Name	Lam Anh Huy
Student ID	s3741230
Course Code	ECON1193
Course Name	Business Statistics
Lecturer	Sienney Liu
Group	7

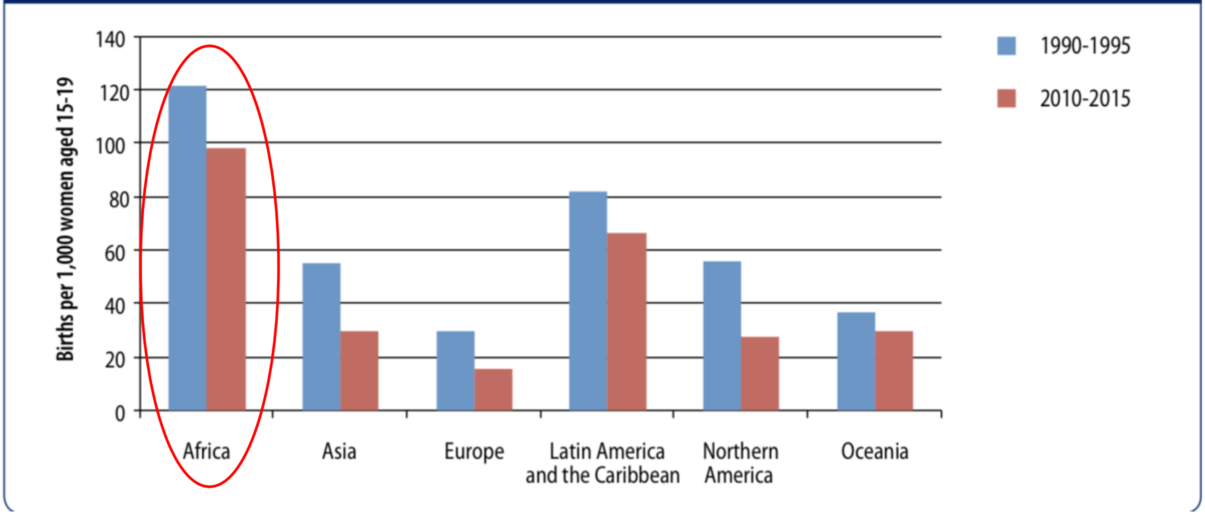
Case study set: Teen Pregnancy – Teen8

Business Statistics

Part 1: Introduction

In 2002, UNICEF estimated over 15% of global births are given by adolebscent girls (Brown 2012). However, it is reported by United Nations (2013) that there are a slightly decrease of global prevaleance of teenage pregnacy (age under 18) by 2.2 percent, from 23.3% to 20,1%. In fact, as United nation (2015) recorded, the adolescent fertility rate in all regions has been downward but there is still occured 'sharp differences' between the levels (Figure 1.1).

Figure 1.1 Adolescent birth rate by region, 1990-1995 and 2010-2015



Specifically, Africa, which is usually considered as ‘low-income’ region, remains the highest adolescent birth rate (Figure 1.1). Moreover, in developing countries, over 20,000 records of 15-18 years old girls **give birth DAILY (excluding the prenancy rate)** (United Nation Population Fund 2013). **One of the major reasons for this to be such a high rate is** unequal and inappropriate behaviors of society among girls. In addition, United Nations (2019) stated that woman and girls are **unable** to freely make their primary decisions especially regarding their sexual and reproductive rights. In fact, 90% of married adolebscent girls in developing countries are mostly under “arranged-marriage” as their parents usually decide whether who or when that their daughters can marry (Brown 2012). In addition, those who are forced to marry under age are more likely to get pregnant due to the highly exposure of unprotected sex and inability of using contraception, which is statiscally proven around 75% of adolescent pregnancy are planned (Hindin & Fatusi 2009; Presler-Marshall & Jones 2012). As United Nations Goal 5 – Gender Equality indicated it is essential that achieving equal rights for women especially for adolescences in order to achieve sustainable development. Meanwhile, adolescent girls who are pregnant have a significantly high proportion of dropping school (Hindin & Fatusi 2009; Hargreaves et al 2008). As a result, this creates the shortage of not only working labour but also creates low level of educated workforce. Moreover, some countries husbands can even legally prohibit their wives to work especially when they’re pregnant as United Nation Goal 5 mentioned. Therefore, WHO urgened that due to lack of education, and employment oppoturnities forces adolescent girls to marry under age. Overall, this leads to a decrease Gross nation income (GNI) in general.

Part 2: Descriptive Statistics

a. Central Tendency:

Central Tendency			
	High-Income	**Middle-Income**	**Low-Income**
Median	8.9614	50.2486	104.2415
Mode	#N/A	#N/A	#N/A
Mean	10.8100444	47.91215714	114.643

Table 2.1 Central Tendency (Births per 1000 woman ages 15-19)

By comparing between 3 country categories, there is a vast difference between the average adolescent fertility rate, which remarkably increases from the lowest (High-income) to the highest rate (Low-Income). This means that low-income countries have the highest average births given of teenagers with approximately 114 adolescent birth cases per 1000 woman.

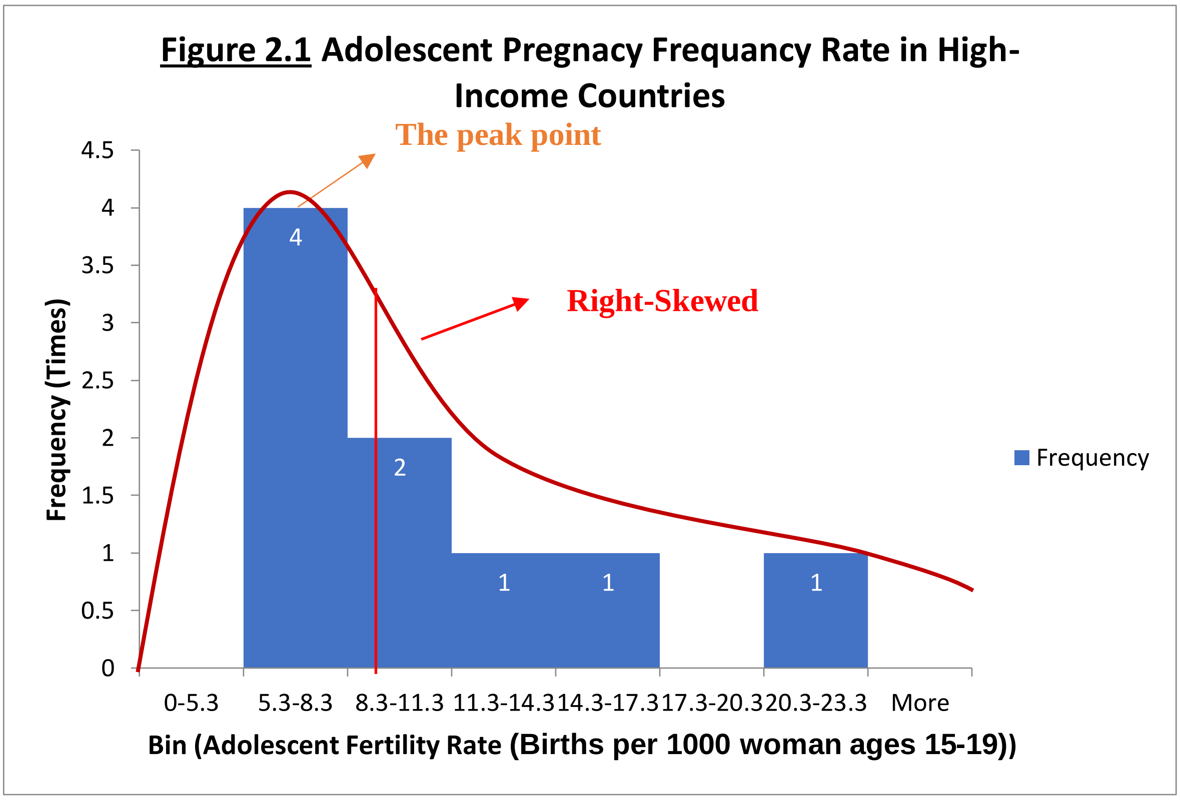
b. Variation

	High Income	**Middle Income**	**Low Income**
Mean	10.810	47.912	114.643
Standard Deviation	5.446	27.689	42.570
Sample Variance (S^2)	29.668	766.699	1812.243
Quartile 1	7.749	22.018	86.874
Quartile 3	14.0274	63.604	132.0097
Minimum	5.308	12.531	77.050
Maximum	22.441	111.218	173.0382
IQR	6.278	41.5857	45.135
CVs (%)	50.38%	57.79%	37.13%

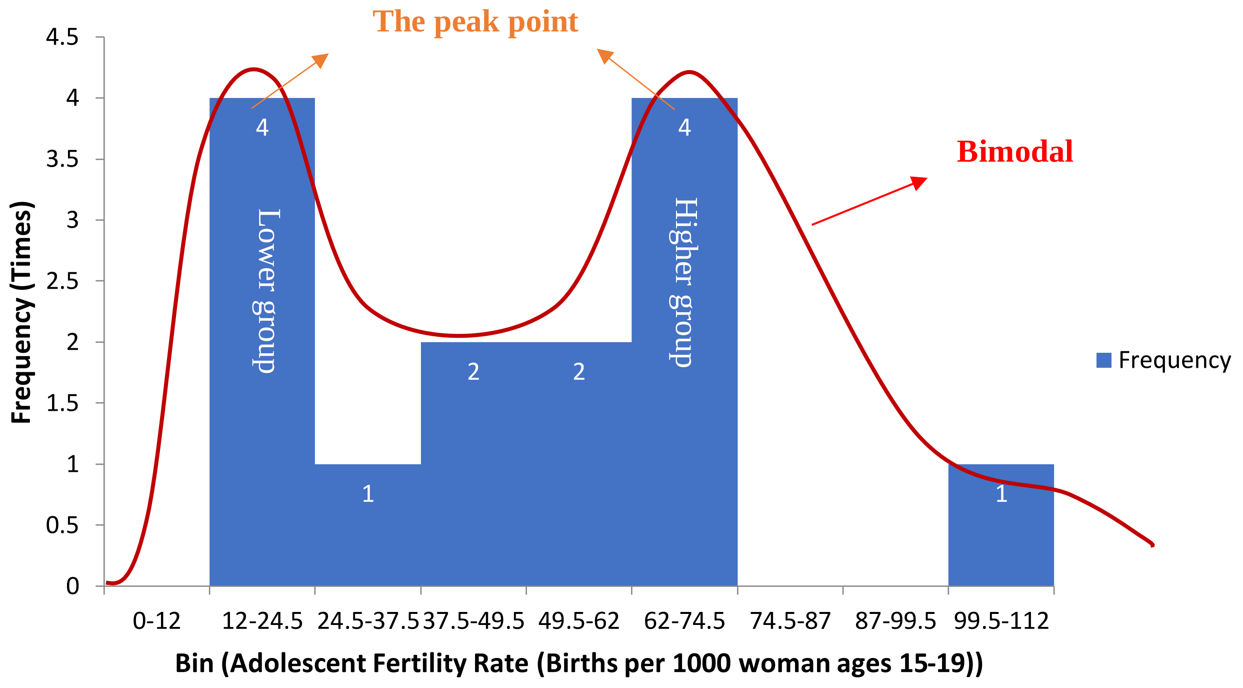
Table 2.2 Variation of each country category of adolescent fertility rate (Births per 1000 woman ages 15-19)

The first significant finding is about the differences in the standard deviations because the sample variance would be squared units; therefore, it would be hard to exactly compare them. The value of standard deviations increase heavily from high-income countries to low income countries, which means that the data from low income countries is less concentrated to the average adolescent fertility rate (**mean**) while high-income's data is more concentrated to the 'mean'. In addition, this simply reflects the data from low-income category is spread out and has diverse records in different countries. Secondly, the **interquartile range** from Middle and Low-income is obviously larger than High-Income. Therefore, it can be concluded that the data of these country categories are much more spread out to the **center value** (median). This again ensures the diversity of the data in middle and low income category. However, the CVs in 3 categories are quite low (around 37% to 57%), which means that most AFR still remain around **the average** values and is not very fluctuate.

c. Shape of each country category



<u>Figure 2.2</u> Adolescent Pregnacy Frequancy Rate in Middle-Income Countries



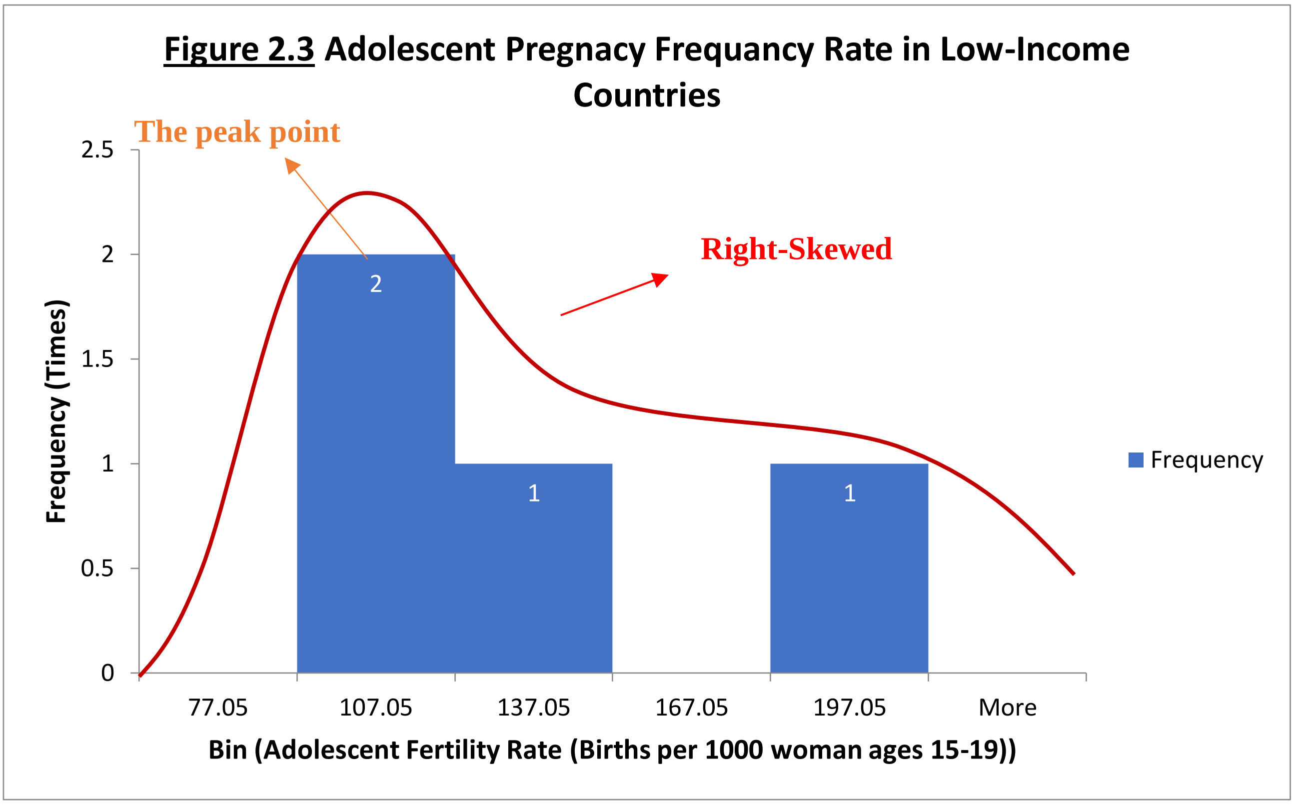


Figure 2.3 Adolescent Pregnacy Frequancy Rate in Low-Income Countries

According to the figures above, both **Figure 2.1** and **2.3** are skewed-right. This indicates that the adolescent fertility rate of **high-income** and **low-income** countries are **positively distributed** to the large amount where the fertility rate always varies **higher** than the most common adolescent fertility rate (the peak point). As a result, there are some countries might not able to follow the 'trend' in decreasing or controlling the pregnacy prevalence in teenagers. However, the most interesting part is that the middle-income rate has 2 different most common points (peak points) or technically called as bimodal distribution. This properly shows that there are 2 different gourp of middle-income countries. One group of middle-income countries (lower peak point) is actually doing very well in controlling the adolescent pregnancy and the others are simply just 'underprepared' or not yet able to decrease the pregnancy rate. However, the trend of the higher peak point group is showing the decreasing trend because the rate **is negatively distributed** (vaies to the lower adolescent fertility rate (opposite to positive distributed)).

Words: 330

<u>Part 3:</u> Confidence Intervals

a. <u>Calculation for Confidence Intervals for the world average:</u>

i. <u>Adolescent Fertility Rate</u>

<u>Level of significant:</u> $\alpha = 0.05$

Sample Size (n) = 27 → d.F = 27-1 = 26 (there are 26 degree of freedom so we could try several possibilities for 26 countries)

Sample Mean ($\overline{X}$)	45.43084
Standard Deviation (S)	41.91386

Table 3.1 Mean and Standard Deviation of Adolescent Fertility Rate (Births per 1000 woman ages 15-19)

$t_{\alpha/2,\, n-1} = t_{0.025,26} = 2.0555$

$\overline{X} \pm t_{\alpha/2,\, n-1}\frac{S}{\sqrt{n}} = 45.43083704 \pm 2.0555.\frac{41.9138562}{\sqrt{27}} = 28.8505 \leq \mu \leq 62.01117$

Intervals extend from 28.805 (Birth rate per 1000 women) to 62.01117 (Birth rate per 1000 women)

Interpretation:

- We are 95% confident is that the true average of adolescent fertility rate is between 28.8 to 62.011 births per 1000 woman.

ii. <u>Gross Nation Income (GNI)</u>

<u>Level of significant:</u> $\alpha = 0.05$

Sample Size (n) = 27 → d.F = 27-1 = 26 (there are 26 degree of freedom so we could try several possibilities for 26 countries)

Sample Mean ($\overline{X}$) ($)	18077.41
Standard Deviation (S) ($)	24922.22

Table 3.2 Mean and Standard Deviation of Gross Nation Income (current US$)

$t_{\alpha/2,\, n-1} = t_{0.025,26} = 2.0555$

$\bar{X} \pm t_{\alpha/2,\,n-1}\frac{S}{\sqrt{n}}$ = 18077.41 $\pm$ 2.0555. $\frac{24922.22}{\sqrt{27}}$ = 8218.649 $\leq \mu \leq$ 27936.17

Intervals extend from 8218.649 (current US dollars) to 27936.17 (current US dollars)

Interpretation:

- We are 95% confident is that the true average of world gross nation income is between 8218.65 to 27936.17 US Dollars per capita.

iii. <u>Domestic general government health expenditure per capita, PPP (current international $)</u>

<u>Level of significant:</u> $\alpha = 0.05$

Sample Size (n) = 27 → d.F = 27-1 = 26 (there are 26 degree of freedom so we could try several possibilities for 26 countries)

Sample Mean ($\bar{X}$) ($)	1287.677035
Standard Deviation (S) ($)	1759.790681

Table 3.3 Mean and Standard Deviation of Domestic general government health expenditure (current international $)

$t_{\alpha/2,\,n-1} = t_{0.025,26}$ = 2.0555

$\bar{X} \pm t_{\alpha/2,\,n-1}\frac{S}{\sqrt{n}}$ = 1287.677 $\pm$ 2.0555. $\frac{1759.79}{\sqrt{27}}$ = 591.58 $\leq \mu \leq$ 1983.81

Intervals extend from 591.58 (International $) to 1983.81 (International $)

Interpertation

- Based on the sample of 27 countries, we are 95% confident is that the true average of domestic general government health expenditure per capita of world is between 591.58 to 1983.81 in international dollars at purchasing power parity.
- Although the actual average may or may not be in this interval, 95% of intervals formed in repeated samples will contain the true 'mean'.

b. <u>Discusss asumption.</u>

Because the sample size is smaller than 30 **(27 countries),** we have to make an assumption that the population data is normal distributed; therefore, the sampling data would be normal distributed as well. In another point, there are no population standard deviation (measurement of how much the population's data spread out) and the mean (average) of population rate. Therefore, we have to use t-table instead of Z-table.

c. Suppose the population standard deviation is known.

In this case, we can change of using t-table to z-table: $\mu = \bar{X} \pm Z.\frac{\sigma}{\sqrt{n}}$

Choosing the same level of significance: $\alpha = 0.05 \rightarrow Z = 1.96 < t_{\alpha/2,\ n-1} = t_{0.025,26} = 2.0555$

The standard deviation (spread of data) of sample tend to be smaller than the population standard deviation due to the smaller degree of freedom (n-1) that eliminates the bias of sample variance (how much the data different from the average) (Taylor 2019).

→ The range of the new confidence intervals would be narrower (decrease).

Upper bound: $\mu = \bar{X} + t_{\alpha/2,\ n-1}\frac{S}{\sqrt{n}} < \mu = \bar{X} + Z.\frac{\sigma}{\sqrt{n}}$

Lower bound: $\mu = \bar{X} - t_{\alpha/2,\ n-1}\frac{S}{\sqrt{n}} < \mu = \bar{X} - Z.\frac{\sigma}{\sqrt{n}}$

→ The test result using population standard deviation would more accurate because:

- The variability of population would be less than sample (proved above).
- The sample sometimes base on the lucky of choosing random data; therefore, the sample could sometimes generat false driven data.

Part 4: Hypothesis Testing

a. According to the United Nation report, the world average adolescent fertility rate births per 1,000 women (ages 15-19) is **44** in 2013. However, the given data set (Teen8) gives the result of average AFR approximately **45.43** births per 1000 women. Both rates are in the calculated confidence intervals ($28.8505 \leq \mu \leq 62.01117$) in part 3a. As a result, we can predict that the trend is growing.

b. Hypothesis testing

- Step 1: Sample size = **27 < 30**. Therefore, we have to make an assumption that the population data is normal distributed; therefore, the sampling data would be normal distributed as well so it can be applied with Central Limit Thereom (CLT).

- Step 2: Null hypothesis and Alternative hypothesis.

H_0: $\mu \leq 44$ | H_a: $\mu > 44$

(The null hypothesis indicates that The AFR decrease or remain unchanged) | (The alternative hypothesis indicates that the AFR increase)

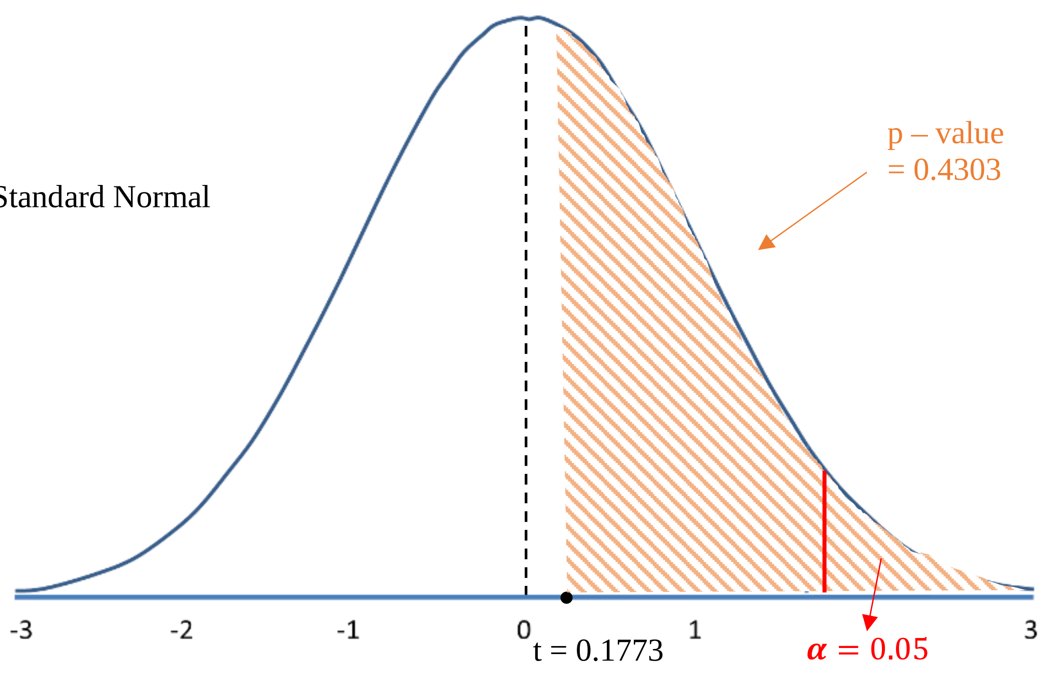
Upper tail test because H_a: $\mu > 44$ (The alternative hypothesis indicates that the AFR increase)

- Step 3: Level of significant: $\alpha = 0.05$

  Sample size: n = 27

- Step 4: Using t-table because standard deviation (σ) of population is unknown.
- The test is determined by p-value test.

Figure 4.1 Standard Normal Distribution



- Step 5: Calculate t stats.

$$t = \frac{\bar{X}-\mu}{\frac{s}{\sqrt{n}}} = \frac{45.43-44}{\frac{41.9}{\sqrt{27}}} = 0.1773$$

- Step 6: Calculate p-value.

→ **p-value** = 0.4303 **>** $\alpha = 0.05$

- Step 7: Statistical decision

We **do not reject** the null average adolescent fertility rate (H_0: $\mu \leq 44$) because the p-value is falling to the non-rejection zone.

- Null

Interpret

No rela -> Brand satis no rela to loyalty -> Thuy should not focus. On satisfac

- No null
- Weak

Weak relat -> Thuy should not focus on satisfac

- Moderate
- ➔ Consider to have an adequate invest to improve brand satisfact
- Strong
- ➔ Focus

<u>Interpretation</u>

The result indicates that the **p – value** (the probability of the test result is true) is larger that $\boldsymbol{\alpha}$ (the chance that the experiment would give a false result)**.** This means that the adolescent fertility rate is smaller than 44 births per 1000 women is true. Therefore, it indicates the decreasing trend of teenage fertility rate.

<u>Explain the possible error</u>

The possible error is type II error. This is because there is a chance of adolescent fertility rate increasing (Which means $H_a : \mu > 44$ could be correct but the data has misleaded the test result).

<u>How to minimize this error</u>

It is type II error because the p – value > α.
The risk of committing a type II error can be decreased by increasing the test power by:

- **Increasing** the sample size. The standard deviation (the broaden of data distribution) of sample (σ) would be **decreased**.
- **Increasing** the level of significance. This would **increase** $\boldsymbol{\alpha}$ (the chance that the experiment would give a false result)**.**

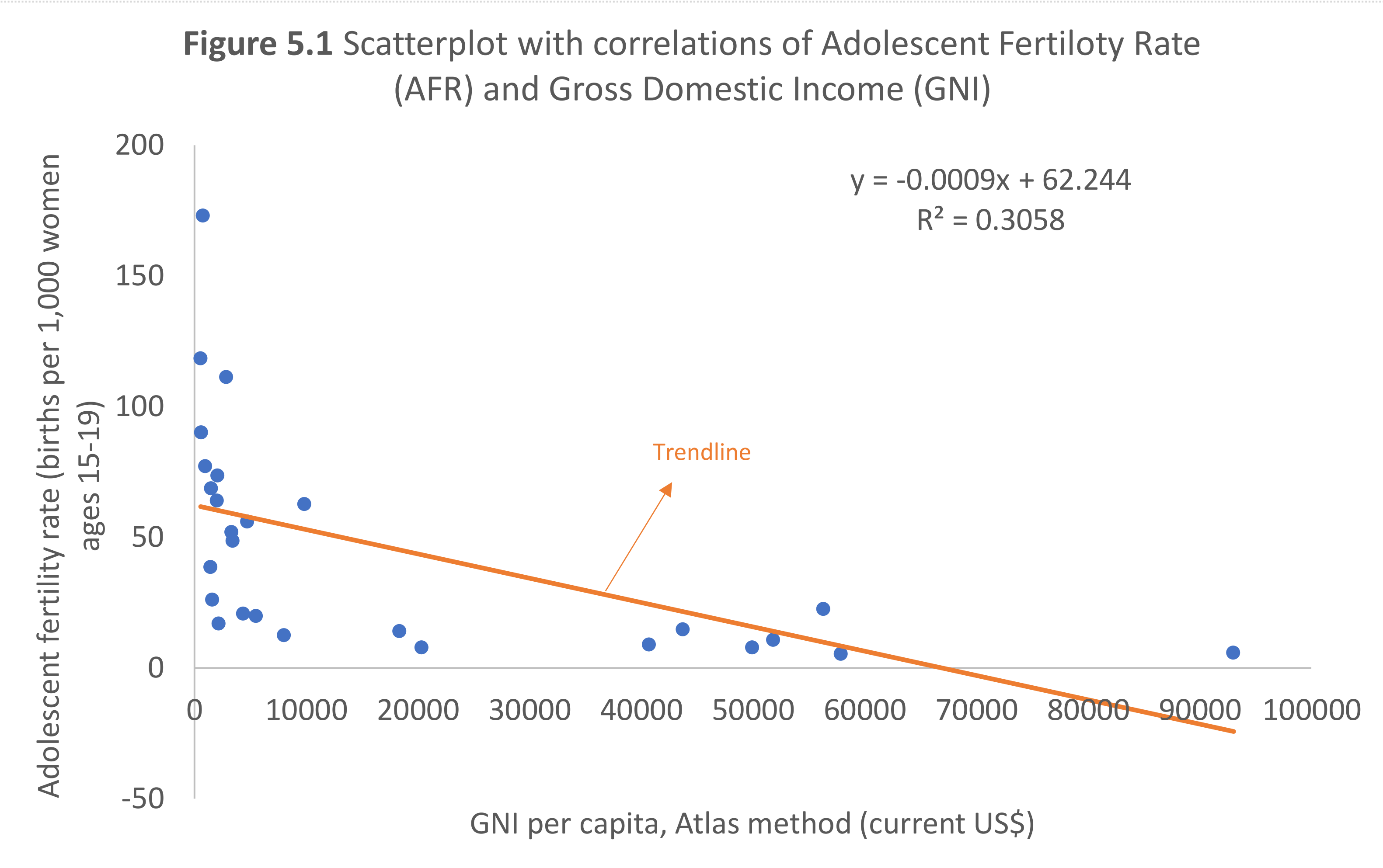
<u>Part 5:</u> Regression Analysis

a. <u>Determine dependent and independent variables.</u>

- Dependent variables (Y): Adolescent Fertili Rate (AFR).
- Independent varibles (X): Gross Nation Income (GNI), Compulsory Education (CE), Domestic General Goverment health expanditure (DGHE), and Life Expectancy (LE).

b. Each independent varibles:

i. Adolescent Fertility Rate (AFR) and Gross Nation Income (GNI).



- <u>Relationship and comment on the scatterplot:</u>

Firstly, the trend line shows AFR decline as GNI increase. Therefore, the scatterplot of AFR and GNI also indicates a sloping downward (declining) pattern as we follow from the left to right. This means that the Adolescent Fertility Rate (AFR) drops while the Gross Nation Income (GNI) increases. Hence, it is **a negative relationship** between birth rate per 1000 woman from 15 to 19 and GNI. However, the scatterplot does not explicitly show linearity and the points distribute quite to the trendline. Therefore, the two varibles seem to have **a negative relationship**.

- Linear Regression:

Regression Statistics	
Multiple R	0.55302
R Square	0.30583
Adjusted R Square	0.27806
Standard Error	35.61297
Observations	27

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	13968.973	13968.973	11.0141	0.0028
Residual	25	31707.082	1268.283		
Total	26	45676.055			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	62.2438	8.5228	7.3032	1.187E-07	44.6908	79.7969	44.6908	79.7969
GNI per capita, Atlas method (c	-0.0009	0.0003	-3.3187	0.0027736	-0.0015	-0.0004	-0.0015	-0.0004

Table 5.1 Regression summary output of adolescent fertility rate and gross nation income (current $$).

- Provide The Simple Linear Equation:

$\boxed{\text{Equation: } \hat{Y}_i = b_0 + b_1 X_i}$

→ SRF: $Adolescent\ \widehat{Fertility}\ Rate\ (AFR)$ = 62.2438 - 0.0009 (Gross Nation Income ($$$))

- Intepretation of regression of coefficients:
  - The coefficients of GNI is < 0; therefore, an increase in GNI would case a decrease in Adolescent Fertility Rate (1% increase in GNI would decrease 0.0009 (0.09%) AFR.

- Interpretation of coefficient of determination:

R Square = 0.3058 = 30.58% → The regression accounts for 30.58% of the variance. This means that there are 30.58% of variation in GNI which fall to the regression line can explicitly explain the variation in AFR. The others remaining are not related to the changes in AFR.

- <u>Test the significance of the independent value (test correlation coefficient is significant):</u>

Null and Alternative Hypothesis:

- H_0: $\beta = 0$ (If H_0 is true, there is no relationship between GNI (X) and AFR (Y))
- H_1: $\beta_1 \neq 0$ (If H_1 is true, it occurs a relationship between GNI (X) and AFR (Y))

d.f = n – k – 1 = 25

Level of significant: $\alpha = 0.05$

→ $\frac{\alpha}{2}$ = 0.025

- Critical value: $t_{\alpha/2,\, n-k-1} = t_{0.025,25} = \pm 2.059$
- Test statistics: $t = \frac{b1 - \beta 1}{SB1}$ = -3.318

→ Reject H_0 because $t_{critical} = \pm$ **2.059** > test statistics = **-3.318**

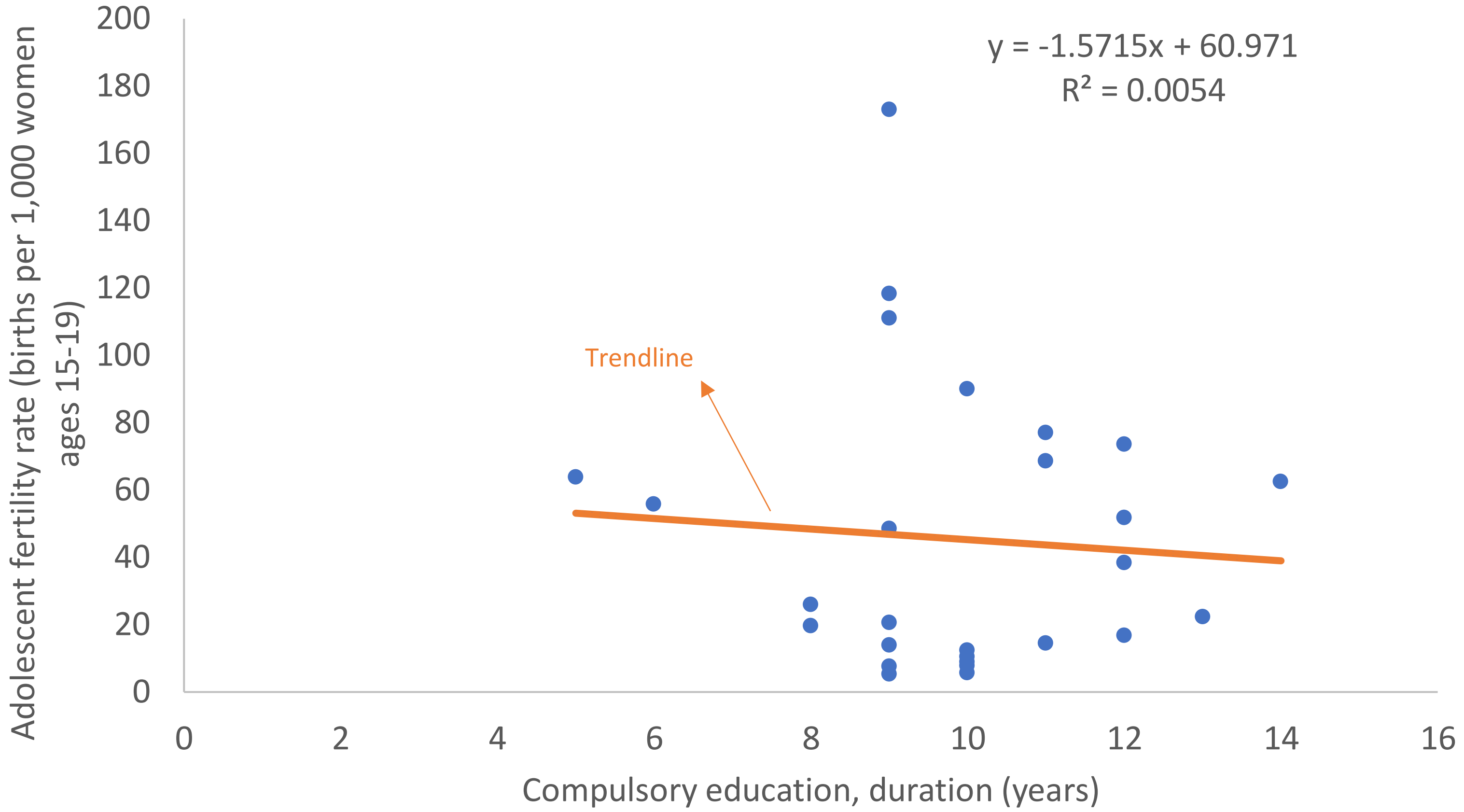
- P-value = 0.002
- $\alpha = 0.05$

→ Reject H_0 because p-value = **0.002** < $\alpha = \mathbf{0.05}$

- <u>Statistical decision:</u> It is proved that Gross Nation Income (GNI) has **a negative relationship** between Adolescent Fertility Rate (AFR) because the p-value (the probability of the test result is true) is larger than $\boldsymbol{\alpha}$ (the chance that the experiment would give a false result)**.**

ii. <u>Adolescent Fertility Rate (AFR) and Compulsory Education (CE).</u>

Figure 5.2 Scatterplots with correlations of Adolescent Fertiloty Rate (AFR) and Compulsory Education, Duration (CE)



- <u>Relationship and comment on the scatterplot:</u>

Firstly, the trend line shows AFR decline as CE increase. Therefore, the scatterplot of AFR and CE also indicates a slight sloping downward (declining) pattern as we follow from the left to right. This means that the Adolescent Fertility Rate (AFR) drops while the Compulsory Education (CE) increases. Hence, it is **a negative relationship** between birth per 1000 women ages 15 to 19 and CE. However, the scatterplot does not explicitly show linearity and distribute widely to the trendline. Therefore, the two varibles seem to have a **very weak negative** relationship.

Regression Statistics	
Multiple R	0.0738
R Square	0.0054
Adjusted R Square	-0.0343
Standard Error	42.6274
Observations	27

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	248.598	248.598	0.137	0.715
Residual	25	45427.457	1817.098		
Total	26	45676.055			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	60.9709	42.8074	1.4243	0.1667	-27.1926	149.1345	-27.1926	149.1345
Compulsory education, duration (years)	-1.5715	4.2486	-0.3699	0.7146	-10.3216	7.1787	-10.3216	7.1787

Table 5.2 Regression summary output of adolescent fertility rate and compulsory education (years).

- Provide The Simple Linear Equation:

Equation: $\hat{Y}_i = b_0 + b_1 X_i$

→ SRF: $\widehat{Adolescent\ Fertility\ Rate}\ (AFR)$ = 60.9709 – 1.5715 (Compulsory Education)

- Intepretation of regression of coefficients:

- The coefficients of GNI is < 0; therefore, an increase in Compulsory Education would case a decrease in Adolescent Fertility Rate (**1 year increase** in Compulsory **1.571 units** of Adolescent births given rate).

- Interpretation of coefficient of determination:

R Square = 0.0054 = 0.54% → The regression accounts for 0.54% of the variance. The R Square is very close to **0**. This means **that approxiamtely none** of the variation of Compulsory Education can explain the changes variation in AFR.

- Test the significance of the independent value(test correlation coefficient is significant):

Null and Alternative Hypothesis:

- Education would decrease **1**
- H_0: $\beta = 0$ (If H_0 is true, there is no relationship between CE (X) and AFR (Y))
- H_1: $\beta_1 \neq 0$ (If H_1 is true, it occurs a relationship between CE (X) and AFR (Y))

d.f = n – k – 1 = 25

Level of significant: $\alpha = 0.05$

→ $\frac{\alpha}{2}$ = 0.025

- Critical value: $t_{\alpha/2,\, n-k-1} = t_{0.025,25} = \pm 2.059$
- Test statistics: $t = \frac{b1 - \beta 1}{SB1}$ = -0.3699

→ Do not reject H_0 because $t_{critical} = \pm$ **2.059** < test statistics = **-0.3699**

- P-value = 0.7146
- $\alpha = 0.05$

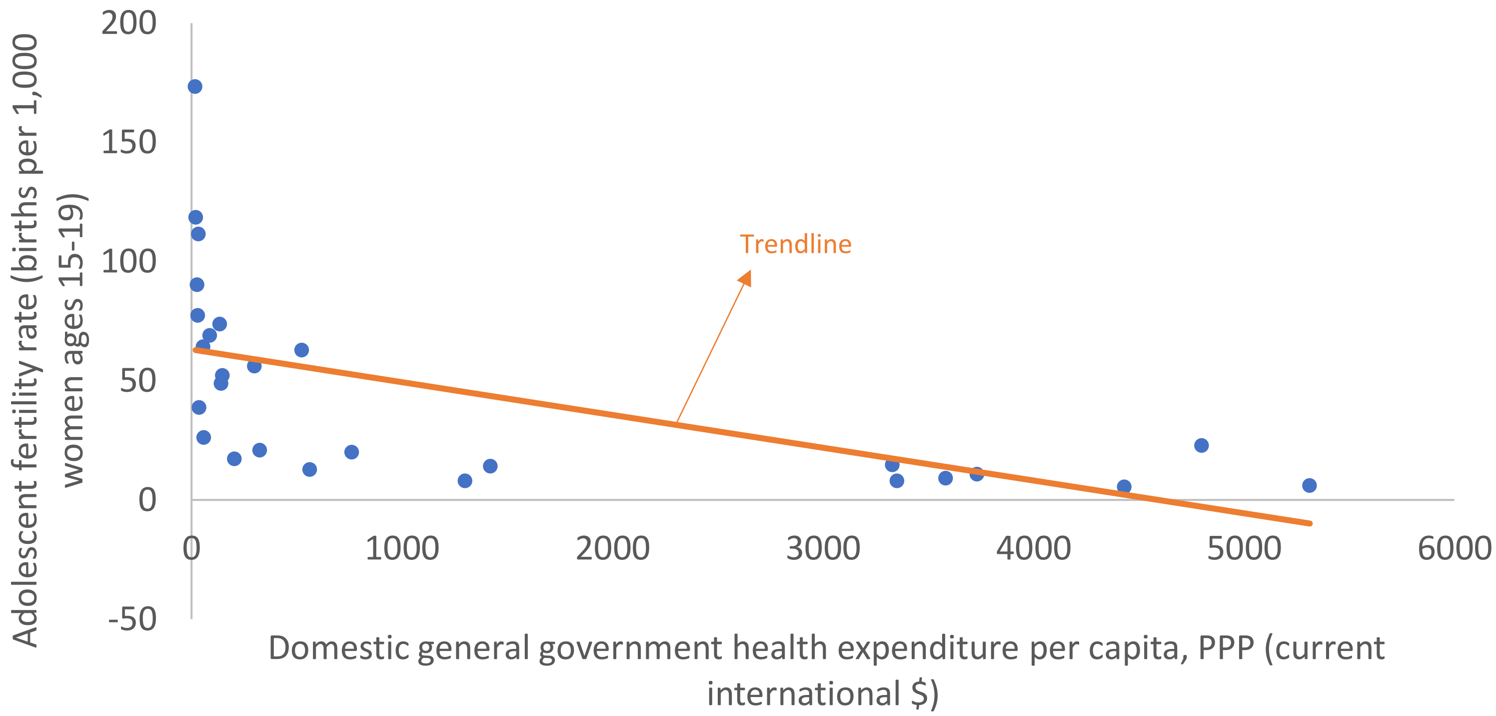
→ Do not reject H_0 because p-value = **0.7146** > $\alpha = \mathbf{0.05}$

- <u>Statistical decision:</u> It is proved that Compulsory Education (CE) has **no relationship** between Adolescent Fertility Rate (AFR) because the p-value (the probability of the test result is true) is **smaller** than $\boldsymbol{\alpha}$ (the chance that the experiment would give a false result)**.**

○

iii. <u>Adolescent Fertility Rate (AFR) and Domestic General Goverment health expanditure (DGHE).</u>

Figure 5.3 Scatterplots with correlations of Adolescent Fertiloty Rate and Domestic General Gov Expanditure per capita



- Relationship and comment on the scatterplot:

Firstly, the trend line shows AFR decline as DGHE increase. Therefore, the scatterplot of AFR and DGHI also indicates a sloping downward (declining) pattern as we follow from the left to right. This means that the Adolescent Fertility Rate (AFR) drops while the Domestic General Goverment health expanditure (DGHE) increases. Hence, it is **a negative relationship** between birth rate 1000 women ages 15 to 19 and DGHE. However, the scatterplot does not explicitly show linearity and distribute **quite closely to the trendline**. Therefore, the two varibles seem to have **a moderate negative** relationship.

Regression Statistics	
Multiple R	0.577
R Square	0.333
Adjusted R Square	0.306
Standard Error	34.910
Observations	27

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	15209.012	15209.012	12.480	0.002
Residual	25	30467.043	1218.682		
Total	26	45676.055			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	63.128	8.380	7.533	6.914E-08	45.868	80.388	45.868	80.388
Domestic general government health expenditure	-0.014	0.004	-3.533	0.0016	-0.022	-0.006	-0.022	-0.006

Table 5.3 Regression summary output of adolescent fertility rate and Domestic General Goverment health expanditure

- Provide The Simple Linear Equation:

$$\boxed{\text{Equation: } \hat{Y}_i = b_0 + b_1 X_i}$$

→ SRF: $\widehat{Adolescent\ Fertility\ Rate\ (AFR)}$ = 63.128 – 0.014 (Domestic general government health expenditure)

- Intepretation of regression of coefficients:
  - The coefficients of DGHE is < 0; therefore, an increase in DGHE would case a decrease in Adolescent Fertility Rate (AFR) (1$ increase in DGHE would decrease 0.014 units of AFR).

- Interpretation of coefficient of determination:

R Square = 0.333 = 33.3% → The regression accounts for 33.3% of the variance. This means that there are 33.3% of variation in DGHE which fall to the regression line can explicitly explain the variation for the changes of AFR. The others remaining are not related to the changes in AFR variation.

- <u>Test the significance of the independent value (test correlation coefficient is significant):</u>

Null and Alternative Hypothesis:

- H_0: β = 0 (If H_0 is true, there is no relationship between DGHE (X) and AFR (Y))
- H_1: $\beta_1 \neq 0$ (If H_1 is true, it occurs a relationship between DGHE (X) and AFR (Y))

d.f = n – k – 1 = 25

Level of significant: $\alpha = 0.05$

→ $\frac{\alpha}{2}$ = 0.025

- Critical value: $t_{\alpha/2,\, n-k-1} = t_{0.025,25} = \pm 2.059$
- Test statistics: $t = \frac{b1 - \beta 1}{SB1}$ = - 3.533

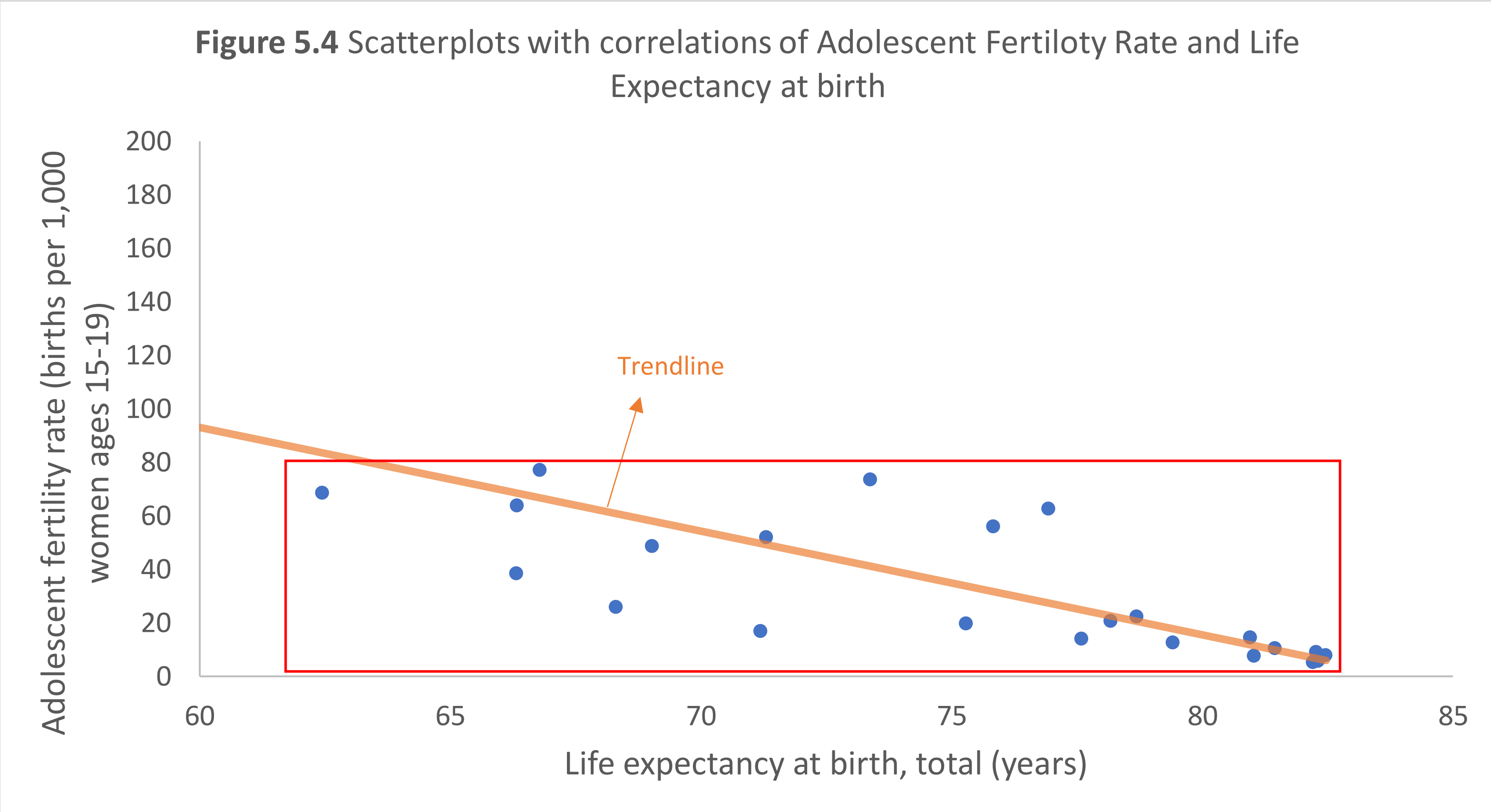
→ Reject H_0 because $t_{critical} = \pm$ **2.059** > test statistics = **-3.533**

- P-value = 0.7146
- $\alpha = 0.05$

→ Reject H_0 because p-value = $\mathbf{0.0016} < \alpha = \mathbf{0.05}$

- <u>Statistical decision:</u> It is proved that Domestic general government health expenditure per capita has **a relationship** between Adolescent Fertility Rate (AFR) because the p-value (the probability of the test result is true) is larger than $\boldsymbol{\alpha}$ (the chance that the experiment would give a false result).

iv. **<u>Adolescent Fertility Rate (AFR) and Life Expectancy at birth (LE)</u>**



- <u>Relationship and comment on the scatterplot:</u>

Firstly, the trend line shows AFR decline as Life Expectancy increase. Therefore, the scatterplot of AFR and LE also indicates a sloping downward (declining) pattern as we follow from the left to right. This means that the Adolescent Fertility Rate (Y-value) drops while the Life Expectancy (X-value) increases. Hence, it is **a negative relationship** between births rate 1000 women ages 15 to 19 and Life Expectancy at birth. Moreover, the points of the scatter plot resemble closely to the rendline. Therefore, the two varibles seem to have a **strong** relationship.

Regression Statistics	
Multiple R	0.859
R Square	0.738
Adjusted R Square	0.727
Standard Error	21.889
Observations	27

ANOVA

	df	SS	MS	F	Significance F
Regression	1	33697.561	33697.561	70.329	9.84101E-09
Residual	25	11978.494	479.140		
Total	26	45676.055			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	325.592	33.672	9.670	6.2892E-10	256.244	394.940	256.244	394.940
Life expectancy at birth, total (years)	-3.876	0.462	-8.386	9.84101E-09	-4.828	-2.924	-4.828	-2.924

Table 5.4 Regression summary output of adolescent fertility rate and Life expectancy at birth (years)

- Provide The Simple Linear Equation:

$$\boxed{\text{Equation: } \hat{Y}_i = b_0 + b_1 X_i}$$

→ SRF: $\widehat{Adolescent\ Fertility\ Rate\ (AFR)}$ = 325.592 – 3.876 (Life Expectancy at Birth (LE))

- Intepretation of regression of coefficients:

- The coefficients of LE is < 0; therefore, an increase in LE would case a decrease in Adolescent Fertility Rate (AFR) (1 year increase in LE would decrease 3.876 units of AFR).

- Interpretation of coefficient of determination:

R Square = 0.738 = 73.8% → The regression accounts for 73.8% of the variance. This means that there are 73.8% of variation in Life Expectancy at birth which fall to the regression line can explicitly explain the variation in AFR. The others remaining are not related to the changes in AFR variation.

- Test the significance of the independent value:

Null and Alternative Hypothesis:

- H_0: $\beta = 0$ (If H_0 is true, there is no relationship between LE (X) and AFR (Y))
- H_1: $\beta_1 \neq 0$ (If H_1 is true, it occurs a relationship between LE (X) and AFR (Y))

d.f = n – k – 1 = 25

Level of significant: $\alpha = 0.05$

→ $\frac{\alpha}{2}$ = 0.025

- Critical value: $t_{\alpha/2,\, n-k-1} = t_{0.025,25} = \pm 2.059$
- Test statistics: $t = \frac{b1 - \beta 1}{SB1}$ = - 8.386

→ Reject H_0 because $t_{critical} = \pm$ **2.059** > test statistics = **-8.386**

- P-value = 6.2892E - 10
- $\alpha = 0.05$

→ Reject H_0 because p-value = **6.2892E - 10** < $\boldsymbol{\alpha = 0.05}$

- Statistical decision: It is proved that Life Expectancy at birth has **a relationship** between Adolescent Fertility Rate (AFR) because the p-value (the probability of the test result is true) is larger than $\boldsymbol{\alpha}$ (the chance that the experiment would give a false result)**.**

c. **Recommendation for furthur research on adolescent fertility rate.**

Based on the simple linear regression above, the varibles of **AFR and Life Expectancy at birth** would be the most suitable because it has the highest **R-Square (0.738);** therefore, our prediction would be most precise. This is because low R Square shows that the model occurs more error (Minitab 2014). Moreover, the points of the scatterplot (figure 5.4) are closer to the trendline, which reflecs less uncertainty in predictions.

Part 6: Conclusion

- **Answer:**

After making several calculations and research, we can anticipate that the global adolescent fertility rate is actually decreasing in genernal. Moreover, the adolescent fertility rate is increasing towards the decrease in income rate. There are 3 significant findings on Adolescent Fertility Rate in the Descriptive Statistics, Hypothesis Testing, and Regression parts.

- **Calculations, Citation and Explanation:**

- Firstly, the main finding in the Descriptive Statistics shows that the average teenage fertility rate increases dramatically from high-income countries to low-income countries. This is due to their inability in investing for controlling fertility rate. As proven by Price (2013), the income rate and fertility has strong negative correlation between each other due to factors such as working ages, capital, healthcare and cultural factors (Ashraf et al 2013). Therefore, the adolescent fertility rate is undoubtedly high in low-income countries as the consequence of poverty and the shortage of capital investment in government.

- Secondly, the Hypothesis testing gives the results of '**Not rejecting** the dereasing rate of global adolescent fertility rate'. This proves that the trend of adolescent fertility rate is actually decreasing in general. According to United Nations (2013), the reduction of global adolescent childbearing due to the increase in school participation, and using contraception. Moreover, the decrease in adolescent females marriage also contributes as a strong factor for the declining in AFR. This is because in the prenant period, adolescent women would prevented from advancing their knowlegde and limit their ability and opportunity for employment (Haub 2013). In addition, as Goal 5 of United Nation mentioned, gender equality is being promoted aggressively in many developing and low income countries, which enhances females' decisions in having babies and changes people's perception.
- Finally, the strong negative relationship between AFR and Life Expectancy at birth also indicates the prediction of global AFR declining trend. According to United Nations Revision (2015), the Life Expectancy at birth goal have achieved by increasing from 67 years to 70 years globally. Moreover, it is project to increase to 77 by 2045 and to 83 by 2100 (UN 2015). Consequently, as proven in the regression analysis (Part 5), the improvement of Life Expectancy at birth would result in decreasing the Adolescent Childbearing. For furthur research, the improvements of LE usually come from factors such as better health service (which reduces child death and death by giving birth in adolescent pregnancies cases), childs infections (pneumonia, diarrhoea and malaria) by giving sufficient vacination and well treated capital.

- **Summary:**

Even though there are differences in AFR by the diffences in income levels, the adolescent childbearing in gernaral is decreasing due to the optimist world situations as well as the development of countries. Moreover, the sample data is statistically proven by different mathmetical equations to be following the population trend as decreasing teenage giving births. As a result, the world trend in AFR is clearly indicated as same as the prediction from regression analysis.

Words: 473

Part 7: Reference List

Ashraf, Q et al 2013, 'The Effect of Fertility Reduction on Economic Growth.', *Population and Development Review*, vol.39, no.1, pp.97–130, *JSTOR*, <https://www.jstor.org/stable/41811954?seq=4#metadata_info_tab_contents>.

Brown, G 2012, 'Out of wedlock, into school: combating child marriage through education', London: The Office of Gordon and Sarah Brown.

Haub, C 2013, 'Trends in Adolescent Fertility Rate Mixed Picture', 20 May, *Population Reference Bureau*, viewed 21 April 2019, <https://www.prb.org/adolescent-fertility/>.

Hindin, M, & Fatusi, A 2009, 'Adolescent sexual and reproductive health in developing countries: An overview of trends and interventions. International Perspectives on Sexual Reproductive Health', *International Perspectives on Sexual and Reproductive Health,* Vol. 35, no.2, pp.58–62. doi:10.1363/ipsrh.35.058.09.

Ogee, A, Ellis, M, Scibilla, B, Pammer, C, and Steelee, C 2014, *How High Should R-squared Be in Regression Analysis?*, blog, viewed 20 April 2019, <https://blog.minitab.com/blog/adventures-in-statistics-2/how-high-should-r-squared-be-in-regression-analysis>.

Price, J 2013, *How Income Affects Fertility*, blog, viewed 19 April 2019, <https://ifstudies.org/blog/how-income-affects-fertility>.

Psaki, R 2014, 'Addressing early marriage and adolescent pregnancy as a barrier to gender parity and equality in education.', *Background Paper for the 2015 UNESCO Education for All Global Monitoring Report*, New York: Population Council.

Taylor, C 2019, ‘Differences Between Population and Sample Standard Deviations’, *ThoughtCo,* 23 January, viewed 20 April 2019, <https://www.thoughtco.com/population-vs-sample-standard-deviations-3126372>.

United Nations Population Funds 2019, *Vast numbers of women lack decision-making power over their own bodies, says UNFPA flagship report,* UNFPA, New York. <https://www.unfpa.org/news/vast-numbers-women-lack-decision-making-power-over-their-own-bodies-says-unfpa-flagship-report>.

United Nations Population Funds 2013, *Adolescent Pregnancy: A Review of the Evidence,* UNFPA, New York, <https://www.unfpa.org/sites/default/files/pub-pdf/ADOLESCENT%20PREGNANCY_UNFPA.pdf>.

United Nations Population Division 2015, *World Fertility Patterns 2015*, Databooklet, Department of Economic and Social Affairs, (ST/ESA/ SER.A/370), <https://www.un.org/en/development/desa/population/publications/pdf/fertility/world-fertility-patterns-2015.pdf>.

United Nation n.d., *Overview of Adolescent Fertility Rate, UNFPA,* <https://www.unfpa.org/adolescent-pregnancy>.

United Nations Population Funds 2013, *State of World Population 2013,* UNFPA, New York, <https://www.unfpa.org/sites/default/files/pub-pdf/EN-SWOP2013.pdf>.

United Nations 2013, *Adolescent Fertility since the International Conference on Population and Development (ICPD) in Cairo*, United Nation, Cairo, <https://www.un.org/en/development/desa/population/publications/pdf/fertility/Report_Adolescent-Fertility-since-ICPD.pdf>.

UNFPA 2015, *Girlhood, not motherhood: Preventing adolescent pregnancy,* UNFPA, New York.

UNICEF 2013, *Ending child marriage: Progress and prospects*, UNICEF, New York.

WHO 2018, *Adolescent Pregnancy,* WHO, <https://www.who.int/news-room/fact-sheets/detail/adolescent-pregnancy>.