Serving The Underserved

PPP Loans in Georgia, USA

Team 32: Danny Rivas & Javan Reuto

Table of contents

Intro

Why?

- The PPP was aimed at small business in the US however, who truly benefited from these loans and was the distribution of loans effective?
- To provide insight beyond the dataset we brought in data regarding ()

Goal: Investigate why applications were removed from the PPP loan applications and whether predictions can be made as to whether removed or not?

About the Data

PPP Georgia Loans

- PPP loans in the state of Georgia spanning from April 2020 to July 2021.
- Heavily imbalanced when combined with the removed and non-removed dataset
- Columns such as forgiveness date and amount had substantial amount of NA values

Other Datasets

To make our analysis more insightful, we decided to include additional data:

- Median Income per Zip Code
- Population per Zip Code
- Number of people under the poverty line per Zip Code
- Race per Zip Code
- Latitude and Longitude for each Zip Code

1

Data Cleaning

- Replaced Null Values
- Cleaned Locations (zip codes & cities)
- Formatted datetime
- Merged additional datasets

2

Data Analysis

After cleaning data, we imported clean data.

1. Explore Removed Dataset
2. Compared Removed Loans to Full Loans

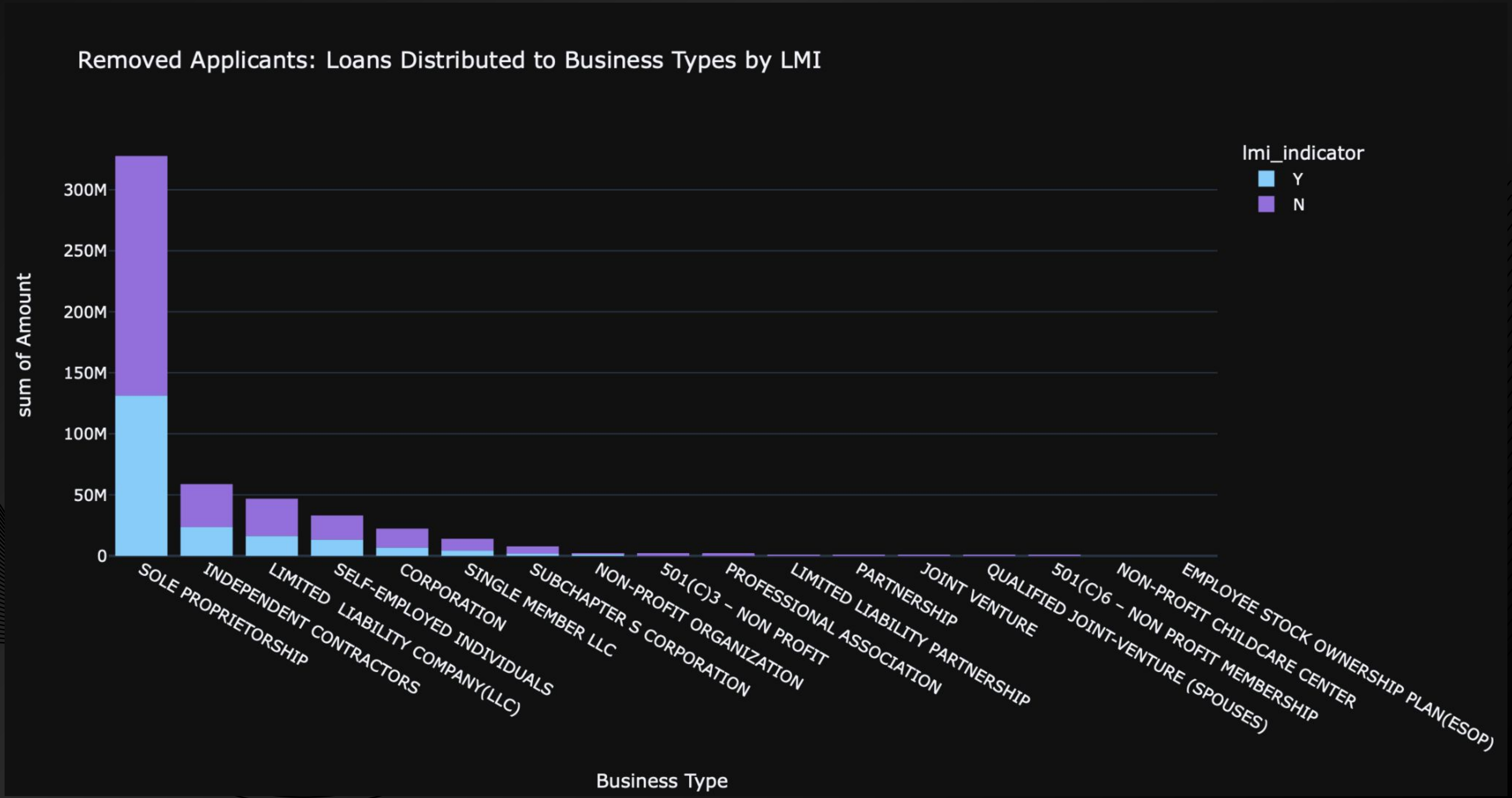
2.1 Jobs Retained by Business Type

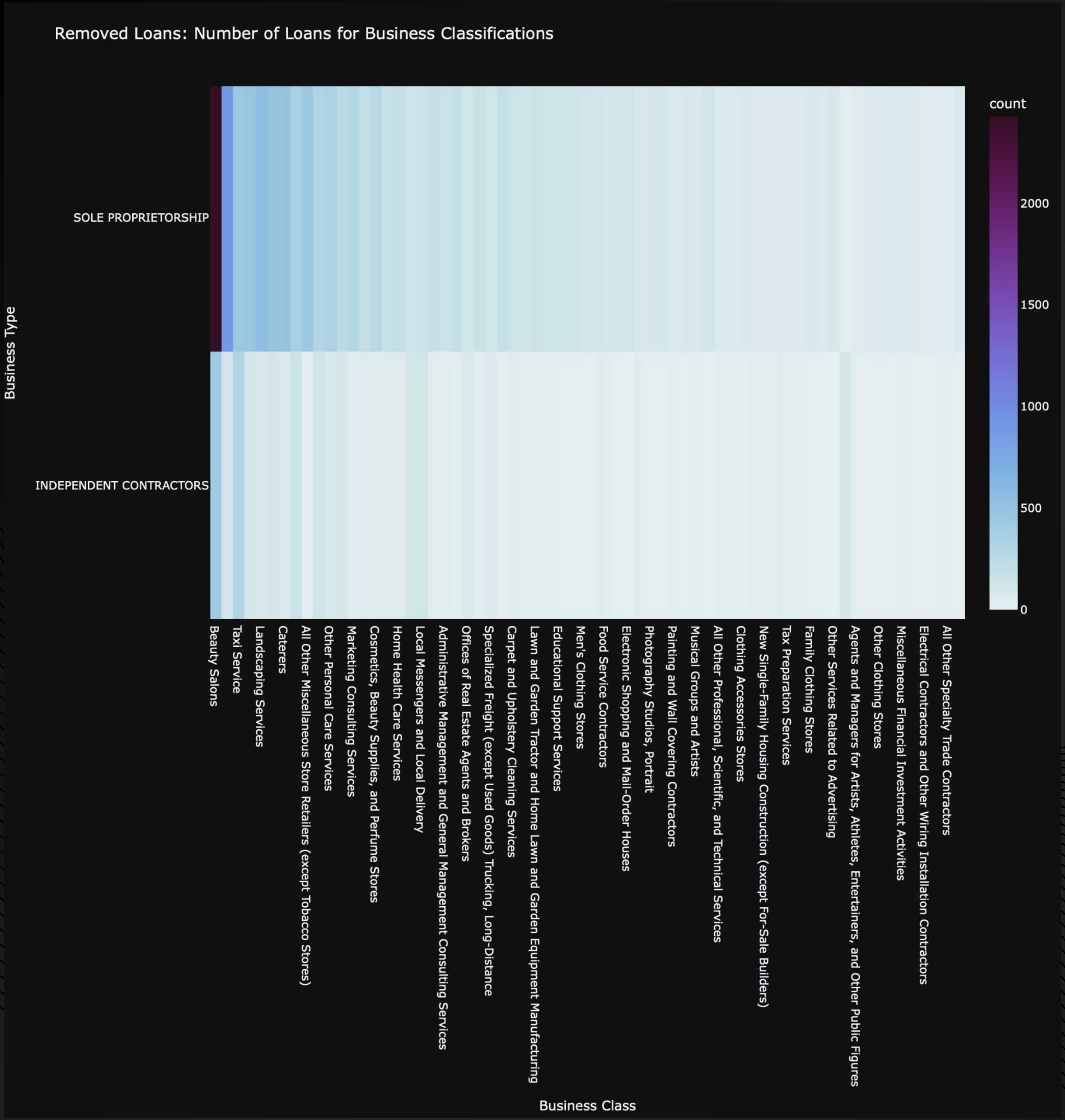
- On average, corps had the highest jobs retained
- Tax-Exempt Nonprofits requested the most amount of money

Out[3]:

	business_type	jobs_retained	amount
0	501(C)3 – NON PROFIT	5.833333	151043.131667
1	CORPORATION	6.673981	69870.952320
2	PROFESSIONAL ASSOCIATION	4.058824	65230.117647
3	NON-PROFIT ORGANIZATION	5.968750	58371.250000
4	SUBCHAPTER S CORPORATION	3.978378	41997.709297
5	LIMITED LIABILITY COMPANY(LLC)	2.872109	31833.373299
6	LIMITED LIABILITY PARTNERSHIP	4.257143	24723.402286
7	PARTNERSHIP	2.347826	24527.479130
8	SINGLE MEMBER LLC	1.633058	23126.942512
9	QUALIFIED JOINT-VENTURE (SPOUSES)	1.000000	20833.000000
10	501(C)6 – NON PROFIT MEMBERSHIP	3.000000	20525.000000
11	SOLE PROPRIETORSHIP	1.076874	19169.341763
12	JOINT VENTURE	4.000000	16669.600000
13	SELF-EMPLOYED INDIVIDUALS	1.001456	16051.311316
14	INDEPENDENT CONTRACTORS	1.000252	14780.551149
15	NON-PROFIT CHILDCARE CENTER	1.000000	3400.000000
16	EMPLOYEE STOCK OWNERSHIP PLAN(ESOP)	1.000000	2103.000000

Business Types & Low-Moderate Income Communities



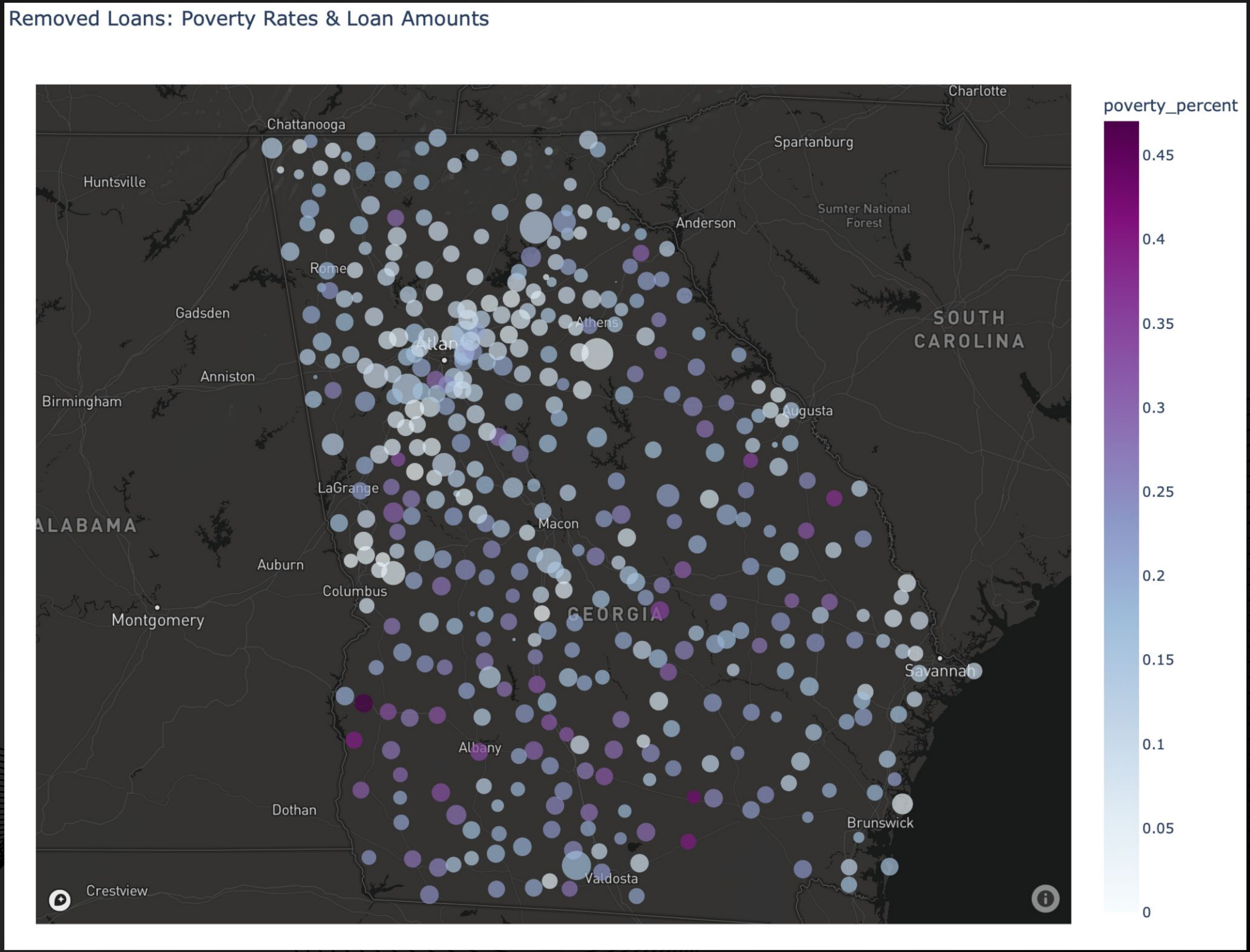


NAICS Names

- Sole proprietorships and Independent Contractors received the most loans
- Highest in business class included cash-basis type companies.
  - Less scrutiny from stakeholders, which make them vulnerable to fraud

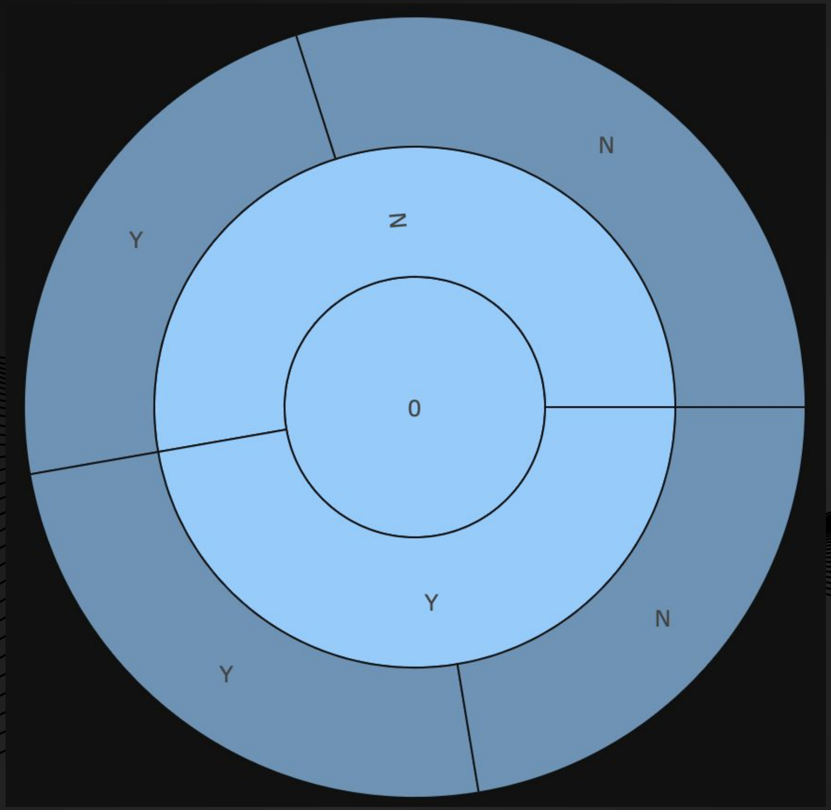
Poverty Rates & Loans Amounts by City

- Cities with lower poverty rates, receive higher loans on average
- Low poverty rate -> more people employed -> healthy economy

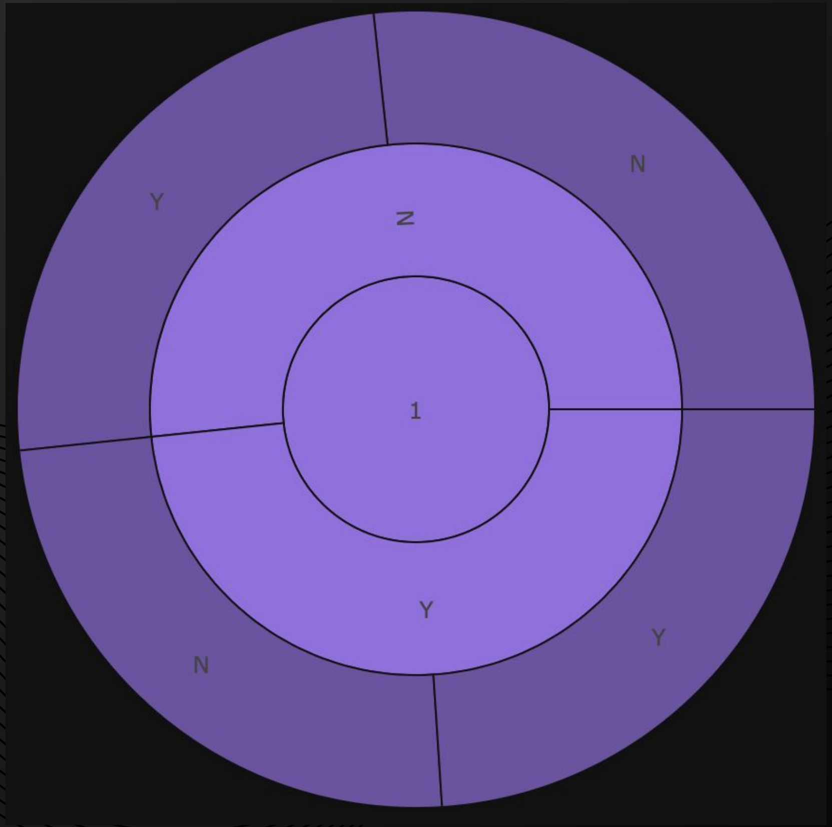


All Loans: Removed, HubZone, LMI

Not Removed



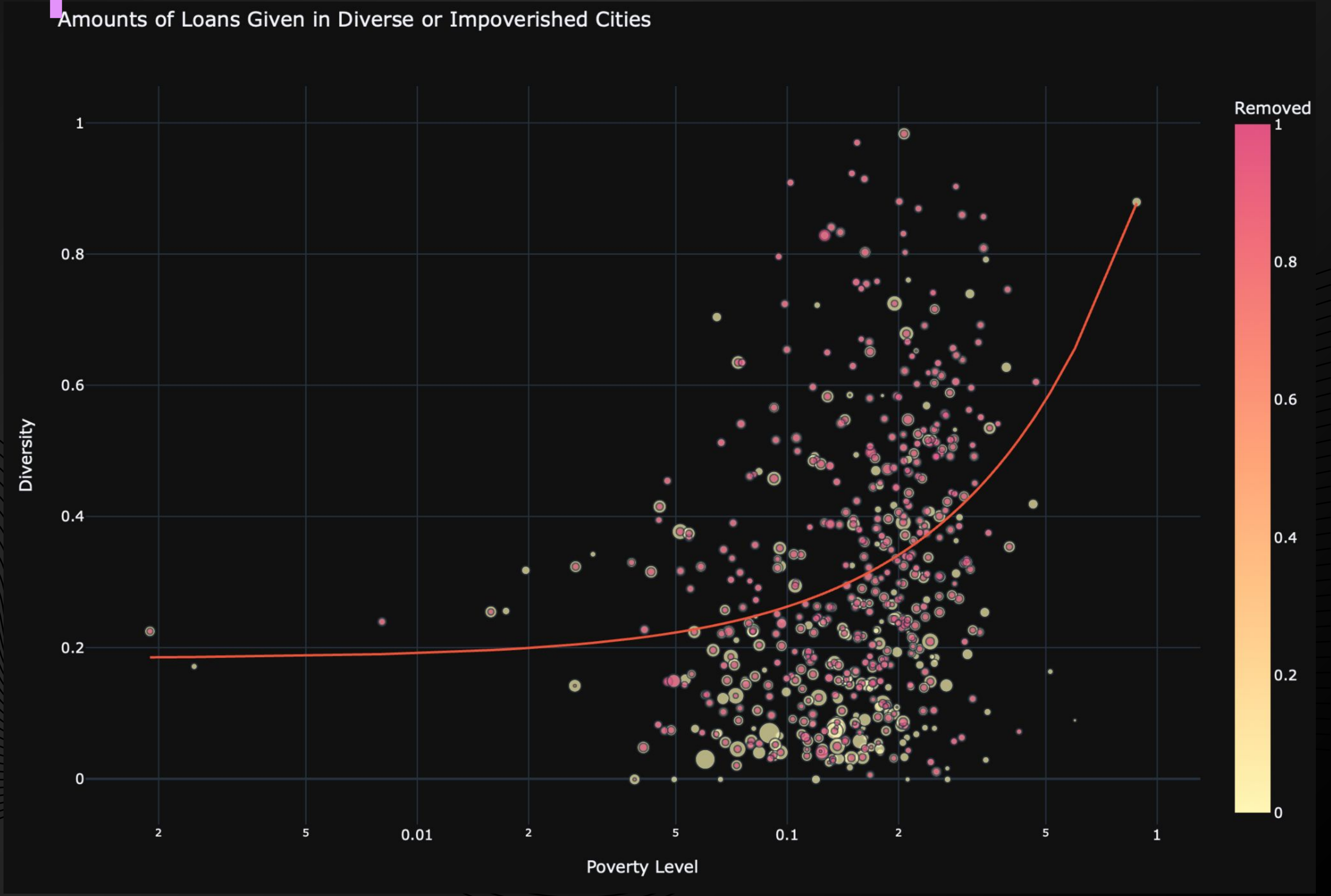
Removed



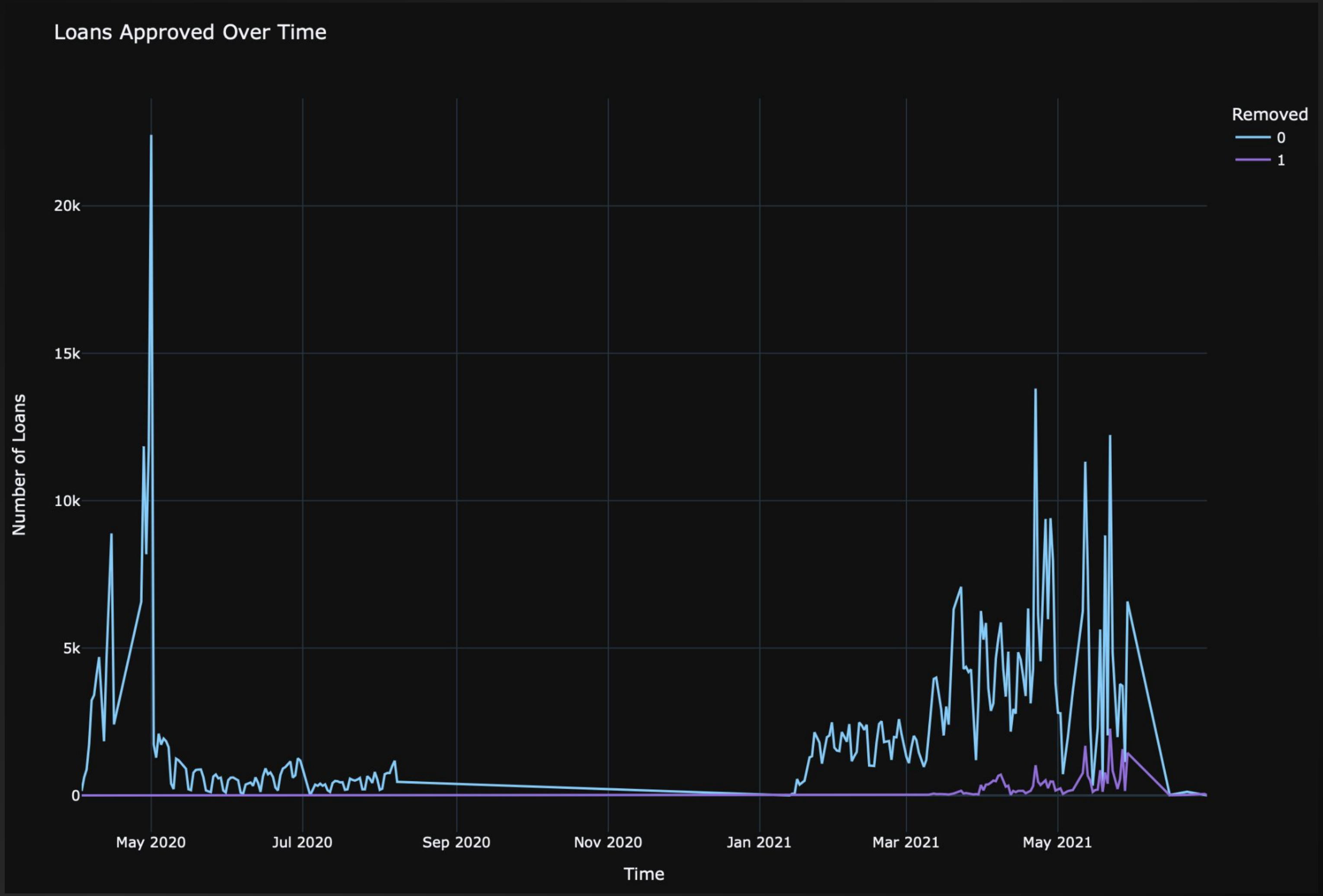
City Demographics

- Strong correlation between diversity levels and poverty levels
- Marker size represents average loan size
- Lower diversity levels received more loans and larger loan size

- Removed loans received less on average



Loans Approved Over Time



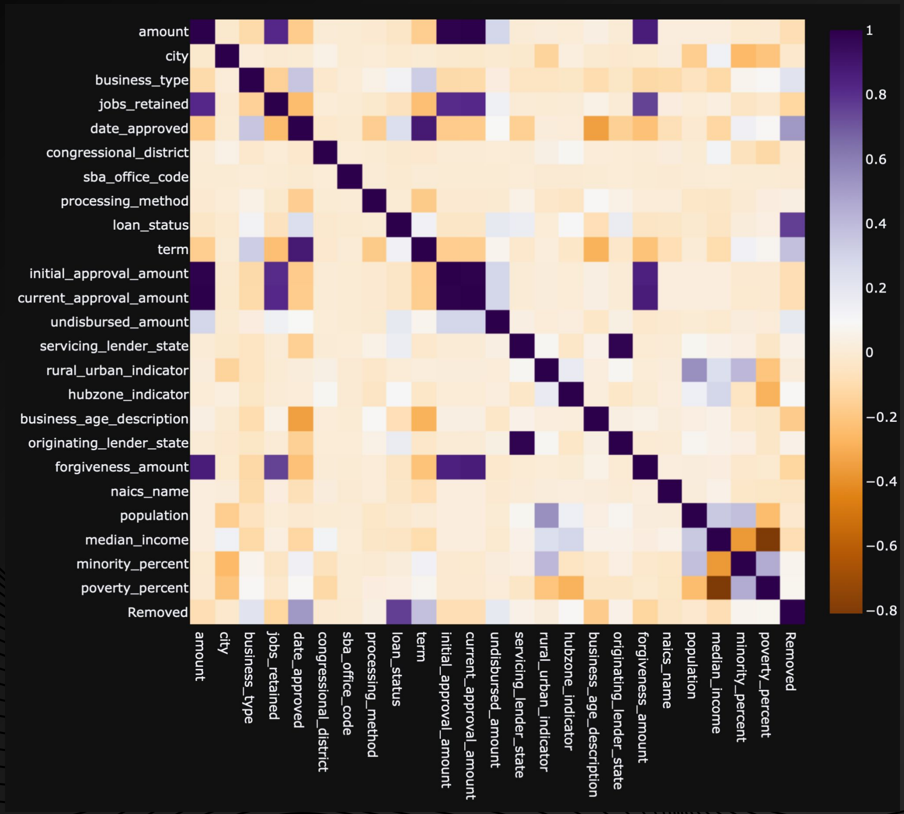
3

Prediction

Is it possible to predict whether or not a loan was removed?

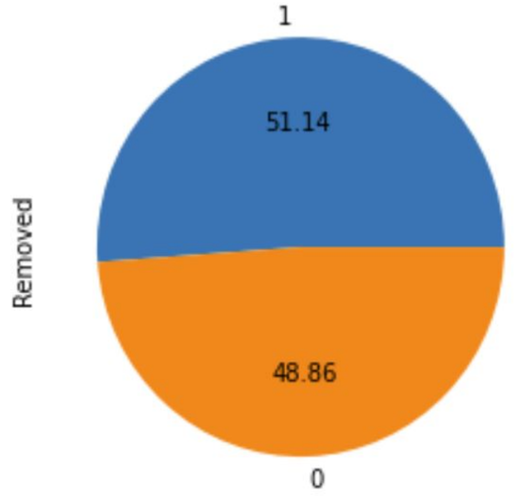
Yes, through Logistic Regression!

Imbalance and Logistic Regression



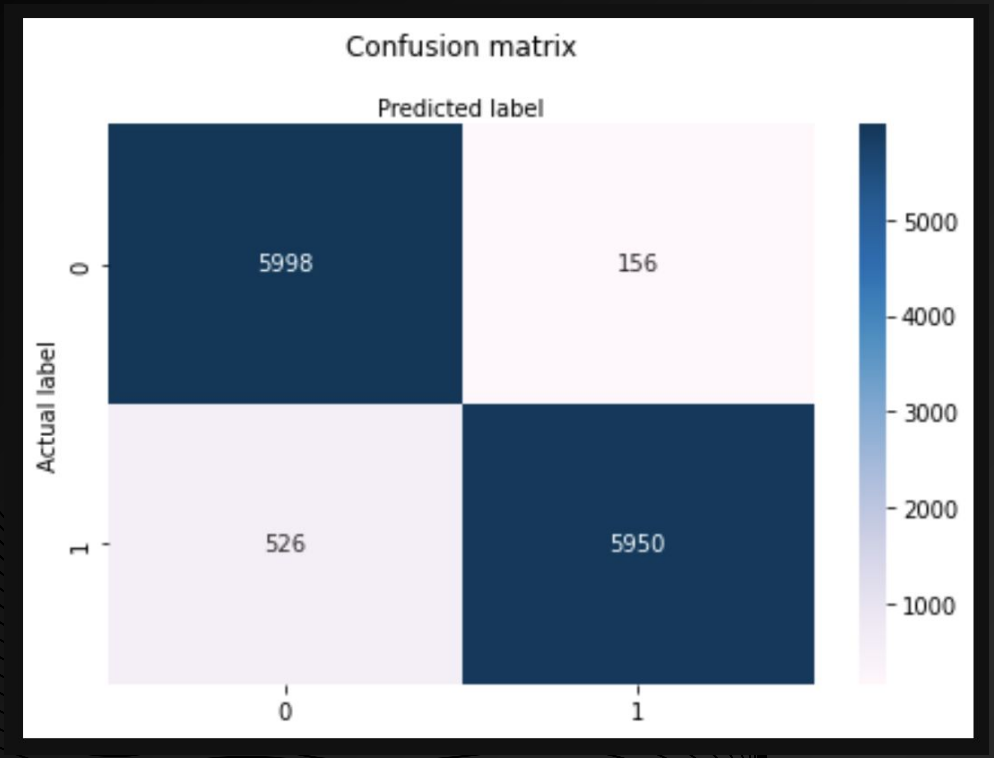

```
[16]: o = ga_1_balanced['Removed']
      o.value_counts()
      o.value_counts().plot.pie(autopct='%.2f')
```

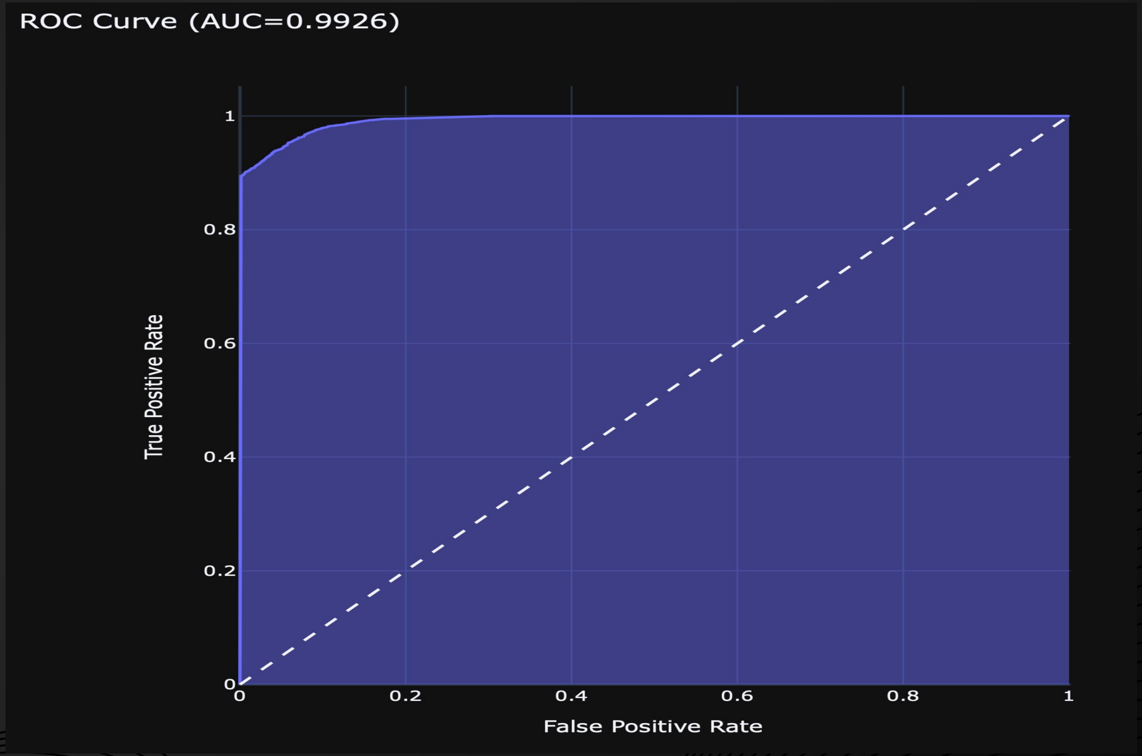
[16]: <AxesSubplot:ylabel='Removed'>



- Removed Correlation : date_approved, loan_status, business_age_description, forgiveness_amount
- Balancing of data resulted in 51/49 split as opposed to 81/19

Performance Metrics






```
[27]: print("Accuracy:",metrics.accuracy_score(y_test, y_pred))
      print("Precision:",metrics.precision_score(y_test, y_pred))
      print("Recall:",metrics.recall_score(y_test, y_pred))

Accuracy: 0.9460015835312747
Precision: 0.974451359318703
Recall: 0.9187770228536133
```

Recommendations

- Characteristics provided in the dataset alone were not enough to determine whether employees also benefited from this loan. Adding employee information provide insight to the true purpose of the PPP being fulfilled.

- Based on the data provided, applicants who received loans were from wealthier areas and companies that were relatively larger benefitted the most.

- There are more metrics to consider when making loan decisions. Hopefully, this can also benefit frontline employees.

Thank you!

Presentation Template: SlidesMania

Sample Images: Unsplash

Fonts used in this presentation: DM Sans