# CS 211 PA #1: Data Analysis

In this assignment, you will write a program to analyze death statistics provided by the US government. Your program will prompt the user for a state, cause of death, and year. Your program must then output the total number of deaths that match the criteria. CSV (comma separated values) is a standardized clear-text format frequently used to represent spreadsheets. All popular spreadsheet applications (e.g. Excel) are capable of opening and manipulating CSV files. The CSV format is pretty straightforward. Each line represents a row in the spreadsheet. Each cell in a row is separated by a comma. 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, the user should be allowed to query death rates for a specific state, cause, and year. For example: (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***  Enter Year (-1 for all years): ***1999***  Total deaths in California from Unintentional Injuries in the year 1999: 9198  Would you like to run another query (Y/N)?: ***N*** |

In addition, the user should be able to specify a wildcard (everything) option, signified by a -1. Here's another example:

|  |
| --- |
| \*\*\*Death Rate Data Analysis\*\*\*  Enter State (-1 for all states): ***California***  Enter Cause (-1 for all causes): ***Unintentional Injuries***  Enter Year (-1 for all years): ***-1***  Total deaths in California from Unintentional Injuries in all years: 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. After lab 1, you will also be provided with code for parsing CSV files. All other code must be written by you.

## 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 Wednesday, February 6, 2019.

## Grading Criteria

Your assignment will be judged by the following criteria:

### Checkin (10pts)

* During Lab 2, you must demonstrate to me your program. In order to receive full credit, you must have made sufficient progress on your assignment. I define significant progress as demonstrating an expenditure of meaningful effort on one large section of code. Your checkin should demonstrate that you are capable of completing the assignment by the due date.

### Reflection essay (10pts)

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

### Program UI Format (10pts)

* Your program's UI must exactly match my sample UI on the first page of this assignment. Your program should first prompt for state, cause, and year. Your program should then ask for another query. Your program **should not** contain any different prompts nor should they be in any different order.

### Test Cases (7 total, 10pts each)

* I will randomly enter 7 combinations of data into your program. Correctly outputting each answer is worth 10 points.

## Grade Distribution

Your final grade for the assignment will be determined based on the number of points earned.

|  |  |
| --- | --- |
| Score (%) | Points Required |
| 100 | 85 |
| 90 | 80 |
| 80 | 70 |
| 70 | 55 |
| 60 | 40 |
| 50 | 25 |
| 40 | 20 |
| 25 | 15 |