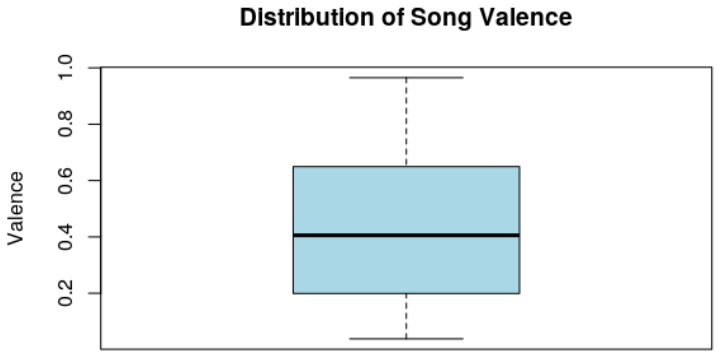
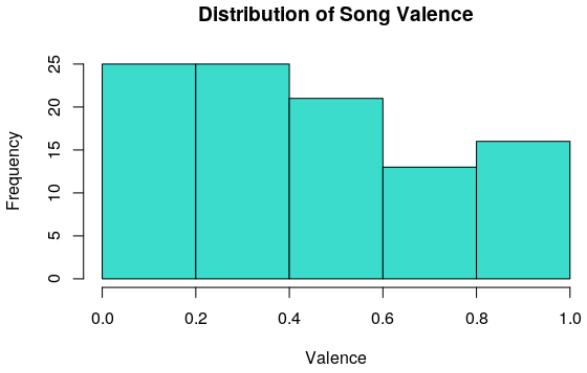
Name: Adetutu B
Date: July 08, 2022

RESPONSE VARIABLE

```
> boxplot(music$valence, ylab='Valence', main='Distribution of Song
Valence', col='light blue')
> mean(music$valence, na.rm = TRUE)
> sd(music$valence, na.rm = TRUE)
> fivenum(music$valence)
```




```
> hist(music$valence, xlab='Valence', main='Distribution of Song Valence',
right=F, col='turquoise', breaks=5)
```



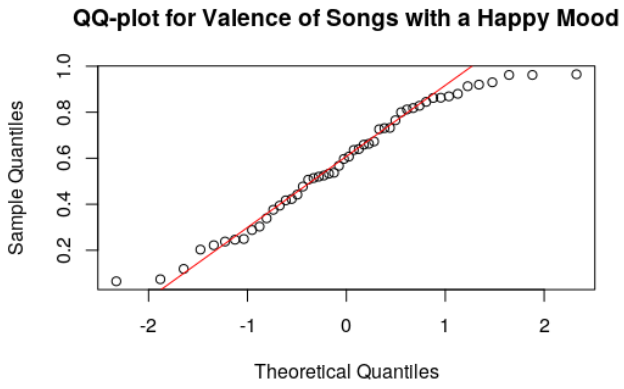

```
> mean(music$valence, na.rm = TRUE)
[1] 0.437801
>
> sd(music$valence, na.rm = TRUE)
[1] 0.2738077
>
> fivenum(music$valence)
[1] 0.0382 0.1990 0.4055 0.6495 0.9650
```

Name: Adetutu B
Date: July 08, 2022

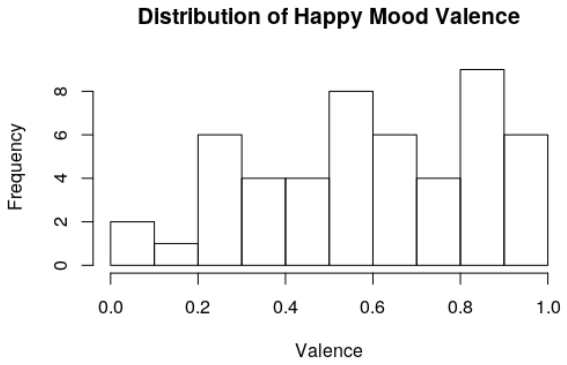
EXPLANATORY VARIABLE 1

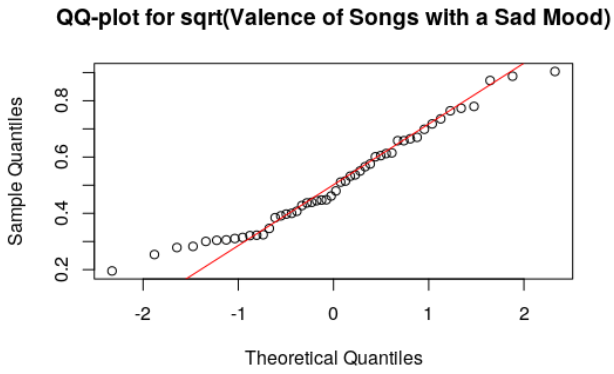
```
> barplot(table(music$mood), ylab='Frequency', main='Counts of Song
Moods', col='dark blue')
> boxplot(music$valence~music$mood, xlab='Mood', ylab='Valence',
main='Valence Distribution of Song Mood', col=c('light blue','light
green'))
> happy <- music[music$mood=='Happy',]
> mean(happy$valence, na.rm = TRUE)
[1] 0.584726
> sd(happy$valence, na.rm = TRUE)
[1] 0.258812
> sad <- music[music$mood=='Sad',]
> mean(sad$valence, na.rm = TRUE)
[1] 0.290876
> sd(sad$valence, na.rm = TRUE)
[1] 0.2010789
```

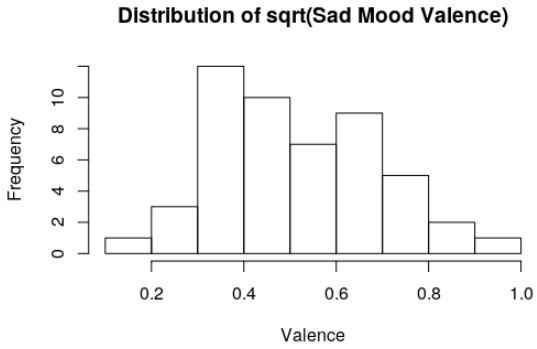
```
> happy_mood <- music$valence[music$mood=='Happy']
> qqnorm(happy_mood,main='QQ-plot for Valence of Songs with a Happy Mood')
> qqline(happy_mood,col='red')
> hist(happy_mood, main='Distribution of Happy Mood Valence',
xlab='Valence', right=F)
> sad_mood <- music$valence[music$mood=='Sad']
> qqnorm(sqrt(sad_mood),main='QQ-plot for sqrt(Valence of Songs with a Sad
Mood)')
> qqline(sqrt(sad_mood),col='red')
> hist(sqrt(sad_mood), main='Distribution of sqrt(Sad Mood Valence)',
xlab='Valence', right=F)
> library(car)
> leveneTest(valence~mood,data=music)
> t.test(valence~mood,data=music)
```



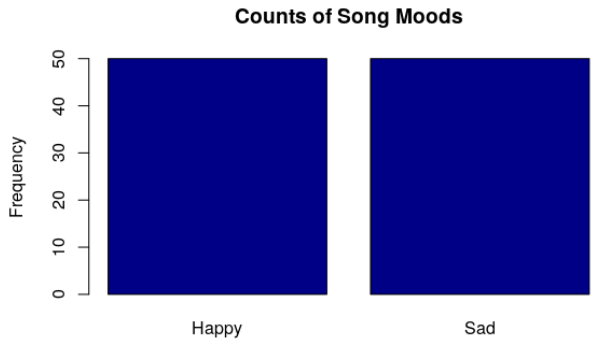
Name: Adetutu B
Date: July 08, 2022



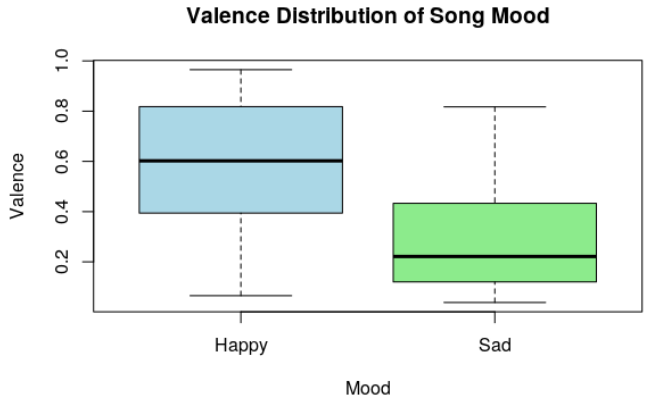




Name: Adetutu B
Date: July 08, 2022



Univariate bar plot




```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  1  4.6234  0.034 *
      98
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

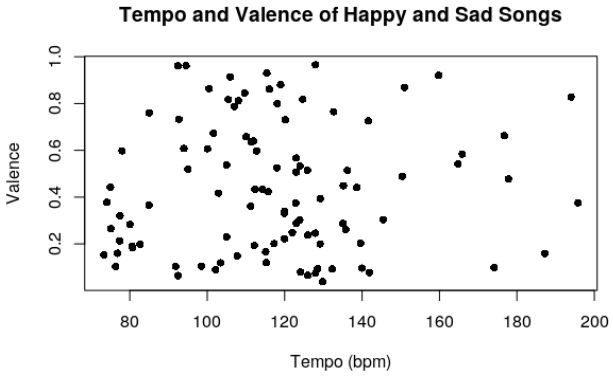
```
Welch Two Sample t-test

data:  valence by mood
t = 6.3398, df = 92.357, p-value = 8.342e-09
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.2017996 0.3859004
sample estimates:
mean in group Happy   mean in group Sad
          0.584726            0.290876
```

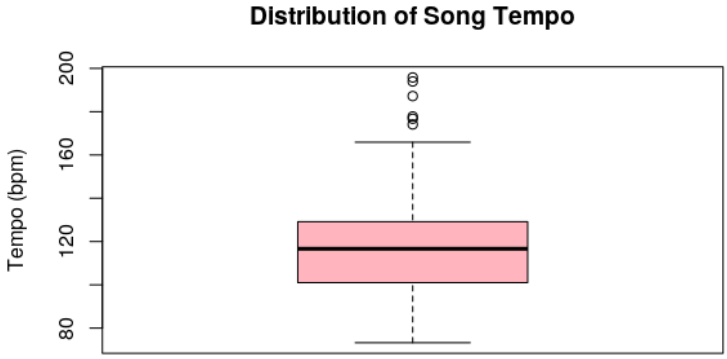
Name: Adetutu B
Date: July 08, 2022

EXPLANATORY VARIABLE 2

```
> boxplot(music$tempo, ylab='Tempo (bpm)', main='Distribution of Song
Tempo', col='light pink')
> mean(music$tempo, na.rm = TRUE)
> sd(music$tempo, na.rm = TRUE)
> fivenum(music$tempo)
> plot(music$tempo, music$valence, xlab='Tempo (bpm)', ylab='Valence',
main='Tempo and Valence of Happy and Sad Songs', pch=16)
```



Bivariate scatterplot




```
> mean(music$tempo, na.rm = TRUE)
[1] 117.6044
> sd(music$tempo, na.rm = TRUE)
[1] 27.0692
> fivenum(music$tempo)
[1]  73.2540 100.9940 116.6675 129.1665 195.8150
```

Name: Adetutu B
Date: July 08, 2022

```
> my_glm <- lm(valence ~ mood + tempo, data=music)
> plot(music$tempo, music$valence, xlab='Tempo (bpm)', ylab= 'Song
Valence', main='Tempo and Valence', pch=20)
> hist(my_glm$residuals, main='Model Residuals', xlab='Residual',
col='light grey', right=F)
> plot(my_glm$fitted.values, my_glm$residuals, xlab= 'Fitted Values',
ylab='Residuals', main='Residual Plot', pch=20)
> abline(h=0, col='red')
> summary(my_glm)

Call:
lm(formula = valence ~ mood + tempo, data = music)

Residuals:
     Min       1Q   Median       3Q      Max
-0.51832 -0.16819 -0.03599  0.17501  0.52279

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.6547675  0.1147120   5.708 1.25e-07 ***
moodSad     -0.3009221  0.0477971  -6.296 8.90e-09 ***
tempo       -0.0005655  0.0008873  -0.637    0.525
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2325 on 97 degrees of freedom
Multiple R-squared:  0.2938, Adjusted R-squared:  0.2792
F-statistic: 20.18 on 2 and 97 DF,  p-value: 4.708e-08

>
> summary(my_glm)$adj.r.squared
[1] 0.2792435
>
> music$tempo_c <- music$tempo - mean(music$tempo)
>
> my_glm_int <- lm(valence ~ mood*tempo_c, data=music)
> summary(my_glm_int)

Call:
lm(formula = valence ~ mood * tempo_c, data = music)

Residuals:
     Min       1Q   Median       3Q      Max
-0.51723 -0.17216 -0.02926  0.17021  0.52474

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
```

Name: Adetutu B
Date: July 08, 2022

```
(Intercept)      0.5914349  0.0341777  17.305  < 2e-16 ***
moodSad         -0.3020223  0.0480519  -6.285  9.6e-09 ***
tempo_c         -0.0010729  0.0014174  -0.757    0.451
moodSad:tempo_c  0.0008389  0.0018224   0.460    0.646
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2334 on 96 degrees of freedom
Multiple R-squared:  0.2954, Adjusted R-squared:  0.2733
F-statistic: 13.41 on 3 and 96 DF,  p-value: 2.209e-07
```