

Group Project Plan

M. Lewis, P. Magee, M. Willis, and E. Wong

Project Title:

Predicting the Forgiveness of Paycheck Protection Program (PPP) Loans in Washington, DC.

Group Members

- Michael Lewis (427)
- Peter Magee (627)
- Margaret Willis (427)
- Ethan Wong (427)

Topic

Using publicly available data on PPP loans for business in DC, we seek to develop models that will explore/predict what factors were important for loan forgiveness – and if so by how much – and whether these things can be used to predict whether a loan was forgiven.

Below is what our dataset looks like.

A tibble: 6 x 29

	Borrower <chr>	Loan amount` <dbl>	Amount forgiven` <dbl>	*Loan Status` <chr>	Date approved` <date>
1	CAVA MEZZE GRI~	10000000	10116389.	Paid in Full	2020-04-08
2	AMERICAN COLLE~	10000000	10115278.	Paid in Full	2020-04-28
3	WILLIAM C. SMI~	9026082	5910159.	Exemption 4	2021-02-11
4	WILEY REIN LLP	8915913	9053194.	Paid in Full	2020-04-28

```

5 NATIONAL COMMI~      8202873.      8309511. Paid in Full   2020-04-09
6 FISCALNOTE HOL~      8000000      7807136. Exemption 4   2020-04-11
# i 24 more variables: `Date forgiven` <date>, Payroll <dbl>, Rent <dbl>,
#   Utilities <dbl>, `Health Care` <dbl>, `Mortgage Interest` <dbl>,
#   `Debt Interest` <dbl>, `Refinancing EIDL` <dbl>, `Not Provided` <dbl>,
#   `Lender name` <chr>, `Loan number` <dbl>, `Business type` <chr>,
#   Industry <chr>, `Industry (detailed)` <chr>, `Age of business` <chr>,
#   `Borrower address` <chr>, `Borrower city` <chr>, `Borrower state` <chr>,
#   `Borrower zip` <chr>, `Borrower county and state` <chr>, ...

```

Questions of Interest

Throughout the COVID-19 pandemic, an estimated \$525 Billion USD was disbursed as part of the federal government’s Paycheck Protection Program (PPP) House (2022). These loans, while crucial for protecting American industry and the workforce, have been met with allegations of fraud and controversy – in some cases spurring litigation Pfeiffer and Fast (2023). Today, the vast majority of the loans have been forgiven Pfeiffer and Fast (2023) “Some Firms Thrived During Covid and Then Got Their PPP Covid Relief Loans Forgiven” (2021). Analyzing this load data presents an opportunity to better understand (1) what business characteristics are associated with higher loan amounts, (2) what factors led to loan forgiveness – and whether the amount forgiven varied based on these things, and finally (3) whether ML techniques can be used to predict whether a loan was forgiven off these considerations.

Regression:

- Assess the relationship between `Amount forgiven` and other variables like `Rent`, `Jobs reported`, `Business type`, `Industry`, `Age of business`, `Utilities`, etc.

Classification:

- What methods have the lowest MSE for the correct prediction of whether loans were forgiven or not.

Approaches

- We propose using logistic regression, KNN, LDA, QDA to predict whether a loan was forgiven in whole or not.
- We also propose using linear regression (with cross-validation) to quantify the numerical effect of business characteristics on amount forgiven.

Group Responsibilities

- Maintain/submit Github, using branching.
- Periodic check-ins to track progress and divide the work.
- Formatting poster.

Deliverable

- Our final deliverable will be in the poster format.

References

- House, The White. 2022. “FACT SHEET: New Data Show Progress on Biden-Harris Commitment to Equity in Emergency Small Business Relief.” *The White House*. <https://www.whitehouse.gov/briefing-room/statements-releases/2022/03/11/fact-sheet-new-data-show-progress-on-biden-harris-commitment-to-equity-in-emergency-small-business-relief/>.
- Pfeiffer, Sacha, and Austin Fast. 2023. “How the Paycheck Protection Program Went from Good Intentions to a Huge Free-for-All.” *NPR*, January. <https://www.npr.org/2023/01/09/1145040599/ppp-loan-forgiveness>.
- “Some Firms Thrived During Covid and Then Got Their PPP Covid Relief Loans Forgiven.” 2021. *NBC News*. <https://www.nbcnews.com/news/firms-thrived-covid-got-ppp-covid-relief-loans-forgiven-rcna5697>.