**EMSE 6586 – Database Management for Data Analysts**

**Team Details:**

Type: Individual project

Members: Aparna Shankar – G49628466

**Project Details: Database Translation**

The objective of this project is to transform a flattened JSON dataset containing information about airline statistics into a format suitable for loading into a SQL database. This dataset contains detailed statistics related to airlines, including flight delays, cancellations, diversions, and on-time performance, organized by airport codes, months, and years. By delving into flight performance metrics such as delays, cancellations, and on-time performance, airlines can identify trends and patterns to streamline operations and minimize disruptions. Python will be employed for data transformation tasks, including parsing the JSON dataset and preparing it for insertion into PostgreSQL. The airlines trend analysis will be analyzed using Python's visualization libraries. PostgreSQL will serve as the SQL database to store the restructured dataset, with a database schema designed based on the normalized structure of the airlines data. The transformed dataset will be loaded into the PostgreSQL database using Python scripts and Psycopg2 library.

**Dataset source:** <https://think.cs.vt.edu/corgis/datasets/json/airlines/airlines.json>