

Exploring the Loan Data from Prosper

Link of the final version of my Tableau Story:

<https://public.tableau.com/profile/sally.ng3979#!/vizhome/ExploringtheLoanDatafromProsper/ExploringtheLoanDatafromProsper?publish=yes>

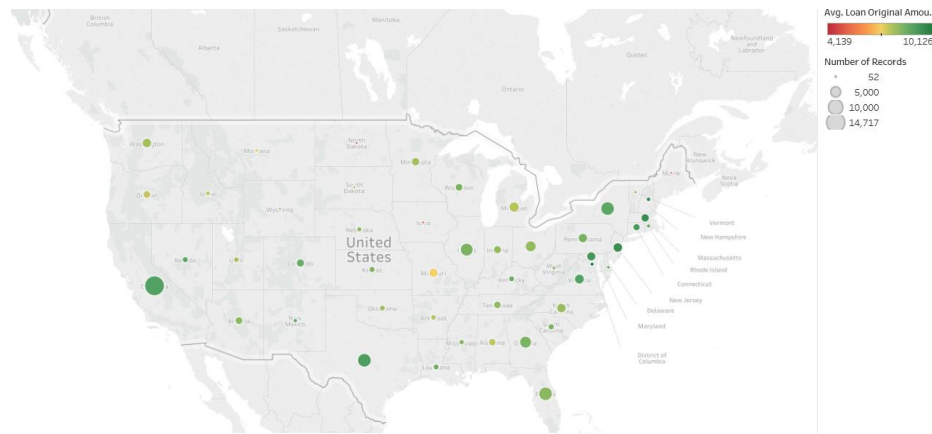
Link of the first version of my Tableau Story:

<https://public.tableau.com/profile/sally.ng3979#!/vizhome/ExploringtheLoanDatafromProsper-V1/LoanDataset?publish=yes>

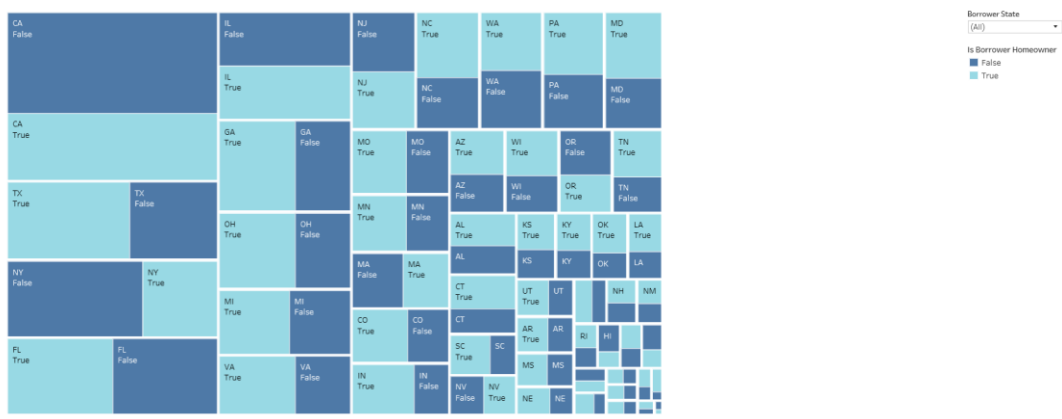
Summary:

The data set contains 81 variables and 113,937 loans information from Prosper, an online peer-to-peer lending marketplace. In this Tableau story, I visualize some of the variables to get a better understanding of the dataset and investigate the correlations among the variables.

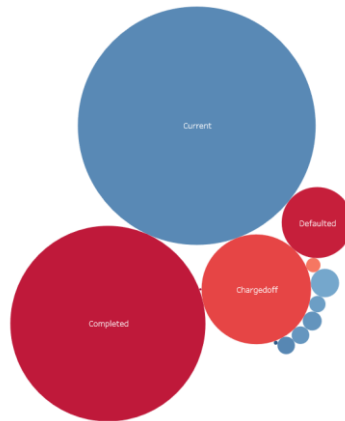
Design:



A symbol map has been used to explore the relationships between the location, number of loans and loan amount as it can present a more vivid picture of how the loan characteristics vary between states.



Treemaps has been used to investigate whether homeowners were more likely to receive a loan because the size of each box provides a very clear insight about the difference between the number of homeowners and non-homeowners who received loans within different states.

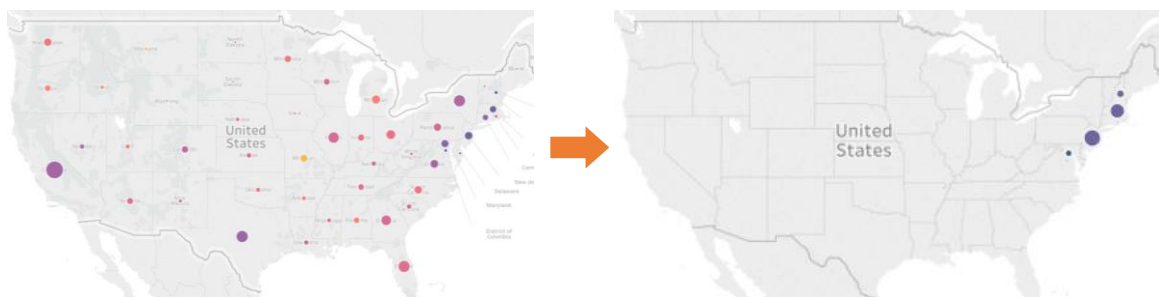


A scatterplot has been used to examine the correlation between the loan status and term. The size of the circles allows viewers to easily get the idea of the status of most of the loans. The great color contrast also provides wonderful insight on how loan term varied within loan status.

Design transformations after receiving the feedbacks:



- Because people with color-blind generally have difficulty distinguishing between reds, greens, browns, and oranges. I changed the color palettes from blue-green to blue-yellow, so people can differentiate the differences better.



- I also adjusted the format of the color, so slides will be more consistent in terms of color palette.

Feedback:

Feedback from Viewer 1:

1. What do you notice in the visualization?
After the financial crisis, the real estate confidence is recovering rapidly.
2. What questions do you have about the data?
The sample region didn't cover the Rust Belt area, which still suffering most.
3. What relationships do you notice?
Homeowner borrower is less affected by the business cycle.
4. What do you think is the main takeaway from this visualization?
We might use the homeowner & loan pattern to predict the housing business cycle.
5. Is there something you don't understand in the graphic?
No.

Feedback from Udacity Reviewer 1:

1. The title of the story doesn't really express your intention of the visualization. Loan Dataset is too general. The title should be more specific and related to what you have explored and what is your goal.
2. The diverging color pallets which has mostly green and red colors is not suitable for color-blinded people. They basically are not able to differentiate these two colors. I recommend colors like blue-orange instead.
3. Slides one and two are not consistent in terms of color palette. For example, in the first slide we have NJ as green but in the second one, it is orange. It's better to be consistent in the charts and since you are comparing one quantitative value so you really don't need to use both color and size to compare them.
4. You did great in terms of explaining your explored data. I can easily understand the charts and communicate with it.
5. I liked the captions (top panels) they are very informative and helped me to understand your story.

Resources:

Tableau Training and Tutorials (<https://www.tableau.com/learn/training>)

Udacity Data Analyst Nanodegree (<https://www.udacity.com/course/data-analyst-nanodegree-2>)