



# **TUNKU ABDUL RAHMAN UNIVERSITY COLLEGE**

## **FACULTY OF COMPUTING AND INFORMATION TECHNOLOGY**

Bachelor of Science in Management Mathematics with Computing

Year 3 Semester 1

RMM (Tutorial Group 2 & 3)

BAMS3043 Mathematical and Statistical Software

### **Assignment 4**

<b>Name</b>	<b>Group</b>	<b>Student ID</b>
1. Lim Hui Jing	3	20WMR09183
2. Yee Jun Hoong	3	20WMR09194
3. Chong Ying Hui	2	20WMR06049

**Lecturer / Tutor's name: Dr. Tan Yan Bin**

## Justification of Selected Independent Variable

In this assignment, we only use the data of 8 countries which are around Malaysia. The 8 countries are Brunei Darussalam, Cambodia, Indonesia, Malaysia, Myanmar, Philippines, Singapore and Thailand.

### 1. Task 1

#### 1.1 Model 1

Independent variable: BMI

References:

- <https://www.medscape.com/answers/123702-11510/does-obesity-reduce-life-expectancy>

### 2. Task 2

#### 2.1 Model 2

##### **Independent variable 1: Alcohol**

References:

- <https://vertavahealthmississippi.com/blog/alcohol-abuse-and-life-expectancy/>
- <https://towardsdatascience.com/correlation-causation-how-alcohol-affects-life-expectancy-a68f7db943f8>
- <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4402015/>

##### **Independent variable 2: HIV/AIDS**

References:

- <https://www.aidsmap.com/about-hiv/life-expectancy-people-living-hiv>
- <https://www.aidsmap.com/news/mar-2020/yes-same-life-expectancy-hiv-negative-people-far-fewer-years-good-health>

##### **Independent variable 3: Schooling**

References:

- <https://www.newscientist.com/article/2166833-more-education-is-what-makes-people-live-longer-not-more-money/>
- <https://genus.springeropen.com/articles/10.1186/s41118-019-0055-0>
- <https://journals.sagepub.com/doi/pdf/10.1177/2372732214549754>

## 2.2 Model 3

### **Independent variable 1: Alcohol**

References:

- <https://vertavahealthmississippi.com/blog/alcohol-abuse-and-life-expectancy/>
- <https://towardsdatascience.com/correlation-causation-how-alcohol-affects-life-expectancy-a68f7db943f8>
- <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4402015/>

### **Independent variable 2: HIV/AIDS**

References:

- <https://www.aidsmap.com/about-hiv/life-expectancy-people-living-hiv>
- <https://www.aidsmap.com/news/mar-2020/yes-same-life-expectancy-hiv-negative-people-far-fewer-years-good-health>

### **Independent variable 3: BMI**

References:

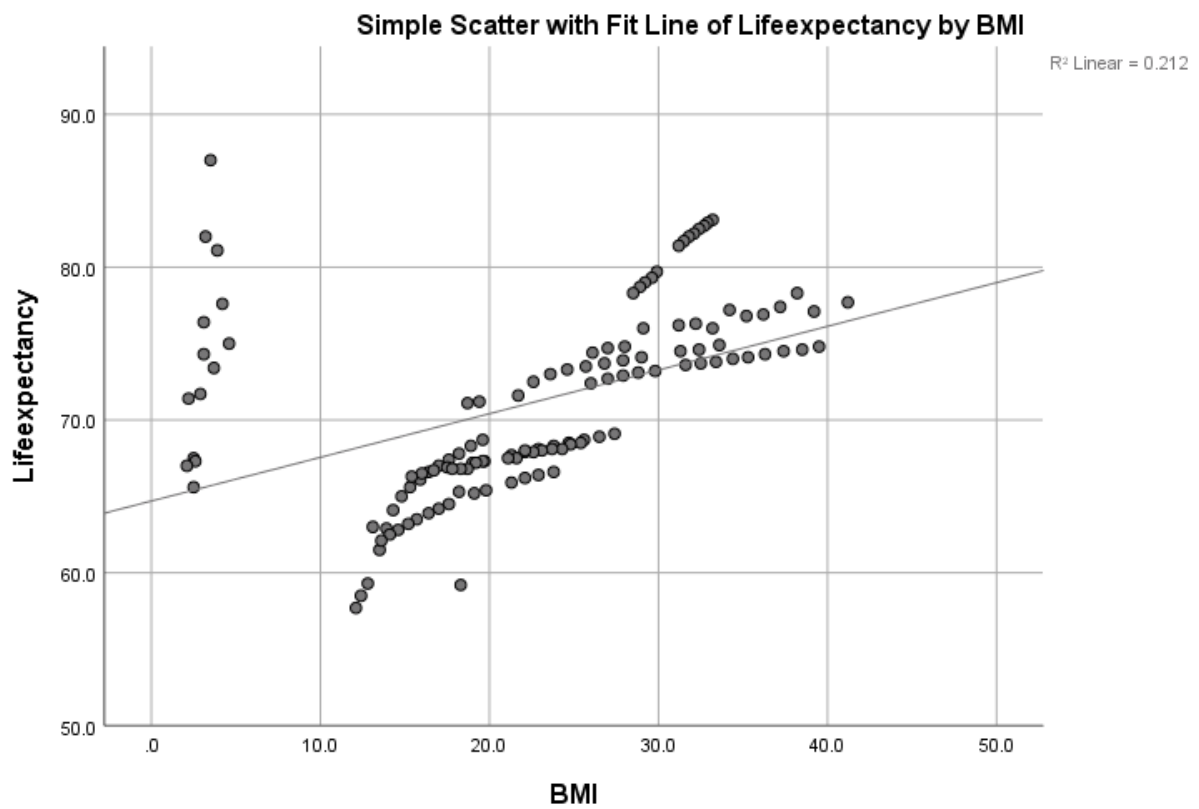
- <https://www.medscape.com/answers/123702-11510/does-obesity-reduce-life-expectancy>

# Task 1

## Model 1

### GGraph

[DataSet1]



## Regression

### Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	BMI <sup>b</sup>	.	Enter

a. Dependent Variable: Lifeexpectancy

b. All requested variables entered.

### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.460 <sup>a</sup>	.212	.205	5.3926

a. Predictors: (Constant), BMI

### ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	983.006	1	983.006	33.803	.000 <sup>b</sup>
	Residual	3664.143	126	29.081		
	Total	4647.149	127			

a. Dependent Variable: Lifeexpectancy

b. Predictors: (Constant), BMI

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	64.699	1.193		54.218	.000
	BMI	.286	.049	.460	5.814	.000

a. Dependent Variable: Lifeexpectancy

### Summary

A simple linear regression was performed to qualify the relationship between BMI and life expectancy. A sample of 8 countries are used in this analysis.

Results showed that there was a statistically significant relationship between BMI and life expectancy (  $t = 5.814$ ,  $p < 0.000$ ) and BMI accounted for 21.2% of explained variability in life expectancy.

Estimated life expectancy in this model =

$$64.699 + 0.286 * (BMI)$$

.

## Task 2

### Model 2

#### → Regression

##### Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	Alcohol, HIVAIDS, Schooling <sup>b</sup>	.	Enter

a. Dependent Variable: Lifeexpectancy

b. All requested variables entered.

##### Model Summary<sup>b</sup>

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.896 <sup>a</sup>	.802	.797	2.7708

a. Predictors: (Constant), Alcohol, HIVAIDS, Schooling

b. Dependent Variable: Lifeexpectancy

##### ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3638.590	3	1212.863	157.977	.000 <sup>b</sup>
	Residual	898.261	117	7.677		
	Total	4536.851	120			

a. Dependent Variable: Lifeexpectancy

b. Predictors: (Constant), Alcohol, HIVAIDS, Schooling

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	40.692	1.994		20.409	.000	36.744	44.641
	Schooling	2.623	.157	.862	16.673	.000	2.311	2.934
	HIVAIDS	-.896	.747	-.062	-1.200	.233	-2.374	.583
	Alcohol	-.157	.121	-.054	-1.294	.198	-.398	.083

a. Dependent Variable: Lifeexpectancy

### Coefficient Correlations<sup>a</sup>

Model			Alcohol	HIVAIDS	Schooling
1	Correlations	Alcohol	1.000	-.135	-.178
		HIVAIDS	-.135	1.000	.598
		Schooling	-.178	.598	1.000
	Covariances	Alcohol	.015	-.012	-.003
		HIVAIDS	-.012	.557	.070
		Schooling	-.003	.070	.025

a. Dependent Variable: Lifeexpectancy



### Summary

A multiple linear regression was performed to quantify the relationship between schooling, HIV/AIDS, alcohol and life expectancy. A sample of 8 countries was used in the analysis.

From the model summary, 80.2% of the variation in life expectancy can be explained by schooling, HIV/AIDS and schooling.

From the coefficients table, the p-value for the explanatory variable schooling is (.000). Since the p-value is less than 0.05, we can conclude that the schooling has a statistically significant association with life expectancy. Next, the p-value for the explanatory variable HIV/AIDS is (.233). Since the p-value is more than 0.05, we cannot conclude that HIV/AIDS has a statistically significant association with life expectancy. The p-value for the explanatory variable alcohol is (.198). Since the p-value is more than 0.05, we cannot conclude that alcohol has a statistically significant association with life expectancy.

Estimated life expectancy in this model =

$$40.692 + 2.623 * (Schooling) - 0.896 * (HIV/AIDS) - 0.157 * (Alcohol)$$

### Model 3

## Regression

### Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	BMI, Alcohol, HIVAIDS <sup>b</sup>	.	Enter

a. Dependent Variable: Lifeexpectancy

b. All requested variables entered.

### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.646 <sup>a</sup>	.417	.402	4.7543

a. Predictors: (Constant), BMI, Alcohol, HIVAIDS

### ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1892.304	3	630.768	27.906	.000 <sup>b</sup>
	Residual	2644.547	117	22.603		
	Total	4536.851	120			

a. Dependent Variable: Lifeexpectancy

b. Predictors: (Constant), BMI, Alcohol, HIVAIDS

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	68.172	1.405		48.515	.000
	Alcohol	.294	.206	.101	1.427	.156
	HIVAIDS	-6.695	1.101	-.461	-6.083	.000
	BMI	.201	.049	.316	4.143	.000

a. Dependent Variable: Lifeexpectancy

### Summary

A multiple linear regression was performed to quantify the relationship between BMI, alcohol, HIV/AIDS and life expectancy. A sample of 8 countries was used in the analysis.

From the model summary, 41.7% of the variation in life expectancy can be explained by BMI, alcohol and HIV/AIDS. .

From the coefficients table, the p-value for the explanatory variable percentage alcohol is (.156). Since the p-value is more than 0.05, we cannot conclude that the alcohol has a statistically significant association with life expectancy. Next, the p-value for the explanatory variable HIV/AIDS is (.000). Since the p-value is less than 0.05, we can conclude that HIV/AIDS has a statistically significant association with life expectancy. The p-value for the explanatory variable BMI is (.000). Since the p-value is less than 0.05, we can conclude that BMI has a statistically significant association with life expectancy.

Estimated life expectancy in this model =




$$68.172 + 0.294 * (Alcohol) - 6.695 * (HIV/AIDS) + 0.201 * (BMI)$$

### **Task 3**

Based on the three models shown above, we can conclude that the model 2 is the best model. The reason is model 2 has the highest R squared value, 80.2%, which means that 80.2% of the variation in life expectancy can be explained by independent variables in model 2. However the R squared value of model 1 is 21.2% and 41.7% for model 3.

## Task 4

By using model 2, we do a prediction of life expectancy by using the mean for each of the independent variables, and we found the 95% confidence interval of life expectancy is  $65.60725 \leq Y(\text{Life Expectancy}) \leq 76.62745$

 PRE_1	 LIC1_1	 UICI_1
71.11735	65.60725	76.62745

We used the frequencies table to find the central tendency of the independent variable. Then, we used the mean value of each independent variable to do the prediction above.

Statistics		Schooling	HIVAIDS	Alcohol
N	Valid	128	128	121
	Missing	0	0	7
Mean		11.820	.311	1.8964
Median		11.900	.100	.8500
Mode		12.9	.1	.01