

Project: Create a Tableau Story

Prosper Loan Data

V1:

https://public.tableau.com/views/TheEffectofEmploymentStatusandOccupationonLoanQuality/TheEffectofEmploymentStatusandOccupationonLoanQuality?:embed=y&:display_count=yes&publish=yes

V2:

https://public.tableau.com/views/TheEffectofEmploymentStatusandOccupationonLoanQualityv2/TheEffectofEmploymentStatusandOccupationonLoanQuality?:embed=y&:display_count=yes&publish=yes

V3:

https://public.tableau.com/views/TheEffectofEmploymentStatusandOccupationonLoanQualityv3/TheEffectofEmploymentStatusandOccupationonLoanQuality?:embed=y&:display_count=yes&publish=yes

Summary:

In my visualization, I wanted to analyze the overall quality of loans that are taken out based on the employment status and occupation of the borrower. The story illustrates full-time/employed workers with stable, well-paying, high education jobs take out the largest and highest quality loans. It also shows that in general, the higher the original amount of the loan, the higher the quality of the loan.

Design:

Looking at this data set was daunting at first. There are so many different variables and I didn't fully understand them. However, once I zoned in on employment status and occupation as a focus the project became much more streamlined. I chose to use sorted bar charts to look at the dataset distribution of employment status and occupation as well as average loan score vs employment status and occupation. I also went with a line-graph with a trendline to display the correlation between employment status duration and prosper score. The prosper score metric gives an easy way to quickly analyze the overall quality of the loans. In the last dashboard I used sorted bar charts to show average loan amount vs occupation and employment status. Finally, I used a line graph to show the relationship between loan amount and average prosper score.

After the feedback session with my friend Jeff, I realized that I could add encode employment status as the color in my Employment Status Duration vs. Average Prosper Score graph to improve it. This also showed both positive and negative correlations based on the

employment status. Further, I realized that I could do a similar color visual encoding in the occupation histogram to give a rough idea to the viewer of employment status by occupation. In addition, adding this as a filter gave another level of interactivity beyond just scrolling through the different jobs. I think there is still possibility to add even more interactivity and break down the different data in new and interesting ways beyond what I have done here.

After the first project submission, I made several changes based on the reviewers comments. In the first comment, the reviewer mentions that their needed to be more commentary on the employment status duration vs prosper score in the captions. I added a few sentences highlighting the notable trends from this graph. The reviewer also mentioned that the graph is harder to interpret due to a lack of units on the horizontal axis. After consulting the variable definition spreadsheet I added the appropriate units to the axis label on the graph. In the second dashboard, the reviewer correctly points out that the 'professional' occupation listing probably contains people from many different professions. For this reason, it would be incorrect to compare this category of occupation vs. the other more specific listings. To correct this, I excluded the professional occupation category from the analysis. The reviewer also mentioned that the titles should be included with every chart. I had actually been doing this by default but recalled one of the videos from the tableau course material in which the person excluded titles on her charts saying that they were redundant and made the dashboards more cluttered. I went back and checked every dashboard to make sure that the titles are included. Finally, the reviewer recommended changing the chart type in the last dashboard. I tried mapping the two variables as a scatterplot but it was massively over-plotted. My mistake was that I had been using the average loan amount by prosper score instead of using average prosper score by loan amount. There was a significant amount of noise in the plot so I used a moving average to show a more clear trend. I also added a trendline to assist the viewer further in noticing the pattern that emerged from the data.

Feedback:

My friend Jeff had the following comments about the story:

- the entire story being uniform blue is boring
- it's cool that you can hover over different graphs but overall the level of interactivity is low compared to other stories that I have seen

Feedback from first project submission:

- Nice work of organizing your charts onto dashboards and tying them together not only by theme but also with filter actions. This really demonstrates the time and effort you put not only into creating your project but also into learning to use Tableau effectively. On the first dashboard, the prosper score by work duration line plot needs to include more explanation in the captions. What should the reader notice about this plot? What do the trends captured by this chart prove about the relationship between loan quality and work duration? Also, the chart is more difficult to interpret because the horizontal axis does not include units of measure. What do the numbers on that axis mean exactly, minutes, hours, weeks?

-The second dashboard also looks well implemented and theme based. Tell me what you think about this hunch: many computer programmers, teachers, accounts, etc. chose to list themselves as "Professionals" instead of giving a more exact occupation title. If this is true, then comparing the "Professional" level of the occupation category to the other levels presents a conflict. I think this needs to be considered and addressed in your analysis.

-Don't forget to include titles on all of your charts. It makes them so much easier to make quick sense of.

-Excellent job with the commentary for the last dashboard! This is exactly the type of summary that we're looking for. It shows depth and proves that you clearly understand what you've plotted in your charts. The line chart at the bottom of that slide needs to be reworked. If you notice there are in some cases more than one prosper score value for a given loan amount. This is a sign that you might not have the right chart type. Try some other chart types, like a histogram, as see if one doesn't work better.

Resources:

-Prosper loan data provided at:

https://docs.google.com/document/d/1w7KhqotVi5eoKE3l_AZHbsxdr-NmcWsLTliZrpxWx4w/pub?embedded=true

-My friend Jeff

-First submission review