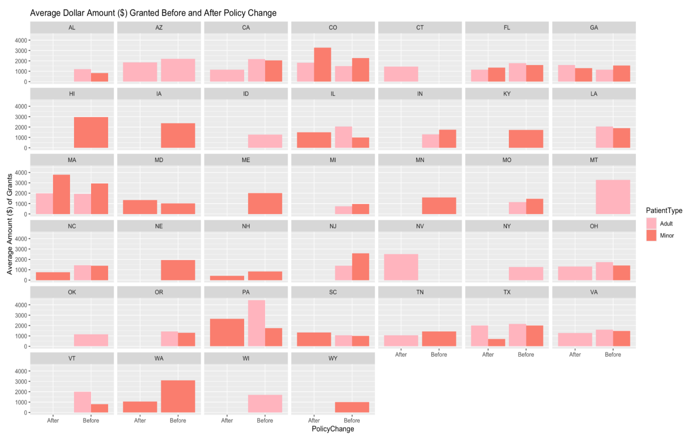
APPENDIX

PATIENTS DATASET:



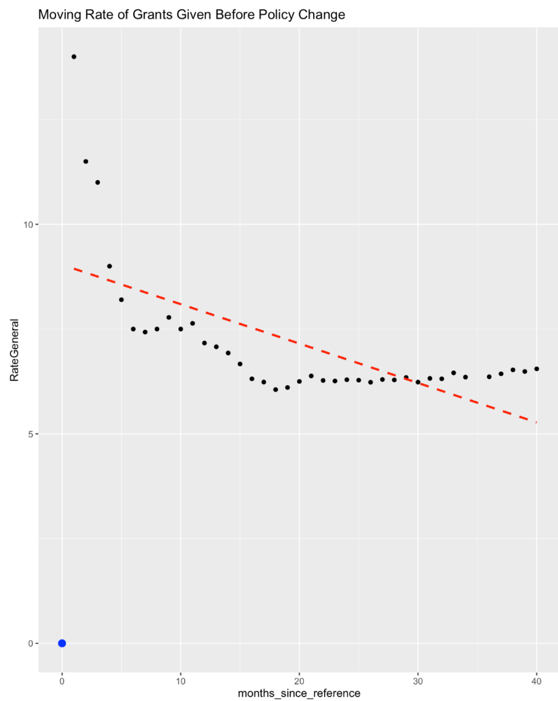
Preliminary explorations were made to analyze the effects of the policy change on minors and adults. We found that these visualizations were hard to interpret and, instead, created tables of our findings. Specifically, the data shows that there was a decrease in the proportion of grants given out to minors. As for how the policy change impacted the dollar amount of how much was granted, there was a bigger drop for adults compared to minors. However, after conducting hypothesis testing, we concluded that the policy change does NOT affect the amount of money granted. This is significant from a business standpoint because one of the motivating factors for conducting this policy change was to lower the average amount granted to patients to increase their reach. Our team will note that there are limitations to this analysis due to the limited data after the policy change (6 months worth) and further investigation is required to confirm this finding.

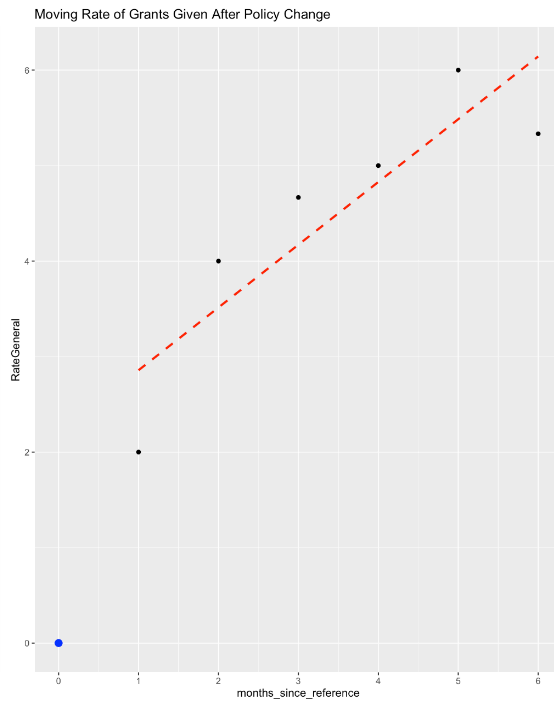
	V1
Prop_Minor_Before	0.5877863
Prop_Minor_After	0.3939394

	Group	before	after
1	Adult	1699.564	1531.291
2	Minor	1711.795	1700.052

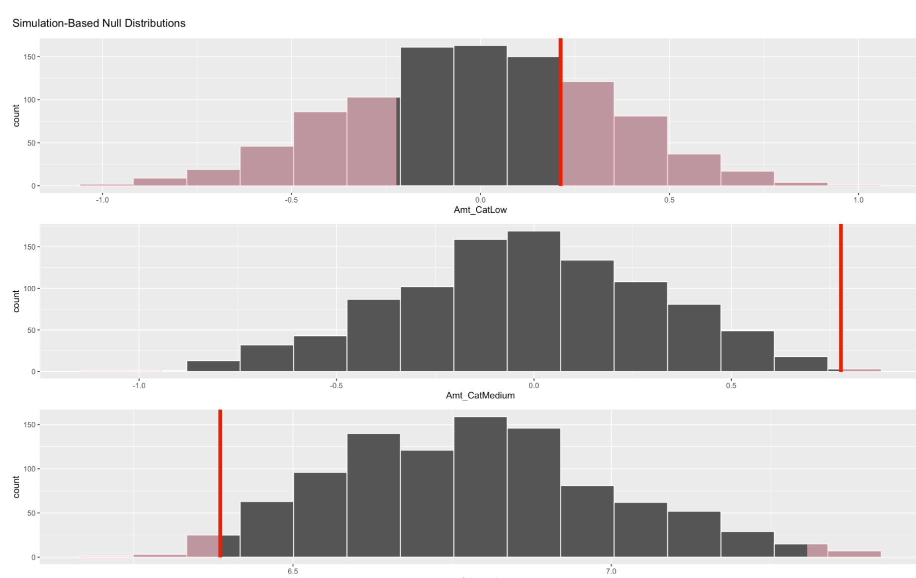
	calculation	Before	After
1	Rate of New	3.6923077	1.1666667
2	General Rate	6.7179487	5.3333333
3	Prop of New	0.5496183	0.2121212

Other findings that we discovered during our preliminary data exploration is that the rate of new applicants, the proportion of new application, and general rate of grants given out (for both returning and new applicants) decreased after the policy change.



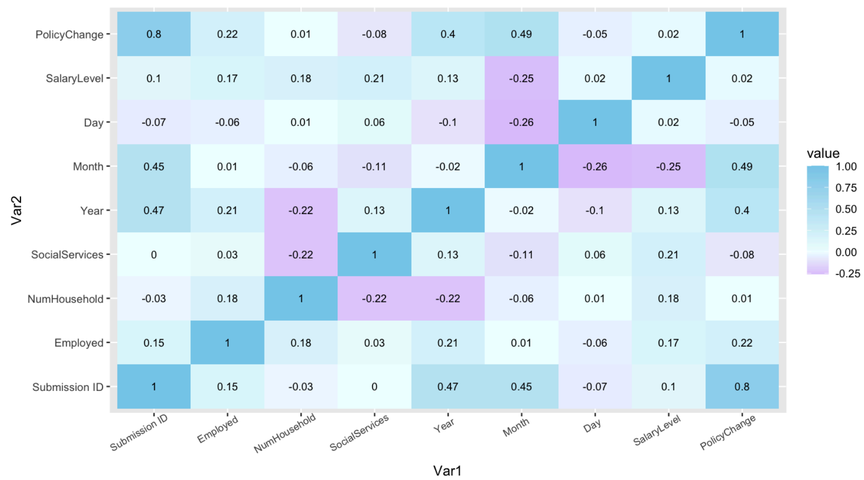


When examining the rate change, there were some difficulties in conducting hypothesis testing to determine the significant of this finding due to the required mutations. To compensate, our team found the confidence interval of the rate before and after the policy change to see if there were overlapping values. With 95% confidence, the rate change after the policy change ranged between 3.44 and 5.39 whereas before the policy change, it ranged from 6.68 and 7.73. Thus, we concluded that this difference was not negligible



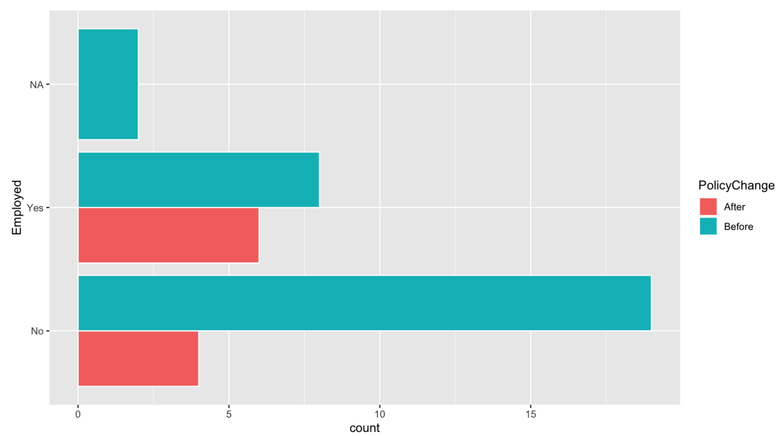
This is a visualization of a hypothesis test that looked at whether the amount category had any effect on the rate of grants given before the policy change. In the end, only the medium amount category had a significant p-value. When the same test was done for data after the policy change, there were no significant p-values.

WPP DATASET:



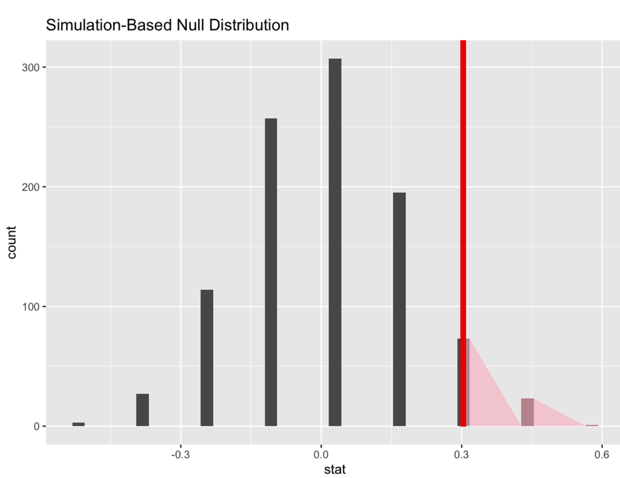
This Heat Map was made to see how employment was affected by other variables. Mostly affected by the number of people in the household, salary level, policy change, and the year. The year and policy change are closely linked.

The visual to the right shows people who completed the program. Unfortunately, they have a couple of NA values (2). This is what we had to filter out for the presentation.



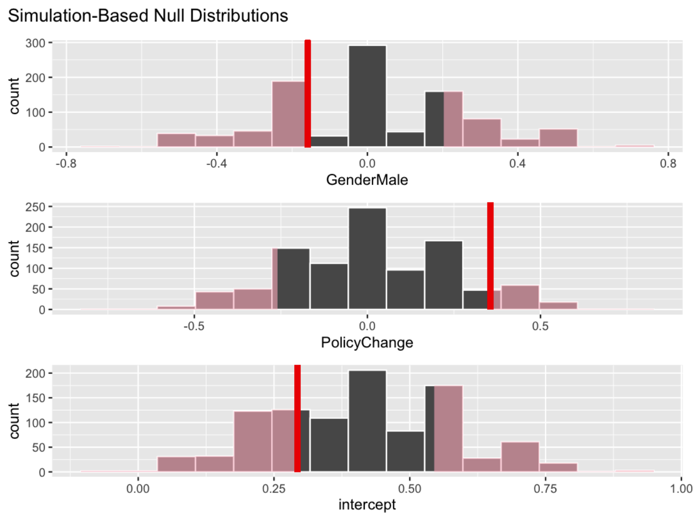
Hypothesis tests conducted removed 2-18 people each time because of respective NA Values.

ALL TESTS EXPLAIN EMPLOYMENT RATES.
IMPACT FOR PEOPLE WITHIN THE PROGRAM:

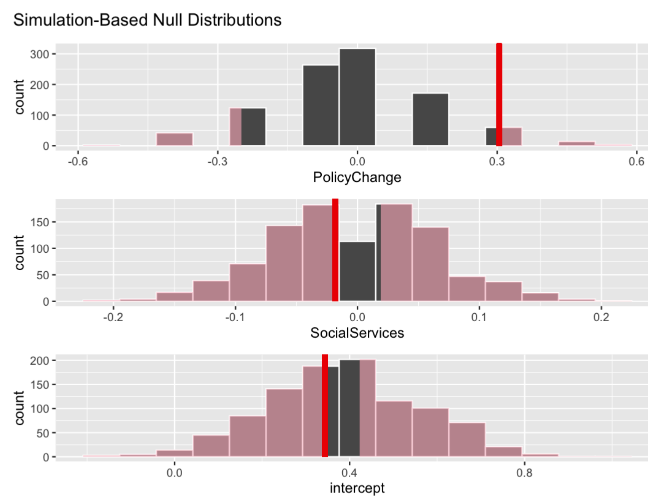


POLICY CHANGE NOT SIGNIFICANT:
Dropped 2 people
Confidence Interval: -0.0536, 0.668
P-value: 0.18
- Difference in proportions test

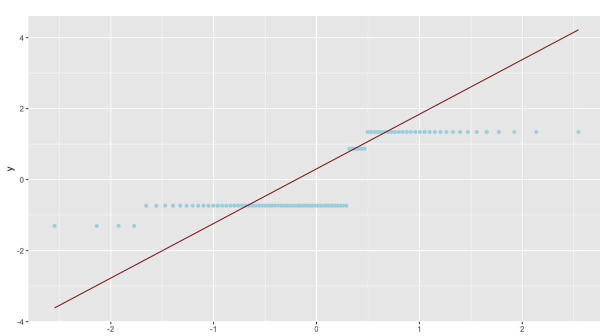
However, we believe this to be due to fewer new applicants after the policy change.



GENDER NOT SIGNIFICANT:
Dropped 18 people
This model attempted to predict employment based on of policy change and gender.
Confidence Interval: Female: -0.0273, 0.697, Male: -0.601, 0.303, PolicyChange -0.102, 0.787
P-value: Female: 0.442, Male: 0.562, PolicyChange: 0.222



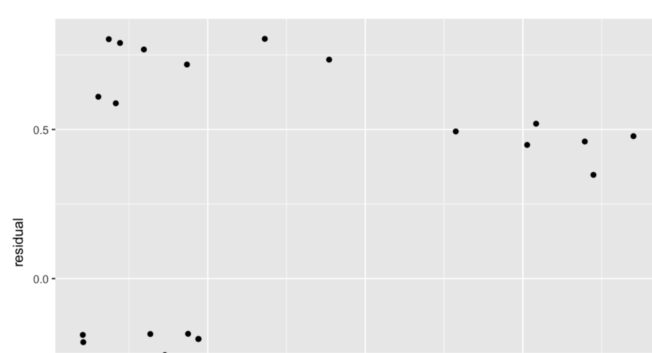
SOCIAL SERVICES NOT SIGNIFICANT: Dropped 18 people
Model made to predict employment based off of policy change and social services.
Confidence Interval: PolicyChange: -0.0867, 0.683, SocialServices: -0.155, 0.115
P-value: PolicyChange: 0.096, SocialServices: 0.724



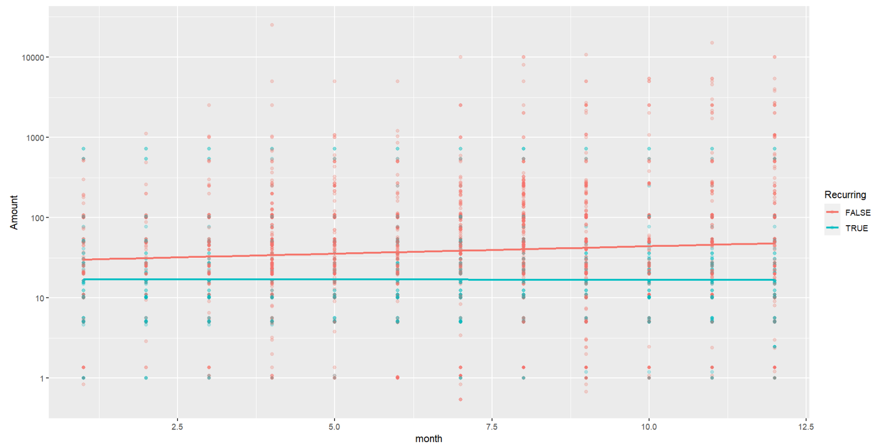
Model shown: Employed ~ SalaryLevel

Could not make an accurate model given that all variables were binary and categorical.
One attempt at a model is shown, but all models attempted were binary and couldn't provide generalizable results.
QQ Plot is shown to the left.

The residual plot for this model is shown to the right.



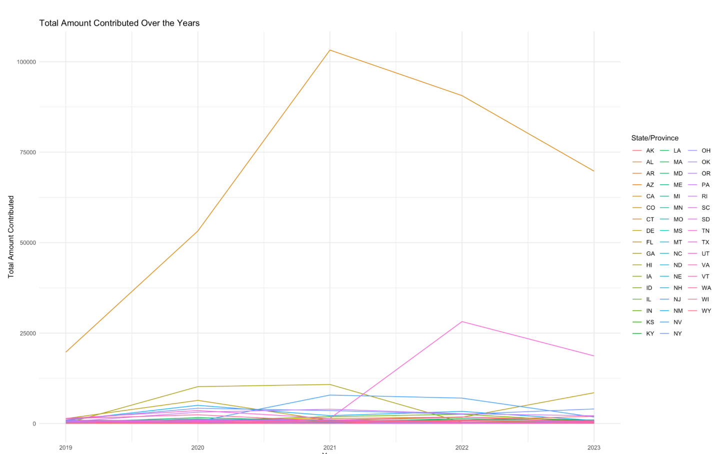
DONATIONS DATASET:



Multiple regression model:
Amount~month*Recurring

To see if the donation amount fluctuated by month for recurring donations, we created this model. The slope for recurring true is essentially 0; this makes since recurring donations are continuous and therefore don't depend on season/month. The slope for one-time donations (non-recurring) increases as the months go on, which also checks out because US donors donate more money closer to the holiday season. Overall, the comparison between recurring and non-recurring donors at Claire's Place is what would be expected from a non-profit and tracks well with the philanthropy sector.

Amount_hat = 9.8 + 22.2(month 1 to 12) + 33.0(if Recurring=TRUE) - 22.5(month 1 to 12)(if Recurring=TRUE)



We wanted to see the amount contributed over the years for each state and their percentage change. Since there were so many states we couldn't see the percentage change so this graph was unsuccessful. This information would be helpful because we could show the decrease of the amount each state donated. However, since it is not visible we focused on the total states that donated each year.

```
ggplot(AllStatesresult, aes(x = Year, y = TotalAmountContributed, color = `State/Province`)) +   geom_line() +
labs(title = "Total Amount Contributed Over the Years",
    x = "Year",    y = "Total Amount Contributed") +   theme_minimal()
```