AIM 5001 Module 13 Assignment

PostgreSQL to MongoDB Data Migration

***** You may work in small groups of no more than three (3) people for this Assignment *****

Your task for the **Module 13 Assignment** is to migrate the content of a PostgreSQL database to MongoDB via Python, and then retrieve the data from MongoDB and load it into Pandas dataframes that conform to the format of the original PostgreSQL database tables. Specifically, you will use Python to retrieve the entire content of the **Customer** and **Invoice** tables from the **Chinook** database we've worked within in Modules 1 & 2 from your PostgreSQL server, and then subsequently load that data into MongoDB. The manner in which you choose to store and manage the Chinook data within MongoDB is up to you as an SQL and NoSQL practitioner to decide.

Then, using Python and your MongoDB data retrieval expertise, you will retrieve the required Chinook data from MongoDB and create and populate a pair of new Pandas dataframes, one for each of the original Chinook **Customer** and **Invoice** database tables. When finished with this process, your new Pandas dataframes should mirror the structure of the SQL tables contained within your PostgreSQL Chinook database.

Your data migration processes need to be 100% reproducible.

Your deliverable for this Assignment is your Jupyter Notebook. It should contain a combination of Python code cells and explanatory narratives contained within properly formatted Markdown cells. The Notebook should contain (at a minimum) the following sections (including the relevant Python code for each section):

1) **Introduction (5 Points)**: Summarize the problem + explain the steps you plan to take to address the problem
2) **PostgreSQL to MongoDB Migration (55 Points)**: Describe + show the steps you have taken to use Python to retrieve the Chinook data from your PostgreSQL server and subsequently load that data into MongoDB. This section should include any Python code used as well as an appropriate explanatory narrative that explains how you have formatted the Chinook data for storage within MongoDB.
3) **MongoDB to Pandas Migration (40 Points)**: Describe + show the steps you have taken in Python to retrieve the Chinook data from MongoDB and create + populate the required Pandas dataframes that mirror the structure of the SQL tables contained within your PostgreSQL Chinook database. This section should include any Python code used for Data Preparation as well as an appropriate explanatory narrative.

Your Jupyter Notebook deliverable should be similar to that of a publication-quality / professional caliber document and should include clearly labeled graphics, high-quality formatting, clearly defined section and sub-section headers, and be free of spelling and grammar errors. Furthermore, your Pythion code should include succinct explanatory comments.

Upload / submit your Jupyter Notebook within the provided M13 Assignment Canvas submission portal. Be sure to save your Notebook using the following nomenclature: **first initial_last name_M13_assn**" (e.g., J_Smith_M13_assn). ***Small groups should identity all group members at the start of the Jupyter Notebook and each team member should submit their own copy of the team's work within Canvas.***