**Week 2 - Assignment 2**

**General Instructions**

1. Download the OPC Database Schema document from the Assignment instructions.
2. Download the Week 2 Assignment Notebook (DSCI504\_Wk2\_Assignment.ipynb).
3. Download the OPC Data Tables (OPC\_Database\_Table\_Archive.zip). Extract this file to a location on your computer that PostgreSQL can access. For Windows users, this is usually a folder on the C:/ drive root directory. For Mac OS users, this can be your Desktop, Documents, or Downloads folder. Try to not use cloud drive folders as they create permission issues.

In this assignment you will create the database schema, tables, and import the OPC data into PostgreSQL while performing basic queries. This assignment is worth 50 points. The most important part of this assignment and the key objective is that you become familiar with importing data into SQL and perform basic DDL and DML querying of the data. You will be evaluated on your ability to properly execute each query using appropriate SQL syntax following the ANSI SQL standards.

You will submit a Jupyter Notebook exported from Azure Data Studio in Canvas for credit.

Assignment Preparation:

1. **Open the Week 2 Assignment Notebook in Azure Data Studio.**
2. **Open the OPC Database Schema document.**
3. **Copy and paste each Part instruction of this assignment to a new text cell before each code execution.**

**Step 1:** You should be working in the default “postgres” database. Create a new schema named ‘dsci\_504’ using the code cell in Step 1 of the Notebook to create new tables in and import the OPC data.

**Step 2:** Create the OPC Database tables using the Schema document. Use a new code cell for each table.

**Step 3:** Import the data into PostgreSQL using the CSV files in the OPC Database Archive.

**Step 4:** View basic data about the table using the following query:

SELECT \* FROM dsci\_504.customers WHERE cus\_id = 48;

**Step 5:** Create primary keys within each of the newly created tables to elevate the database to 2NF. Some tables already have unique key columns that simply need to be identified as a Primary Key. Others may require a new column to do this. Run a summary query after each change to verify id placement.

**Step 6:** CREATE foreign key associations for each table to elevate the database to 3NF. These should be done within each respective table.

**Step 7:** Execute the following SQL queries in separate cells to test your new key relationships:

SELECT warehouse\_name, sum(order\_tot) AS TotalOrders FROM dsci\_504.warehouse, dsci\_504.orders

WHERE orders.id = warehouse.warehouse\_id GROUP BY warehouse\_name

LIMIT 10

SELECT cus\_state, sum(tot\_ord\_value) AS TotalByCus, warehouse\_name FROM dsci\_504.customer

LEFT JOIN dsci\_504.warehouse ON customer.id = warehouse.id WHERE cus\_state = 'OH'

GROUP BY warehouse\_name, cus\_state;

**Step 8:** Save your Notebook, export to html using the instructions provided and upload to Canvas for credit.