İZMİR EKONOMİ ÜNİVERSİTESİ

SE 375

Systems Programming

2022-2023 Spring

Lab Assignment #2

1. Introduction

Assume that each district (Konak, Gaziemir etc) of İzmir reports their data on public transport usage and the overall data set is collected for analysis. The data set from each district contains the daily number of bus passengers and number of bus trips. An example view of a data set from one district (say Konak) is as following;

Date	Day of the Week	Number of Trips	Number of Passengers
01.02.2023	Wednesday	34	1242
02.02.2023	Thursday	51	4589
03.02.2023	Friday	30	2568
04.02.2023	Saturday	48	3874
05.02.2023	Sunday	36	2514
06.02.2023	Monday	44	6257
07.02.2023	Tuesday	27	2674

2. Task

In this lab assignment, you are going to develop a Java Application that will calculate statistics on given data. Your application will read each data file, compute and store the values and report to the user.

a. Specifications

- Each data file contains the daily number of bus trips and number of bus passengers in a CSV (comma-separated values) format for each district. The first line of the file has the headers in the order of:

  Date, *Day of the Week*, *Number of Trips* and *Number of Passengers*

  and the remaining lines include data for each date of the month.
- `DD.MM.YYYY` is the date format.
- Day of the week is one of the below;

  Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday

- The data type of *Number of Trips* and *Number of Passengers* is integer. It is guaranteed that the data files do not have missing values.
- Your application should use a single `HashMap` to store the statistics where keys are the district names and the values are the total number of passengers and the average number of passengers over the month.

- Your application should create a thread for each data file and give the file name as an argument.
- Each thread should **read** the given data file.
- Each thread should **calculate** the total number of passengers over the month.
- Each thread should **calculate** the average number of passengers over the month.

$$Average\ Number\ of\ Passengers\ = \frac{Total\ Number\ of\ Passengers}{Number\ of\ Days}$$

- Each thread should **display** the results on the console.
- Each thread should **return** both results as a list of numbers.
- Once threads are finished, the results they return must be **aggregated** and **stored** in the `HashMap` correctly.
- The `HashMap` should also keep the total and average number of passengers for all districts combined. You may assume that every file has the same number of records within, when computing the average in this case.

b. Example Runs

The result should look like the following:

```
Thread finished: Balcova - Total: 35357.0, Average: 1140.55
Thread finished: Konak - Total: 141996.0, Average: 4580.52
Thread finished: Gaziemir - Total: 96934.0, Average: 3126.9
Thread finished: Bornova - Total: 86956.0, Average: 2805.03
Thread finished: Karsiyaka - Total: 62241.0, Average: 2007.77
Thread finished: Seferihisar - Total: 24193.0, Average: 780.42
Thread finished: Urla - Total: 25069.0, Average: 808.68
Thread finished: Karabaglar - Total: 132160.0, Average: 4263.23

Threads are complete. Aggregating results...

Here are the statistics for the number of passengers:

District            Total Passengers     Avg Passengers
--------            ----------------     --------------
Konak               141996               4580.52
Karsiyaka           62241                2007.77
Seferihisar         24193                780.42
Balcova             35357                1140.55
Urla                25069                808.68
Gaziemir            96934                3126.9
Karabaglar          132160               4263.23
Bornova             86956                2805.03

All districts       604906               19513.1
```

These are the correct results; you may compare your results with these numbers to check whether they are correct.

The results displayed by the individual threads might be in a different order since threads may have different running times at each run of the program.

3. Regulations

- You should use Java.
- Late submission is **not** allowed.
- This is an **individual** assignment. Plagiarism, copying, cheating, outsourcing the exam to another person or organization for with or without pay are considered as actions of academic dishonesty. Failure to maintain academic honesty may result in disciplinary

action according to the Izmir University of Economics' disciplinary bylaw for students of institutions of higher education (https://www.ieu.edu.tr/en/bylaws/type/read/id/13).

- Submissions should be done to Blackboard. Submissions via e-mail will not be accepted.
- You should make sure to obey the specifications.

4. Submission

- You should submit a zip file to Blackboard containing your Java files.
- By submitting you pledge the following; *"On my honor, as an Izmir University of Economics student, I affirm that I have not given or received any unauthorized help on this exam, and that this work is my own."*