


How To Give A **Best Shoot** of your Paycheck Protection Program



Agenda

- 
- Background introduction
 - Research objective
 - Data description
 - Analysis Methodology
 - Key Finding
 - Conclusion
 - Recommendation

Background Introduction



- Paycheck Protection Program (PPP) provide about **800 million** funds to help small business forgive their loans
- The PPP applications are submitted by private lenders and supervised by the Small Business Administration (SBA)
- **Three** possible reasons for removed applications: withdraw by applicants/insufficient information/canceled by lender/ considered as fraudulent.

Research Objective

- Summarize the defining characteristics of removed loans.
- Evaluate the differences between removed applications and approved applications (comparison groups)
- What characteristic has more impact on the evaluation result
- Build classification model to predict whether or not a loan was removed from the data

Data Description

- We worked with two datasets periodically released from Georgia local government. The data consisted of **579,664** loan information of Georgia business.
- 4.46% of all applications is removed (25,836); **95.54%** is approved (553,828)
- Among **41** variables, we have 31 string variables, 7 numeric variables and 3 date variables
- The main attributes include the demographic and geographic info of borrower/lender/project, business type, business age and industry category.

Analysis Methodology



Platform: Google Colab



Statistical Programming tool: Python



Translate string
variables to numeric



Translate string
variables to numeric



Create new
variables



Descriptive analysis and
correlation evaluation



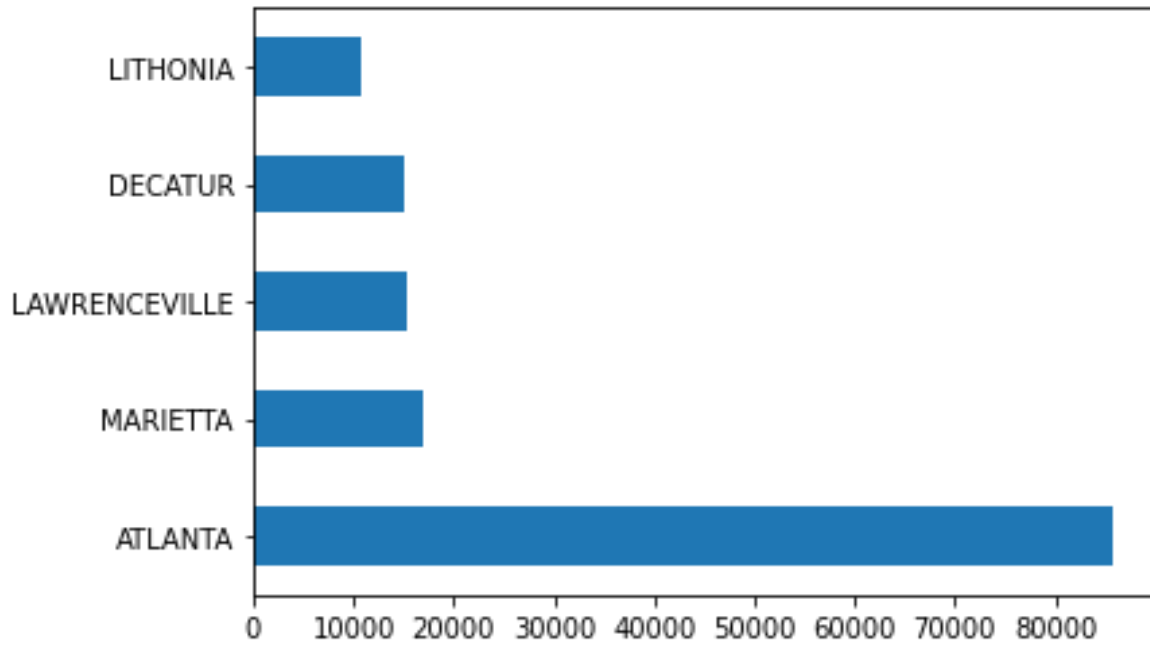
Select variables and build
classification model



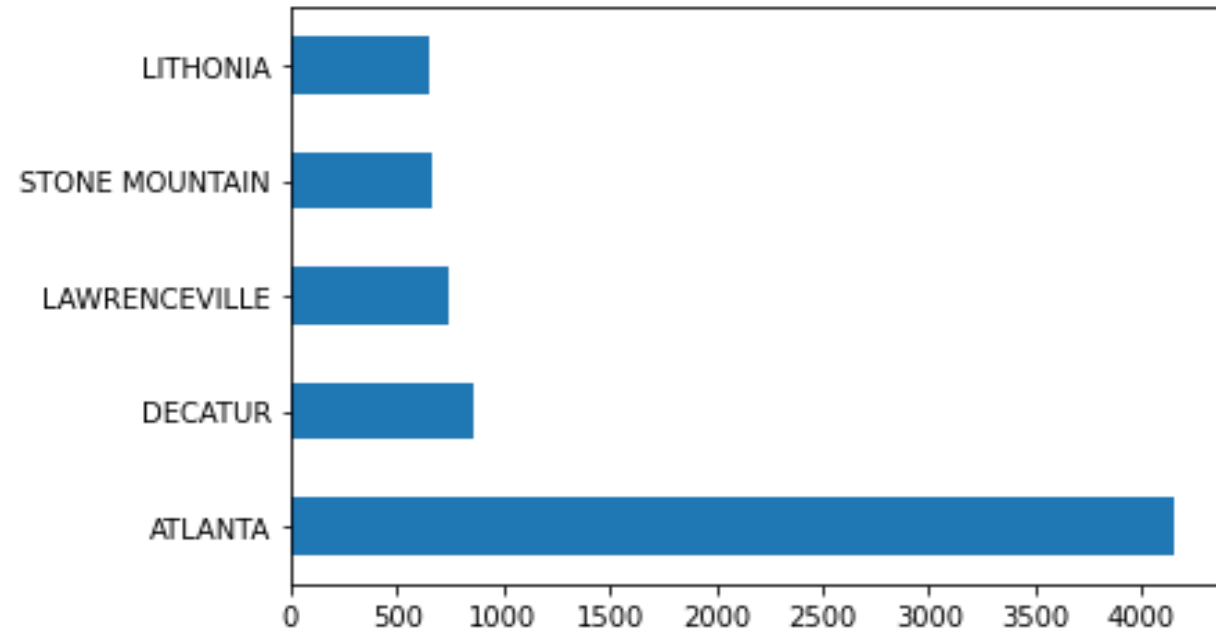
Identify the
best-fit model

Key Finding

Geographic Difference in Borrowers of Removed and Retained Loans by **City**



Cities of Retained Borrowers



Cities of Removed Borrowers

Key Finding

Industry category is another difference between removed and retained applicants



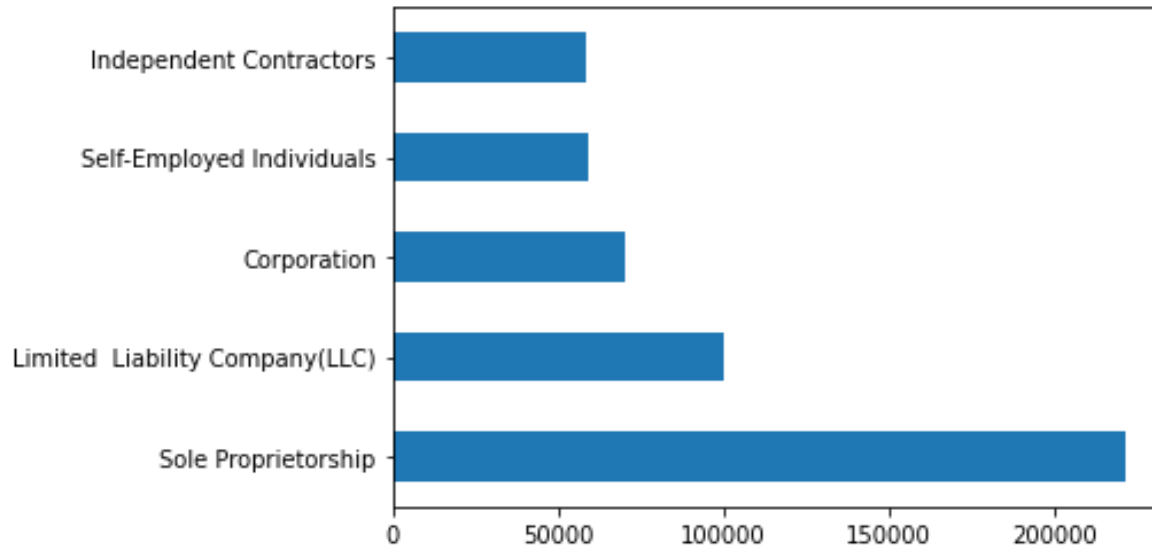
Industry category of Removed Borrowers



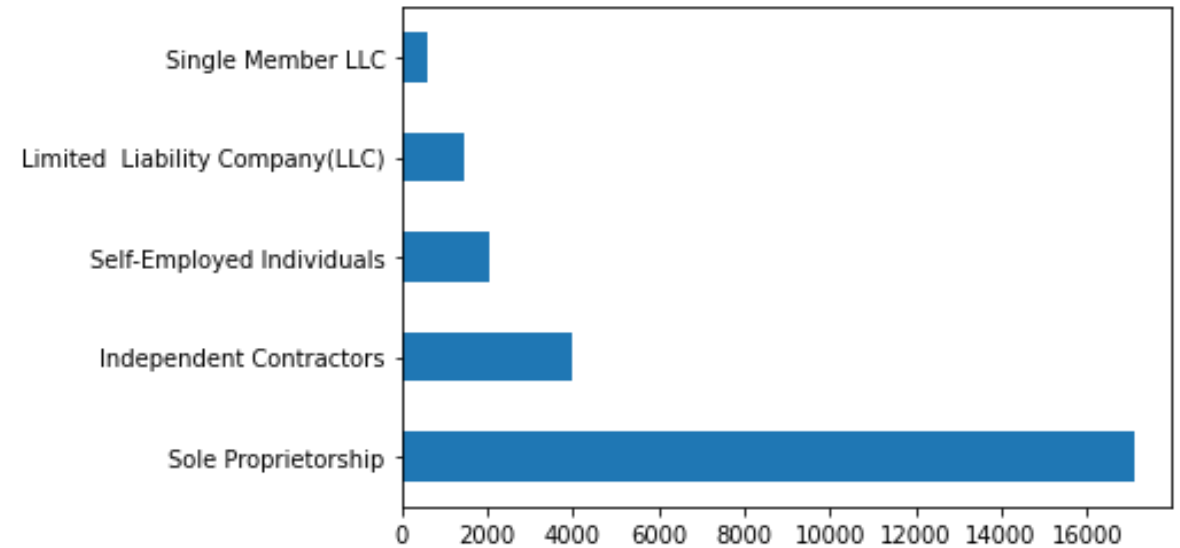
Industry category of Retained Borrowers

Key Finding

The Difference of Removed and Approved Applications by **Business Type**



Business Type of Retained Borrowers



Business Type of Removed Borrowers

- ***Business Size***

Compared to approved applications that retained about **4.93** positions on average, the removed applications were smaller businesses with **1.28** jobs retained.

- ***Undisbursed Amount***

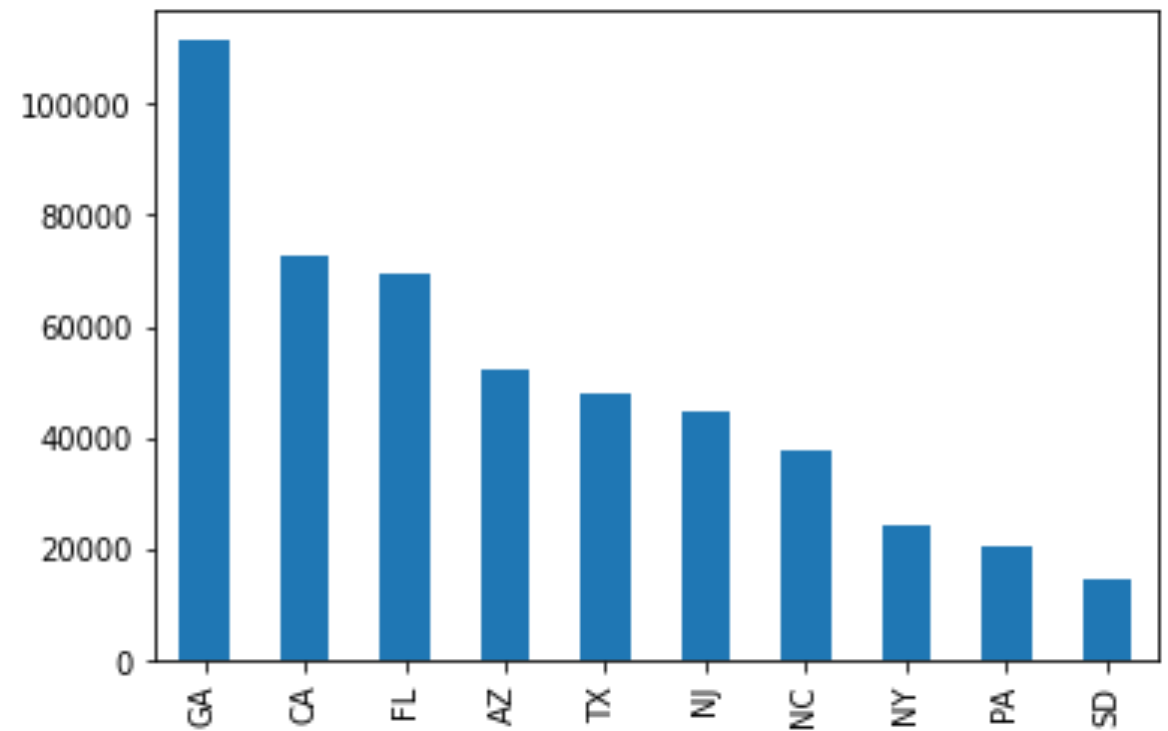
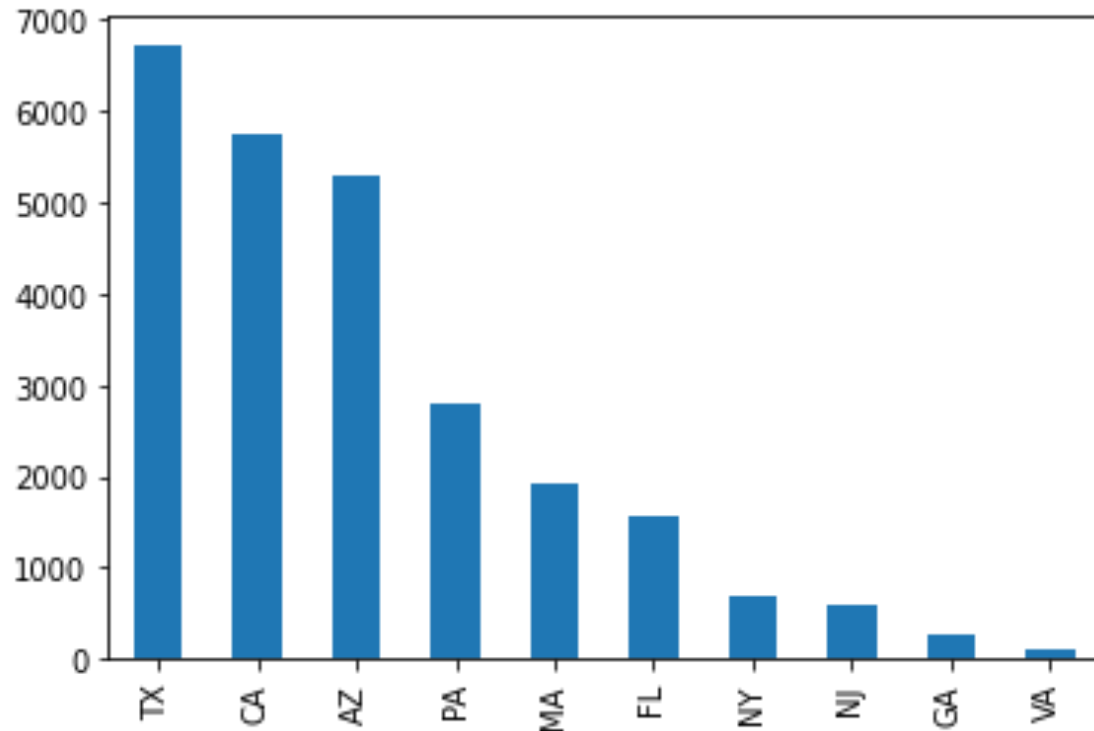
About **99.92%** of approved applicants had no undisbursed funds but only **9.69%** of the opposite group received full funding.

Key Finding



Key Finding

The Difference of Removed and Approved Applications by **Lender's Location**



Classification Model

Variables Selection
correlation between
isRemoved and other variables

term	isDisbursed
0.118474	-0.940874
undisbursed_amount	approvedyear
0.251989	0.135325
isSP	loan_statusmonth
0.110134	-0.281288

Randomly separate the datasets for training dataset and testing dataset

Train
80%

Test
20%

XGBoost

```
classification report:
      precision    recall  f1-score   support

     0       1.00      1.00      1.00     110780
     1       0.99      0.90      0.94       5153

 accuracy: 0.99
macro avg: 0.99      0.95      0.97     115933
weighted avg: 0.99      0.99      0.99     115933

confusion matrix:
[[110713    67]
 [   513   4640]]

ROC AUC:
0.9499207698129247
```

Recommendation



In order to get forgiveness for their loans:

1. Professional, Scientific and Technical Services

2. Bigger Business Size

...