Pierre Hentges, Udacity DAND
Term 2 Project 4 report

Create a Tableau Story :
Prosper P2P lending

Summary

This visualisation investigates the P2P lending company Prosper and how borrowers use this novel finance facility. The principal interest is the large proportion of borrowers who use Prosper for debt consolidation – usually successfully in the sense that they subsequently manage to pay off the loan. This use of Prosper loans for debt consolidation and other ways of overcoming economic hurdles became more viable after their introduction of stricter lending criteria in July2009. The way borrowers use P2P lending has to be understood as a response to financial difficulties many faced after the 2008 global depression and particularly the credit crunch which made unsecured loans difficult to obtain for many individuals.

First version:
https://public.tableau.com/profile/pierre.hentges#!/vizhome/ProsperStory_15544624601900/ProsperStoryver1

Final version after feedback:
https://public.tableau.com/profile/pierre.hentges#!/vizhome/ProsperStory_15544624601900/ProsperStoryver2

Design

The key insight I want to get across is that Debt consolidation accounts for the largest amount of loands, and this is often a successful strategy for borrowers, in the sense that defaults are rare as even high-risk borrowers usually manage to keep up their payments.
This insight has to be understood in the context of the global depression starting in 2007/8 and the ensuing credit crunch that made loans inaccesible for many individuals. These economic conditions also shaped P2P lending, with increased government regulation from 2009.
The second insight is that Prosper changed fundamntally in 2009.

[1] First dashboard: how does P2P lending at Prosper work?
Introduction aiming to convey basic concepts of how P2P lending works, enabling the audience to understand the story being told.
Top chart : to get across what a typical loan looks like, I decided to use a histogram of the different loan amounts. This shows what the most common loans are as well as their range.

The second concept I wanted to introduce is the loan term, as this is what makes loan repayments predictable for borrowers, so I decided to use stacked bars in the histogram. However, I didn't want to draw attention to term too much, so chose shades of greys to group them.

Bottom chart: introducing the key concept that there are different groups of credit risk, which are linked to the loan interest, whose range of distributions depend on it. Boxplots are an excellent way to compare different categories in terms of the distribution found in each.

<ins>*[2] Timeline of how economic and regulatory pressures shaped Prosper*</ins>
The way Prosper functions changed fundamentally in 2009: government regulation caused lending to collapse, until Prosper relaunched in July 2009 with an improved lending model. Other factors shaping Prosper were a lawsuite and the impact of the 2008 global depression/credit crunch. The appropriate way to visualise these developments is a timeline. I made two charts – one for the total amount lent ($) and a second for the number of loans. To show how credit deliquency levels fell after the 2009 relaunch, loan status is displayed as stacked areas. Having two charts allows for 4 explanatory text boxes without excessive clutter.

<ins>*[3] Prosper borrowers are located across US states*</ins>
The purpose of this dashboard is twofold: 1 - give a 'flavour' of where borrowers are and their demographics while consolidating some of the concepts used in the introduction ; 2 - introduce the idea that the audience can use filters by clicking on a chart to zoom in on demographics, while also gently introducing a couple of new metrics (borrower income range, employment status).

I chose a map as the visual because they look good and are satisfying to use as a filter. The 4 demographics charts below are easy to understand, two are vertical, two horizontal bars. Colours used are minimal to avoid distraction.

<ins>*[4] Borrower demographics – social class*</ins>
The purpose of the dashboard is to encourage the audience to further investigate who the borrowers are, particularly in terms of social class. This has implications for what their lending needs are – which will be directly addressed in the next dashboard. A bubble chart of occupations is used as the filter to find out about income range, employment, home ownership etc.

<ins>*[5] How borrowers use Prosper loans*</ins>
To understand how borrowers use loans, I wanted to look at Prosper's 'ListingCategory' variable. This as 20 different categories – too many to visualise easily, so I grouped some into larger categories that express how people use them. The area in the tree map represents the fraction of loans, which is displayed as a percentage, along with the median loan amount. Further metrics are available via mouseover – these enable the audience to compare the different ways borrowers use loans.
A timeline using the broad categories as colour is used to look at how the purpose of loans has evolved over the years. Growth of debt consolidation is particularly pronounced - to draw attention to this, several relevant points are labeled.

<ins>*[6] Successful repayment rate of loans used for debt consolidation*</ins>
60% of borrowers use Prosper loans for debt consolidation – how often do they manage to pay off their debts?
To visualise this, I decided to use stacked bars, with loan status as area, by Prosper risk grade as category. Values were adjusted to make up 100% - this allows the different groups of Prosper risk grade to be compared, even though the number of loans in each varies. The colours for the 3 statuses of delinquency are in order of gravity – red for default, orange for chargedoff, yellow for 'past due'.
Essentially the same chart was used for pre-2009 debt consolidation loans – though all loans in this group have matured, so they dont contain loans that are 'current' and 'past due'.
The difficulty with this chart is that the viewer needs to understand what the loan status means – but this is specialised terminology which many may not be familiar with. The terms are also specific for US legislation – for example 'chargedoff' does not exist internationally but is legally defined for the

US finance industry as failing to make loan repayments for over 6 months. This problem can only be solved by explaining the terms in the text caption, hopefully succinctly.

[7]The cost of debt consolidation loans
Looking further into the benefits of using Prosper loans for debt consolidation, I wanted to check what typical monthly repayments are, and how they differ across the Prosper risk groups. Monthly repayment amount is obviously dependent on the loan amount, so the whole distribution needs to be considered. The most appropriate visual here is a box plot. Given that the term also determines the monthly repayment, the term, the Prosper risk group are variables used for the x-axis.

Finally I wanted to study the total cost of a typical loan for debt consolidation amounts to, i.e. the total sum of repayments compared to the original loan amount. Again this makes most sense by considering the different groups of Prosper risk grade, since these affect interest rates. The median values were used for the visualisation, using data from borrowers who had completely paid off their loan.

Feedback

I received feedback from 2 people. The first person made detailed comments on most dashboards, which are quoted and addressed below.

The second person shared most of the specific criticism, and also gave the following overall feedback:

"Takeaway message: As somebody who knows/understands very little about economy, this visualization has helped me understand that over the past 5 years there has been an enormous increase of people who needed a loan from prosper, and this was mostly due to a need to pay off depts. Even people who were classed as being 'high risk' were able to access a loan and appear to be able to have been able to pay this off/are paying it off.
I struggled to make sense of the 'who borrows through propser' graph and to draw useful conclusions from it. It will be more useful if this information could be presented differently.
The visualization tells an intersting and clear story leaving no questions unanswered."

[1] Comments from the first person:

The x- scale and bins don't match – 14K/15K. There is no legend explaining the credit term, and the term is variously expressed as months or as years - either add a legend with consistent units, or get rid of stacked bars.
Boxplot: Need to explain what the categories (Prosper credit grade) are to a non-specialist audience. The terms related to this are used very inconsistently throughout the viz, which is pretty confusing.

changes made in v2
- added legend for term, to make the unit of the term consistent throughout the viz, I added 1/3/5 as alias for 12/36/60 months.
- fixing histrogram bins – Tableau chose a 1100 bin size, which also displaced axis labels, ticks. Changed bin size to 1K, imposed fixed labels (5K) and ticks (1K)
- introduced consistent naming: Prosper risk score
- add explanatory captions of what P2P lending is
- shotened text of story point to fit available space

[2] Feedback person 1:
The legend of the bottom chart is at the top – it down.
Explanatory text box needs to point to year. Emphasise what happened in 2009.
consider info boxes for additional info on lending jargon such as chargedoff.
The two charts are a bit redundant - combine into one?
Dubious about having boxes inside chart.

changes made in v2
- Removed chart with total loan amount lent, leaving only chart of timline with quarterly number of loans
- added bullet list for timeline
- only left 2 text boxes in chart area
- added explanation of loan status terminology

[3] Feedback person 1:
It's misleading to uniformly colour a whole state as it is interpreted as related to area – better to use e.g. a circle for each. Using the loan terms in the histogram is pointless.

changes made in v2
- changed map
- removed loan term from histogram

[4] Feedback person 1:
size of bubble charts are problematic - bar chart might be better
is it justified as a filter handle??

Feedback person 2:
The bubble chart is difficult to use due to the number of bubbles and small size of many of them

changes made in v2
- using income range as filter – this has a much smaller number of categories which are well defined

[5] Feedback person 1:
too much info on mouseover in the treemap - top category needs to be different

Feedback person 2:
Caption and categories are hard to follow

changes v2:
- improved the tooltip data
- moved legend close to treemap
- improved caption
- improved labels in timeline

[6]
Feedback person 1:
the use of percentage markers may be legitimate but I always find it slightly strange for stacked bars... the size of the bar is a proportion and can't be derived from the percentage axis... without addition and/or subtraction. you can only really compare the first group/category. they have advantages if the categories have some order to them, which (moving from v.bad to ok) this does.

the debt consolidation claim doeesn't leap off the page. I'm seeing some bars, the larger one 'current' (what do these bars mean?), which change according to demographic. What has that to do with debt consolidation?

changes v2:
- tried to explain loan status better
- changed the ordering of the loan status of the stacked bars: 'default' at the top, then 'chargedoff' then 'past due' – this way, delinquency can easily be guessed but subtracting these 3 categories from 100%

[7]
Feedback person 1:
There is quite a lot of uniformity in repayments until you get to D, E, HR... slightly surprising the wealthier bands don't pay more. Those are big whiskers on the top plot. Second chart does what it says on the tin.

changes in v2:
- removed the series of 12 month term repayments – these are a small minority but significantly distract from the 3 and 5 year ones
- improved caption

Resources

- https://en.wikipedia.org/wiki/Prosper_Marketplace
- https://en.wikipedia.org/wiki/Peer-to-peer_lending
- https://prosper.zendesk.com/hc/en-us/articles/208500586-What-are-Prosper-Ratings-
- https://www.lendingmemo.com/risk-grades-lending-club-prosper/
- consulted https://onlinehelp.tableau.com/ and some threads on https://community.tableau.com/community/forums