

Assignment 1

Due Date: Oct 23, 2020 at 11:59PM

Notes:

- This assignment should be finished in Java.
- Zip all your files (code files, readme file, test data files, and others if needed) together into one file named “CPSC473_a#_yourlastname_yourfirstname_student#.zip” (or .rar, or .7z, or .tar), then submit this package to blackboard.
- The format of input data will be discussed in class.

Assignment Description:

The main task of this assignment is straightforward: implement an Apriori algorithm for frequent pattern mining.

Detail requirements for Apriori:

1. Your program should be able to read transaction database from a .txt file.
2. The format of transaction database file will be discussed in class.
3. All possible input variables should be entered by command-line arguments.
 - a. For example, if, 1) the program is named “apriori”, 2) the file path of database is “./testdata.txt”, 3) the minimum support threshold is set to 50%. The command to run your program can be:

```
java apriori ./testdata.txt 50
```
 - b. If you are not sure how this can be done, I can briefly talk about it in class.
4. You should provide a README file for program compilation and execution.

About bonus marks:

- To get the bonus marks of this assignment, you need to document any special enhancement you included in the implementation. For example, you used some special implementation tricks to deal with candidate generation, which, can improve the performance compare with the traditional method.
- The document does not have to be long, as long as you clearly explain a) what you did, and b) why do you think this method can improve the performance.
- The document does not have to be formal, an informal report should be good enough.
- The more special enhancements you include in your implementation the better chance you will get full bonus marks.
- The document should be in PDF, and named as “BonusMarkDocumentation.pdf”.