Project 4 Group 1 Proposal:

Roles:

Machine Learning: Shauntel and Ally
Tableau: Fabi and Uzor

Dataset:

- https://www.kaggle.com/datasets/samuelcortinhas/credit-card-classification-clean-data/data
- https://www.kaggle.com/datasets/itssuru/loan-data

Inspiration:

- We chose this dataset because the data is clean, and it has a high usability rating.
- We are curious about how different demographics will affect loan risk.
- We added a second dataset about loan information to further explore this subject.

Machine Learning:

- Our target is loan risk (high or low)

Visualizations:

- Income/career story:
  - Income (create bins) - bar chart
  - Occupation - bubble chart
  - Education correlation
  - Debt to income ratio
- Personal demographics story:
  - Gender correlation
  - Marital status correlation
  - Age (create bins) - bar chart
- Loan purpose visualization (donut chart?)
- FICO score correlation

Predictions:

- Age and income will be predictors of high risk. Marital status and gender will be the least effective in predicting risk.

GitHub repo:

- https://github.com/sphilli/P4.Group1-

Similar Kaggle Links:

- https://www.kaggle.com/code/upadorprofzs/credit-card-customer-analysis-ml-xgbtree#machine-learning-xgbtree

- https://www.kaggle.com/code/muhammadahmed68/credit-card-approval-predictions-85-accuracy

Similar Tableau Links:

- https://public.tableau.com/app/profile/caesarmario/viz/CreditCardApplicantDashboard/ApplicantDashboard
- https://public.tableau.com/app/profile/vikrant.shah/viz/DataAccuracyAnalysis-CreditCardApplications/DataAccuracy-DonutChart

Color Palette: