Raspberry Jelly Jam

James Smith, Joelle Lum, Ryan Siu

APCS1 pd5

1/6/17

Final Project Proposal

Terminal CSV Tool

Description: CsvTool is a tool used on the command line to work with and create csv data. This tool allows a user to sort columns, perform inner and outer joins on two csv files or columns, search for words or numbers in a csv, selection of specific columns for output, and perform mathematical operations on columns such as addition, subtraction, multiplication, division, and operations relating to statistical analysis. CsvTool will also allow for a user to print out csv's in a nice ASCII format (see below). The idea behind this project is to give a user a quick and easy way to work with CSVs through the terminal so they can also utilize the functionality and convenience of terminal commands.

In order to read/write the data from/to the CSV file, we will use the FileInputStream and FileOutputStream classes from the java.io package.

Get
- Returns the value at a cell given input coordinates

Set
- Changes the value at a cell given input coordinates and the new value
- Returns the old value

Delete
- Accepts a coordinate and changes the value at that cell to null
- Accepts a row number and shifts all the rows below it up 1
- Accepts a column number and shifts all the columns to the right of it left one

Sort - Allows a user to sort columns in ascending or descending order. Sorts strings alphabetically and by capitalization.
- Outputs a sorted version of the CSV

Inner Join - Takes in two columns from the same or different CSV files and outputs a column that is the intersection of the two columns.

Outer Join - Takes in two columns from the same or different CSV files and outputs a column that is the union of two columns

Search
- Takes in a digit or string and outputs the entire row that has the given input in them.

● Takes in a digit or string and outputs the coordinates that have the given input in them.

Mathematical Operations -
- Addition: Will output the sum of 2 columns(adds across rows)
  ○ If given only a single column it will output the sum of the column
- Subtraction: Will output the difference of two columns
- Multiplication: Will output a column that is the outcome of multiplying across the rows of 2 columns.
- Division: Will output a column that is the outcome of dividing across the rows of 2 columns.
- Statistics:
  ○ 5 Number Summary: Outputs the minimum, 1st quartile, median, 3rd quartile and maximum of a column
    ■ Box and whisker graph
  ○ Average: Outputs the average of a column
  ○ Standard Deviation: Outputs the standard deviation of a column
  ○ Correlation between two data sets, as specified by the user

Add Column - Accepts a CSV column and adds it to a given CSV file and outputs the new CSV
- Can be added at a specific position or will default to the end

Add Row - Accepts a CSV row and adds it to a given CSV file and outputs the new CSV
- Can be added at a specific position or will default to the end

Pretty Printing - Outputs a CSV in an organized ASCII based grid (see below for an example)

Example syntax: java ourCSVTool oldFile.csv --select_column 2,3,4 > mynewfile.csv
- This would select columns 2, 3, and 4 from oldFile.csv and then redirect the output from the program to mynewfile.csv where mynewfile.csv would now be a CSV file containing columns 2,3,4.

Action Plan:
csvtool
- code DataTable class that represents the csv file
- make tests - make some small csv files that are the input/output of each
  of the functions below (csvtool --tests)

- prettyprint(DataTable)
- join(DataTable 1, Datable 2, int column1, int column2)
- sort(Datatable, List<int> columns)
- search
- select
- arithmetic operations

- code argument parsing
- addColumn

--- basic string handling
- string handling - go back to each function and make it give
  a nice error message if it can't handle strings (arithmetic)
  or make sure it works with strings (e.g. search, sort)

--- better string handling
- better escaping (look online for csv standards, escaping, regex)

prettyPrint output:

```
== | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S |
== |===============================================================================|
1  |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
== |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
2  |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
== |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
3  |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
== |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
4  |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
== |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
5  |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
== |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
6  |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
== |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
7  |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
== |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
8  |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
== |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
9  |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
== |___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
~
```