Programação

2020/2021

Project 6

Data limite de entrega: 17 de maio, 23h:59m; Peso de 1 valor em 20 valores na nota final

Description

Consider a CSV file with various information about books (at least the following information should be included: ISBN of the book, title, author(s), genre, number of pages, year of publication). Also consider another file with only two fields: the ISBN of the book and the respective number of words it contains.

The first file can be manually constructed or consider a small part from those available at https://github.com/zief0002/epsy-8251/raw/master/data/goodreads.csv or at https://www.kaggle.com/jealousleopard/goodreadsbooks. For the construction of the second file, you may consider information such as that found at https://blog.fostergrant.co.uk/2017/08/03/word-counts-popular-books-world/.

Perform the following tasks (you should always use matrices; the use of lists in any situation will be considered invalid):

Read the content of the two files into two matrices and display them on the screen;

Cross the information from the first file with that of the second, obtaining a single matrix, which should be saved in a new CSV file. If the second file does not

contain information about the word count of a specific book listed in the first file, ask the user to input this value.

Construct a matrix containing information about the minimum, maximum, and average number of words of books for each genre and save it in a CSV file;

Graphically display the information about the number of books published each year.

Grading

The maximum grade for the practical work is 1 point (out of 20 points). Each group (two students) must develop a solution and submit it on Inforestudante by 11:59 PM on May 17, 2021. They must submit a .zip file containing a .py or .ipynb file and a .pdf of the code.

There are four grading levels:

0 – the solution is incorrect;

0.5 - intermediate situation, i.e., a solution with a partially incorrect result, or with an incorrect result, but an algorithm very close to being correct;

0.75 - intermediate situation, i.e., a solution with a correct result, but with unoptimized code or that does not make adequate use of the materials taught in class;

1 - the solution is correct, with optimized code and makes appropriate use of the knowledge taught in class;

Defense rules:

each work is subject to a defense, and each group member must be able to explain the solution and answer questions from the faculty;

defenses take place in the first class after the submission deadline;

in online classes, students must have their camera and sound activated during the defense;

after the defense, the grade awarded to each student may be lower than the grade for the work, depending on individual performance during the defense.