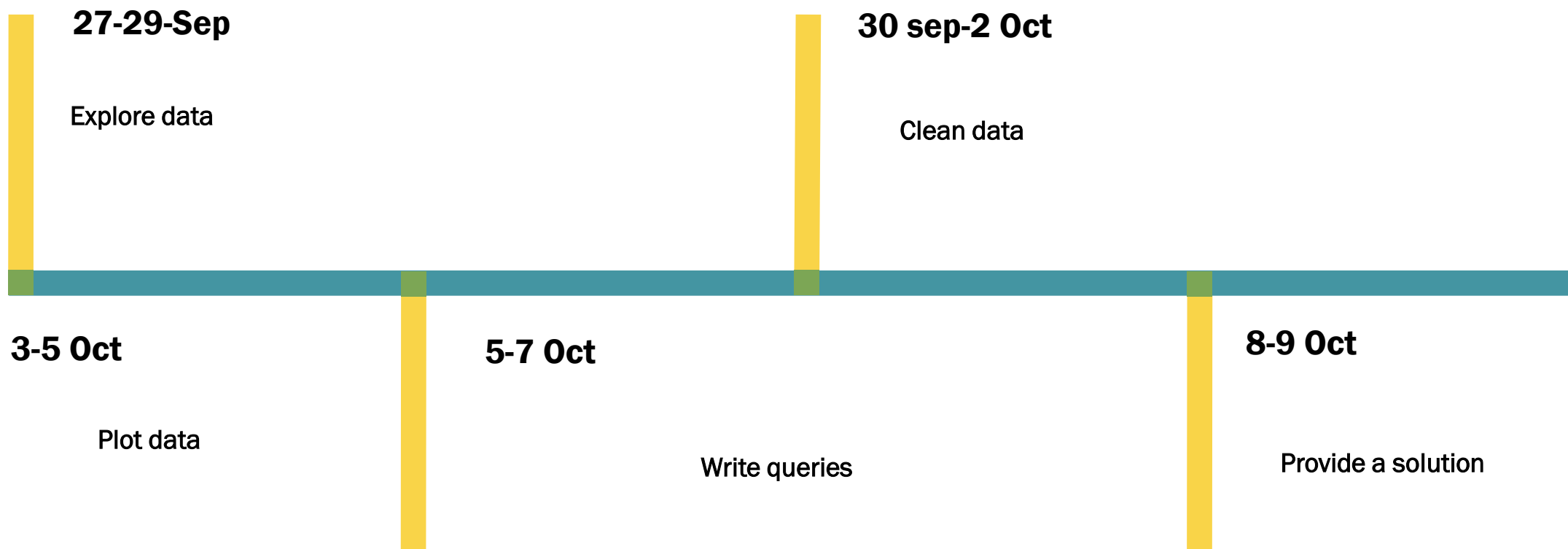
NumPy



pandas



# Timeline



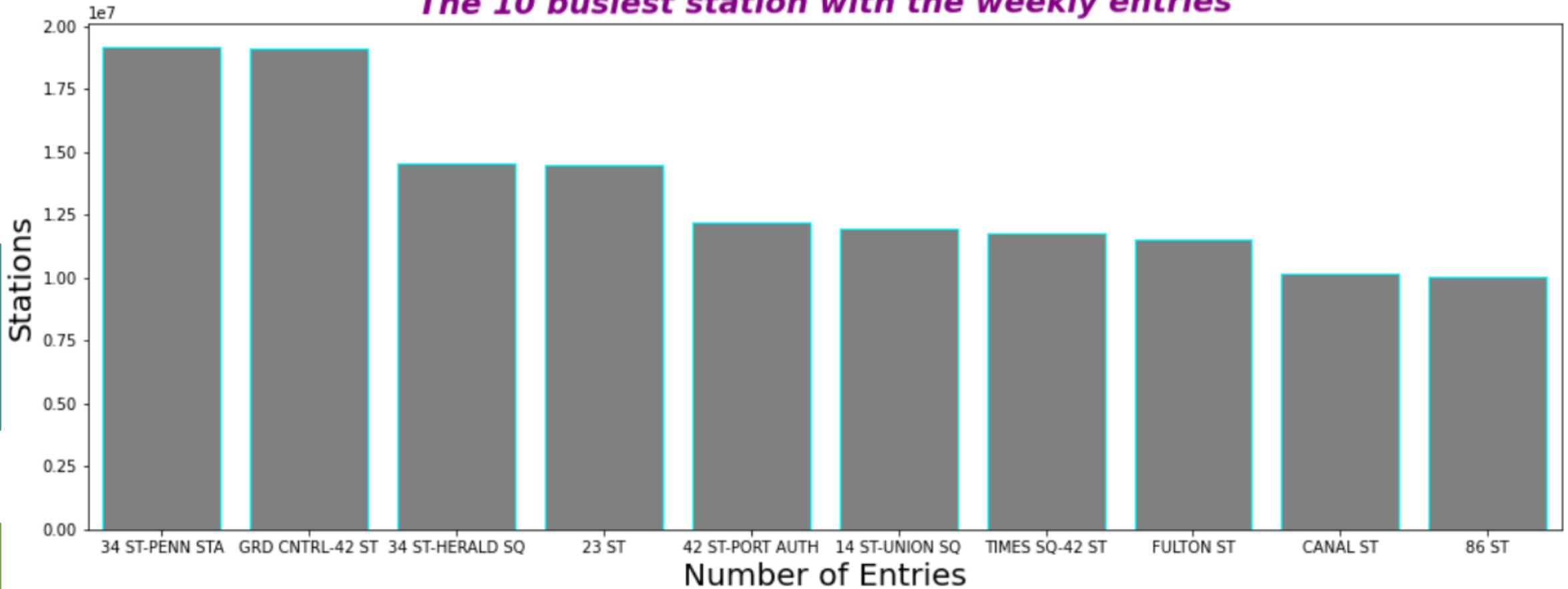




## Analysis Result

*Cookies*

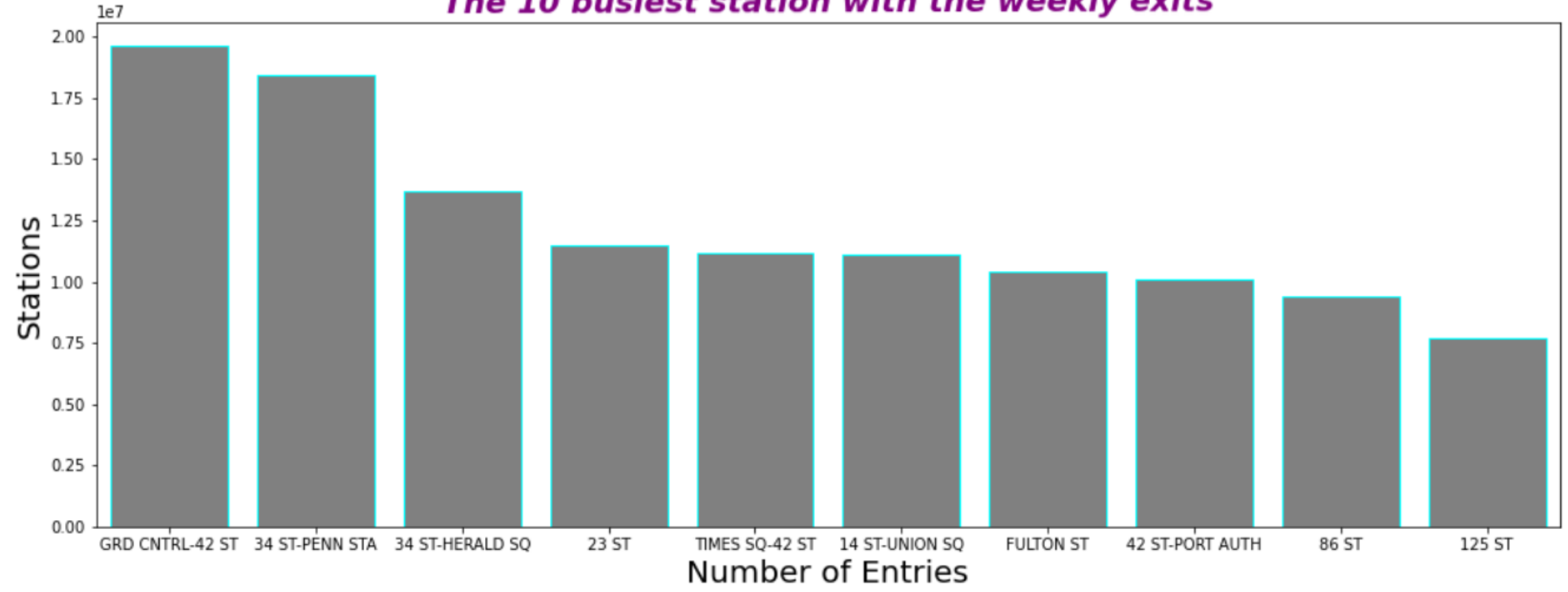
*The 10 busiest station with the weekly entries*





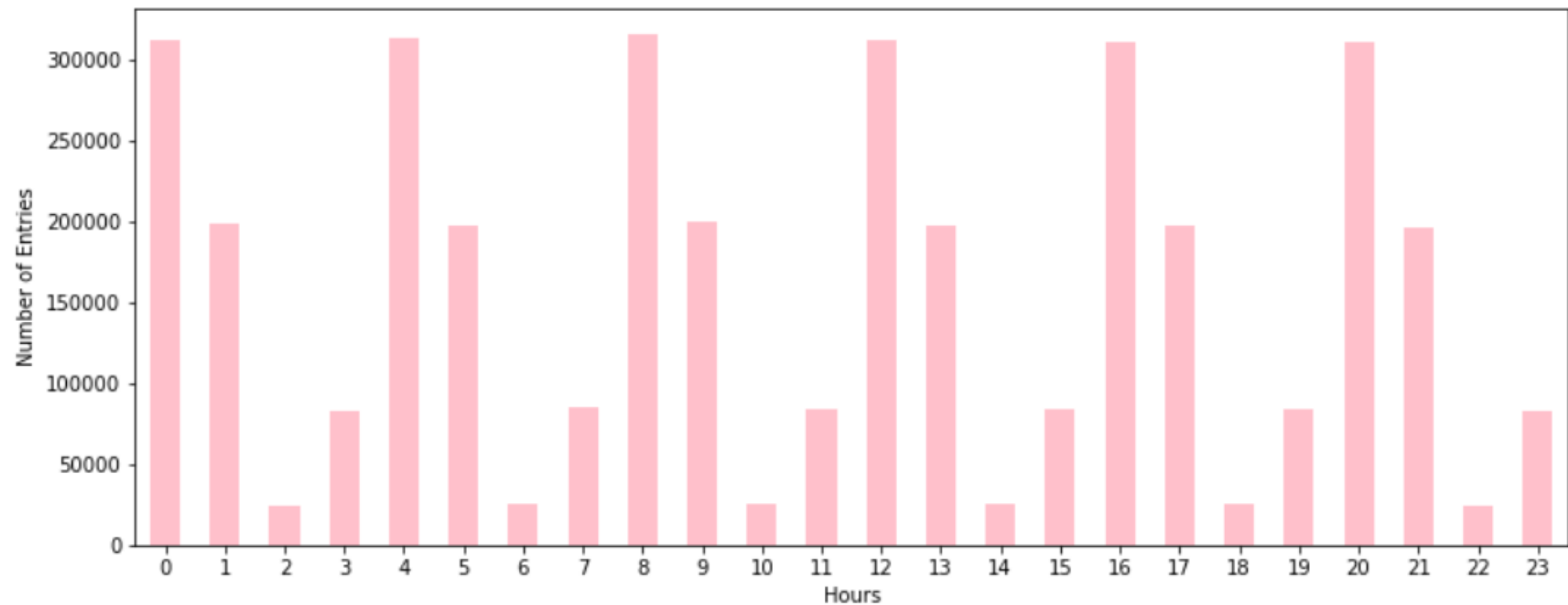
## Analysis Result

*The 10 busiest station with the weekly exits*



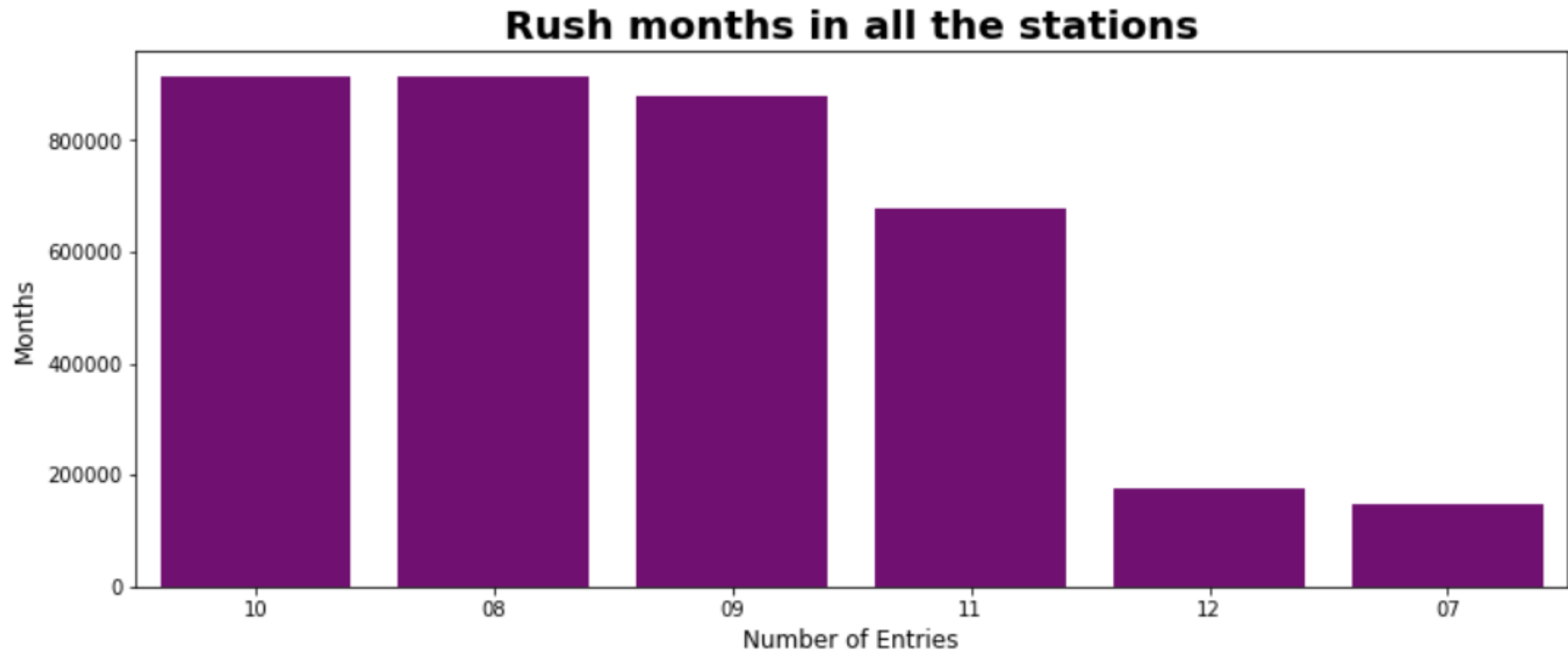
# “ Analysis Result

Rush Hours in the busiest station



The rush hours 4,8,12 A.M & 12,4,8 P.M

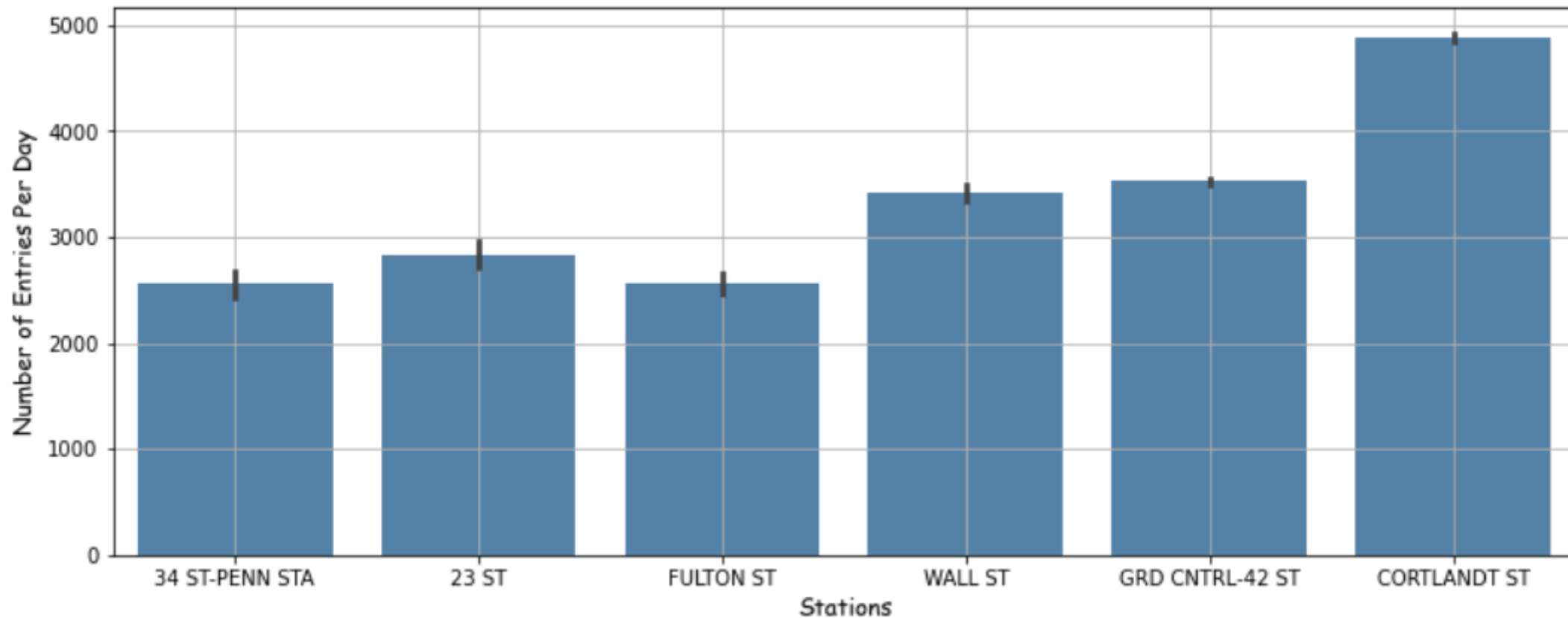
## “ Analysis Result



- The busiest months are August and October of 2019

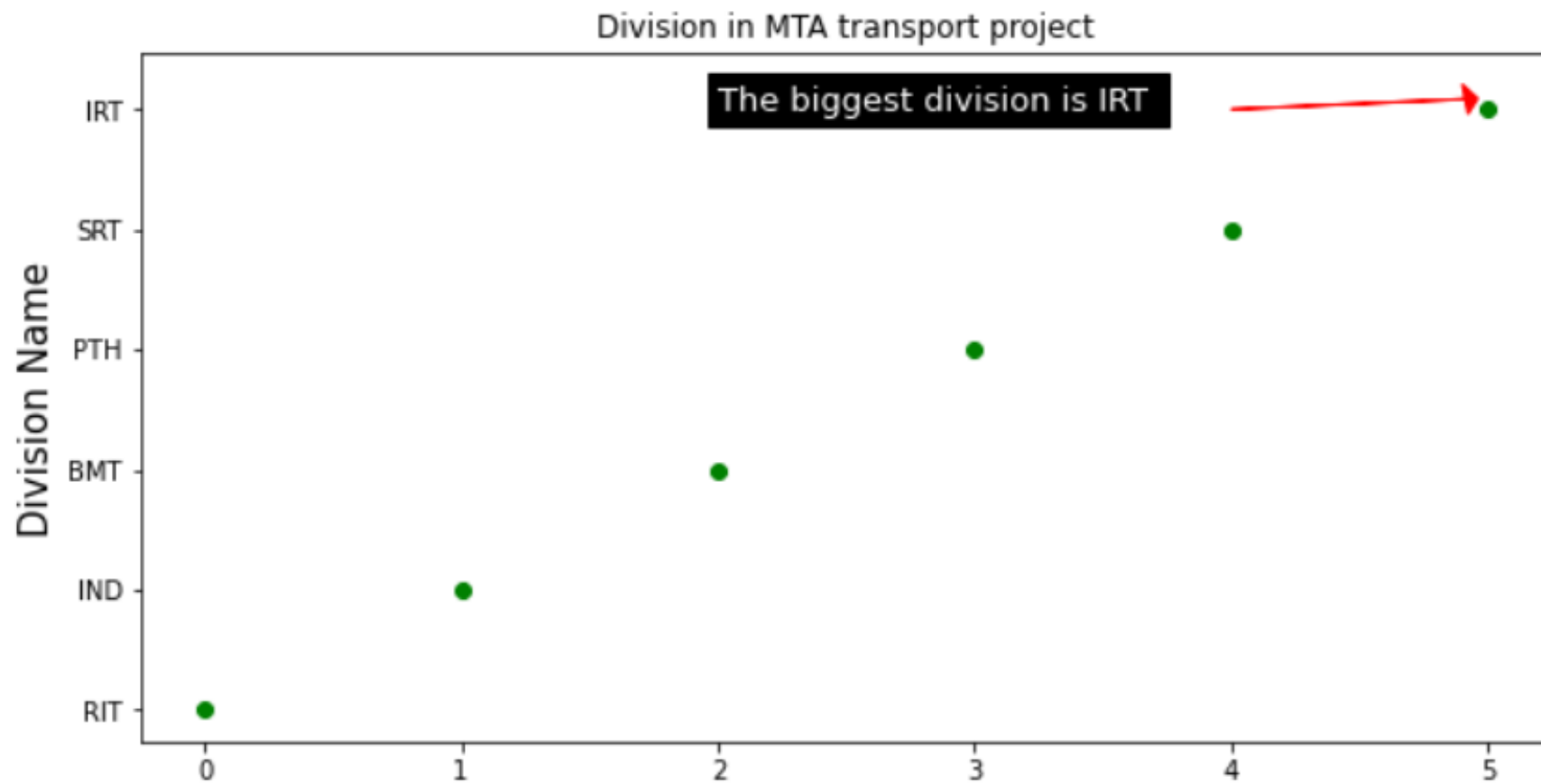
# “ Analysis Result

The Six Most Crowded Stations



*Cookies*

# “ Analysis Result





*Cookies*

**Thank you**

---