# CS 211 PA #6: Data Analysis

In this assignment, you will write a program to analyze death statistics provided by the US government. This assignment was developed with the intent of showing how the skills you've acquired thus far in your academic career can be leveraged to address a real-world problem that you might be asked to complete by an employer. Just like in the real-world, I will provide overall requirements, but how you address my requirements is entirely up to you. Each line in the CSV file is formatted as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Year*** | ***Cause Name*** | ***State*** | ***Deaths*** | Age-Adjusted Death Rate |

The bolded columns are the cells that you will need to collate for this assignment. From this data, you must allow the user to make several queries:

1. List total deaths for a given state
2. List deaths by cause for a given state
3. List death rates by year for a given state and a given cause
4. List total death rates by year for a given state
5. List total deaths for each cause for all states
6. List total deaths for all causes for all states
7. List death rates by year for all states and a given cause
8. List total death rates by year for all states

Here is an example interface (bolded text was entered by the user):

|  |
| --- |
| \*\*\*Death Rate Data Analysis\*\*\*  Enter State (-1 for all states): ***California***  Enter Cause (-1 for all causes): ***Unintentional Injuries***  Do you want this data broken down by year (Y/N)?: ***N***  Total deaths in California from Unintentional Injuries in the years 1999 to 2015: 180908  Would you like to run another query (Y/N)?: ***N*** |

## Starter Code

You are provided with death\_rates.csv, which contains the death rates that you will use in your analysis. Again, I provide StringSplitter.h, which will greatly aid in your parsing of CSV files. All other code must be written by you.

## Possible Strategy for Getting Started

1. Again, the use of hash tables will be extremely useful in this assignment
2. If you want to maintain sorted order on a particular data set, consider using STL *map* as it ensures order. Note that *map* and *unordered\_map* are programmed in exactly the same manner.
3. Consider developing some sort of DataPoint class to contain a data entry. This class would have year, cause, state, and deaths properties.
   1. Furthermore, consider how such a DataPoint class might be used in an additional State or Injury class.

## Header Comment, and Formatting

1. Be sure to modify the file header comment at the top of your script to indicate your name, student ID, completion time, and the names of any individuals that you collaborated with on the assignment.
2. Remember to follow the basic coding style guide. A basic list of rules is included with this document.

# Reflection Essay

In addition to the programming tasks listed above, your submission must include an essay that reflects on your experiences with this homework. This essay must be at least 350 words long. Note that the focus of this paper should be on your reflection, ***not*** on structure (e.g. introductory paragraph, conclusion, etc.). The essay is graded on content (i.e. it shows deep though) rather than syntax (e.g. spelling) and structure. Below are some prompts that can be used to get you thinking. Feel free to use these or to make up your own.

* Describe a particular struggle that you overcame when working on this programming assignment.
* Conversely, describe an issue with your assignment that you were unable to resolve.
* Provide advice to a future student on how he or she might succeed on this assignment.
* Describe the most fun aspect of the assignment.
* Describe the most challenging aspect of the assignment.
* Describe the most difficult aspect of the assignment to understand.
* Provide any suggestions for improving the assignment in the future.

## Deliverables

You must upload your program and reflection as a ZIP file through Canvas no later than midnight on Friday, May 11, 2018.

## PA #6 Checkins

This assignment has no checkins.

## Grading Criteria

Your assignment will be judged by the following criteria:

### Reflection essay (10pts)

* Your reflection meets the minimum requirements as specified earlier in this document.

### Style (10pts)

* Your project contains good structure and implements the required classes. Your program intelligently uses classes when appropriate and generally conforms to good OOP design (i.e. everything isn't slapped into main).

### Tasks (10pts each; 80pts total)

* List total deaths for a given state
* List deaths by cause for a given state
* List death rates by year for a given state and a given cause
* List total death rates by year for a given state
* List total deaths for each cause for all states
* List total deaths for all causes for all states
* List death rates by year for all states and a given cause
* List total death rates by year for all states