

# Problem and Goal

- Problem:
  - Bank Examiners must rely on off-site risk-modeling tools to monitor banks, particularly smaller institutions that do not have as frequent on-site examinations
- Goal:
  - Create a model that helps with early detection of financial distress and/or failure of a bank. The early detection system can help stem losses and save taxpayers from potentially costly government bailouts.
  - Utilize predictive analytics and further data analysis to detect individual bank weakness and vulnerability to failure.

# Brief Summary of Data Cleaning and Analysis

1. Split the dataset into training and testing.
2. Built a logistic regression model
3. Narrowed down number of variables to 15 from 33
  - a. Chosen by taking 3 training and testing sets on 1500 points
4. Utilized area under the curve to confirm accuracy

## Results:

1,500:

Average accuracy on training

86.79%

Average accuracy on testing

86.59%

Average false negative on training

14.67%

Average false negative on testing

15.56%

Average false positive on training

12.48%

Average false positive on testing

12.33%

15,000:

Average accuracy on training

86.47%

Average accuracy on testing

86.43%

Average false negative on training

23.35%

Average false negative on testing

22.51%

Average false positive on training

8.62%

Average false positive on testing

9.10%

# Market Adaptability and Insight

- Scalability of model for updates and iterative revisions
- Ease of model adoption for large array of regulatory bodies (FDIC, FRS, CSBS)
- Robust design for ease of manipulating variable importance for improved identification
- Simplicity focused with binary output for flagging forecasted failure