CS 166: Lab   Clustered Index Performance

In this lab, you are expected to study the difference between clustered and unclustered indexes in PostgresSQL.

You can find material about clustered indexing from these links:

1. PostgresSQL documentation

For this assignment you need to download and unzip lab10.zip from iLearn (make sure you unzip it to /extra/$USER/cs166/data directory, because you can easily exceed you disk quota).

For this assignment we are using a short version of DB schema from Lab7 with only single table (part_nyc). As usual *create_tables.sql* script is used to create initial table structure and populate table with initial data.

The purpose of the lab is to study the performance difference between clustered and unclustered indexes. To do that you need to create a secondary index and put corresponding SQL code to *create_indexes.sql*.

Also you would need to write SQL for clustering part_nyc table according to the index, you've just created.

Finally you have to write the query, which will be testing performance of clustered vs. unclustered index lookup. Note that to observe the difference you need to use range query, which will return significant number of results (order of 100 thousands).

Use *measure.sh* to gather performance numbers.