

## Programming Languages - Project

WJ Inc., is a company with offices around the world. The HR team wants to build reports on the number of employees around the world, in different departments. But they do not want to use the raw data for the reporting purpose. They want some consolidation and aggregation to be done prior to being loaded into their reporting system. The data is provided to the software engineer in CSV files. There are 3 CSV files, each holding the department, location and employee headcount data.

The department and location information is provided as a flattened hierarchy. Given below is a snapshot of the department file content:

level1	level2	level3	level4	level5	level6	level7	bottom_level
WJ Group	WJ Inc.	DEPT_1	DEPT_11	DEPT_111	DEPT_1111	DEPT_11111	DEPT_11111
WJ Group	WJ Inc.	DEPT_1	DEPT_12				DEPT_12
WJ Group	WJ Inc.	DEPT_1	DEPT_13	DEPT_133			DEPT_133

The department hierarchy has 7 levels. The file has columns level1 - level7 and an additional column to indicate the the bottom level. The department hierarchy is a jagged hierarchy meaning not all hierarchies have 7 levels.

Similarly the location data is also provided. The screen shot is given below:

level1	level2	level3	level4	level5	level6	level7	level8	bottom_level
World	North America	USA	California	Northern California	San Jose	Main Campus	SJC 1	SJC 1
World	North America	USA	California	Northern California	San Jose	Main Campus	SJC 2	SJC 2
World	North America	USA	California	Southern California	Los Angeles	LA Techspace	LA 1	LA 1

The head count data includes the department, location. Each row in the file gives the number of employees for a particular department at a particular location. The headcount data is only available at bottom levels. Given below is a screenshot of the data:

A	B	C
location	department	head_count
SJC 1	DEPT_11111	12
SJC 1	DEPT_12	42
SJC 1	DEPT_133	122
SJC 1	DEPT_2222	3
SJC 1	DEPT_2244	12
SJC 1	DEPT_22455	5

The headcount data needs to be aggregated to level5 (district) for locations, and departments. The output should also be in CSV format for being easily uploaded to the reporting system. Should look something like:

location	department	head_count
Northern Californ	DEPT_111	58
Southern Californ	DEPT_111	94
Kuching District	DEPT_111	25

**Instructions:**

1. You are free to implement the application in any of the languages studied in the course, and also in addition Java and Python are also acceptable.
2. Code should be well commented to explain the logic being used.
3. Functional programming paradigm needs to be used (use recursive functions, pattern matching, higher order functions, etc. No imperative programming - no loops)
4. Clear instructions should be included in submission on how to run the application. (.txt file)

**Rubric:**

Criteria	Weightage
Application runs and produces expected output	45%
Functional programming paradigm is used consistently	20%
Code is well commented and understandable	15%
Code is well formatted and readable	10%
Meaningful names used for variables, functions, classes, etc	10%
	100%

**\* Clear instructions should be included in submission on how to run the application. If not included, I might consider this as “application does not run.”**

**\* If the application does not run, I will give credit based on comments given and my understanding of the comments and code.**