```
## Project:  Base class script for STA 215, spring 2022
# Located:   Kline Posit Cloud
# File Name: Kline(2023_2_27)STA215_base
# Date:      2023_2_27
# Who:       Zachary D. Kline

### Settings + Packages
# Install packages
install.packages('readr')
install.packages("dplyr")
install.packages("tidytext")
install.packages("tidyverse")
install.packages('ggplot2')
install.packages("haven")
install.packages("psych")
install.packages("readxl")

  ##### load data #####
data <- read_delim(raw_data)


##############################################################################
#################### Examine descriptive statistics          ####################
##############################################################################

# EXAMINE plays_TS
table(data$plays_TS)
mean(data$plays_TS)
sd(data$plays_TS)
summary(data$plays_TS)

# EXAMINE plays_TV
table(data$plays_TV)
mean(data$plays_TV)
sd(data$plays_TV)
summary(data$plays_TV)

# EXAMINE type_song
table(data$type_song)

# EXAMINE genre
table(data$genre)

#####Step 2: Table

table (data$genre, data$type_song)


####### Step 3: Perform Chi-Squared Test

chisq.test(data$genre, data$type_song)

###### Step 4: Perform ANOVA

# Perform Anova
anova <- aov(data$plays_TS ~ data$type_song)
# Summarize ANOVA results
summary(anova)
# get R2
# between/total
# OR between/(between+within)

###### Step 5: Correlation
```

```
cor(data$plays_TS, data$plays_TV)

##### STEP 6: Calculate linear regression line (i.e., slope) and add to scatter plot
linear_relationship <- lm(data$plays_TS ~ data$plays_TV)
summary(linear_relationship)

linear_relationship <- lm(data$plays_TS ~ data$plays_TV)
summary(linear_relationship)

plot(data$plays_TS ~ data$plays_TV)

abline(linear_relationship, col = "red")


##### Step 7: Scatterplot ######

abline(h=6909522, col="blue")

abline(v=8580276, col="green")

#### Step 8 Residuals #####

plot(residuals.lm(linear_relationship))

abline(linear_relationship, col = "red")

abline(h= -871255 , col="blue")

abline(v=25 , col="green")
```