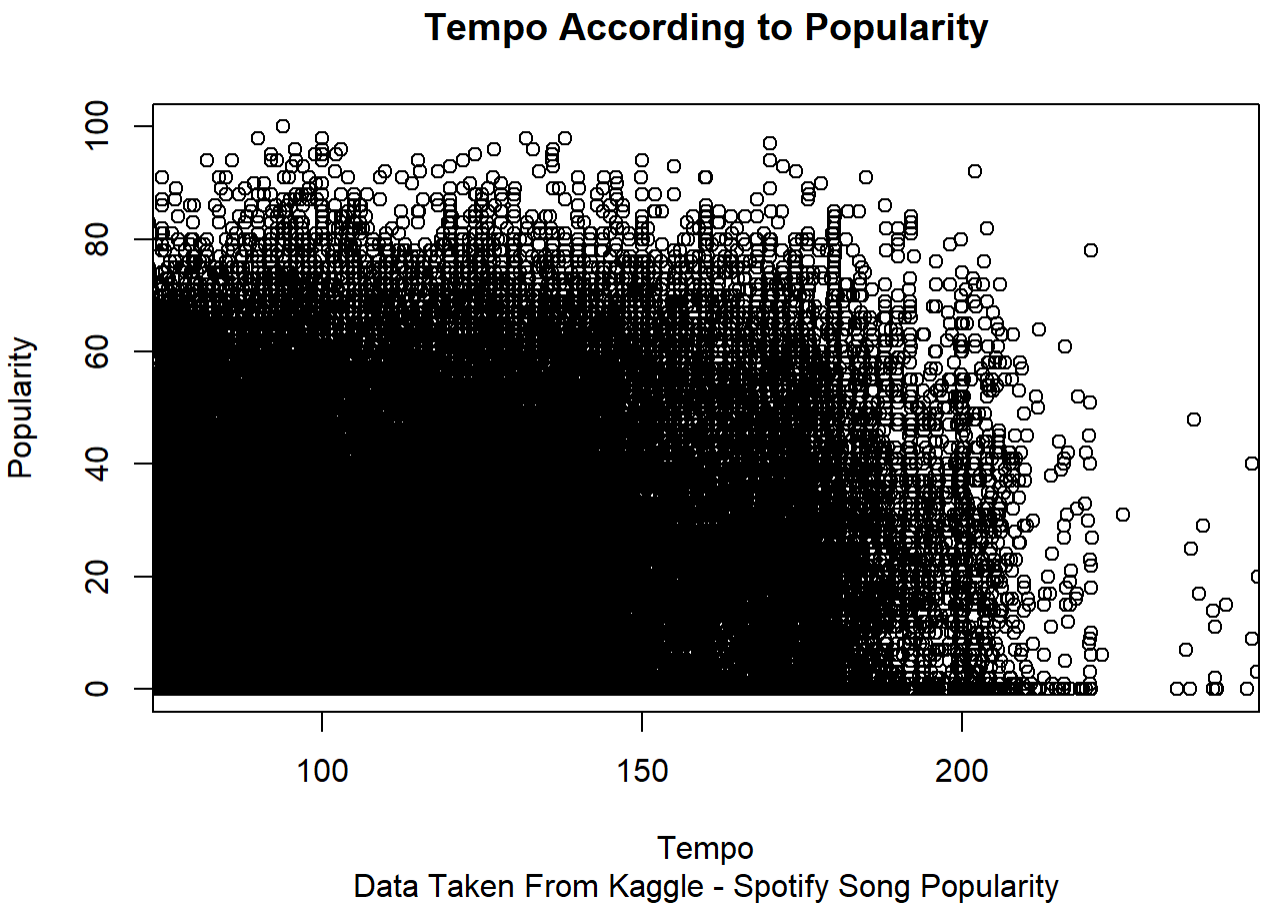
R Notebook

This is an R Markdown (http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Ctrl+Shift+Enter*.

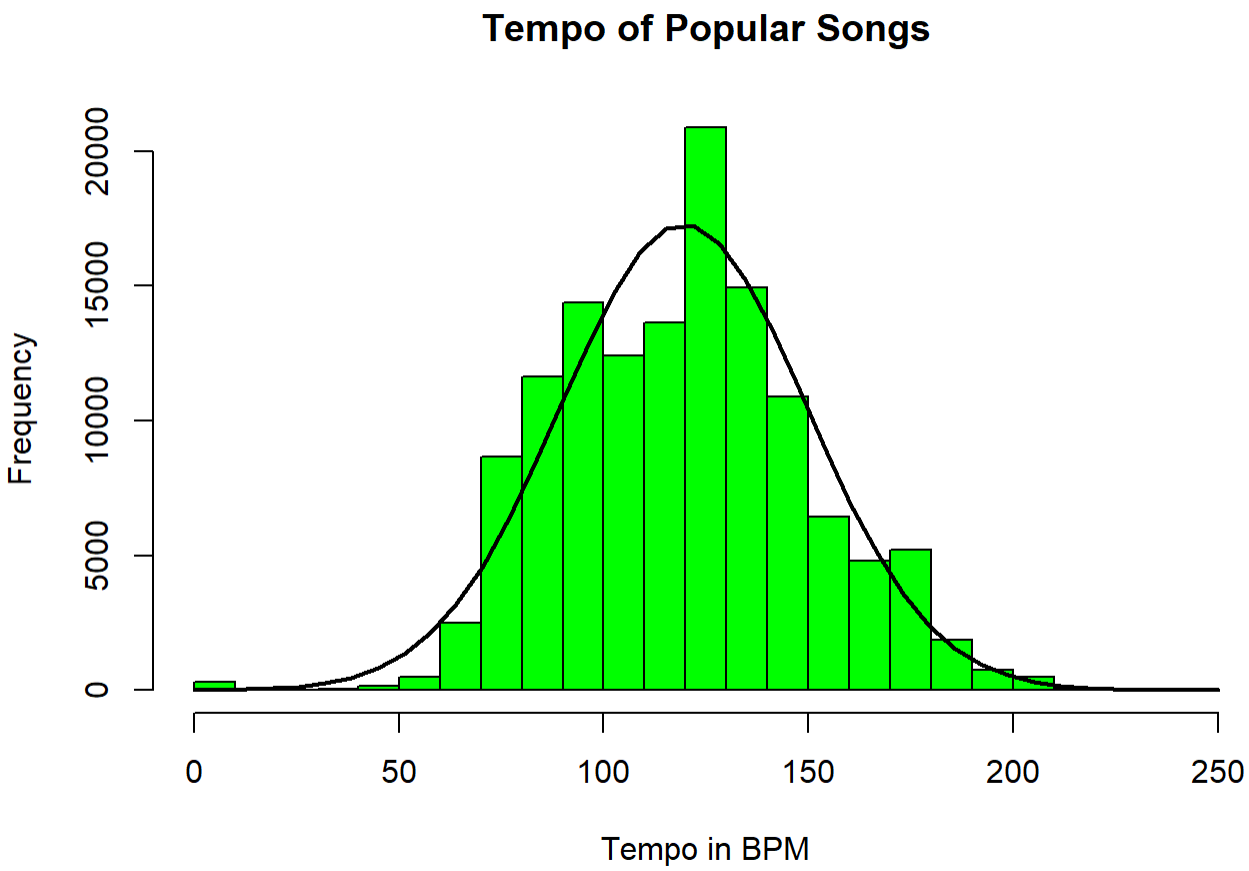
```
df <- read.csv("C:/Users/pokem/OneDrive/Documents/CS 4375/data/spotifySongPopularity.csv")

tempo <- df$tempo
pop <- df$popularity

plot(tempo, pop, type = "p", xlim=c(80,240), main="Tempo According to Popularity", sub="Data
Taken From Kaggle - Spotify Song Popularity", xlab="Tempo", ylab="Popularity" )
```




```
cor(tempo,pop,use="complete")
```

```
## [1] 0.03707521
```

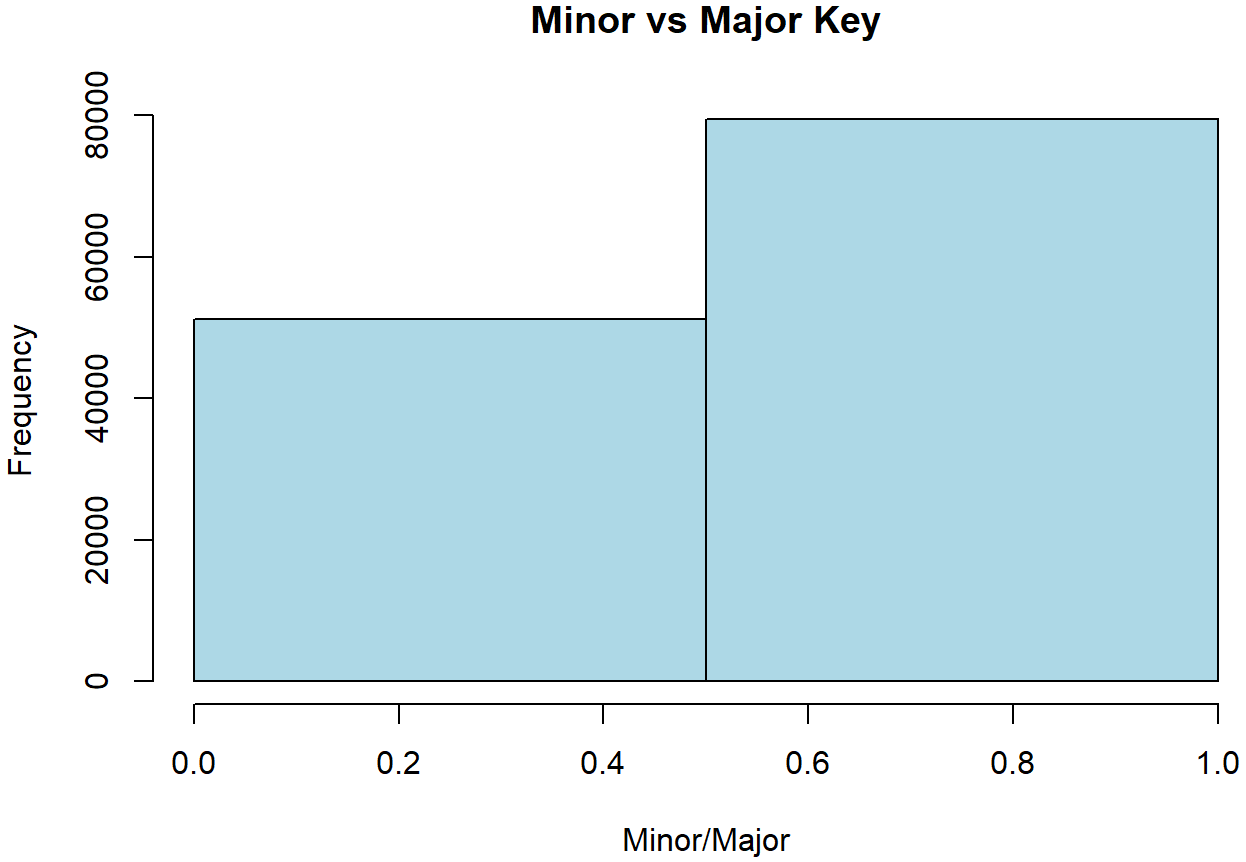
```
h <- hist(tempo, breaks=20, col="green", xlab="Tempo in BPM", main="Tempo of Popular Songs")
xfit <- seq(min(tempo), max(tempo), length=40)
yfit <- dnorm(xfit, mean=mean(tempo), sd=sd(tempo))
yfit <- yfit*diff(h$mids[1:2]*length(tempo))
lines(xfit, yfit, col="black", lwd=2)
```



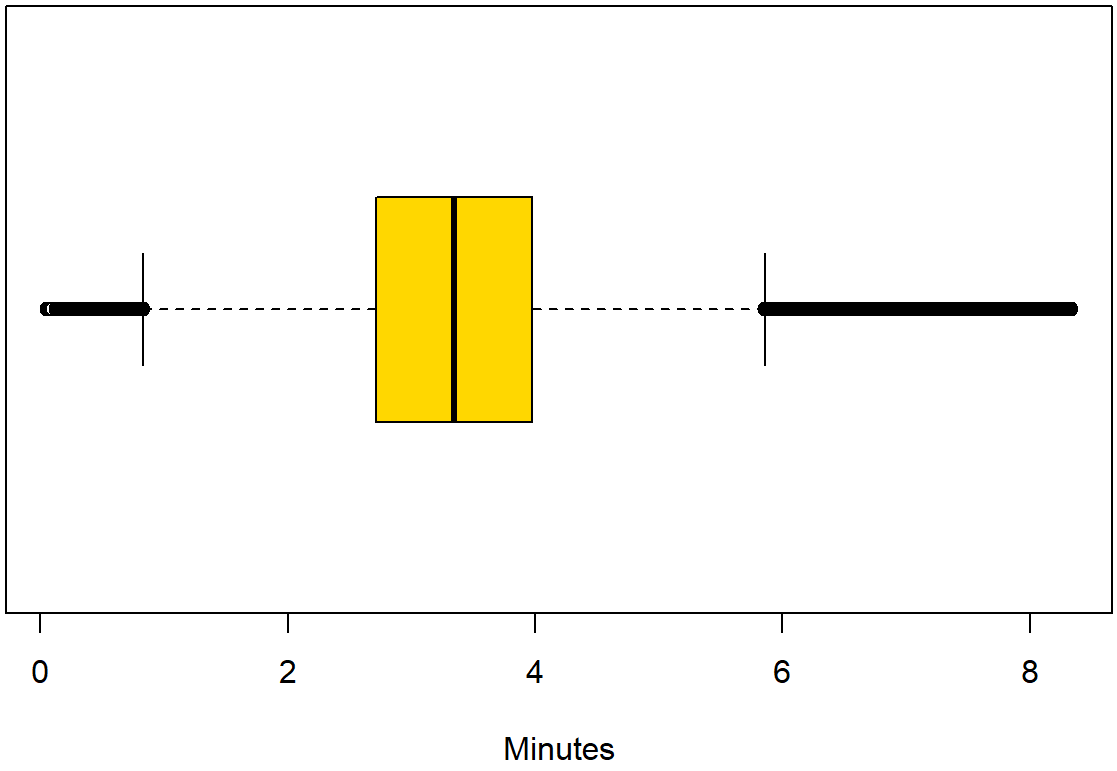

```
median(tempo)
```

```
## [1] 120.027
```

```
hist(df$mode, breaks=2, col="lightblue", xlab="Minor/Major", main="Minor vs Major Key")
```




```
duration <- subset(df, df$duration_ms < 500000)

boxplot(duration$duration_ms/60000, col="gold", main="Song Duration in Minutes", xlab="Minute
s", horizontal=TRUE)
```




```
mean(duration$duration_ms, na.rm=TRUE)/60000
```

```
## [1] 3.392507
```