Assignment 4

Matthias Rathbun

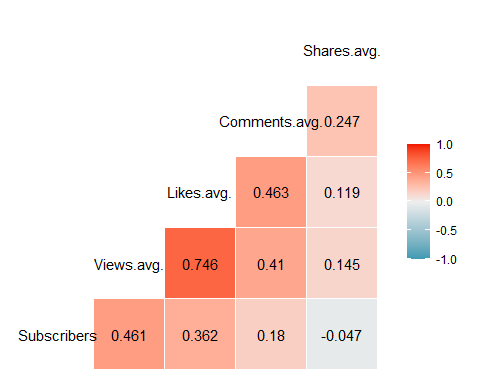
2022-10-10

df <-read.csv("tiktok.csv", header=TRUE)  
library(stringr)  
df$Subscribers <- as.numeric(str\_replace\_all(df$Subscribers, setNames(c("e3", "e6"), c("K", "M"))))  
df$Views.avg. <- as.numeric(str\_replace\_all(df$Views.avg., setNames(c("e3", "e6"), c("K", "M"))))  
df$Likes.avg. <- as.numeric(str\_replace\_all(df$Likes.avg., setNames(c("e3", "e6"), c("K", "M"))))  
df$Comments.avg. <- as.numeric(str\_replace\_all(df$Comments.avg., setNames(c("e3", "e6"), c("K", "M"))))  
df$Shares.avg. <- as.numeric(str\_replace\_all(df$Shares.avg., setNames(c("e3", "e6"), c("K", "M"))))  
library(GGally)

## Loading required package: ggplot2

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

ggcorr(df[4:8], label = TRUE, label\_round = 3)



The Highest Correlation is Between Views and Likes