CS 46B

Introduction to Data Structures

Summer Semester 2015

Department of Computer Science
San José State University
Instructor: Ron Mak

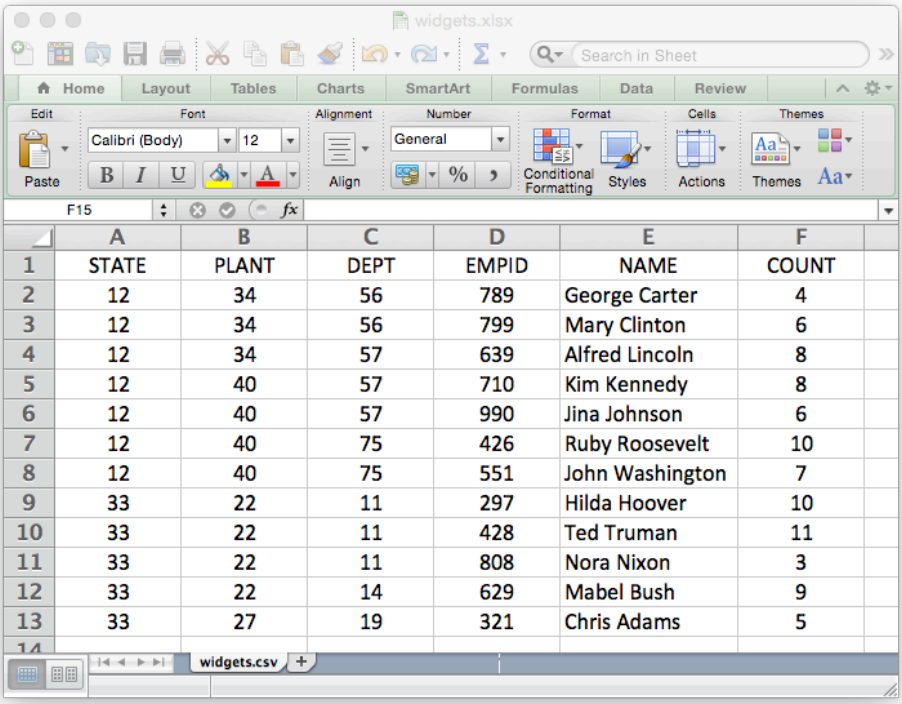
Homework #3

Generate a Detail Report from a CSV File

Assigned:	Tuesday, June 9
Draft due:	Saturday, June 13 at 11:59 PM
Codecheck URL:	http://codecheck.it/codecheck/files/150611070629f45sczkm70ffwuedde6p8ed
Canvas:	Homework 3 Draft
Points:	6 points max
Final due:	Monday, June 15 at 11:59 PM
Codecheck URL:	http://codecheck.it/codecheck/files/1506120526avo37t7yjlh6dyh21xjloss7e
Canvas:	Homework 3 Final
Points:	18 points max

This assignment will give you practice reading a CSV (comma-separated values) file in order to generate a detail report using Java's **`Scanner`** and **`PrintWriter`** classes.

CSV is a text file format saved from a spreadsheet. Here's the Excel spreadsheet **widgets.xlsx** that is the source of the CSV file that we'll use:



	A	B	C	D	E	F
1	STATE	PLANT	DEPT	EMPID	NAME	COUNT
2	12	34	56	789	George Carter	4
3	12	34	56	799	Mary Clinton	6
4	12	34	57	639	Alfred Lincoln	8
5	12	40	57	710	Kim Kennedy	8
6	12	40	57	990	Jina Johnson	6
7	12	40	75	426	Ruby Roosevelt	10
8	12	40	75	551	John Washington	7
9	33	22	11	297	Hilda Hoover	10
10	33	22	11	428	Ted Truman	11
11	33	22	11	808	Nora Nixon	3
12	33	22	14	629	Mabel Bush	9
13	33	27	19	321	Chris Adams	5

After the headers row, each row contains detail information about an employee: state code, plant (factory) code, department code, employee ID, name, and the count of widgets the employee made. The plants in a state are separate from the plants of another state, and the departments in a plant are separate from the departments of another plant. (Department 57 in Plant 34 is separate from Department 57 in Plant 40.)

The detail rows are already sorted first by state, then by plant, then by department, and finally by employee id.

Here's the CSV text file **widgets.csv** saved from the spreadsheet:

```
STATE,PLANT,DEPT,EMPID,NAME,COUNT
12,34,56,789,George Carter,4
12,34,56,799,Mary Clinton,6
12,34,57,639,Alfred Lincoln,8
12,40,57,710,Kim Kennedy,8
12,40,57,990,Jina Johnson,6
12,40,75,426,Ruby Roosevelt,10
12,40,75,551,John Washington,7
33,22,11,297,Hilda Hoover,10
33,22,11,428,Ted Truman,11
33,22,11,808,Nora Nixon,3
33,22,14,629,Mabel Bush,9
33,27,19,321,Chris Adams,5
```

widgets.xlsx: http://www.cs.sjsu.edu/~mak/CS46B/assignments/3/widgets.xlsx
widgets.csv: http://www.cs.sjsu.edu/~mak/CS46B/assignments/3/widgets.csv

Draft

Write a Java program that uses the **`Scanner`** class to read the CSV file and the **`PrintWriter`** class to output a simple detail report **widgets.out**:

```
STATE PLANT DEPT EMPID COUNT NAME

   12    34   56   789     4 George Carter
   12    34   56   799     6 Mary Clinton
   12    34   57   639     8 Alfred Lincoln
   12    40   57   710     8 Kim Kennedy
   12    40   57   990     6 Jina Johnson
   12    40   75   426    10 Ruby Roosevelt
   12    40   75   551     7 John Washington
   33    22   11   297    10 Hilda Hoover
   33    22   11   428    11 Ted Truman
   33    22   11   808     3 Nora Nixon
   33    22   14   629     9 Mabel Bush
   33    27   19   321     5 Chris Adams
```

Draft **widgets.out:** http://www.cs.sjsu.edu/~mak/CS46B/assignments/3/draft/widgets.out

Draft logic tips

You can use a file **`Scanner`** object to read the CSV file one line at a time. Construct a string **`Scanner`** object from the line you just read. For example, for the column headers:

```
String detailRecord = in.nextLine();
Scanner detail = new Scanner(detailRecord);
```

Now read the values from the string **`detailRecord`** using the string **`Scanner`** object **`detail`**. What delimiters should you use?

You might also consider using the **`split()`** method of the **`String`** class.

Draft submission

Complete this file:
http://www.cs.sjsu.edu/~mak/CS46B/assignments/3/draft/Widgets.java
The Codecheck URL:
http://codecheck.it/codecheck/files/150611070629f45sczkm70ffwuedde6p8ed

Canvas: Homework 3 Draft
Due Saturday, June 13 at 11:59 PM

Final

Generate a detail report **widgets.out** that contains **multi-level totals**:

```
STATE  PLANT  DEPT  EMPID  COUNT  NAME

   12     34    56    789      4  George Carter
   12     34    56    799      6  Mary Clinton

                              10  TOTAL FOR DEPT  56 *

   12     34    57    639      8  Alfred Lincoln

                               8  TOTAL FOR DEPT  57 *
                              18  TOTAL FOR PLANT 34 **

   12     40    57    710      8  Kim Kennedy
   12     40    57    990      6  Jina Johnson

                              14  TOTAL FOR DEPT  57 *

   12     40    75    426     10  Ruby Roosevelt
   12     40    75    551      7  John Washington

                              17  TOTAL FOR DEPT  75 *
                              31  TOTAL FOR PLANT 40 **
                              49  TOTAL FOR STATE 12 ***

   33     22    11    297     10  Hilda Hoover
   33     22    11    428     11  Ted Truman
   33     22    11    808      3  Nora Nixon

                              24  TOTAL FOR DEPT  11 *

   33     22    14    629      9  Mabel Bush

                               9  TOTAL FOR DEPT  14 *
                              33  TOTAL FOR PLANT 22 **

   33     27    19    321      5  Chris Adams

                               5  TOTAL FOR DEPT  19 *
                               5  TOTAL FOR PLANT 27 **
                              38  TOTAL FOR STATE 33 ***

                              87  GRAND TOTAL        ****
```

The report shows the total widgets made by each department, plant, and state, and a grand total.

Final **widgets.out:** http://www.cs.sjsu.edu/~mak/CS46B/assignments/3/final/widgets.out

Final logic tips

Since the input detail records are already sorted, you should be able to generate the detail report line by line as you read the input CSV file line by line. You should not need to read and store all the input lines into an array list, which would be impractical if the input file contains millions of records.

Look carefully at the final detail report. What is the trigger to print a department total? a plant total? a state total? In what order should you check these triggers?

A state change implies a plant change, and a plant change implies a department change. How should your code exploit this relationship for the multiple levels of totals?

Final submission

Complete this file:
http://www.cs.sjsu.edu/~mak/CS46B/assignments/3/final/Widgets.java
The Codecheck URL:
http://codecheck.it/codecheck/files/1506120526avo37t7yjlh6dyh21xjloss7e

Canvas: Homework 3 Final
Due Monday, June 15 at 11:59 PM