

DSCI 6607 – Python Data Analysis Project Guidelines

To demonstrate your understanding of programmatic data analysis with Python, you are encouraged to perform a data analysis project with Python. Please note that

- **Each person should submit their own project.** While you may discuss ideas and troubleshoot code with each other, each person must submit a unique project.
- You have to do your data analysis project by Python and report your project in python notebook (.ipynb). Otherwise, your project is not accepted.

Project Minimum Requirements:

- 1 .ipynb file and the .html file (or its .pdf file).
- While your project report should not be skimpy, there is flexibility for the layout of your report. To help you in this case, based on statistical topics we learned in this course, I encourage you to organize your report as follows
 1. An ‘Introduction’ section describing the dataset (and its background), describe statistically the types of the variables and what graphical displays and statistical measures are appropriate for the variables.
 2. A ‘Data analysis’ section where you show your python codes and graphical displays and data processing. Code should have accompanying text that presents what you claimed in the introduction section.
 3. A ‘Conclusion’ section to summarize the results of the analysis, as well as any deviations from the project.
 4. Note that, in data analysis, your interpretations are by far more important than just graphical displays and python codes. You, as a data scientist, must view and present the graphical displays and codes as the support of your interpretation.

Project Deadline:

The project report is due by **Wednesday December 13, 2023 at 5:00 pm**. You should submit your project as a zip file (including .ipynb file and its .html/.pdf file) to D2L dropbox.