Create a Tableau Story (Prosper Loan Data)

First version of story:
https://public.tableau.com/profile/kashfia.faruque#!/vizhome/prosperloandata_15532873403920/Story1

Summary and Design: Prosper is a peer to peer to lending company. It rates its borrowers as AA, A, B, C, D, E and HR. With AA being the best and HR being worst in terms of prosper score. We plotted a box plot for the amount of loan vs the prosper rating. We decided to use boxplot as for each rating there was a lot of variation in the amount of loan. From the plot we see that the loan amounts sanctioned for borrowers seem to depend on what rating they belong to.

So the next question is what variables contribute to the rating. Plotting the average of total inquiries, open trades, current delinquencies, bank card utilization, debt to income ratio against the categorical variable prosper rating, we see that as the average value of these variables decreases, we move from HR to AA. Whereas for Available Bank card credit, stated monthly income and credit score decrease we move from AA to HR. Here we used a bar plot since the length of the bar plot indicates the average value for each rating group. We used the prosper rating as the filter as there were null values that needed to be removed.

We next wanted to see which states had the highest percent of investors/borrowers? We see that California had both highest percent of investors and borrowers. Also, by plotting the amount invested, we see that California topped here as well. CA received the highest percent of invested amount from prosper. We used a map and then colored it based on the number of investors/borrowers/invested amount per state. Lastly, we converted the sum value to percent so that we can find out which state has what amount of percentage of these variables. We also have legends that shows the color the state will take on depending on the value.

Next to determine if investing money in prosper is beneficial for the lenders or not we plot the average annualized return per day and the average estimated return per day together in a dual axis plot. Here we only used to data where the loan status is completed. We see that the average annualized daily return is less than the estimated return. What is interesting is that the annualized return was close to the estimated return in earlier years and the difference seems to have increased over time. Also, we see the return being estimated higher than it actually is. For this visualization we used a scatter plot. We decided to use a scatter plot as we were plotting data for each day. Also, two different colors were used to identify each of the different averages in the legend. Also, we have a filter to select for loan status equal to completed.

When we plotted the annualized return and estimated return for each state, we see that Utah had the highest annualized return, although it had the lowest estimated return. Here since we were interested in observing the values of each state so we drew the map and then colored for those variables. The meaning of the color gradient is shown in the legend.

Here we had filtered using listing creation date between 1st August 2009 to 31st December 2013. This is because prosper launched their new system from 1st August 2009.

Feedback:

1st feedback:

I have two points:

- Please avoid acronyms in your charts.
- In some charts we have null as one filter. Please remove the null values from the filter or rename them.

2nd feedback:

page 1 : half of the page = blank

page 2 : lots of plots. which one do we look at ?

page 3 : 3 maps which look the same. Is it the message ?

page 4 : what is your answer about divergeing lines ?

page 5 : 2 maps, but what's the conclusion ?

Sorry, I don't work on this dataset and don't understand loan field well

Final version of story:

https://public.tableau.com/profile/kashfia.faruque#!/vizhome/prosperloandataversion2/Story1

Improvement on the visualization based on the feedback received:

I got really good feedbacks on improving my story. I implemented all of the suggestions. I took out the last slide from my previous story as I felt the map didn't really add any value to the analysis.

I removed acronyms from my chart. I removed the null options from my filters. I added the box plot in a dashboard first and then added the dashboard to the story. This made the figure position at the center of the page. This was not possible by worksheet previously. And I finally added descriptive captions for each of my charts to better explain the user what they are observing.

References:

https://community.tableau.com/thread/241670

https://community.tableau.com/thread/200647

https://community.tableau.com/thread/250240

https://onlinehelp.tableau.com/current/pro/desktop/en-us/buildexamples_boxplot.htm

https://www.credera.com/blog/business-intelligence/tableau-workaround-part-2-create-an-include-exclude-filter-on-the-same-field/