

Prosper Loan Data

Visualization URL

First iteration

https://public.tableau.com/profile/sandeep2389#!/vizhome/Prosper_Data/ProsperLoanStory?publish=yes
(https://public.tableau.com/profile/sandeep2389#!/vizhome/Prosper_Data/ProsperLoanStory?publish=yes)

Second/Final iteration after receiving feedback

https://public.tableau.com/profile/sandeep2389#!/vizhome/Prosper_Data_Revision_01/ProsperLoanStory?publish=yes
(https://public.tableau.com/profile/sandeep2389#!/vizhome/Prosper_Data_Revision_01/ProsperLoanStory?publish=yes)

Summary

This data set contains 113,937 loans with 81 variables on each loan, including loan amount, borrower rate (or interest rate), current loan status, borrower income, and many others. This has data from 2004 to 2014 . What I am trying to do here in this visualization is to explore the data with single variables intially like time series or geographocal distribution and later moving on to multi variable analysis like how Prosper rating / score was a good predictor of loan performance. I lalso could notice a strong correlation between APR and the performance of loans. The details would be viewed in the visualization itself.

Design

Format I have decided that the I will present the exploration as a Tableau story . This presents an easy transition between the worksheets and adds a contexts to it since I can add a text box explaining the main idea/finding pertains to the sheet .

Visualization Elements I have decided to use **Maps** to represent the geographical distribution of the loan data .

Histograms/barcharts to represent the distribution of data . As suggested in Udacity guide for comparing groups, **bar graphs** have been used . **Color** is often used to encode a third dimension since it is really difficult to interpret three dimensional data on a two dimensional plane. For E.g to show performing / non-performing loans over a quarter or anyother classification.

Also have decided to hide all the **worksheets** for the story to get undivided attention from the user.

Later after the initial feedback I have used uniform color palettes across the worksheets for similiar visualizations (E.g Bar charts) .

Feedback

Feedback 1

Hi Sandeep,

I have gone through the viz you have shared . The main idea of exploring the single variables and towards the end exploring relationship between different elements (APR Vs Performance) sounds solid . I have few feedback points which I guess it would make the viz better.

Convert the prosper score values to discrete. Right now prosper score values are continuous which makes the distribution looks bigger because of the lengthier X-axis. After going through the values I could see that the prosper score values are discrete , so it will be better to convert them to discrete on Analysis of Prosper score and Were Prosper score a good predictor for loan performance ?

Label state in the geographic distribution of loans. Geographic distribution of map looks very helpful to find out the loan distribution across United States. But it would be good if you can detail /label about the state names

Correct the sort order of prosper ratings. I got a clear idea about the correlation between prosper ratings and loan performance. Although the default sorting of the loan data seems little off. Because A rating comes before AA. Will you be able to change that? That will give a great representation of the data

Make sure the color palette is common for the similar visualizations. I can see that you have switched from a blue palette to red – yellow (?) palette while the story progress. Is there any reason for this? Can you make sure the color palette is common for the similar visualizations a different color palette for geographical data is fine.*

Reply 1

Thanks for that super quick response . All the suggestions were valid and I have incorporated them . I wish I could think of them in the first go . Please find the link below .

https://public.tableau.com/profile/sandeep2389#!/vizhome/Prosper_Data_Revision_01/ProsperLoanStory?publish=yes
(https://public.tableau.com/profile/sandeep2389#!/vizhome/Prosper_Data_Revision_01/ProsperLoanStory?publish=yes)

Also additionally just to check whether as a reader you got the same ideas which I have tried to convey, I am asking few questions about the viz . Please reply back to me with your understanding and I will compare .

What relationships do you notice?

What do you think is the main takeaway from this visualization?

You have been very helpful with all this . Thanks again .

Feedback 2

I could notice relationships between APR(borrower's rate) and Loan performance . I also could notice the uneven regional distribution of loans across US . The dip in total numbers during the US 2008 financial crisis also made lots of sense .

I guess the main take away from the visualization was the predictability of score and points against the Loan performance and the relationships between APR(borrower's rate) and Loan performance.

A PDF version of the mail communication would be shared among the submission documents

Resources

http://rstudio-pubs-static.s3.amazonaws.com/181962_e93d8068394b4b97ace91c74d5c4915a.html (http://rstudio-pubs-static.s3.amazonaws.com/181962_e93d8068394b4b97ace91c74d5c4915a.html)
<https://www.tableau.com/learn/training> (<https://www.tableau.com/learn/training>)

In []: