Algorithms and Data Structures

Processing Election Results with Functional Interfaces and Streaming

Introduction

Being impressed by the efficient processing of the votes in the US elections of 2020, Dutch officials decide to make a prototype of similar continuous counting and publishing of votes and results. As a starting point they have acquired the intermediate results of all counties and for every candidate running for president.

One of the goals of the prototype is to investigate the design that makes outscaling to massive parallel processing of the counting process achievable. This time the software will be ready for a pandemic!

Exercise

The application must perform some basic data gathering about the ongoing voting process. Reading and converting of the prototype election data has been implemented already.

A dataset from the US 2020 elections is provided as a project resource with the name "president_county_candidate.csv" The provided framework converts the election data into a iterable table of rows. The table is implemented by the class CsvTable and the rows are implemented by the class CsvRow. The CsvRow can be converted into a business object using the class CandidateCountyVotes. This class represents the counted votes for one candidate in one county.

You must implement the methods of the StatisticsPrinting class in order to realize the functionality.

The following requirements must be implemented.

1. Print the total number of votes counted.
   - Convert the CsvTable collection into a stream of CsvRows.
   - Convert the stream of CsvRows into a stream of CandidateCountyVotes instances.
   - Summerize the getTotalVotes() of the resulting stream and print that number to the standard output.
2. Print the total number of votes given to the winner of a county.
3. Print the total number of votes which are given to the other candidates.
4. Print the votes for each party summed over all counties and states.
   - You are required to use at least one merge() in this method.

The output must be as follows.

```
Total votes all counties of USA: 153.667.887
Total votes for the winning candidates: 96.807.421
Total votes for the losing candidates: 56.860.466
Votes per party: {ALI=80169, APV=407, ASP=24093, BAR=5386, BFP=143, BMP=127,
CST=61233, DEM=78220235, GOP=1157, GRN=378697, IAP=809, IND=120765, LIB=1808321,
LLC=668, LLP=1366, NON=13419, OTH=546, PRG=5278, PRO=4678, PSL=77504,
REP=72691128, SEP=192, SWP=6723, UNA=3399, UTY=6279, WRI=155165}
```

Design Constraints

- It is forbidden to use a for-loop. Otherwise it is not possible to achieve scalability.