Final Project (Group 2)

Group 2

2024-05-03

```
library(readxl)
library(dplyr)
```

```
library(dplyr)
data <- read_excel("WHR_2015.xlsx", col_names = TRUE)
data <- data %>%
  rename(
    HappinessScore = `Happiness Score`,
    GDPPerCapita = `Economy (GDP per Capita)`,
  )
colnames(data)
```

```
##  [1] "Country"                       "Region"
##  [3] "Happiness Rank"                "HappinessScore"
##  [5] "Standard Error"                "GDPPerCapita"
##  [7] "Family"                        "Health (Life Expectancy)"
##  [9] "Freedom"                       "Trust (Government Corruption)"
## [11] "Generosity"                    "Dystopia Residual"
```

```
colnames(data) <- c("Country", "Region", "Happiness Score",
                    "Happiness Rank","Economy (GDP per Capita",
                    "Health (Life Expectancy)", "Freedom",
                    "Trust (Government Corruption)", "Family", "Generosity",
                    "Dystopia Residual", "Standard Error")
```

```
head(data)
```

Country	Region	Happiness Score	Happiness Rank	Economy (GDP per Capita)	Health (Life Expectancy)	Freedom	Trust (Government Corruption)	Family	Generosity	Dystopia Residual	Standard Error
Switzerland	Western Europe	1	7.587	0.03411	1.39651	1.34951	0.94143	0.66557	0.41978	2.29678	2.51738
Iceland	Western Europe	2	7.561	0.04884	1.30232	1.40223	0.94784	0.62877	0.14145	2.43630	2.70201
Denmark	Western Europe	3	7.527	0.03328	1.32548	1.36058	0.87464	0.64938	0.48357	2.34139	2.49204
Norway	Western Europe	4	7.522	0.03880	1.45900	1.33095	0.88521	0.66973	0.36503	2.34699	2.46531
Canada	North America	5	7.427	0.03553	1.32629	1.32261	0.90563	0.63297	0.32957	2.45811	2.45176
Finland	Western Europe	6	7.406	0.03140	1.29025	1.31826	0.88911	0.64169	0.41372	2.23351	2.61955

```
data <- read_excel("WHR_2015.xlsx", col_names = TRUE)
data$`Economy (GDP per Capita)` <- as.factor(data$`Economy (GDP per Capita)`)
```

```
summary_stats <- data %>%
  summarize(
    Mean = mean(`Happiness Score`, na.rm = TRUE),
    Median = median(`Happiness Score`, na.rm = TRUE),
    Standard_Deviation = sd(`Happiness Score`, na.rm = TRUE),
    Minimum = min(`Happiness Score`, na.rm = TRUE),
    Maximum = max(`Happiness Score`, na.rm = TRUE)
  )

print(summary_stats)
```

```
## # A tibble: 1 x 5
##    Mean Median Standard_Deviation Minimum Maximum
##   <dbl>  <dbl>              <dbl>   <dbl>   <dbl>
## 1  5.38   5.23               1.15    2.84    7.59
```

```
data %>%
  summarize(center = median(`Happiness Score`, na.rm = TRUE))
```

center
5.2325