

SUMMARY:

We want to open shop for luxuries for NBA(national Basketball association) on the final day and use MTA turnstile data to choose which stations are crowding

Data:

I'll use the 3 months data to prepare for the final day

11 row's

| feature | definition |
|----------|---|
| C/A | Control area (A002) |
| UNIT | Remote unit for a station |
| STATION | Represent the station name the device |
| 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 data (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 comulative entry register value for a device |
| EXIST | the comulative exit register value for a device |
| SCP | represents an specific address for a device |

Tools:

I use numpy,panda,matplotlib,datetime and SQL Browser ,jupyter

Conclusion:

After studying the data I can choose which devices receive the largest number of visitors and I can choose which station the store is placed in