# **PostgreSQL Query Assignment**

# Introduction

PostgreSQL is an open source relational database management system. Large datasets can be created and maintained using PostgreSQL and queries to retrieve data can be integrated into R. This assignment is intended to introduce you to building tables in PostgreSQL and writing a query to select the set of cases that you would like to use in R.

## Directions

Use the lecture material from Lesson 7 as your guide for how to create tables and run queries in PostgreSQL.

The completed assignment should adhere to the following guidelines:

* 1. Include your answers on the assignment document.
  2. Write your answers using complete sentences with correct punctuation, grammar, and spelling.
  3. Submit your completed assignment through the Blackboard portal in Lesson 7.

Open pgAdminIII and right click on the database icon and create a new database. Name the database. Within the database, create a table for the City Health Dashboard data.

1. Create a table in pgAdmin for the City Health Dashboard data in the chdbtc.csv file (the City Health Dashboard data that was transformed in Lesson 2). Populate the table with data from the chdbtc.csv file. Save the CREATE TABLE code and the COPY code together as one .sql file. Paste the complete code below. (5 points)

CREATE TABLE chdbtc\_master

(

state character(2) NOT NULL,

city character varying NOT NULL,

absenteeism numeric,

pollution numeric,

binge numeric,

"breast.cancer" numeric,

"cardio.deaths" numeric,

"child.poverty" numeric,

"colo.deaths" numeric,

"dental.care" numeric,

diabetes numeric,

"frequent.mental" numeric,

"frequent.physical" numeric,

"blood.pressure" numeric,

highschool numeric,

"housing.costs" numeric,

"housing.lead" numeric,

inequality numeric,

"lead.risk.index" integer,

"life.expectancy" numeric,

"limited.healthy.foods" numeric,

birthweight numeric,

segregation numeric,

obesity numeric,

overdose numeric,

parks numeric,

inactivity numeric,

"premature.deaths" integer,

prenatal numeric,

preventive numeric,

diversity numeric,

smoking numeric,

"teen.births" numeric,

thirdgrade numeric,

unemployment numeric,

uninsured numeric,

"violent.crime" numeric,

walkability numeric,

CONSTRAINT city\_state\_index PRIMARY KEY (state, city)

)

WITH (

OIDS=FALSE

);

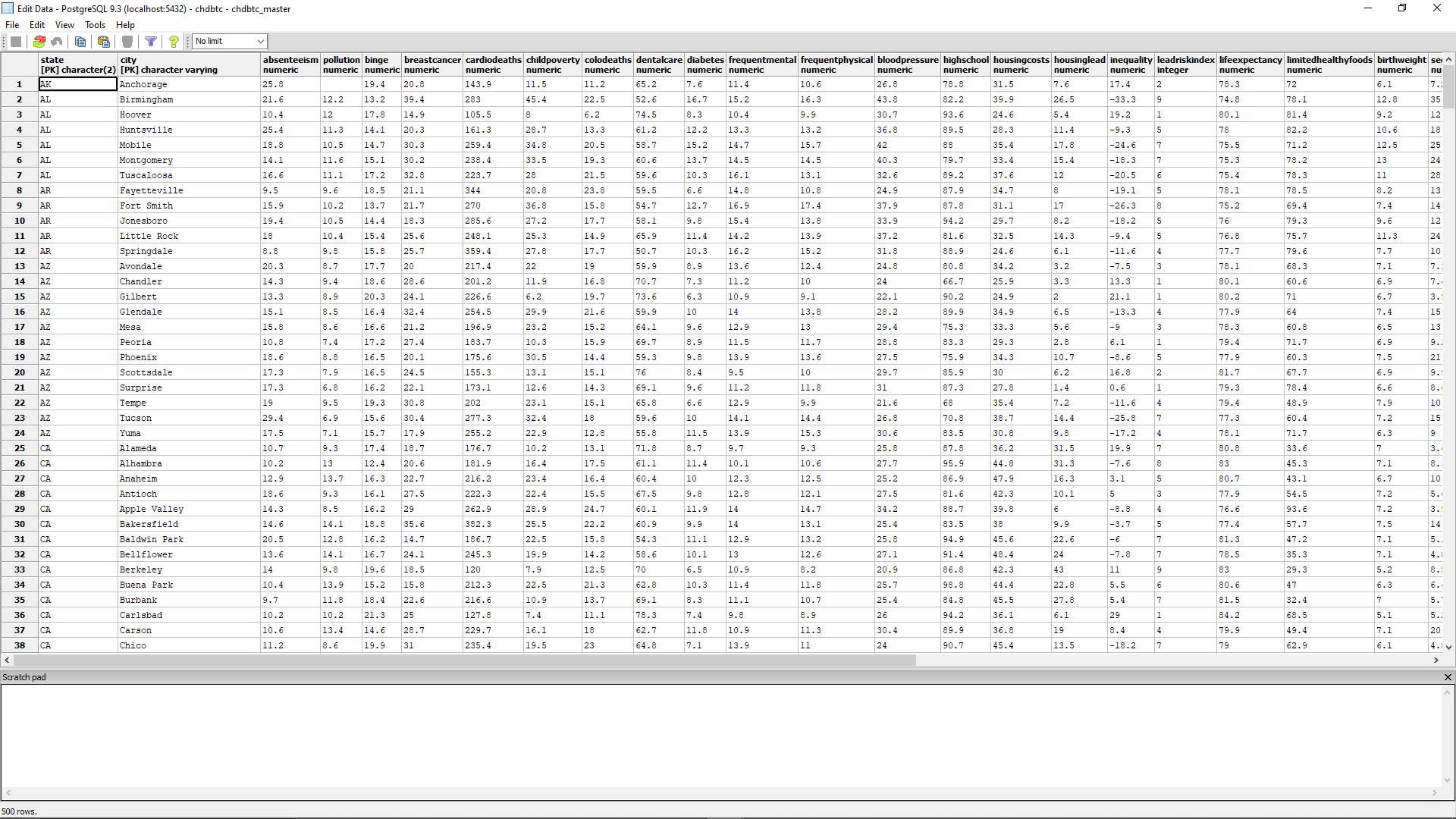
ALTER TABLE chdbtc\_master

OWNER TO postgres;

COPY chdbtc\_master FROM 'D:\PostgreSQL\Data\chdbtc.csv' DELIMITER ',' CSV HEADER NULL 'NA';

1. Use the spreadsheet icon to open the table. Create a screenshot of the data and paste it into your assignment. (2 points)

(next page)



1. Select the full table with a query and save as a CSV file named chdb.csv. Open the chdb.csv file in R and run dim(chdb) to confirm that you have correctly created all of the rows and columns in the dataset. Paste your results into your assignment. (3 points)

> chdb <- read.csv("C:/Users/carlb/Desktop/Week 7/chdb.csv")

> View(chdb)

>

> dim(chdb)

[1] 500 38

## Scoring

The assignment is worth 10 points. You must have a correct and complete answer to receive full credit.