

FAQ final assignment

April 22, 2024

Final assignment, Frequently Asked Questions

Thank you for asking, this helps me, you and everyone!

1) Q: May I select a dataset with a “Usability Rating” on Kaggle lower than 8? I have found a dataset with a rating of 2.4.

A: Yes, the decision is yours. Selecting a suitable dataset is part of the assignment. But, it is not recommended because a low usability rating means it might be more difficult to gain meaningful insights and apply Machine Learning Algorithms on the data.

You are encouraged to look at other notebook submissions on Kaggle and see for yourself what others have already done with this dataset.

2) Q: Another student has selected a similar subject. Is this Ok?

A: Yes, you may even select the exact same subject, as long as you use different datasets.

3) Q: Besides the list of datasets that we should not use, are there any other datasets that we should not use?

A: You may not use any dataset that has been used as a tutorial example. The list in the guidelines may not be updated, or extensive.

4) Q: Every year some students select datasets without reading the guidelines carefully. Some even datasets that are explicitly mentioned in the “Don’t select” sections. What can I do as an educator?

A: Write a question, include it in the FAQ and write the reply too. This will highlight the need to read the instructions carefully once more, and at the same time is like making a small joke about it. Jokes draw attention and are slightly more memorable, even bad jokes like this one. By the way, the **Don’t Repeat Yourself (DRY)** principle applies in programming. In education, there should be a principle called: **Repeat Everything At Least Thrice (REALT)** and more if you have to.

Ps. FAQ will be updated, thanks to you!