# Lab_1

Brooke Wheeler

1/12/2022

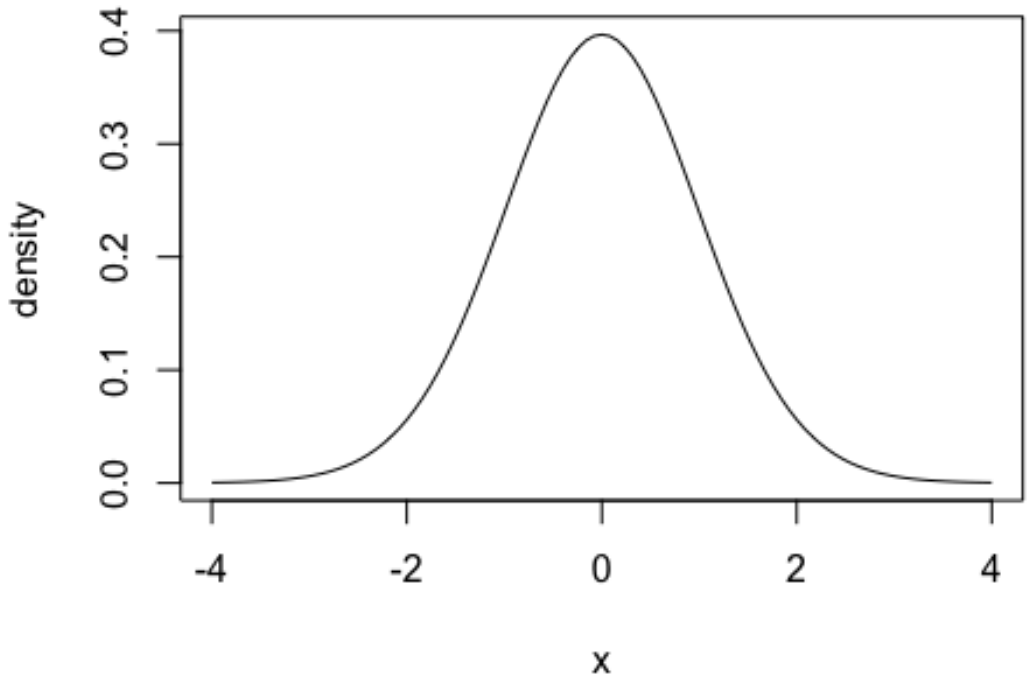
```
library(tidyverse)
```

```
## ── Attaching packages ─────────────────────────────────────── tidyverse
1.3.1 ──

## ✓ ggplot2 3.3.5     ✓ purrr   0.3.4
## ✓ tibble  3.1.6     ✓ dplyr   1.0.7
## ✓ tidyr   1.1.4     ✓ stringr 1.4.0
## ✓ readr   2.1.1     ✓ forcats 0.5.1

## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

## Part One Working with Probability Density Functions

### 1 Plot the density function for the t-distribution with 44 degrees of freedom, or t(44).

```
curve(dt(x,44), from = -4, to = 4, ylab = "density")
```




**2. Find the 75th percentile for the t-distribution with 44 degrees of freedom (df=44). Report your answer as t(p;df)=q, filling in values for p,df and q.**

```
qt(p= .75, df= 44)
```

```
## [1] 0.6801065
```

```
# t(.75;44)=.680
```

**3. Consider a t-distribution with df=44. 96% of values will fall between which two values of t (that is find the values that satisfy P(_____<t(44)<_____=0.96)? Report both answers as t(p;df)=q.**

```
qt(p= .02, df= 44)
```

```
## [1] -2.116438
```

```
# t(.02;44)=-2.11
```

```
qt(p= .98, df= 44)
```

```
## [1] 2.116438
```

```
#t(.98;44)=2.11

# P(-2.11 <t(44)< 2.11= .96)
```

### 4. Consider a t-distribution with df=44. Find P(t(44)≤-1.5).

```
pt(q=-1.5, df=44)
```

```
## [1] 0.07037815
```

### 5. Consider a t-distribution with df=44. Find P(t(44)≥2).

```
pt(q=2, df= 44)
```

```
## [1] 0.9741517
```

## Part 2: Descriptive Statistics and Basic Graphs

```
getwd()
```

```
## [1] "/Users/brookewheeler/Desktop/Regression/Labs"
```

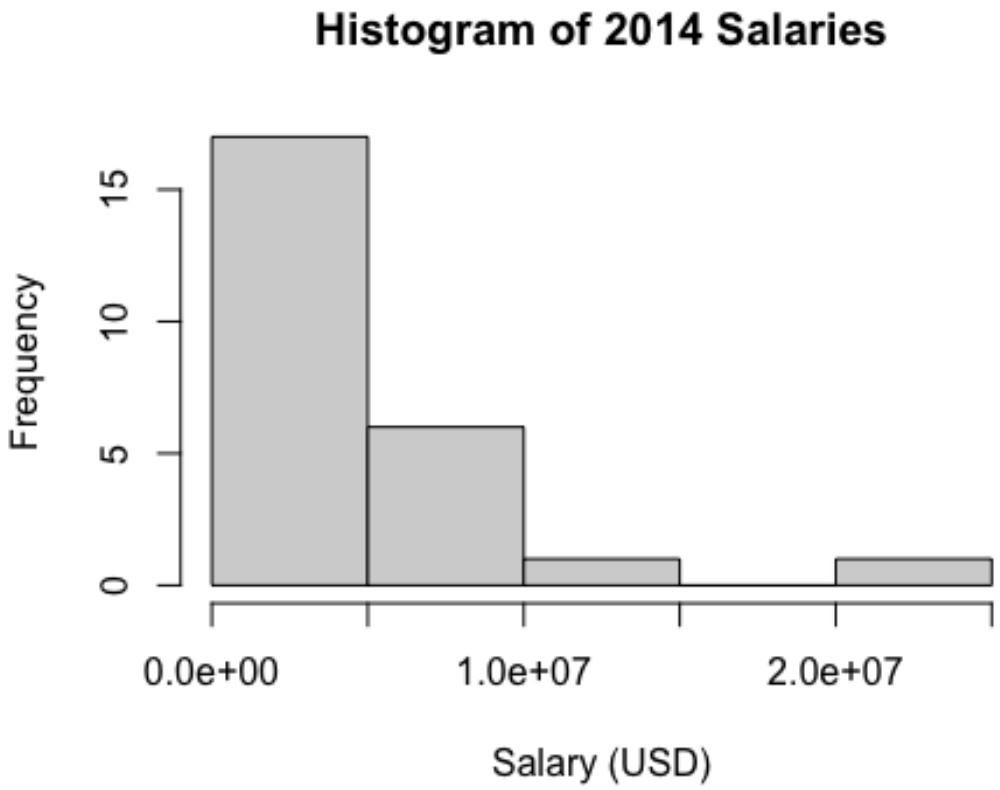
```
nationals <- read.table("../Data/nationalsdata2014.csv", header = TRUE, sep =
",")

nationals %>%
  select(Salary)  -> salary
salary
```

```
##      Salary
## 1  20571429
## 2  14000000
## 3   8600000
## 4   7500000
## 5   7200000
## 6   6500000
## 7   6500000
## 8   5875000
## 9   5000000
## 10  3975000
## 11  3500000
## 12  3450000
## 13  3000000
## 14  2700000
## 15  2150000
## 16  2095000
## 17  1675000
## 18  1350000
## 19  1250000
## 20   950000
## 21   900000
## 22   540850
```

```
## 23   506100
## 24   504300
## 25   501400
```

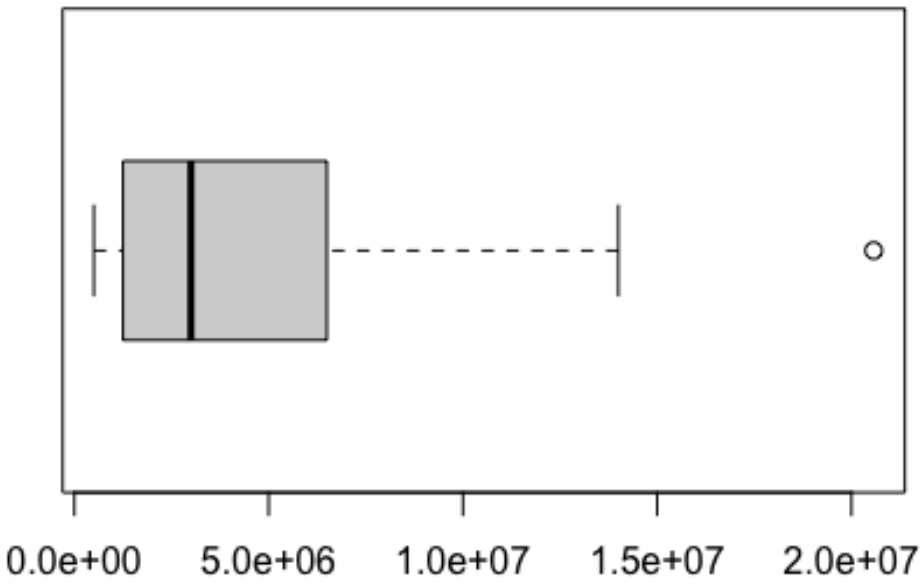
### histogram

```
summary(salary)

##      Salary
##  Min.   :  501400
##  1st Qu.: 1250000
##  Median : 3000000
##  Mean   : 4431763
##  3rd Qu.: 6500000
##  Max.   :20571429

hist(nationals$Salary, xlab = "Salary (USD)", main = "Histogram of 2014
Salaries")
```




### boxplot

```
boxplot(nationals$Salary, horizontal = TRUE, main = "Boxplot of 2014
Salaries")
```


Boxplot of 2014 Salaries
0.0e+00
5.0e+06
1.0e+07
1.5e+07
2.0e+07