



Project – Multivariate Analysis
(MEDS & MMA, 1st Semester, 2019/2020)

Handed out on December 01, 2019.

To be handed back by January 3, 2020.

Presentations and discussion on January 20, 2020.

1. Find a labelled dataset that you have interest in its analysis.
2. Make a preliminary analysis of the data and discuss what you have learned from this analysis.
3. Solve your classification problem using supervised learning methods. Have in mind that some of the input variables may be irrelevant to the classification problem and that you may need to do some preprocessing methodologies of your data set e.g. dimensionality reduction techniques.
4. Apply unsupervised methods to your data. Interpret the results. Compare the obtained partitions with the true class each object belongs to.
5. Repeat the classification study using for classes the partition clusters obtained in (4). Compare the results with the ones obtained in 3. Discuss potential advantages and drawbacks of each strategy. Which would you recommend to analyze your dataset?

Include in your discussion all options that you have made, advantages and disadvantages of each alternative.
6. Imagine you are going to meet the researcher who contacted you. Report to him/her what you have learned about the problem. Discuss limitations of the analysis you have done and provide suggestions for future work.

- **About the groups:**

- Students should organize themselves in groups of 4 persons, and fill their groups in the following Google form: <https://forms.gle/idYNWwVpB3GTTNyN6>

- **About the presentation:**

- Duration of 15 minutes plus 10 minutes for discussion.

- **About the report:**

- The report should not exceed 10 pages, in the form of a scientific paper.

- Do not forget topics such as:

1. Description of the problem under study;
2. Objectives;
3. Estimation and validation methods;
4. Discussion of the results and interpretation of the findings;
5. Conclusions;
6. References.

- The R code, the presentation slides as well as the report must be send by email to:
`rosario.oliveira@tecnico.ulisboa.pt`