**A Story About Student Loan Data**

**Author: Mason Krug Udacity Nanodegree Project**

**Story Version 1, Story Version 2, GitHub**

**Summary**

The Prosper loan data set contains 113,937 loans with 81 variables on each loan, including loan amount, borrower rate (or interest rate), current loan status, borrower income, and many others. The dataset and data-dictionary can be found on below links.

* [Prosper Loan Dataset](https://s3.amazonaws.com/udacity-hosted-downloads/ud651/prosperLoanData.csv)
* [Loan Data Variable Definitions](https://docs.google.com/spreadsheets/d/1gDyi_L4UvIrLTEC6Wri5nbaMmkGmLQBk-Yx3z0XDEtI/edit#gid=0)

**Decisions About**

I’m very interested in how data can vary by regions. People have a specific amount of variance based on genetics. However, differences greatly manifest themselves when geographic regions are taken into account. Culture, resources distribution, socioeconomic factors, any many other variables can be very interesting to look at when latitude and longitude are explored. Coastal states, such as California, have wildly different economies than Nebraska. This is why I chose maps as the primary way to explore relationships. Additionally, I chose scatter plots to explore relationships. I chose them because they can be used to explore many variables at once.

**References**

2014 Population Statistics - census.gov/quickfacts/fact/table/US/PST040219#PST040219