Lab 02. Spatial Data

**Due date:** Thursday, Feb 13 submitted as Word document to Canvas ***Lab03*** link. This lab counts 9 % toward your total grade.

**Objectives:**

1. Understand and apply logical operators to filter datasets.

2. Utilize Data Query Language (DQL) to retrieve and summarize data.

3. Perform Data Manipulation Language (DML) operations to update and modify records.

4. Use Data Definition Language (DDL) to alter table structures.

5. Implement subqueries to perform complex queries.

**Format of answer:** Submit your answers as a **Word document or pdf** with graphs (screenshot of your result table) and answer in SQL.

**Notice:** All SQL commands are in blue color

**In-Class Exercise (4pts)**

**On Your Own (5pts)**

Using the nyc\_census\_blocks table, answer the following questions .

Here is some helpful information to get started.

|  |  |
| --- | --- |
| **blkid** | A 15-digit code that uniquely identifies every census **block**. (“360050001009000”) |
| **popn\_total** | Total number of people in the census block |
| **popn\_white** | Number of people self-identifying as “white” in the block |
| **popn\_black** | Number of people self-identifying as “black” in the block |
| **popn\_nativ** | Number of people self-identifying as “native american” in the block |
| **popn\_asian** | Number of people self-identifying as “asias” in the block |
| **popn\_other** | Number of people self-identifying with other categories in the block |
| **hous\_total** | Number of housing units in the block |
| **hous\_own** | Number of owner-occupied housing units in the block |
| **hous\_rent** | Number of renter-occupied housing units in the block |
| **boroname** | Name of the New York borough. Manhattan, The Bronx, Brooklyn, Staten Island, Queens |
| **geom** | Polygon boundary of the block |

1. Logical Operators and Data Types

Question: Retrieve all census blocks where the total population exceeds 1000 and the number of owner-occupied housing units is greater than renter-occupied ones.

2. Data Query Language (DQL)

Question: List the total population and total housing units for each borough. Aggregate popn\_total and house\_total to each borough

3. Data Manipulation Language (DML)

Question: Increase the number of renter-occupied housing units by 10% in all blocks where the borough is 'Manhattan'.

4. Data Definition Language (DDL)

Question: Add a new column to the table to store the percentage of white population in each block.

5. Data Manipulation Language (DML)

Question: Populate the new column 'perc\_white' with the percentage of the white population relative to the total population.

6. Advanced Logical Operators

Question: Retrieve blocks where the total population is less than 500 or the percentage of the white population exceeds 70%.

7. Combining DQL and Spatial Queries

Question: Find all census blocks in 'Brooklyn' with a total population density greater than 10,000 people per square kilometer.

Hint: Using ST\_Area(geom) to calculate the Area of each census tract

8. Creating a New Table (DDL)

Question: Create a new table ‘borough\_summary’ to store summarized population data by borough.

|  |  |
| --- | --- |
| Column name | Data Type |
| boroname | VARCHAR(50) |
| population | INTEGER |
| housing | INTEGER |

9. Inserting Data (DML)

Question: Insert data into the new ' borough\_summary ' table.

Insert name of borough, Sum of popn\_total for each borough, Sum of hous\_total for each borough

10. Deleting Records (DML)

Question: Delete records from the ' borough\_summary ' table where the total population is less than 500,000.

Bonus: Using Subqueries

Question: Retrieve blocks where the total population is above the average population of all blocks.

Hint: To solve this, think about how to calculate the average population first using an aggregate function. Then, use a subquery to embed this average into the main query's WHERE clause to compare each block's population against it.

Hint: subquery syntax:

SELECT column\_name  
FROM table\_name  
WHERE column\_name expression operator   
 (SELECT column\_name FROM table\_name WHERE ...);