

Data Science Assignment

Welcome to your first step along your path inside iGenius!

The following assessment has been created to evaluate your ability to deal with a scenario you may encounter here as a Data Scientist.

Together with this document, there is a `.csv` file containing the data you need to answer the questions you will find in the next pages.

Information

Questions: 3

Points: 10

Good luck!

iGenius

1 Introduction

It has always been your dream, and now it has become true. You have opened your brand new winery in Milan, *WineGenius!*

To do a proper *grand opening*, you have collected data about wines from all over the world. It is time to exploit your analytical mind to understand which are the best wines you should order.

2 Details

The zip file contains a `.csv` with several information about wines, e.g., names, varieties, and price. Specifically, we have:

1. wine information:
 - title: name of the product;
 - variety: variety of grapes used to make that wine;
 - designation: specific geographical location in which the wine is produced;
 - country: country of production;
 - province: province of production;
 - price: price of the wine;
 - winery: winery that produces that wine;
2. tasting ¹:
 - taster name: name of the taster;
 - description: description provided by the taster to the wine;
 - points: points assigned from the the taster to the wine.

Given these data, your goal is to answer the following questions *at best*.

Question 1 (4 points) Some winemakers have asked you to buy their wines, which are new products not present in your list. How would you exploit the data you have to estimate the best price to *buy* them?

¹It may happen that more tasters evaluate the same wine, giving different descriptions and points.

Please, notice that the prices reported in the file are related to the cost at which the corresponding wineries sell them.

Question 2 (4 points) Since you will have customers from different countries, you want to create a wine list that presents wines from different parts of the world. How could you identify which areas are similar for production? Select the best wines from such areas to create the best wine list.

Please, take into account that you will have different types of customers, from wine lovers to wealthy people, able to invest an interesting amount of money in your bottles.

Question 3 (2 points) One of the most interesting project you want to carry out is a *blind tasting* night every Wednesday. Describe which approach would you follow to identify the characteristics of a wine, e.g., variety or denomination, given their description.

Please, in your description specify how you would structure the project.

3 Delivery

You are required to deliver a `.zip` containing the following documents:

- a *python* notebook containing the code you developed to answer the above questions;
- a `.pdf` file containing the explanation of the steps that you have undertaken to obtain your solution.

Please, name your `.zip` as

surname_name_assignment.zip

Notes

- For *Questions 1, 2*, you should provide code and explanation of your solution, while for Question 3 you should provide a description of your approach (*no implementation is needed*).
- If you need to create some functions and import them into your notebook, be sure to include the `.py` files so that we can reproduce your code.
- If specific (versions of) libraries you used are required for your project, please specify them so that we can reproduce your code.
- If you want to add images or visualizations to your notebook, feel free to do it.
- If you want to use additional data that you might find online, feel free to do it, but be sure that they are free. Add a brief description of them, specifying how you used them.