203 Group Project: CEO Salary Compared to Company Performance

Cheng Cheng, Jessica Economou, and Mona Iwamoto

February 2, 2017

Abstract

This paper summarizes an exploratory analysis of the relationship of CEO salary and company performance. The data was provided in (ceosal_w203.RData). In this analysis, we found no correlation between CEO salary and performance as measured by return on equity in the given sample. However, we did find a relationship in the more extreme ROE measures; and salaries tended to be slightly lower. Further, we explored the relationship of other variables to both salary and performance. Our study was limited, as we were missing key variables and we were unable to claim that this was a representation of the current population.

Introduction

This analysis is motivated by the following research question:

Is company performance is related to CEO salary?

Disproportionately high executive compensation is very topical. On average, it has been shown that CEOs make 300 times what their workers. Additionally, compensation routinely includes bonuses and other performance based incentives - even when targets are not met. (Pyke, 2014) Worse, many of the highest-paid CEOs in the past two decades have ended their careers as failures, bailed out by taxpayers, or indicted for criminal behavior. (Pyke, 2014)

Are the high salaries justified? In fact, a recent study suggests a negative correlation between CEO salary and financial performance. A group at the University Utah, performed a study on broad cross-section of Execucomp S&P1500 firms over the 1994-2015 period and found that a negative relation holds in the cross-section of firms over this time period(Cooper, 2016).

We addressed this question using exploratory data analysis techniques with CEO salary data from 1990. The data was provided to us with the following codebook:

1. salary | 1990 compensation, $1000s
2. age | in years
3. college | =1 if attended college
4. grad | =1 if attended graduate school
5. comten | years with company
6. ceoten | years as ceo with company
7. profits | 1990 profits, millions
8. mktval | market value, end 1990, mills.

An overview of the raw data revealed that several factors were missing. Thus, this analysis proved difficult with the available dataset. Notably, compensation was incomplete, as salary is only one component of executive compensation. This sample set also did not consider year over year growth, an important factor of performance. A company in the dataset could have generated $50 million in profit in 1990, but if they did $70 million the previous year, this could be an indicator of poor performance. Company size was not measured either- apart from market share, and industry sector was not specified. Comparing a company like Google's profits or market value with a company that simply produces milk is not a good measure of

the company's relative performance. So, this limited our evaluation of performance to return on equity, and comparisons could be made only broadly.

For this exploratory analysis, we defined performance as a return on equity (profits / market value * 100) within the dataset.

We also considered secondary relationships:

- Did any patterns emerge where CEOs at certain salary thresholds perform better?
- Was there a correlation between tenure and performance?

Setup

```
load("ceo_w203.RData")
```

Data Selection

Initial inspection of the data showed that we have 185 observations and 8 variables.

```
nrow(CEO)
```

```
## [1] 185
```

```
str(CEO)
```

```
## 'data.frame':    185 obs. of  8 variables:
##  $ salary : num  1033 879 971 567 1336 ...
##  $ age    : num  62 63 72 56 60 59 46 59 51 56 ...
##  $ college: num  1 1 1 1 1 1 1 1 1 1 ...
##  $ grad   : num  1 1 1 0 1 1 1 1 0 1 ...
##  $ comten : num  30 21 33 31 21 2 7 3 8 9 ...
##  $ ceoten : num  1 9 24 10 13 2 3 3 8 3 ...
##  $ profits: num  478 212 69 65 562 401 44 257 13 34 ...
##  $ mktval : num  7300 4900 609 1700 4300 10700 533 3900 458 6700 ...
```

Our first examination considered the valid ranges of the numeric values. The fields - age, conten, ceoten and mktval - must have only non-negative values.

Age:

```
summary(CEO$age)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   21.00   51.00   57.00   55.78   61.00   86.00
```

Years at Company:

```
summary(CEO$comten)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    2.00    9.00   21.00   21.66   33.00   58.00
```

Years as CEO:

```
summary(CEO$ceoten)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   3.000   5.000   7.681  11.000  37.000
```

Market Value:

```
summary(CEO$mktval)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      -1     567    1200    3450    3200   45400
```

Data Trimming

We determined that any values related to time must must be non-negative. Additionally, market value, as defined as the product of the market price of the company's shares and the total number of outstanding shares could not be negative. We assumed that these were errors in data collection, as the values for these were -1 every time. For valid analysis, we omitted records of companies with a market value less than 0.

```
CEO_trimmed = subset(CEO, mktval >= 0)
nrow(CEO_trimmed)
```

```
## [1] 180
```

After eliminating invalid rows 180 observations remained. We examined a summary.

```
summary(CEO_trimmed)
```

```
##      salary            age           college            grad
##  Min.   : 100.0   Min.   :21.00   Min.   :0.0000   Min.   :0.0000
##  1st Qu.: 469.2   1st Qu.:51.00   1st Qu.:1.0000   1st Qu.:0.0000
##  Median : 705.5   Median :57.00   Median :1.0000   Median :1.0000
##  Mean   : 861.3   Mean   :55.96   Mean   :0.9611   Mean   :0.5389
##  3rd Qu.:1110.0   3rd Qu.:62.00   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :5299.0   Max.   :86.00   Max.   :1.0000   Max.   :1.0000
##      comten          ceoten          profits            mktval
##  Min.   : 2.00   Min.   : 0.000   Min.   :-463.00   Min.   :  303.0
##  1st Qu.:10.75   1st Qu.: 3.000   1st Qu.:  33.75   1st Qu.:  606.5
##  Median :22.00   Median : 6.000   Median :  61.50   Median : 1200.0
##  Mean   :22.18   Mean   : 7.861   Mean   : 204.72   Mean   : 3545.8
##  3rd Qu.:33.00   3rd Qu.:11.000   3rd Qu.: 206.50   3rd Qu.: 3275.0
##  Max.   :58.00   Max.   :37.000   Max.   :2700.00   Max.   :45400.0
```

Corporate performance Metrics - Profit, Market Share or Return on Equity?

To answer our initial question as to the relationship between CEO salary and performance, it was necessary to define how to measure performance.

Wall Street investment firms and executives tend to focus on return on equity as their primary measure of company performance. (Hagill, 2010) However, there are many other metrics often used such as return on assets or year over year growth to measure financial performance. With this limited dataset, we lacked that information to compute these other key indicators.

Because this cross-sectional dataset does not include year over year changes in profits or market value, in order to evaluate the company's perfomance used return on equity (ROE) as a performance metric, as we defined earlier:

```
CEO_trimmed$roe = (CEO_trimmed$profits / CEO_trimmed$mktval * 100)
head(CEO_trimmed)
```

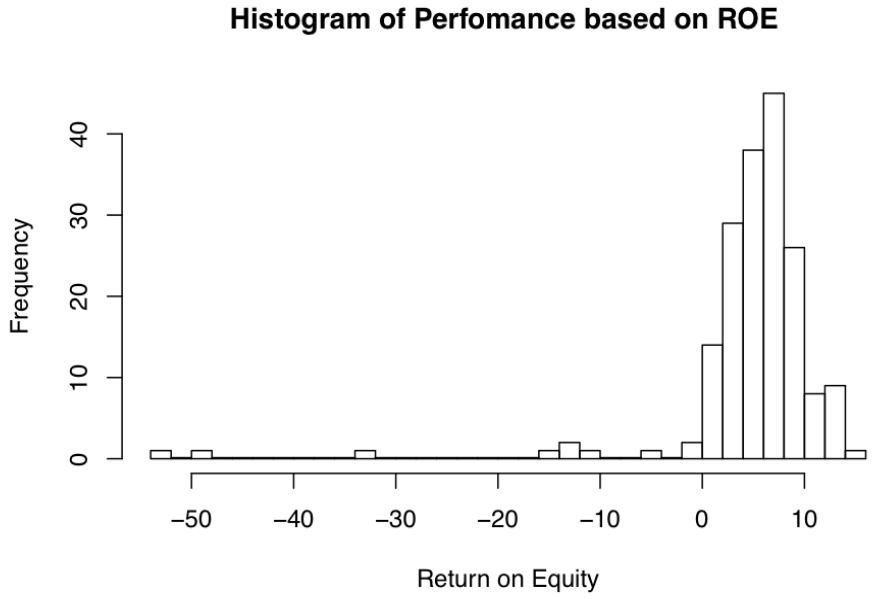
```
##     salary age college grad comten ceoten profits mktval       roe
## 154   1033  62       1    1     30      1     478   7300  6.547945
## 79     879  63       1    1     21      9     212   4900  4.326531
## 19     971  72       1    1     33     24      69    609 11.330049
## 115    567  56       1    0     31     10      65   1700  3.823529
## 36    1336  60       1    1     21     13     562   4300 13.069767
## 153   1444  59       1    1      2      2     401  10700  3.747664
```

```
summary(CEO_trimmed$roe)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -53.610   3.831   5.929   4.903   7.956  15.740
```

We noted that the return on equity ranges from -53.6107711 to 15.7435897, The histogram below shows the distribution of performance based on ROE.

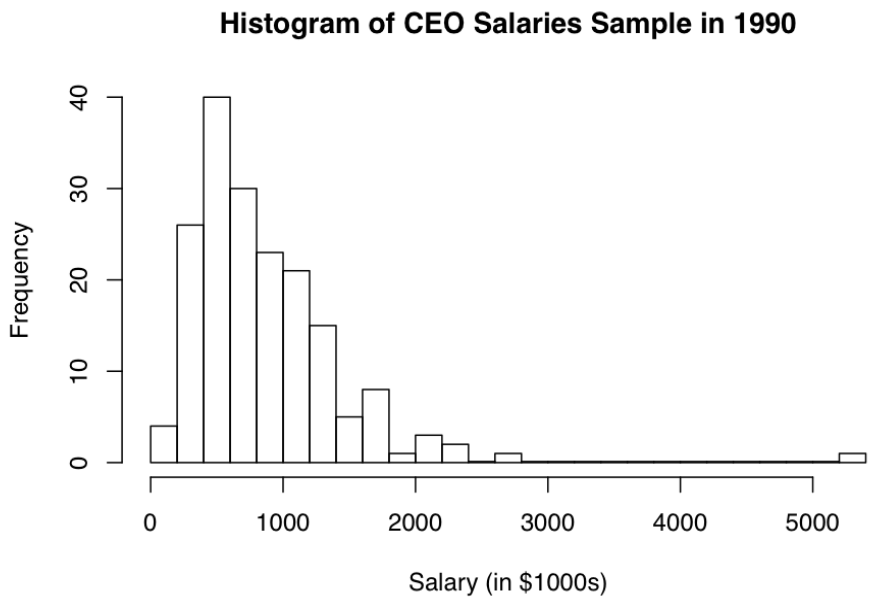
```
hist(CEO_trimmed$roe,
     main = "Histogram of Perfomance based on ROE ",
     xlab = "Return on Equity", breaks = 30)
```



Salary transformation

Using the transformed data, we delved further into the key independent variable in the dataset, CEO salary. We set up a histogram that would allow us to further visualize how these were distributed:

```
hist(CEO_trimmed$salary,
     main = "Histogram of CEO Salaries Sample in 1990",
     xlab = "Salary (in $1000s)", breaks = 20)
```

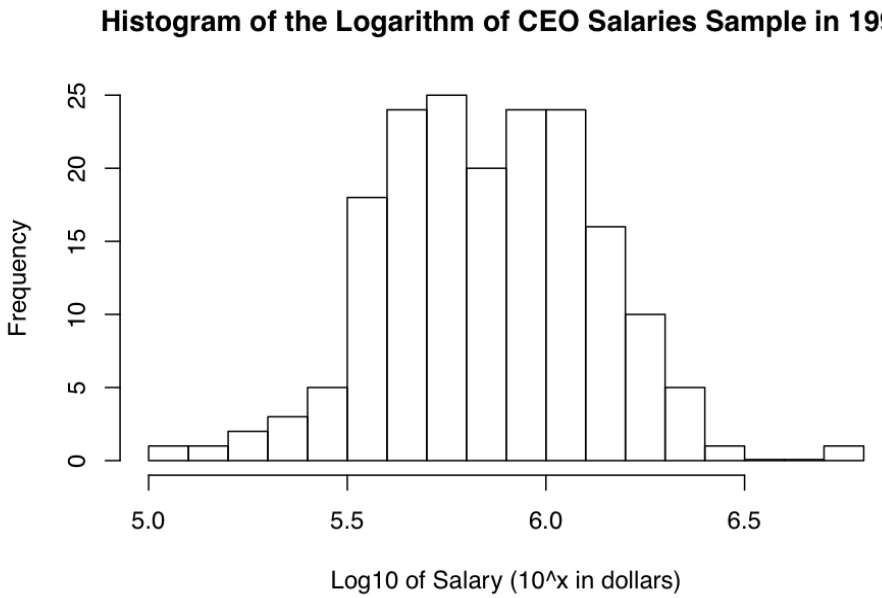



```
summary(CEO_trimmed$salary)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   100.0   469.2   705.5   861.3  1110.0  5299.0
```

These salaries ranged from $100k to $5.2 million, with the maximum salary being an outlier. No null values existed, and there was a clear positive skew. Because of the strong positive skew in the histogram, we considered computing the salary on a logarithmic scale. This would allow us to further visualize and model out marginal changes in where the bulk of the salaries were distributed. As shown by the histogram below, the computed sal_log field is the log base 10 of the salary and appears to approximate a log-normal distribution. This facilitated observation of possible relationships more clearly.

```
CEO_trimmed$sal_log = (log(CEO_trimmed$salary * 1000,10))
hist(CEO_trimmed$sal_log,
     main = "Histogram of the Logarithm of CEO Salaries Sample in 1990",
     xlab = "Log10 of Salary (10^x in dollars)", breaks = 20)
```



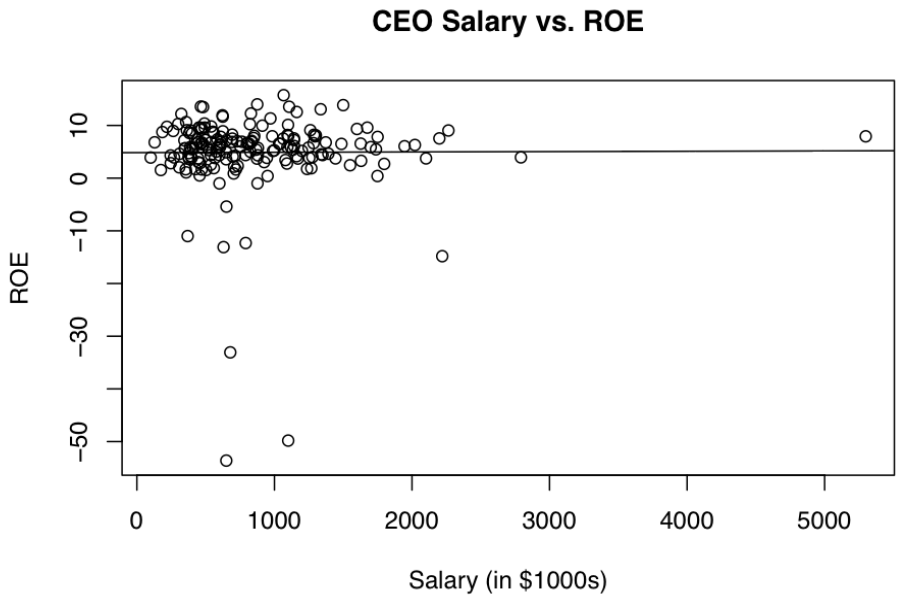

```
summary(CEO_trimmed$sal_log)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   5.000   5.671   5.848   5.857   6.045   6.724
```

Exploration of Primary Relationships

Plotting the relationship between the non-logarithmic salary compared to the return on equity (ROE) showed the necessity of looking at the salary logarithmically. The salary data points are very clustered together because most of the CEOs in this dataset have a similar range in income. Without seeing the line of best fit, it's difficult to make any insightful observations about their relationship. Here we have an example of the two:

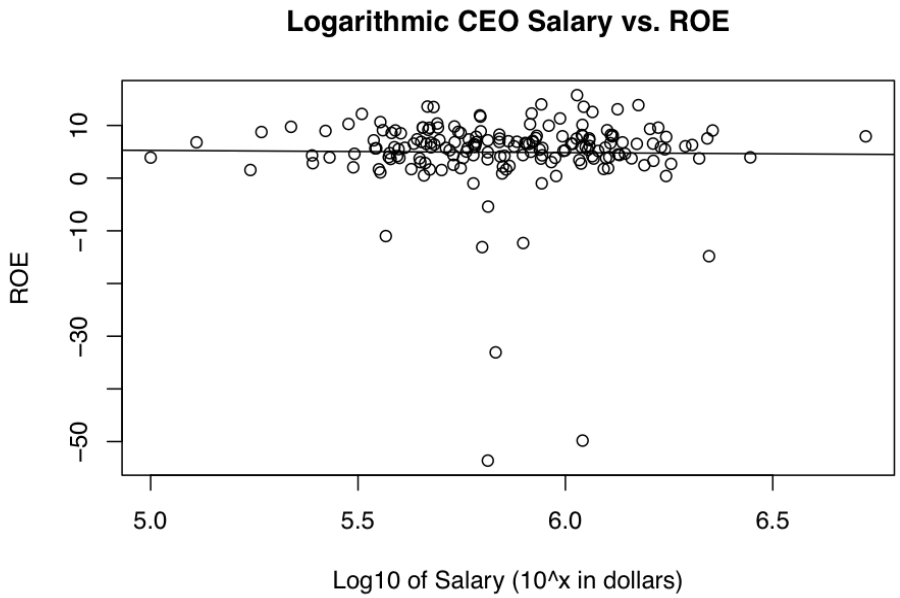
```
plot(CEO_trimmed$salary, CEO_trimmed$roe,
     main = "CEO Salary vs. ROE",
     xlab = "Salary (in $1000s)",
     ylab = "ROE")
abline(lm(CEO_trimmed$roe~CEO_trimmed$salary))
```




```
cor(CEO_trimmed$roe,CEO_trimmed$salary)
```

```
## [1] 0.005186691
```

```
plot(CEO_trimmed$sal_log, CEO_trimmed$roe,
     main = "Logarithmic CEO Salary vs. ROE",
     xlab = "Log10 of Salary (10^x in dollars)",
     ylab = "ROE")
abline(lm(CEO_trimmed$roe~CEO_trimmed$sal_log))
```




```
cor(CEO_trimmed$roe,CEO_trimmed$sal_log)
```

```
## [1] -0.01437027
```

CEO salary and return on equity did not appear to be correlated in this sample of data.

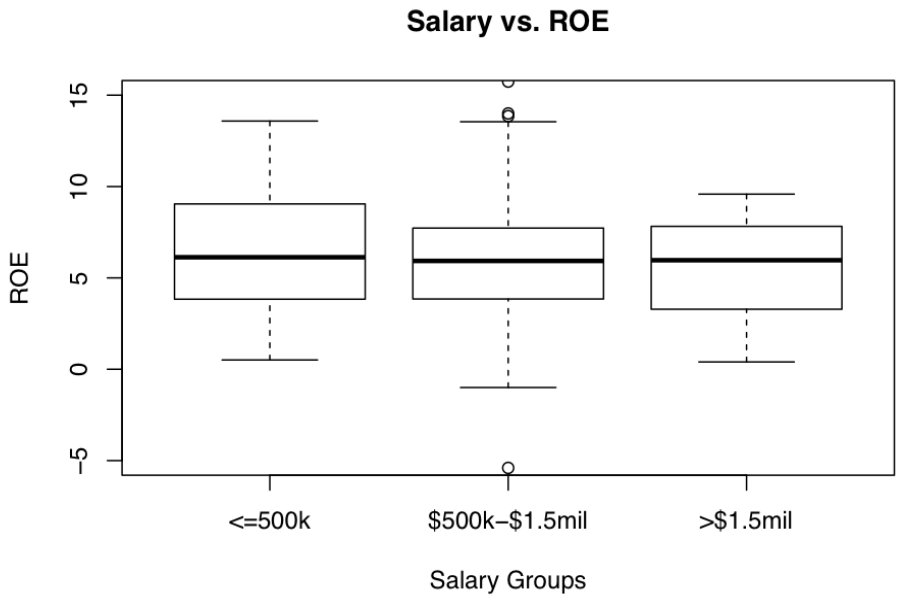
Distributions

Taking the logarithm of salaries was sufficient for visualizing the data points and their relationships, but we were also interested in quantifying salaries into meaningful levels. To further analyze salary compared to ROE, we decided it would be helpful to break out the CEO salaries into separate groups, sorted by lowest to highest salary. We expected that a CEO making under $500,000 in 1990 could have different behavior than a CEO making over $1.5mil, and grouping these up would allow us to see their differences if they did exist. A box plot of these three show similar means (note: some data points are missing because they are too high or low for the graph):

```
CEO_trimmed$sal_groups<-cut(CEO_trimmed$salary, breaks = c(0,500,1500,Inf),
                            labels = c("<=500k", "$500k-$1.5mil", ">$1.5mil"))
summary(CEO_trimmed$sal_groups)
```

```
##        <=500k $500k-$1.5mil      >$1.5mil
##            54           108            18
```

```
plot(CEO_trimmed$sal_groups, CEO_trimmed$roe,
     main = "Salary vs. ROE",
     xlab = "Salary Groups",
     ylab = "ROE",
     ylim = c(-5,15))
```

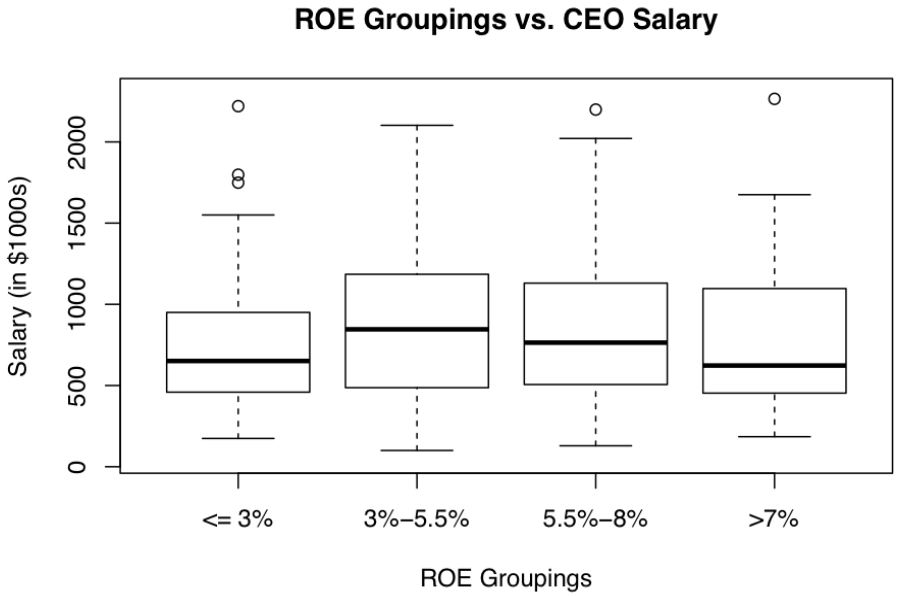


However, their distributions were visibly different, especially for the salaries over $1.5mil. We dove into this further by analyzing the ROEs as low, medium-low, medium-high, and high levels, and then compared them to salary (note: the box plot omits the one data point where the salary was $5.2 to show a clearer picture of the bulk of the data set):

```
CEO_trimmed$roe_groups<-cut(CEO_trimmed$roe, breaks = c(-100,3,5.5,8,Inf),
                            labels = c("<= 3%", "3%-5.5%", "5.5%-8%", ">7%"))
aggregate(CEO_trimmed$salary~CEO_trimmed$roe_groups, FUN = median)
```

```
##   CEO_trimmed$roe_groups CEO_trimmed$salary
## 1                 <= 3%                651
## 2               3%-5.5%                846
## 3               5.5%-8%                764
## 4                   >7%                623
```

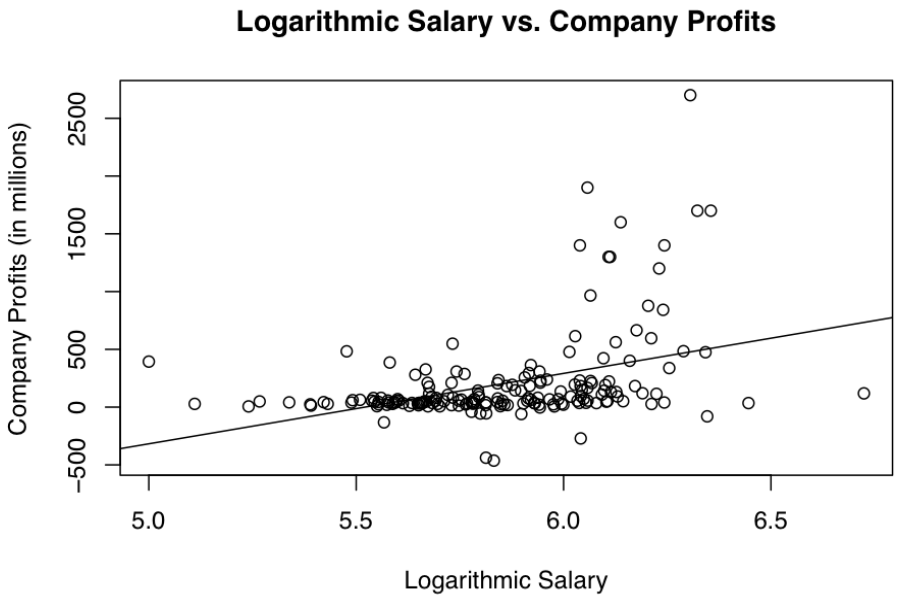
```
plot(CEO_trimmed$roe_groups,CEO_trimmed$salary,
     ylim = c(50,2300),
     main = "ROE Groupings vs. CEO Salary",
     ylab = "Salary (in $1000s)",
     xlab = "ROE Groupings")
```



When ROE was either below 3% or above 7% in this sample, the more extreme values in ROE, the values tended to be lower.

We also considered the relationship between profits and CEO salary, but decided a confounding factor would be too much at play here. The reason this is moderately positively correlated could be because of company size. Larger companies tend to make more in profit dollars to begin with, and would therefore have more money to pay out their CEOs. Below is a graph showing their relationship, which shows that most companies making over $500,000,000 in profits have higher CEO salaries than the rest of the group, skewing the correlation:

```
plot(CEO_trimmed$sal_log, CEO_trimmed$profits,
     main = "Logarithmic Salary vs. Company Profits",
     xlab = "Logarithmic Salary",
     ylab = "Company Profits (in millions)")
abline(lm(CEO_trimmed$profits~CEO_trimmed$sal_log))
```

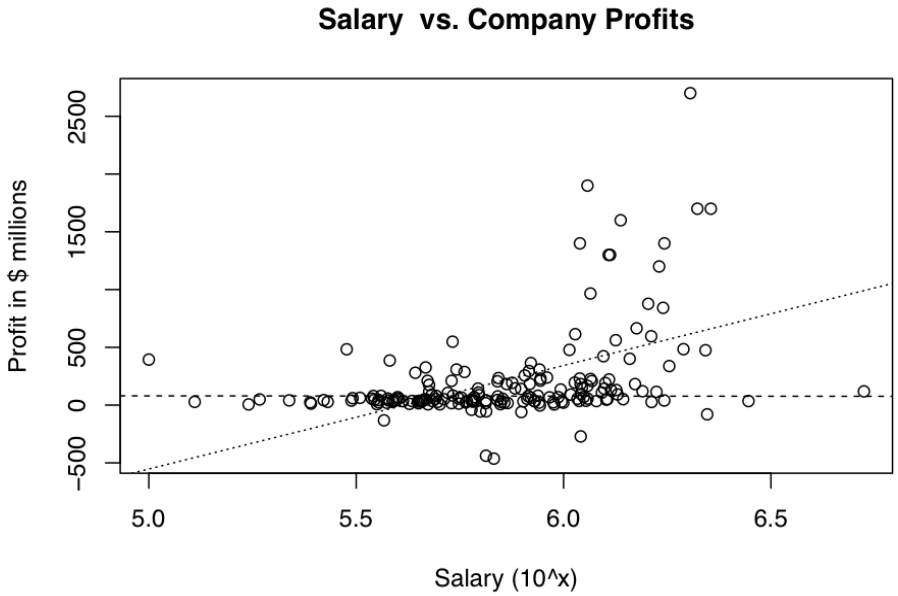



```
cor(CEO_trimmed$profits,CEO_trimmed$sal_log)
```

```
## [1] 0.3973178
```

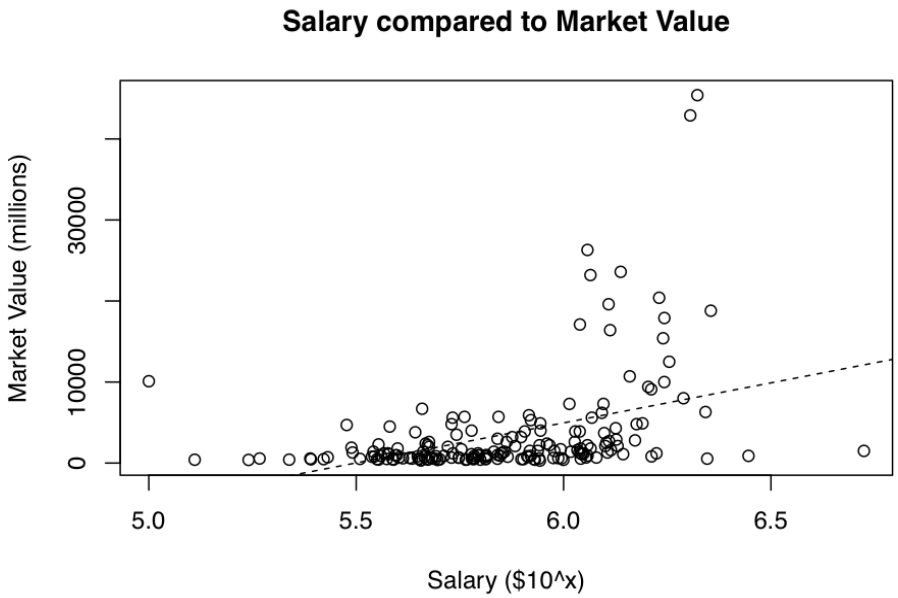
Similarly, we noted that below a 7 figure salary that there was little correlation between profits and CEO salaries. However, above that figure, there seemed to be a positive correlation as indicated below. The dashed line was fitted to the CEO salaries below $1mil, and the dotted line was fitted to the higher earners.

```
CEO_topearners = subset(CEO_trimmed, sal_log >= 6)
CEO_bottom = subset(CEO_trimmed, sal_log < 6)
plot(CEO_trimmed$sal_log, CEO_trimmed$profits,
     main = "Salary  vs. Company Profits", ylab = "Profit in $ millions", xlab="Salary (10^x)")
abline(lm(CEO_bottom$profits~CEO_bottom$sal_log), lty="dashed")
abline(lm(CEO_topearners$profits~CEO_topearners$sal_log), lty="dotted")
```



Looking at CEO salaries with market value generated similar results as profits vs. salary, because of the same confounding factor indicated above:

```
plot(CEO_trimmed$sal_log, CEO_trimmed$mktval,
     main = "Salary compared to Market Value", xlab = "Salary ($10^x)", ylab = "Market Value (millions)"
abline(lm(CEO_trimmed$mktval~CEO_trimmed$sal_log), lty = "dashed")
```




```
cor(CEO_trimmed$mktval,CEO_trimmed$sal_log)
```

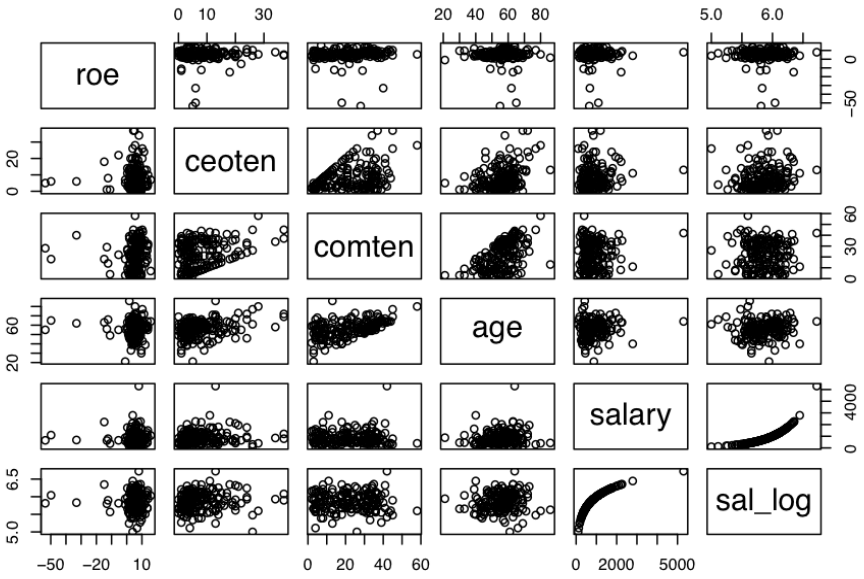
```
## [1] 0.404188
```

We deduced from the comparisions above that there was a positive correlation between market value to CEO salary and profits with CEO salary within the sample. However, there was no correlation to performance, as measured by return on equity.

Secondary Variables

In order to gain a high level view of some of the bivariate relationships in our existing dataset, we created a scatterplot matrix for our key variables that were able to be plotted in this manner:

```
plot(~roe+ceoten+comten+age+salary+sal_log, data = CEO_trimmed)
```

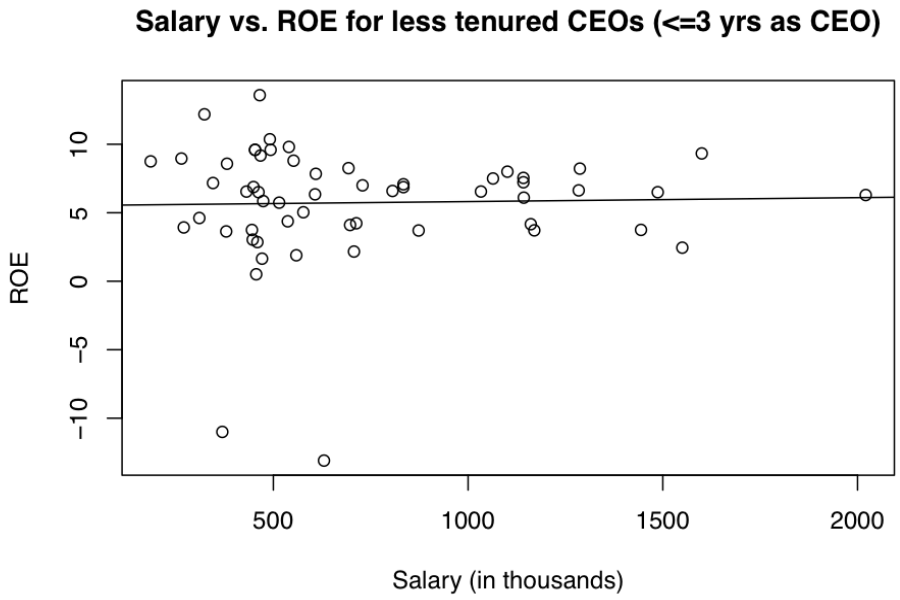


The top row did not display anything notable about the linear relationship between ROE compared to these other variables yet. The strongest correlations were from factors that made sense to be related in the first place: CEO tenure compared to age, salary compared to company tenure, company tenure compared to age, and salary compared to the log of the salary. A closer look needed to be done on on each of these variables.

CEO Tenure (CEOTEN)

CEO tenure could have played a role in what the CEOs salary was in the first place. In order to isolate the impact of prolonged tenure, we looked at a subset of companies who only had their CEOs in place for 3 years or less to see if there was a relationship between ROE and salary. Salary wasn't looked at on a logarithmic scale here because this subset of data has a much less drastic positive skew. Even when the analysis was restricted to looking at CEOs with 3 or fewer years as a CEO, there does not appear to be much of a linear relationship between ROE and CEO salary.

```
CEO_tenured3 <- subset(CEO_trimmed, ceoten <= 3)
plot(CEO_tenured3$salary, CEO_tenured3$roe,
     main = "Salary vs. ROE for less tenured CEOs (<=3 yrs as CEO)",
     xlab = "Salary (in thousands)",
     ylab = "ROE")
abline(lm(CEO_tenured3$roe~CEO_tenured3$salary))
```

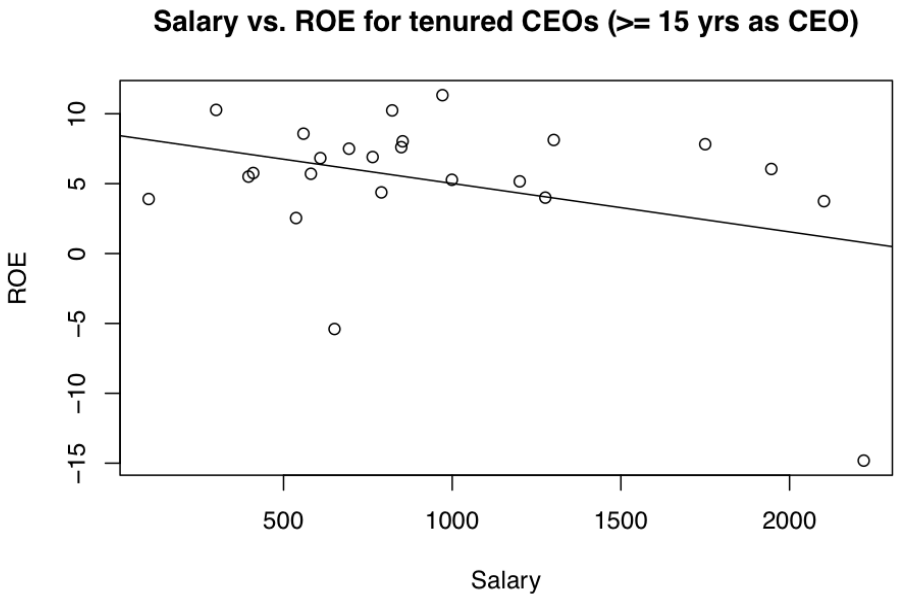



```
cor(CEO_tenured3$roe,CEO_tenured3$salary)
```

```
## [1] 0.02666566
```

We ran the same analysis for CEOs who have been in place for 15 years or more and found something very interesting:

```
CEO_tenured15 <- subset(CEO_trimmed, ceoten >= 15)
plot(CEO_tenured15$salary, CEO_tenured15$roe,
     main = "Salary vs. ROE for tenured CEOs (>= 15 yrs as CEO)",
     xlab = "Salary",
     ylab = "ROE")
abline(lm(CEO_tenured15$roe~CEO_tenured15$salary))
```




```
cor(CEO_tenured15$roe,CEO_tenured15$salary)
```

```
## [1] -0.3660428
```

There was a weak negative correlation between CEO salary and ROE for tenured CEOs (in place for 15 or more years by our definition), but by looking at the graph it was visibly being skewed by one data point. Of all the tenured CEOs in the sample set, the company with the highest CEO salary had the lowest return on equity.

It would also be preemptive to assume that more tenured CEOs have better ROE's than the newer CEOs. Surprisingly, the median ROE in the dataset of the CEOs who had been in their position 3 years or less was slightly higher than the median ROE of the CEOs:

```
summary(CEO_tenured15$roe)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -14.810   4.281   5.906   5.210   7.873  11.330
```

```
summary(CEO_tenured3$roe)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -13.100   3.927   6.545   5.737   8.222  13.580
```

Company Tenure (COMTEN)

What if a CEO was brand new to the company when they started? Did this have an impact on ROE? Subsetting the data down to companies where the CEO was a part of the company before they became the CEO result in the following ROEs:

```
COM_ceobefore <- subset(CEO_trimmed, comten > ceoten)
COM_ceobefore$cat <- "Did not start as CEO"
summary(COM_ceobefore$roe)
```

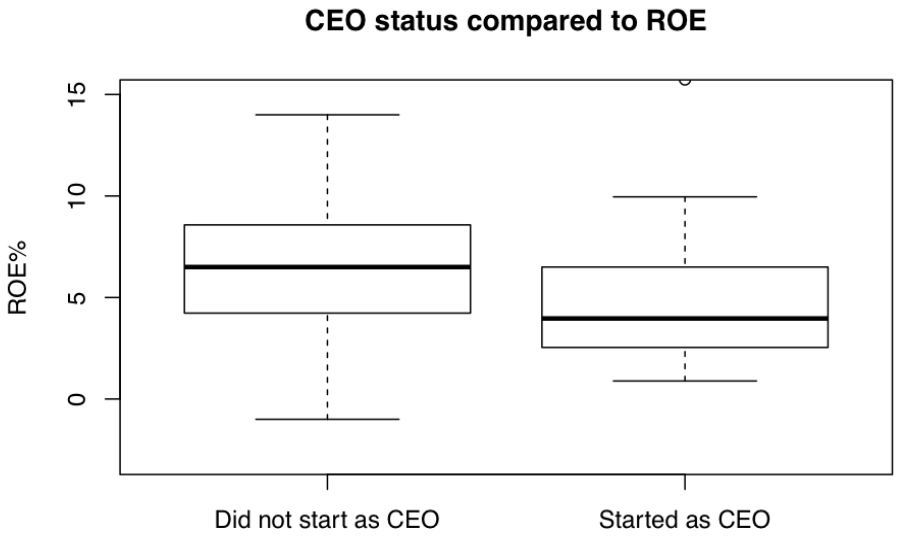
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -53.610   4.231   6.500   5.187   8.576  14.000
```

And CEOs who started with the company as CEOs have the following ROE:

```
COM_ceostarted <- subset(CEO_trimmed, comten == ceoten)
COM_ceostarted$cat <- "Started as CEO"
summary(COM_ceostarted$roe)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -14.810   2.598   3.967   4.002   6.302  15.740
```

```
combined = data.frame(rbind(COM_ceobefore,COM_ceostarted))
combined$cat <- as.factor(combined$cat)
plot(combined$cat,combined$roe,
     ylim = c(-3,15),
     main = "CEO status compared to ROE",
     ylab = "ROE%")
```



CEOs who were with the company prior to starting as a CEO tended to have higher ROEs than those who started as CEO in this sample. While this was interesting for us to find out in relation to ROE and tenure, we also delved into whether or not salary and ROE was related for people who did not start as a CEO in addition to those who did start as a CEO, and did not notice anything worth publishing.

CEO Age (age)

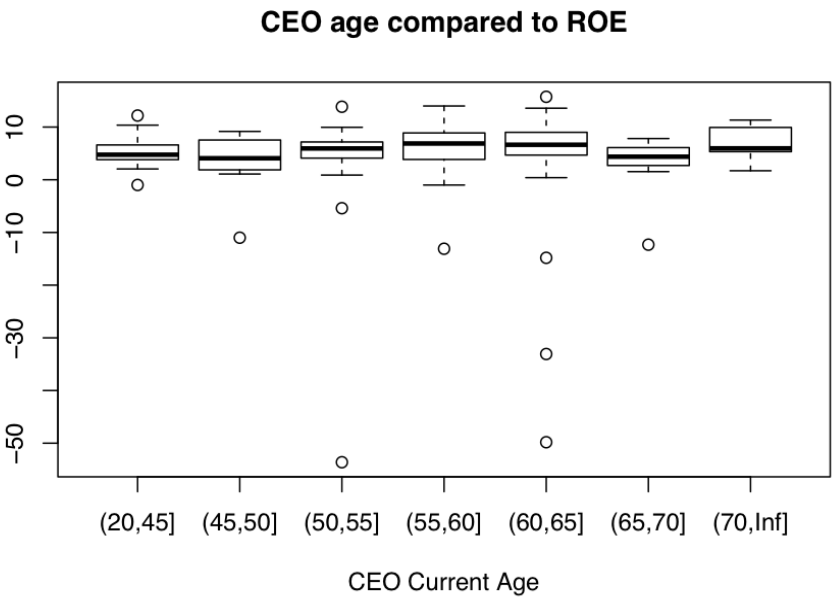
We considered age as it related to performance. To find a relationship between age and company performance, we binned our age variable into the intervals. In our analysis, we did not observe a clear relationship between ROE and CEO age.

```
summary(CEO_trimmed$age)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   21.00   51.00   57.00   55.96   62.00   86.00
```

ROE had a slight positive correlation from 45 year-old to 65 year-old CEOs, then fell in the 65 to 70 age group and increase again over 70. So, this data remained ambiguous.

```
CEO_age_bin <- cut(CEO_trimmed$age, breaks = c(20,45,50,55,60,65,70,Inf))
boxplot(CEO_trimmed$roe~ CEO_age_bin, xlab="CEO Current Age", main = "CEO age compared to ROE")
```

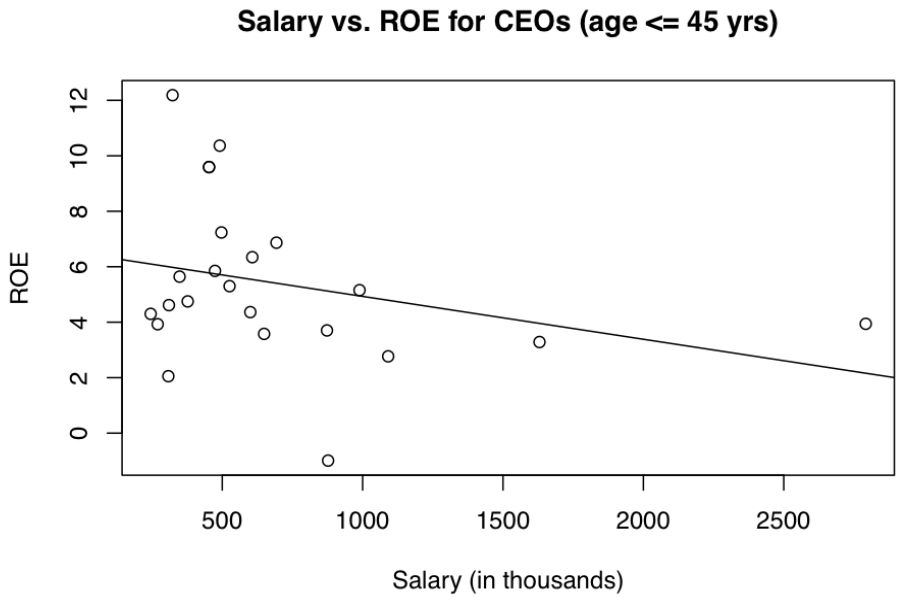



```
summary(CEO_age_bin)
```

```
##  (20,45]  (45,50]  (50,55]  (55,60]  (60,65]  (65,70] (70,Inf]
##       23       18       37       49       32       13        8
```

Interestingly, when the analysis was restricted to looking at CEOs at 45 years old or younger, there appeared to be a weak negative correlation between ROE and CEO salary. It showed us that when the CEOs were younger, company got less return on equity with higher CEO salary.

```
CEO_age45 <- subset(CEO_trimmed, age <= 45)
plot(CEO_age45$salary, CEO_age45$roe,
     main = "Salary vs. ROE for CEOs (age <= 45 yrs)",
     xlab = "Salary (in thousands)",
     ylab = "ROE")
abline(lm(CEO_age45$roe~CEO_age45$salary))
```

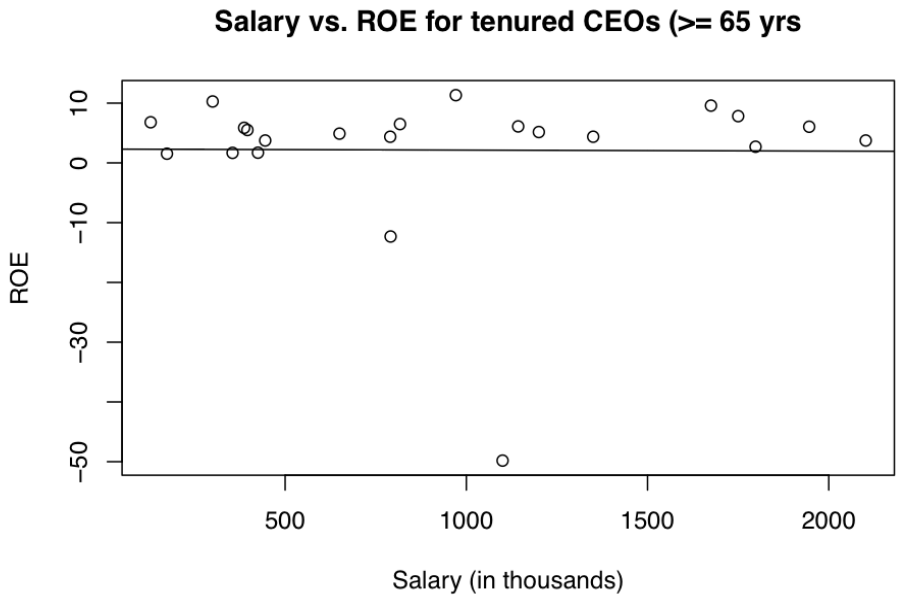



```
cor(CEO_age45$roe,CEO_age45$salary)
```

```
## [1] -0.2950547
```

We ran the same analysis for CEOs at 65 years old or older and found no correlation.

```
CEO_age65 <- subset(CEO_trimmed, age >= 65)
plot(CEO_age65$salary, CEO_age65$roe,
     main = "Salary vs. ROE for tenured CEOs (>= 65 yrs",
     xlab = "Salary (in thousands)",
     ylab = "ROE")
abline(lm(CEO_age65$roe~CEO_age65$salary))
```




```
cor(CEO_age65$roe,CEO_age65$salary)
```

```
## [1] -0.008709638
```

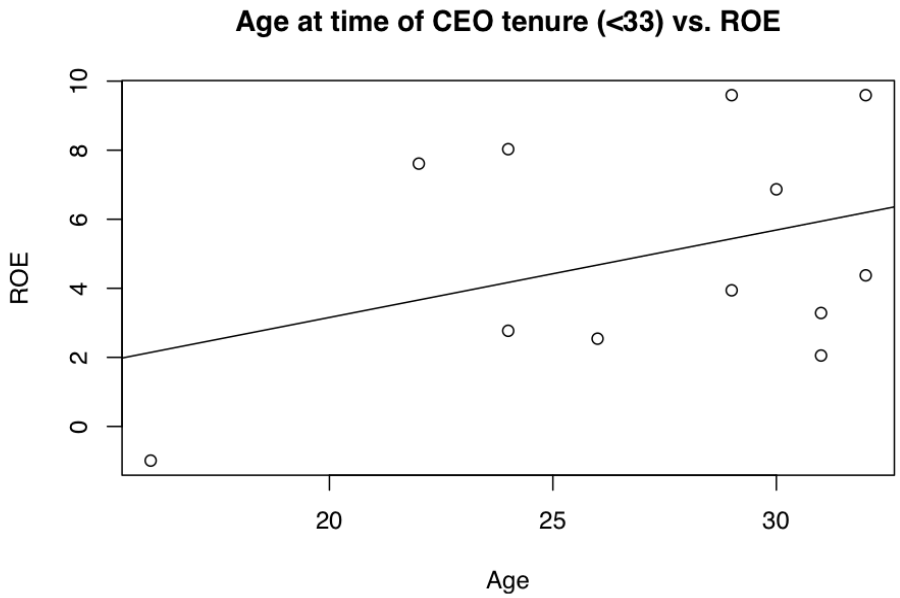
We also considered another data point that was not provided. The age at which the CEO assumed the executive position was calculated by subtracting the years of tenure from the CEOs age. We stored this new variable as ageten.

```
CEO_trimmed$ageten = CEO_trimmed$age - CEO_trimmed$ceoten
summary(CEO_trimmed$ageten)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   16.00   42.00   49.50   48.09   55.00   73.00
```

For companies where the executives assumed the position of CEO at an age younger than 33, we observed a positive correlation between ageten and company ROE.

```
CEO_ageten33 <- subset(CEO_trimmed, ageten<33)
plot(CEO_ageten33$ageten, CEO_ageten33$roe,
     main = "Age at time of CEO tenure (<33) vs. ROE",
     xlab = "Age",
     ylab = "ROE")
abline(lm(CEO_ageten33$roe~CEO_ageten33$ageten))
```

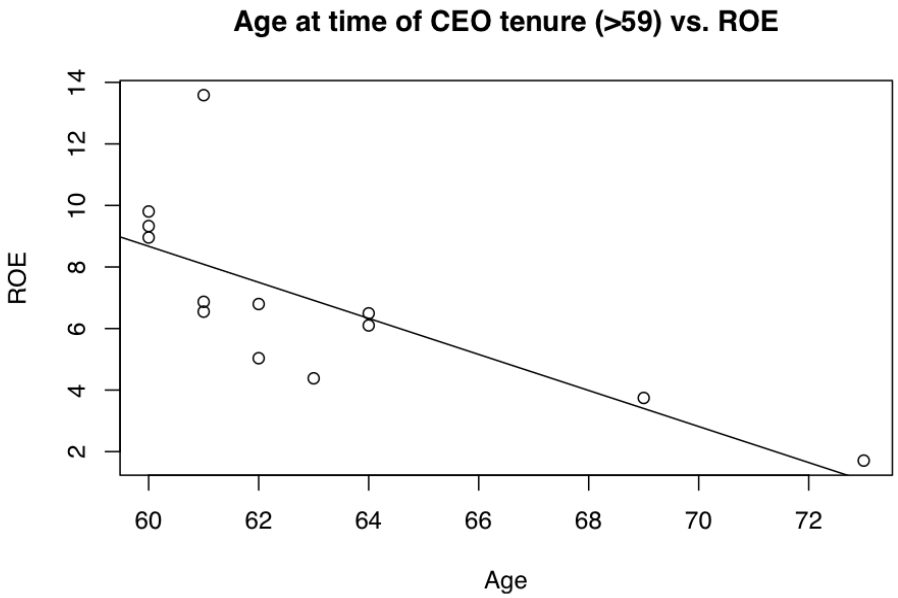



```
cor(CEO_ageten33$roe,CEO_ageten33$ageten)
```

```
## [1] 0.3729781
```

However, for companies where the CEO started at an age older than 59 years old, there was a negative correlation between ageten and company ROE.

```
CEO_ageten59 <- subset(CEO_trimmed, ageten>59)
plot(CEO_ageten59$ageten, CEO_ageten59$roe,
     main = "Age at time of CEO tenure (>59) vs. ROE",
     xlab = "Age",
     ylab = "ROE")
abline(lm(CEO_ageten59$roe~CEO_ageten59$ageten))
```




```
cor(CEO_ageten59$roe,CEO_ageten59$ageten)
```

```
## [1] -0.7442295
```

This bimodal relationship indicated that there may be interesting factors that contribute to CEO success including starting the position at a young age. However, this relationship cannot be deemed causal. The law of small numbers may be in play here as well, as there are only 13 data points being measured here.

Education

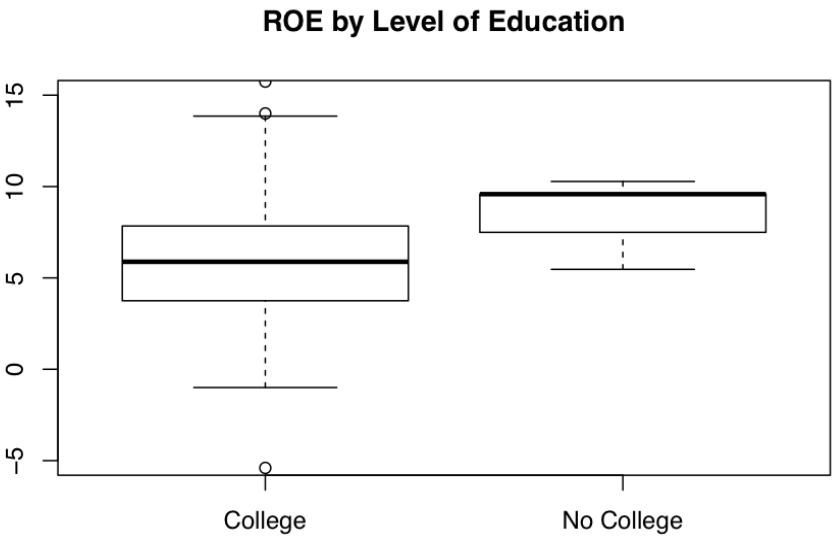
Education was another important factor that may have affected the CEO's salary and company's return on equity. We might have predicted that more highly educated CEOs would correlate to higher ROEs. Surprisingly, we found the inverse. The performance of companies with college educated CEO was lower than those without college educated CEOs. We observed no obvious difference between CEOs who have taken graduate education between those who have not.

Attended College (college)

We binned the college variable and plotted ROE by level of education.

```
CEO_gradCollege <- subset(CEO_trimmed, college == 1)
CEO_gradCollege$cat <- "College"
CEO_gradNoCollege <- subset(CEO_trimmed, college == 0)
CEO_gradNoCollege$cat <- "No College"

combined1 = data.frame(rbind(CEO_gradCollege,CEO_gradNoCollege))
combined1$cat <- as.factor(combined1$cat)
plot(combined1$cat,combined1$roe,
     ylim = c(-5,15),
     main = "ROE by Level of Education")
```



Note that the sample size is small (n= 7), therefore this may not be reliable.

```
sum(CEO_trimmed$college==0)
```

```
## [1] 7
```

Attended Graduate School (grad)

Examining whether the CEO attended graduate school, we binned the grad variable and again plotted ROE by level of education. We found that CEOs who have taken graduate level eduation had only a slightly better median ROE than CEO who have not.

```
CEO_NogradCollege <- subset(CEO_trimmed, grad == 1)
summary(CEO_NogradCollege$roe)
```

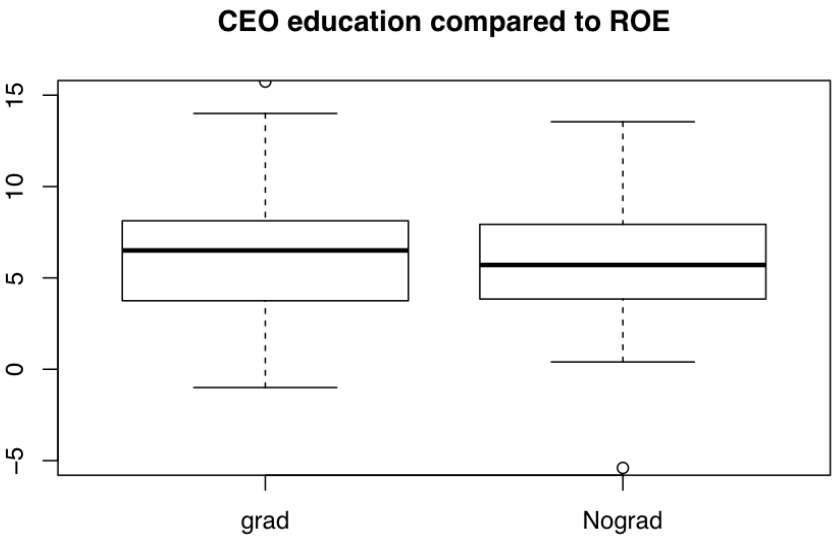
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -53.610   3.748   6.502   5.339   8.125  15.740
```

```
CEO_NogradCollege$cat <- "grad"
CEO_NogradNoCollege <- subset(CEO_trimmed,  grad == 0)
summary(CEO_NogradNoCollege$roe)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -49.820   3.843   5.707   4.393   7.930  13.550
```

```
CEO_NogradNoCollege$cat <- "Nograd"

combined2 = data.frame(rbind(CEO_NogradCollege,CEO_NogradNoCollege))
combined2$cat <- as.factor(combined2$cat)
plot(combined2$cat,combined2$roe,
     ylim = c(-5,15),
     main = "CEO education compared to ROE")
```



Conclusion

In the given sample, there was no linear correlation between salary and return on equity. Even if there was, it would be a stretch to say a correlation between salary and performance existed, since ROE was an oversimplified measure of company performance. We did find that more extreme ROEs (under 3% or above 7%) tended to yield slightly lower salaries than ROEs in between these two percentages. This particular study was very limited because of the lack of other the supplementary variables that were missing in the dataset, as we discussed earlier (year over year growth, company size, etc.).

It's important to note that this study was not causal; we could not conclude that paying the CEOs a certain amount would cause them to perform a certain way, we could only comment on their relationship. This was only a sample of the population, so we were unable to make conclusions about the entire population of companies from it. We were unaware if the original researchers used random selection in choosing the companies in his sample, or if the sample size was sufficient. Considering this study was done in 1990, it may not be reflective of the current environment anyways.

This said, our analysis did generate some interesting findings after subsetting some of our data with the attributes of our secondary variables. Had the dataset included a current, larger sample size with additional performance data, our analysis on CEO salary vs. company performance would have been much more comprehensive.

Works Cited

Canipe, C., & Slobin, S. (2015, June 24). CEO Pay vs Performance - WSJ.com. Retrieved January 29, 2017, from http://graphics.wsj.com/ceopay-2015/

Cooper, Michael J. and Gulen, Huseyin and Rau, P. Raghavendra, Performance for Pay? The Relation Between CEO Incentive Compensation and Future Stock Price Performance (November 1, 2016). Available at SSRN: https://ssrn.com/abstract=1572085 or http://dx.doi.org/10.2139/ssrn.1572085

Hagill, J., III, Brown, J. S., & Davison, L. (2010, March 4). The Best Way to Measure Company Performance. Harvard Business Review. Retrieved January 29, 2017, from https://hbr.org/2010/03/

the-best-way-to-measure-compan

Pyke, A. (2016, September 23). The More A Company Pays Its CEO, The Worse Its Shareholders Do. Retrieved January 29, 2017, from https://thinkprogress.org/the-more-a-company-pays-its-ceo-the-worse-its-shareholders-do-9b33
.ua9gkhgnu