Fall 2018 : CSCI 271 : Project 3   Working with Data

This project involves working with data in csv files modified from originals downloaded from the NYC Open Data website (https://data.cityofnewyork.us/browse). CSV (or comma separated values) files can be opened in a text editor, or in a program like Excel. Each line in the file represents a record of values, separated by commas.

Student Collaboration: Sharing ideas is fine, but sharing code is not. Do not copy code from other students, or let other students use your code. Projects that are identical except for superficial changes like different variable names may receive a 0.

Part 1:

a) Modify readcsv.cpp file to use the columns in your dataset. See readcsv_example.cpp for a program that works with a dataset of FDNY response times. (*If you need this to work with an online compiler, use readcsv_online.cpp and readcsv_example_online.cpp as a guide.*)

b) Modify your program by adding a function to analyze some aspect of the data. What you do will depend on the particular dataset you choose. You might add a function to add up values in a column, find an average value, find a count of some particular value, or report on some other aspect of the data.

Part 2:
- Create a short presentation of your results. (~ 3 slides)
  - What dataset you chose
  - Example of data (output of part 1a)
  - What you were able to find out in your analysis of the dataset (output of part 1b)

Due Date:

Wednesday, December 12, end of day: Submission of .cpp and pdf file

The following data sets are available. The data file should go in the same directory as your C++ source code file.

NYC Dog License (size = 37893)

2015 Tree Census (size = 25857)

2017-2018 Demographic Snapshot Schools (size = 1000)

DHS Daily Report (size = 1876)

NYPD Complaint Data (size = 6053)