

Multivariate Analysis

Final Project Timeline and Notes (**updated version**)

5/14/2021 (Organize your team):

- Send me the **Names and IDs** of your team members by email **before 5/14**.
The number of members for each team is **4** (except for two teams with size **3**).

5/30/2021 (Find Your Data and Prepare a Project Proposal):

- Each team must **send me a proposal (in pdf format, limited to 2 pages)** of your final project **before 5/30**. In addition, the proposal must include:
 - (1) A brief description of **data set** you will use in your final project. The data set must come from **an eligible source**, has **more than 1000 observations** and **at least 10 variables** (with both **categorical** and **numerical** attributes).
 - (2) Describe the goals and problems of interest in terms of **both supervised** (i.e. classification, **do not perform regression analysis**) and **unsupervised learning** (i.e. clustering, PCA, CA/MCA, FA, etc) based on the data.

6/24/2021 (Turn in the Final Report):

- The final report should be **limited to 12 pages** (not including the title page). **Please send me your report by email (hungy@nccu.edu.tw, in pdf format) before 12:00 noon of 6/24.**
- **The 15-min power point presentation is cancelled.**

The write up of your final report should be formal and include:

- A title page with detailed information of your team members (names and ID).
- Introduction, motivation and a quick summary of the problems you explore.
- Notation, assumptions and necessary data exploration/preprocessing.
- Description of techniques you use for solving the problems of interest. **The techniques should be those you have learned from class – unless it has better performance than all you have learned.**
- The key results and necessary numerical and graphical outputs.
- Conclusion, comments on the meaning of the results, and its implications. If possible, highlight some open questions.

Grading Policy: The grade of your final report depends on:

- Importance/Showcase/Innovation/Influence of the problem introduced
- Quality of writing (do you follow all the guidelines mentioned above?)