

# lab 1

Mason Wong

15th March 2022

```
library(tidyverse)
```

Note that when loading the data, `rugby.txt` should be in the same directory as the R Markdown document.

You can change the `html_document` to `pdf_document` for the **pdf** output but you need the latex installed. If it does not work for you then you should open your **html** output and print it to a pdf. Be sure you know how to do this before the assessment. Ask your lab tutor for help if you do not know how to get it done.

```
rugby = read.table("rugby.txt", header = TRUE)
rugby[1:3, ]
```

```
##   Game Time
## 1    A 39.2
## 2    A  2.7
## 3    A  9.2
```

```
head(rugby)
```

```
##   Game Time
## 1    A 39.2
## 2    A  2.7
## 3    A  9.2
## 4    A 14.6
## 5    A  1.9
## 6    A 17.8
```

```
head(rugby, 2)
```

```
##   Game Time
## 1    A 39.2
## 2    A  2.7
```

```
head(rugby$Game)
```

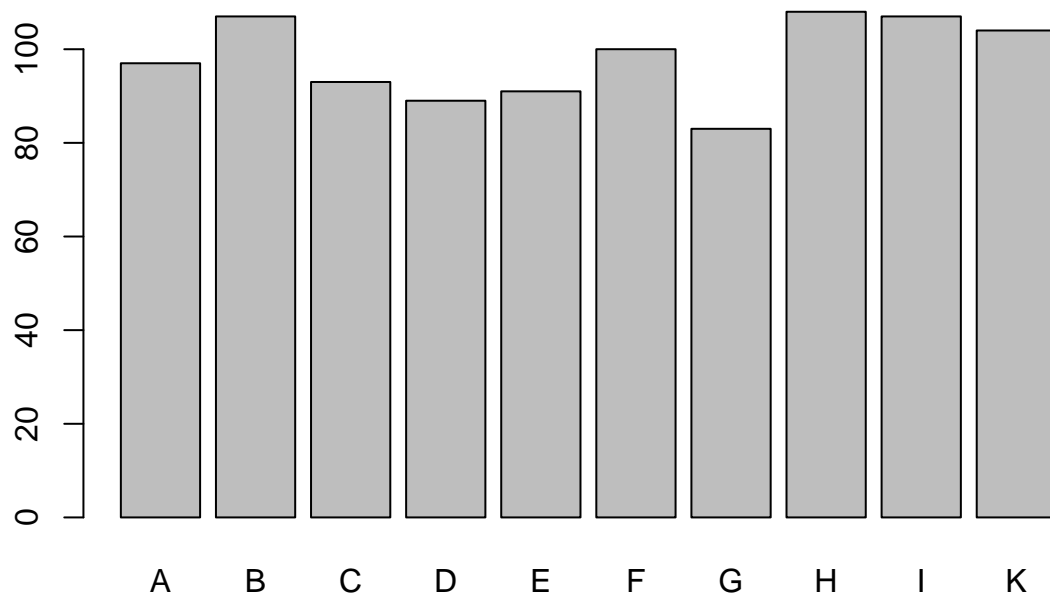
```
## [1] "A" "A" "A" "A" "A" "A"
```

```
(tt <- table(rugby$Game))
```

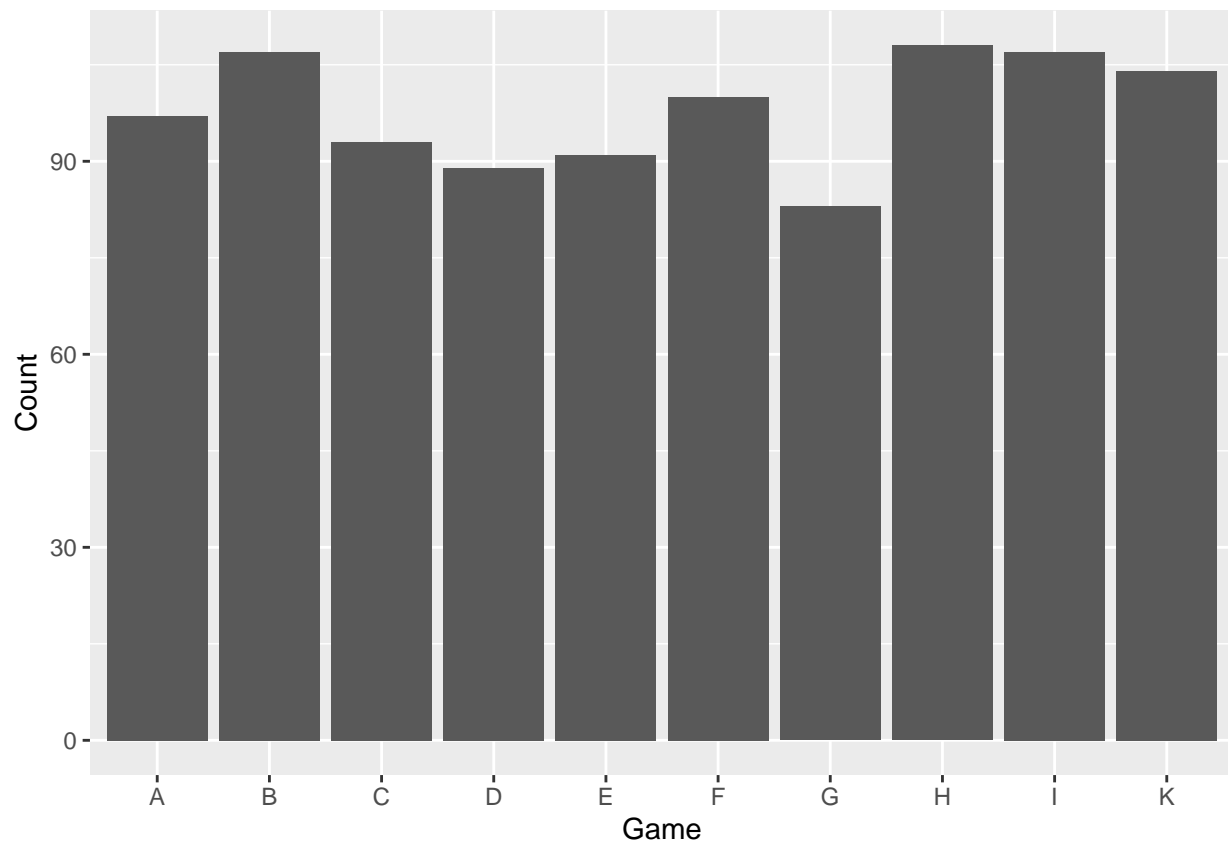
```
##
##   A   B   C   D   E   F   G   H   I   K
## 97 107 93 89 91 100 83 108 107 104
```

Game H had the most separate passages of play.

```
barplot(tt)
```

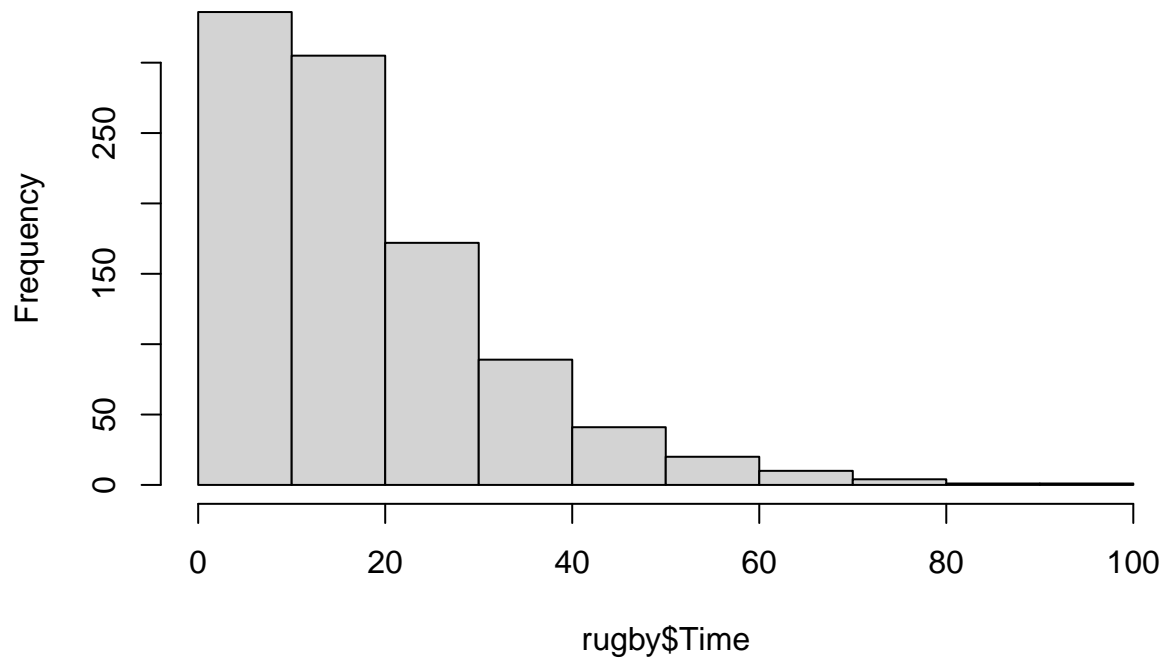


```
as.data.frame(tt) %>% # change to data frame first
  rename(Game=Var1, Count=Freq) %>% # rename the column names
  ggplot(aes(Game, Count)) +
  geom_bar(stat="identity")
```



```
hist(rugby$Time)
```

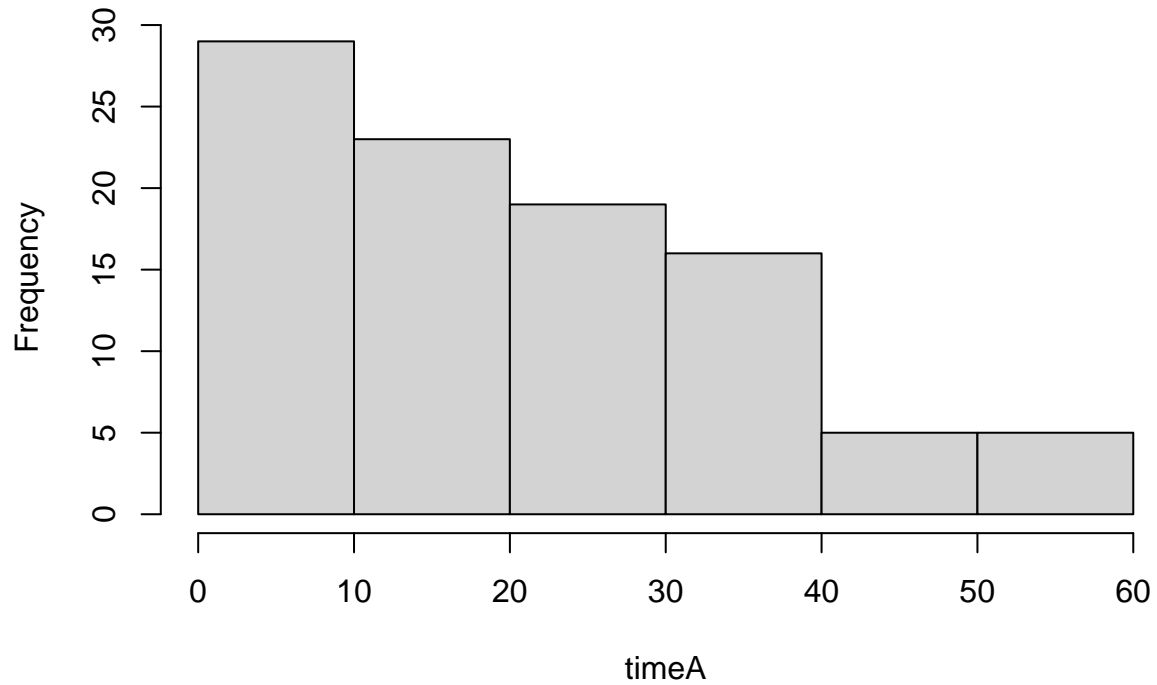
## Histogram of rugby\$Time



The histogram is right skewed and does not appear to be normally distributed.

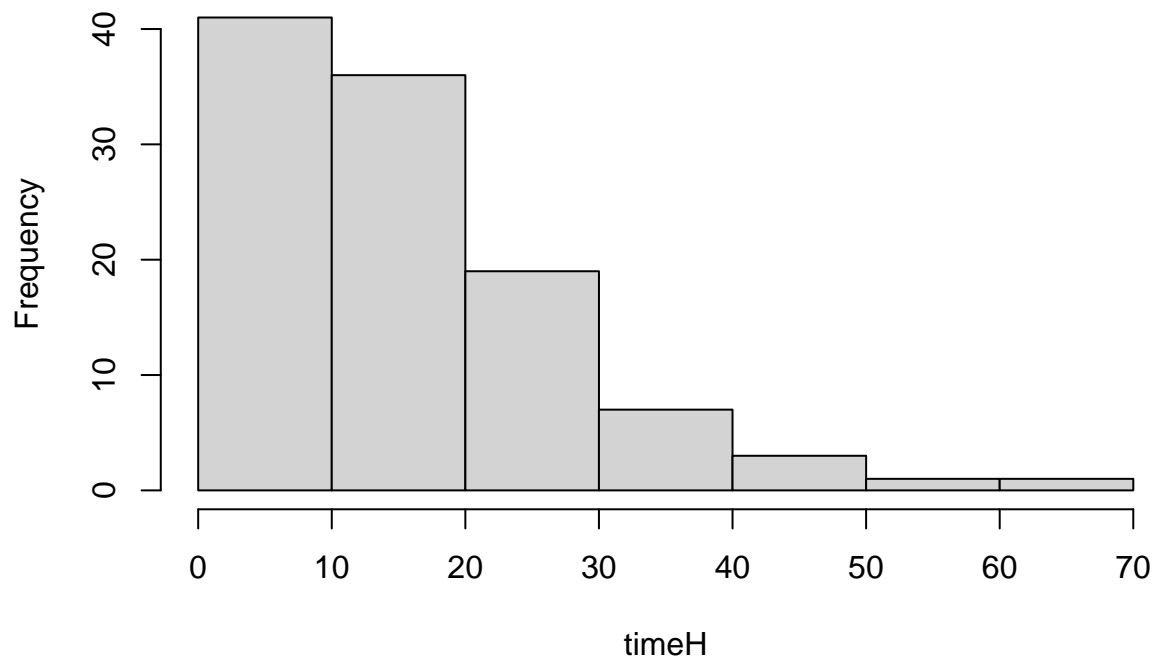
```
timeA <- rugby %>%  
  filter(Game=="A") %>%  
  pull(Time)  
timeH <- rugby %>%  
  filter(Game=="H") %>%  
  pull(Time)  
hist(timeA)
```

**Histogram of timeA**



```
hist(timeH)
```

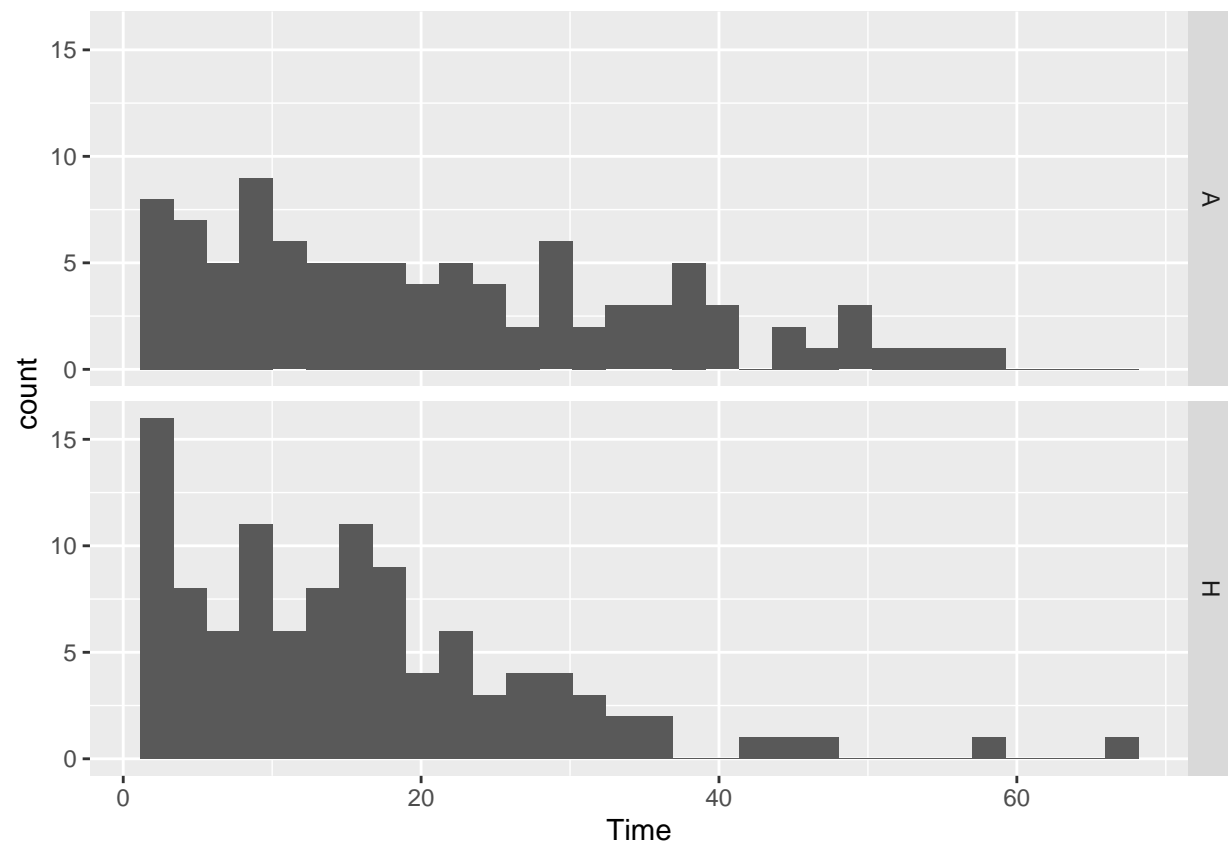
**Histogram of timeH**



```
rugby %>%  
  filter(Game %in% c("A", "H")) %>%  
  ggplot(aes(Time)) +  
  geom_histogram() +
```

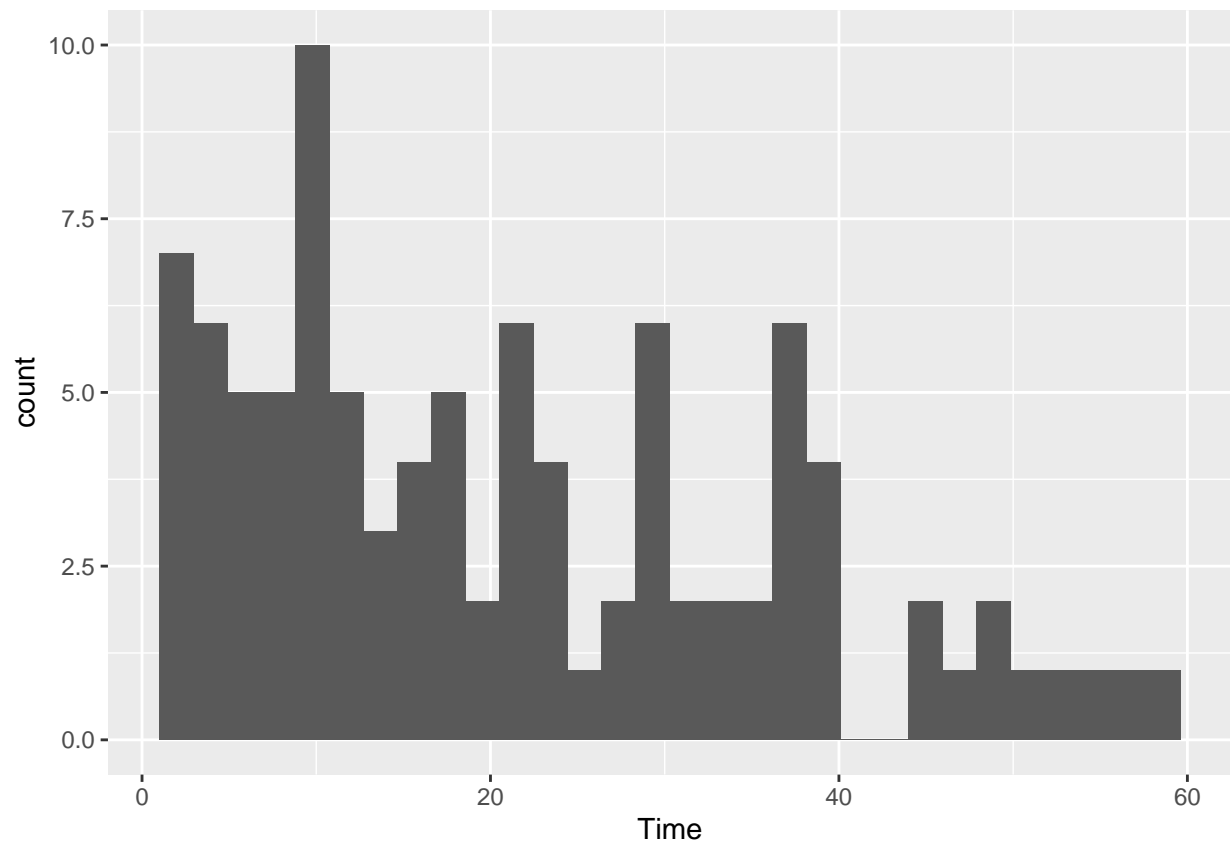
```
facet_grid(Game ~ .)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



You could alternatively code as follows (credit to Connor Smith) or any other way to achieve separate histograms.

```
rugby_plot <- ggplot(rugby, aes(Time)) + geom_histogram()  
rugby_plot %>% subset(x = rugby, Game == "A")
```



```
rugby_plot %+% subset(x = rugby, Game == "H")
```

