Lab 1: Exploratory Analysis of CEO Salary Data

Anamika Sinha, Asha Anju, Noah Randolph

June 4, 2017

Introduction

An exploratory analysis of CEO and company data was conducted to explore CEO salary, with a particular interest in answering the question of whether company performance is related to CEO salary.

Exploratory data analysis techniques were used and will be presented with commentary on the findings. We first explore data in univariate fashion, followed by multivariate analyses selected for their contribution to the exploration of CEO salary. In general, CEO salary is treated as a response variable, while other datasets are treated as predictor variables.

Setup

We will utilize the "car: Companion to Applied Regression" package in R and load it, as well as the data file of interest.

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

Exploratory Analysis

To get an initial understanding of the dataset, we look at the structure of its data frame and a summary of its variables.

```
str(CEO)
```

```
## 'data.frame':    184 obs. of  8 variables:
##  $ salary : num  1276 925 2199 369 218 ...
##  $ age    : num  64 56 52 49 57 62 86 59 54 43 ...
##  $ college: num  1 1 1 1 1 1 1 1 1 1 ...
##  $ grad   : num  0 1 1 1 1 0 1 0 0 1 ...
##  $ comten : num  41 26 8 4 33 40 13 35 31 10 ...
##  $ ceoten : num  17 12 8 1 5 6 13 10 4 10 ...
##  $ profits: num  52 67 475 -132 41 -463 11 24 46 48 ...
##  $ mktval : num  1300 2200 6300 1200 421 1400 644 623 812 1100 ...
```

```
summary(CEO)
```

```
##      salary            age           college           grad        
##  Min.   : 100.0   Min.   :21.00   Min.   :0.000   Min.   :0.0000  
##  1st Qu.: 470.8   1st Qu.:51.00   1st Qu.:1.000   1st Qu.:0.0000  
##  Median : 700.5   Median :57.00   Median :1.000   Median :1.0000  
##  Mean   : 856.0   Mean   :55.98   Mean   :0.962   Mean   :0.5489  
##  3rd Qu.:1102.5   3rd Qu.:61.25   3rd Qu.:1.000   3rd Qu.:1.0000  
##  Max.   :5299.0   Max.   :86.00   Max.   :1.000   Max.   :1.0000  
##      comten          ceoten          profits            mktval
```

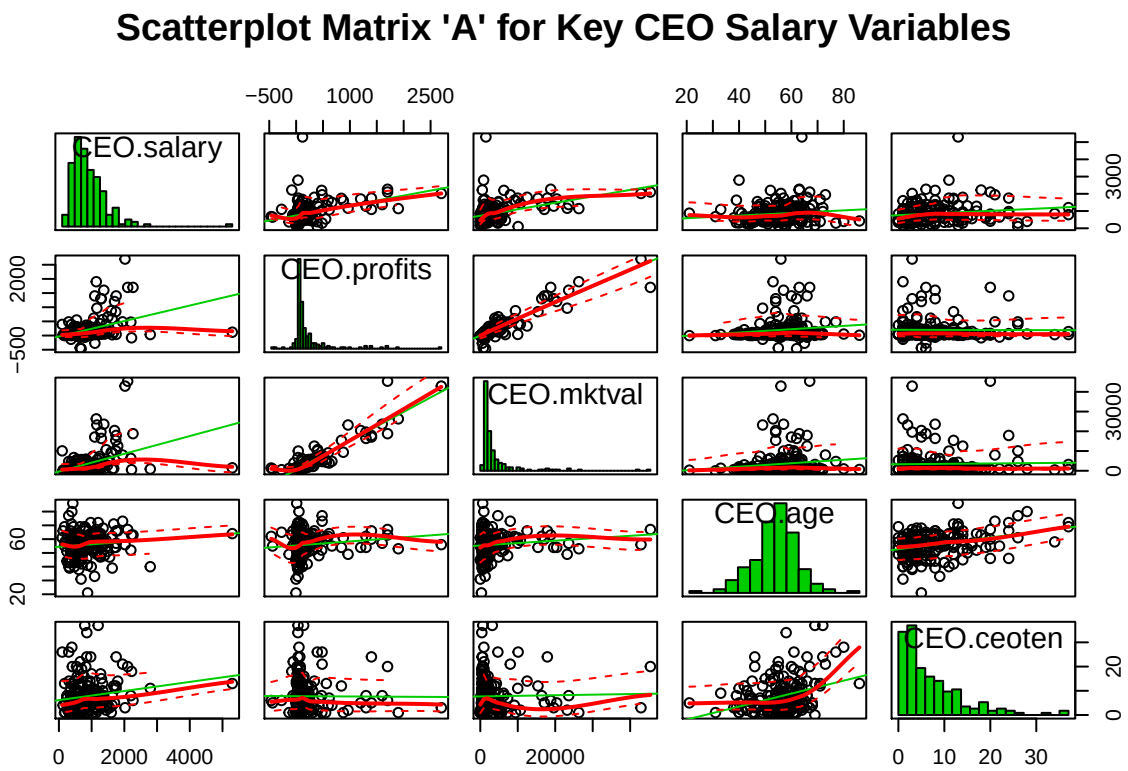
```
##  Min.   : 2.00   Min.   : 0.000   Min.   :-463.00   Min.   :   -1.0
##  1st Qu.: 9.00   1st Qu.: 3.000   1st Qu.:  31.75   1st Qu.:  578.2
##  Median :21.50   Median : 5.500   Median :  57.00   Median : 1200.0
##  Mean   :21.77   Mean   : 7.755   Mean   : 199.89   Mean   : 3466.1
##  3rd Qu.:33.00   3rd Qu.:11.000   3rd Qu.: 197.75   3rd Qu.: 3200.0
##  Max.   :58.00   Max.   :37.000   Max.   :2700.00   Max.   :45400.0
```

Now we have an idea of the size of our dataset (184 observations) and the scale of each variable. The data appears to be all integers. We can see that college and grad appear to be variables of just 0 and 1 outcomes, so it will be useful to convert those to factors for labeling. Both profits and market value (mktval) have negative values.

Given that we are most interested in company performance and CEO salary, market value and profits appear to be our primary variables of interest, along with salary, of course. The rest of which are secondary variables that could contribute to a better understanding of the primary relationships.

Given the large number of variables, scatterplot matrices are printed below to provide an overview of relationships between variables and identify areas of interest.

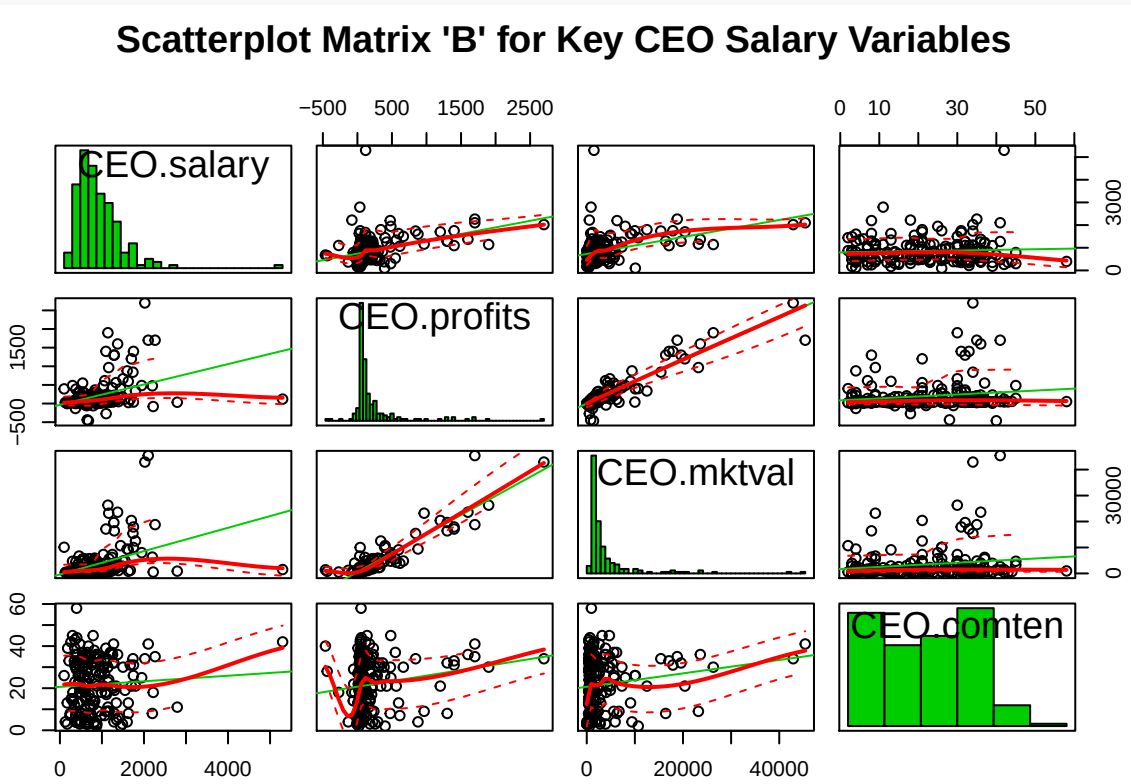
```
scatterplotMatrix(~ CEO$salary + CEO$profits + CEO$mktval + CEO$age +
                          CEO$ceoten, data = CEO, diagonal = "histogram",
                  main = "Scatterplot Matrix 'A' for Key CEO Salary Variables")
```



In the upper diagonal half of the above matrix, it appears that the two company performance variables, profits and market value, both show a positive trend against CEO salary, although the data for both appear to be highly clustered at the low end. The CEO age variable does not appear to have a strong influence on profits and market value, but there do appear to be slighter more high values to the right of the median age. The variable CEO tenure (ceoten) looks to have higher values of profits and market value clustered at the low end of tenure.

```
scatterplotMatrix(~ CEO$salary + CEO$profits + CEO$mktval + CEO$comten,
                  data = CEO, diagonal = "histogram",
```

```
                main = "Scatterplot Matrix 'B' for Key CEO Salary Variables")
```



In the upper diagonal half of the matrix above, the variable company tenure (comten) appears to have slight positive relationships between profits and market value. Salary appears to be fairly evenly distributed across the span of company tenures.

Univariate Analyses

We investigate the variables independently to understand their frequency distributions and to observe any anomalies, apparent data errors, and missing values. The dataset describing college attendance will not be analyzed due to the small number of CEOs who did not attend college (<4%), as can be seen below.

```
CEO$college <- factor(CEO$college, labels = c("Did Not Attend College",
                                              "Attended College"))
summary(CEO$college)
```

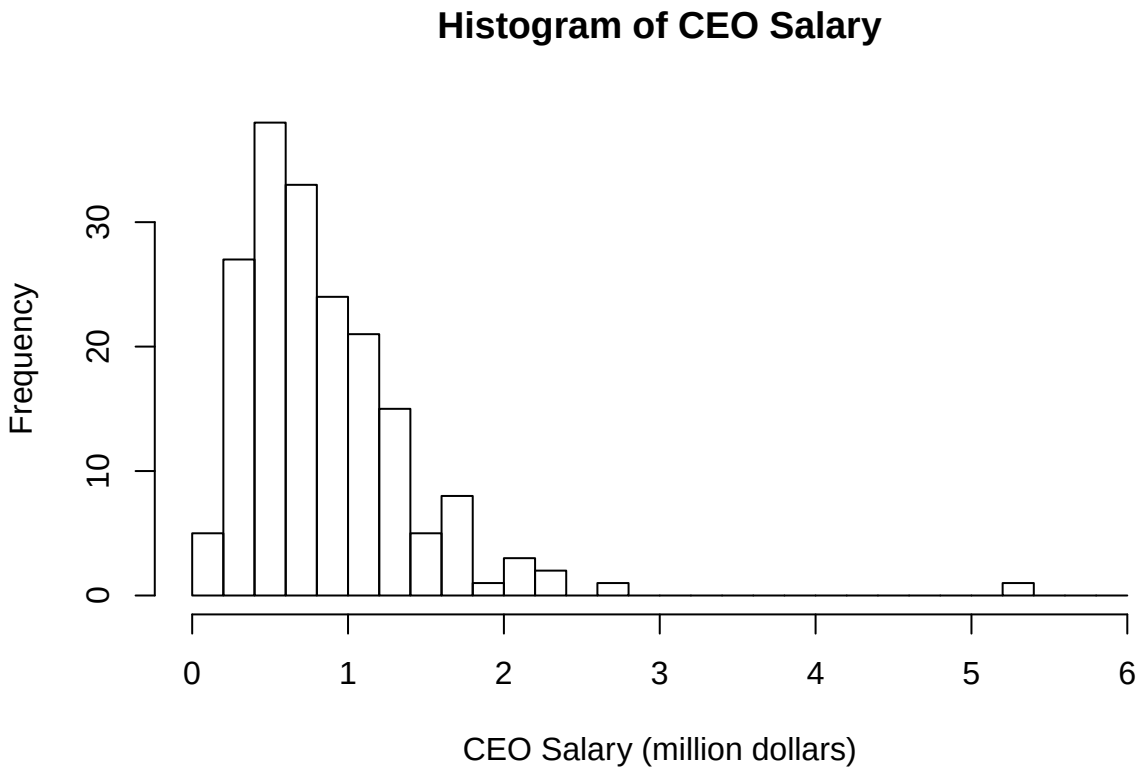
```
## Did Not Attend College       Attended College
##                      7                    177
```

Univariate Analysis of Salary Data

We start with the analysis of CEO salary, our response variable of interest. We first examine the histogram of salary data. The individual salary amounts have been scaled to reflect the value in millions. We have set the break points manually but the bin width is same as that obtained using Freedman and Diaconis rule.

```
hist(CEO$salary/1000, breaks = seq(0,6,.2),
     xaxt = "n",
     main = "Histogram of CEO Salary",
```

```
     xlab = "CEO Salary (million dollars)")
axis(1, seq(0,6,1))
```



Next we look at the summary statistics for the salary data.

```
summary(CEO$salary/1000)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.1000  0.4708  0.7005  0.8560  1.1025  5.2990
```

The following observations are apparent from our analysis of CEO salaries:

1. The distribution of CEO salaries is positively skewed with a peak around $0.6 million.
2. Except for the maximum value, all salaries are below $3.0 million. Most of the salaries are below $2.0 million.
3. The distance of maximum value from the 3rd quartile value is more than three times the interquartile range, making the data point an outlier.

Eventhough the maximum salary is an outlier, there is no strong evidence to conclude that it should be removed from our analysis.

Univariate Analysis of Market Value

Market value is commonly associated with company performance, which is of primary interest as it relates to CEO salary. Thus, we include an assessment of its data in the variable, "mktval", univariately. First, a summary of the data:

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

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    -1.0   578.2  1200.0  3466.1  3200.0 45400.0
```

The market value range is vast, with the maximum value far greater than the values in the interquartile range. In fact, the mean is greater than not just the median, but also the 3rd quartile. Since a stock price cannot go below $0, and market value for public companies is composed of the sum of the values of each individual stock, any negative values should be scrutinized.

```
CEO[CEO$mktval < 0, ]
```

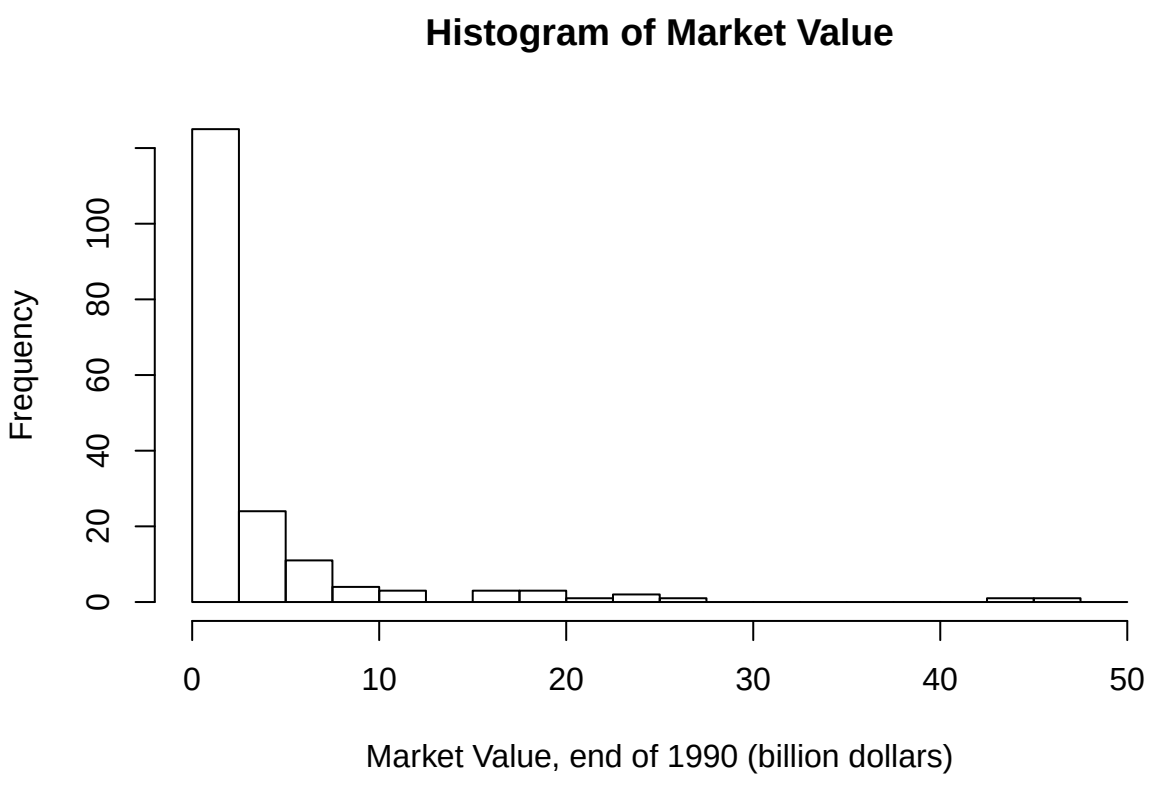
```
##     salary age          college grad comten ceoten profits mktval
## 182    637  45 Attended College    1      3      1      -1     -1
## 179    677  31 Attended College    1      3      1      -1     -1
## 180    173  55 Attended College    1      3      1      -1     -1
## 178    379  55 Attended College    1      4      2      -1     -1
## 181    873  61 Attended College    1      3      1      -1     -1
```

A subset of the data less than zero reveals that a coding error is apparent. The -1 values in both market value and profits likely represent data points that were unavailable for particular companies. Further analyses in this report on market value or profits will be on subsets of the data excluding the values of -1.

```
cleanCEO <- subset(CEO, CEO$mktval > 0)
```

Next, we create a histogram of the cleaned up market value data. The dataset is transformed by a factor of 1/1,000 to make for easier reading of the X-axis. The bins are set with even widths at a number of 20, which is an amount between two common rules, one of which is the square root of the number of data points, and the other is the Freedman and Diaconis rule.

```
mktvalbillions <- cleanCEO$mktval/1000
hist(mktvalbillions, breaks = seq(0, 50, 2.5),
     main = "Histogram of Market Value",
     xlab = "Market Value, end of 1990 (billion dollars)")
```



The histogram reveals the following characteristics:

1. Market value is skewed toward the positive and frequencies drop steeply between $2.5 billion and $5

billion, for a decaying profile of frequencies as market value increases.

2. Although market value does not take on negative values, integer effects are not of concern due to its high dispersion.
3. Most of the market values fall below $5 billion.
4. There is a large gap in data between roughly $27.5 billion and $42.5 billion.

Although the market values between $40 billion and $50 billion fall far outside of the bulk of the data, there is no strong evidence to conclude that they should be removed from the analysis as outliers.

Univariate Analysis of Profits

Next, we investigate company profits as it is also an indicator of company performance.

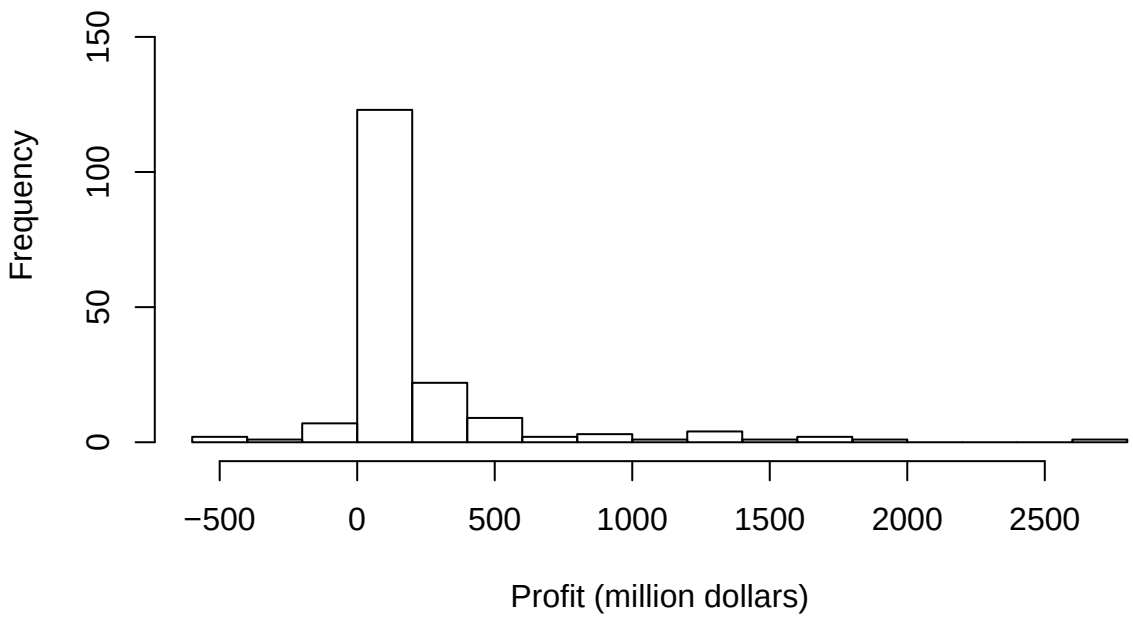
```
summary(cleanCEO$profits)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  -463.0    34.0    63.0   205.5   207.0  2700.0
```

We see that there are negative profits but the number of companies with negative profits may be low as the first quartile value is still positive. There is a wide range in values with the maximum profit going as high as 13 times the 3rd quartile value. The mean being much higher than the median and even more than the third quartile values indicates the presence of a few very high profit values.

We use the cleanCEO dataset as explained above for market value (5 values with coding errors have been dropped).

```
hist(cleanCEO$profits,
     breaks = 20,
     ylim = c(0,175),
     main = "Profit for Companies in 1990",
     xlab = "Profit (million dollars)")
```



Observations from company profit histogram:

1. Most companies have profits below $250 million.
2. The histogram has a slight positive skew due to the presence of some outliers. They seem to be pushing the mean to even beyond the 3rd quartile. The mean is much higher than the median value.
3. Since most of the profits are from $0 to $250 million, that bin is worth some further analysis. We split this bin later, for more detail, to see if there are any apparent trends.
4. It will be interesting to see the salaries of CEOs of the 10 companies with negative profits.

Univariate Analysis of CEO Age

We next examine the age of CEOs to understand the distribution and its features.

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

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   21.00   51.00   57.00   55.98   61.25   86.00
```

```
hist(CEO$age, breaks = "FD",
     main = "Histogram of CEO Age",
     xlab = "Age of CEO (years)")
```

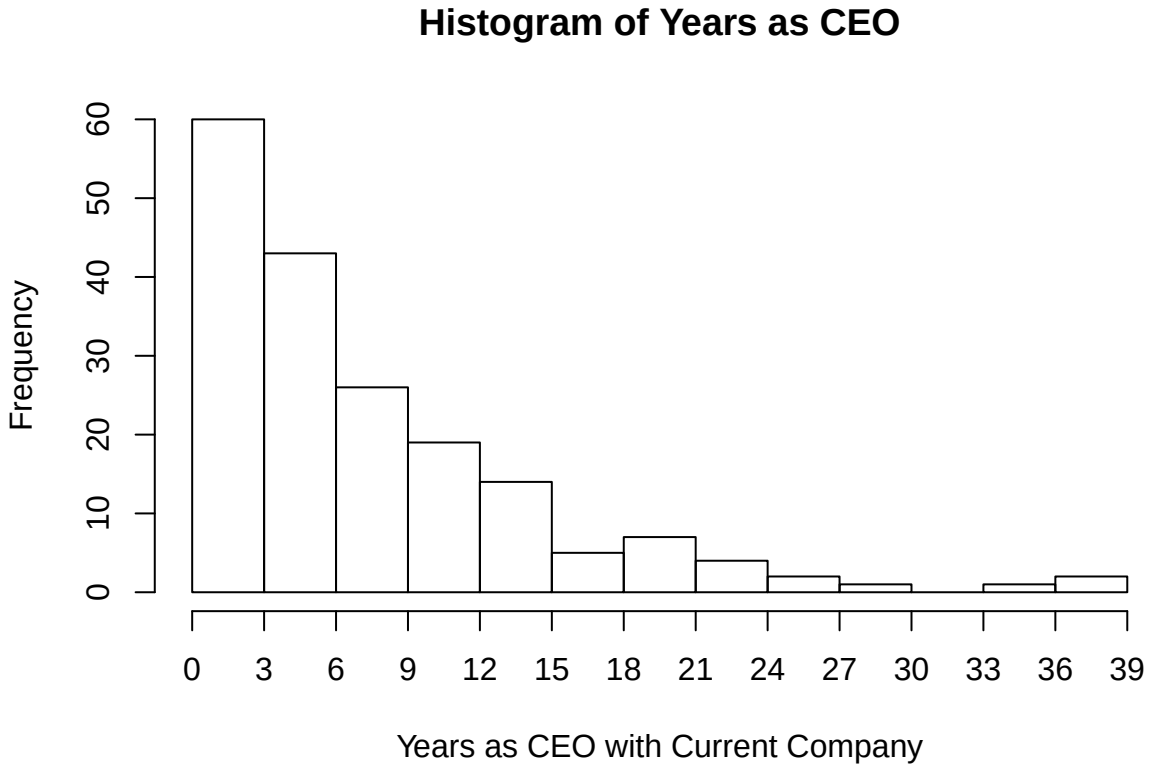
The following observations are apparent from our analysis of age information:

1. The ages cover a large range with a minimum age of 21 years and maximum age of 86 years.
2. The distribution peaks peak around 55-60 years of age and have a slight negative skew.
3. 50% of the CEOs in the dataset have ages in a 10 year range between 51 and 61.25 years.

Univariate Analysis of CEO Tenure

We next examine years of CEO tenure, which may relate to CEO salary as any normal salary typically grows with tenure. We manually set the cut points for our histogram to make 13 bins, which is roughly the square root of the number of data points.

```
hist(CEO$ceoten, breaks = seq(0, 39, by=3),
     main = "Histogram of Years as CEO",
     xlab = "Years as CEO with Current Company",
     xaxt = "n")
axis(1, seq(0, 39, by=3))
```



The following observations are apparent in the histogram of CEO tenures:

1. The lower bound is at zero years (indeed, negative years would not make sense in this context).
2. The data is skewed positively, with more than one data point in the rightmost bin of 36 to 39 years.
3. CEO tenure frequencies peak in the lowest number of years and then mostly drop in a decaying form as the number of years increases.
4. There is a gap in tenures from 30 to 33 years and a decrease from 15 to 18 years, but these are most likely due to dispersion of lower frequencies of higher year values.

Univariate Analysis of Company Tenure

We analyzed the total number of years CEOs stayed with the company including years before and after becoming CEO.

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

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    2.00    9.00   21.50   21.77   33.00   58.00
```

There is a wide range in the number of years CEOs have stayed with their companies. The mean and median are almost equal, thus there should be no skew of the data.

```
hist(CEO$comten,
     ylim = c(0,40),
     main = "Histogram of Years in Company",
     xlab = "Number of Years in Company"
     )
```

The following observations are apparent from the histogram:

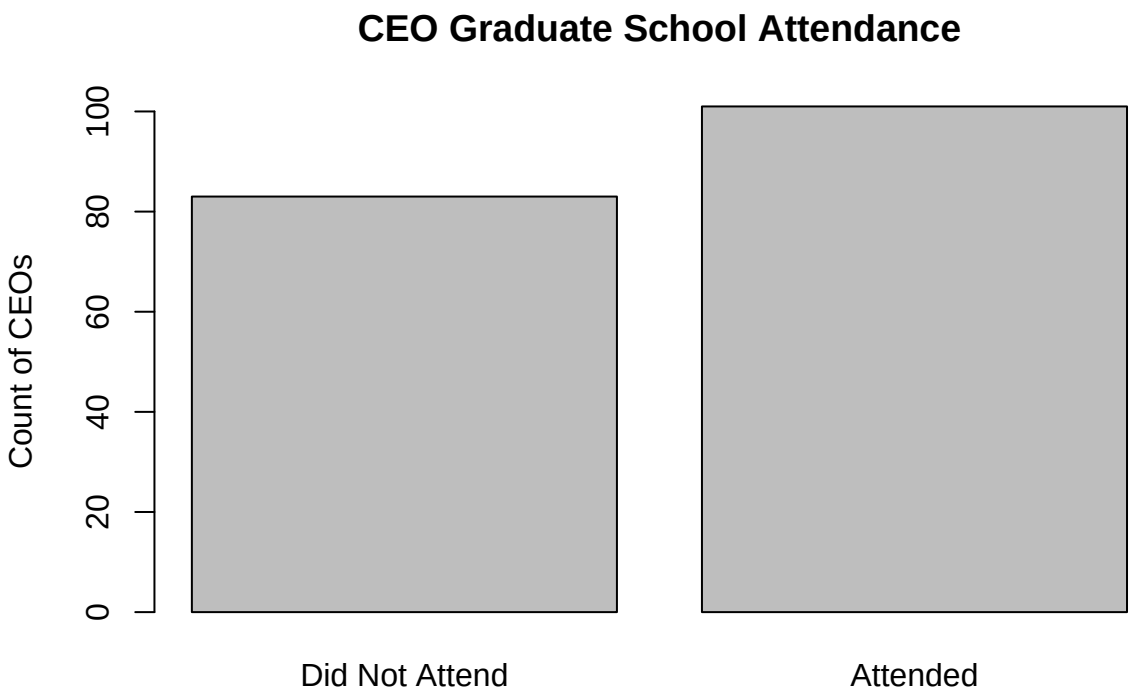
1. We see that there is almost a uniform spread of data across the number of years from zero to forty years, hence the close agreement of mean and median.
2. There is a small spike at 30-35 years and 1 observation in the 55 to 60 years.

Univariate Analysis of Graduate School Attendance

Since graduate school (grad) is a categorical variable, we create a factor variable and assign levels of "Attended" and "Did Not Attend" to it.

```
CEO$grad <- factor(CEO$grad,
                   labels = c("Did Not Attend", "Attended"))
```

```
plot(CEO$grad, ylab = "Count of CEOs",
     main = "CEO Graduate School Attendance")
```




```
summary(CEO$grad)
```

```
## Did Not Attend       Attended
##             83            101
```

We see that the number of CEOs who did not attend graduate school is pretty comparable to the number of CEOs who did.

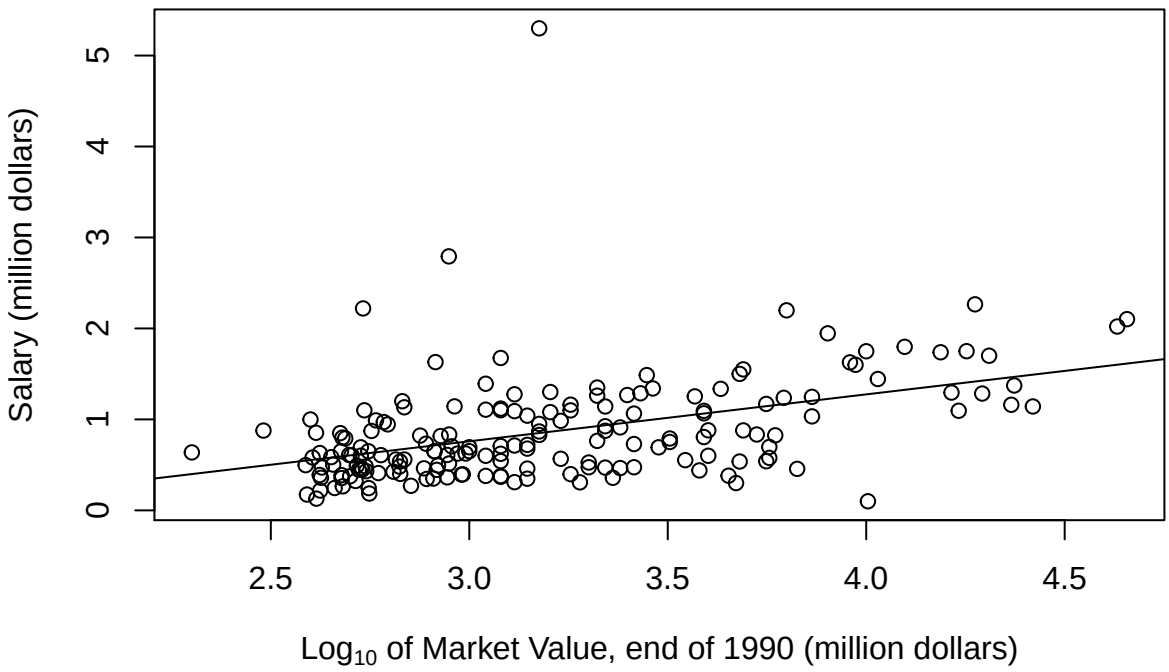
Multivariate Analyses

Now that we understand the type and general nature of the data in each variable, we turn to multivariate analyses to begin to answer the question of whether CEO salary is related to company performace. We are using market value and profits as proxies for company performance for the primary assessment of salary as it relates to company performance.

Market Value Versus CEO Salary

Because of the high clustering of market value data seen in the scatterplot matrix of the initial exploration above, market value will be plotted using a log_{10} scale. We exclude observations (whole rows of data) including the -1 values found in the univariate analysis of market values (though they would not plot on a log scale even if included).

```
plot(log10(cleanCEO$mktval), cleanCEO$salary/1000, main = "Market Value Versus CEO Salary",
     xlab = expression(paste(Log[10],
                             " of Market Value, end of 1990 (million dollars)")),
     ylab = "Salary (million dollars)")
abline(lm(cleanCEO$salary/1000 ~ log10(cleanCEO$mktval)))
```



The plot above, with its linear regression line, shows a clear positive relationship between salary and the log_{10} of market value. The correlation, below, also shows that indeed some relation exists. However, the relationship cannot be said to be linear, since market value was plotted in log_{10} scale.

```
cor(cleanCEO$mktval, cleanCEO$salary)
```

```
## [1] 0.4066159
```

Salary Versus Profit

High clustering of profit data in the scatterplot matrix tempted us to plot a log transformation of profits against salary to look for a relationship. However, due to the presence of 10 negative profit values, we decided against a transformation.

Instead, we grouped the profit data to see if salaries for CEOs showed a trend across profit bins.

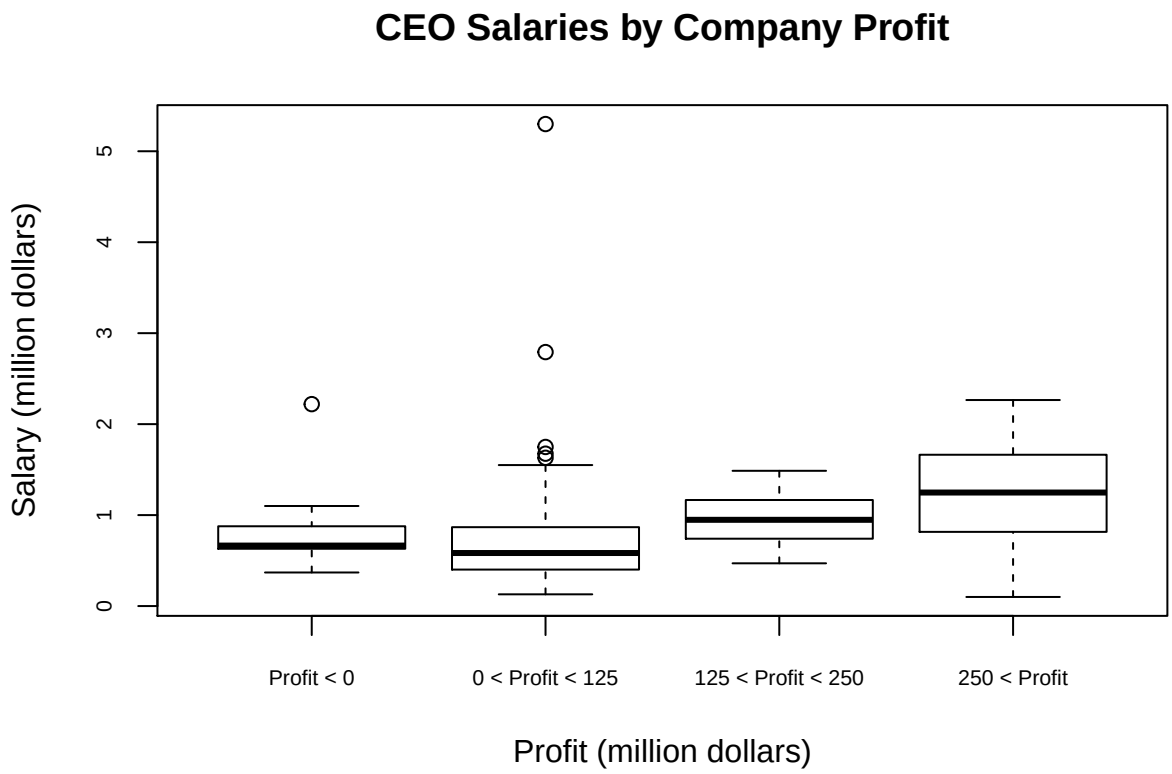
The histogram analysis of profits, earlier, led us to create a bin for negative profits. We created two more bins of profits in the ranges of $0 to $125 million and $125 to $250 million. Finally, we created another bin of values greater than $250 million to bundle higher values of profit together.

```
profit_bin = cut(cleanCEO$profits, breaks = c(-500,0,125,250, Inf), labels =
      c("Profit < 0", "0 < Profit < 125","125 < Profit < 250", "250 < Profit"))
summary(profit_bin)
```

```
##         Profit < 0   0 < Profit < 125 125 < Profit < 250
##                 10                106                 28
##       250 < Profit
##                 35
```

```
boxplot(salary/1000 ~ profit_bin, data = cleanCEO, cex.axis = .7,
        main = "CEO Salaries by Company Profit",
```

```
        ylab = "Salary (million dollars)",
        xlab = "Profit (million dollars)")
```



From the graph above, we notice an upward trend in median salary with increasing profits, for all positive profits . There are some outliers with high salaries in the zero to $250 million bucket. The wider range of salaries in the $0 to $250 million as well as greater than $250 million bucket is also noticeable.

It is also interesting to note that CEOs in companies with negative profits did not get paid the lowest. In fact, the median value in that bucket is slightly higher than the $0 to $250 million bucket.

```
cor(cleanCEO$salary, cleanCEO$profits, use = "complete.obs")
```

```
## [1] 0.3942312
```

The correlation value of 0.39 indicates a positive, moderate, correlation.

Other Influences on Salary

This section looks beyond the company performance variables of profits and market value in relation to salary. We seek to find the extent to which other variables relate to salary, which could help make clearer the primary relationship with company performance.

CEO Tenure Versus Salary

The variable ceoten captures the number of years of service as CEO. 75% of the values are concentrated in the first one third of the overall tenure spread. This clustering makes scatter plots less informative. Therefore to analyse the influence of ceo tenure on the primary variables, we group them into bins. Fifty percent of the tenure values are below 6 years and therefore we create two bins in this range to better understand the data.

```
ceoten_bins <- cut(cleanCEO$ceoten,
                        c(0,3,6,10,20,40),
```

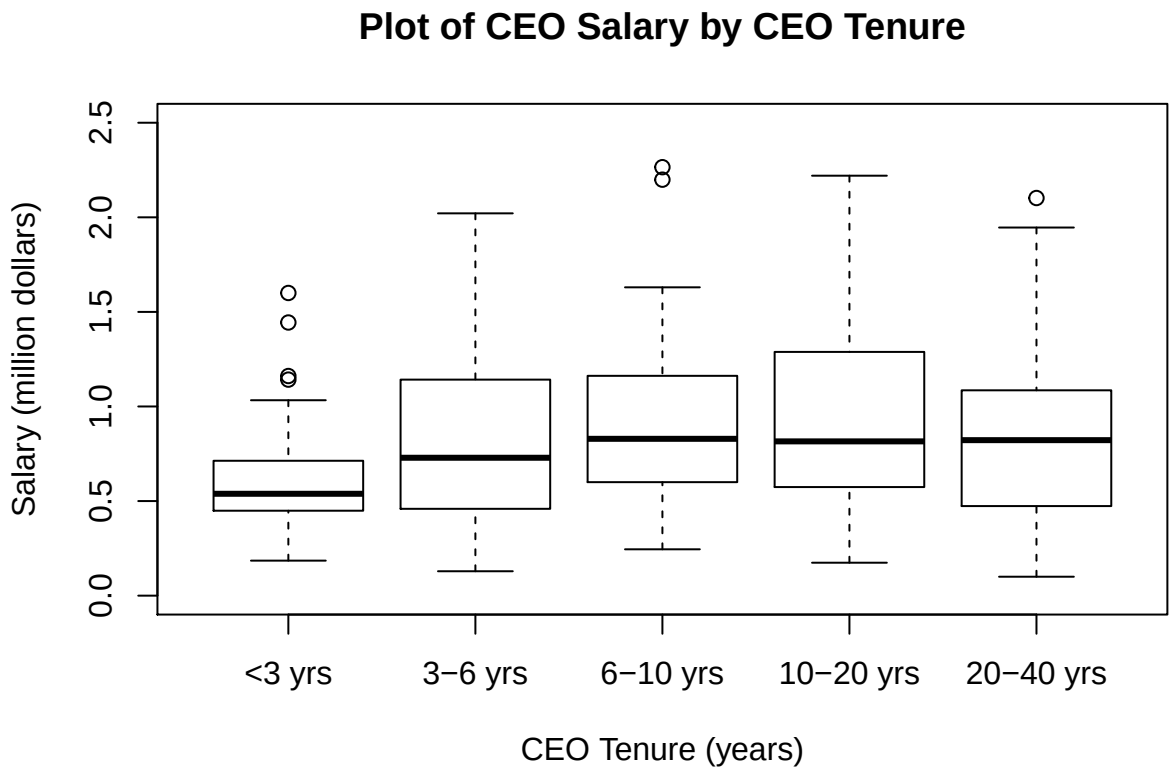
```
                        include.lowest = TRUE,right = FALSE,
                        labels = c("<3 yrs","3-6 yrs", "6-10 yrs",
                                   '10-20 yrs','20-40 yrs'))

# Gives no of datapoints in each bin
summary(ceoten_bins)
```

```
##    <3 yrs   3-6 yrs  6-10 yrs 10-20 yrs 20-40 yrs
##        34        53        37        40        15
```

We create a boxplot of salary for the various tenure groups. The y axis has been limited to salary of $2.5 million.

```
boxplot(cleanCEO$salary/1000~ ceoten_bins,
     main = "Plot of CEO Salary by CEO Tenure",
     xlab = "CEO Tenure (years)",
     ylab = "Salary (million dollars)",
     ylim = c(0,2.5))
```



The median salaries for CEOs appear to increase with tenure for tenure less than 10 years. Beyond that the effect of years of tenure seem to reduce.

Company Tenure Versus Salary

The variable comten captures the number of years of service with the company. The data is fairly uniformly distributed across the years. To simplify our analysis, we group the data based on years of tenure at the company. We have created bins of size 10 years except for the last bin which captures data for a twenty year period. This is because of the low frequency of data with such high tenure.

```
comten_bins <- cut(cleanCEO$comten,
                   c(0,10,20,30,40,60),
```

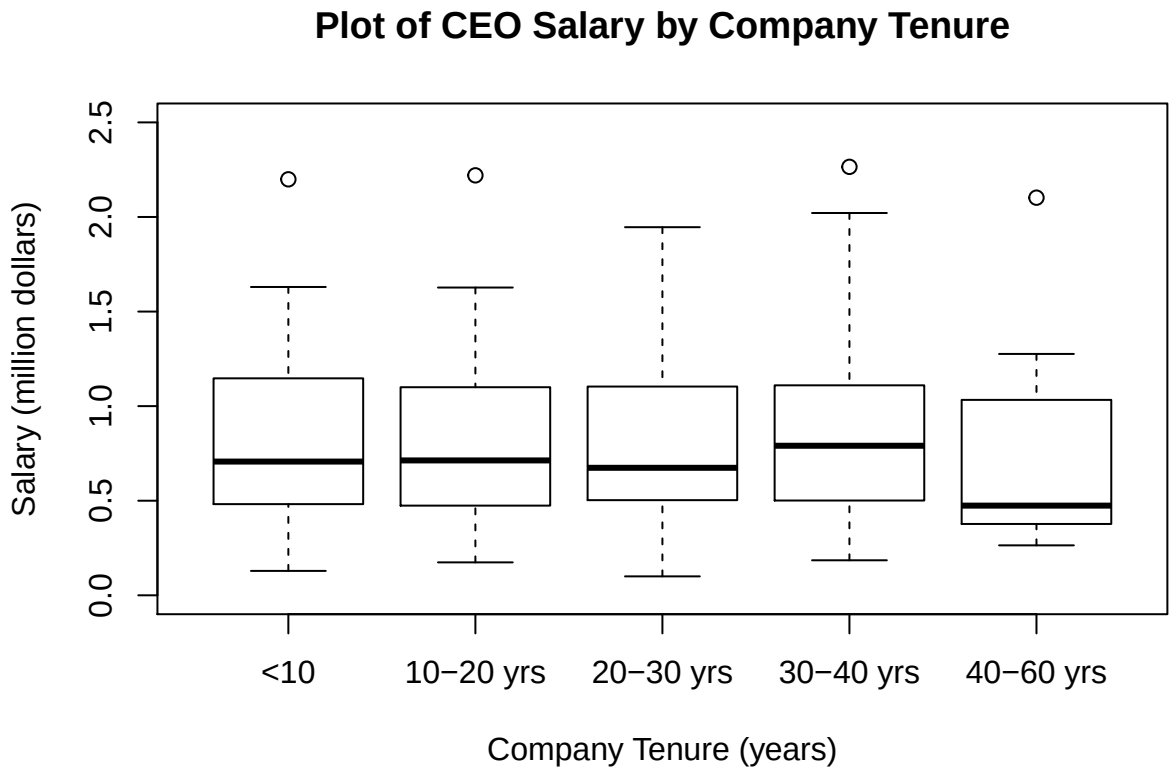
```
                   labels = c('<10', '10-20 yrs','20-30 yrs',
                              '30-40 yrs', '40-60 yrs'),
                   right = FALSE)

summary(comten_bins)
```

```
##       <10 10-20 yrs 20-30 yrs 30-40 yrs 40-60 yrs
##        43        33        36        56        11
```

We create a boxplot of salary for the various tenure groups. The y-axis has been limited to salary of $2.5 million.

```
boxplot(cleanCEO$salary/1000~ comten_bins,
     main = "Plot of CEO Salary by Company Tenure",
     xlab = "Company Tenure (years)",
     ylab = "Salary (million dollars)",
     ylim = c(0,2.5))
```



The median salary and its distribution appears fairly similar for tenure less than 40 years. The low median salary for the highest tenure group comes as a surprise. Also, we would expect to see at least some rise in salary with a rise in tenure.

Age Versus Salary

Age data is clustered with nearly 50% of the data spread on a 10 year period. This makes drawing conclusions using a scatterplot difficult. Therefore we apply grouping to age data as well.

```
age_bins <- cut(cleanCEO$age,
                    c(0,40,50,60,70,90),
               labels = c('<40 yrs', '40-50 yrs', '50-60 yrs',
                          '60-70 yrs', '70-90 yrs'),
```

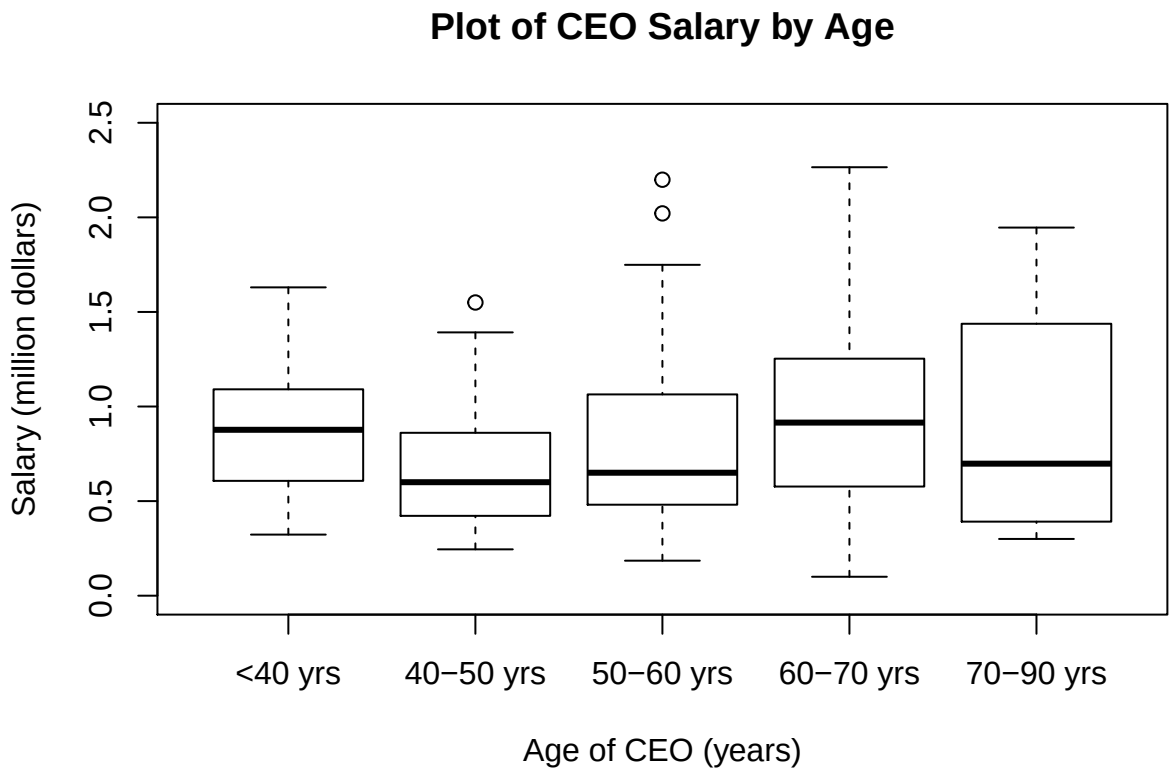
```
                    right = FALSE)

summary(age_bins)
```

```
##   <40 yrs 40-50 yrs 50-60 yrs 60-70 yrs 70-90 yrs
##         5        31        81        54         8
```

We create a boxplot of salary for the various tenure groups. The y axis has been limited to salary of $2.5 million.

```
boxplot(cleanCEO$salary/1000~ age_bins,
     main = "Plot of CEO Salary by Age",
     xlab = "Age of CEO (years)",
     ylab = "Salary (million dollars)",
     ylim = c(0,2.5))
```



For age in range 40 to 70 years, the median salary shows an increasing trend with age. The lower median salary for the highest age bin agrees with the same median for the highest company tenure bin.

Other Influences on Market Value

We analyse the impact of secondary variables on market value. Due to the large range of market values and clustering at the lower level values, we perform grouping on the market value variable. Due to the shape of the market value distribution we have more bins at the lower end of the market value spectrum.

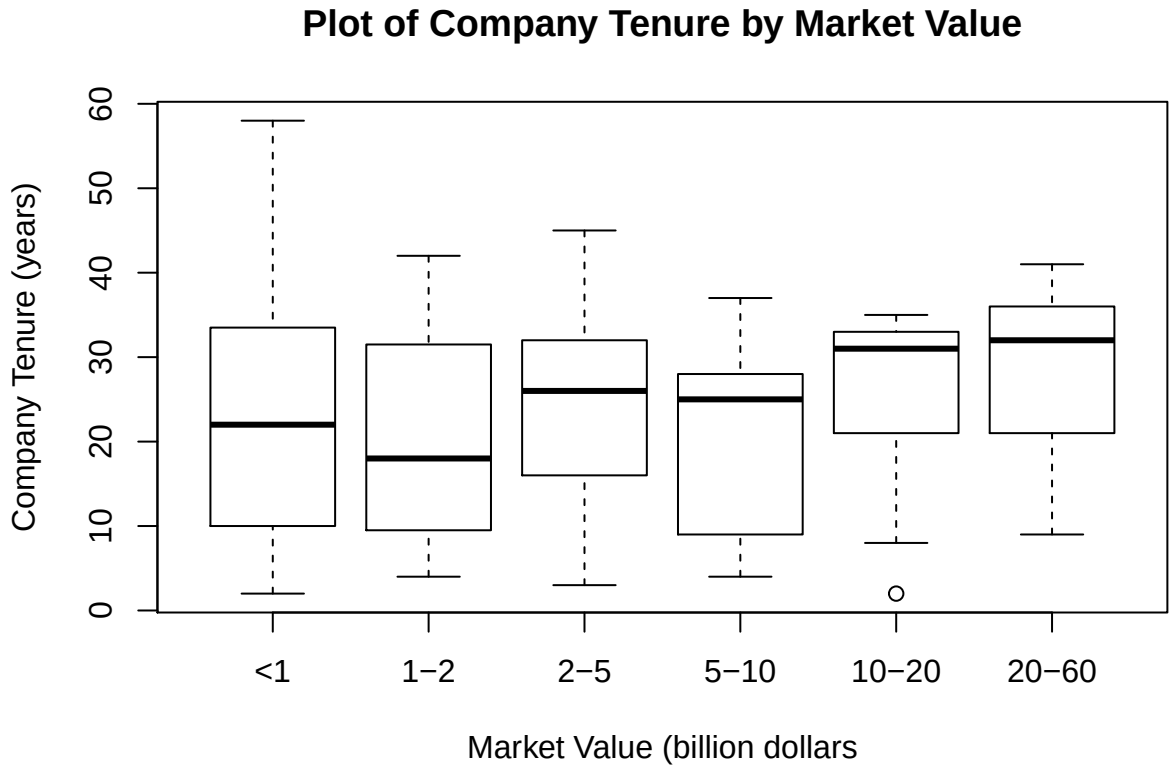
```
mktval_bins <- cut(cleanCEO$mktval/1000,
                        c(0,1,2,5,10,20,60),
                   labels = c('<1','1-2','2-5','5-10','10-20','20-60'))
summary(mktval_bins)
```

```
##    <1   1-2   2-5  5-10 10-20 20-60
```

```
##     79     35     35     15      9      6
```

Company Tenure Versus Market Value

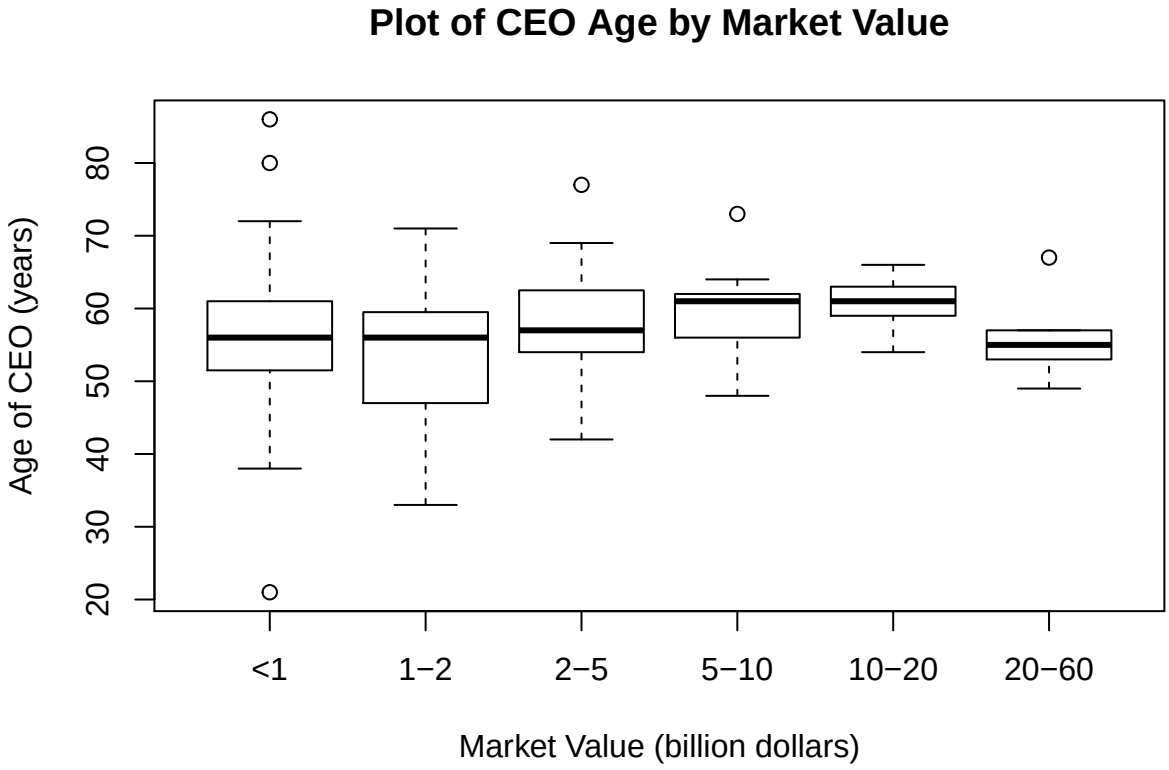
```
boxplot(cleanCEO$comten ~ mktval_bins,
     main = "Plot of Company Tenure by Market Value",
     ylab = "Company Tenure (years)",
     xlab = "Market Value (billion dollars")
```



The companies in different market value ranges tend to behave slightly differently. For companies with a market value below $2 billion, the median tenure is around 20 years. For companies with market values in $2 billion to $10 billion, the median tenure is around 26 years. The median company tenure is closer to 30 years for companies with market values above $10 billion.

CEO Age Versus Market Value

```
boxplot(cleanCEO$age ~ mktval_bins,
     main = "Plot of CEO Age by Market Value",
     ylab = "Age of CEO (years)",
     xlab = "Market Value (billion dollars)")
```



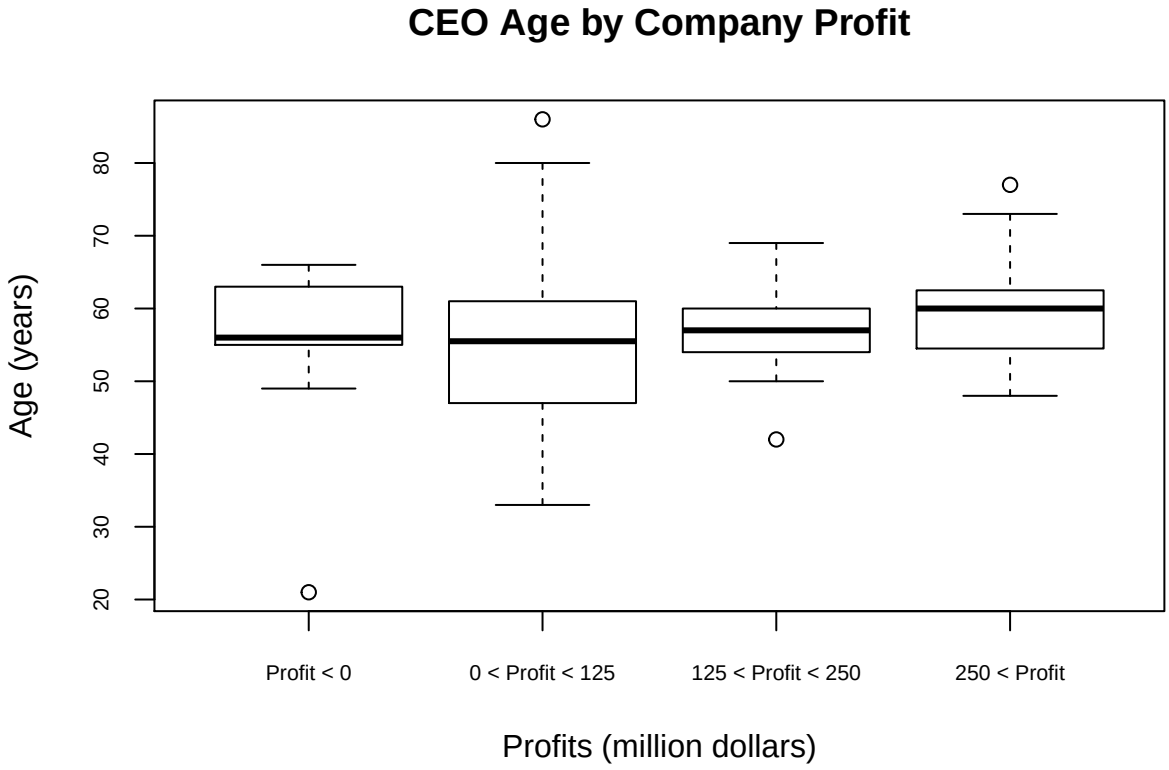
The median age of CEOs appear to increase with increase in market value of the company. The exception is the group with the largest market value, which has the lowest median age.

Other Influences on Profit

As with the previous two sections, this section assesses what other variables might influence profit, which would help understand better profit's primary relationship with salary.

CEO Age Versus Profit

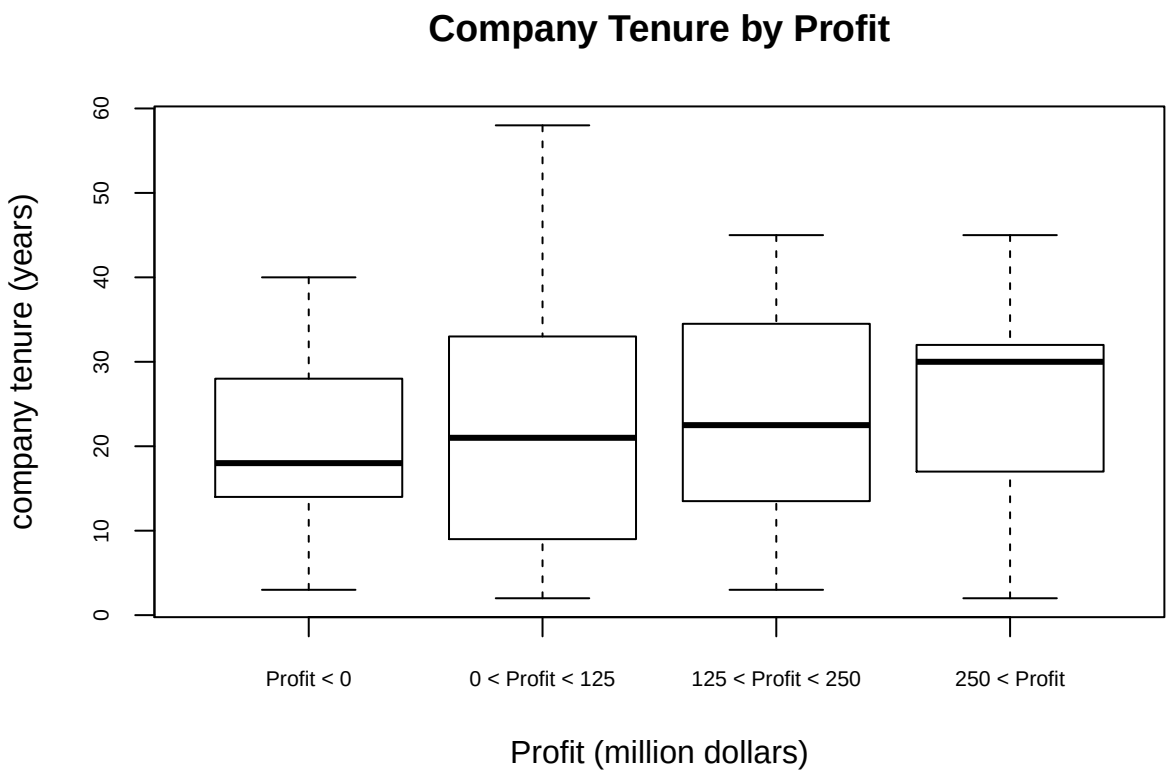
```
boxplot(age ~ profit_bin, data = cleanCEO, cex.axis = .7,
        main = "CEO Age by Company Profit",
        ylab = "Age (years)",
        xlab = "Profits (million dollars)"
        )
```



The graph above shows a very slight increasing trend in median age of CEOs with profit when profit is positive. But since most of our profits are in the \$0 to \$125 million range, the wide variablity of in this range age shouldn't be used to draw conclusions.

Company Tenure Versus Profit

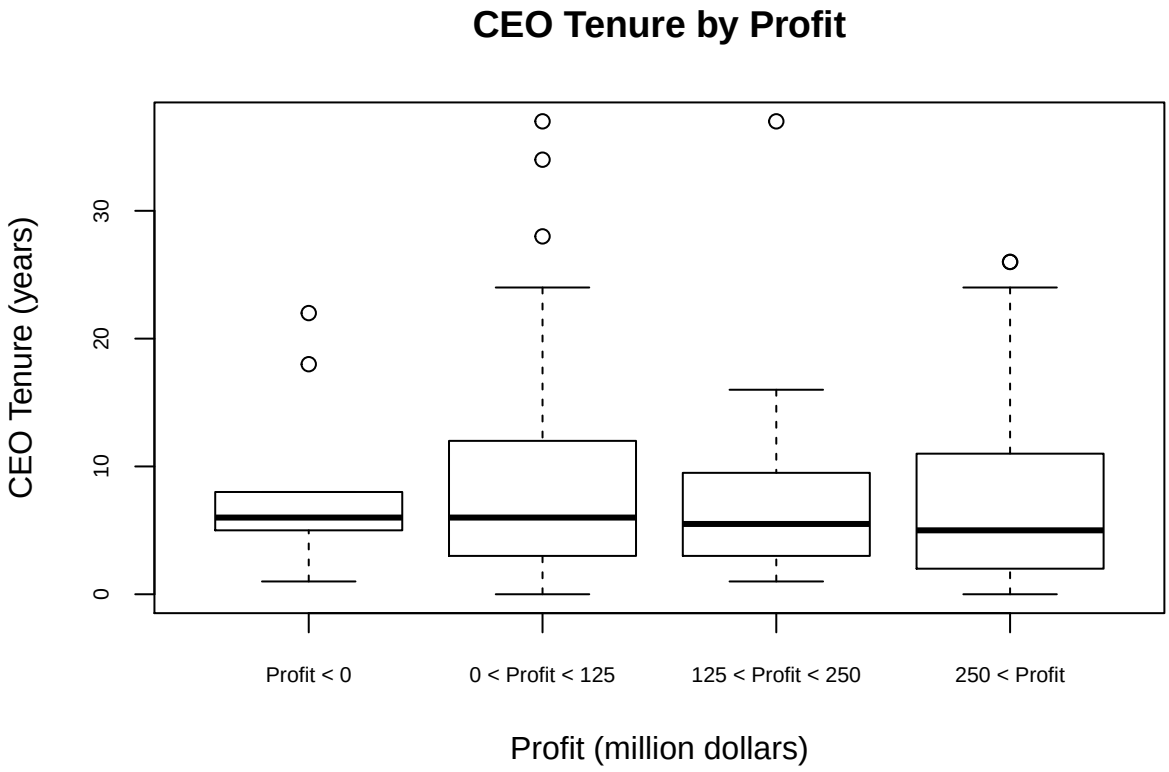
```
boxplot(comten ~ profit_bin, data = cleanCEO, cex.axis = .7,
        main = "Company Tenure by Profit",
        ylab = "company tenure (years)",
        xlab = "Profit (million dollars)")
```



From the graph above, we see an increasing trend in median company tenure as profit increases. This is prominent in companies with very high profits. One could argue that the number of years the CEO has been in a company does have a relationship with high company profit. Also, companies with negative profits also show to have CEOs with the lowest company tenure.

CEO Tenure Versus Profit

```
boxplot(ceoten ~ profit_bin, data = cleanCEO, cex.axis = .7,
        main = "CEO Tenure by Profit",
        ylab = "CEO Tenure (years)",
        xlab = "Profit (million dollars)")
```

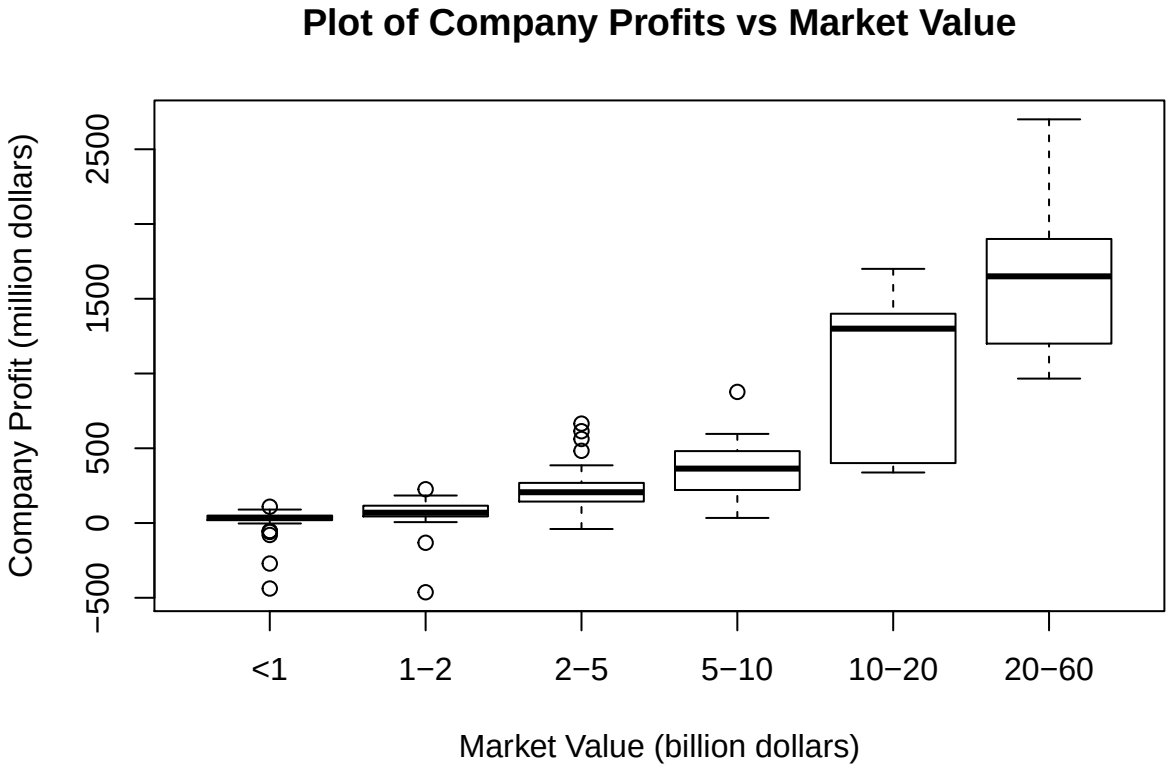


From the graph above, median CEO tenure seems to have a very slight decreasing trend with profit increase. It is interesting to note that companies with high profits have the lowest median CEO tenure, even a bit lower than compnies with negative profits. This may indicate a slightly high CEO turnover at companies with higher profits.

Confounding Effects

We have used profits and market values as proxies for company performance. We have explored the relationship between these individual variables and salary. However, the two variables themselves exhibit a positive relationship, as seen in the plot below. Therefore, at least some of the positive relationship of market value with salary may be explained by profits, and the same is true in where some of the positive relationship of profit with salary may be explained by market value.

```
boxplot(cleanCEO$profits ~ mktval_bins,
        main = "Plot of Company Profits vs Market Value",
        xlab = "Market Value (billion dollars)",
        ylab = "Company Profit (million dollars)")
```



Additionally, as seen in the section "Age Versus Salary," for 90% of the age data, or from 40 to 70 years, the median salary increases with age. Except for the 6 highest market values, median age shows an increase with market value. Therefore, at least some of the relationship between market value and salary is explained by CEO age.

Missing Values, Coding Errors, Etc.

In addition to the coding error found in market value and profits, where several data points were coded as -1, there is another apparent error in the data. For two data points, the company tenure is less than the CEO tenure. Since these two data points did not change the results of our analysis, we left them in our dataset.

Conclusion

An exploratory analysis of the CEO dataset was insightful in understanding the variables and establishing some preliminary relationships between the variables in the dataset. Some behaviors we found were positive skews in salary, market value, profits, and CEO tenure, as well as a negative skew in age. A sharp drop in frequency as well as data clustering were also prominent in the histograms for market value and profits. Outliers were present in the salary, market value, and profit variables. Of the key relationship variables, market value showed most data below $5 billion, profit data was mostly below $250 million, and most CEO salaries were below $1.4 million. Larger companies, in terms of market value, tend to have more CEOs who attended graduate school. Overall median salaries did not change based on whether CEOs attended college.

Company performance, in the form of profit and market value, does appear to hold a positive relationship with CEO salary. As both profit and market value increase, salary generally increases as well. However, the relationships appear to be non-linear and, with correlations only around 0.4, vary widely from CEO to CEO. As profit and market value have a positive relationship amongst themselves, it is not yet clear whether one drives CEO salary more than the other. Across the dataset, higher age is associated with higher median salaries and market values. Therefore, age might explain some of the positive relationship between

salary and market value. In other relationships, a clear influence on profit is company tenure of CEOs. The longer a CEO is with a company, the higher the profits the company tends to show. A similar, yet less clear, relationship held for market value, but CEO's with highest company tenure did not to make the highest salaries. From 0 to 10 years, CEO tenure and median salary show a positive relationship, while after 10 years, median salary levels off with increases CEO tenure. Companies with the highest profits have the lowest median CEO tenure yet companies with the longest median company tenures have the highest profits.