Dr. Sandro Andreotti

Programming for Data Science

Winter term 2023/2024

Assignment 2
Issued: 2023-10-18

1 Summary

Write a python program that reads a csv file and:

- for every column computes the min, max, mean, median and variance
- writes a new csv file containing the summary statistics (see example file for output format)

You can expect that the input file contains only numeric columns, has no missing values and a header row.

2 GroupBy

Write a python program that reads a csv file and receives three string arguments

- <agg>: string argument from the set `(min, max, mean, sum)`
- <gcol>: string defining the grouping column name
- <acol>: string defining the aggregation column name

Your program will group the rows according to the unique values in the grouping column and aggregate the values in the aggregation column according to the provided function. The group values together with their aggregation results shall be printed to the command line in a two column user friendly format. See example below!

Remark:

For Exercises 1 and 2 your program should **not** use numpy, pandas or any other third part library directly providing the required functionality.

Table 1: GroupBy example

col1	col2	...
”A”	5	...
”B”	7	...
”C”	4	...
”B”	3	...
”A”	3	...

→ <gcol> = col1, <acol> = col2, <agg> = `sum` →

col1	col2
”A”	8
”B”	10
”C”	4

3 Palindrome detector

Write a python program that reads a text file and returns all words in the file that are palindromes together with their occurence count. The palindrome check should be case insensitive!