1. Correlation heatmap between key music features and stream counts.

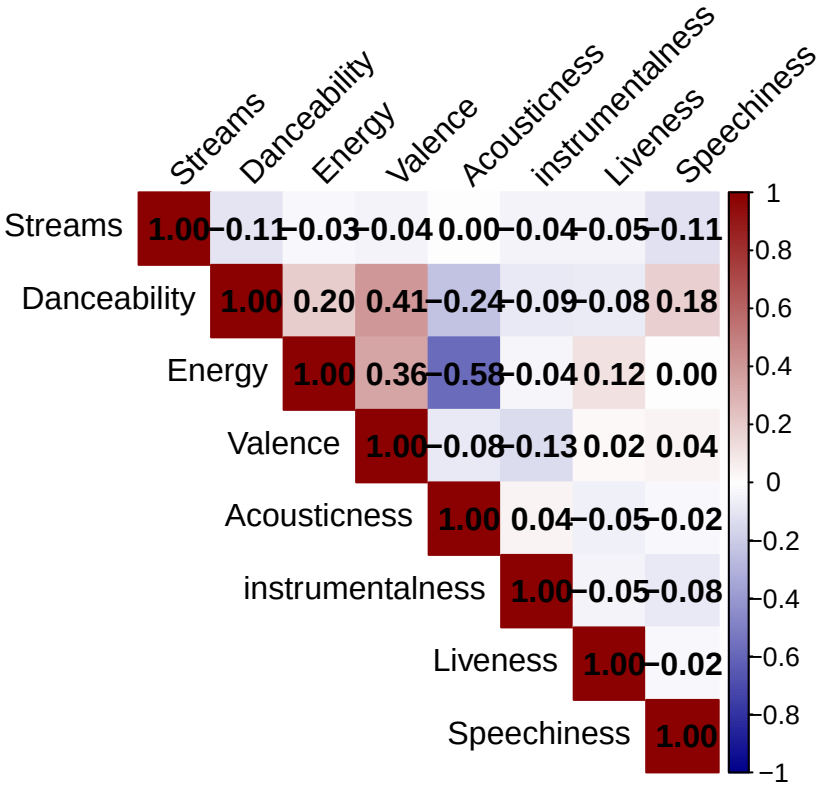
```
# install.packages("ggplot2")
# install.packages("reshape2")
# install.packages("RColorBrewer")
# install.packages(corrplot)


library(ggplot2)
library(reshape2)
library(RColorBrewer)
library(corrplot)
```

```
## corrplot 0.95 loaded
```

```
# Load the data
spotify_data <- read.csv("Spotify Most Streamed Songs.csv")


# Select relevant features for correlation analysis
correlation_features <- spotify_data[, c("streams", "danceability_.", "energy_.", "valence_.", "acoustic

colnames(correlation_features) <- c("Streams", "Danceability", "Energy", "Valence", "Acousticness","inst


# Convert columns to numeric if necessary and handle NA values
correlation_features <- data.frame(lapply(correlation_features, function(x) as.numeric(as.character(x))
```

```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```

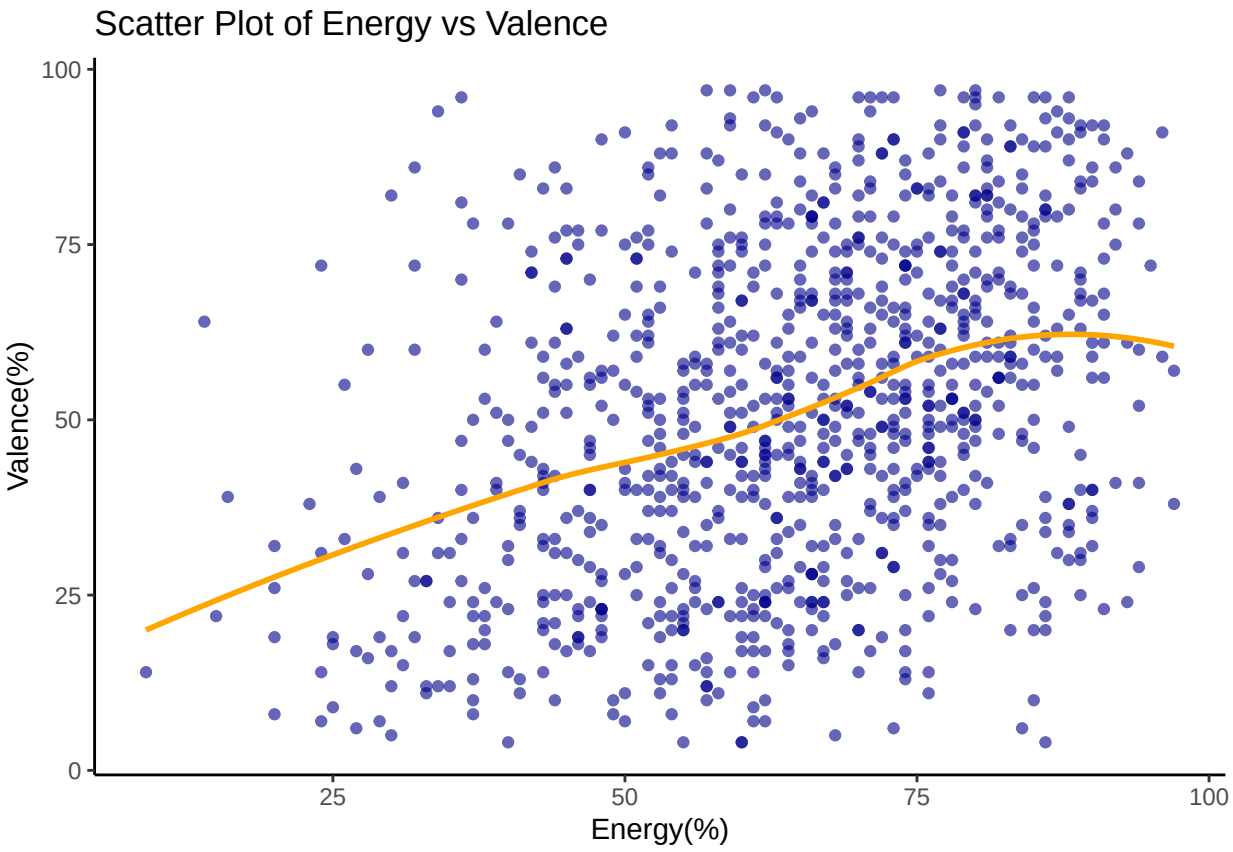
```
# Remove rows with NA values to ensure complete pairs for correlation
correlation_features <- na.omit(correlation_features)


# Calculate the correlation matrix
correlation_matrix <- cor(correlation_features, use = "complete.obs")


# Create a heatmap using corrplot
heatmap_plot<-corrplot(correlation_matrix, method = "color", type = "upper", tl.col = "black", tl.srt =
```



2. Scatterplot of energy vs valence, danceability vs valance, and energy vs acounstiness.

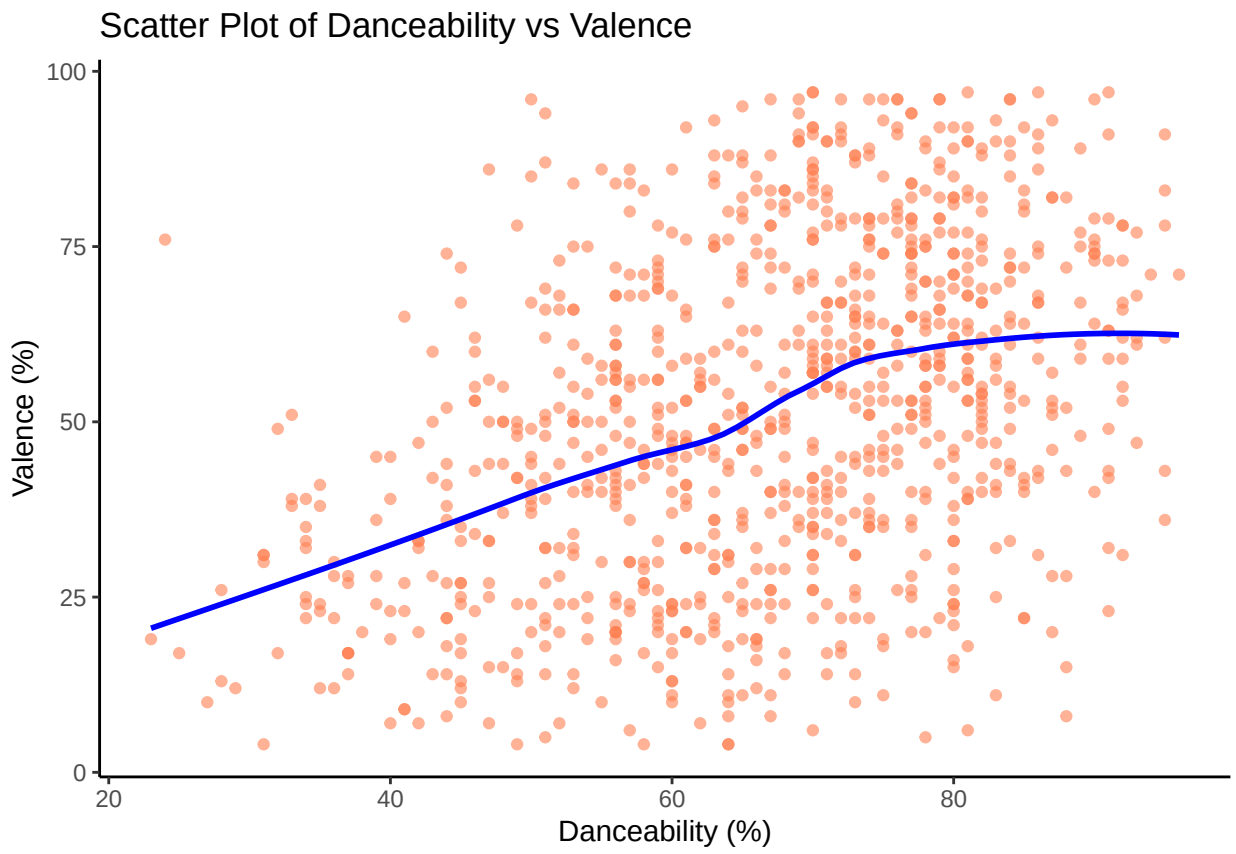
```
# Visualizations to Explore Relationships Between Highly Correlated Variables
# Scatter plot of energy vs valence
scatter_plot_energy_valence <- ggplot(data = spotify_data, aes(x = energy_., y = valence_.)) +
  geom_point(alpha = 0.6, color = "darkblue") +
  geom_smooth(method = "loess", color = "orange", se = FALSE) +
  theme_classic()+
  labs(title = "Scatter Plot of Energy vs Valence",
       x = "Energy(%)",
       y = "Valence(%)")

print(scatter_plot_energy_valence)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



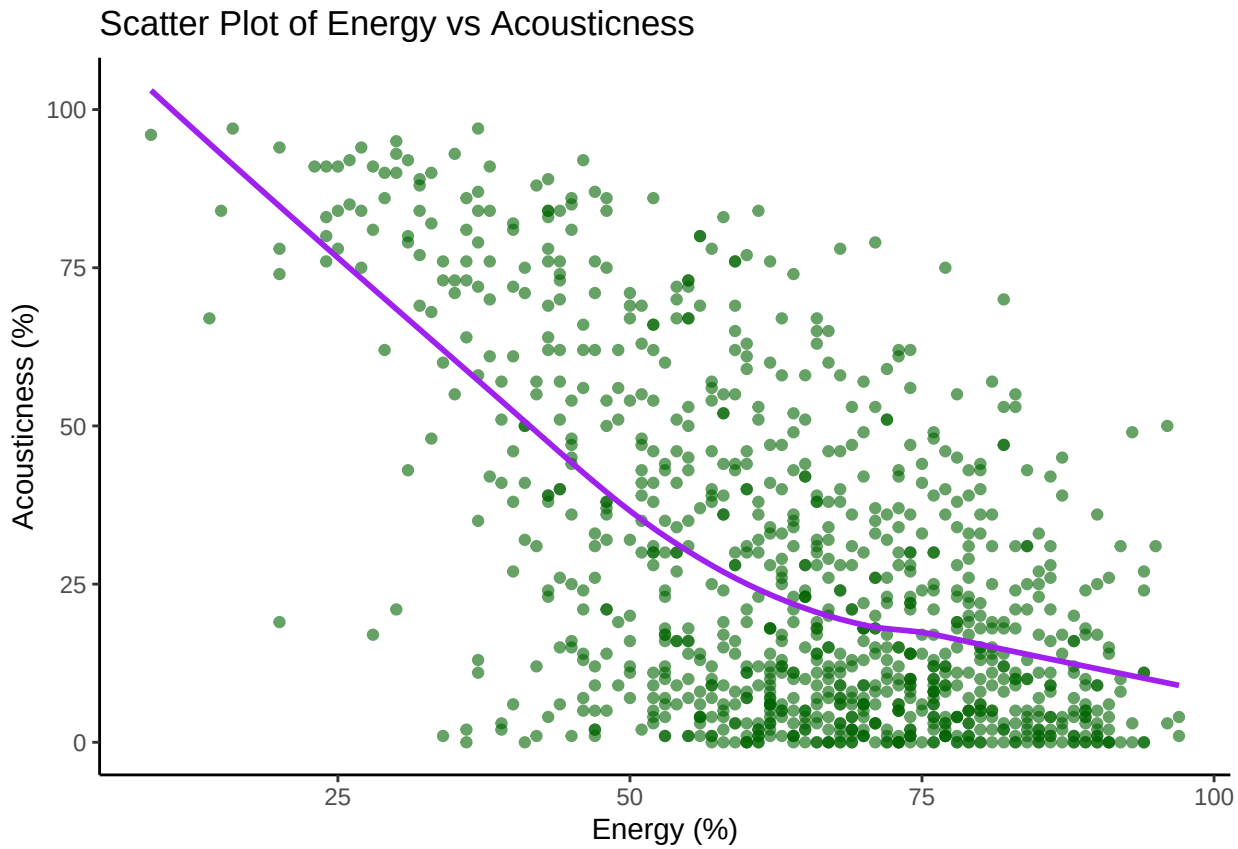

```
# Scatter plot of danceability vs valence
scatter_plot_danceability_valence <- ggplot(data = spotify_data, aes(x = danceability_., y = valence_.))
  geom_point(alpha = 0.6, color = "coral") +
  geom_smooth(method = "loess", color ="blue", se = FALSE) +
  theme_classic() +
  labs(title = "Scatter Plot of Danceability vs Valence",
       x = "Danceability (%)",
       y = "Valence (%)")

print(scatter_plot_danceability_valence)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
# Scatter plot of energy vs acousticness
scatter_plot_energy_acousticness <- ggplot(data = spotify_data, aes(x = energy_., y = acousticness_.)) +
  geom_point(alpha = 0.6, color = "darkgreen") +
  geom_smooth(method = "loess", color = "purple", se = FALSE) +
  theme_classic() +
  labs(title = "Scatter Plot of Energy vs Acousticness",
       x = "Energy (%)",
       y = "Acousticness (%)")
print(scatter_plot_energy_acousticness)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```


Scatter Plot of Energy vs Acousticness
Acousticness (%)
100
75
50
25
0
25
50
75
100
Energy (%)