

```

install.packages("readxl")
install.packages("ggplot2")
library(readxl)
library(ggplot2)
library(dplyr)

getwd()
setwd("C:/Users/shiva/OneDrive/Desktop/Capstone")

data <- read_excel("ConsolidatedData.xlsx")

df <- data.frame(data)

# Check for missing values in both columns
missing_in_rank <- sum(is.na(data$`MergedRank`))
missing_in_women <- sum(is.na(data$`MergedWomenonboard`))
missing_in_rank
missing_in_women

# Remove rows with missing values in either column
data_cleaned <- na.omit(data[c("MergedRank", "MergedWomenonboard")])

# Fit a linear regression model
model <- lm(`MergedWeightedsalary` ~ `MergedWomenonboard` +
`MergedFemalefaculty` + `MergedInternationalstudents` + `MergedInternationalfaculty` + `MergedInternationalboard` + `MergedFemalestudents`,
data = data)

# Print a summary of the regression model
summary(model)

# Make predictions using the model
new_data <- data.frame(`MergedWomenonboard` = 100, `MergedFemalefaculty` = 28, `MergedInternationalstudents` = 25,
`MergedInternationalfaculty` = 28, `MergedInternationalboard`=28, `MergedFemalestudents` = 25)# Provide values for new data

predictions <- predict(model, newdata = new_data)

predictions

```