Top Song Predictor

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
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
```

```
library(ggplot2)
library(pROC)
```

```
## Warning: package 'pROC' was built under R version 3.6.2
```

```
## Type 'citation("pROC")' for a citation.
```

```
##
## Attaching package: 'pROC'
```

```
## The following objects are masked from 'package:stats':
##
##     cov, smooth, var
```

Helper Functions

Data Analysis

```
data <-read.csv('MusicData.csv')
data[is.na(data)]
```

```
## character(0)
```

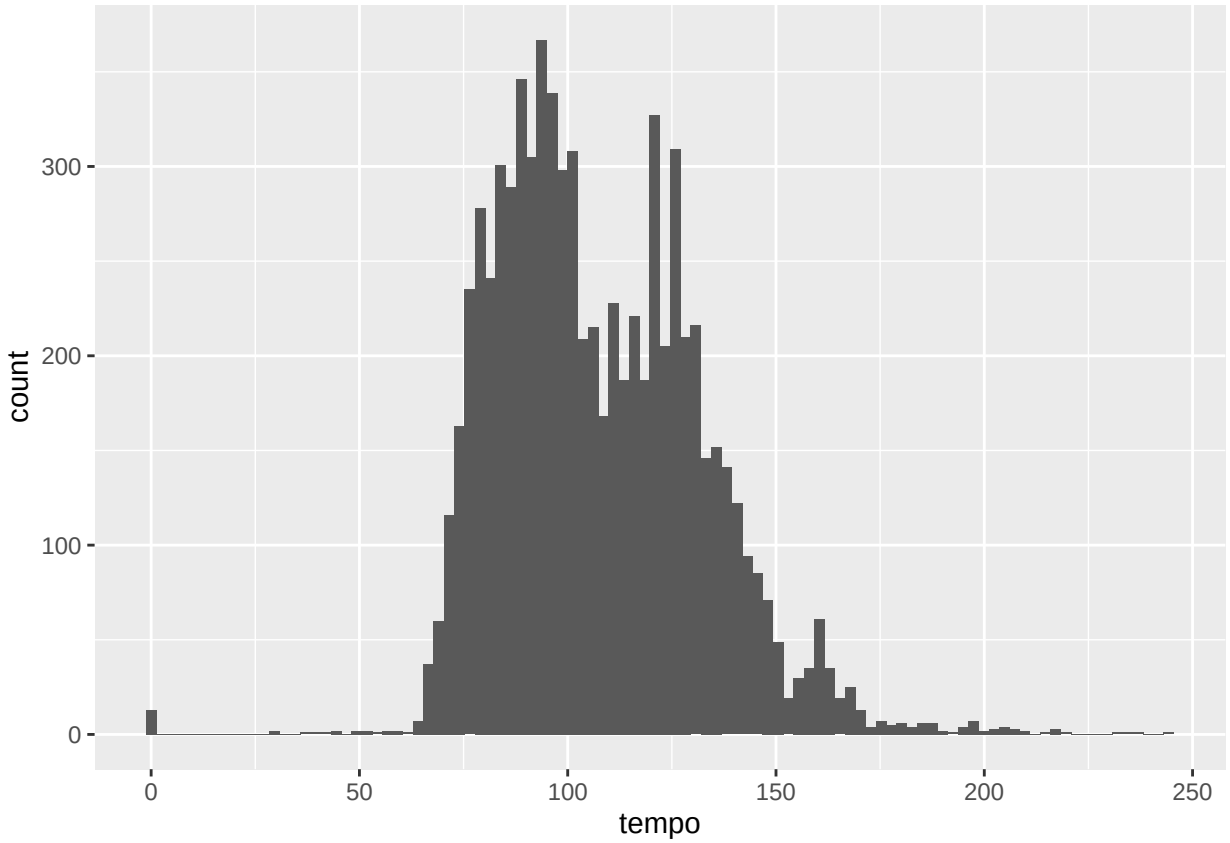
There is no missing data

```
dim(data)
```

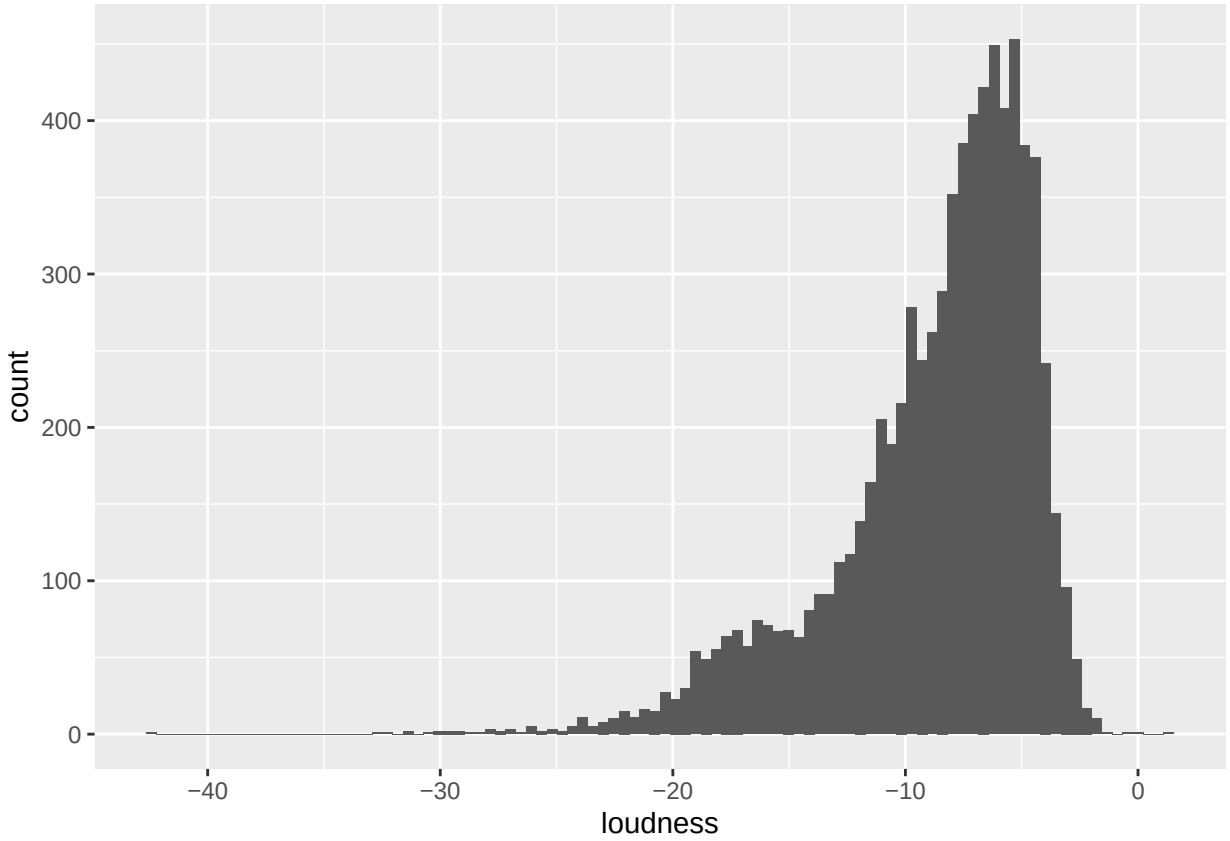
```
## [1] 7574   39
```

There's 7574 rows and 39 columns

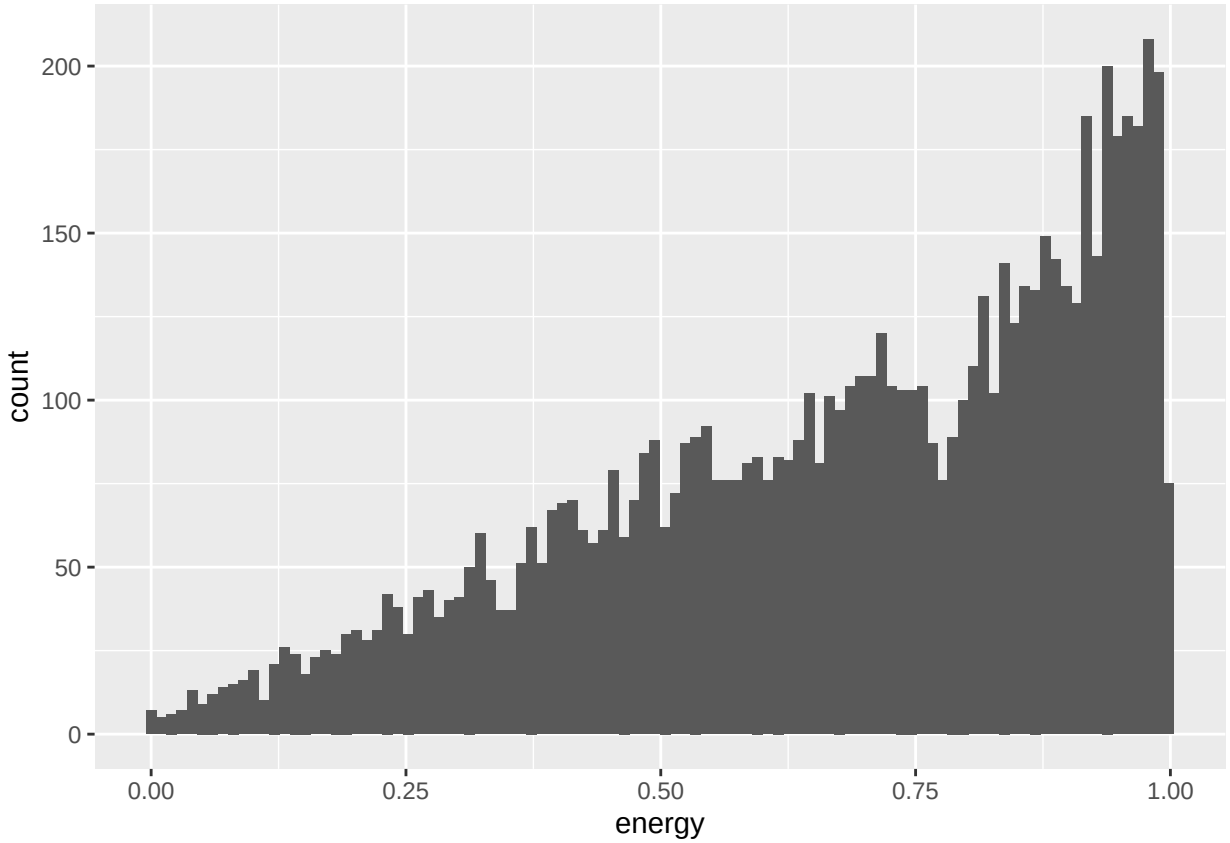
```
p <- ggplot(data, aes(x=tempo)) + geom_histogram(bins = 100)
print(p)
```




```
p <- ggplot(data, aes(x=loudness)) + geom_histogram(bins = 100)
print(p)
```




```
p <- ggplot(data, aes(x=energy)) + geom_histogram(bins = 100)
print(p)
```


200
150
100
50
0
count
0.00
0.25
0.50
0.75
1.00
energy