

## R Module 1 Rubric

This R module is pretty simple; there aren't that many difficult questions. Students should mostly just be able to load and run R Studio.

```
# Load data
library(MASS)

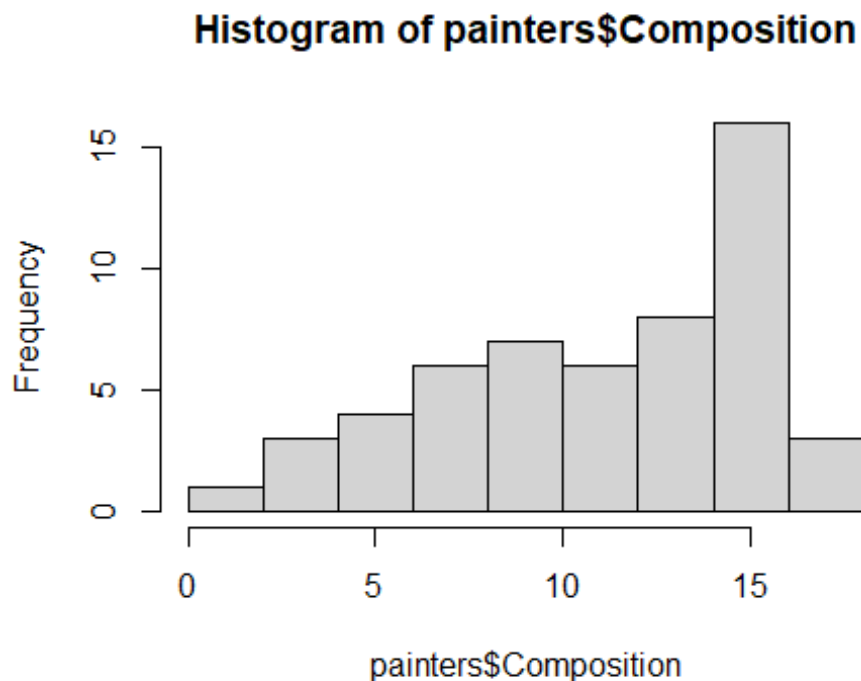
head painters)
```

**Use R to produce the following simple data summaries and visualizations commonly used for qualitative data:**

### Question 1:

*A histogram of the Composition variable from the painters dataset, using the hist() function.*

```
hist(painters$Composition)
```

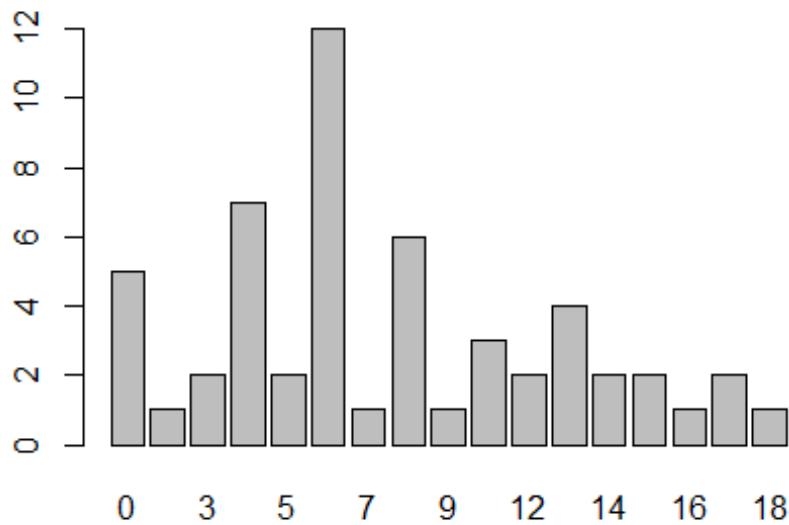


### Question 2:

*A bar chart of the Expression variable, using the barplot() function.*

- \*Tip: use `?barplot()` to see which argument you need to change to get a horizontal bar plot\* Also, wrap the `painters$Expression` with the `table()` function, then use the `barplot()` function

```
barplot(table(painters$Expression))
```

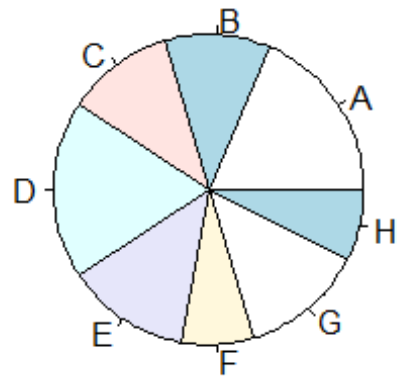


### Question 3:

*A pie chart of the `Schools` variable, using the `pie()` function.*

- \*Hint: you need to wrap the `pie()` function around the `table()` function, because you need **counts** of each school, rather than just their label\*

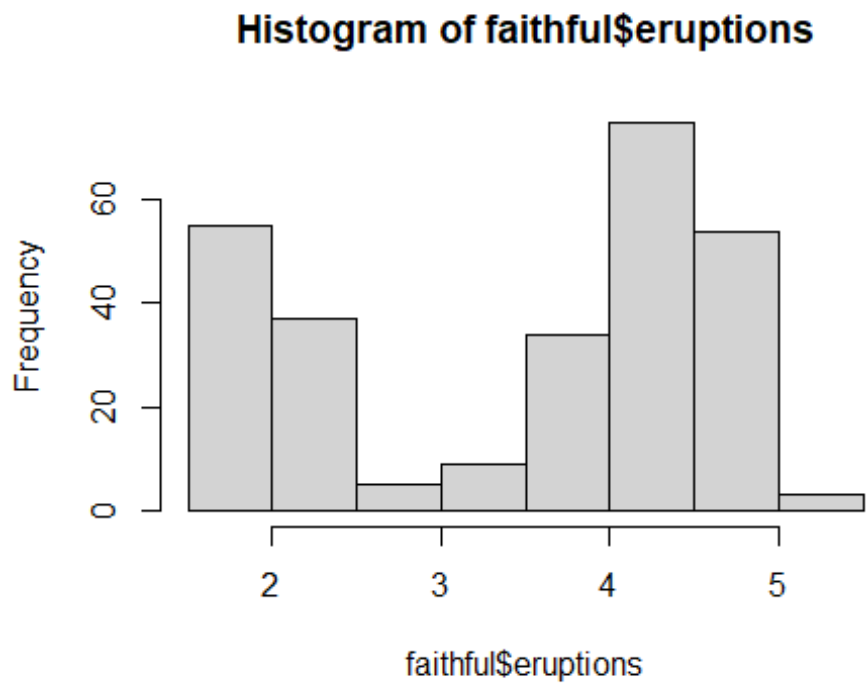
```
pie(table(painters$School))
```



#### Question 4:

*A histogram for the eruptions. R will automatically decide how to group your observations. Remember to plot **only** the eruptions variable.*

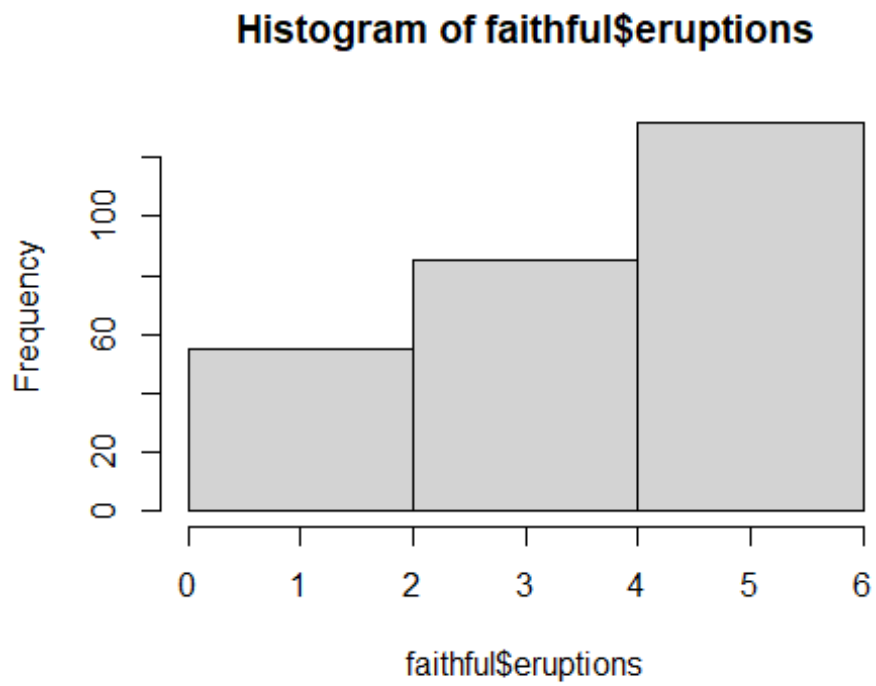
```
hist(faithful$eruptions)
```



#### Question 5:

*Another histogram of eruptions, but specify 2 data classes instead of the default values.*

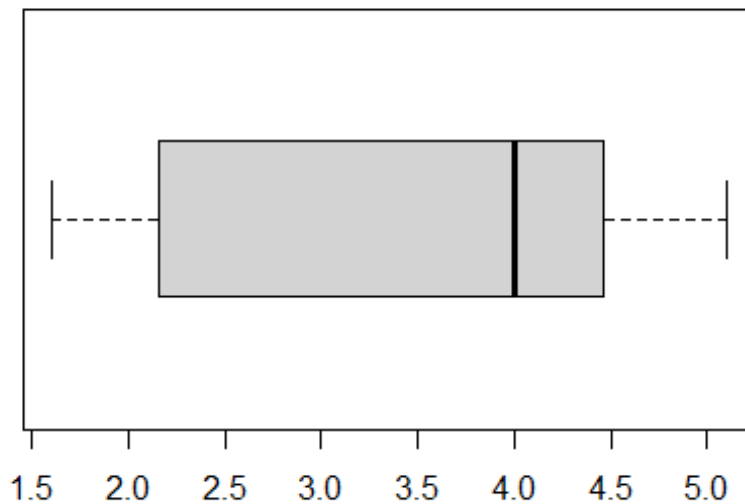
```
hist(faithful$eruptions, nclass = 2)
```



#### Question 6:

*A boxplot of eruptions that is horizontal, rather than vertical.*

```
boxplot(faithful$eruptions,  
        horizontal = T)
```



### Question 7:

Repeat steps 4-6 (histogram, histogram with `nbreaks = 2`, and horizontal boxplot) on a dataset of your own choosing and interpret the data. You can also use functions such as `mean`, `median`, `quantile`, etc. to help interpret your data. Give each of your new plots x- and y-axis labels and a title

```
mydata <- mtcars$mpg

# Students don't have to use the par(mfrow); I've just used it to tidy the
# plots
# a bit.

par(mfrow = c(2, 2))

hist(mydata,
      xlab = "MPG",
      ylab = "Frequency",
      main = "Histogram of MPG")

hist(
  mydata,
  nclass = 2,
  xlab = "MPG",
  ylab = "Frequency",
  main = "Histogram of MPG")
```

```
)
boxplot(mydata,
        horizontal = TRUE,
        xlab = "MPG",
        main = "Boxplot of MPG")

par(mfrow = c(1, 1))
```

