PROJECT MILESTONE 2

PACE: Plan Stage

- How can I best prepare to understand and organize the provided information?

I will begin by exploring the structure of the provided dataset and consider reviewing the provided Data Dictionary

- What follow-along and self-review codebooks will help me perform this work?

I will use the follow-along notebooks from Milestone 2 to help me perform this work.

- What are some additional activities a resourceful learner would perform before starting to code?

I would recall the Python documentation to repeat some crucial aspects briefly.

PACE: Analyze Stage

- Will the available information be sufficient to achieve the goal based on your intuition and the analysis of the variables?

Before starting the analysis of the given dataset, I thought it would be enough. But, as I "dive" into analysis, I discovered some insights for additional information to the stakeholders.

- How would I build summary data frame statistics and assess the min and max range of the data?

By using the powerful pandas and NumPy libraries.

- Do the averages of any of the data variables look unusual?

No, the mean values didn't look unusual.

PACE: Construct Stage

Note: The Construct stage does not apply to this workflow. The PACE framework can be adapted to fit the specific requirements of any project.

PACE: Execute Stage

- Given my current knowledge of the data, what would I initially recommend to my client to investigate further prior to performing EDA analysis?

I will recommend to my client-financial institution to investigate further the loan history of its clients to reveal some insights for the future.

- What data initially presents as containing anomalies?

I didn't find any anomalies in the raw data.

- What additional types of data could strengthen this dataset?

Maybe the information on the spending habits and financial literacy (overuse of credit cards or payday loans, or lack of savings or emergency funds). Additionally, the industries prone to layoffs or low job security would be a valuable information in predicting loan default.