Programming Workshop Presentation

Assignment Description

1. The purpose of this assignment is to apply what you have each learned during the DataCamp courses to a real data set
2. This assignment is to be done in groups of 2-3 students
3. Each group must choose a data set with up to two time periods and up to three variables
4. The data is to be analyzed as follows:
    - A cluster analysis involving all variables at each period
    - Box plots of each variable at both the aggregate and sub-sample levels
    - A scatterplot matrix of all variables at both the aggregate and sub-sample levels
    - Data manipulation and generation of graphs are to be done in Python
5. Each presentation must involve...
    - an explanation of any data types/structures, control flow commands, functions, and modules used in the code as well as how they work
    - a successful demonstration of the code
    - commentary on the results of the plots generated
    - Bonus: a comparison of the R and Python code used
6. Do not forget to also cite all works that you reference/use
7. The group must also submit their presentation slides with a correctly filled in *Statement of Authorship* attached through WhatsApp or e-mailed to lukman.hanif@ui.ac.id
8. Slides are to be in pdf form with the following file name format: [groupname]_slides.pdf (e.g. "01_slides.pdf")

Statement of Authorship

"I/we the undersigned declare that the attached assignment is purely our own work. It does not contain any works by others without our giving them full credit with proper citation.

This assignment is not/has never been presented/used as part of an assignment for another subject unless I/we clearly state that I/we have done so.

I/we understand that the this assignment I/we submit may be reproduced or communicated for the purposes of plagiarism detection."

Name(s)   :

NPM(s)   :

Signature(s)   :

Subject   :

Assignment Title   :

Date   :

Subject Lecturer   :

(To be filled in by all group members)