Set your working directory using the function `setwd()` to where your data is stored and download the data directly from the website. (https://wimr-genomics.vip.sydney.edu.au/AMED3002/data/frmgham.csv

- Set your working directory using the function `setwd()` to where your data is stored and download the data directly from the website. (https://wimr-genomics.vip.sydney.edu.au/AMED3002/data/frmgham.csv)
- Check the size of your data. Think about what the number of rows actually means.

Code

- How does R classify your R object `heartData`?

Hide

```
class(heartData)
```

- How does R classify each of the variables in your R object `heartData`?

Hide

```
str(heartData)
```

- You may like to redefine the classification of a variable differently. This component is more critical later in the course.

Hide

```
sex <- heartData$SEX
sexC <- as.character(sex)
sexF <- factor(sexC, levels = c(1, 2), labels = c("Men", "Women"))
class(sex)
class(sexC)
class(sexF)
```

- Produce numerical summaries

Hide

```
head(heartData$SYSBP)
summary(heartData$SYSBP)
table(heartData$SEX)
```

- Try using some 'base R' graphics.

Hide

```
hist(heartData$SYSBP, prob = T)
boxplot(heartData$SYSBP)
```

- Customise the plots.

Hide

```
hist(heartData$SYSBP, freq = FALSE, main = "Histogram", ylab = "Probabilities", col = "green")
boxplot(heartData$SYSBP, horizontal = TRUE, col = "red")
```

- Looking at relationship between Age and Sex.

Hide

```
boxplot(SYSBP ~ SEX, data = heartData)
```

- Try this using ggplot.

Hide

```
library(tidyverse)
## Graphical summary of the variable 'Age'
ggplot(heartData, aes(x = 1, y = SYSBP)) + geom_boxplot()

## Relatoinship between Age and Sex.
ggplot(heartData, aes(x = factor(SEX), y = SYSBP)) + geom_boxplot()

## Adding some colors
ggplot(heartData, aes(x = factor(SEX), y = SYSBP, fill = SEX)) + geom_boxplot()

ggplot(heartData, aes(x = factor(SEX), y = SYSBP, col = SEX)) + geom_boxplot()
```

- (Extension) Generate a violin plot, does it take longer to plot ?