Data Analysis Of Prosper Online Peer to Peer Loan Business

Visualisation:

Summary:

In my Tableau story I have done exploration on the relationship between loans and the people acquiring them, the reasons for taking out a loan and who defaults the most.
I have given a brief insight into the people acquiring loans by looking at the debt levels by state as well as the amount of money borrowed over the past few years looking at whether or not people were homeowners.
I explored the reasons that people have for taking out loans as well as who is most likely to default based on their occupation and income range.
Breaking down the number of defaulters by occupation it seems that college student have the highest borrower and default rates.

Design

I began by looking at what data was available and deciding what variables might yield an insight into the trends of loans over the past few years. I decided to focus on what made a person more likely to take out a loan by looking into where they live, whether or not they owned a house and their occupation an income range. I was also interested in seeing who was likely to default on their loans.
I decided on the following graph types:

- For showing the Income:Debt ratio by state it made sense to use a map to best convey this information.
- Scatter plots were used for showing the loan amount (indicated by size) broken down by:
  - loan status
  - category
  - occupation.
    Each of which was categorised by colour.
- Bar plots for depicting whetehr or not a person was a homeowner, their income range and Prosper score.

Feedback:

I shared my first Tableau design with friends and received the following feedback:

- Showing the amount of money by state as a line graph was too crowded and difficult to take in information.
- The plot with occupation has quite a lot of information and it was difficult to see this plot because it was quite small since it was on a worksheet with two other plots.
- For the loan categories graph shown by quarter some people thought it would be easier to read if it was done by year so that there would be less poibts on the scatterplot.

Resources:

- Tableau tutorials
- Information on Prosper
- Udacity course on Tableau

Data Files

This Data set contains 113,937 loans with 81 variables on each loan.
The original data can be found here: