

Udacity Project - Visualization using Tableau Write Up

#Visualization

#Tableau

#Udacity/DAND/Term2

Link to Tableau Public workbooks:

First Version: [Tableau Public - First Version](#)

Final Version: [Tableau Public - Final Version](#)

Summary

This data visualization mainly focuses on exploring and explaining various aspect of American Loan from 2005 to 2014. Originally, there were 81 variables that describe many perspectives of a loan borrower. After simple data wrangling, I kept 32 variables for further analysis. Here are the most important ones that I found useful for analysis:

- Listing Creation Date: The date that listing was created. This variable provide the time variable for analysis.
- Loan Status: The current status of the loan: Cancelled, Completed, Charged off, Current, Past due. Defaulted
- Borrower rate: The Borrower's interest rate for this loan.
- Borrower state: The two letter abbreviation of the state of the address of the borrower at the time the Listing was created. Important variable for geographical analysis.
- Stated Monthly income
- Credit Rating: The letter represents the credit quality of the borrower: AA, A, B, C, D, E, HR. AA stands for best credit rating and HR is the worst.
- Credit Rating (Numeric): The numeric version of credit rating. 7 means AA and 1 represents HR. The number decreases while credit rating decreases.

The explanatory data analysis aims on providing insight about these factors to the audience. The interactive visualization story will enable the audience to uncover insights in the historical data of the Loan of America by themselves.

Design

- I used several maps to see the geographical difference of some variables including: credit rating, verifiable monthly income, loan original amount and debt to income ratio. This process intends to give the audience an overall idea of the visualization.
- By using the variable "ListingCreationDate", I created line plot for some variables including

the average debt to income ratio and the original loan amount. There was a financial crisis in the 10 year period, so by examining some variables, the audience will be able to see the abnormal data before the crisis and how things changed in the post-crisis stage.

- I used packed bubble plot to measure the most popular loan type overall and among different credit rating groups.
- I used color visual encoding to differentiate elements of continuous variables or categorical variables, mostly I used them as a additional dimension to existing plots to see the effect of other factors on the variable.
- I used heat map instead of line chart to avoid overwhelming information in a single plot.

Change log Feedback #1:

1. Changed all the stated issue of the previous version in the feedback.
2. Added a log-transformed scatter plot for the number of records and inquiries in last 6 months.
3. Added head map describing the change of average loan original amount for different credit rating groups.
4. Added a dual axis plot for credit rating, average monthly loan payment and the delinquent amount.
5. Added the yearly change of the credit score for each credit rating group, added color encoding for better interpretation, too.

Feedback

Overall:

- Revise the headings of story points and comments.

First story point (Credit rating map):

1. Change the starting limit of the horizontal axis to 2 to make the visualization become more differentiated.
2. Reduce the decimal spaces of the average recommendation.

6th story point (Original loan amount map):

1. Remove the heading of the highlight table and move the color encoding section up to make better sense.

7th story point (Is borrower a home owner):

1. Adjust the width of the box plot.

9th story point (The most popular loan type of different credit rating groups):

1. Improve the color contrast of the color encoding.

12th story point (Bankcard utilization broken down by credit rating):

1. Try to arrange horizontally instead of vertically.

13th story point (Upper and lower credit score of each credit rating group):

1. Add the yearly change of the credit score for each credit rating group.

Resources:

NA.