TidyTuesday #48: Student Loan Debt

@SchackartK

The Data

Data comes from the US Department of Education. Information is provided for several financing agencies of student loans in the United States. The information used is the `Total Payments` to each lending `Agency` (when available) for the fiscal years 2016-2018. This total is broken down into 4 categories:

- `Consolidation`
- `Rehabilitation`
- `Voluntary Payments`
- `Wage Garnishments`

Visualization

Agency Category Breakdown

Here is how the `Total Repayment` for each `Agency` breaks down by `Repayment Type`, averaged over all years (and quarters) that data were present for that `Agency`.
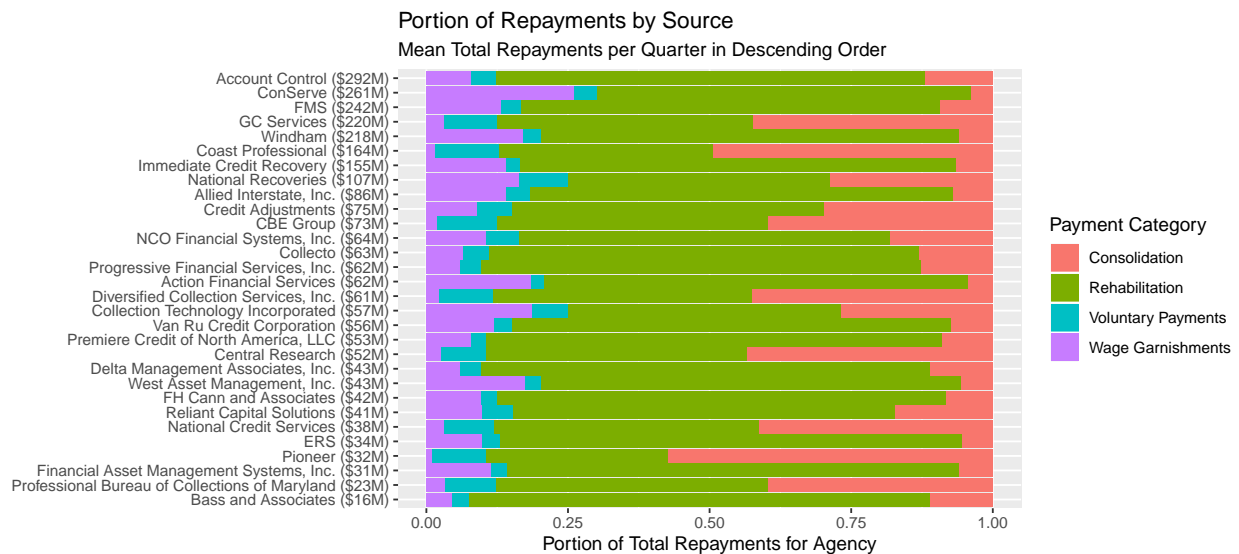
While this is interesting if you are curious about each individual `Agency`, it does not give much insight into any over all trends. I was hoping that you may see a trend in repayment composition as a function of `Total Repayments`, which is why I sorted it that way; there doesn't seem to be such a trend.

Repayment Category Proportion by Total Repayment

Next, let's look at how the proportion of each `Repayment Type` varies with `Total Repayment`. Each point represents an `Agency`.

Portion of Total Repayments by Category for Each Agency
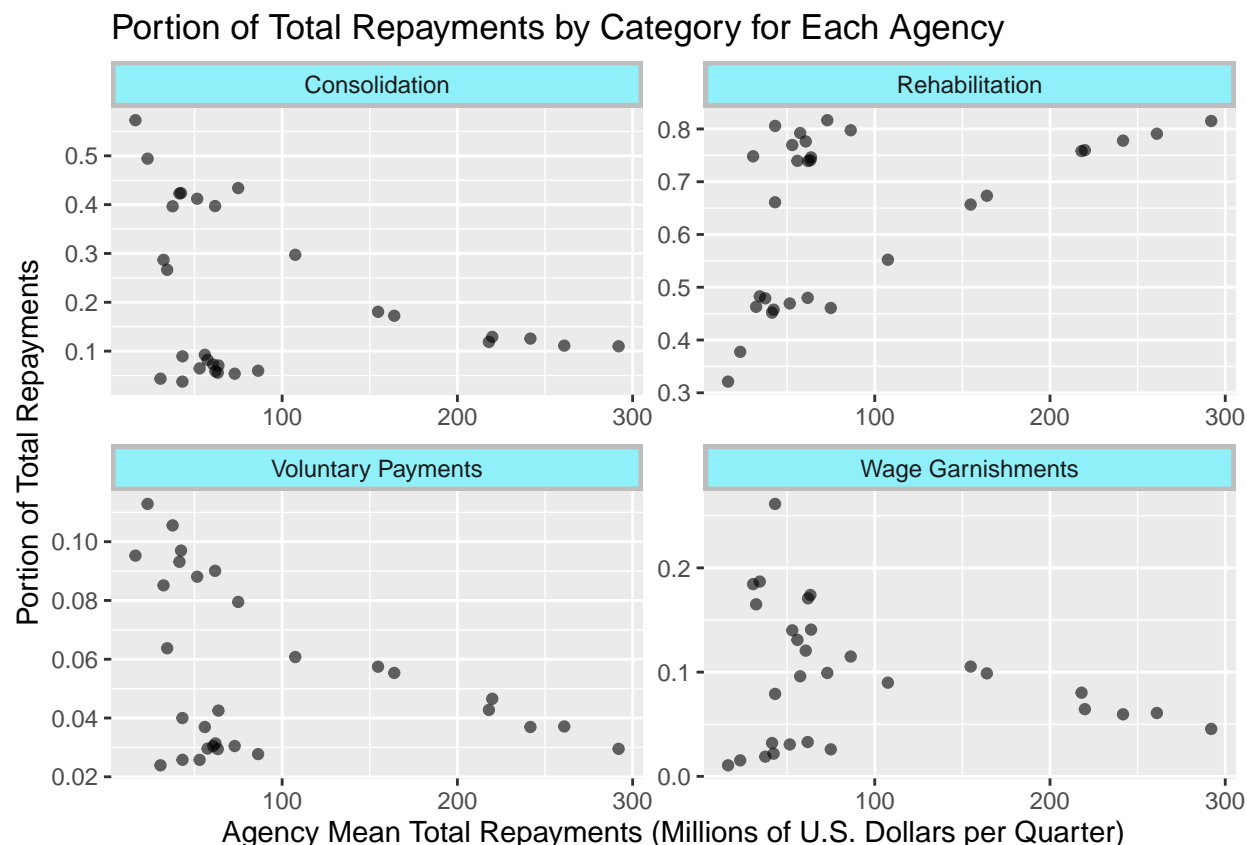
Observations

From this plot, we can see that there are some general trends between the `Repayment Type` proportions and `Total Repayment` amounts:

- For `Consolidation`, `Voluntary Payments`, and `Wage Garnishment`, proportion decreases with increasing `Total Repayment`
- Proportion `Rehabilitation` increases with increasing `Total Repayment`

Outliers

The more interesting thing to note though, is that in all of the graphs there is a group of points that defy this trend, where proportion of `Consolidation`, `Voluntary`, and `Wage Garnishment` is low for low `Total Repayment`, and proportion `Rehabilitation` is high for low `Total Repayment`.

Upon examination however, the outlier groups do not consist of the same `Agency`s across plots. In fact, no single `Agency` is present in all four outlier groups. However, the outliers in all plots except `Wage Garnishment` consist of these 12 `Agencies`:

Allied Interstate, Inc.	CBE Group	Collection Technology Incorporated
Delta Management Associates, Inc.	Diversified Collection Services, Inc.	Financial Asset Management Systems, In
Premiere Credit of North America, LLC	Progressive Financial Services, Inc.	Van Ru Credit Corporation

Covariation Between Repayment Categories

A final analysis that was investigated was how do the proportion of the different `Payment Types` vary together.
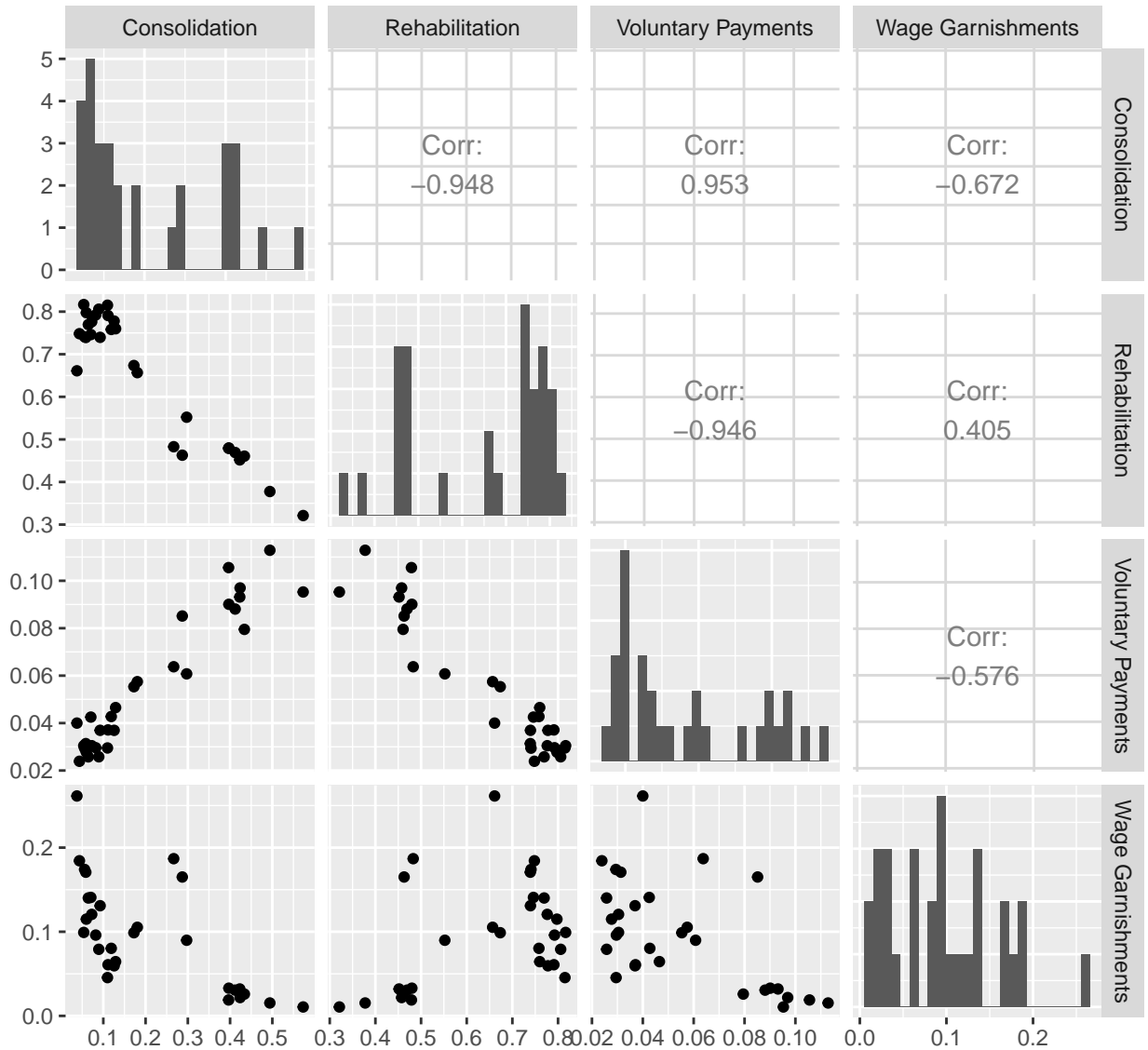
From this analysis two main observations were made:

- `Wage Garnishment` does not strongly correlate with any of the other payment types
- `Consolidation`, `Rehabilitation`, and `Voluntary Payments` do show correlations with each other, with the following relationships:

- Consolidation and `Voluntary Payments` are negatively correlated with `Rehabilitation` (r2 = -0.948 & r2 = -0.946)
  - `Consolidation` is positively correlated with `Voluntary Payments` (r2 = 0.953)

Conclusions

- Overall repayment category composition for each `Agency` does not depend on `Total Repayment` amount.
- `Rehabilitation` is positively correlated with `Total Repayment`, while all other categories are negatively correlated.
- All repayment types have strong correlations with one another, with the exception of `Wage Garnishments`.

Acknowledgements

These data were fetched from the tidytuesday Github repo. Thank you to the R for Data Science Community for the support and ideas to improve my data analysis and visualization skills.

Authorship

Kenneth Schackart

- Twitter: @SchackartK
- Github: schackartk