

# Disaster Loans, Market Measures, and Unemployment

By Christopher Ng, Jake Thompson, Mikhaela Anderson, Nick Kartschoke, Ryan Lippe

# Outline

---

Our Process

---

About Our Data

---

Data Preprocessing

---

Hypothesis / Exploratory Questions

---

Data Insights

---

Machine Learning and Results

---

Recommendations

---

Questions



# Our Process

- Initially, we knew we wanted to say something about unemployment and started our search for data with that topic in mind
- We first found the Small Business Administration (SBA) disaster loans and became curious to see if we could find a relationship between disasters and unemployment
- As unemployment is complex, we knew we needed more data sources to consider and chose local government finance data from the U.S. Census and data from the S&P 500

# About Our Data

- Unemployment Data
  - The dataset was built by scraping the U.S. Labor Bureau of Labor Statistics website and includes data of unemployment by county. It represents the Local Area Unemployment statistics from 1990-2016
- SBA Disaster Loans
  - Following a declared disaster, the SBA provides low-interest, long-term disaster loans for damages not covered by insurance. The data includes information on a county level, SBA Disaster Numbers to refer to the specific disaster, verified reported damages, and awarded loan amounts
- Local Government Finance Historical Datasets from the U.S. Census
  - The State & Local Government Finance Historical Datasets include statistics on revenue, expenditure, debt, and assets. For our research, we included columns such as “Total Salaries & Wages”, “Property Tax”, “Interest Revenue”, and “Fines and Forfeits” to maximize the amount of data we had in lieu of the amount of nulls found in other columns
- S&P 500
  - The Standard and Poor’s 500 (S&P 500) is a stock market index tracking the performance of 500 large companies listed on stock exchanges in the U.S. The stock market index is a measurement of a portion of the stock market and is a large indicator of the health of the economy and we knew we would be remiss if we did not include it while predicting unemployment

# Data Preprocessing



Using our four datasets from different sources, our goal was to transform the data into a uniform format to ultimately merge it into one table and feed it to our ML model

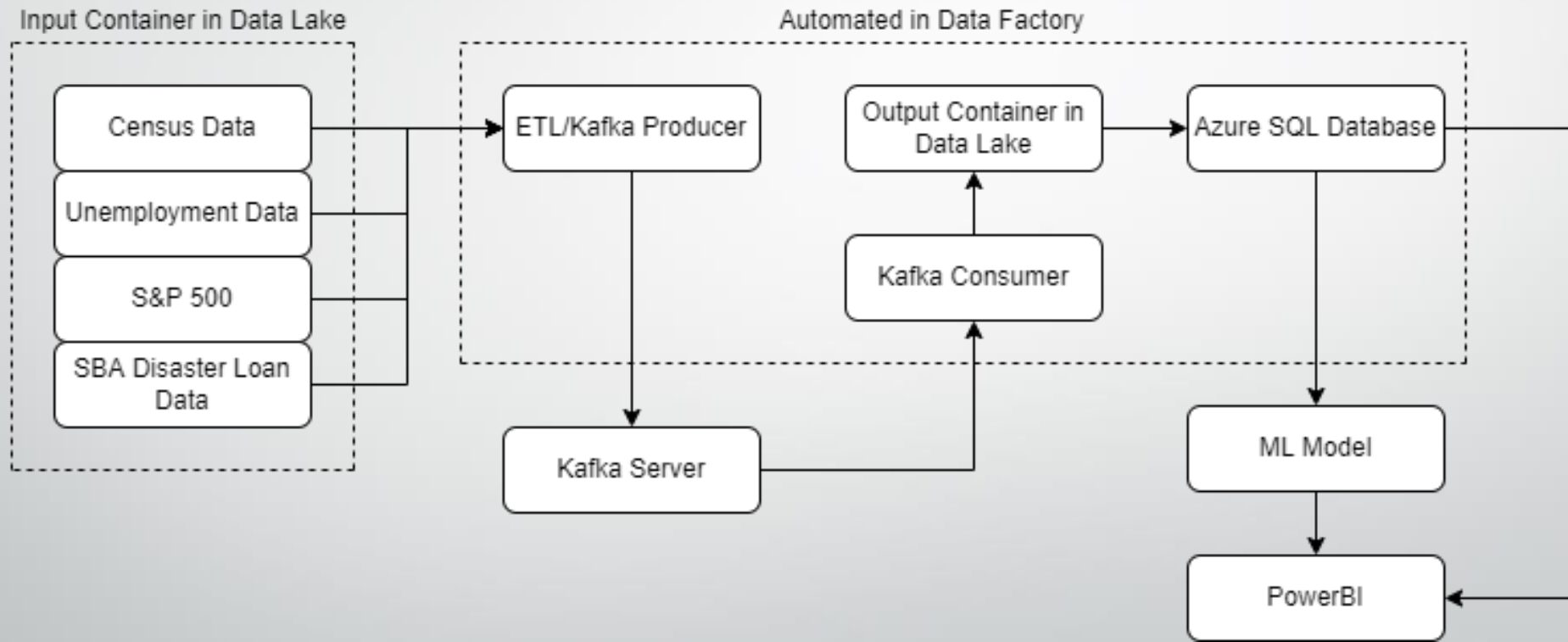


To get the datasets into a uniform format, each set had to undergo an ETL process that excluded null values, included county level data, and were within the years 2008-2015



It was then merged into a single table to be fed into a Kafka server and ultimately be stored in a relational SQL database which we'll refer to for machine learning and other analysis

# Cloud Diagram



# Hypothesis & Procedure

- Unemployment can be predicted by a combination of disaster loans, local government finances, and market data
- We used data points between 2008 and 2015 on a county level
- We used an elastic net regression model using cross validation
- We judged which factors were the best predictors of unemployment by looking at coefficient weights of the finished model

# Exploratory Questions

1. Are disaster loans a significant predictor of unemployment?

2. Does the amount of the loan affect mitigating unemployment?

3. Which disasters saw the most loan money awarded?

4. Were local government revenues correlated with increased unemployment rate?

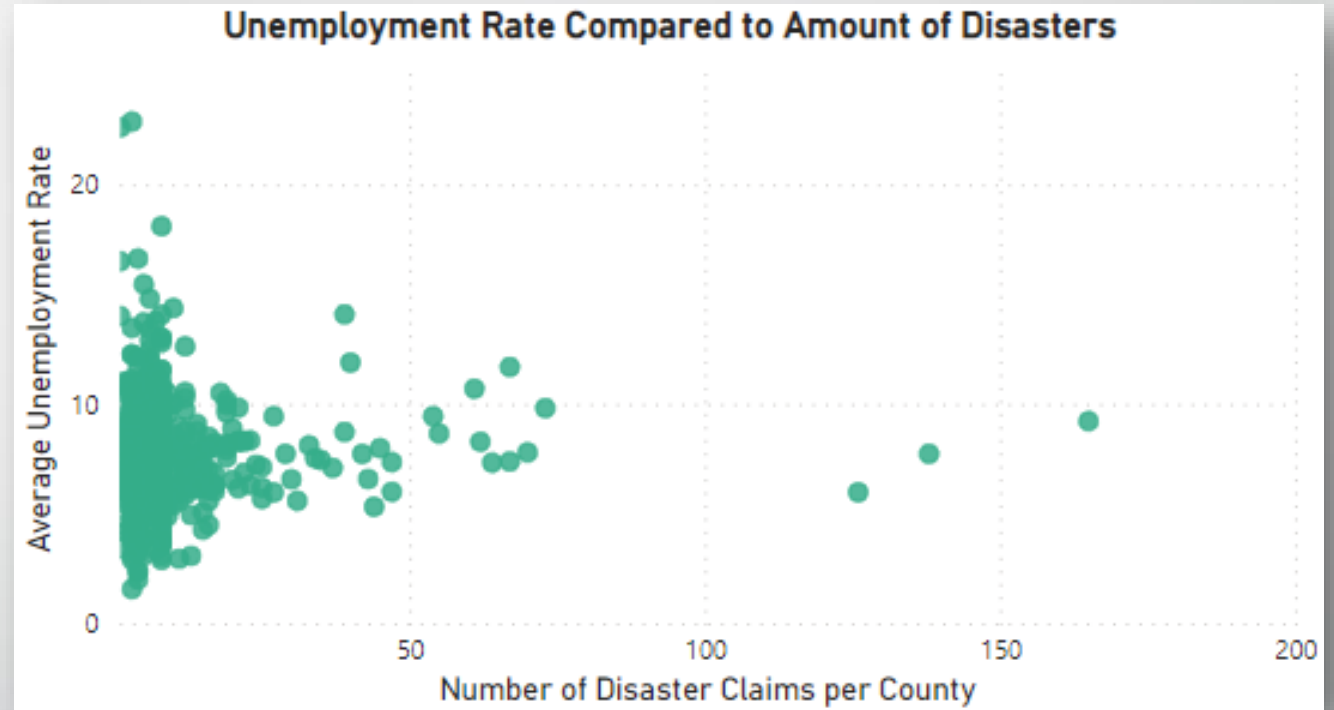
5. What states received the most disaster loans?

6. Is the S&P 500 a significant predictor for unemployment?

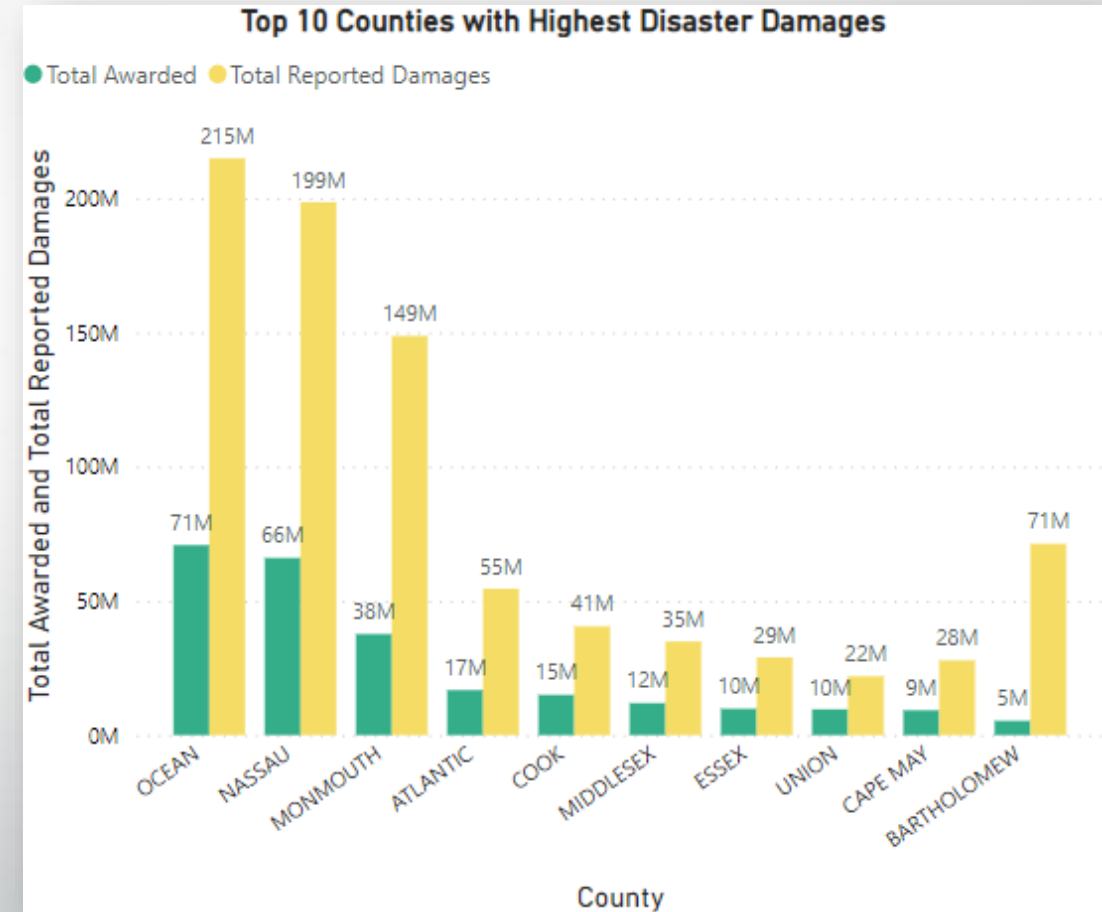
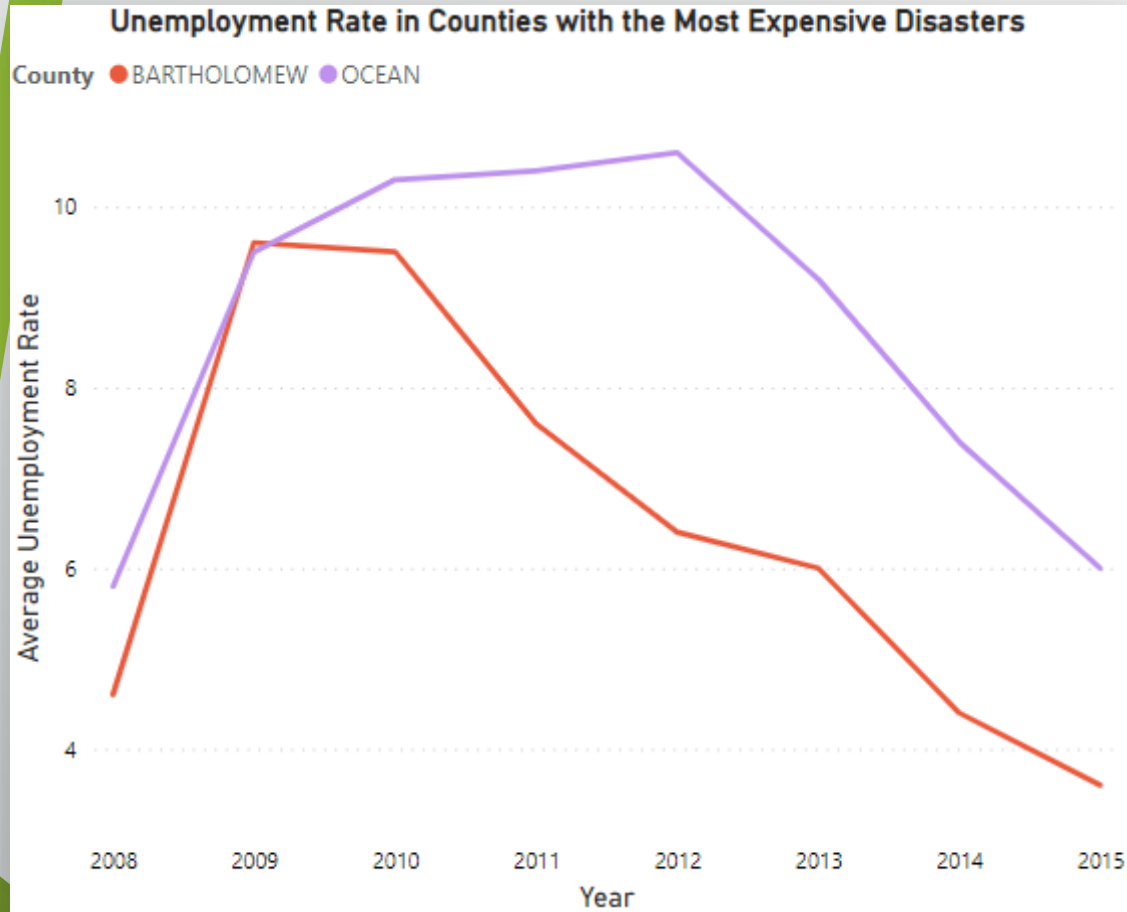


# Disasters as Predictors of Unemployment

- The average unemployment was aggregated by year
- Weak correlation between unemployment and quantity of disasters
- The 3 outliers with over 100 disaster claims are Nassau NY, Monmouth NJ, and Ocean NJ

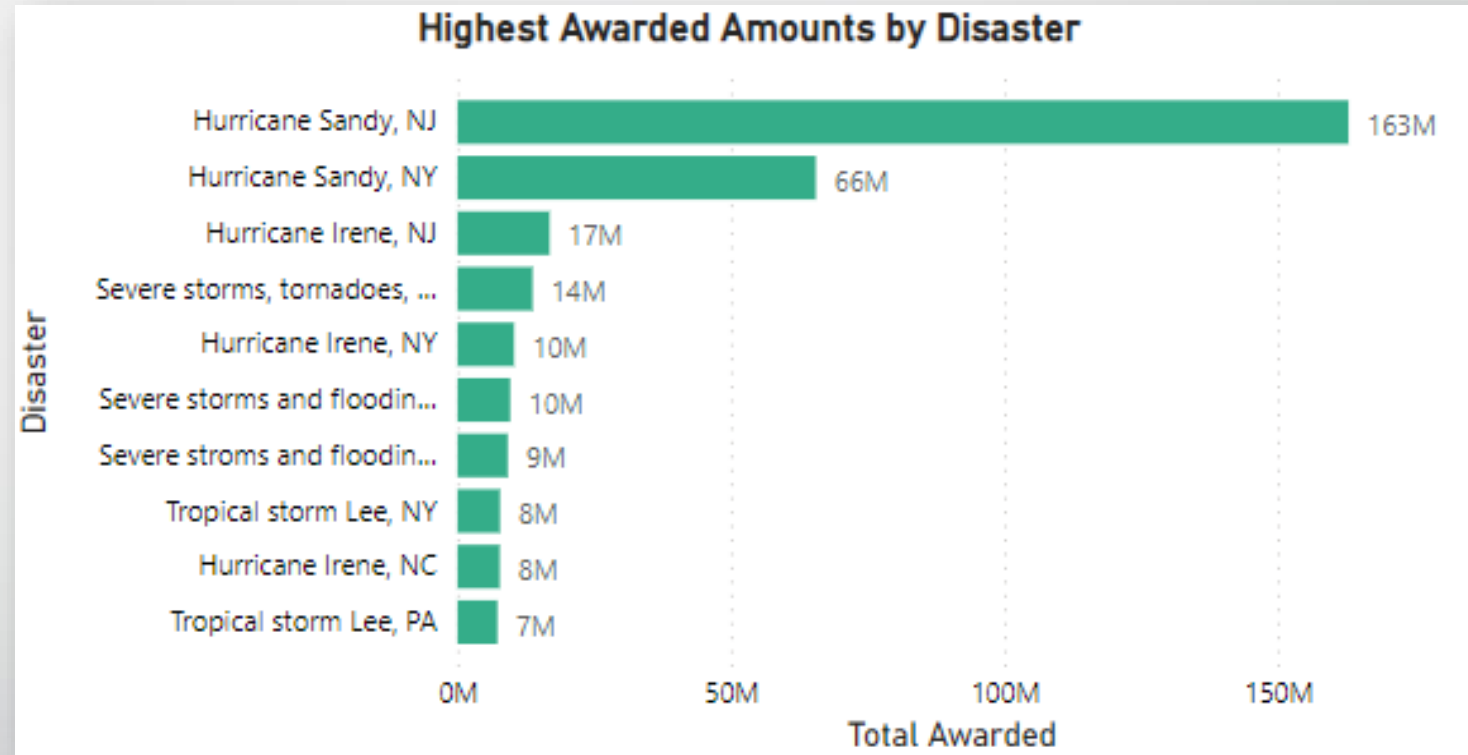


# Loan Amount as a Predictor



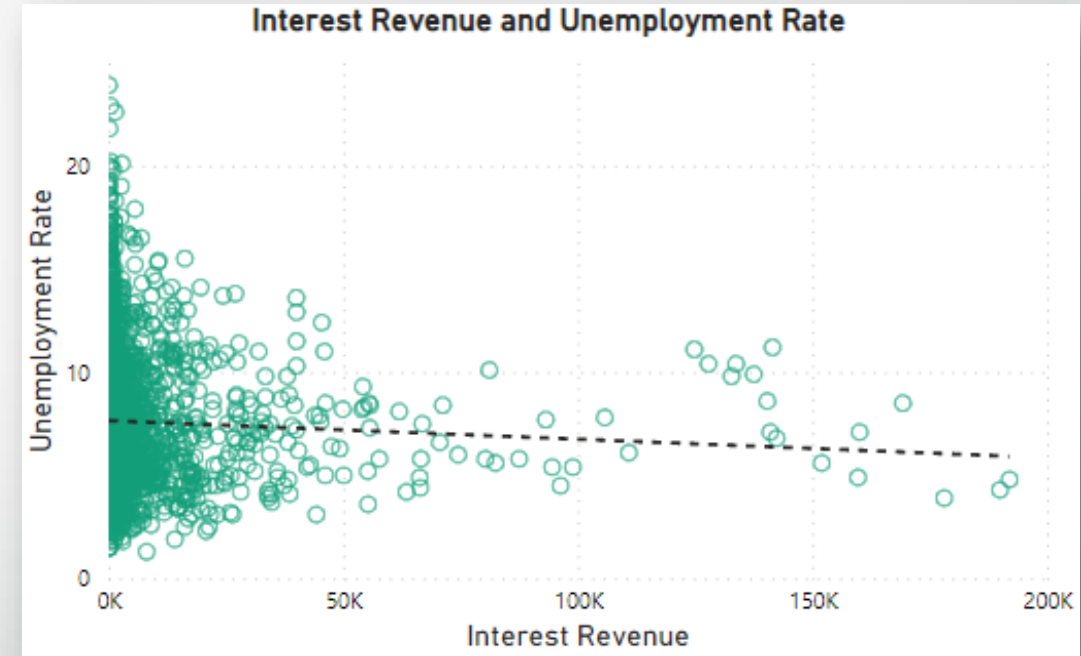
# Most Expensive Disasters

- Hurricane Sandy was the most expensive disaster by far
- The top 3 states hit by disasters loans were NJ, NY, and IL
- NJ received \$180,667,100
- NY received \$88,670,330
- IL received \$17,681,200



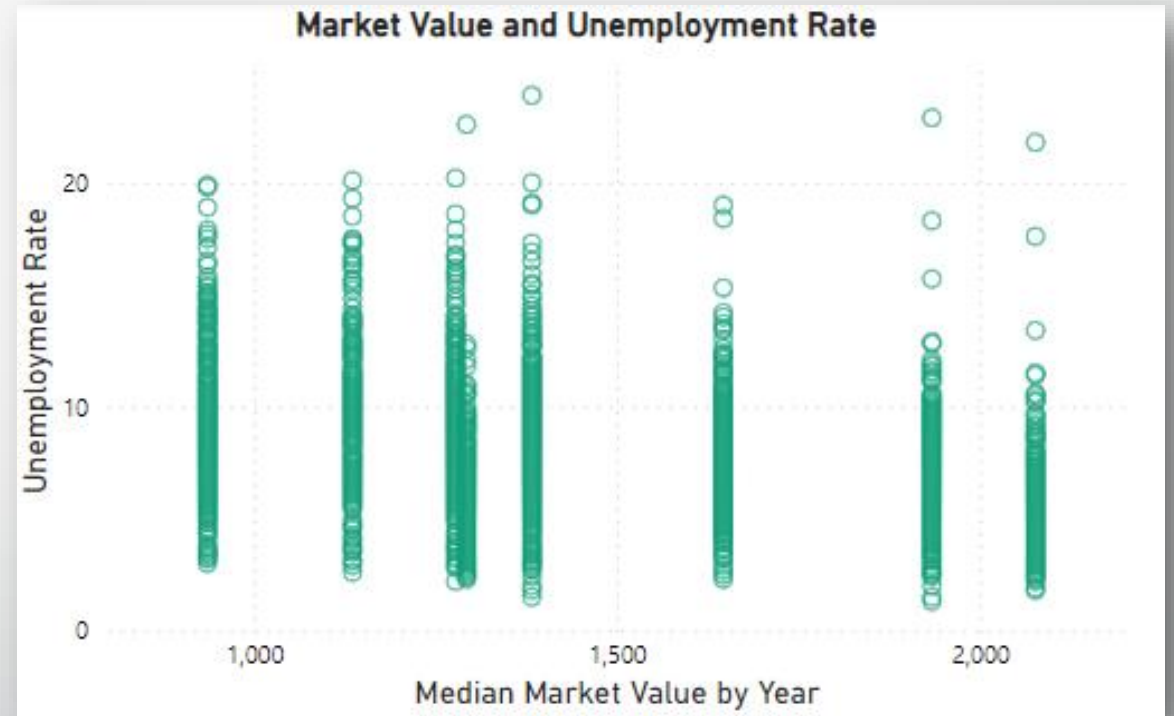
# Local Government Finances as a Predictor

- Fines & Forfeits was our second-best predictor
- Interest Revenue was our third-best predictor



# S&P 500 As a Predictor

- Market value was our strongest predictor
- Interestingly, it had a lower correlation with unemployment (-0.11) than local government finances

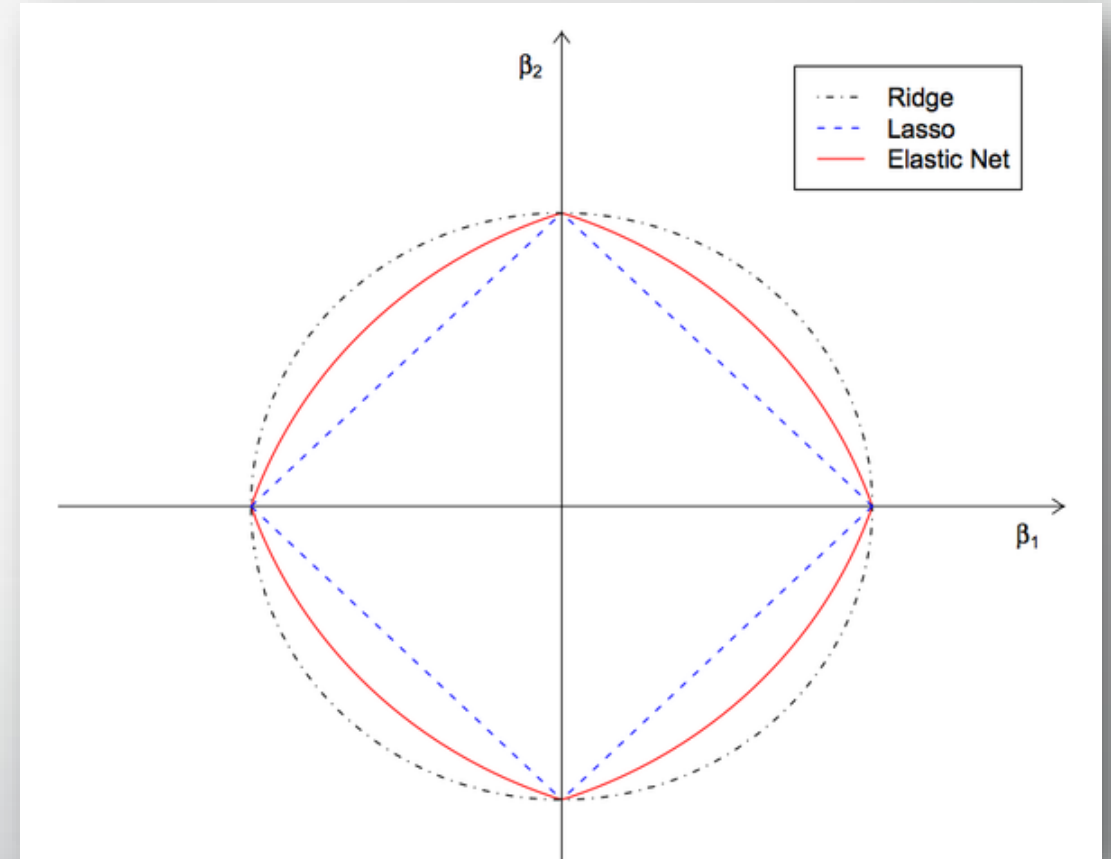


# Correlation Matrix

	Approved Amount Content	Approved Amount EIDL	Approved Amount Real Estate	Damaged Property Zip Code	Fines and Forfeits	Interest Revenue	Market Value	Property Tax	Total Salaries & Wages	Verified Loss Content	Verified Loss Real Estate	Year	Unemployment
Approved Amount Content	1.000000	0.416740	0.659065	-0.017202	-0.016806	-0.015025	0.039631	-0.009469	-0.014500	0.583050	0.649920	-0.041560	-0.048577
Approved Amount EIDL	0.416740	1.000000	0.359718	0.039691	-0.012043	-0.003591	-0.057778	-0.016077	-0.023734	0.223797	0.287088	-0.077939	-0.116359
Approved Amount Real Estate	0.659065	0.359718	1.000000	0.027232	-0.041942	-0.033125	0.040617	-0.033432	-0.040234	0.417810	0.633326	-0.066324	-0.066842
Damaged Property Zip Code	-0.017202	0.039691	0.027232	1.000000	0.136964	0.051875	-0.060827	0.047237	0.017746	0.007310	0.032540	-0.285273	-0.127382
Fines and Forfeits	-0.016806	-0.012043	-0.041942	0.136964	1.000000	0.659474	0.090892	0.723044	0.696462	-0.012186	-0.025699	-0.065824	-0.160826
Interest Revenue	-0.015025	-0.003591	-0.033125	0.051875	0.659474	1.000000	0.003518	0.589251	0.543271	-0.015774	-0.016960	-0.151293	-0.142499
Market Value	0.039631	-0.057778	0.040617	-0.060827	0.090892	0.003518	1.000000	0.092355	0.128464	0.025304	0.038948	0.628099	-0.113223
Property Tax	-0.009469	-0.016077	-0.033432	0.047237	0.723044	0.589251	0.092355	1.000000	0.954393	-0.019972	-0.022222	-0.004724	-0.045300
Total Salaries & Wages	-0.014500	-0.023734	-0.040234	0.017746	0.696462	0.543271	0.128464	0.954393	1.000000	-0.020373	-0.021845	0.036101	-0.034038
Verified Loss Content	0.583050	0.223797	0.417810	0.007310	-0.012186	-0.015774	0.025304	-0.019972	-0.020373	1.000000	0.923536	-0.061022	-0.068573
Verified Loss Real Estate	0.649920	0.287088	0.633326	0.032540	-0.025699	-0.016960	0.038948	-0.022222	-0.021845	0.923536	1.000000	-0.055770	-0.061441
Year	-0.041560	-0.077939	-0.066324	-0.285273	-0.065824	-0.151293	0.628099	-0.004724	0.036101	-0.061022	-0.055770	1.000000	0.364116
Unemployment	-0.048577	-0.116359	-0.066842	-0.127382	-0.160826	-0.142499	-0.113223	-0.045300	-0.034038	-0.068573	-0.061441	0.364116	1.000000

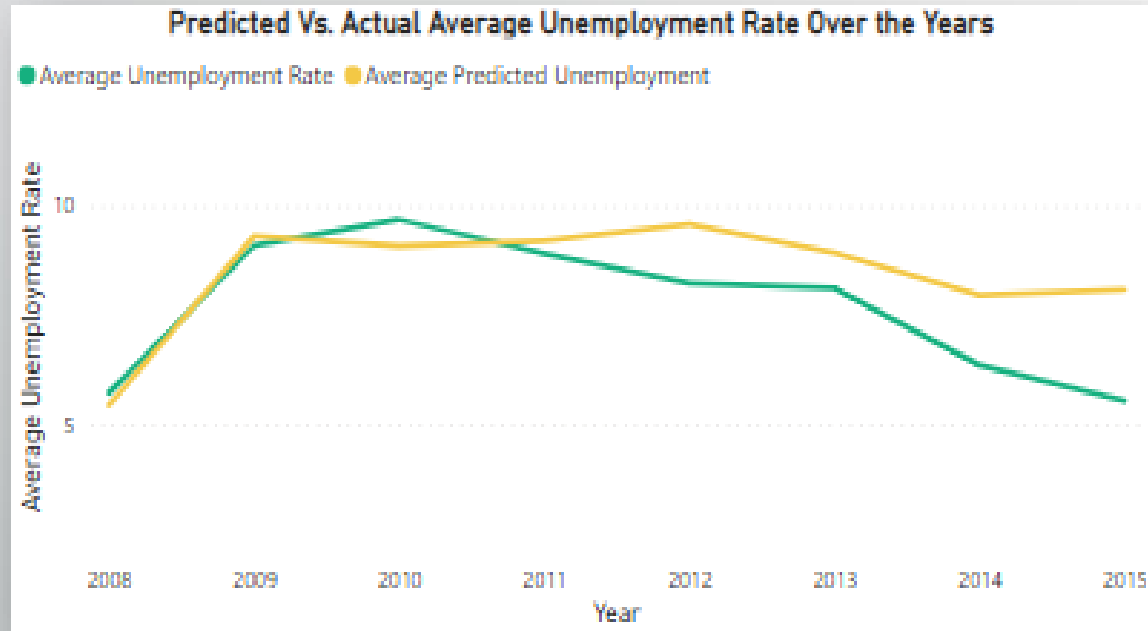
# Machine Learning Model

- After consideration, we decided to use Elastic Net for our unemployment prediction
- Elastic Net linear regression uses the penalties from both the lasso and ridge techniques to regularize regression models



# Model Results

Elastic Net(alpha=0.1,l1\_ratio=0.954766,max\_iter=10000)



- $R^2 = 0.403$
- $MSE = 1.989$
- Top 3 Numerical Attributes:
  1. S&P 500 Market Value
  2. Fines and Forfeits
  3. Interest Revenue
- Bottom 3 Numerical Attributes:
  1. Verified Loss Real estate
  2. Approved Amount Real estate
  3. Verified Loss Content





# Power BI Report

# Recommendations

- More attributes from different data sources should be considered as predictors for unemployment rate in the future
- Ultimately, we recognize that unemployment is very complex and believe that adding more features would increase the accuracy of unemployment rate predictions
- Consider including more market data as features because the S&P 500 was our best predictor
- Research the effect of disasters and disaster loans on future unemployment, they could have a delayed impact



Questions?

# Data Sources

- U.S. Small Business Administration. *Disaster Loan Data*. <https://data.sba.gov/dataset/disaster-loan-data>. 2008-2015.
- United States Census Bureau. *State & Local Government Finance Historical Datasets and Tables*. <https://www.census.gov/programs-surveys/gov-finances/data/datasets.html>.
- Ravaliya, Jay. *US Unemployment Rate by County, 1990-2016*. <https://www.kaggle.com/datasets/jayrav13/unemployment-by-county-us>
- Willden, Chase. *Sp500*. <https://data.world/chasewillden/stock-market-from-a-high-level>
- Corporate Finance Institute. *Elastic Net*. <https://corporatefinanceinstitute.com/resources/knowledge/other/elastic-net>