

Araos, Gabrielle
Gonida, Sean
Ilicito, Norence
Bacoy, Jeremiah

CS 129.1 Final Project - Big Data Problem

I. Approved Big Data problem

A. Question:

- a. Which time of the day does the MRT Line 3 stations experience the most traffic?
 - i. Comparing the average traffic per hour of the day
- b. Does the day of the week / month affect rider traffic?
 - i. Comparing average traffic per day of the week and week of the month.
- c. Which stations receives the most amount of traffic on average?

II. Description of source dataset

A. Description of source + Links

The dataset is sourced from Open Data Philippines (data.gov.ph), a website that collects data from different government websites and agencies. The Department of Transportation, which manages the MRT Line 3, provided the data.

The dataset holds a record of the hourly number of riders that alight from and board trains stopping in each station of the MRT Line 3.

B. Procedure on how to get data set

- a. Access the link:
<http://data.gov.ph/dataset/metro-rail-transit-line-3-passenger-traffic-daily/resource/ee0e3699-62a5-444c-a651>
- b. Download the data set as a csv file entitled
“2013_mrt_hourly_daily_ridership” , which will be used for the project.

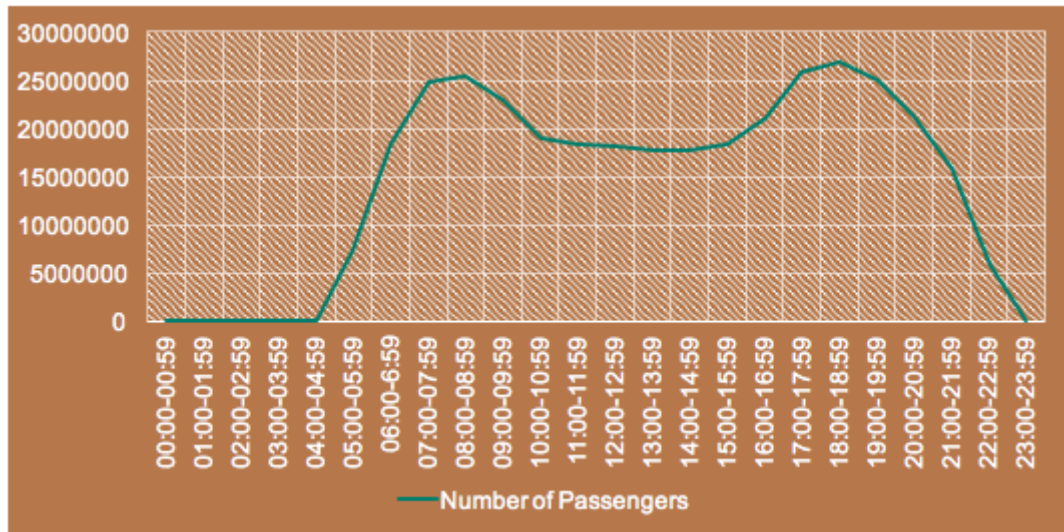
III. Description of the output, explanation and discussion of the output in relation to the Big Data problem + Visualization of the output

A. The MRT-3

- A rapid transit system located in Metro Manila, Philippines
- 16.9 km track length
- 40-60 km/h operating speed

B. Questions

- a. Which time of the day does the MRT receive the most traffic?



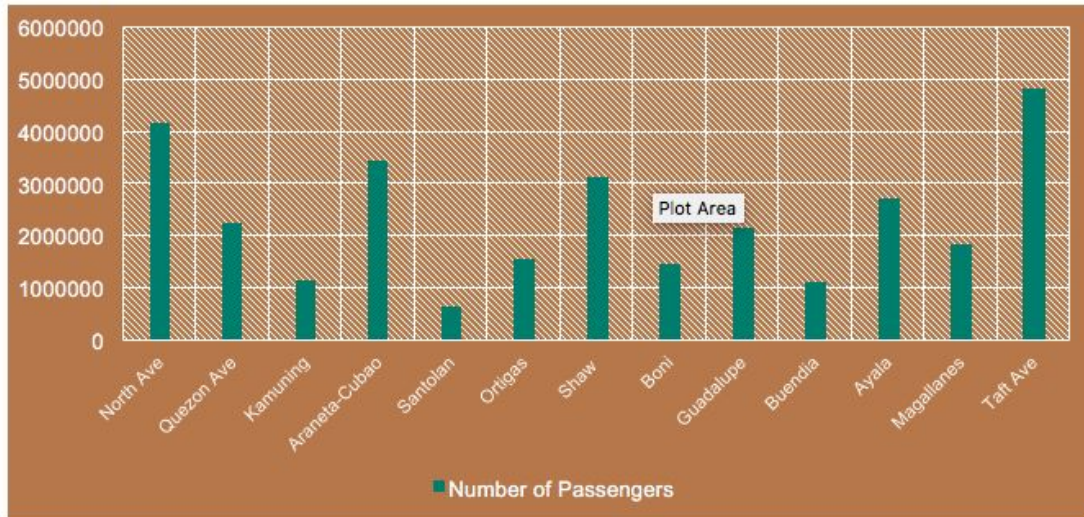
Based on the graph above, the time of day that has the most traffic is **6:00 - 6:59 PM (afternoon rush hour)**.

- b. Does the day of the month affect rider traffic? At what month does the MRT-3 receive the most traffic?

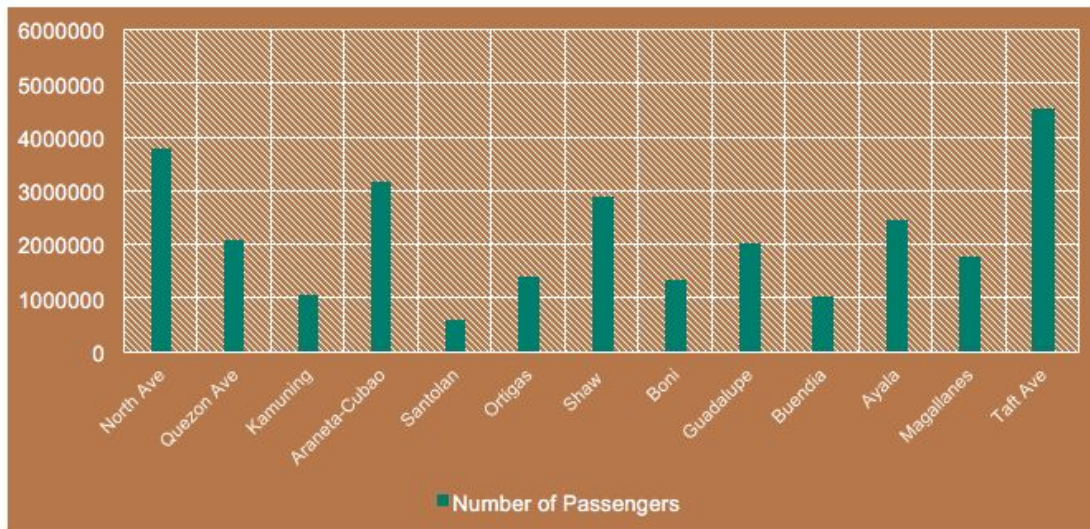


Based on the graph above, the month that has the most rider traffic falls between July and August, or the usual start of a new school year. However, this doesn't differ as much as from the rest of the months.

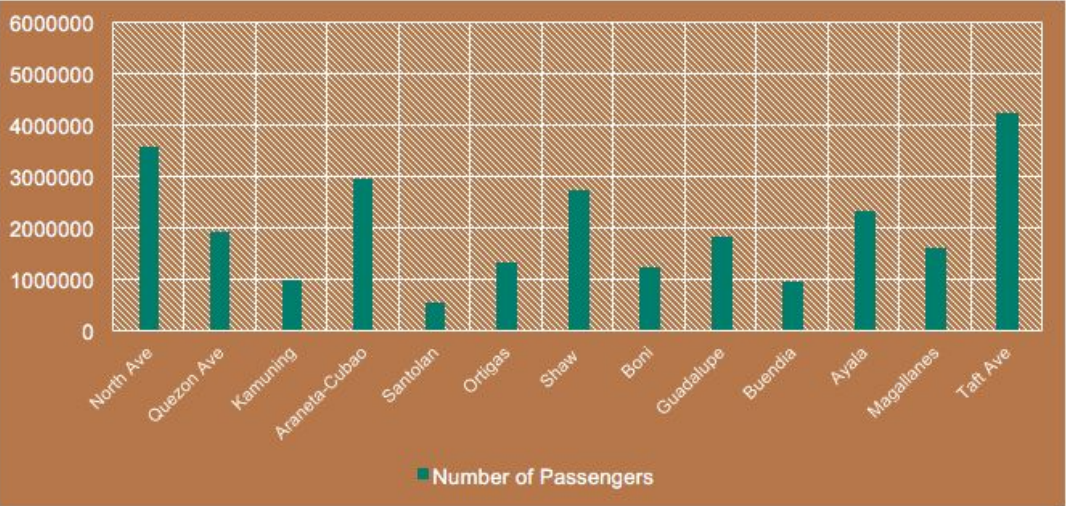
January Traffic Per Station



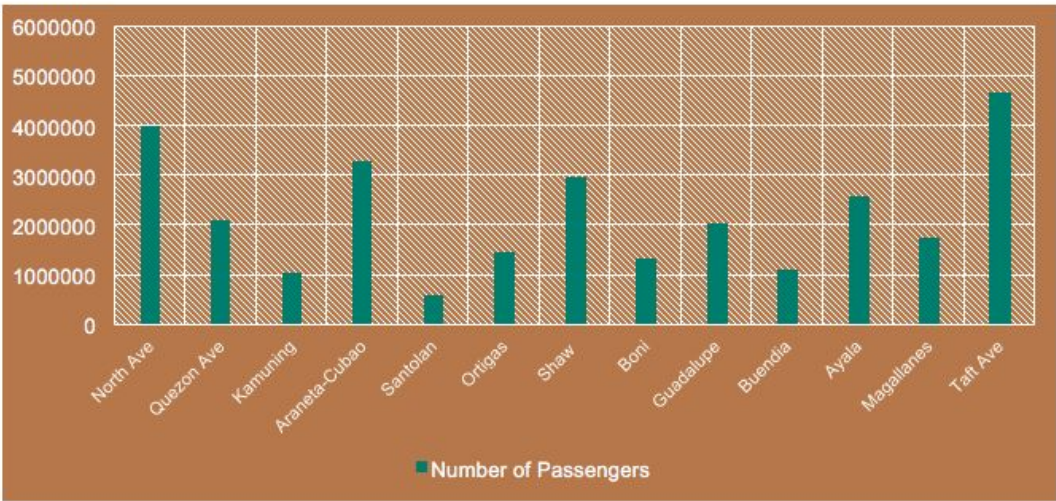
February Traffic Per Station



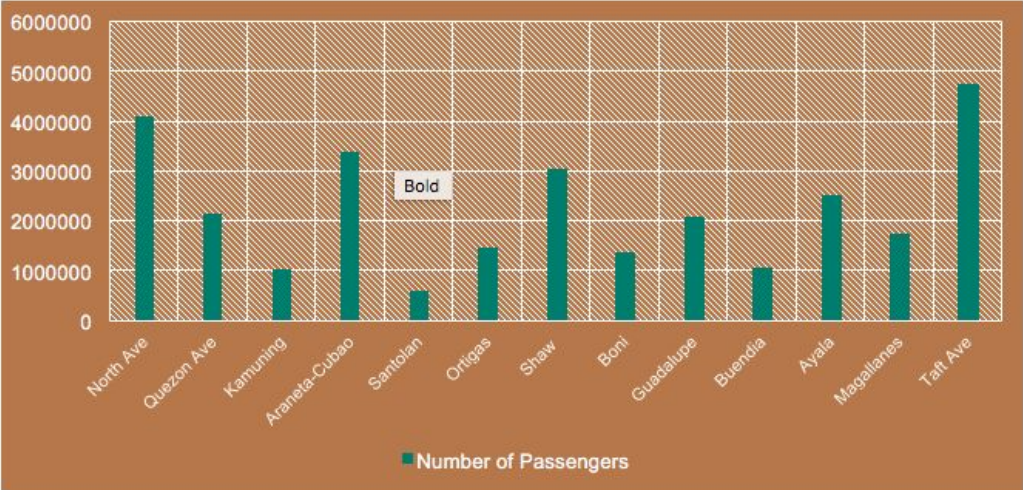
March Traffic Per Station



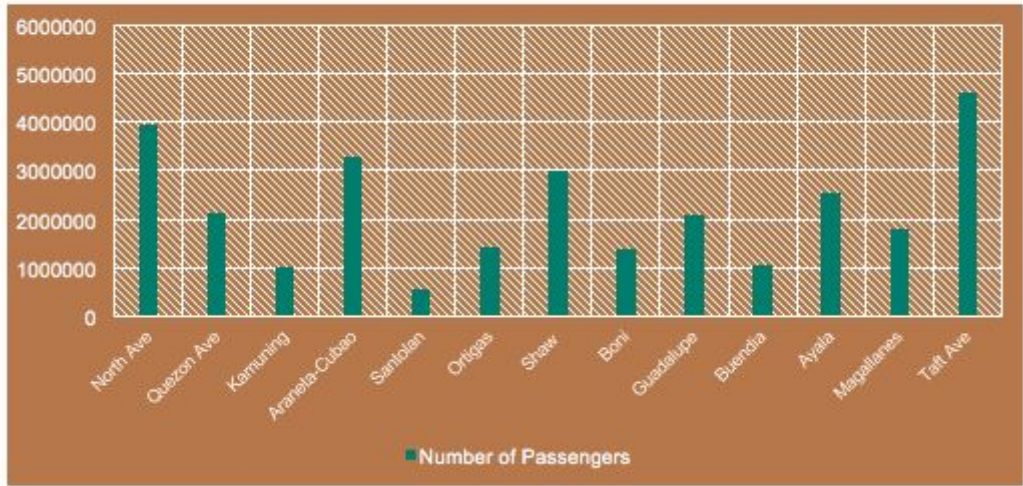
April Traffic Per Station



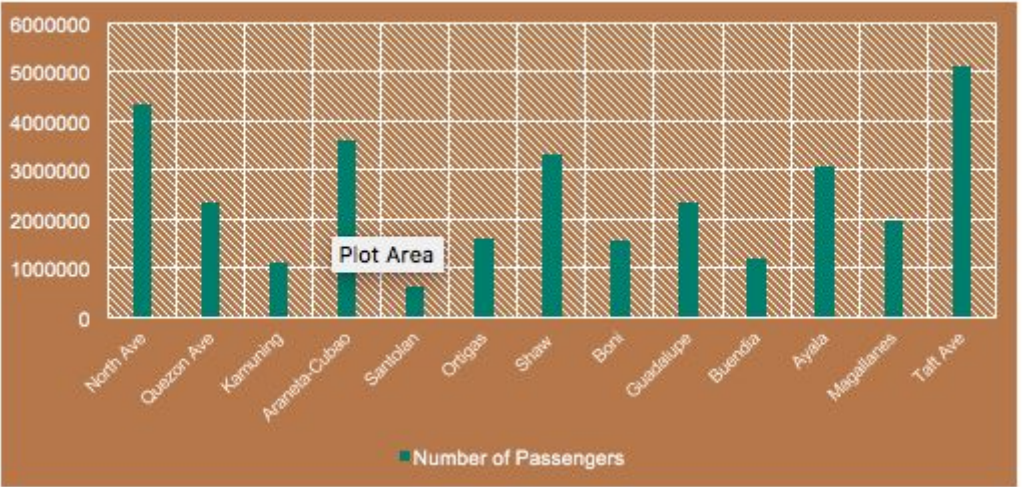
May Traffic Per Station



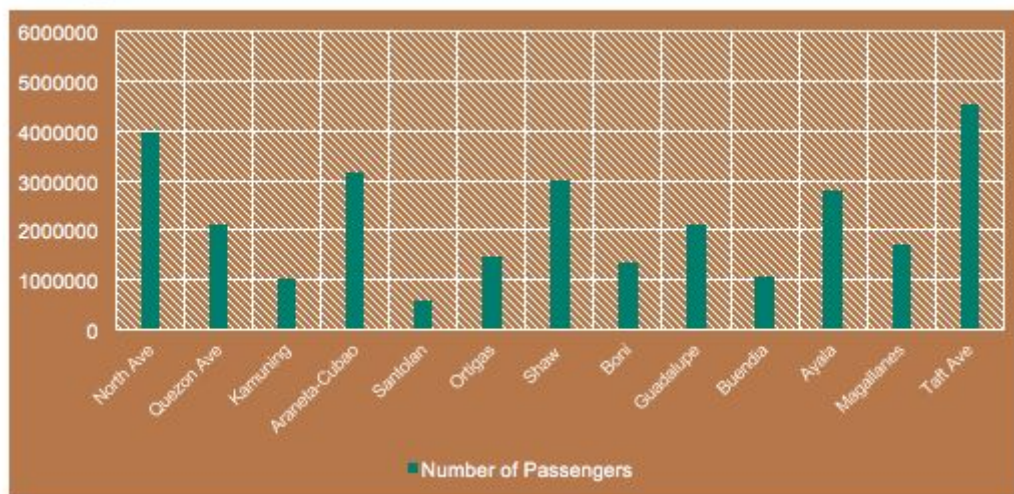
June Traffic Per Station



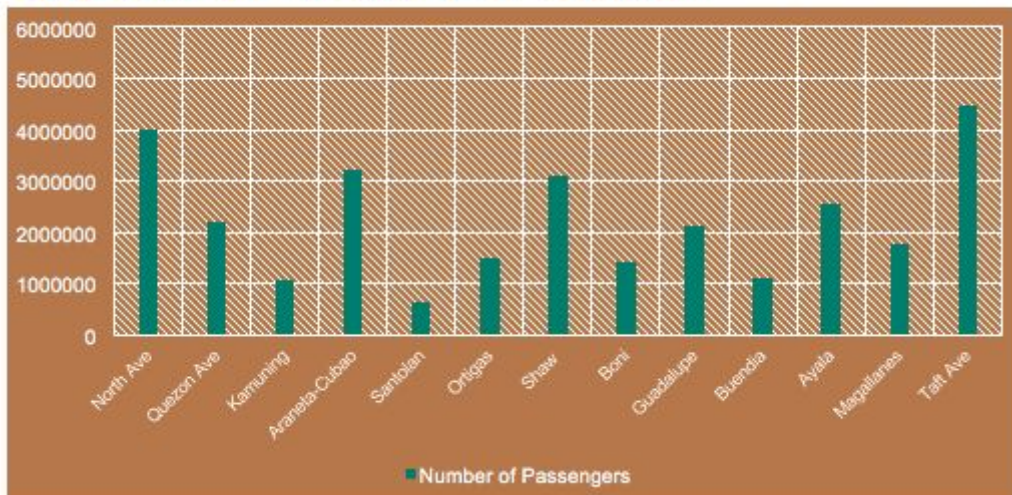
July Traffic Per Station



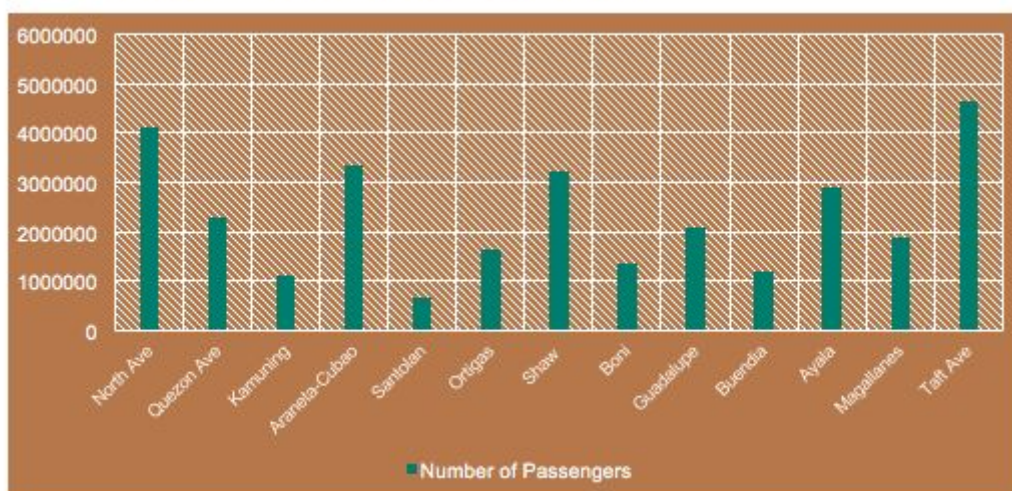
August Traffic Per Station



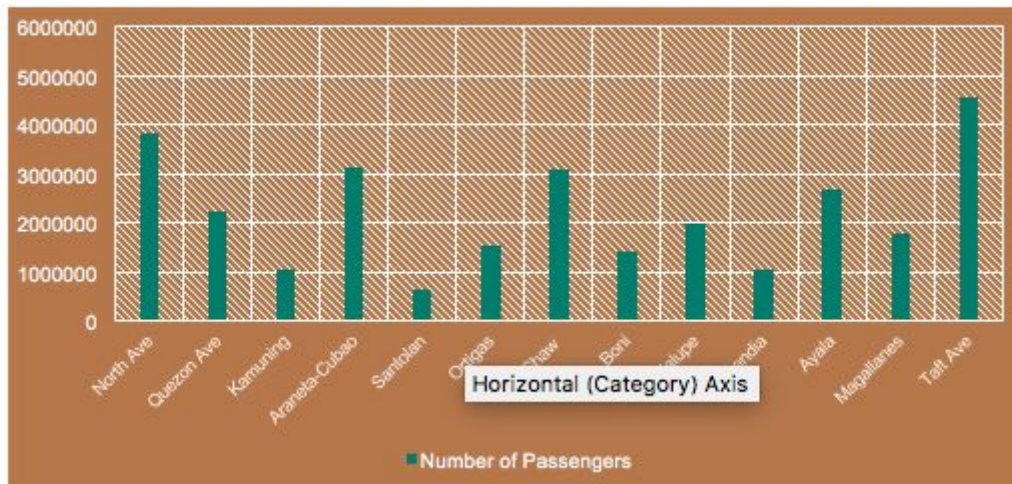
September Traffic Per Station



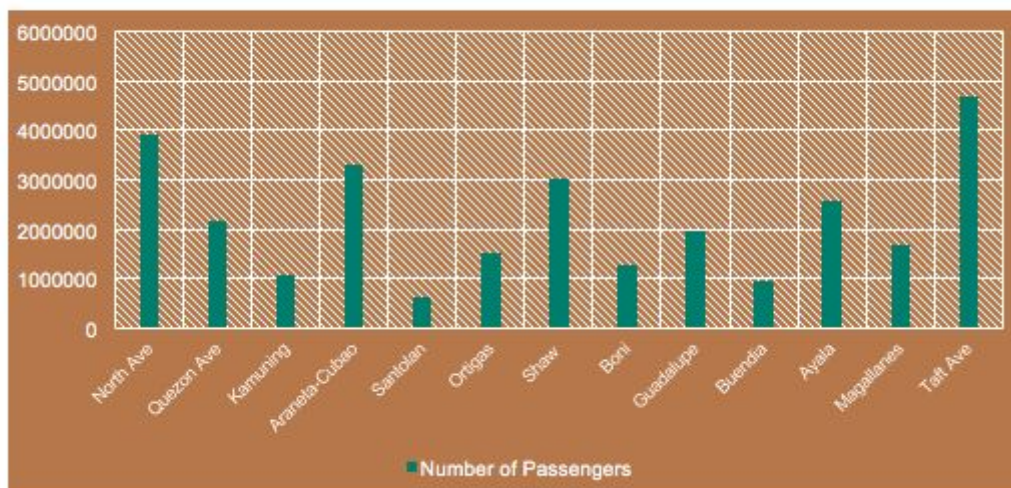
October Traffic Per Station



November Traffic Per Station



December Traffic Per Station



Based on the visualizations above, regardless of the month in the year, the station that has the greatest ridership traffic is at Taft Avenue, followed by North Avenue, which are the ends of the route taken by the MRT-3.