DTSC660: Data and Database Management with SQL
Module 4
Assignment 3

Purpose

For this assignment, you have been asked to provide some informational statistics on Bitcoin transactions from 2009. Note that many, but not necessarily all, of the tools you learned in Module 4 will be applied in this assignment.To complete this assignment, download and import the dataset and then create queries that respond to each prompt. Please make sure that you only use postgres language conventions. **Each question is all or nothing. Graders will not attempt to correct or interpret malformed SQL queries.** You will be responsible for testing your code on the provided data set before submission. The question will be graded based on whether or not it generates the correct output and addresses all requirements specified in the question. Extraneous columns will not count against you as long as correct results are obtained.

Submission

You will submit a total of **1** sql files. files to CodeGrade. Files should be named appropriately and be in .sql format. Each file must use the postgres standards taught in the course. Use of other SQL languages such as T-SQL will result in an automatic 0 for the assignment. **Each question is all or nothing. Ensure your file runs in its entirety in pgAdmin. This means ensuring each query ends in a semicolon (;). Graders will not attempt to correct or interpret malformed SQL queries.** You will have **one** submission attempt for this assignment.

- ***File 1***: You must submit a SQL document called <LastName>_Assignment3. This document must include ALL queries requested in the instructions below.

- You will submit the file to the Assignment 3 link to CodeGrade.

Instructions

As in the practice assignments you will be querying a large dataset to gather insights about that data. You have been provided a SQL file with all the necessary data. Make sure you take a moment to familiarize yourself with the columns, their data types, and any other pertinent information. I would recommend printing out a query with the first few records using the following command to help you get started. It is expected that if you have questions or difficulties with any portion of this assignment that you utilize the assignment discussion board or email the GAs to gain clarity (dtsc_ga_660@eastern.edu).

PART 1 Creating the Table and Importing the Data

1. In the assignment 3 folder, download the ***bitcoin_data.csv*** file
2. Place this file in a public folder on your computer
3. Take note of the path to this file (copy the path)
4. Utilize the COPY command you learned in Module 3 (revisit this module if necessary) to import the data.
5. Run a basic select statement that verifies the data is present and matches what is in the csv file.

Part 2 (Queries)

This is the part you will be graded on. To complete this part, download the assignment_3_template.sql file from the assignment 3 folder. Rename this file using the naming convention: <LastName>_Assignment3. Complete each query in the identified space in this document. Once you are done, submit the document to the Assignment 3 folder.

Please note that some questions are intentionally vague. Remember that you will not be penalized for extraneous columns, but you will be for extraneous rows (data). There may be multiple correct ways to solve some queries.

1. Create the table with appropriate data types
    a. Name the table ***bitcoin_data***
    b. Reuse the column titles from the csv but remove any capitalization
    c. Change the date column to trans_date
    d. Change size to code_size
2. Write the copy statement to bring the code into the database
    a. Remember that if you choose an incompatible data type, you can enter the command ***DROP TABLE bitcoin_data*** to remove the table and restart.
3. Write a query that returns all the data in the table
4. Write a query that returns the transaction date and the code size divided by the transactions with the column name difficulty.
5. Write a query that returns the transaction date and the product of median transaction fee and transactions with the column name daily cost.
6. Write a query that returns the transaction date and the amount sent in USD divided by the number of transactions with the column name average transaction. Also include the median transaction value.
7. Write a query that returns the average price in USD with the column name avg price.
8. Write a query that returns the total number of transactions with the column name total transactions.
9. Write a query that returns the largest value from the market cap column and call it max cap.
10. Write a query that returns the mean number of tweets and call it the avg daily tweets.

**********GRADING RUBRIC ON NEXT PAGE**********

This assignment will be graded on the following rubric. Remember that questions are ALL OR NOTHING. Incorrect syntax, extraneous results, or incorrectly addressing all question requirements will result in loss of points for that question. Graders will NOT attempt to correct malformed sql code. :

Question Number	**Points**
1	10
2	5
3	5
4	5
5	10
6	10
7	10
8	15
9	15
10	15
Total	100