

```

rm(coef_df)

library(tidyverse)
library(ggcorrplot)
library(ggplot2)
library(ggExtra)
library(car)
library(lattice)
library(plotly)
library(broom)

red_sox_players_all <- read.csv("Red_Sox_Hitters.csv")
str(red_sox_players_all)
summary(red_sox_players_all)

red_sox_players_wk <- red_sox_players_all

red_sox_players_wk <- red_sox_players_wk %>%
mutate(Player = str_remove_all(Player, "[*#]"))
red_sox_players_wk$Player

SLG_HR_Cor_value <- cor(red_sox_players_wk$SLG, red_sox_players_wk$HR)
SLG_HR_Cor_value

base_plot <- ggplot(red_sox_players_wk, aes(x = SLG, y = HR)) +
geom_point(color = "firebrick", size = 3, alpha = 0.8) +
geom_smooth(method = "lm", se = FALSE, color = "gray20", linetype = "dotted") +
theme_minimal() +
labs(title = "Slugging vs Home Run with Marginal Distributions",
x = "Slugging % (SLG)",
y = "Home Runs (HR)")

ggMarginal(base_plot, type = "density", fill = "red", alpha = 0.4)

ggplot(red_sox_players_wk, aes(x = SLG, y = HR)) +
geom_smooth(method = "loess", color = "darkgreen", fill = "lightgreen") +
theme_minimal() +
labs(title = "Smoothed Trend: Slugging vs Home Run",
x = "Slugging % (SLG)",
y = "Home Runs (HR)")

red_sox_players_wk <- red_sox_players_wk %>%
mutate(SLG_quartile = ntile(SLG, 4))

ggplot(red_sox_players_wk, aes(x = factor(SLG_quartile), y = HR)) +
geom_boxplot(fill = "skyblue") +
theme_minimal() +
labs(title = "Home Run Distribution by Slugging Quartile",

```

```

x = "Slugging Quartile",
y = "Home Runs")

reg_dataframe <- data.frame(
  SLG = red_sox_players_wk$SLG,
  HR = red_sox_players_wk$HR,
  OBP = red_sox_players_wk$OBP,
  WAR = red_sox_players_wk$WAR
)

reg_model <- lm(WAR ~ OBP + SLG + HR, data = reg_dataframe)
summary(reg_model)

plot_ly(red_sox_players_wk,
  x = ~SLG, y = ~HR, size = ~WAR, color = ~OBP,
  type = "scatter", mode = "markers", text = ~Player,
  marker = list(sizemode = 'diameter', opacity = 0.7)) %>%
  layout(title = "Offensive Output Bubble Plot",
  xaxis = list(title = "Slugging % (SLG)"),
  yaxis = list(title = "Home Runs (HR)"))

plot_ly(data = reg_dataframe,
  x = ~fitted(reg_model),
  y = ~resid(reg_model),
  type = "scatter", mode = "markers",
  text = ~paste("WAR:", WAR, "<br>HR:", HR),
  marker = list(size = 8, color = 'blue', opacity = 0.6)) %>%
  layout(title = "Interactive Residuals vs Fitted",
  xaxis = list(title = "Fitted Values"),
  yaxis = list(title = "Residuals"))

```