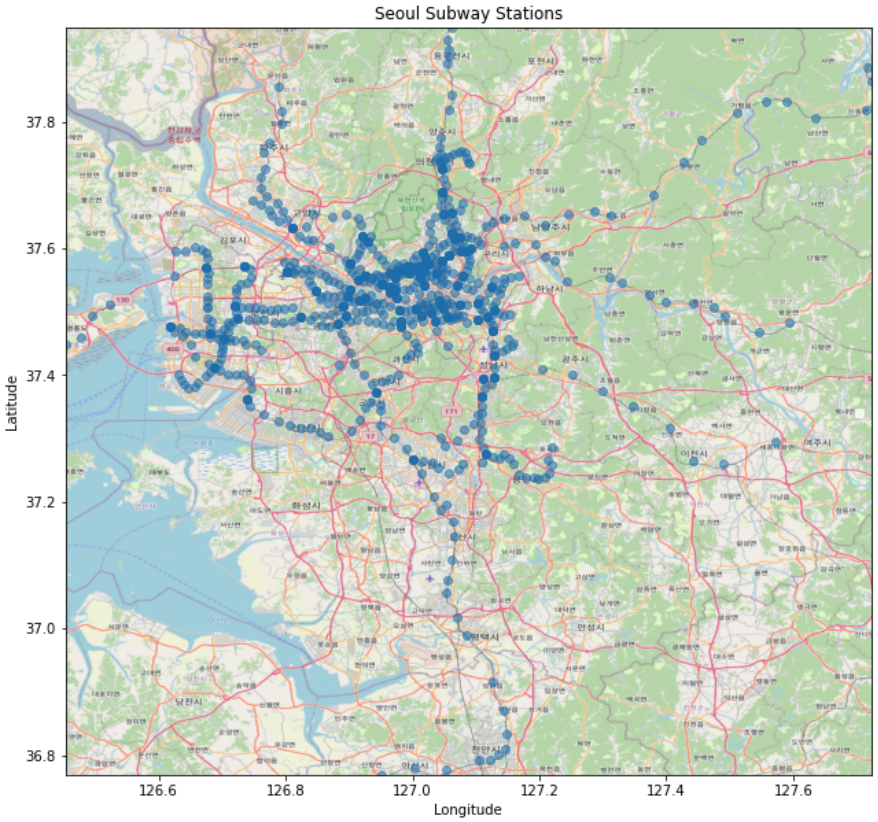
Amy Tan
Lingfei Wu
INFSCI 1520
12 April 2023

Seoul Subway Station Traffic

Figure 1

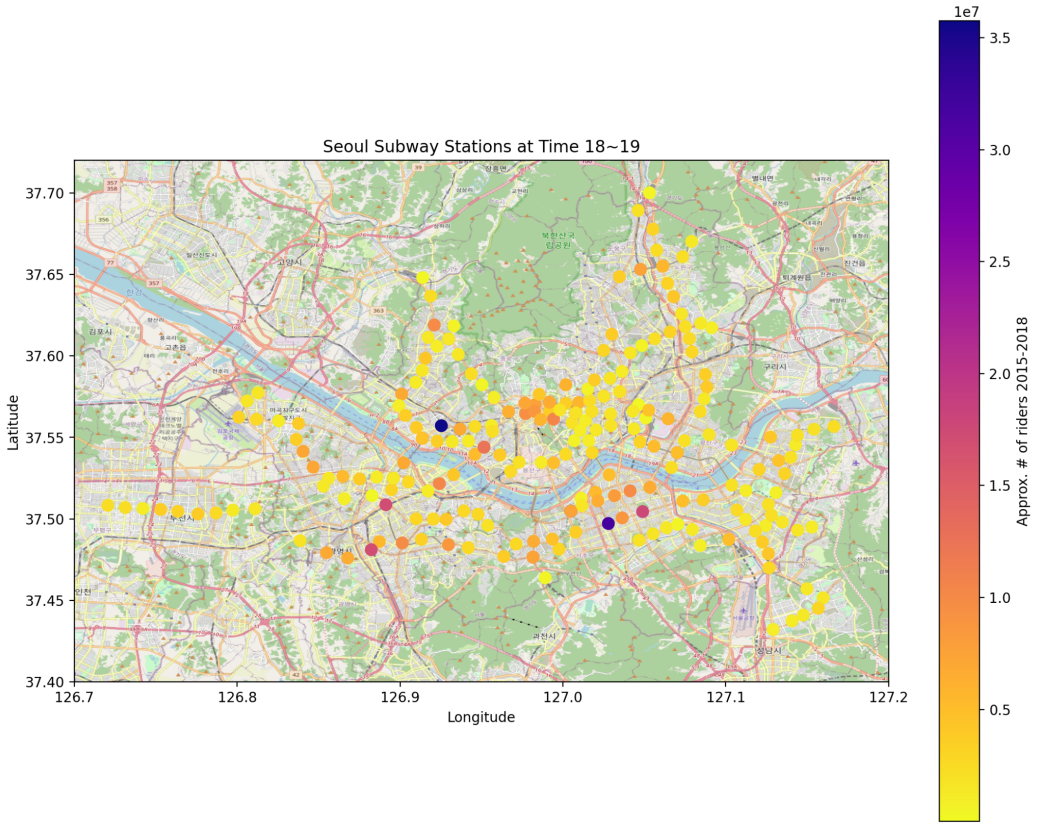


Findings & Highlights (Figure 1)

- Map of subway stations in South Korea, primarily around the capital city of Seoul
- Visualize the higher volume and density of stations closer to the metropolitan area of the city
- Easily see how the routes branch out into the surrounding cities of the province

https://github.com/amyjtan/1520finalreport

Figure 2



Legend explained (Figure 2)

- Each dot represents a station
- Color of each dot indicates how many riders visited that station during the specified hours from the years 2015-2018 → Darker color indicates more passengers at that station
- Click the figure or this link to view how it changes throughout the day by hour

Findings & Highlights (Figure 2)

- Takes a closer look at the metropolitan area
- Shows relative busyness of subway stations in the Seoul metropolitan area
- Visualize which stations have more passenger traffic at a given time
    - Easily see what are possibly business districts or where a lot of transfers occur
    - It can clearly be seen where some of the busiest stations are (likely Hongdae and Gangnam)
- I was curious to see if there were any stations that were really busy for just certain times, but in general, the stations that were busier consistently had more traffic than other stations. All of the stations seemed to get gradually busier as it got closer to rush hour times.

Data & Methods

https://github.com/amyjtan/1520finalreport

- Data from https://www.kaggle.com/datasets/gyejr95/subway-boarding-and-station-information-in-seoul
- Background map images from https://www.openstreetmap.org/
- Merged two datasets
    - First dataset contained general information about the subway stations, such as the name and latitude/longitude coordinates
    - Second dataset contained daily timely boarding information: for a given day and station, the number of riders that visited that station each hour was recorded
        - I summed all of the rows, grouping by each station, so that I had a total number of passengers for each hour over approximately 3 years
- Jupyter Notebook
    - pandas to prepare the data
    - matplotlib to create the visualizations by plotting the points on the image

Significance

Why is the presented figure an important topic?

- From the figure, one can easily see where the city of Seoul is most connected and how the passenger traffic varies between the stations and time.