

This is the summary of a project I did in my Econometrics class. I retrieved data from FDIC's website, which is a government website to insure the data integrity. Then I utilized SAS to input and analyzed the data to calculate frequencies and percentages, also, I used PROC GMAP to visualize the findings.

Where and When do Banks Fail?

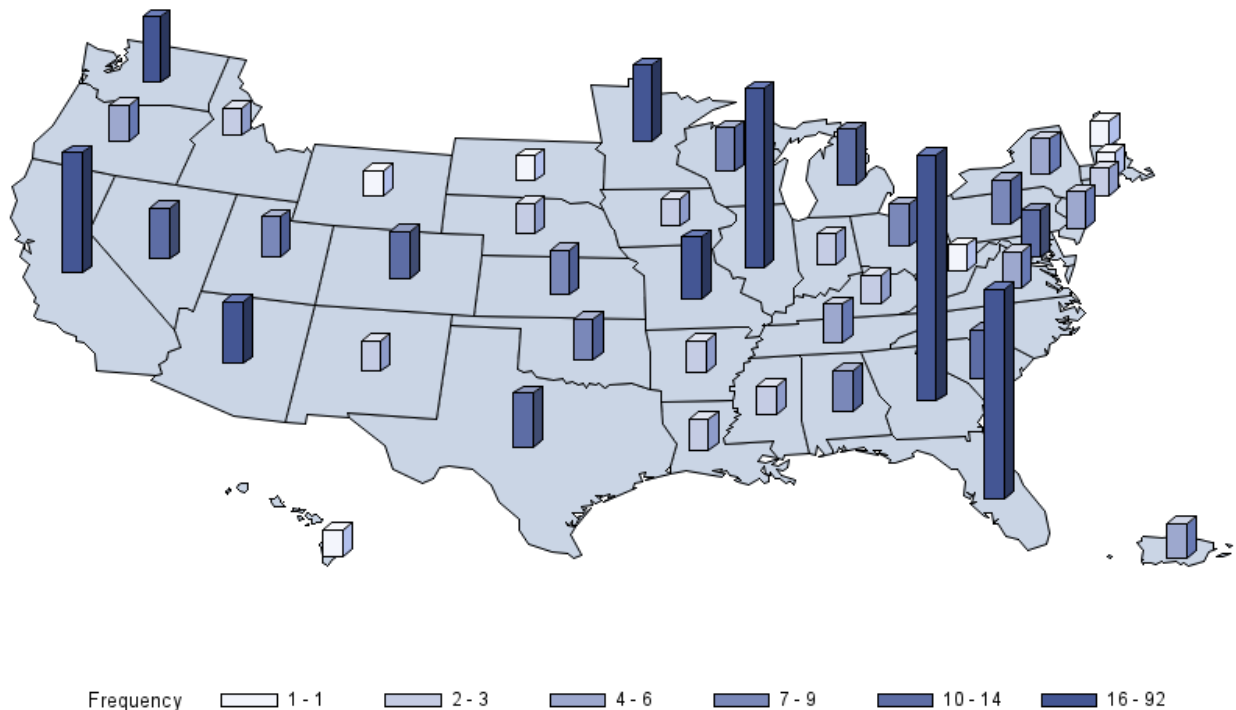
By Maya Yan Zhao

There are banks failing every year in every state in the United States. In this exercise, I would like to determine if some states have higher bank failure rate than the others, also, which year has the highest bank failure rate since 2000.

By using PROC GMAP in SAS, a map shows bank failure numbers for all the states is generated. As we can see in the map, Georgia, Florida and Illinois have the highest numbers of bank failures since 2000.

Data source: <https://www.fdic.gov/bank/individual/failed/banklist.html>

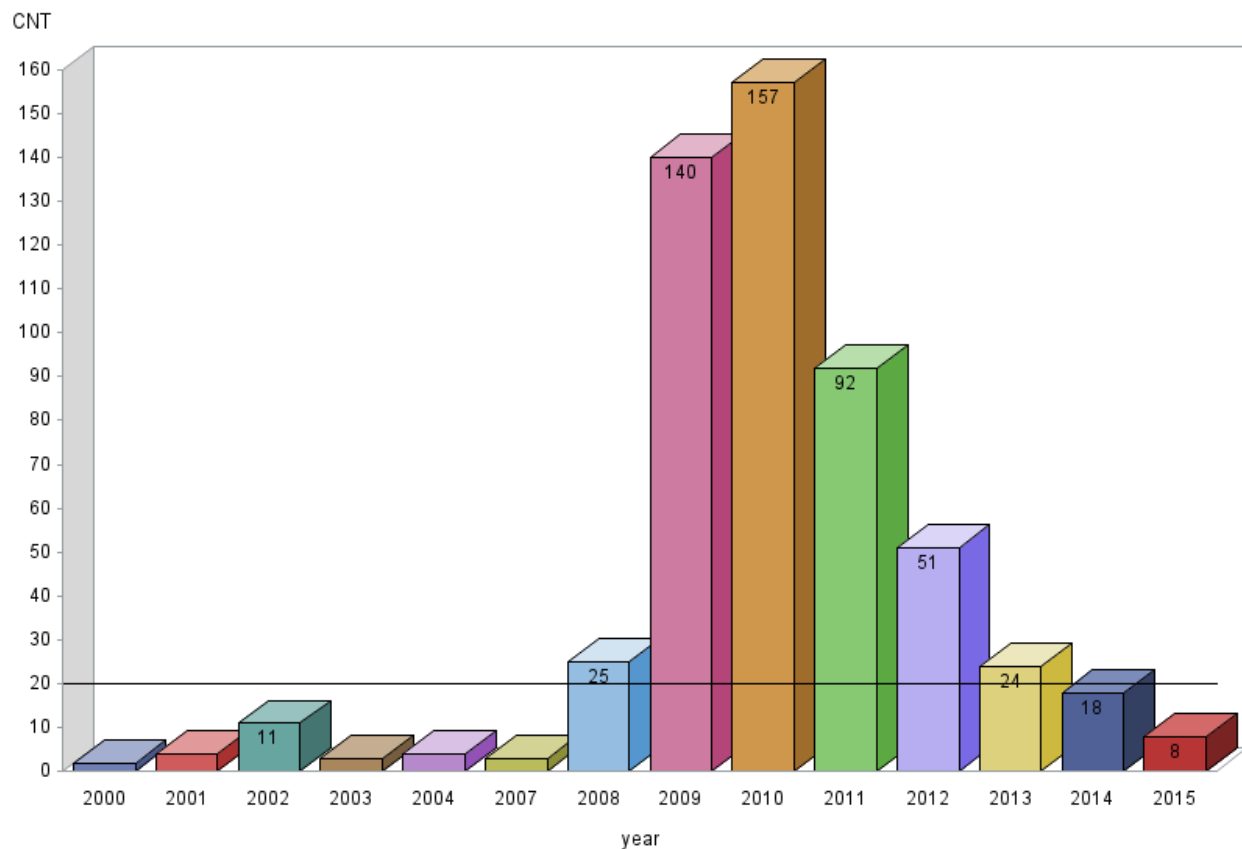
numbers of FDIC-insured failed banks across the States (2000-2015)



Georgia is the top one on the bank failing list with 93 failures from 2000 to 2015. This is not surprising. It has an outdated state law that favors small community banks over large multi-branch banks. Therefore, a large number of small banks found homes in Georgia, especially Metro Atlanta due to the constant population growth. However, some of these small banks just did not have the cash flow required to survive the financial crisis compared to their multi-branch competitors.

From the following bar chart, we can see that starting from 2008 to 2013, the average of bank failures is 20/year. In 2010, 157 banks failed and in 2009, 140 banks failed. This is not surprising either, since that was when the Subprime mortgage crisis took place.

Numbers of Bank Failures by Years (2000-2015)
With line of average failing number per year



SAS Code for graph 1

```
*i imported the banklist.csv first;
proc freq data=bank;
    tables st;
run;

/* Set the graphics environment */
goptions reset=all border cback=white htitle=13pt;

*I copied the data from the previous proc freq table;
data StFreq;
    input statecode $ Frequency Percent CumFreq CumPercen;
    datalines;
AL 7 1.29 7 1.29
AR 3 0.55 10 1.85
AZ 16 2.95 26 4.80
CA 41 7.56 67 12.36
CO 10 1.85 77 14.21
CT 2 0.37 79 14.58
FL 75 13.84 154 28.41
GA 92 16.97 246 45.39
HI 1 0.18 247 45.57
IA 2 0.37 249 45.94
ID 2 0.37 251 46.31
IL 66 12.18 317 58.49
IN 3 0.55 320 59.04
KS 9 1.66 329 60.70
KY 2 0.37 331 61.07
LA 3 0.55 334 61.62
MA 1 0.18 335 61.81
MD 10 1.85 345 63.65
MI 14 2.58 359 66.24
MN 23 4.24 382 70.48
MO 16 2.95 398 73.43
MS 2 0.37 400 73.80
NC 7 1.29 407 75.09
NE 3 0.55 410 75.65
NH 1 0.18 411 75.83
NJ 6 1.11 417 76.94
NM 3 0.55 420 77.49
NV 12 2.21 432 79.70
NY 5 0.92 437 80.63
OH 8 1.48 445 82.10
OK 7 1.29 452 83.39
OR 6 1.11 458 84.50
PA 9 1.66 467 86.16
PR 4 0.74 471 86.90
SC 10 1.85 481 88.75
SD 1 0.18 482 88.93
TN 6 1.11 488 90.04
TX 12 2.21 500 92.25
UT 7 1.29 507 93.54
VA 5 0.92 512 94.46
WA 19 3.51 531 97.97
WI 9 1.66 540 99.63
```

```

WV 1 0.18 541 99.82
WY 1 0.18 542 100.00
;
run;

title1 'numbers of FDIC-insured failed banks across the States (2000-2015)';

/* Display the block map */
proc gmap map=maps.us data=StFreq;
    id statecode;
    block Frequency / cblkout=black;
run;
quit;

```

SAS Code for graph 2

```

*I want to extract years from the closing_data variables;
data Year;
    set bank;
    Year = year (Closing_Date);
run;

PROC freq data=year;
    tables year;
run;

data YearFailed;
input year freq perc CumFreq CumPerc;
datalines;
2000 2 0.37 2 0.37
2001 4 0.74 6 1.11
2002 11 2.03 17 3.14
2003 3 0.55 20 3.69
2004 4 0.74 24 4.43
2007 3 0.55 27 4.98
2008 25 4.61 52 9.59
2009 140 25.83 192 35.42
2010 157 28.97 349 64.39
2011 92 16.97 441 81.37
2012 51 9.41 492 90.77
2013 24 4.43 516 95.20
2014 18 3.32 534 98.52
2015 8 1.48 542 100.00
;
run;

```

```

goptions reset=all border cback=white htitle=12pt;

title1 'Numbers of Bank Failures by Years (2000-2015)';
title2 'With line of average failing number per year';

axis1 minor=none label=('CNT') offset=(0,0);

/* Use the FRONTREF option to place reference lines */
/* in front of the bars */
proc gchart data=YearFailed;
    vbar3d YEAR / sumvar=Freq discrete raxis=axis1
                  subgroup=year nolegend
                  coutline=black width=8 space=3
                  vref=20 cref=black inside=sum frontref;
run;
quit;

```