Final Project (Group 2)

Group 2

2024-05-02

```
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 C
```

```
colnames(data)
```

```
##  [1] "Country"                       "Region"
##  [3] "Happiness Score"               "Happiness Rank"
##  [5] "Economy (GDP per Capita"       "Health (Life Expectancy)"
##  [7] "Freedom"                       "Trust (Government Corruption)"
##  [9] "Family"                        "Generosity"
## [11] "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	0.29678	2.51738	
Iceland	Western Europe	2	7.561	0.04884	1.30232	1.40223	0.94784	0.62877	0.43630	2.70201	
Denmark	Western Europe	3	7.527	0.03328	1.32548	1.36058	0.87464	0.64938	0.34139	2.49204	
Norway	Western Europe	4	7.522	0.03880	1.45900	1.33095	0.88521	0.66973	0.34699	2.46531	
Canada	North America	5	7.427	0.03553	1.32629	1.32261	0.90563	0.63297	0.45811	2.45176	
Finland	Western Europe	6	7.406	0.03140	1.29025	1.31826	0.88911	0.64169	0.23351	2.61955	

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

```
library(readxl)
library(dplyr)

data <- read_excel("WHR_2015.xlsx")

print(colnames(data))
```

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

```
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
```

```
library(readxl)
data <- read_excel("WHR_2015.xlsx")
print(colnames(data))
```

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

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

center
5.2325

```
library(readxl)
library(dplyr)

data <- read_excel("WHR_2015.xlsx")

print(colnames(data))
```

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

```r
summary_stats <- data %>%
  summarize(
    Mean = mean(`Economy (GDP per Capita)`, na.rm = TRUE),
    Median = median(`Economy (GDP per Capita)`, na.rm = TRUE),
    Standard_Deviation = sd(`Economy (GDP per Capita)`, na.rm = TRUE),
    Minimum = min(`Economy (GDP per Capita)`, na.rm = TRUE),
    Maximum = max(`Economy (GDP per Capita)`, na.rm = TRUE)
  )

print(summary_stats)
```

```
## # A tibble: 1 x 5
##    Mean Median Standard_Deviation Minimum Maximum
##   <dbl>  <dbl>              <dbl>   <dbl>   <dbl>
## 1 0.846  0.910              0.403       0    1.69
```

```r
library(readxl)
library(ggplot2)
library(dplyr)
data <- read_excel("WHR_2015.xlsx")
print(colnames(data))
```

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

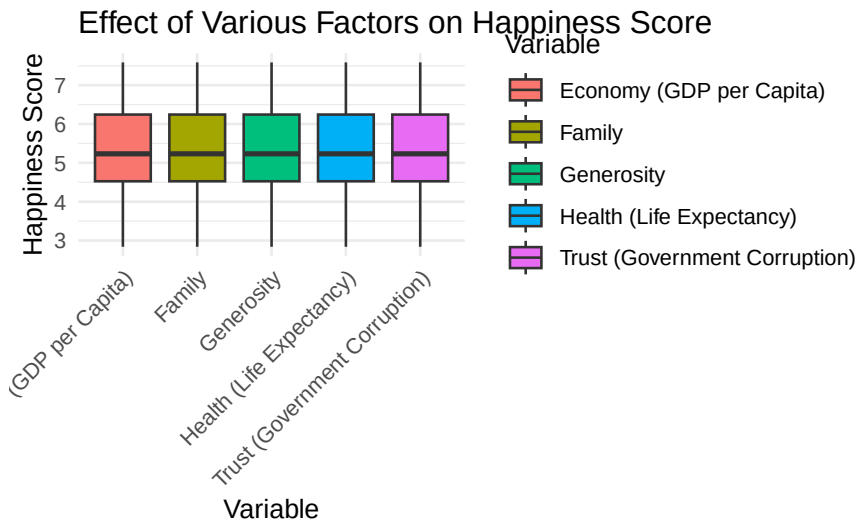
```r
summary(data)
```

```
##    Country             Region          Happiness Rank   Happiness Score
##  Length:158         Length:158         Min.   :  1.00   Min.   :2.839
##  Class :character   Class :character   1st Qu.: 40.25   1st Qu.:4.526
##  Mode  :character   Mode  :character   Median : 79.50   Median :5.232
##                                        Mean   : 79.49   Mean   :5.376
##                                        3rd Qu.:118.75   3rd Qu.:6.244
##                                        Max.   :158.00   Max.   :7.587
##  Standard Error    Economy (GDP per Capita)     Family
##  Min.   :0.01848   Min.   :0.0000           Min.   :0.0000
##  1st Qu.:0.03727   1st Qu.:0.5458           1st Qu.:0.8568
##  Median :0.04394   Median :0.9102           Median :1.0295
##  Mean   :0.04788   Mean   :0.8461           Mean   :0.9910
##  3rd Qu.:0.05230   3rd Qu.:1.1584           3rd Qu.:1.2144
##  Max.   :0.13693   Max.   :1.6904           Max.   :1.4022
```

```
##  Health (Life Expectancy)    Freedom       Trust (Government Corruption)
##  Min.   :0.0000           Min.   :0.0000   Min.   :0.00000
##  1st Qu.:0.4392           1st Qu.:0.3283   1st Qu.:0.06168
##  Median :0.6967           Median :0.4355   Median :0.10722
##  Mean   :0.6303           Mean   :0.4286   Mean   :0.14342
##  3rd Qu.:0.8110           3rd Qu.:0.5491   3rd Qu.:0.18025
##  Max.   :1.0252           Max.   :0.6697   Max.   :0.55191
##    Generosity     Dystopia Residual
##  Min.   :0.0000   Min.   :0.3286
##  1st Qu.:0.1506   1st Qu.:1.7594
##  Median :0.2161   Median :2.0954
##  Mean   :0.2373   Mean   :2.0990
##  3rd Qu.:0.3099   3rd Qu.:2.4624
##  Max.   :0.7959   Max.   :3.6021
```

```r
data_long <- data %>%
  select(`Happiness Score`, `Economy (GDP per Capita)`, Family, `Health (Life Expectancy)`, `T
  pivot_longer(cols = -`Happiness Score`, names_to = "Variable", values_to = "Value")

ggplot(data_long, aes(x = Variable, y = `Happiness Score`, fill = Variable)) +
  geom_boxplot() +
  labs(title = "Effect of Various Factors on Happiness Score",
       x = "Variable",
       y = "Happiness Score") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```r
ggplot(data, aes(x = factor(1), y = `Happiness Score`)) +
  geom_boxplot(aes(fill = `Economy (GDP per Capita)`)) +
  scale_fill_gradient(low = "blue", high = "red") +
  labs(title = "Box Plot of Happiness Score by Economy (GDP per Capita)",
```

```
       x = "Economy (GDP per Capita)",
       y = "Happiness Score") +
  theme_minimal() +
  theme(legend.position = "none")
```

```
## Warning: The following aesthetics were dropped during statistical transformation: fill.
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a 'group' aesthetic or to convert a numerical
##   variable into a factor?
```


```
ggplot(data, aes(x = factor(1), y = `Happiness Score`)) +
  geom_violin(aes(fill = `Economy (GDP per Capita)`)) +
  scale_fill_gradient(low = "blue", high = "red") +
  labs(title = "Violin Plot of Happiness Score by Economy (GDP per Capita)",
       x = "Economy (GDP per Capita)",
       y = "Happiness Score") +
  theme_minimal() +
  theme(legend.position = "none")
```

```
## Warning: The following aesthetics were dropped during statistical transformation: fill.
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a 'group' aesthetic or to convert a numerical
##   variable into a factor?
```


Violin Plot of Happiness Score by Economy (GDP per Cap
Happiness Score
7
6
5
4
3
1
Economy (GDP per Capita)