STAT 511A: Homework 1

Kathleen Wendt

8/29/2019

Load packages

```
library(tidyverse)
library(magrittr)
```

Question 1 (Textbook 3.30)

Read data

```
lumber_data <- readr::read_csv("../data/OTT_Final/ASCII-comma/CH03/ex3-30.txt",
                               quote = "'")
```

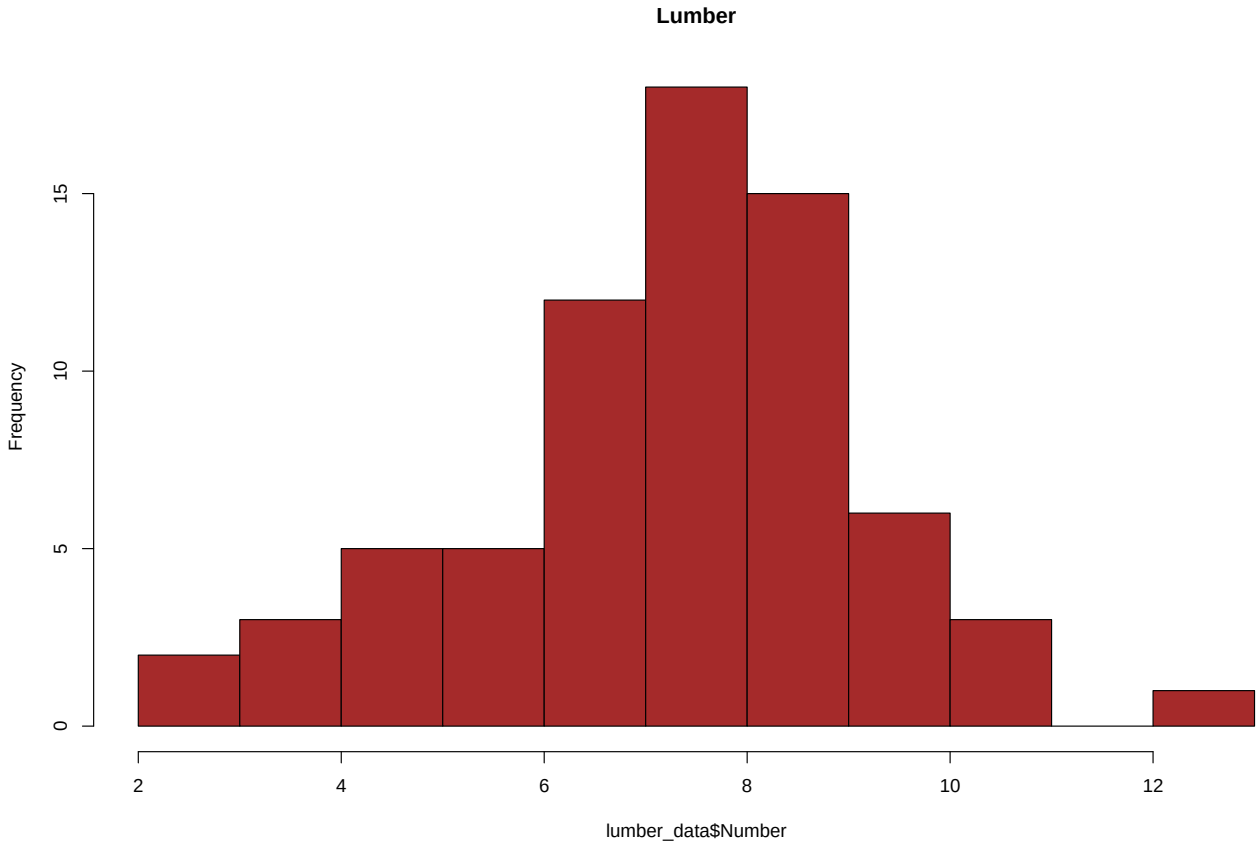
Data structure

```
tibble::glimpse(lumber_data)
```

```
## Observations: 70
## Variables: 1
## $ Number <dbl> 7, 8, 6, 4, 9, 11, 9, 9, 9, 10, 9, 8, 11, 5, 8, 5, 8, 8...
```

Distribution

```
lumber_hist <- hist(lumber_data$Number,
     col = "brown",
     main = "Lumber")
```



The distribution of the lumber is approximately normal.

Measures of Central Tendency

```
mean(lumber_data$Number)
```

```
## [1] 7.728571
```

```
median(lumber_data$Number)
```

```
## [1] 8
```

Question 2 (Textbook 3.7)

Read data

```
survival_data <- readr::read_csv("../data/OTT_Final/ASCII-comma/CH03/ex3-7.txt",
                                 quote = "'")
```

Data structure

```
tibble::glimpse(survival_data)
```

```
## Observations: 28
## Variables: 2
## $ StandardTherapy <dbl> 4, 14, 29, 6, 15, 2, 6, 13, 24, 16, 12, 21, 10...
## $ NewTherapy      <dbl> 5, 17, 27, 9, 20, 15, 14, 18, 29, 19, 10, 33, ...
```

Measures of Central Tendency

Standard Therapy

```
mean(survival_data$StandardTherapy)
```

```
## [1] 15.67857
```

```
sd(survival_data$StandardTherapy)
```

```
## [1] 9.630405
```

New Therapy

```
mean(survival_data$NewTherapy)
```

```
## [1] 20.71429
```

```
sd(survival_data$NewTherapy)
```

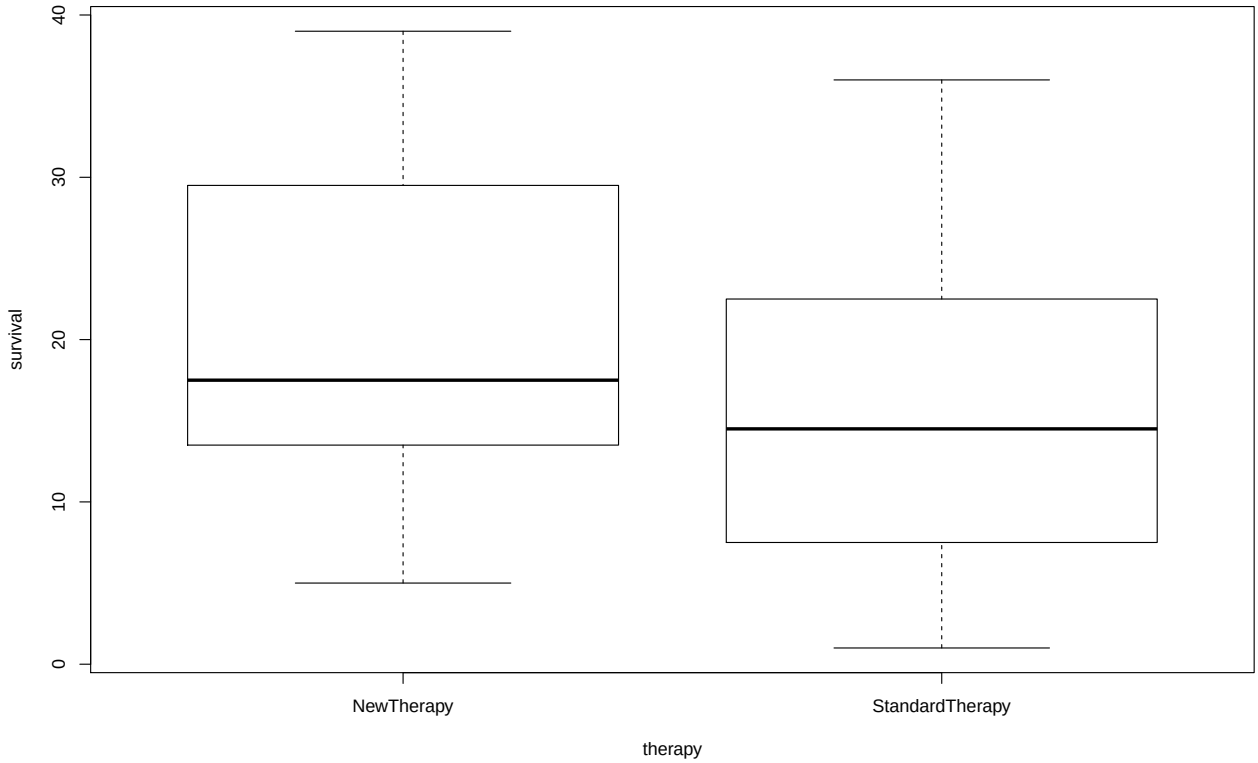
```
## [1] 9.808753
```

Boxplots

Convert data

```
survival_data <- survival_data %>%
  tidyr::gather(key = "therapy",
                value = "survival")
```

```
boxplot(survival ~ therapy, data = survival_data)
```


survival
0
10
20
30
40
NewTherapy
StandardTherapy
therapy