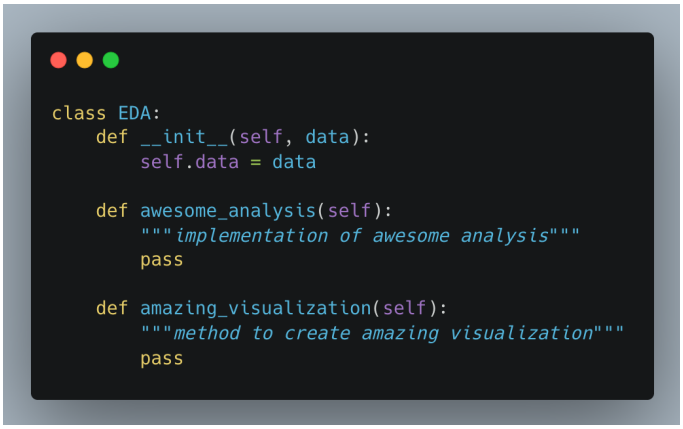
Data Analysis Exercise

Lending Club Dataset

https://www.kaggle.com/wordsforthewise/lending-club

You will be performing an Exploratory Data Analysis (EDA) of the popular Lending Club dataset and share your findings via a Jupyter Notebook posted on a GitHub repository or a similar platform (e.g. GitLab, Google Colab, nbviewer). The repository should contain a class called EDA which takes in the data as input and runs different analyses via its methods. Your Jupyter Notebook will leverage your EDA class (and other common open source data analysis packages) to find insights about the data. The implementation of the class can be either done within the notebook, or preferably in a separate module which gets imported into the notebook. For example, the skeleton class structure might look like:



```python
class EDA:
    def __init__(self, data):
        self.data = data

    def awesome_analysis(self):
        """implementation of awesome analysis"""
        pass

    def amazing_visualization(self):
        """method to create amazing visualization"""
        pass
```

After setting up your class, perform your EDA and report up to 10 key insights from the dataset through visualizations accompanied by short descriptions. You do not need to build a predictive model for the purposes of measuring risk on this dataset, just explore it. The exercise is intentionally left nonspecific to see how you will approach the problem. Feel free to browse other notebooks, blogs, Kaggle kernels, etc. for inspiration.

What we will be looking for:

- How you creatively utilize your knowledge of statistics to explore the data and what you choose to present as key findings.
- Your code quality (readability, organization and cleanliness of contents, proper use of statistics and visualization packages).
- How efficiently you communicate through visualization and convey information (less is more).
- A clear and concise notebook which preferably has a brief conclusion section in which you summarize your findings.

What bums us out:

- Long notebooks that we have to scroll through forever before we see a plot.
- That's it… We're generally very happy people!

Additional Notes:

- You are allowed to use any open source python package such as numpy, pandas, matplotlib, etc.
- If you are posting your notebook on GitHub, name the repository *lending-club-analysis*.
- When submitting your solution to Zest, only send the path to the repository; do not attach your notebook (.ipynb file) as it will not be reviewed.