

Data Analyst Exercise

Objective:

You have been assigned by Newark International Airport's Engineering Manager to conduct an analysis of the airport's current capabilities and identify key metrics that can improve the airport's administration. You have been provided with several data files: [nyc_airlines.csv](#), [nyc_airports.csv](#), [nyc_flights.csv](#), [nyc_planes.csv](#), and [nyc_weather.csv](#). Your task is to develop a Dashboard using a Notebook and/or BI tool to present the following information:

Overall Pane:

1. Determine the number of distinct destinations connected to the airport.
2. Calculate the rounded average of distinct destinations for each day.
3. Calculate the average number of destinations per day of the week.
4. Identify the month with the highest number of flights.

Dealing with Delays:

5. Identify the month with the highest accumulated departure delay.
6. Analyze the correlation between departure delay and weather conditions by cross-referencing the departure delay data with the weather dataset.

Airlines:

7. Create a ranking of airlines based on the number of flights they operate.
8. Create a ranking of airlines based on the daily availability of seats.

Airplanes:

9. Determine the most common plane manufacturer.
10. Identify the most common plane model.

Note: Feel free to use any Notebook and/or BI tool of your choice to develop the Dashboard. We recommend that you download the files beforehand and load them to your preferred data exploration platform.

This exercise aims to assess your data analysis skills and ability to work with multiple data sources. You will be analyzing airport data to derive meaningful insights and create visualizations that address various aspects such as overall connectivity, delays, airline rankings,

and airplane statistics. The exercise requires you to use SQL queries, data manipulation, statistical calculations, and visualization techniques to present your findings effectively. Good luck!