**SDAIA T5 Boot Camp Project Proposal**

**Question/Need:**

What are the most used subway stations and the features of their location (e.g: near workplaces, tourist areas, housing area, etc..) then by using this information try to predict which stations will be used the most in the new Riyadh train project.

This could help the government in planning which stations to prioritize in the future.

**Data Description:**

Data is from the MTA open data set, to achieve the project’s goal, information about how many people enter a certain station in what time and date is needed.

Also a subway map for both MTA and Riyadh train project with the subway stations laid on top of them, from each organization website.

**Tools:**

SQLite and python.

**MVP Goal:**

Names of the stations that will be most likely crowded in Riyadh train project, based on the analysis done on the MTA.