



សាភលខិន្យាល័យ ភូមិន្លភ្លំពេញ ROYAL UNIVERSITY OF PHNOM PENH

DEPARTMENT OF INTERNATIONAL BUSINESS MANAGEMENT-IBM

APPLIED RESEARCH METHODOLOGY

BACHELOR PROGRAM

4[™] YEAR, SEMESTER 1

MATERIAL INFO

Authors: Veasna SOU., Sambath PHOU., & Phichhang OU Published online at <u>www.spss.site</u>: 13-Sept 2024 Revised version: 17-Dec-2024 © 2024 Published by Research Office, (RUPP) All rights reserved.

TUTORING SERIES-07: Simple and Multiple Regression Analysis

SIMPLE REGRESSION

AND

MULTIPLE REGRESSION ANALYSIS

PURPOSES:

Simple and multiple regression analysis in SPSS is used to understand relationships between variables and make predictions based on those relationships (Groebner et al., 2018). Both analyses can be performed using the built-in regression functions, allowing users to easily input data, choose variables, and interpret results through output tables and graphs. Used in various fields such as social sciences, marketing, health research, and any discipline where understanding or predicting relationships among variables is important.

SIMPLE REGRESSION:

In simple linear regression, each observation consists of two variables: one for the independent variable and one for the dependent variable (Anderson et al., 2014).

- To assess the relationship between one independent variable (predictor) and one dependent variable (outcome).
- Key Uses:
 - Predicting the value of the dependent variable based on the independent variable.
 - Understanding the strength and direction of the relationship (positive or negative).
 - Calculating the effect size and assessing the significance of the predictor variable.
- Rules and Guidelines:
 - **Sample Size:** A minimum of 15-20 observations for each predictor variable is often recommended. However, more data is preferable to achieve reliable results.
 - **Linearity:** The relationship between the independent and dependent variable should be linear. This can be visually assessed using scatter plots.
 - **Normality:** Residuals (errors) should be approximately normally distributed, especially for inference tests (like t-tests on regression coefficients).
 - **Homoscedasticity:** The variance of residuals should be constant across all levels of the independent variable(s). This means that the spread of residuals should not increase or decrease as predicted values increase.
 - **No Multicollinearity:** Since there is only one predictor in simple regression, this rule does not directly apply, but you should ensure the predictor is not perfectly correlated with another variable if it's included in a subsequent analysis.

MULTIPLE REGRESSION:

Multiple regression analysis is the study of how a dependent variable y is related to two or more independent variables (Anderson et al., 2014).

- To examine the relationship between one dependent variable and multiple independent variables.
- Key Uses:
 - o Identifying the most impactful predictors on the dependent variable.
 - Evaluating how each independent variable contributes to the prediction of the dependent variable, adjusting for the effects of other predictors.
 - Understanding interaction effects among independent variables.
 - Conducting hypothesis tests and obtaining confidence intervals for the regression coefficients.
- Rules and Guidelines:



- Sample Size: Ideally, at least 10-15 observations per predictor variable in the model. However, more data can help increase the robustness of the model.
- **Multicollinearity:** Check for high correlations among independent variables. Variance Inflation Factor (VIF) values greater than 10 may indicate problematic multicollinearity.
- **Linearity:** The relationship between each independent variable and the dependent variable should remain linear. Consider transformations if necessary.
- **Normality of Residuals:** The residuals should be normally distributed. This is particularly important for hypothesis testing on regression coefficients.
- **Homoscedasticity:** Residuals should exhibit constant variance. Assess this with a plot of residuals versus predicted values.
- **Independence of Residuals:** The observations should be independent of one another, especially in time series data or grouped data scenarios.

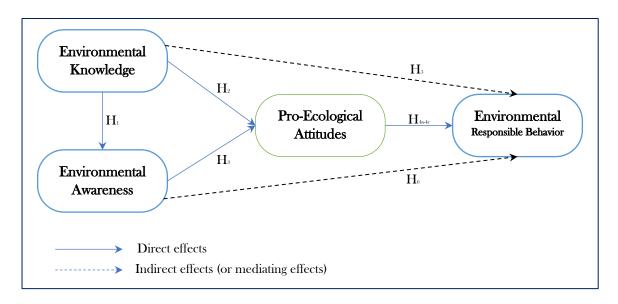
These rules serve as guidelines to help ensure that the regression analyses yield valid and interpretable results.

1. CONCEPTUAL FRAMEWORK AND HYPOTHESES DEVELOPMENT:

This conceptual framework has key roles as follow:

Table 1. Sumn	nary of Hypothes	ses Development
---------------	------------------	-----------------

Research Constructs	Independent Variables (X)	Mediating Variables (MEV)	Dependent Variables (Y)
Direct Effects			
H₁: ENK→ENA	Environmental		Environmental
	Knowledge		Awareness
H₂: ENK→PEA	Environmental Knowledge		Pro-Ecological Attitudes
H₂: ENA→PEA	Environmental Awareness		Pro-Ecological Attitudes
H₄: PEA →ENR	Pro-Ecological		Environmentally
	Attitudes		Responsible Behavior
Indirect Effects			
H _s : ENK \rightarrow PEA \rightarrow ENR	Environmental	Pro-Ecological Attitudes	Environmentally
H.: ENA \rightarrow PEA \rightarrow ENR	Knowledge Environmental	Pro-Ecological Attitudes	Responsible Behavior Environmentally
	Awareness		Responsible Behavior





2. RULE OF THUMBS: REGRESSION ANALYSIS

Criterion	Threshold Values
1. \mathbf{R}^2 (R -square)	≥ 0.10 (10%)
2. Adjusted- \mathbf{R}^2	≥ 0.10 (10%)
3. F-value	≥4
4. t-value	≥ 1.96
5. p-value	< 0.05
6. Durbin-Watson (D-W)	[1.50-2.50]
7. VIF	≤ 2.5

Table 2. The Rule of Thumbs: Regression Analysis

Sources: (i.e., Hair Jr et al., 2019; Hair Jr et al., 2021; Johnston et al., 2018)

3. EQUATIONS OF SIMPLE AND MULTIPLE REGRESSION ANALYSIS

3.1. EQUATION OF SIMPLE REGRESSION

An equation of simple regression as written below:

 $Y_1 = \boldsymbol{a}_1 X_1 + \boldsymbol{b} + \boldsymbol{\varepsilon}_0 \text{ or } Y_1 = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 X_1 + \boldsymbol{\varepsilon}_0$ (Equation 1.1)

Where:

- β_1 : is a slope of regression line (or curve) direction
- β_0 : is an intercept (or constant value)
- X_i: is an independent variable
- Y₁: is a dependent variable
- E: is the standardized error or random error term

Then,
$$Y_1 = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 X_1 + \boldsymbol{\varepsilon}_0$$

(Equation 1.2)

Example 1:

In the above conceptual framework, our research hypothesis (H₁) is to predict the relationship between "Environmental Knowledge" and "Environmental Awareness". Therefore, the following equation of simple regression will be:

$$ENA = \beta_0 + \beta_1 ENK + \varepsilon_0$$
 (Equation 1.3)

Where:

- β_1 : is a slope of regression line (or curve) direction
- β_0 : is an intercept (or constant value)
- X1: is Environmental Knowledge: ENK (Independent Variable)
- Y1: is Environmental Awareness: ENA (Dependent Variable)
- **ɛ**: is the standardized error or random error term

3.2. EQUATION OF MULTIPLE REGRESSION

An equation of multiple regression as written below:

$$Y_1 = \mathbf{b} + (\mathbf{a}_1 X_1 + \mathbf{a}_2 X_2 + \dots + \mathbf{a}_n X_n) + \mathbf{\epsilon}_0 \text{ or } Y_1 = \mathbf{\beta}_0 + (\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n) + \mathbf{\epsilon}_0$$
(Equation 2.1)

Where:

- β_1 : is a slope of regression line (or curve) direction
- β_0 : is an intercept (or constant value)



- X1...: is the independent variables
- Y₁: is a dependent variable
- E: is the standardized error or random error term

Then, $Y_1 = \boldsymbol{\beta}_0 + (\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n) + \boldsymbol{\varepsilon}_0$ (Equation 2.2)

Example 2:

In the above conceptual framework, our research hypotheses (H_2) and (H_3) are to predict the relationship between "Environmental Knowledge" and "Environmental Awareness" on "Pro-Environmental Attitude". Therefore, the following equation of simple regression will be:

$$\mathbf{PEA} = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 ENK + \boldsymbol{\beta}_2 ENA + \boldsymbol{\varepsilon}_0$$

(Equation 2.3)

Where:

- β_1 : is a slope of regression line (or curve) direction of relationship for Hypothesis 1
- β_2 : is a slope of regression line (or curve) direction of relationship for Hypothesis 2
- β_0 : is an intercept (or constant value)
- X1: is Environmental Knowledge: ENK (Independent Variable)
- X: is Environmental Awareness: ENA (Independent Variable)
- Y1: is Pro-Environmental Attitude: PEA (Dependent Variable)
- E: is the standardized error or random error term

Overall: The SPSS process must illustrate the slope and constant value to test the relationship between the independent variable (X) and the independent variable (Y).

4. STEP BY STEP...

PART I. SIMPLE REGRESSION ANALYSIS FOR A SINGLE FACTOR

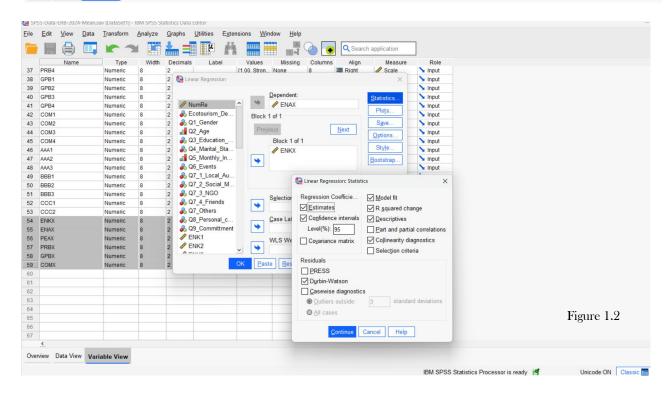
Go to Analyze >> Regression >> Linear (Figure 1.1) >> [move the mean score of "Environmental Knowledge— ENKX" to "Block box = Independent Variable(s) box" and mean score of "Environmental Awareness— ENAX" must move to "Dependent Variable" box (Figure 1.2) and >> Statistics (check: R-square change, Part and partial correlations, Collinearity diagnostics, Confident Interval (95%), and Durbin-Watson) and rest o of other function just let it defaults, then click Continue (Figure 1.2) >> OK. Then, you will see the following outputs of simple regression analysis for the research hypothesis (H1). Refer to the outputs (Figure 1.3) and shown in (Table 3) of the H1 below:

Hypothesis (H1): Environmental Knowledge has a positive influence on Environmental Awareness.

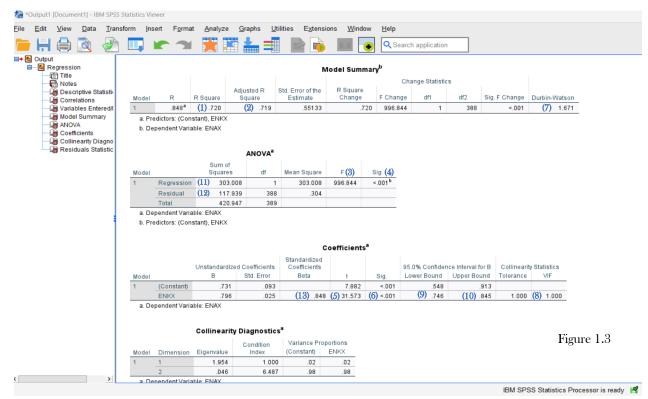


Edit View									
Equit Alexa	<u>D</u> ata	Transform	Analyze Graphs Utilities Exte	nsions	Windo	w <u>H</u> elp			
		10.3	Power Analysis	>			l 🕐 🗖	Q Search app	lication
		- 1	Meta Analysis	>					
			Reports	>	issing				Role
				,	-				> Input
					-	-			> Input
			-	>	6	8			> Input
			Ta <u>b</u> les	>	6	8			> Input
		8	Compare Means and Proportions	>	5	8			> Input
	Numeric	8	General Linear Model	>	>	8	Right	Ordinal	> Input
	Numeric	8			>	8	Right		> Input
	Numeric	8				8	Right	Ordinal	> Input
	Numeric	8	-	>		8	Right	J Ordinal	🔪 Input
ENA4	Numeric	8	Correlate	>	>	8	Right	Ordinal	S Input
ENA5	Numeric	8	Regression	>	Auto	matic Lin	ear Modeling		🔪 Input
ENA6	Numeric	8		>	_				🔪 Input
ENA7	Numeric	8							🔪 Input
ENA8	Numeric	8		,	Line	ar OLS A	ternatives	>	🔪 Input
ENA9	Numeric	8	Classify	>	Curv	e Estimat	ion		🔪 Input
PEA1	Numeric	8	Dimension Reduction	>		00500.4	0.1	Maria	🔪 Input
	Numeric	8	Mapping	>				. nayes	🔪 Input
	Numeric	8		>	👬 Part	ial Lea <u>s</u> t S	Squares		🔪 Input
	Numeric	8	-		🔡 Bina	ary Logisti	3		🔪 Input
				,		in a mini la	alatia		🔪 Input
		-	Forecasting	>	_		gistic		> Input
			Survival	>	or <u>d</u> i	nal			S Input
		•	Multiple Response	>	Erot	pit			S Input
		-			Mon	linear			> Input
		-							> Input
			Multiple Imputation	>	Wei Wei	ght Estima	ition		Input
			Complex Samples	>	2-St	age Least	Squares		Input
			Simulation			ntilo			Input
				,					Input
				,	opti Opti	mal Scalir	ig (CATREG)		Input
				>	Kerr	nel Ridge			> Input
			Direct Marketing	>			Right	A Scale	> Input
							- agin	4 000ie	put
								-	-
	ENK4 ENA1 ENA2 ENA3 ENA3 ENA4 ENA5 ENA6 ENA7 ENA8 ENA7 ENA8 ENA9 PEA1	21_Gender Numerc 22_a00 Numerc 23_Martial_ Numerc 23_Martial_ Numerc NRN Numerc NRN Numerc NRN Numerc NRA Numerc NRA	21_Gendor Numeric 8 22_A091 Numeric 8 23_Martial	Name Type Watte Annumeric Matte Annumeric Matte Annumeric Matte Reports 22_Aoge Numeric 0 Bayesian Bayesian Bayesian Statistics 32_MATIAL Numeric 0 Bayesian Statistics Tables NNR Numeric 0 Compare Means and Proportions Compare Means and Proportions NNR Numeric 0 General Linear Model NNA Numeric 0 General Linear Model NNA Numeric 0 General Linear Model NNA Numeric 0 Compare Means and Proportions NAN Numeric 0 General Linear Model NAN Numeric 0 Compare Microson NAN Numeric 0 Compare Microson NAN Numeric 0 Compare Microson NAN Numeric 0 Sala NAN Numeric 0 Sala NAN Numeric	Name Type Work Meta Analysis > 21_Gender Numeric 0 Reports > 22_ode Numeric 0 Dgs.criptive Statistics > 23_MARTIAL_ Numeric 0 Dgs.criptive Statistics > 23_MARTIAL_ Numeric 0 Compare Means and Proportions > NN3 Numeric 0 General Linear Model > NN4 Numeric 0 Gorrelate > Numeric 0 Loginear > Numeric 0 Classify > Numeric 0 Mapping > R44 Numeric 0 Mapping > R44 Numeric 0 Mapping > R44 Num	Name Type Wolf 21_Gendriv Numeric 8 22_00e Numeric 8 22_00e Numeric 8 22_00e Numeric 8 22_00e Numeric 8 23_04112 Numeric 8 2010 Numeric 8 2011 General Linear Model 9 2012 Numeric 8 2014 Numeric 8 2014 Numeric 8 2014 Regession 1 2014 Numeric 8 2014 Numeric 1 2014 Numeric 1	Name Type With Math Analysis > ising Columnation 0.1.Gender Numeric 8 Reports > > 0 22.0.40 Numeric 0 Descriptive Statistics > > 0 22.0.40 Numeric 0 Descriptive Statistics > > 0 23.0.4014 Numeric 0 Tables > 0 0 NMA Numeric 0 Compare Mans and Proportions > 0 0 NMA Numeric 0 Generalized Linear Models > 0 0 NMA Numeric 0 Generalized Linear Models > 0 0 NMA Numeric 0 Correlate > 0 0 0 NMA Numeric 0 Carelate No Numeric Nationality 0 0 0 0 NMA Numeric 0 Carelate Numeric Nationality 0 0 </td <td>Name Descriptive Statistics Image: Statistics Statistics 22_0de Numeric 0 Right 23_Matta Numeric 0 Right NN3 Numeric 0 Compare Means and Proportions 0 NN4 Numeric 0 Generalized Inter Model 0 NN4 Numeric 0 Right Numeric 0 Right 0 Numeric 0 Right Numeric 0 Right Numeric 0 Right Numeric 0 Right Numeric 0 Right</td> <td>Name Type Wittin Measure asing Columns Main Measure 22_06 Numeric 0 Reports Reports Numeric 0 Reports Numeric Numeric</td>	Name Descriptive Statistics Image: Statistics Statistics 22_0de Numeric 0 Right 23_Matta Numeric 0 Right NN3 Numeric 0 Compare Means and Proportions 0 NN4 Numeric 0 Generalized Inter Model 0 NN4 Numeric 0 Right Numeric 0 Right 0 Numeric 0 Right Numeric 0 Right Numeric 0 Right Numeric 0 Right Numeric 0 Right	Name Type Wittin Measure asing Columns Main Measure 22_06 Numeric 0 Reports Reports Numeric 0 Reports Numeric Numeric

Overview Data View Variable View







Independent Variable	Dependent Variable
	Environmentally Responsible
	Behavior(Y ₁)
	β_1 (Model-1)
Environmental Knowledge (X ₁)	(13) 0.848
\mathbf{R}^2 (R -square)	(1) 0.720
$Adjusted-R^2$	(2) 0.719
F-value (Significant of p-value)	(3) 996.844 (p<0.001) (4)
t-value	(5) 31.573
p-value	(6) <0.001
Durbin-Watson (D-W)	(7) 1.671
VIF	(8) 1.00
Confident Interval (CI)	(9) [0.746 - 0.845] (10)
D.F (Regression)	(11) 1
D.F (Residual)	(12) 388

 Table 3. The Result of Hypothesis (H1)

Note: "p<0.001, "p<0.05, p<0.10 and significant at t-value > [1.96]. d.f = degree of freedom

In Table 3, simple linear regression analysis was conducted to evaluate the extent to which "Environmental Knowledge" could predict "Environmental Awareness" of local tourism community. A significant regression was found F (1,388) = 996.844, p = < 0.001). The R² was 0.720, indicating "Environmental Knowledge–ENK" explained approximately 72.0% of the variance in "Environmental Awareness–ENA". The simple regression equation in Model-1 was:

$Y_1(ENK) = \beta_0 + \beta_1 X_1(ENA) + \varepsilon_0$	(Equation 1)
Then: $Y_1(ENK) = 0.731 + 0.848PEA + 0.093$	(Equation 1)



That is, for increase in "Environmental Knowledge", the predicted "Environmental Awareness" increased by approximately 0.746 or 74.6%. Confidence intervals indicated that we can be 95% certain that the slope to predict "Environmental Awareness" from "Environmental Knowledge "is between 0.746 and 0.845. The regression coefficients showed that for "Environmental Awareness" increased by an average of 0.848(84.8%) ($\beta_1 = 0.848$, SE = 0.093, t = 31.573, p < 0.001). This result indicated that "Environmental Knowledge" are a significant predictor of "Environmental Awareness", supporting the hypothesis that increased "Environmental Knowledge" is associated with "Environmental Awareness". Overall, the findings suggest that encouraging local tourism community to engage in more "Environmental Knowledge" could enhance their "Environmental Awareness".

PART II. SIMPLE REGRESSION ANALYSIS FOR A MULTIPLE FACTORS

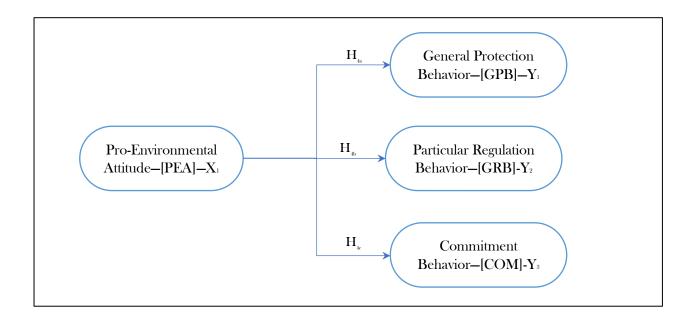
Go to **Analyze** >> **Regression** >> **Linear** (Figure 2.1) >> [move the mean score of "Pro-Environmental Attitude— PEA" to **"Block box = Independent Variable(s) box"** and mean score of Environmental Responsibility—(i.e., GPB, GRB, and COM) must move to "Dependent Variable" box (Figure 2.2) and >> **Statistics (**check: Rsquare change, Part and partial correlations, Collinearity diagnostics, Confident Interval (95%), and Durbin-Watson) and rest o of other function just let it defaults, then click **Continue** (Figure 2.2) >> **OK.** Then, you will see the following outputs of simple regression analysis for the research hypothesis (H₄₊₊). Refer to the outputs (shown in **Table 4**) of the H4a, H4b, and H4c below:

ATTENTION: In this case, a variable of "Environmentally Responsible Behavior "consists of three subdimensions (1)- General Protection Behavior–GPB, (2)-Particular Regulation Behavior–GRB, and (3)-Commitment Behavior–COM. We already computed mean score of these factors in SPSS data set. Therefore, we will have three simple regression results with Hypothesis (H_{40}), (H_{40}), and (H_{40}). Let's break down the relationship of this hypothesis below. Based on this break down relationships, we need to run the simple regression twice. First, to test the relationship of H_{40} . Second, to the test the relationship of H_{40} , and third, test the relationship of H_{40}

Hypothesis (H4a): "Pro-Environmental Attitude" has a positive influence on "General Protection Behavior."

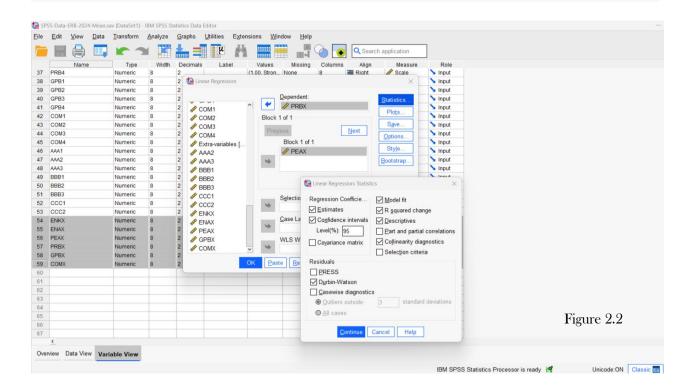
Hypothesis (H4b): "Pro-Environmental Attitude" has a positive influence on "Particular Regulation Behavior."

Hypothesis (H4c): "Pro-Environmental Attitude" has a positive influence on "Commitment Behavior."





			K 3	Power Analysis	>	🔛 📲 🍙 💽 🔍 Sear	h application						
	Name	Туре	Width	Meta Analysis	>	Label	Values	Missing	Columns	Align	Measure	Role	
1 0		Numeric	8	Reports	>		{0, Male}	None	8	Right	🙈 Nominal	> Input	
2 0	2_Age	Numeric	8	Descriptive Statistics	>		None	None	8	I Right	Scale	S Input	
3 0	3_Marital	Numeric	8	Bayesian Statistics	>		{1, Single}	None	8	Right	🚓 Nominal	S Input	
4 E	NK1	Numeric	8		,		{1, Not reall	None	8	🚟 Right	Ordinal	S Input	
5 E	NK.2	Numeric	8	Tables			{1, Not reall	None	8	The Right	I Ordinal	🥆 Input	
6 E	NK3	Numeric	8	Compare Means and Proportions	>		{1, Not reall	None	8	🚟 Right	I Ordinal	S Input	
7 E	NK4	Numeric	8	General Linear Model	>		{1, Not reall	None	8) 居 Right	J Ordinal	🥆 Input	
8 E	NA1	Numeric	8	Generalized Linear Models	>		{1, Very dis	None	8	🚎 Right	Ordinal	🦒 Input	
		Numeric	8	Mixed Models	>		{1, Very dis	None	8	🚟 Right	Ordinal	S Input	
		Numeric	8				{1, Very dis	None	8	🚟 Right	J Ordinal	🔪 Input	
		Numeric	8	Correlate	>		{1, Very dis	None	8	🚟 Right	📑 Ordinal	S Input	
		Numeric	8	Regression	>	Automatic Linear Modeling	ery dis	None	8	🚟 Right	I Ordinal	S Input	
		Numeric	8	Loglinear	>	Linear	ery dis	None	8	灌 Right	J Ordinal	S Input	
		Numeric	8	Neural Networks	>		ery dis	None	8	🚟 Right	d Ordinal	S Input	
		Numeric	8		Ś	Linear OLS Alternatives	> ery dis	None	8	🗏 Right	Ordinal	S Input	
		Numeric	8	Classify	,	Curve Estimation	ery dis	None	8	Right Right	d Ordinal	🥆 Input	
		Numeric	8	Dimension Reduction	>	PROCESS v4.2 by Andrew F. Haves	ery dis	None	8	Right	Ordinal	S Input	
		Numeric	8	Mapping	>		ery dis	None	8	🗃 Right	J Ordinal	> Input	
		Numeric	8	Scale	>	🚻 Partial Lea <u>s</u> t Squares	ery dis	None	8	The Right	I Ordinal	S Input	
		Numeric	8	Nonparametric Tests		Binary Logistic	ery dis	None	8	🖷 Right	Ordinal	S Input	
		Numeric	8	-		Multinomial Logistic), Very	None	8	遭 Right	J Ordinal	S Input	
		Numeric	8	Forecasting	>		Imost	None	8	灌 Right	Grdinal	S Input	
		Numeric	8	Survival	>	Kanal	Imost	None	8	Right	Ordinal	> Input	
		Numeric	8	Multiple Response	>	Erobit	Imost	None	8	I Right	Ordinal	> Input	
		Numeric	8	Missing Value Analysis		Nonlinear	Imost	None	8	a Right	Ordinal	S Input	
		Numeric	8				Imost	None	8) 居 Right	Ordinal	> Input	
		Numeric	8	Multiple Imputation	>	Weight Estimation	Imost	None	8	Right Right	Ordinal	> Input	
		Numeric	8	Complex Samples	>	2-Stage Least Squares	Imost	None	8	Right	Ordinal	Input	
		Numeric	8	P Simulation		Quantile	Imost	None	8	Right	Ordinal Scale		
		Numeric	8		>		0	None	10	Right Right		> Input	
		Numeric	8	Quality Control		Deptimal Scaling (CATREG)	e	None	10	Right Right	Scale Scale	Input	
		Numeric Numeric	8	Spatial and Temporal Modeling	>	Kernel Ridge	e	None	10	遷 Right 澤 Right	Scale	> Input	
		Numeric	8	Direct Marketing	>		e None	None	10	Right III Right	Scale Scale	Input	
34 0	IND .	Numeric	0				None	None	10	rught	er ocale	* mput	
36													
37									-		-		
38													
39								-				Fim	ıre 2.1
40												rigt	nc 2.1
41													
42													





4.1. SIMPLE REGRESSION: OUTPUTS

4.1.1. Simple Regression for Hypothesis (H4a)

					Model	Summa	ury ^b								
								Ch	nange Sta	atistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate		iquare nange	F	Change	df1		df2	Sig. F C	hange (Durbin	-Watson
1	.570ª	(1) .324	(2) .323	.75159		.324	t i	186.280		1	388		<.001	(7)	1.624
		nstant), PEAX able: PRBX					_							_	
					ANG	DVAª									
Mode	el			um of Juares		df		Mea	n Sqi	Jare		F		Sig	
1	Re	gressio	n (11)	105.227			1		105.	227	(3)18	6.28	0 (4)	<.0	01 ^b
	Re	esidual	(12)	219.175		38	8			565					
	То	tal		324.401		38	9								
			iable: PRE nstant), Pl		Coef	ficients	a								
		Unstandard	lized Coefficien	Standardiz Coefficien					95.0% (Confider	nce Interva	al for B	Colline	arity St	atistics
Model		В	Std. Error	Beta		t	5	Sig.	Lower	Bound	Upper B	Bound	Tolerand	e	VIF
	(Constant)	1.75	i5 .13	-		13.222		<.001	(0)	1.494	(1.5)	2.016			
1 _		.53	2 .03	(13)	570 (5)	13 6 10	(6)	<.001	(9)	.456	(10)	.609	1.00	0 15	1.000



4.1.2. Simple Regression: Outputs-For Hypothesis (H4b)

					lodel Summar							
			Adjusted R	Std. Error of the	R Square	Cr	nange Statistics					
Nodel	R	R Square	Square	Estimate	Change	F Change	df1	df2	Sig. F Cl	hange	Durbin-W	/atson
1	.691 ^a	.478	.477	.70628	.478	355.104	1	388		<.001		1.507
a. Prec	dictors: (Cor	nstant), PEAX		-								
b. Dep	endent Vari	able: GPBX										
					ANOVA ^a							
			Su	m of								
Mod	el		Sq	uares	df	Mea	n Square		F		Sig.	
1	Re	egressio	n	177.138	1		177.138	35	5.104	4	<.00	1 ^b
	Re	esidual		193.548	388	}	.499					
	То	tal		370.686	389)						
a.	Depen	dent Var	iable: GPB	X				_				
			nstant), PE									
w.	ricale	.013. (00	notanty, r c									_
					Coefficients ^a	1						
		Unstandar	dized Coefficients	Standardized Coefficients			95.0% Confide	nce Interv	al for B	Colline	arity Stati	istics
Model		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper		Toleran		/IF
1	(Constant)	1.20	51 .12	5	10.111	<.001	1.016		1.506			
1	PEAX	.69	90 .03	7	1 18.844	<.001	.618		.763	1.0	00	1.000



4.1.3. Simple Regression: Outputs–For Hypothesis (H4c)

				M	lodel Summar	У ^Б					
						CI	nange Statistics				
odel	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	Durbin-W	atson
	.638 ^a	.407	.406	.76664	.407	266.787	1	388	<.001		1.578
		nstant), PEAX able: COMX									
					ANOVA ^a						
			Su	im of							
Mode	el		Sq	uares	df	Mea	in Square		F	Sig.	
1	Re	egressio	n	156.800	1		156.800	26	6.787	<.00	1 ^b
	Re	esidual		228.040	388	}	.588				
	То	tal		384.840	389)					
a.	Depen	dent Var	iable: CON	ΛX							
b.	Predict	tors: (Co	nstant), Pl	AX							
					Coefficients ^a						
			dized Coefficient				95.0% Confider			nearity Statis	
odel		В	Std. Error	Beta	t	Sig.	Lower Bound	Upper B		ance V	IF
	(Constant) PEAX			-	9.873	<.001	1.070		1.603 .728 1	000	0.000
		b	50 .04	0.63	8 16.334	<.001	.571		1/8 1	.000	1.000

4.2. TABLE FORMAT OF SIMPLE REGRESSION

Table 4. The Result of Hypothesis H4a, H4b, and H4c

Independent variable	Dependent variable: Environm	nentally Responsible Be	havior
	GPB (Y _i)	GRB (Y ₂)	COM (Y ₃)
	β_1 (Model-1)	β ₂ (Model-2)	β₃ (Model-3)
Pro-Environmental Attitudes (X ₁)	(13) 0.570 ^{***}	0.691***	0.638
R ² (R-square)	(1) 0.324	0.478	0.407
Adjusted- R ²	(2) 0.323	0.477	0.406
F-value (Significant of p-value)	(3) 186.280 (p<0.001) (4)	355.104 (p<0.001)	266.787 (p<0.001)
t-value	(5) 13.648	18.844	16.334
p-value	(6) <0.001	<0.001	<0.001
Durbin-Watson (D-W)	(7) 1.624	1.507	1.578
VIF	(8) 1.00	1.00	1.00
Confident Interval (CI)	(9) [0.456 - 0.609] (10)	[0.618 - 0.763]	[0.571-0.728]
D.F (Regression)	(11) 1	1	1
D.F (Residual)	(12) 388	388	389

Note: p<0.001, *p*<0.05, *p*<0.10 and significant at t-value > [1.96]

4.3. RESULTS AND INTERPRETATION

4.3.1 For Hypothesis (H4a)

In Table 4, simple linear regression analysis was conducted to evaluate the extent to which "Pro-Environmental Attitudes" could predict "Green Protected Behavior" of local tourism community. A significant regression was found F (1,388) = 186.280, $p \le 0.001$). The R² was 0.324, indicating "Pro-Environmental Attitudes" explained



approximately 32.4% of the variance in "Green Protected Behavior". The simple regression equation in Model-1 was:

$Y_1(GPB) = \beta_0 + \beta_1 X_1(PEA) + \varepsilon_0$	(Equation 2)
<i>Then:</i> $Y_1(GPB) = 1.755 + 0.570PEA + 0.133$	(Equation 2)

That is, for increase in "*Pro-Environmental Attitudes*", the predicted "*Green Protected Behavior*" increased by approximately 0.456 or 45.6%. Confidence intervals indicated that we can be 95% certain that the slope to predict "*Green Protected Behavior*" from "*Pro-Environmental Attitudes*" is between 0.456 and 0.609. The regression coefficients showed that for "Green Protected Behavior" increased by an average of 0.570 (57.0%) ($\beta_1 = 0.570$, SE = 0.133, t = 13.648, p < 0.001). This result indicated that "Pro-Environmental Attitudes" are a significant predictor of "Green Responsible Behavior", supporting the hypothesis that increased "Pro-Environmental Attitudes" is associated with "Green Responsible Behavior". Overall, the findings suggest that encouraging local tourism community to engage in more "Pro-Environmental Attitudes" could enhance their "Green Protected Behavior".

4.3.2 For Hypothesis (H4b)

In Table 4, simple linear regression analysis was conducted to evaluate the extent to which "Pro-Environmental Attitudes" could predict "Green Responsible Behavior" of local tourism community. A significant regression was found F (1,388) = 355.104, p = < 0.001). The R² was 0.478, indicating "Pro-Environmental Attitudes" explained approximately 47.8% of the variance in "Green Responsible Behavior". The simple regression equation in Model-2 was:

$Y_2(GRB) = \beta_0 + \beta_2 X_1(PEA) + \varepsilon_0$	(Equation 3)
Then: $Y_2(GRB) = 1.261 + 0.691PEA + 0.125$	(Equation 3)

That is, for increase in "*Pro-Environmental Attitudes*", the predicted "Green Responsible Behavior" increased by approximately 0.618 or 61.8%. Confidence intervals indicated that we can be 95% certain that the slope to predict "*Green Protected Behavior*" from "*Pro-Environmental Attitudes*" is between 0.618 and 0.763. The regression coefficients showed that for "Green Responsible Behavior" increased by an average of 0.691 (69.1%) ($\beta_1 = 0.691$, SE = 0.125, t = 18.844, p < 0.001). This result indicated that "Pro-Environmental Attitudes" are a significant predictor of "Green Responsible Behavior", supporting the hypothesis that increased "Pro-Environmental Attitudes" is associated with "Green Responsible Behavior". Overall, the findings suggest that encouraging local tourism community to engage in more "Pro-Environmental Attitudes" could enhance their "Green Responsible Behavior". Therefore, the results show a very weak relationship between "pro-environmental attitudes" and "green responsible behavior," with an 8.1% correlation. This means that among 390 participants, there are only 186 respondents (i.e., [390x0.478] = 186) who have the concepts of "green responsible behavior" in their local tourism community.

4.3.3 For Hypothesis (H4c)

In Table 4, simple linear regression analysis was conducted to evaluate the extent to which "Pro-Environmental Attitudes" could predict "Commitment Behavior" of local tourism community. A significant regression was found F (1,389) = 266.787, p = < 0.001). The R² was 0.407, indicating "Pro-Environmental Attitudes" explained approximately 40.7% of the variance in "Commitment Behavior". The simple regression equation in Model-3 was:

$Y_3(COM) = \beta_0 + \beta_3 X_1(PEA) + \varepsilon_0$	(Equation 4)
<i>Then:</i> $Y_3(COM) = 1.337 + 0.638PEA + 0.135$	(Equation 4)

That is, for increase in "*Pro-Environmental Attitudes*", the predicted "Green Responsible Behavior" increased by approximately 0.571 or 57.1%. Confidence intervals indicated that we can be 95% certain that the slope to predict "*Green Protected Behavior*" from "*Pro-Environmental Attitudes*" is between 0.571 and 0.728. The regression coefficients showed that for "Green Responsible Behavior" increased by an average of 0.638 (63.8%)



 $(\beta_1 = 0.638, SE = 0.135, t = 16.334, p < 0.001)$. This result indicated that "Pro-Environmental Attitudes" are a significant predictor of "Commitment Behavior", supporting the hypothesis that increased "Pro-Environmental Attitudes" is associated with "Commitment Behavior". Overall, the findings suggest that encouraging local tourism community to engage in more "Pro-Environmental Attitudes" could enhance their "Commitment Behavior". Therefore, the results show a moderate relationship between "pro-environmental attitudes" and "Commitment Behavior," with an 40.7% correlation. This means that among 390 participants, there are only 159 respondents (i.e., [390x0.407] = 159) who have the concepts of "Commitment Behavior" in their local tourism community.

In summary, all sub-dimensions of "Environmentally Responsible Behavior" are well-supported by the prediction of a key independent research variable of "Pro-Ecological Attitudes". Therefore, H4a, H4b, and H4c are accepted.

PART III. MULTIPLE REGRESSION

In conceptual model, we have one multiple linear regression for Hypothesis 2 and Hypothesis 3. The following step-by-step is provided below:

Go to **Analyze** >> **Regression** >> **Linear** (Figure 3.1) >> [move the mean scores of Environmental Knowledge– (ENK) and Environmental Awareness–(ENA) to **"Block box = Independent Variable(s) box"** (i.e., ENK for Block 1 and ENA for Block 2 by clicking on next) and mean score of "Pro-Environmental Attitude–PEA" must move to **"Dependent Variable"** box (Figure 3.2) and **>> Statistics (**check: R-square change, Part and partial correlations, Collinearity diagnostics, Confident Interval (95%), and Durbin-Watson) and rest o of other function just let it defaults, then click **Continue** (Figure 3.2) **>> OK.** Then, you will see the following outputs of multiple linear regression for the research hypotheses (H₂) and (H₃). Refer to the outputs (shown in **Table 5**) of the H₂ and H₃ below:

This study proposes the research hypotheses, as followed:

Hypothesis (H₂): "Environmental Knowledge" has a positive influence on "Pro-Environmental Attitude."

Hypothesis (H₃): "Environmental Awareness" has a positive influence on "Pro-Environmental Attitude."

e	Edit ⊻iev	_	Transform	Analyze Graphs Utilities Extensi Power Analysis	ions	Window Help		_					
_				Meta Analysis	>	i 📰 🖬 🌜 💽 🔽 Search							
	Name Q1 Gender	Type Numeric	Width	Reports	>	Label	Values {0, Male}	Missing	Columns 8	Align	Measure	Role	
1 2	Q2_Age	Numeric	8	-	>		{0, Male} None	None	8	Right	Nominal Scale	Input	
2	Q2_Age Q3_Marital	Numeric	8	Descriptive Statistics			{1, Single}	None	8	Right Right	Scale	> Input	
3	ENK1	Numeric	8	Bayesian Statistics	>		(1, Single) {1, Not rea		8	層 Right	Ordinal	> Input	
5	ENK2	Numeric	8	Tables	>		{1, Not rea		8	温 Right	Ordinal	> Input	
6	ENK3	Numeric	8	Compare Means and Proportions	>		{1. Not rea		8	I Right	Ordinal	> Input	
7	ENK4	Numeric	8	General Linear Model	>		{1, Not rea		8	温 Right	Ordinal	> Input	
8	ENA1	Numeric	8	-			{1. Very di		8	I Right	Ordinal	> Input	
9	ENA2	Numeric	8	Generalized Linear Models	>		{1, Very dis		8	Right	Ordinal	> Input	
10	ENA3	Numeric	8	Mixed Models	>		{1, Very dis		8	温 Right	Ordinal	> Input	
11	ENA4	Numeric	8	Correlate	>		{1, Very dis		8	I Right	Ordinal	> Input	
12	ENA5	Numeric	8	Regression	>	Automatic Linear Modeling	ery dis		8	Right	Ordinal	> Input	
13	ENA6	Numeric	8	Loglinear	>		ery dis		8	Right	Ordinal	> Input	
14	ENA7	Numeric	8	=0		🔛 Linear	ery dis		8	Right	Ordinal	> Input	
15	ENA8	Numeric	8	Neural Networks	>	Linear OLS Alternatives	> ery dis		8	Right	Ordinal	> Input	
16	ENA9	Numeric	8	Classify	>	Curve Estimation	ery dis		8	Right	Ordinal	> Input	
17	PEA1	Numeric	8	Dimension Reduction	>		ery dis		8	Right	Ordinal	> Input	
18	PEA2	Numeric	8	Mapping	>	PROCESS v4.2 by Andrew F. Hayes	ery dis	None	8	遍 Right	J Ordinal	S Input	
19	PEA3	Numeric	8			🔠 Partial Least Squares	ery dis	None	8	I Right	Ordinal	S Input	
20	PEA4	Numeric	8	Sc <u>a</u> le	>	Binary Logistic	ery dis	None	8	I Right	Ordinal	S Input	
21	PEA5	Numeric	8	Nonparametric Tests	>), Very	None	8	酒 Right	J Ordinal	🥆 Input	
22	ENR1	Numeric	8	Forecasting	>	Multinomial Logistic	Imost	None	8) 居 Right	Ordinal	🔪 Input	
23	ENR2	Numeric	8	Survival	>	Grdinal	Imost	None	8)温 Right	🛃 Ordinal	🔪 Input	
24	ENR3	Numeric	8	Multiple Response	>	Probit	Imost	None	8	I Right	Ordinal	🍾 Input	
25	ENR4	Numeric	8				Imost	None	8)温 Right	Ordinal	🔪 Input	
26	ENR5	Numeric	8	🔛 Missing Value Analysis		Nonlinear	Imost	None	8	🚟 Right	Ordinal	🔪 Input	
27	ENR6	Numeric	8	Multiple Imputation	>	Weight Estimation	Imost	None	8	3 Right	Ordinal	🔪 Input	
28	ENR7	Numeric	8	Complex Samples	>	2-Stage Least Squares	Imost	None	8) 居 Right	Ordinal	🔪 Input	
29	ENR8	Numeric	8	Simulation			Imost	None	8	🗃 Right	J Ordinal	🥆 Input	
30	ENK	Numeric	8			Guantile	0	None	10	灌 Right	I Scale	🥆 Input	
31	ENA	Numeric	8	Quality Control	>	Optimal Scaling (CATREG)	e	None	10) 居 Right	I Scale	S Input	
32	PEA	Numeric	8	Spatial and Temporal Modeling	>	Kernel Ridge	e	None	10	滬 Right	I Scale	🥆 Input	
33	GPB	Numeric	8	Direct Marketing	>	Terret Kidde	e	None	10	Right Right	Scale	🔪 Input	
34	GRB	Numeric	8				None	None	10)温 Right	I Scale	🥆 Input	
35													
36							_		_				
37			_				_		_				
38			_				_		_				
39			_				-		_				E' 0
40			_						_				Figure 3.
41			_						_				3
42													
	<	_											



le	<u>E</u> dit <u>V</u> iew <u>D</u> ata	Iransform Analyze	<u>Graphs</u> <u>Utilities</u> Extensions <u>Window</u>	Help	
		K J K		Search application	
	Name	Type Width		sing Columns Align Measure	Role
	PRB4	Numeric 8	2 (1.00. Stron None	8 🗃 Right 🖋 Scale	N Input
38	GPB1	Numeric 8 Numeric 8	2 🔄 Linear Regression	— ×	N Input
39	GPB2		2 Depend	ent .	> Input
10	GPB3	Numeric 8			Input
11	GPB4 COM1	Numeric 8 Numeric 8	- COM1	Plots	S Input
2			2 SCOM2 Block 2 of 2	Save	S Input
13 14	COM2 COM3	Numeric 8 Numeric 8	2 COM3 Previous	Next	input
15	COM4	Numeric 8	- 01014	2 of 2	N Input
16	AAA1	Numeric 8	CAtia-valiables [N Input
47	AAA2	Numeric 8		Bootstrap	s input
18	AAA3	Numeric 8		Doorstrap	N Input
19	BBB1	Numeric 8	2 # BBB1 2 # BBB2	Linear Regression: Statistics	×
50	8882	Numeric 8	2 & BBB3	-	
51	BBB3	Numeric 8	2	Regression Coefficie Model fit	
52	CCC1	Numeric 8	2 Selecti	Estimates R squared cha	ance
	CCC2	Numeric 8	2 2 ENKX	Confidence intervals Descriptives	
54	ENKX	Numeric 8	2 PENAX Case L		a su
55	ENAX	Numeric 8	2 / PRBX		
56	PEAX	Numeric 8	2 SPBX WLS V	Covariance matrix Collinearity dia	agnostics
57	PRBX	Numeric 8	2 / COMX	Selection crite	eria
58	GPBX	Numeric 8	2	Residuals	
59	COMX	Numeric 8	2 OK Paste R	PRESS	
50				Durbin-Watson	
51				Casewise diagnostics	
52					and deviations
53					no demanorità
54				All cases	
55					
36				Continue Cancel Help	
37					Figure 3.2
21	<				i iguit 0.2

Independent variables	Dependent variable: Pro-Environmental Attitudes-(Y1)							
	β ₁ (Model-1)	β ₂ (Model-2)						
Environmental Knowledge (X ₁)	(13) 0.295							
Environmental Awareness (X3)		(M) 0.441						
R ² (R-square)	(1) 0.448	(A) 0.502						
Adjusted- \mathbf{R}^2	(2) 0.446	(B) 0.499						
F-value (Significant of p-value)	(3) 314.284 (p<0.001) (4)	(C) 194.971 (p<0.001) (D)						
t-value	(5) 4.355	(E) 6.500						
p-value	(6) <0.001	(F) <0.001						
Durbin-Watson (D-W)	(7) 1.510	(G) 1.510						
VIF	(8) 3.569	(H) 3.569						
Confident Interval (CI)	(9) [0.143 - 0.377] (10)	(I) [0.289 – 0.539] (J)						
d.f (Regression)	(11) 1	(K) 2						
d.f (Residual)	(12) 388	(L) 387						

Table 5. The Result of Hypothesis (H₂) and (H₃)

Note: "p<0.001, "p<0.05, p<0.10 and significant at t-value > [1.96], d.f= degree of freedom

The interpretation of multiple regression is similar to simple regression because simple regression is a special case of multiple regression (Doane & Seward, 2016). In Table 5, multiple linear regression analysis was conducted to evaluate the extent to which "Environmental Knowledge" and "Environmental Awareness" could predict "Pro-Environmental Attitudes" of local tourism community. In Model-1, a significant regression was found F (1,388) = 314.284, p = < 0.001). The R² was 0.448, indicating "Environment Knowledge" explained approximately 44.8% of the variance in "Pro-Environmental Attitudes". In Model-2, a significant regression was found F (2,387) = 194.971, p = < 0.001). The R² was 0.502, indicating "Environment Awareness" explained approximately 50.2% of the variance in "Pro-Environmental Attitudes". Thus, the multiple regression equation in was:

$Y_1(EPA) = \beta_0 + \beta_1 X_1(ENK) + \beta_2 X_2(ENA) + \varepsilon_0$	(Equation 5)
Then: $Y_1(EPA) = 0.891 + 0.295ENK + 0.441ENA + 0.125$	(Equation 5)



That is, for increase in "*Environmental Knowledge*", the predicted "*Pro-Environmental Attitudes*" increased by approximately 0.143 or 14.3%. Confidence intervals indicated that we can be 95% certain that the slope to predict "*Environmental Knowledge*" from "*Pro-Environmental Attitudes*" is between 0.143 and 0.377. The regression coefficients showed that for "*Pro-Environmental Attitudes*" increased by an average of 0.446 (44.6%) ($\beta_1 = 0.295$, SE = 0.060, t = 4.355, p <0.001). This result indicated that "Pro-Environmental Attitudes" are a significant predictor of "*Environmental Knowledge*", supporting the hypothesis that increased "*Environmental Knowledge*" is associated with "Pro-Environmental Attitudes". Overall, the findings suggest that encouraging local tourism community to engage in more "*Environmental Knowledge*" could enhance their "Pro-Environmental Attitudes".

Similarly, for increase in "*Environmental Awareness*", the predicted "*Pro-Environmental Attitudes*" increased by approximately 0.289 or 28.9%. Confidence intervals indicated that we can be 95% certain that the slope to predict "*Environmental Awareness*" from "*Pro-Environmental Attitudes*" is between 0.289 and 0.539. The regression coefficients showed that for "*Pro-Environmental Attitudes*" increased by an average of 0.502 (50.2%) ($\beta_1 = 0.441$, SE = 0.064, t = 6.500, p < 0.001). This result indicated that "Pro-Environmental Attitudes" are a strongly significant predictor of "*Environmental Awareness*", supporting the hypothesis that increased "*Environmental Awareness*" is associated with "Pro-Environmental Attitudes". Overall, the findings suggest that encouraging local tourism community to engage in more "*Environmental Awareness*" could enhance their "Pro-Environmental Attitudes". In summary, the goal of multiple regression analysis is to identify which independent variables have the greatest influence on predicting the dependent variable. Thus, the results indicate that "environmental awareness" is the most significant factor influencing "pro-environmental" attitudes, accounting for 50.2%. This means that if researchers wish to change the behavior or attitudes of the local tourism community, they would gain more knowledge and awareness about the environments.



						Мо	del Summa	ry ^c								
								CI	hange Sta	itistics						
Model	R	R Square	Adjuste Squa		Std. Error of th Estimate	е	R Square Change	F Change	df1		df2	Sig. F C	hange	Durbir	n-Wats	
1	.669 ^a	(1) .448	(2)	.446	.7273	6	.448	314.284		1	388	<.001				
2 .708 ^b (A) .502 (B) .499		.499	.6915	3	.054	42.247		1	387		<.001	(7)+(G) 1.51				
b. Predic	ctors: (Cor	istant), ENKX istant), ENKX, able: PEAX	ENAX			_								_	_	
						А	NOVA ^a									
					im of		-16					-			_	
Mode	1			Sq	uares	_	df	Mea	in Squ	lare		F		Sig	-	
1	Re	gressio	n		166.274		(11)	1	166.274			(3) 314.284			(4) <.001 ¹	
Residual				205.274		(12) 38	8	.529								
	То	Total			371.548 389		9									
2	Re	gressio	n		186.477	,	(K)	2	93.239		(C)194.971		1 (D)	(D) <.001		
	Re	sidual	185.070 (L)		(L) 38	7	.478									
	То	Total			371.548 38		9									
a. C	Depen	dent Var	iable:	: PEA	х											
b. F	redict	ors: (Co	nstar	nt), EN	VKX											
c. P	redict	ors: (Co	nstan	t). EN	NKX, ENA	x										
							oefficients	a								
						-	benncients									
	Unstandardized Coefficients Coefficients					95.0% Confiden		nce Interv	Colline	linearity Statisti						
Model		B Std. Erro			Beta		t	Sig.	Lower Bound				Toleran	ce	VIF	
((Constant)	1.19	94	.12	2		9.759	<.001		.953		1.434				
	NKX	.58		.03		.669	17.728	<.001		.524		.655	1.0	00	1.0	
((Constant)	.89		.12			7.115	<.001	(0)	.645	(10	1.137			2)	
	NKX	20	50	.06	o (13)	295	(5) 4.355	(6) <.001	(9)	.143	(10	.377	.2	80 (З) з.е	
E		. 21														

REFERENCES

- Anderson, D. R., Sweeney, D. J., Williams, T. A., Freeman, J., & Shoesmith, E. (2014). Statistics for business and economics (12th ed.). Cengage Learning, Inc.
- Doane, D. P., & Seward, L. W. (2016). Applied statistics in business and economics. Mcgraw-Hill.
- Groebner, D. F., Shannon, P. W., & Fry, P. C. (2018). *Business statistics: A decision-making approach*. Pearson.
- Hair Jr, J., Black, W., Babin, B., & Anderson, R. (2019). *Multivariate data analysis: A global perspective*. Prentice Hall and Pearson.
- Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). A primer on partial least squares structural equation modeling (PLS-SEM). Sage publications.
- Johnston, R., Jones, K., & Manley, D. (2018). Confounding and collinearity in regression analysis: a cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. *Quality & Quantity*, 52(4), 1957-1976. https://doi.org/10.1007/s11135-017-0584-6

GLOSSARY

17 | Veasna SO U., Sambath PHOU., & Phichhang Ou. (2024). Simple and multiple regression analysis. IBM-RUPP, Cambodia, 1-16.



- **Regression Equation:** An equation that expresses the linear relationship between independent variable and dependent variable.
- **Regression Slope Coefficient:** The average change in the dependent variable for a unit change in the independent variable. The slope coefficient may be positive, negative, or zero, depending on the relationship between the two variables.
- **Coefficient of Determination:** The portion of the total variation in the dependent variable that is explained by its relationship with the independent variable. The coefficient of determination is also called R-squared and is denoted as R².
- **R-square (R^{*}):** The squared correlation coefficient (R^{*}) is a very important statistic to explain the strength of the relationship we have between two variables.
- Adjusted R-Squared: A measure of the percentage of explained variation in the dependent variable in a multiple regression model that takes into account the relationship between the sample size and the number of independent variables in the regression model.
- **Simple Linear Regression:** The method of regression analysis in which a single independent variable is used to explain the variation in the dependent variable.
- **Multiple Regression:** extends simple regression to include several independent variables (called predictors).
- **Correlation Coefficient:** A visual display is a good first step in analysis, but we would also like to quantify the strength of the association between two variables. It is a quantitative measure of the strength of the linear relationship between two variables. The correlation ranges from -1.0 to +1.0. A correlation of {1.0 indicates a perfect linear relationship, whereas a correlation of 0 indicates no linear relationship.
- Variance Inflation Factor—(VIF): A measure of how much the variance of an estimated regression coefficient increases if the independent variables are correlated. A VIF equal to 1.0 for a given independent variable indicates that this independent variable is not correlated with the remaining independent variables in the model. The greater the multicollinearity, the larger the VIF.