Simple Linear Model Report

Far Eastern University
APM1205: Linear Model
Mr. Teodolfo Bonitez
February 22, 2024

ROSALES, Frances Aneth C.

Scenario No.1:

- Read DataBase Code:

```
whrdata <- read_excel("C:\\Users\\asus\\Documents\\ALL FEU FILES\\FEU FOLDER 6\\LINEAR MODEL\\FA1\\WHR2023.xlsx")
colnames(whrdata) <- c("Country_name", "Index", "Healthy_life_expectancy", "Freedom_to_make_life_choices", "Generosity", "Perceptions_of_corruption","Ladder_score_in_Dystopia","Log_GDP_per_capita")
View(whrdata)
```

No. 1:

Construct a SLM and try to assess the movement in the Happiness Index of a country if Perception of Corruption changes

- CODE:

```
happiness_vs_corruption <- lm(Index ~ Perceptions_of_corruption, data = whrdata)

summary_corruption <- summary(happiness_vs_corruption)
summary_corruption
```

- Output:

```
```{r 1.1}
happiness_vs_corruption <- lm(Index ~ Perceptions_of_corruption, data = whrdata)

summary_corruption <- summary(happiness_vs_corruption)
summary_corruption
```

Call:
lm(formula = Index ~ Perceptions_of_corruption, data = whrdata)

Residuals:
    Min   1Q Median   3Q   Max
-3.3111 -0.7077 0.2721 0.7524 1.8803

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.7450   0.3649 21.225 < 2e-16 ***
Perceptions_of_corruption -3.0400   0.4888 -6.219 5.84e-09 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.009 on 135 degrees of freedom
Multiple R-squared: 0.2227,	Adjusted R-squared: 0.2169
F-statistic: 38.68 on 1 and 135 DF, p-value: 5.84e-09
```

No. 2:

Determine the β_0 and the β_1

- CODE:

beta0_corruption <- coef(happiness_vs_corruption)[1]

print(paste("Beta0:", beta0_corruption))

beta1_corruption <- coef(happiness_vs_corruption)[2]

print(paste("Beta1:", beta1_corruption))

Output:

```
```{r 1.2}
beta0_corruption <- coef(happiness_vs_corruption)[1]
print(paste("Beta0:", beta0_corruption))

beta1_corruption <- coef(happiness_vs_corruption)[2]
print(paste("Beta1:", beta1_corruption))
```

[1] "Beta0: 7.74501135762855"
[1] "Beta1: -3.03999351977372"
```

No. 3:

Determine the variance of β_0 and the β_1

- CODE:

beta0_corruption <- coef(happiness_vs_corruption)[1]

print(paste("Beta0:", beta0_corruption))

beta1_corruption <- coef(happiness_vs_corruption)[2]

print(paste("Beta1:", beta1_corruption))

Output:

```
```{r 1.3}

var_beta0_corruption <- vcov(happiness_vs_corruption)[1, 1]
print(paste("Variance of Beta0:", var_beta0_corruption))

var_beta1_corruption <- vcov(happiness_vs_corruption)[2, 2]
print(paste("Variance of Beta1:", var_beta1_corruption))

```

[1] "Variance of Beta0: 0.133156982564878"
[1] "Variance of Beta1: 0.238935725997191"
```

No. 4:

Determine the mean value of e_i

- CODE:

```
mean_residuals_corruption <- mean(residuals(happiness_vs_corruption))

print(paste("Mean value of e_i:", mean_residuals_corruption))
```

Output:

```
```{r 1.4}
mean_residuals_corruption <- mean(residuals(happiness_vs_corruption))
print(paste("Mean value of e_i:", mean_residuals_corruption))
```

[1] "Mean value of e_i: 4.64133496437444e-17"
```

No. 5:

Determine the Coefficient of Determination

- CODE:

```
r_squared_corruption <- summary_corruption$r.squared

print(paste("Coefficient of Determination:", r_squared_corruption))
```

Output:

```
```{r 1.5}
r_squared_corruption <- summary_corruption$r.squared
print(paste("Coefficient of Determination:", r_squared_corruption))
```

[1] "Coefficient of Determination: 0.222699564323429"
```

FEU

No. 6:

Write the final SL Model

- CODE:

```
finalmodel_corruption <- paste("Happiness Index vs Perception of Corruption =", round(beta0_corruption, 3), "+", round(beta1_corruption, 3), "* Perceptions of Corruption", "+",(mean_residuals_corruption) )

print(paste(" Final SL Model: ", finalmodel_corruption))
```

Output:

```
```{r 1.6}

finalmodel_corruption <- paste("Happiness Index vs Perception of
Corruption =", round(beta0_corruption, 3), "+",
round(beta1_corruption, 3), "* Perceptions of Corruption",
"+",(mean_residuals_corruption) )

print(paste(" Final SL Model: ", finalmodel_corruption))
```

[1] " Final SL Model: Happiness Index vs Perception of Corruption
= 7.745 + -3.04 * Perceptions of Corruption + 4.64133496437444e-17"
```

- CODE of Visualizations

```
ggplot(whrdata, aes(x = Perceptions_of_corruption, y = Index)) +
geom_point() +geom_smooth(method = "lm", se = FALSE, color = "violet") +
labs(title = "Happiness Index vs Perceptions of Corruption",
x = "Perceptions of Corruption",
y = "Happiness Index",
caption = "Figure 1.1") + theme(plot.caption = element_text(hjust = 0.5))
```

No. 7 & 8 : Explain the result of your model and prepare some visualizations.

As we determine the β0 accumulated **7.7450**, Since the Perception of Corruption is unlikely to be exactly zero in our circumstances, it may not have a clear practical interpretation using SLM.

Since β1 is the predicted shift in the Happiness Index in response to a rise of one unit in the perception of corruption.

In our situation, it is anticipated that the Happiness Index will drop by **3.0400** units for every unit that the Perception of Corruption rises.

Therefore, the bad indication of shows a negative correlation between the Happiness Index and the Perception of Corruption. The Happiness Index tends to decrease when corruption occurs to happen.

However, it is also needed to consider that the residual **standard error (1.009)** provides an indication of the normal residual size and a measure of the Happiness Index variability not explained by the Perception of Corruption.

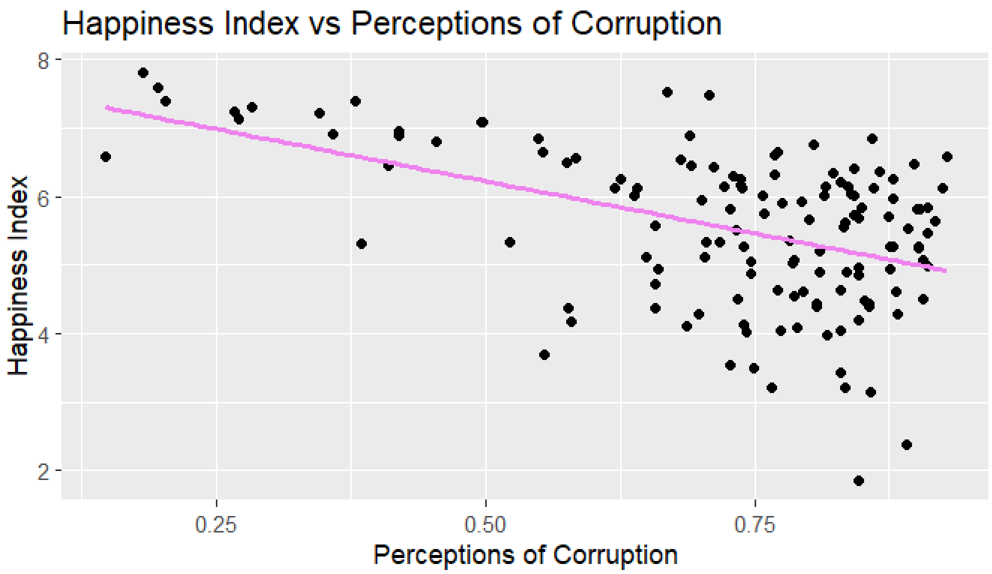

Figure 1.1

As assumed, the index of Happiness indeed decreases as the Perception of Corruption increases as shown in Figure 1.1.

Scenario No.2:

- Read DataBase Code:

```
whrdata <- read_excel("C:\\Users\\asus\\Documents\\ALL FEU FILES\\FEU FOLDER 6\\LINEAR
MODEL\\FA1\\WHR2023.xlsx")
colnames(whrdata) <- c("Country_name", "Index", "Healthy_life_expectancy", "Freedom_to_make_life_choices",
"Generosity", "Perceptions_of_corruption","Ladder_score_in_Dystopia","Log_GDP_per_capita")
View(whrdata)
```

No. 1:

Construct a SLM and try to assess the movement in the Happiness Index of a country if life expectancies at birth changes

- CODE:

```
happiness_vs_life_expectancy <- lm(Index ~ Healthy_life_expectancy, data = whrdata)

summary_life_expectancy <- summary(happiness_vs_life_expectancy)

summary_life_expectancy
```

- Output:

```
```{r 2.1}
happiness_vs_life_expectancy <- lm(Index ~ Healthy_life_expectancy,
data = whrdata)
summary_life_expectancy <- summary(happiness_vs_life_expectancy)
summary_life_expectancy
```

Call:
lm(formula = Index ~ Healthy_life_expectancy, data = whrdata)

Residuals:
    Min   1Q Median   3Q   Max
-3.3278 -0.4584 0.1201 0.5081 1.4043

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)   -4.09971   0.74452 -5.507 1.8e-07 ***
Healthy_life_expectancy 0.14845   0.01142 13.004 < 2e-16 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7627 on 134 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared: 0.5579,   Adjusted R-squared: 0.5546
F-statistic: 169.1 on 1 and 134 DF, p-value: < 2.2e-16
```

No. 2:

Determine the β_0 and the β_1

- CODE:

beta0_life_expectancy <- coef(happiness_vs_life_expectancy)[1]

beta1_life_expectancy <- coef(happiness_vs_life_expectancy)[2]

print(paste("Beta0:", beta0_life_expectancy))

print(paste("Beta1:", beta1_life_expectancy))

Output:

```
```{r 2.2}
beta0_life_expectancy <- coef(happiness_vs_life_expectancy)[1]
beta1_life_expectancy <- coef(happiness_vs_life_expectancy)[2]
print(paste("Beta0:", beta0_life_expectancy))
print(paste("Beta1:", beta1_life_expectancy))
```

[1] "Beta0: -4.09971381920294"
[1] "Beta1: 0.148445535821299"
```

No. 3:

Determine the variance of β_0 and the β_1

- CODE:

var_beta0_life_expectancy <- vcov(happiness_vs_life_expectancy)[1, 1]

var_beta1_life_expectancy <- vcov(happiness_vs_life_expectancy)[2, 2]

print(paste("Variance of Beta0:", var_beta0_life_expectancy))

print(paste("Variance of Beta1:", var_beta1_life_expectancy))

Output:

```
```{r 2.3}

var_beta0_life_expectancy <- vcov(happiness_vs_life_expectancy)[1, 1]
var_beta1_life_expectancy <- vcov(happiness_vs_life_expectancy)[2, 2]
print(paste("Variance of Beta0:", var_beta0_life_expectancy))
print(paste("Variance of Beta1:", var_beta1_life_expectancy))

```

[1] "Variance of Beta0: 0.554305362098477"
[1] "Variance of Beta1: 0.000130313883139216"
```

FEU

No. 4:

Determine the mean value of e_i

- CODE:

mean_residuals_life_expectancy <- mean(residuals(happiness_vs_life_expectancy))

print(paste("Mean value of e_i:", mean_residuals_life_expectancy))

Output:

```
```{r 2.4}
mean_residuals_life_expectancy <-
mean(residuals(happiness_vs_life_expectancy))
print(paste("Mean value of e_i:", mean_residuals_life_expectancy))

```

 [1] "Mean value of e_i: -2.13906710971552e-17"
```

No. 5:

Determine the Coefficient of Determination

- CODE:

r_squared_life_expectancy <- summary_life_expectancy$r.squared

print(paste("Coefficient of Determination:", r_squared_life_expectancy))

Output:

```
```{r 2.5}
r_squared_life_expectancy <- summary_life_expectancy$r.squared
print(paste("Coefficient of Determination:",
r_squared_life_expectancy))

```

 [1] "Coefficient of Determination: 0.557901676640821"
```

No. 6:

Write the final SL Model

- CODE:

final_model_life_expectancy <- paste("Happiness Index vs Life Expectancy =", round(beta0_life_expectancy, 3), "+", round(beta1_life_expectancy, 3), "* Healthy Life Expectancy ","+",(mean_residuals_life_expectancy))

print(paste(" Final SL Model ", final_model_life_expectancy))

Output:

```
```{r 2.6}
final_model_life_expectancy <- paste("Happiness Index vs
Life Expectancy =", round(beta0_life_expectancy, 3), "+",
round(beta1_life_expectancy, 3), "* Healthy Life Expectancy
","+",(mean_residuals_life_expectancy) )
print(paste(" Final SL Model ",
final_model_life_expectancy))

```

[1] " Final SL Model Happiness Index vs Life Expectancy
= -4.1 + 0.148 * Healthy Life Expectancy +
-2.13906710971552e-17"
```

- CODE of Visualizations

```
library(cowplot)
ggplot(whrdata, aes(x = Healthy_life_expectancy, y = Index)) +
geom_point() + geom_smooth(method = "lm", se = FALSE, color = "violet") + labs(title = "Happiness Index vs Life
Expectancy", x = "Life Expectancy", y = "Happiness Index", caption = "Figure 2.1") + theme(plot.caption =
element_text(hjust = 0.5))
```

No. 7 & 8 : Explain the result of your model and prepare some visualizations.

As we determine the β0 accumulated **−4.09971** , Since the Life Expectancy is also not possible to be exactly zero in real-life circumstances, interpreting data using Simple Linear Model would not be practical for this scenario.

This on the otherhand predicted (β1) the shifting of Happiness Index corresponds to the Life Expectancy from our Data Base.

We can now see as Happiness Index will increases by **0.14845** units whenever the Life Expectancy increases, then it pertains that the result has a correlate proportionally with the index.

Therefore, now we can see a good correlation of proportion increment of Life Expectancy and the Happiness Index.

However, just we assumed on our preceding assumption on Happiness Index vs Perception of Corruption, our data accumulated a **standard error of (0.7627)** thus we cannot also say that the Happiness Index has an indeed correlation between Life Expectancy.

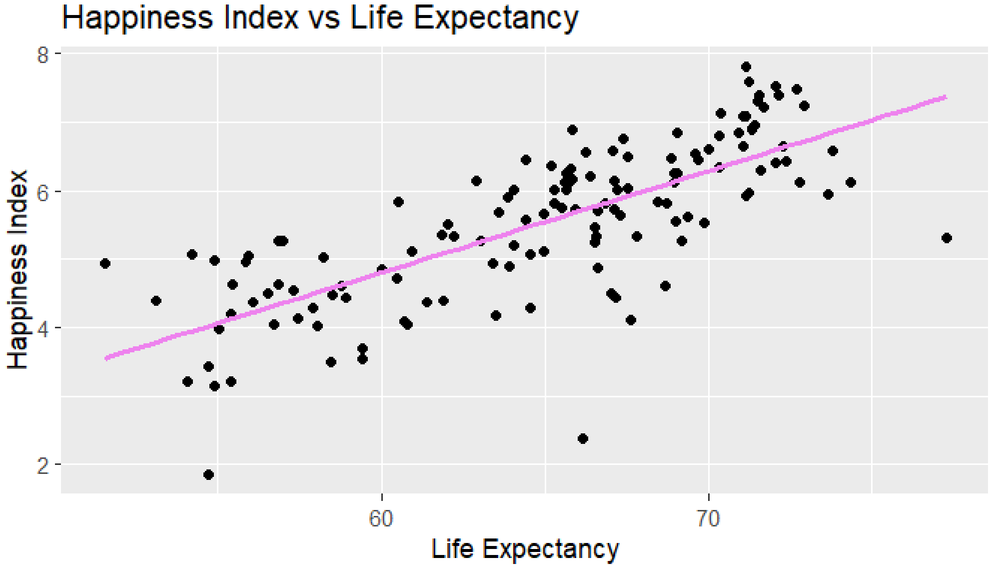

Figure 2.1

As assumed, the index of Happiness indeed increases as the Life Expectancy increases as shown in Figure 2.1.