#Respondent Survey Time Averages

#Average with outliers includes
> time <- (X2022FSULibSurveyDataFinalVersion$`Duration (in seconds)`)
> minutetime <- time/60
> mean(minutetime)
[1] 26.74554

#The average across all respondents, regardless of whether or not they finished, and whether they left their screen for an extended amount of time or not, is 26.75 minutes.

#It should be noted that the *smallest* outlier for survey duration (in seconds) was 1474 seconds long, or 24.5667 minutes. As this is incredibly lengthy in comparison to the majority of responses that took less than 500 seconds, I have done an outlier replace function to get an average duration time for surveys that took less than 1000 seconds to get an idea of how long the majority of respondents took.
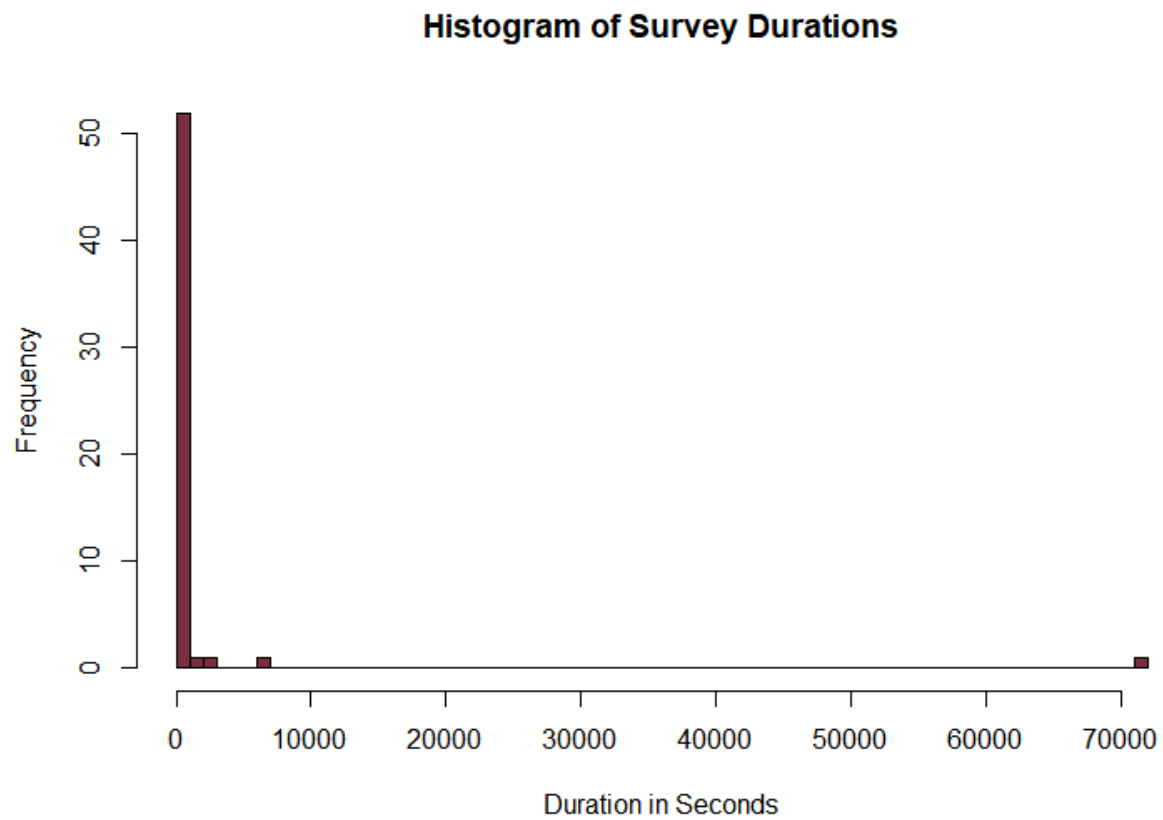> library(data.table)
data.table 1.14.2 using 4 threads (see ?getDTthreads). Latest news: r-datatable.com
> outlierReplace = function(dataframe, cols, rows, newValue = NA) {
+   if (any(rows)) {
+   set(dataframe, rows, cols, newValue)
+   }
+ }
> outlierReplace(X2022FSULibSurveyDataFinalVersion, "Duration (in seconds)", which(X2022FSULibSurveyDataFinalVersion$`Duration (in seconds)` > 1000, NA))
> View(X2022FSULibSurveyDataFinalVersion)
> mean(na.omit(X2022FSULibSurveyDataFinalVersion$`Duration (in seconds)`))
[1] 151.9808
> mean(na.omit(X2022FSULibSurveyDataFinalVersion$`Duration (in seconds)`))/60
[1] 2.533013

After accounting for outliers, the average survey response time was 2.53 minutes, which fits within the 2-5 minute survey response time.
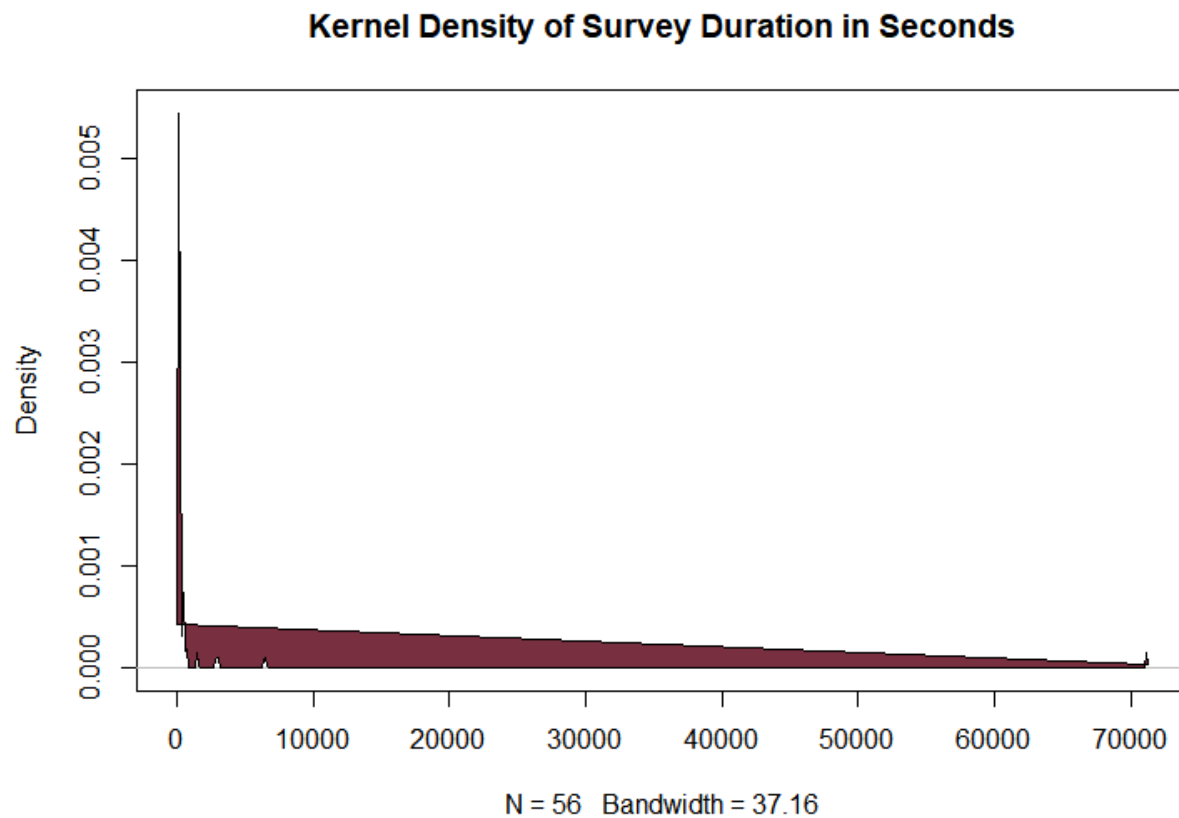
#After restarting R, I made these graphs
#Histogram of the survey data in seconds (with outliers included)
> h <-hist(time, breaks = 100, col="#782F40", xlab = "Duration in Seconds", main = "Histogram of Survey Durations")
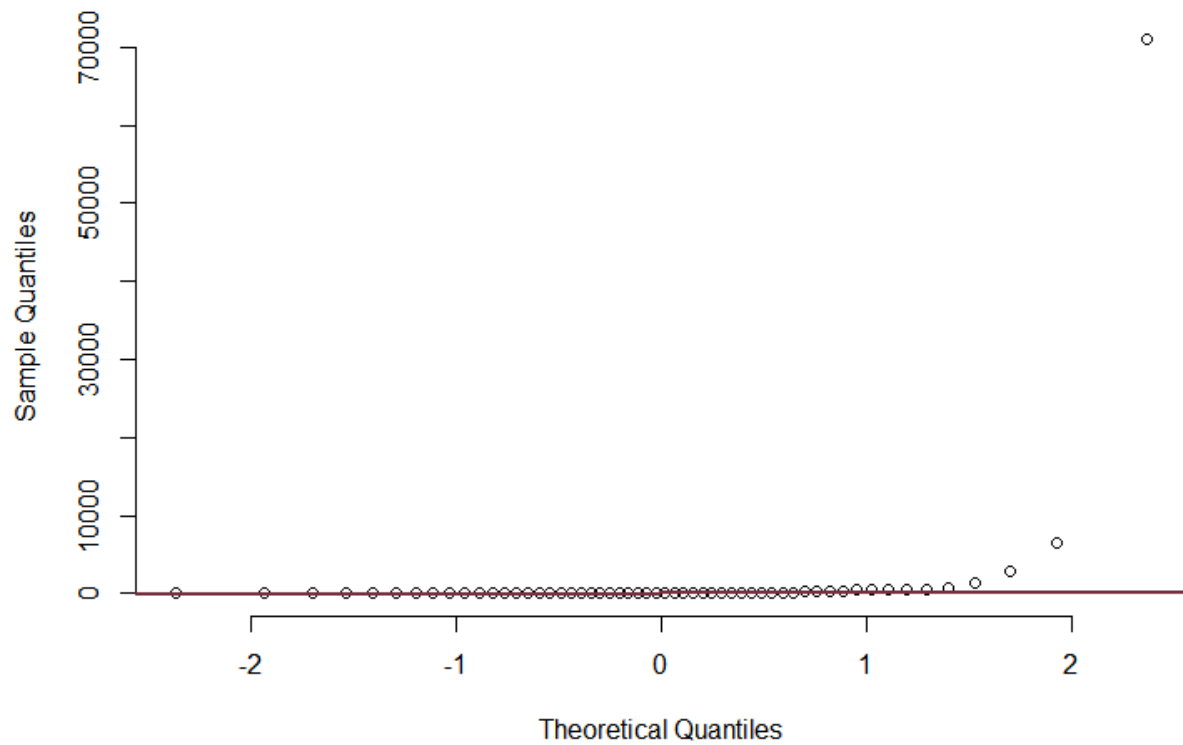

Histogram of Survey Durations

Kernel Density of Survey Duration in Seconds

N = 56   Bandwidth = 37.16
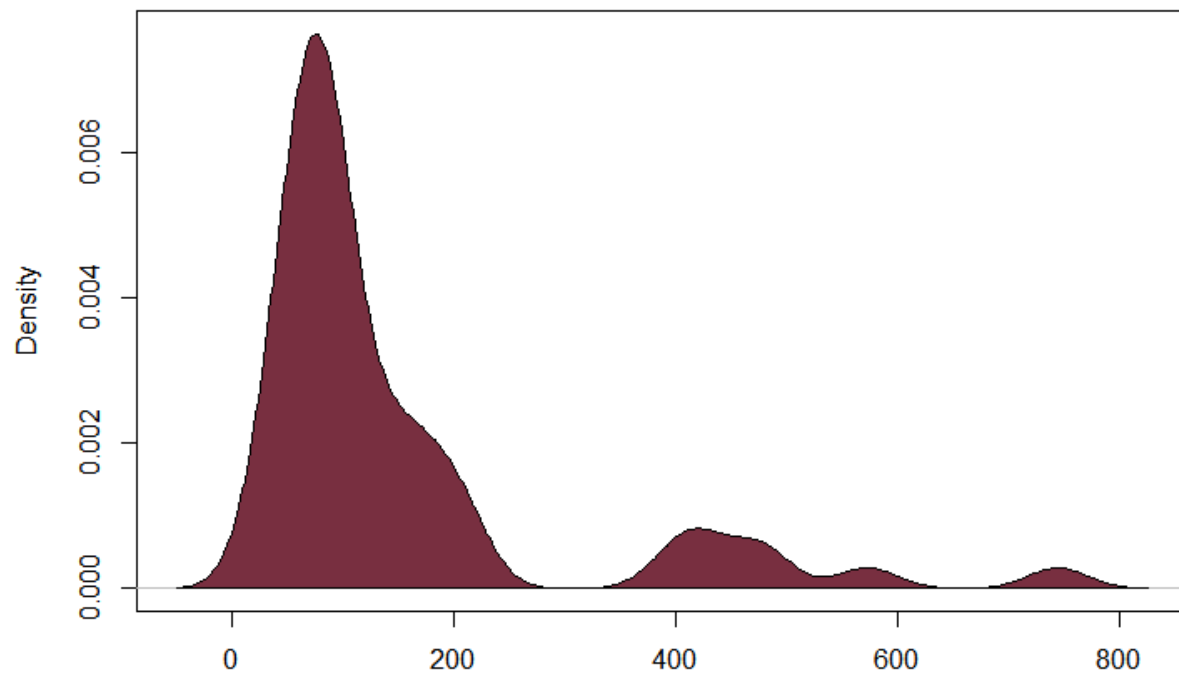
Q-Q Plot For Survey Duration in Seconds


```
>library(data.table)
> outlierReplace = function(dataframe, cols, rows, newValue = NA) {if (any(rows))
{set(dataframe, rows, cols, newValue)}}
> outlierReplace(X2022FSULibSurveyDataFinalVersion, "Duration (in seconds)",
which(X2022FSULibSurveyDataFinalVersion$`Duration (in seconds)` > 1000, NA))
> View(X2022FSULibSurveyDataFinalVersion)
> h2 <- hist(na.omit(X2022FSULibSurveyDataFinalVersion$`Duration (in seconds)`), breaks=10,
col="#782F40", xlab = "Duration in Seconds", main = "Histogram with Normal Curve (No
Outliers)")
```

Histogram with Normal Curve (No Outliers)
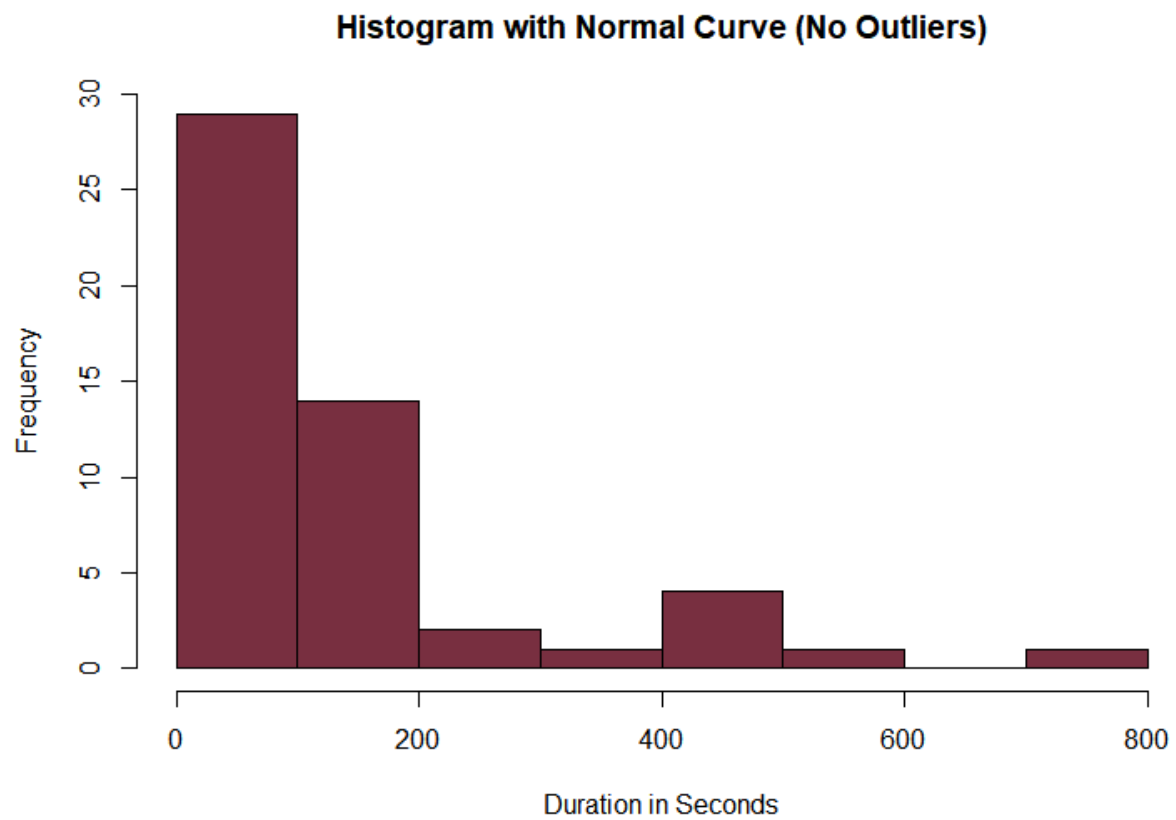
> d2 <-density(na.omit(X2022FSULibSurveyDataFinalVersion$`Duration (in seconds)`))
> plot(d2, main="Kernel Density of Survey Duration in Seconds (No Outliers)")
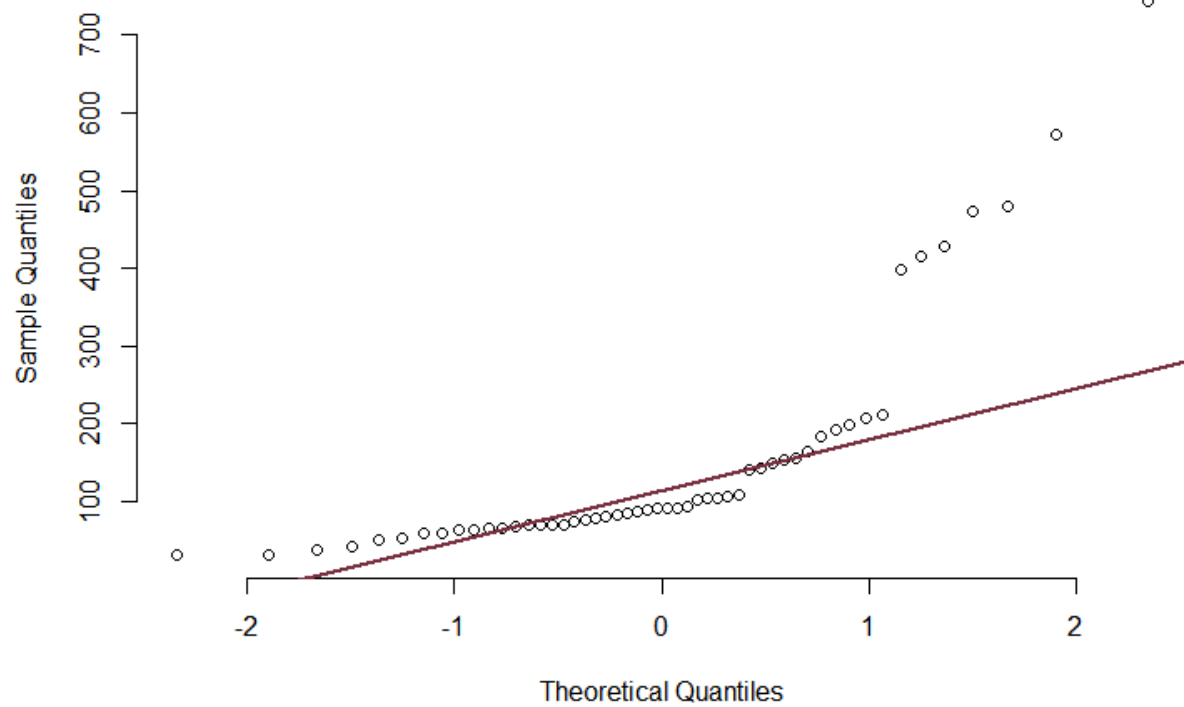> polygon(d2, col="#782F40", border="black")

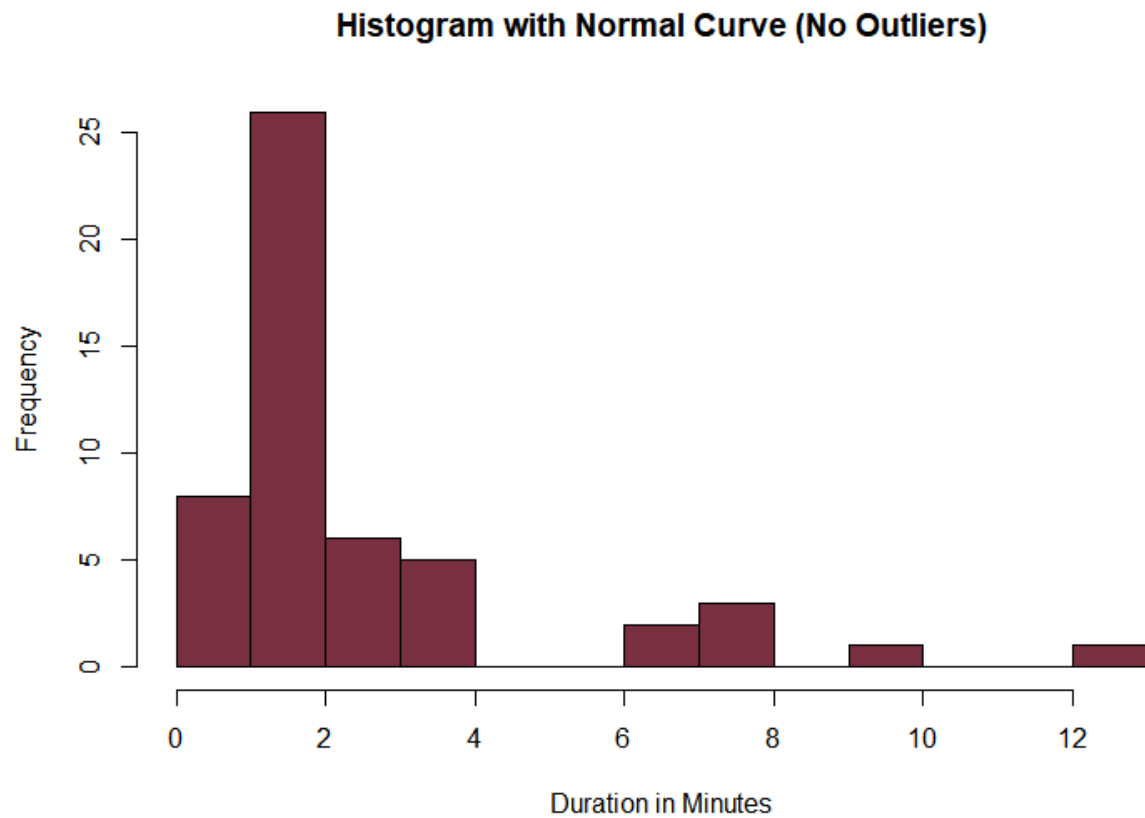## Kernel Density of Survey Duration in Seconds (No Outliers)

N = 52   Bandwidth = 26.97

> qqnorm(na.omit(X2022FSULibSurveyDataFinalVersion$`Duration (in seconds)`), pch=1,
frame = FALSE, main = "Q-Q Plot For Survey Duration in Seconds (No Outliers)")
> qqline(X2022FSULibSurveyDataFinalVersion$`Duration (in seconds)`, col="#782F40", lwd=2)

Q-Q Plot For Survey Duration in Seconds (No Outliers)

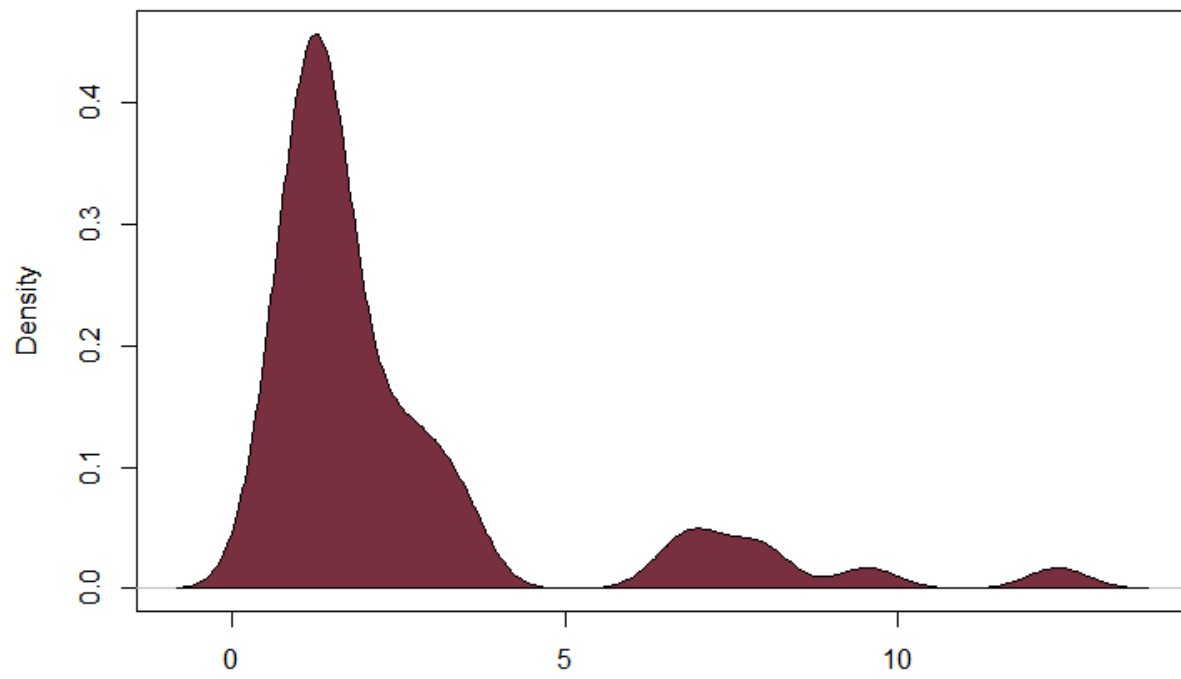#Time to scale the above information and graphics into minutes instead of seconds
> omittedtime <- na.omit(X2022FSULibSurveyDataFinalVersion$`Duration (in seconds)`)
> OmittedMinuteTime <- omittedtime/60

> h3 <- hist(OmittedMinuteTime, breaks=10, col="#782F40", xlab = "Duration in Minutes", main = "Histogram with Normal Curve (No Outliers)")

Histogram with Normal Curve (No Outliers)

> d3 <-density(OmittedMinuteTime)
> plot(d3, main="Kernel Density of Survey Duration in Minutes (No Outliers)")
> polygon(d3, col="#782F40", border="black")

Kernel Density of Survey Duration in Minutes (No Outliers)

N = 52   Bandwidth = 0.4495