

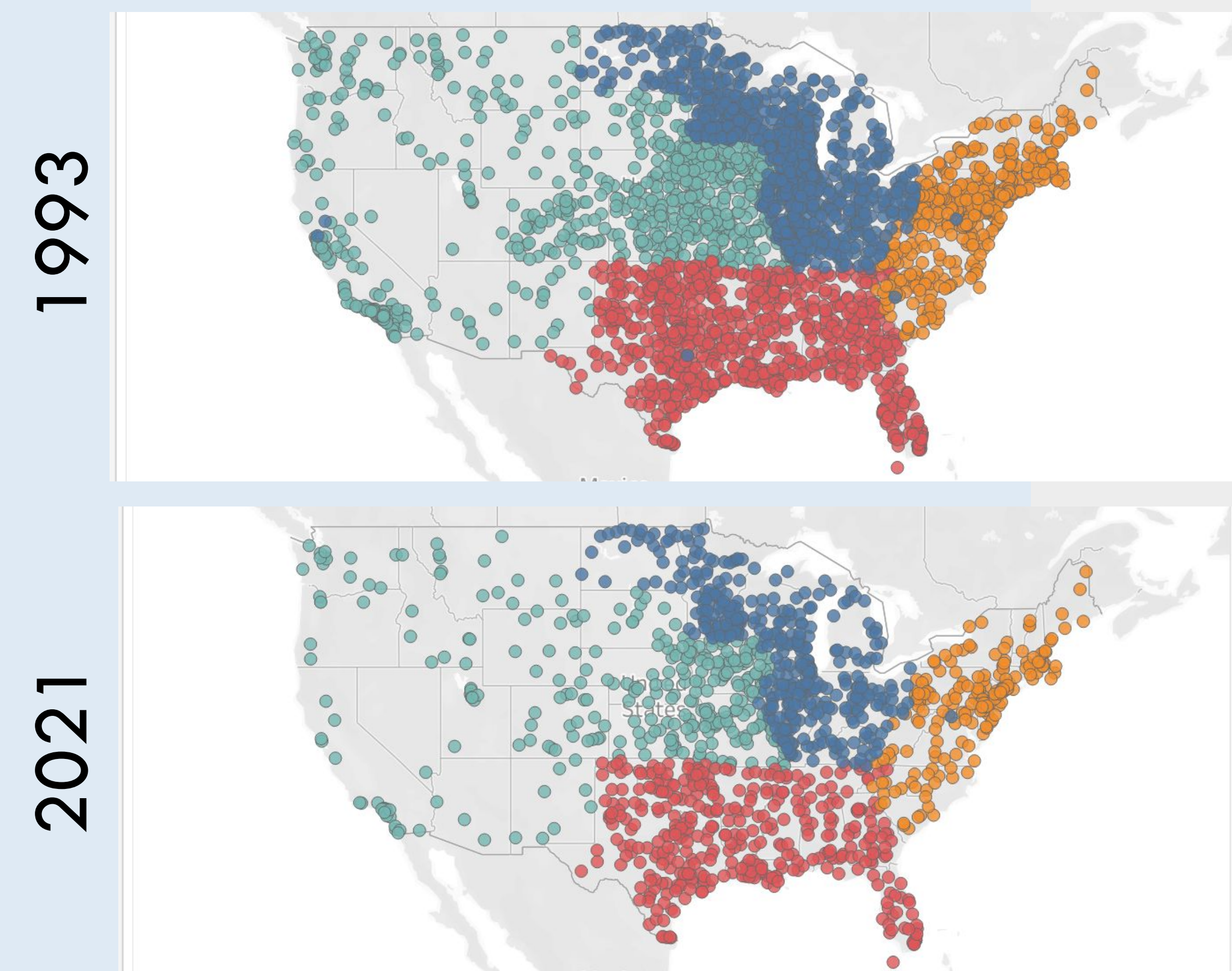
# (HOW EARLY) CAN WE PREDICT BANK FAILURE?

NAOMI CANNELL, ELISA IMPARA, AVNI MALHOTRA, ESIN ONAL, MINYOUNG RHO, DESSLAVA VELTCHEVA, XINYI XU

## Background

Large decline in the number of US banks in the past three decades due to

- mergers and acquisitions
- Failure of the banks → negative impact on economy



## Questions

- Can we predict bank failures?
- If so, how early can we predict failure?

## Methods

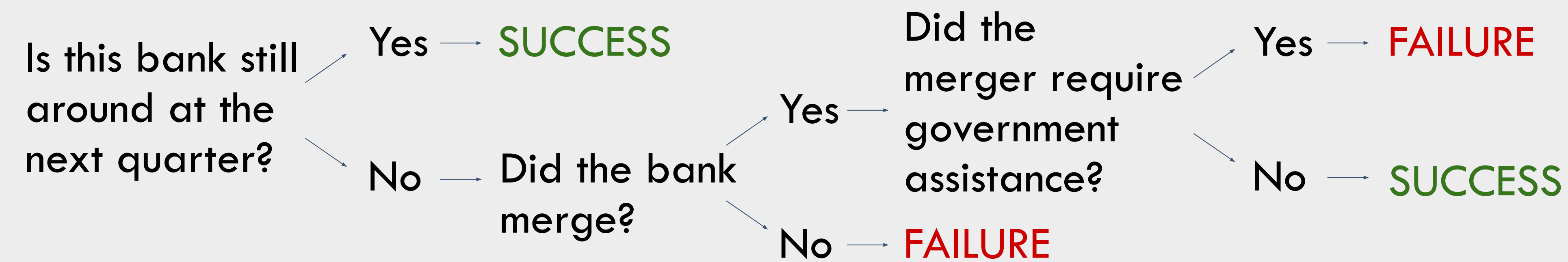
- Data Science (Machine Learning) tools for prediction
  - Random Forest
    - Flexibility to learn non-linear relationship
    - Highest accuracy rate at the time of default
  - Logistic Regression
  - Clustering (k-prototype)

→ We report the results from the **random forest** analysis

## Data

- Data from **ALL** US commercial banks from 1993-2015 on
  - Mergers and acquisitions (M&A) of banks across the US
  - **Quarterly** call reports on financial health status (assets, loans, etc.)

- **SUCCESS VS. FAILURE: y**



- Factors predicting failure of the bank: **X**
  - Only used ratios to avoid having to control for confounders (e.g., inflation)
  - For each bank at each quarter, included variables (ratios) up to previous 5Q

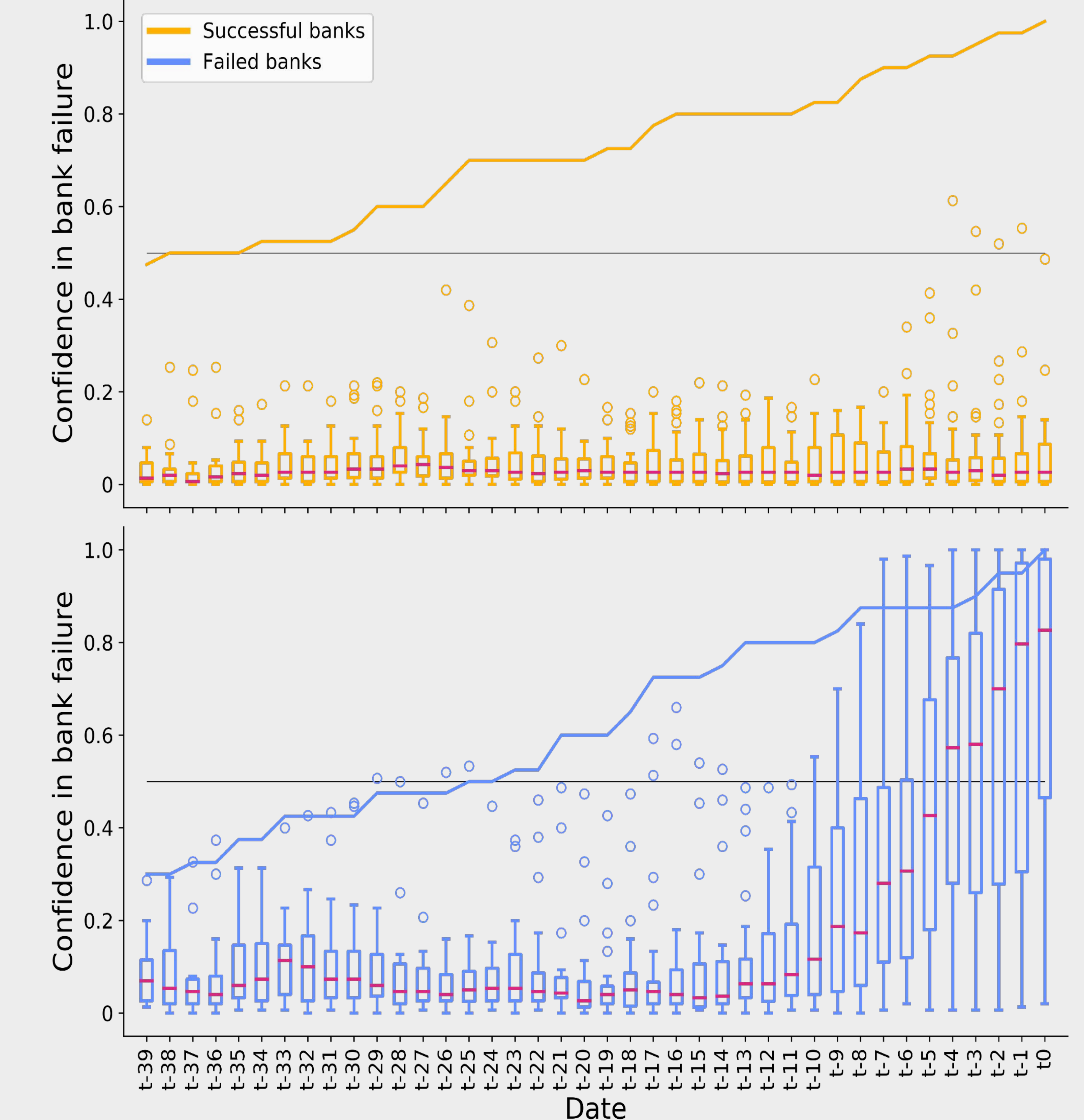
**NOTICE:** The unit of each data point is (a bank, at each quarter); each bank has multiple data points, which causes some challenges to our analysis

## Challenges and Solutions

Challenge	Solution
Success metrics not accurate for failed banks for all quarter leading up to the ultimate failure	Visualize the likelihood of failure at each time period
Traditional train-and-test split would randomize at (bank, quarter), but want to test at a bank level	Test: randomly selected 40 successful and 40 failed banks over their whole lifetime Train: rest of the data points
Imbalance on classifier	SMOTE oversampling method to produce balanced data

## Results and Conclusions

**Yes, we can predict bank failure  
At least 5Q (year+) in advance**



- Y-axis: predicted likelihood of bank failure
- X-axis
  - t0: time of failure for failed banks; most recent report for successful banks
  - tN: quarters prior to t0
- Line represents the sample size at each timepoint as a proportion of max sample size (40 banks)

## Future Work

Explain factors leading to bank failure, answer “why” banks fail

Data:  
USA bank merger and acquisition data from The Federal Reserve Bank of Chicago:  
<https://www.chicagofed.org/banking/financial-institution-reports/merger-data>  
Call Reports from Federal Financial Institutions Examination Council (FFIEC) from 1993 to 2019:  
<https://cdr.ffiec.gov/public/ManageFacsimiles.aspx>