Instructions for practical test

Data Scientist

- Problems should be solved using Python as programming language.
- Elaborate a report with your insights, supporting them with formatted data.
- Save the results on git hub and send us the link.

1. For this activity you will use the data set " Air Quality", which is available for download on: Air Quality - UCI Machine Learning Repository
   This dataset presents readings of 5 metal oxide sensors, extracted between years 2004 and 2005, from a road area in Italy. The data is collected hourly.
   a. present you exploratory data analysis.
   b. Estimate Relative Humidity behavior based on its answer to other parameters.

2. For this activity you will use the date set “Car Evaluation”, which is available for download on: Car Evaluation - UCI Machine Learning Repository
   The Car Evaluation Database contains basic car features such as price, technical and comfort information, as well as the market acceptability of the vehicle.
   a. Provide some insights on the data such as shape, distribution and cross-category comparisons (data exploration)
   b. Given Logistic Regression, Random Forest Classifier and Decision Tree, which model performs better when predicting car class? Justify your answer with data.
   c. Rank feature importance with respect to Random Forest Model and share your insights.
   d. Present a visualization of the Decision Tree and share your insights.