DATA STORY USING TABLEAU

For this project, exploratory data visualizations were created using Tableau that tells a story highlighting some key trends in the data.

The project workflow is presented in the sections below.

Data Set

Data exploration and visualizations was conducted on Prosper Loan data. Prosper loan is a peer to peer loan lending company that provide loan at very low rates. These loans are typically funded by multiple people within United States. The investors who fund these loans use prosper ratings, a proprietary system developed by prosper that assigns ratings to every loan application to determine the applicant's level of risk. The ratings represent an estimated range of annualized loss rate range to the investor. Higher the ratings, lower the risk.

Summary

Link below is to the draft data story

https://public.tableau.com/profile/varsha.raj#!/vizhome/ProperLoanAnalysis/DataStory

Link below is to the final data story

https://public.tableau.com/profile/varsha.raj#!/vizhome/ProperLoanAnalysis/DataStoryFinal?publish=yes

Data exploration was conducted on prosper loan prior to developing a story in tableau. The data exploration suggested that there was a strong link between economy growth/employment and number of investors. The data story in tableau works around this idea to generate visualizations that correlate trend in number of borrowers particularly borrowers with 'employed status' with number of investors. The data story also explores the connect between prosper ratings and lender yield and the extent of their influence on the investors.

Design

The idea behind developing the data story was to incorporate different types of graphics to explore relationship between borrower count and investor count. To correlate the numbers of investors with prosper ratings and lender yield for the available year of records, a packed bubble chart seemed ideal. This is a three-dimensional graph (consist of three variables) and the best way to visualize this was to reduce as much annotations as possible.

The main theme of this data story is to assess how trend in number of borrower's influence number of investors. To first provide a general idea, the best practice was developing multiple graphs that each play a role in developing a complete story. The graphs look at borrowers count particularly borrowers with status as 'employed' and how the trends are when compared to the number of investors.

Initially separate bar charts were made to show borrower count and employed status borrower count and the idea was to place them side by side. Post feedback, it made sense to keep the bar chart for only for borrower count and create a line graph indicating the percentage of employed status borrowers with the borrower population. This modified graph avoids redundancy.

Median monthly income trend was also explored to see if that played any part in influencing the trends in the number of investors. This graph was included to add an extra dimension to the story. While this graph did not play a role in the data story conclusions, it does open some interesting discussions.

The observed trends between number of borrowers and number of investors were then assessed using a state by state analysis.

While the design choices were made to look at the overall trends combining all years, as per the feedback (shown below), a filter with the year of records was provided that gives the flexibility to assess the trends by either selecting individual years or by looking at all years combined.

Assessing the trends year by year does provide some interesting trends which was otherwise masked.

Feedback

1. In the Lender Yield graph to assess number of investors in each ratings. Would it make sense to assess this by year, may be by drilling down while you select each prosper rating?
2. In the Prosper Loan growth correlations with economy growth. I don't think you need two graphs one showing Total Number of Borrowers and the other the number of Employed Borrowers. Maybe you can plot the employed borrowers on the same graph as total borrowers indicating as a line graph the employed borrowers as a percentage may be using a secondary axis.
3. In the data story with map showing number of borrowers as a direct link to number of investors in states, wouldn't it make sense to break down by year, may be add a filter to the graph so specific period can be selected?

Resources

Prior data exploration conducted on prosper loan was used to develop the story. The visualizations however were chosen from the options available within tableau that best suited to present an effective data story.

Data Files

Raw prosper data csv file was directly uploaded into tableau and filtering was dome based on the visualizations.