

Pre Practical Session #5

Data Analytics and Visualization

Tim Materi Pengenalan Komputasi 2020/2021

11 Desember 2020

Instructions

1. Do this module according to the tested material (Data Analytics and Visualization). Don't use other materials that have not been tested.
2. Pay attention to the file extension (*.ipynb). Source code without extension may not be graded.
3. Make sure your program can be compiled and run.
4. For every source code, provide your identity, at least:

# SID/Name	:
# Date	:
# Description	:

5. All source code should be compressed with filename format **H05.SID.zip**.
6. You may assume user inputs are valid, except if it is explicitly stated in the statement.
7. It will be better if your code has good indentation and comments (usage of variable, control, loop, function, or procedure), so that the debugging process will be easier.
8. Cheating (copying and pasting) from other participant(s) or other source(s) will not be tolerated and you will get a serious penalty.
9. If there are any differences between the module's instructions and your assistant's instructions, please follow your assistant's.
10. Good luck!

Problem 1

Save with filename: **H05.SID_01.ipynb**.

You have the final exam scores of ITB students in `nilai_uas.csv`. Write:

1. How many data are there?
2. First 10 data.
3. The 50th data until 60th (inclusive).
4. Number of students for each faculty.
5. Correlation coefficient of the calculus score and the physics score. Is it negatively correlated, positively correlated, or not correlated at all? Explain.

The description of data:

- "name": name of the student
- "fakultas": the faculty of the student
- "nilai_kal": the calculus score
- "nilai_fis": the physics score
- "nilai_kim": the chemistry score

Problem 2

Save with filename: **H05.SID_02.ipynb**.

You have the final exam scores of ITB students in `nilai_uas.csv`. Write:

1. Calculus, physics, and chemistry scores of Tuan Mor.
2. Student with highest physics score. (if there are multiple students, show all of them).
3. 10 students with highest chemistry score.
4. Number of calculus scores below 50.
5. Number of students inside FMIPA faculty.

The description of data:

- "name": name of the student
- "fakultas": the faculty of the student
- "nilai_kal": the calculus score
- "nilai_fis": the physics score
- "nilai_kim": the chemistry score

Problem 3

Save with filename: **H05.SID_03.ipynb**.

You have the final exam scores of ITB students in `nilai_uas.csv`. Write / plot:

1. Histogram distribution of calculus scores.
2. Horizontal bar chart, showing number of students for each faculties.
3. Pie diagram, showing number of students for each faculties.
4. Based on the pie and horizontal bar chart, which faculty has the most students? Which chart is better at showing the faculty with most students? Why?
5. Scatter plot with chemistry as the x axis and physics as the y axis.

The description of data:

- "name": name of the student
- "fakultas": the faculty of the student
- "nilai_kal": the calculus score
- "nilai_fis": the physics score
- "nilai_kim": the chemistry score