# Q1 Descriptive Stats FOR 1986. I filtered the data first. THIS ISNOT FOR ALL OF THE DATA. ONLY 1986.

## Bin Pengs Way – Only gives this shit, we probably need more to describe data with more sophistication.

BANK\_ID ASSETS COST P\_LABOR\_w1

Min. : 1155 Min. : 9062 Min. : 183 Min. : 3.706

1st Qu.: 268556 1st Qu.: 68116 1st Qu.: 2488 1st Qu.: 8.556

Median : 509726 Median : 130616 Median : 5187 Median :13.903

Mean : 519105 Mean : 1263994 Mean : 60337 Mean :14.529

3rd Qu.: 785062 3rd Qu.: 303048 3rd Qu.: 12825 3rd Qu.:19.641

Max. :1017032 Max. :151228264 Max. :12129396 Max. :61.015

P\_FUNDS\_w2 P\_DEPOSI\_w3 CON\_LOAN\_y1 NONC\_LOA\_y2

Min. :0.0000447 Min. :0.007643 Min. : 80 Min. : 3020

1st Qu.:0.0031743 1st Qu.:0.023356 1st Qu.: 5668 1st Qu.: 25895

Median :0.0073762 Median :0.037168 Median : 13762 Median : 55638

Mean :0.0112286 Mean :0.039456 Mean : 95652 Mean : 653931

3rd Qu.:0.0146360 3rd Qu.:0.051394 3rd Qu.: 34453 3rd Qu.: 123170

Max. :0.2948620 Max. :0.335618 Max. :6048776 Max. :87124480

SECURITI\_y3 t POLICY

Min. : 2901 Min. :1.00 Min. :1.000

1st Qu.: 32007 1st Qu.:1.75 1st Qu.:2.000

Median : 57676 Median :2.50 Median :2.000

Mean : 516859 Mean :2.50 Mean :1.969

3rd Qu.: 131438 3rd Qu.:3.25 3rd Qu.:2.000

Max. :62132000 Max. :4.00 Max. :3.000

## My Way – Additional Stuff like st dev, skew kurtosis

**Check q1\_desc\_stats.csv**

item name

item number

number of valid cases

mean

standard deviation

trimmed mean ----- IGNORE THIS. NOT APPLICABLE.

median

mad: median absolute deviation (from the median)

minimum

maximum

skew

kurtosis

standard error

**Also, its only 2 marks so I’ve not made histograms with normal plot overlays. Would be overkill for this Question**

# Q2 how many banks are running under each policy regime respectively in 1986 at each quarter.

**Wrote a convenient loop to go over everything asked with ease. He is I think a bit of a dick for stating “**each policy regime respectively in 1986 at each quarter.” To confuse students. The number of banks operating all policies has been constant. Except 4 wasn’t introduced. Policy 4 started in 1994 – As per the PDF and confirmed by data by yours truly “Yashvir Surana”

The proportion of policy 1 at Quarter 1 is 0.2167382

The proportion of policy 2 at Quarter 1 is 0.6008584

The proportion of policy 3 at Quarter 1 is 0.1824034

The proportion of policy 4 at Quarter 1 is 0

The proportion of policy 1 at Quarter 2 is 0.2167382

The proportion of policy 2 at Quarter 2 is 0.5965665

The proportion of policy 3 at Quarter 2 is 0.1866953

The proportion of policy 4 at Quarter 2 is 0

The proportion of policy 1 at Quarter 3 is 0.2167382

The proportion of policy 2 at Quarter 3 is 0.5965665

The proportion of policy 3 at Quarter 3 is 0.1866953

The proportion of policy 4 at Quarter 3 is 0

The proportion of policy 1 at Quarter 4 is 0.2167382

The proportion of policy 2 at Quarter 4 is 0.5965665

The proportion of policy 3 at Quarter 4 is 0.1866953

The proportion of policy 4 at Quarter 4 is 0

# Q3 Regressions

3 Regressions (policy 4 does NOT exist for our given year 1986) – Each on grouped data by policy. Quarterly is irrelevant because we’re evaluating impact/significance whatever of each policy. Across time. Which is 1986.

HE’s A **dick** for wording a hint primarily to confuse students. “Note that you DEFINITELY have four quarters in 1986, and you MAYBE have four policies in each quarter? So how many regressions should you run in order to answer Q3?”

In my opinion that does not mean one regression for each policy and each quarter. That would give us 12 sets of results and its definitely pointless. Grouped across time makes sense.

I’ve modified the code a little to allow us to instantly run 12 of them if you guys think so. But here’s the 3 for each policy – across the entire year.

Comment on significance of variables, coefficients and what do they mean. No idea why R square is so high, possibly because banks are very similar in terms of RTS etc. COEFFICIENTS ARE IN THE EXACT SAME ORDER AS THE PDF. AND LOGGED AND EXACTLY HOW THE PDF SAYS.

Write about WHY LOG. Without log, model would be up, tehres exponsnetial scaling. I plotted them. GOOGLE REASONS WHY WE TAKE LOG of our variables before fitting a regression model. THEN seee the plots included. Possibly appendix them. HIGH R-square confirms with the fact that some of our variables are linear (after taking log and ratio) and fittinga line is easy, comment using significance .

ALSO NOTE X1, X2 arent just variables .. according to PDF they are w1/w3, etc.

Don’t forger

**Policy 1**

Residuals:

Min 1Q Median 3Q Max

-0.25789 -0.04996 -0.00658 0.04275 0.62105

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -0.133570 0.110500 -1.209 0.227

X1 0.215928 0.017848 12.098 <2e-16 \*\*\*

X2 0.001937 0.003736 0.518 0.605

X3 0.140183 0.005951 23.557 <2e-16 \*\*\*

X4 0.422289 0.009313 45.346 <2e-16 \*\*\*

X5 0.442634 0.008790 50.354 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.08171 on 398 degrees of freedom

Multiple R-squared: 0.993, Adjusted R-squared: 0.993

F-statistic: 1.136e+04 on 5 and 398 DF, p-value: < 2.2e-16

RTS is 0.9949209

**Policy 2**

Residuals:

Min 1Q Median 3Q Max

-0.37491 -0.05535 -0.01010 0.04141 0.51613

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.103953 0.071756 1.449 0.148

X1 0.229674 0.012489 18.390 < 2e-16 \*\*\*

X2 0.013701 0.003036 4.513 7.08e-06 \*\*\*

X3 0.138311 0.004266 32.424 < 2e-16 \*\*\*

X4 0.362404 0.006363 56.956 < 2e-16 \*\*\*

X5 0.476018 0.006417 74.185 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1005 on 1108 degrees of freedom

Multiple R-squared: 0.9946, Adjusted R-squared: 0.9945

F-statistic: 4.046e+04 on 5 and 1108 DF, p-value: < 2.2e-16

RTS is 1.023821

**Policy 3**

Residuals:

Min 1Q Median 3Q Max

-0.47981 -0.07077 -0.00680 0.06147 0.36856

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -0.363032 0.133754 -2.714 0.00698 \*\*

X1 0.390819 0.020335 19.219 < 2e-16 \*\*\*

X2 0.038818 0.006824 5.688 2.77e-08 \*\*\*

X3 0.242570 0.009207 26.346 < 2e-16 \*\*\*

X4 0.329429 0.012700 25.939 < 2e-16 \*\*\*

X5 0.379490 0.011476 33.067 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1161 on 340 degrees of freedom

Multiple R-squared: 0.9953, Adjusted R-squared: 0.9952

F-statistic: 1.439e+04 on 5 and 340 DF, p-value: < 2.2e-16

RTS is 1.050985