test

2024-03-28

## R Markdown

library(tidyverse)

## Warning: package 'ggplot2' was built under R version 4.3.3

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.3 ✔ readr 2.1.4  
## ✔ forcats 1.0.0 ✔ stringr 1.5.0  
## ✔ ggplot2 3.5.0 ✔ tibble 3.2.1  
## ✔ lubridate 1.9.3 ✔ tidyr 1.3.0  
## ✔ purrr 1.0.2   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(factoextra)

## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa

library(dplyr)  
library(GGally)

## Warning: package 'GGally' was built under R version 4.3.3

## Registered S3 method overwritten by 'GGally':  
## method from   
## +.gg ggplot2

library(inspectdf)

## Warning: package 'inspectdf' was built under R version 4.3.3

library(ggiraphExtra)

## Warning: package 'ggiraphExtra' was built under R version 4.3.3

library(factoextra)  
library(tidyr)  
df\_original = read\_csv("SpotifyFeatures.csv")

## Rows: 232725 Columns: 18  
## ── Column specification ────────────────────────────────────────────────────────  
## Delimiter: ","  
## chr (7): genre, artist\_name, track\_name, track\_id, key, mode, time\_signature  
## dbl (11): popularity, acousticness, danceability, duration\_ms, energy, instr...  
##   
## ℹ Use `spec()` to retrieve the full column specification for this data.  
## ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

# Select only numeric variables  
df = df\_original %>%  
 select(where(is.numeric))  
# scale the data  
df <- data.frame((scale(df)))  
# plot the correlation matrix  
ggcorr(df, label=T, label\_size = 2.9, hjust = 1)

