Big Data Analytics

Description

Please carefully read the following instructions:

1. The assignment starts on May 30th and ends on June 5th at 23:59h.

2. The notebook is available on Moodle and contains the full set of instructions. You will need to import this notebook to databricks and perform all exercises in databricks to complete the assignment.

3. It is acceptable that you discuss with your colleagues' different approaches to solve each step of the problem set, but the assignment is individual. That is, you are responsible for writing your own code, and analyzing the results. Clear cases of cheating will be penalized with 0 points in this assignment.

4. After review of your submission files, and before a mark is attributed, you might be called to orally defend your submission.

5. Please read all instructions in the notebook carefully, especially instructions on how to correctly deliver your completed assignment.

6. The statement of authorship required is included on the next page.

Questions – Any questions about this assignment should be posted in the Forum@Moodle so that everybody can see and get inspiration from your questions.

Statement of Authorship

I, Beatriz Gonçalves, with the student number 20210695, attending the Big Data course as part of the master program in Data Science and Advanced Analytics at the NOVA IMS, state that:

1. I fully understand the rules of the assignment.

2. All the materials I have submitted have been prepared by myself, even if I had discussed their preparation with my colleagues.

3. All the analysis and answers in the attached document have been conducted and written by me.

4. I understand that this is an individual assignment and that anyone caught cheating will obtain a mark of 0 points for this assignment.

Lisbon, May 30th

Beatriz Gonçalves