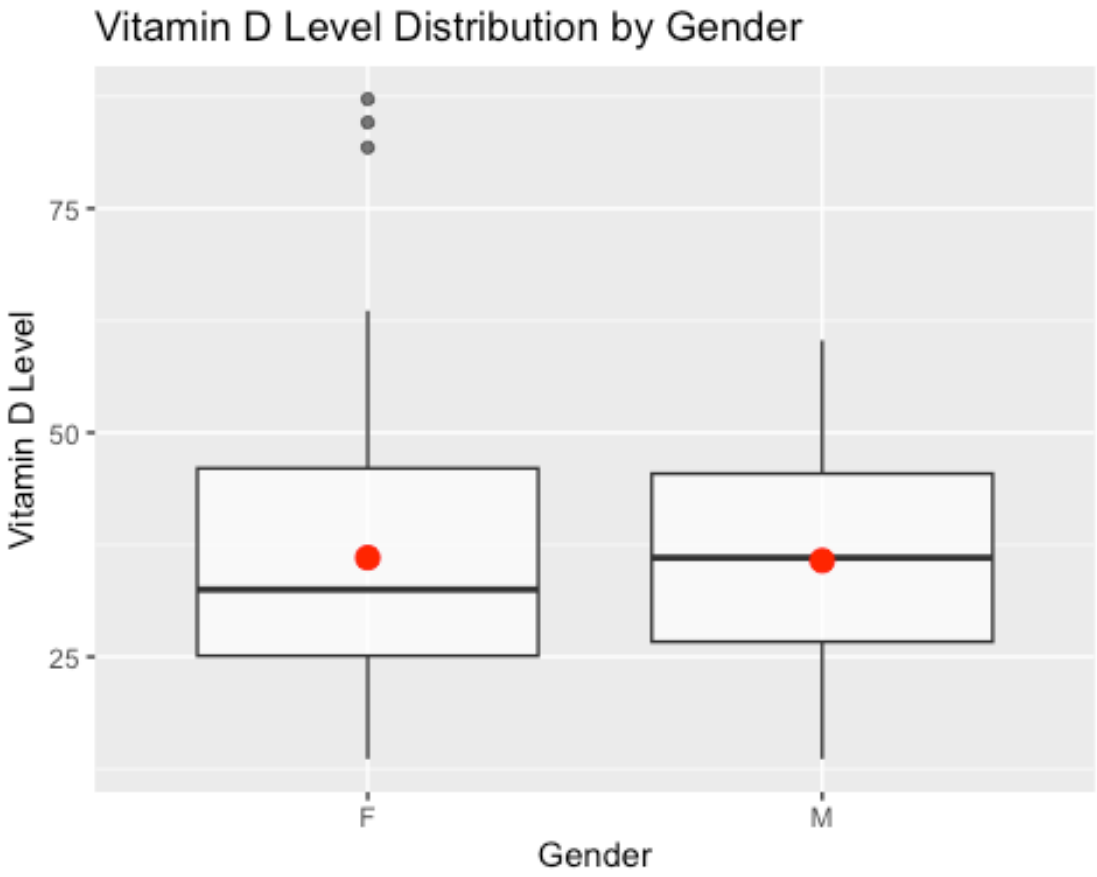
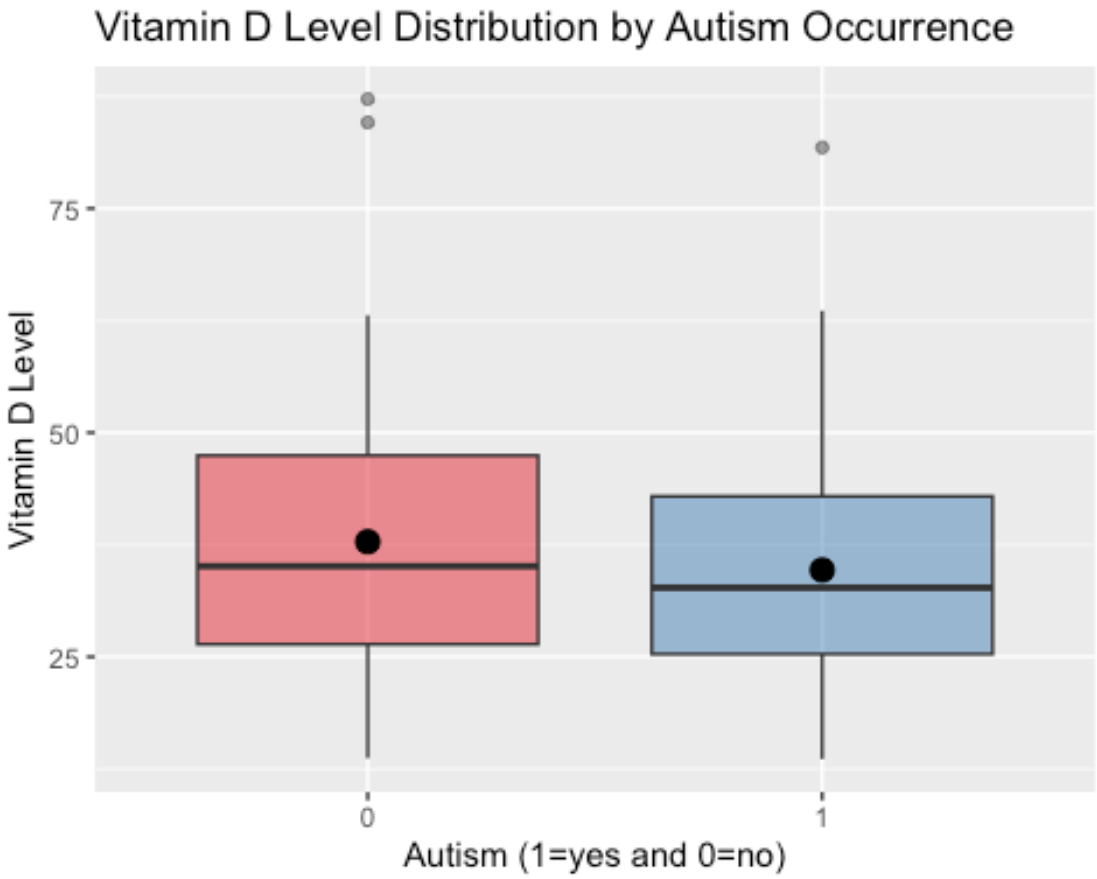
Module 2 Report: Emalee Schuler

Boxplots

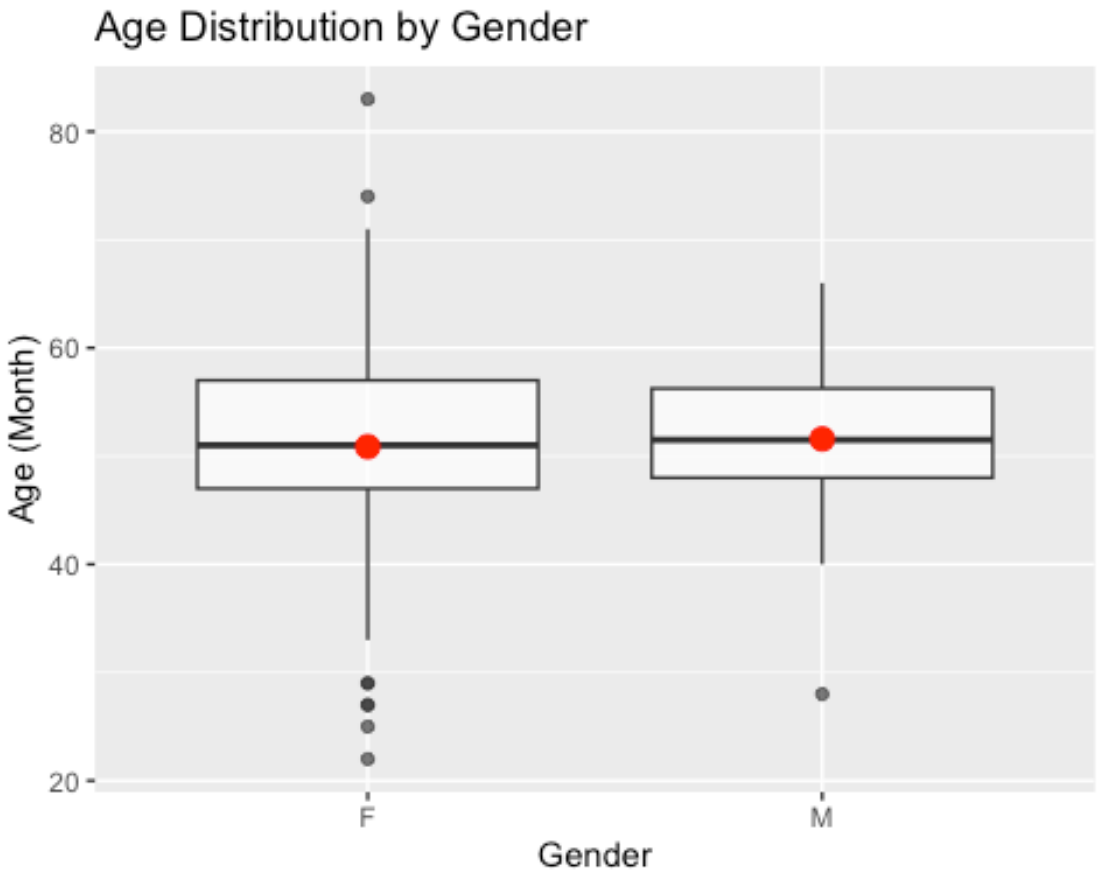
```
library(ggplot2)
library(dplyr)


Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union

data1_LSC598 <- read.csv("~/Downloads/data1_LSC598.txt", sep="")

data1_LSC598 <- data1_LSC598 %>% filter(!is.na(vitD_level))


ggplot(data1_LSC598, aes(x = gender, y = vitD_level)) +
    geom_boxplot(alpha = 0.7) +
  stat_summary(fun = mean, geom = "point", shape = 20, size = 5, color =
"red", fill = "red") +
    theme(legend.position = "none") +
     scale_fill_brewer(palette = "Set1") + ggtitle("Vitamin D Level
Distribution by Gender") +
  xlab("Gender") +
    ylab("Vitamin D Level")
```



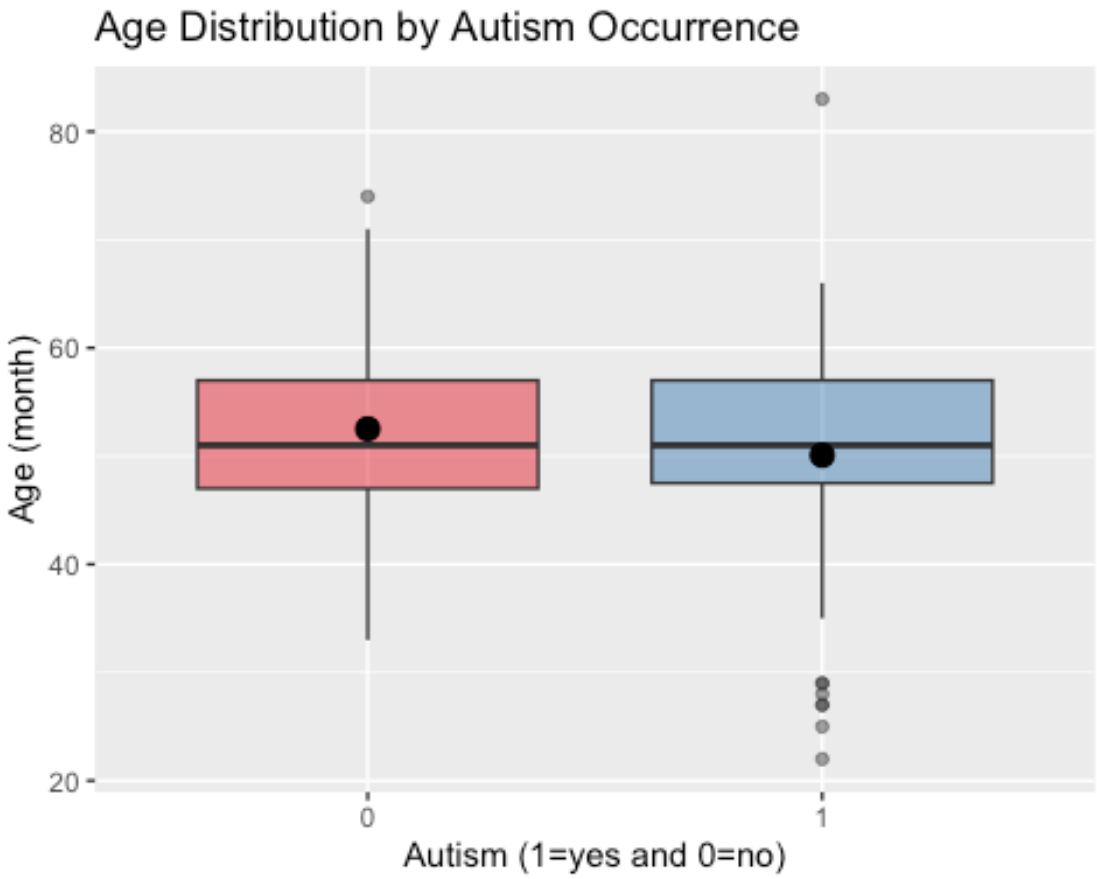

```
ggplot(data1_LSC598, aes(x = factor(group), y = vitD_level)) +
    geom_boxplot(aes(fill = factor(group)), alpha = 0.5) +
    stat_summary(fun = mean, geom = "point", shape = 20, size = 5, color =
"black", fill = "red") +
    theme(legend.position = "none") +
    ggtitle("Vitamin D Level Distribution by Autism Occurrence") +
    scale_fill_brewer(palette = "Set1") +
    xlab("Autism (1=yes and 0=no)") +
    ylab("Vitamin D Level")
```




```
ggplot(data1_LSC598, aes(x = gender, y = age_month)) +
    geom_boxplot(alpha = 0.7) +
  stat_summary(fun = mean, geom = "point", shape = 20, size = 5, color =
"red", fill = "red") +
    theme(legend.position = "none") +
     scale_fill_brewer(palette = "Set1") + ggtitle("Age Distribution by
Gender") +
  xlab("Gender") +
    ylab("Age (Month)")
```




```
ggplot(data1_LSC598, aes(x = factor(group), y = age_month)) +
    geom_boxplot(aes(fill = factor(group)), alpha = 0.5) +
    stat_summary(fun = mean, geom = "point", shape = 20, size = 5, color =
"black", fill = "red") +
    theme(legend.position = "none") +
    ggtitle("Age Distribution by Autism Occurrence") +
    scale_fill_brewer(palette = "Set1") +
    xlab("Autism (1=yes and 0=no)") +
    ylab("Age (month)")
```



T-Tests

```
t.test(data1_LSC598$vitD_level, mu = 0)
```

```

    One Sample t-test

data:  data1_LSC598$vitD_level
t = 21.418, df = 85, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 32.62538 39.30253
sample estimates:
mean of x
 35.96395
```

```
t.test(data1_LSC598$age_month, mu = 0)
```

```

    One Sample t-test

data:  data1_LSC598$age_month
t = 43.315, df = 85, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
```

```
95 percent confidence interval:
 48.72556 53.41398
sample estimates:
mean of x
 51.06977
```

For the "age" variable:

- Hypothesis: The mean age is not equal to 0 ($\mu \neq 0$).
- Null Hypothesis (H0): The mean age is equal to 0 ($\mu = 0$).
  - Age (in months) mean is equal to 51.06977. Therefore, we reject the null hypothesis for 'age' because the mean is not equal to 0.

For the "vitD_level" variable:

- Hypothesis: The mean vitamin D level is not equal to 0 ($\mu \neq 0$).
- Null Hypothesis (H0): The mean vitamin D level is equal to 0 ($\mu = 0$).
  - Vitamin D level mean is equal to 35.96395. Therefore, we reject the null hypothesis for 'vitD_level' because the mean is not equal to 0.