# Project report : Introduction to Python for Data Sciences
## Kaggle Challenge AAA

Student 1          Student 2

Date

# 1  Challenge description

In this section, it is expected that you describe your chosen Kaggle challenge. Some important elements that you should clearly present:

- Overview of the problem: which task do we want to perform, and how do we measure the performance of a model.

- The problem type (supervised/unsupervised, regression/classification/clustering).

- Available dataset: type (images/texts/tabular), size (how many training/validating instances), special fields/features.

# 2  Model choices

You describe and explain your choice of statistical models and their corresponding hyper-parameters.

# 3  Data preprocessing and engineering

In this section, you explain how do you choose features for your learning task. You can present them using the following guideline: how do you

- Fill missing values in the dataset?

- Remove features/fields that you feel irrelevant for the task?

- Create/engineer new features/fields?

- Handle special features/fields?

# 4   Results and assessment

You present your obtained models and their performances. You should discuss about:

- The current best results if a leaderboard is available.

- Your results and comparison to other existing models.

- Interesting behavior of your models (e.g., it predicts well if the instance has some special properties, which feature is the most important, . . .).

- Possible improvement that can further boost the performance.

- Optional: if you have a partner, explain how do you share the workload.

## Further notes

Please abide to the following rules:

1. Page limit: 3 pages.

2. Do not use Large Language Models to generate your report. Practice your writing skills.