

# Python Final Project

May 3, 2021

## 1 NJ PPP Loan Data Analysis

Data acquired from [nj.data.gov](https://nj.data.gov).

The Paycheck Protection Program (PPP) loans provide small businesses with the resources they need to maintain their payroll, hire back employees who may have been laid off, and cover applicable overhead. This data set includes businesses in New Jersey who received PPP funding, how much funding the employer received & how many jobs the employer claims they saved. The NAICS (National Industry Classification) was provided by the loan recipient.

This dataset was used to analyze the distribution of Payment Protection Program loans within New Jersey. In this notebook is a breakdown consisting of loan bracket, business owner race, among others.

```
[1]: import pandas as pd
from pandas_profiling import ProfileReport
import seaborn as sns
import numpy as np

import matplotlib as mpl
from matplotlib import rcParams
from matplotlib import style
mpl.style.use('fivethirtyeight')
import matplotlib.pyplot as plt
%matplotlib inline

import csv
import math
```

## 1.1 Data Preview

```
[2]: data = pd.read_csv('PPP_Data_fixed.csv')
data.transpose()
```

```
[2]:
```

	0	\
Loan Range	\$150000-350000,RIDGETOP CORPORATION,PO Box 644...	
Business Name	D&G LLC	
Address	TBD	
City	ENGLEWOOD	
State	NJ	
Zip	7631	
NAICS Code	722511.0	
Business Type	Limited Liability Company(LLC)	
Race Ethnicity	Unanswered	
Gender	Male Owned	
Veteran	Non-Veteran	
Non-Profit	NaN	
Jobs Retained	30.0	
Date Approved	04/27/2020	
Lender	NVE Bank	
Congressional District	NJ - 05	
Georeference	NaN	

	1	2	\
Loan Range	\$350000-1 million	\$150000-350000	
Business Name	DATA DYNAMICS INC	CHAVOND-BARRY ENGINEERING CORP	
Address	STE 102 101 CEDAR LN	400 County Route 518	
City	TEANECK	BLAWENBURG	
State	NJ	NJ	
Zip	7666	8504	
NAICS Code	518210.0	541330.0	
Business Type	Corporation	Corporation	
Race Ethnicity	Unanswered	Unanswered	
Gender	Unanswered	Unanswered	
Veteran	Unanswered	Unanswered	
Non-Profit	NaN	NaN	

Jobs Retained	25.0	10.0
Date Approved	04/14/2020	04/27/2020
Lender	Cross River Bank	Peapack-Gladstone Bank
Congressional District	NJ - 05	NJ - 07
Georeference	NaN	NaN

3 \

Loan Range	\$150000-350000
Business Name	MONMOUTH AIRCRAFT SERVICE INC
Address	RTE 530
City	BAYVILLE
State	NJ
Zip	8721
NAICS Code	488190.0
Business Type	Partnership
Race Ethnicity	Unanswered
Gender	Male Owned
Veteran	Non-Veteran
Non-Profit	NaN
Jobs Retained	0.0
Date Approved	04/15/2020
Lender	TD Bank, National Association
Congressional District	NJ - 03
Georeference	NaN

4 \

Loan Range	\$1-2 million
Business Name	BAY RIDGE AUTOMOTIVE MANAGEMENT CORP.
Address	7500 Westside Ave
City	NORTH BERGEN
State	NJ
Zip	7047
NAICS Code	441110.0
Business Type	Corporation
Race Ethnicity	Unanswered
Gender	Unanswered
Veteran	Unanswered

Non-Profit	NaN
Jobs Retained	147.0
Date Approved	04/16/2020
Lender	Citibank, N.A.
Congressional District	NJ - 08
Georeference	NaN

	5	\
Loan Range	\$1-2 million	
Business Name	DODGE OF PARAMUS, INC.	
Address	315 STATE RT 4 West	
City	PARAMUS	
State	NJ	
Zip	7652	
NAICS Code	441110.0	
Business Type	Subchapter S Corporation	
Race Ethnicity	Unanswered	
Gender	Unanswered	
Veteran	Unanswered	
Non-Profit	NaN	
Jobs Retained	NaN	
Date Approved	04/13/2020	
Lender	Ally Bank	
Congressional District	NJ - 05	
Georeference	NaN	

	6	\
Loan Range	\$350,000-1 million	
Business Name	FOURANS LLC	
Address	1031 Route 22 Suite 301	
City	BRIDGEWATER	
State	NJ	
Zip	8807	
NAICS Code	541512.0	
Business Type	Corporation	
Race Ethnicity	Unanswered	
Gender	Unanswered	

Veteran	Unanswered
Non-Profit	NaN
Jobs Retained	0.0
Date Approved	04/29/2020
Lender	TD Bank, National Association
Congressional District	NJ - 07
Georeference	NaN

7 \

Loan Range	\$150,000-350,000
Business Name	FATHER NATURE LANDSCAPING & LAWN INC.
Address	PO Box C
City	MOUNT HOLLY
State	NJ
Zip	8060
NAICS Code	561730.0
Business Type	Corporation
Race Ethnicity	Unanswered
Gender	Unanswered
Veteran	Unanswered
Non-Profit	NaN
Jobs Retained	26.0
Date Approved	04/10/2020
Lender	Cornerstone Bank
Congressional District	NJ - 03
Georeference	NaN

8 \

Loan Range	\$150,000-350,000
Business Name	FRANKLIN ALARM COMPANY, INC
Address	4294 Tuckahoe Road
City	FRANKLINVILLE
State	NJ
Zip	8322
NAICS Code	561621.0
Business Type	Corporation
Race Ethnicity	Unanswered

Gender	Unanswered
Veteran	Unanswered
Non-Profit	NaN
Jobs Retained	19.0
Date Approved	04/16/2020
Lender	Newfield National Bank
Congressional District	NJ - 02
Georeference	NaN

	9	...	\
Loan Range	\$350,000-1 million	...	
Business Name	FRED BEANS MOTORS OF FLEMINGTON, INC.	...	
Address	174 Route 202	...	
City	FLEMINGTON	...	
State	NJ	...	
Zip	8822	...	
NAICS Code	441110.0	...	
Business Type	Corporation	...	
Race Ethnicity	White	...	
Gender	Male Owned	...	
Veteran	Veteran	...	
Non-Profit	NaN	...	
Jobs Retained	67.0	...	
Date Approved	04/09/2020	...	
Lender	The Huntington National Bank	...	
Congressional District	NJ - 07	...	
Georeference	NaN	...	

	21846	\
Loan Range	\$150,000-350,000	
Business Name	THE COR BEHAVIORAL GROUP LLC	
Address	601 FLAGHOUSE DRIVE	
City	HASBROUCK HEIGHTS	
State	NJ	
Zip	7604	
NAICS Code	551112.0	
Business Type	Limited Liability Partnership	

Race Ethnicity	Unanswered
Gender	Unanswered
Veteran	Unanswered
Non-Profit	NaN
Jobs Retained	20.0
Date Approved	05/01/2020
Lender	JPMorgan Chase Bank, National Association
Congressional District	NJ - 05
Georeference	POINT (-74.061306 40.869695)

	21847 \
Loan Range	\$150,000-350,000
Business Name	VT SECURITY LLC
Address	1010 EASTPARK BLVD
City	CRANBURY
State	NJ
Zip	8512
NAICS Code	561621.0
Business Type	Limited Liability Company(LLC)
Race Ethnicity	Unanswered
Gender	Unanswered
Veteran	Unanswered
Non-Profit	NaN
Jobs Retained	18.0
Date Approved	05/03/2020
Lender	PNC Bank, National Association
Congressional District	NJ - 12
Georeference	POINT (-74.500556 40.341215)

	21848 \
Loan Range	\$350,000-1 million
Business Name	TOTAL PIPE CONNECTION INC
Address	919 Highway 33 Suite 35
City	FREEHOLD
State	NJ
Zip	7728
NAICS Code	238220.0

Business Type	Corporation
Race Ethnicity	Unanswered
Gender	Unanswered
Veteran	Unanswered
Non-Profit	NaN
Jobs Retained	20.0
Date Approved	04/12/2020
Lender	ConnectOne Bank
Congressional District	NJ - 04
Georeference	POINT (-74.226691 40.237303)

	21849 \
Loan Range	\$350,000-1 million
Business Name	WHIBCO INC
Address	87 E COMMERCE ST
City	BRIDGETON
State	NJ
Zip	8302
NAICS Code	212322.0
Business Type	Subchapter S Corporation
Race Ethnicity	Unanswered
Gender	Male Owned
Veteran	Unanswered
Non-Profit	NaN
Jobs Retained	47.0
Date Approved	04/10/2020
Lender	Truist Bank d/b/a Branch Banking & Trust Co
Congressional District	NJ - 02
Georeference	POINT (-75.233698 39.428951)

	21850 \
Loan Range	\$150,000-350,000
Business Name	TETHERVIEW, LLC.
Address	1000 SANGER AVE SUITE 315
City	OCEANPORT
State	NJ
Zip	7757



NAICS Code	531190.0
Business Type	Limited Liability Company(LLC)
Race Ethnicity	Unanswered
Gender	Unanswered
Veteran	Unanswered
Non-Profit	NaN
Jobs Retained	19.0
Date Approved	04/07/2020
Lender	OceanFirst Bank, National Association
Congressional District	NJ - 04
Georeference	NaN

	21851 \
Loan Range	\$150,000-350,000
Business Name	ZOO KISS LOGISTICS, LLC.
Address	575 BEECH ST
City	HAWORTH
State	NJ
Zip	7641
NAICS Code	484110.0
Business Type	Limited Liability Company(LLC)
Race Ethnicity	Unanswered
Gender	Unanswered
Veteran	Unanswered
Non-Profit	NaN
Jobs Retained	78.0
Date Approved	04/28/2020
Lender	Queensborough National Bank and Trust Company
Congressional District	NJ - 05
Georeference	POINT (-74.000844 40.959387)

	21852 \
Loan Range	\$150,000-350,000
Business Name	TREZZCO WAREHOUSE INC.
Address	570 South Avenue East
City	CRANFORD
State	NJ

Zip	7016
NAICS Code	541614.0
Business Type	Corporation
Race Ethnicity	Unanswered
Gender	Male Owned
Veteran	Non-Veteran
Non-Profit	NaN
Jobs Retained	19.0
Date Approved	04/28/2020
Lender	Valley National Bank
Congressional District	NJ - 07
Georeference	POINT (-74.29003 40.656512)

	21853 \
Loan Range	\$150,000-350,000
Business Name	UNITED TALMUDICAL ACADEMY OF LAKEWOOD INC
Address	485 Locust St
City	LAKEWOOD
State	NJ
Zip	8701
NAICS Code	611110.0
Business Type	Corporation
Race Ethnicity	Unanswered
Gender	Unanswered
Veteran	Unanswered
Non-Profit	NaN
Jobs Retained	30.0
Date Approved	04/28/2020
Lender	Harvest Small Business Finance, LLC
Congressional District	NJ - 03
Georeference	POINT (-74.20136 40.046054)

	21854 \
Loan Range	\$150,000-350,000
Business Name	VITAL INSULATION SERVICES, INC
Address	3125 WOODBRIDGE AVE STE 2A
City	EDISON

State	NJ
Zip	8837
NAICS Code	238310.0
Business Type	Subchapter S Corporation
Race Ethnicity	Unanswered
Gender	Male Owned
Veteran	Unanswered
Non-Profit	NaN
Jobs Retained	5.0
Date Approved	05/01/2020
Lender	JPMorgan Chase Bank, National Association
Congressional District	NJ - 06
Georeference	POINT (-74.328121 40.529783)

	21855
Loan Range	\$150,000-350,000
Business Name	TRANSMODAL CORPORATION
Address	48 S Franklin Turnpike
City	RAMSEY
State	NJ
Zip	7446
NAICS Code	488510.0
Business Type	Corporation
Race Ethnicity	Unanswered
Gender	Unanswered
Veteran	Unanswered
Non-Profit	NaN
Jobs Retained	0.0
Date Approved	04/27/2020
Lender	Bank Leumi USA
Congressional District	NJ - 05
Georeference	POINT (-74.131799 41.056986)

[17 rows x 21856 columns]

## 1.2 White Owned

```
[3]: WhtOwned = data.copy()
whtNum = WhtOwned[WhtOwned['Race Ethnicity'].str.contains("White")]
num = str(whtNum.shape[0])
print("There are " + num + " White owned businesses")
whtNum.head()
```

There are 973 White owned businesses

```
[3]:
```

	Loan Range	Business Name \
9	\$350,000-1 million	FRED BEANS MOTORS OF FLEMINGTON, INC.
13	\$150,000-350,000	PHOENIX PRECISION INC
20	\$350,000-1 million	L&M ARCHITECTURAL GRAPHICS, INC
21	\$150,000-350,000	LEPPERT GROUP
120	\$150,000-350,000	PBKK ENTERPRISES INC

	Address	City State	Zip	NAICS Code \
9	174 Route 202	FLEMINGTON NJ	8822	441110.0
13	2963 Route 23 South	NEWFOUNDLAND NJ	7435	332710.0
20	62 Parsppany Blvd	PARSIPPANY NJ	7054	339950.0
21	ONE COUNTY ROAD 560	SANDYSTON NJ	7826	541211.0
120	1670 Route 34 N Suite 1C	WALL NJ	7727	722513.0

	Business Type	Race Ethnicity	Gender	Veteran \
9	Corporation	White	Male Owned	Veteran
13	Subchapter S Corporation	White	Male Owned	Non-Veteran
20	Corporation	White	Male Owned	Non-Veteran
21	Subchapter S Corporation	White	Male Owned	Unanswered
120	Corporation	White	Male Owned	Non-Veteran

	Non-Profit	Jobs Retained	Date Approved	Lender \
9	NaN	67.0	04/09/2020	The Huntington National Bank
13	NaN	19.0	04/28/2020	Lakeland Bank
20	NaN	30.0	04/04/2020	Columbia Bank
21	NaN	14.0	04/14/2020	SB One Bank
120	NaN	60.0	04/14/2020	Lakeland Bank

	Congressional District	Georeference
9	NJ - 07	NaN
13	NJ - 05	NaN
20	NJ - 11	NaN
21	NJ - 05	NaN
120	NJ - 04	NaN

### 1.3 Black/Afr. Amer. Owned

```
[4]: AfrOwned = data.copy()
blkNum = AfrOwned[AfrOwned['Race Ethnicity'].str.contains("Black")]
num = str(blkNum.shape[0])
print("There are " + num + " Black/African American owned businesses")
blkNum.head()
```

There are 26 Black/African American owned businesses

```
[4]:
```

	Loan Range	Business Name \
966	\$150,000-350,000	SADAT ASSOCIATES, INC
1159	\$350,000-1 million	ARG MEDIA
1369	\$150,000-350,000	NEXT MILE LLC
2218	\$150,000-350,000	SPECIALIZED THERAPY ASSOCIATES
3014	\$150,000-350,000	DIVERSITY FOOD BRANDS, LLC

	Address	City State	Zip	NAICS Code \
966	1545 LAMBERTON RD	TRENTON NJ	8611	541620.0
1159	29 Columbia Turnpike Suite 102B	FLORHAM PARK NJ	7932	541613.0
1369	109 GARNER AVENUE	BLOOMFIELD NJ	7003	484110.0
2218	4 Forest Avenue	PARAMUS NJ	7652	621111.0
3014	41 Watchung Plaza 339	MONTCLAIR NJ	7042	236220.0

	Business Type	Race Ethnicity	Gender \
966	Corporation	Black or African American	Male Owned
1159	Limited Liability Company(LLC)	Black or African American	Unanswered
1369	Limited Liability Company(LLC)	Black or African American	Male Owned

2218		Sole Proprietorship	Black or African American	Unanswered
3014	Limited Liability Company(LLC)		Black or African American	Male Owned

	Veteran	Non-Profit	Jobs Retained	Date Approved	\
966	Non-Veteran	NaN	16.0	04/11/2020	
1159	Unanswered	NaN	20.0	04/28/2020	
1369	Veteran	NaN	52.0	06/29/2020	
2218	Unanswered	NaN	33.0	04/28/2020	
3014	Non-Veteran	NaN	29.0	05/05/2020	

	Lender	Congressional District	\
966	Berkshire Bank	NJ - 04	
1159	Harvest Small Business Finance, LLC	NJ - 11	
1369	Cross River Bank	NJ - 09	
2218	Harvest Small Business Finance, LLC	NJ - 05	
3014	JPMorgan Chase Bank, National Association	NJ - 10	

	Georeference
966	POINT (-74.752342 40.186607)
1159	POINT (-74.375149 40.7818)
1369	POINT (-74.17655 40.837934)
2218	POINT (-74.053367 40.918209)
3014	POINT (-74.206159 40.830265)

## 1.4 Asian Owned

```
[5]: AsianOwned = data.copy()
asianNum = AsianOwned[AsianOwned['Race Ethnicity'].str.contains("Asian")]
num = str(asianNum.shape[0])
print("There are " + num + " Asian owned businesses")
asianNum.head()
```

There are 269 Asian owned businesses

```
[5]: Loan Range Business Name \
26 $150,000-350,000 MYONG DONG OF NEW JERSEY INC
```

102	\$150,000-350,000	RATAN HOTEL PLAZA LLC
159	\$150,000-350,000	SMILE DENTAL CARE
246	\$150,000-350,000	GENTLE HEART CARE LLC
265	\$1-2 million	DATA GLOVE INCORPORATED

	Address	City	State	Zip	NAICS Code \
26	365 Route 17 South	PARAMUS	NJ	7652	722511.0
102	120 Evergreen Place	EAST ORANGE	NJ	7018	721110.0
159	666 Plainsboro Rd	PLAINSBORO	NJ	8536	621111.0
246	333 Sylvan Ave #205	ENGLEWOOD CLIFFS	NJ	7632	621610.0
265	125 VILLAGE BLVD Suite 270	PRINCETON	NJ	8540	541519.0

	Business Type	Race	Ethnicity	Gender \
26	Corporation	Asian	Unanswered	
102	Limited Liability Company(LLC)	Asian	Male Owned	
159	Corporation	Asian	Female Owned	
246	Limited Liability Company(LLC)	Asian	Female Owned	
265	Cooperative	Asian	Female Owned	

	Veteran	Non-Profit	Jobs Retained	Date Approved \
26	Unanswered	NaN	39.0	04/15/2020
102	Non-Veteran	NaN	29.0	04/28/2020
159	Unanswered	NaN	17.0	04/29/2020
246	Non-Veteran	NaN	90.0	04/11/2020
265	Non-Veteran	NaN	79.0	04/08/2020

	Lender	Congressional District \
26	Bank of Hope	NJ - 05
102	Flushing Bank	NJ - 10
159	Piermont Bank	NJ - 12
246	Hanmi Bank	NJ - 09
265	Peapack-Gladstone Bank	NJ - 07

	Georeference
26	NaN
102	POINT (-74.222042 40.762078)
159	POINT (-74.573829 40.32821)

```
246 POINT (-73.953005 40.873966)
265 POINT (-74.612035 40.355811)
```

```
[6]: csv = pd.DataFrame(pd.read_csv('PPP_Data_fixed.csv'))
      print('\n')
      profile = ProfileReport(csv, title='CSV Pandas Profiling Report', explorative = True, minimal = True)
```

```
[7]: profile.to_notebook_iframe()
```

```
Summarize dataset: 0%|          | 0/25 [00:00<?, ?it/s]
Generate report structure: 0%|          | 0/1 [00:00<?, ?it/s]
Render HTML: 0%|          | 0/1 [00:00<?, ?it/s]
<IPython.core.display.HTML object>
```

```
[8]: # Imports csv and selects the columns of interest
      # Note: data are already of type float
      df = pd.read_csv('PPP_Data_fixed - Jess.csv')
      df = df[["NAICS Code", "Jobs Retained"]]

      # Remove rows with empty data
      df.dropna(inplace = True)

      # Make sorted list of unique NAICS codes
      codes = sorted(list(set(np.array([238220, 541110, 541511, 621111, 722511]))))
      decodes = sorted(list(set(np.array(["Plumbing/HVAC", "Law Offices", "Computer Programming", "Physicians", "Restaurants"]))))
      jobs = []

      # Sum jobs retained for each job code
      for code in codes:
          total = df.loc[df["NAICS Code"] == code, "Jobs Retained"].sum()
          jobs.append(total)
```



```

# Make ticks evenly spaced despite their values
x_pos = np.arange(len(codes))

fig, ax = plt.subplots(figsize=(8,3))

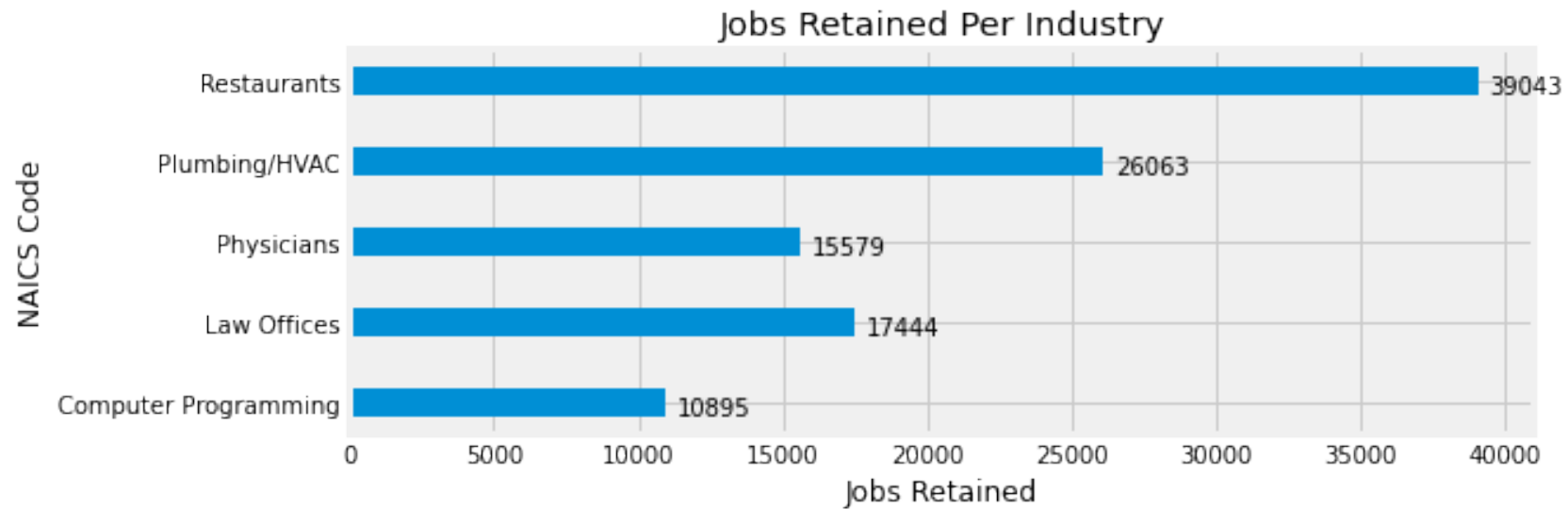
# Add chart labels
plt.barh(x_pos, jobs, .35)
plt.title("Jobs Retained Per Industry")
plt.ylabel("NAICS Code")
plt.xlabel("Jobs Retained")

plt.yticks(x_pos, decodes)
plt.setp(ax.yaxis.get_majorticklabels(), rotation=0, ha="right", rotation_mode="anchor")

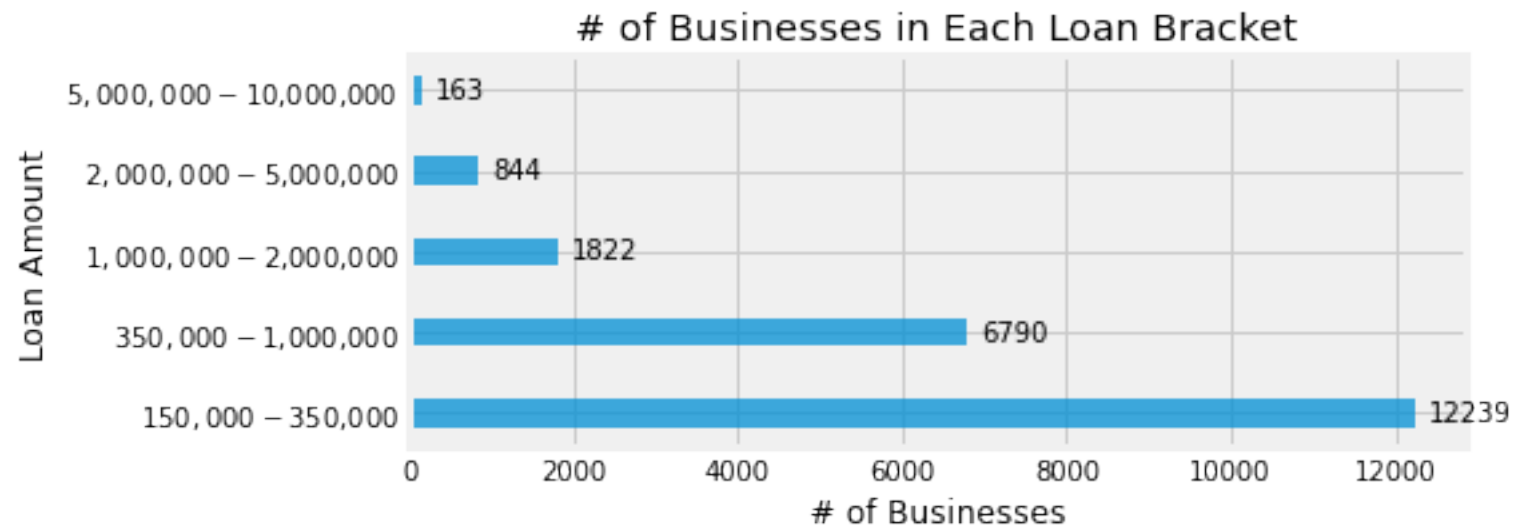
for p in ax.patches:
    # ax.annotate(str(p.get_x()), (p.get_width() * 1.005, p.get_x() * 1.005))
    ax.annotate(str(p.get_width().astype(int)), (p.get_x() + p.get_width(), p.get_y()), xytext=(5, 1), textcoords='offset points')

plt.show()

```



```
[9]: pd.set_option('precision', 0)
fields = ['Loan Range', 'NAICS Code']
order = ['$5,000,000-$10,000,000', '$2,000,000-$5,000,000', '$1,000,000-$2,000,000', '$350,000-$1,000,000', '$150,000-$350,000']
dfJob = pd.DataFrame(pd.read_csv('PPP_Data_fixed - Jess.csv', skipinitialspace=True, usecols=fields))
ax = dfJob.groupby('Loan Range').size().reindex(order)\
.plot(kind='barh', alpha=0.75, figsize=(8,3),\
      title="# of Businesses in Each Loan Bracket",\
      xlabel="Loan Amount", width=.35)
plt.gca().invert_yaxis()
for p in ax.patches:
    ax.annotate(str(p.get_width()), (p.get_x() + p.get_width(), p.get_y()), xytext=(5, -8), textcoords='offset points')
plt.xlabel("# of Businesses")
plt.tight_layout()
```



[ ]: