

Analysis Report

This report starts with the execution of two correlation analyses, followed by a multiple regression model.

Correlation Analysis

Correlation coefficients are indicators of associations between variables (Pallant, 2010). Values between 0.10 and 0.29 indicate a small degree of association, while values between 0.30 and 0.49 are considered medium and values higher than 0.50 represent a high degree of association (Cohen, 1988). This test will be used to measure the association between different types of engagement and several other variables.

There are a number of different statistics available, depending on the level of measurement and the nature of your data. Pearson's coefficient 'r' is designed for interval level (continuous) variables, whereas Spearman's 'rho' is designed for use with ordinal level or ranked data and is particularly useful when the data does not meet the criteria for Pearson correlation (Pallant, 2010). Pearson's correlation is particularly used for normally distributed data. In this study, the data is non-normal, but according to the "central limit theorem", if we have a large enough sample (what is the case), such violations become negligible; we can assume, that our data (population) normally distributed and we can use a normal tests. It opens new possibilities. Therefore, Pearson's coefficients were calculated. The tests were performed using SPSS version 27.

The association of Transportation and Sympathy and Empathy

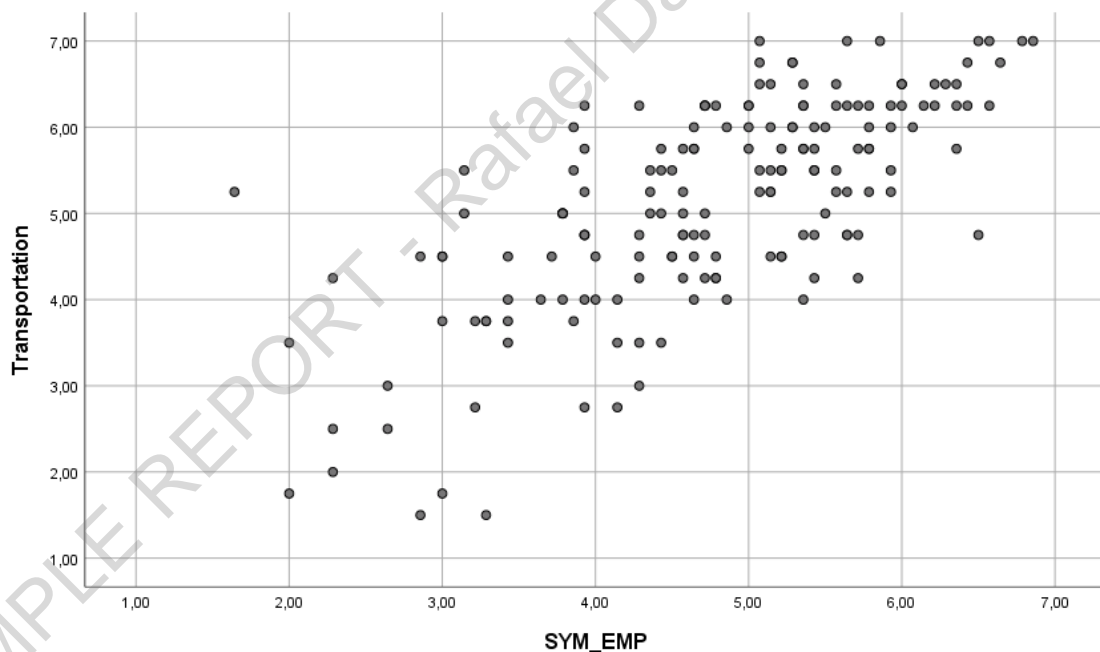
'Sympathy and Empathy' is significantly correlated with 'Transportation' ($r = .706$, $p < .001$). The sign of the coefficients is positive. That is, when the score of Sympathy and Empathy increases by one-unit, the score of Transportation increases by 0.706 (figure below).

Correlations

		Transportation	SYM_EMP
Transportation	Pearson Correlation	1	0,706**
	Sig. (2-tailed)		0,000
	N	170	170
SYM_EMP	Pearson Correlation	0,706**	1
	Sig. (2-tailed)	0,000	
	N	170	170

** . Correlation is significant at the 0.01 level (2-tailed).

In order to better visualize the association between these two variables, a scatterplot was created. The figure below shows the relationship between both scores. Each point is a subject of the sample. An increase pattern can be noticed in the graph. As we go along the X-axis, the points tend to be on the higher side of the graph, which is also what happened when we move up the Y-axis.



Another interesting pattern is that the points tend to be more 'concentrated' on the right-side of the graph. This suggests that the relationship can be 'stronger' when values of transportation and Sympathy and Empathy are higher.

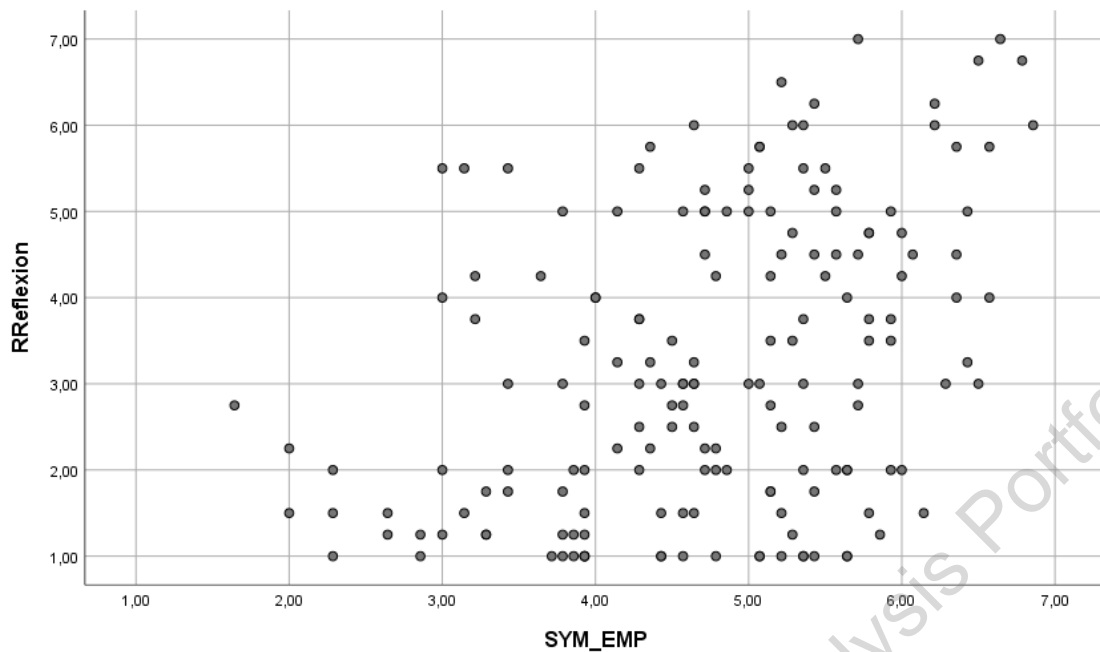
The association of Retrospective Reflection and Sympathy and Empathy

The same test was performed to examine the association between Retrospective Reflection and Sympathy and Empathy. 'Sympathy and Empathy' is significantly correlated with 'Retrospective Reflection' ($r = .396$, $p < .001$). The magnitude of the association is relatively lower than for Transportation, but still statistically significant. When the score of Sympathy and Empathy increases by one-unit, the score of Retrospective Reflection is expected to increase by 0.396 (figure below).

<i>Correlations</i>		RReflexion	SYM_EMP
RReflexion	Pearson Correlation	1	0,396**
	Sig. (2-tailed)		0,000
	N	170	170
SYM_EMP	Pearson Correlation	0,396**	1
	Sig. (2-tailed)	0,000	
	N	170	170

** . Correlation is significant at the 0.01 level (2-tailed).

Similar to what was done on the previous section, a scatterplot was generated to illustrate the relationship between Retrospective Reflection and Sympathy and Empathy. It is also possible to visualize an 'increase' pattern on the graph, but not as clear as on the analysis for Transportation (previous section). This is reflected on the correlation coefficient, which was lower (as shown above).



The next sections present the tests of the assumptions of linear regression, which are: absence of significant outliers, normality, linearity and homoscedasticity of residuals (errors) and absence of multicollinearity.

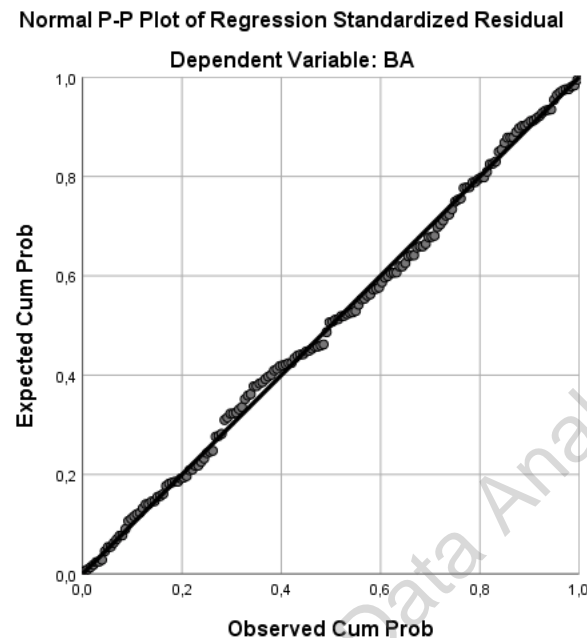
Outlier analysis

An assumption of many statistical methods is that there are no significant multivariate outliers in the data, which might distort the models. A pragmatic approach to identify outliers is suggested by Hair et al. (2014): Mahalanobis distances. These are calculated for the variables to be entered on the multiple regression analysis and their results are divided by the number of variables. When sample sizes are large (100+), coefficients above 3.5 or 4.0 can be considered outliers (Hair et al., 2014). In this study, the highest value was 3.87. Since the sample size for the Mahalanobis calculation was 170, it can be concluded that there are no extreme values in the data that would need to be deleted.

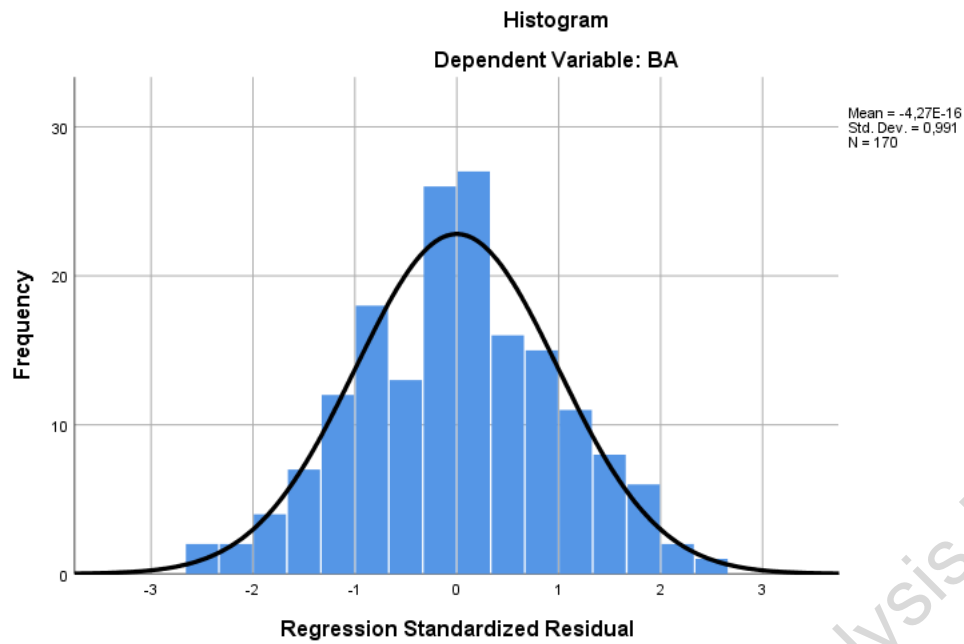
Normality, linearity and homoscedasticity of residuals (errors).

Lastly, violations of the assumptions of normality, linearity and homoscedasticity of residuals (errors) were examined for the regression model. The next figure shows a P-P plot, which is used to assess the normality of residuals. The observations should follow a

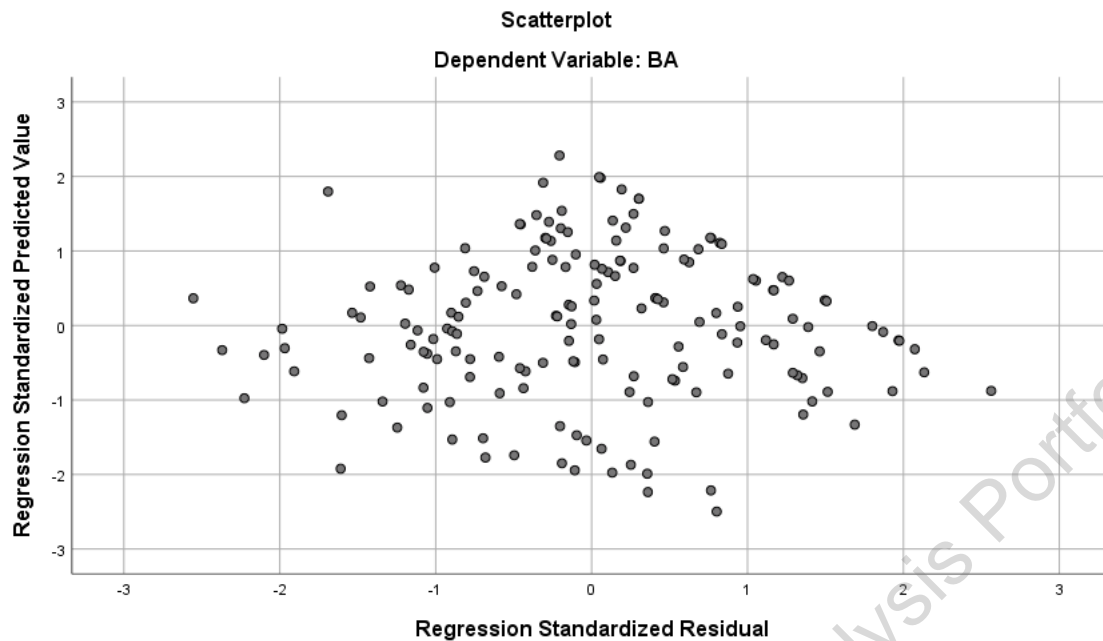
diagonal pattern to suggest normality of residuals (Tabachnick and Fidell, 2014). The tested model had Brand Attitude as outcome variable and Familiarity, Involvement and Transportation as predictor variables.



The graph nicely follows a diagonal pattern, suggesting that residuals are normally distributed. The next graph is a histogram of residuals. A ‘bell-shaped’ pattern would indicate normality, which is the case (figure below). This provides additional evidence that residuals are normally distributed.



The next figure shows a scatterplot of standardized residuals and standardized predicted values of the dependent variable. If points are well distributed along the X and Y axes, this would suggest homoscedasticity and linearity. Nonlinearity is indicated when most of the residuals are above the zero line on the plot at some predicted values and below the zero line at other predicted values. Lack of homoscedasticity is indicated if values are more dispersed for a given predicted values than at other values (Tabachnick and Fidell, 2014).



The points on the graph are well distributed across the figure and shows no particular patterns, which indicates that the assumptions of linearity and homoscedasticity were not violated.

Multicollinearity

Multicollinearity refers to excessive correlations among the predictor variables, which is not desirable.

A method to test for multicollinearity is looking at Variance Inflation Factors (VIFs). When values are below 10.0, multicollinearity can be assumed to be absent (Hair et al., 2014).

Coefficients ^a		Collinearity Statistics	
Model		Tolerance	VIF
1	Transportation	0,963	1,039
	Involvement	0,854	1,171
	Familiarity	0,828	1,207

a. Dependent Variable: BA

Regression Results

Multiple Regression analysis is a technique used to explore the relationships between a continuous dependent variable and two or more independent (or predictor) variables (Pallant, 2010). The objective of multiple regression analysis is to use the independent variables whose values are known to predict the single dependent value selected by the researcher. Each independent variable is weighted by the regression analysis procedure to ensure maximal prediction from the set of independent variables. The weights denote the relative contribution of the independent variables to the overall prediction and facilitate interpretation as to the influence of each variable in making the prediction, although correlation among the independent variables complicates the interpretative process. The set of weighted independent variables forms the regression variate, a linear combination of the independent variables that best predicts the dependent variable (Hair et al., 2014).

The ANOVA test for the model was significant ($p < .05$) (figure below), meaning that the model is valid and the regression coefficients can be interpreted.

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	115,520	3	38,507	43,281	0,000 ^b
	Residual	147,688	166	0,890		
	Total	263,208	169			

a. Dependent Variable: BA

b. Predictors: (Constant), Familiarity, Transportation, Involvement

Next figure shows the adjusted R-square for the model ($\text{Adj } R^2 = .429$). This value means the total variance in the variable 'Brand Attitude' that is explained by all the variables that were inserted in the model. In other words, all variables explain 42.9% of the variance in 'Brand Attitude'.

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0,662 ^a	0,439	0,429	0,94323

a. Predictors: (Constant), Familiarity, Transportation, Involvement

b. Dependent Variable: BA

Next, we look at beta coefficients and their significance (columns 'Beta' and 'Sig.', below).

A p-value (Sig.) of less than 0.05 indicated the model is significant at the 95% confidence level (or 5% significance level). This means that the regression variate (set of independent variables) has a significant effect on the outcome variable.

While the t-statistics is not directly interpretable (same as the F-statistic), the p-value represents the results of the significance test of the variable. Values below 0.05 represent statistical significance of the predictor. That is, the variable significantly influences the outcome variable. The beta coefficient refers to the absolute change in the outcome variable that happens for a one-unit change on the independent variable, while keeping the values of the additional variables inserted in the model constant. So, negative beta coefficients indicate an inverse relationship between the variables, whereas positive values represent the opposite.

The figure below shows the model coefficients.

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients		Sig.	95,0% Confidence Interval for B	
		B	Std. Error	Beta	t		Lower Bound	Upper Bound
1	(Constant)	1,613	0,376		4,286	0,000	0,870	2,355
	Transportation	0,320	0,061	0,310	5,224	0,000	0,199	0,441
	Involvement	0,141	0,046	0,194	3,077	0,002	0,051	0,232
	Familiarity	0,336	0,049	0,442	6,918	0,000	0,240	0,432

a. Dependent Variable: BA

Transportation shows a positive effect on brand attitude, while keeping familiarity and involvement constant ($\beta = .320$, $p < .001$). A one-unit increase on Transportation is expected to have a positive effect of 0.320 on the score of brand attitude. The 95% confidence interval of the beta coefficient is between 0.199 and 0.441. This means that

we can be 95% sure that the 'real' regression parameter lies anywhere between 0.199 and 0.441.

Involvement has a lower effect on brand attitude, compared to transportation, but still significant ($p < .01$). A one-unit increase on involvement is expected to generate an increase on the score of brand attitude of 0.141, while holding transportation and familiarity constant ($\beta = .141$, $p < .01$). The 95% confidence interval is between 0.051 and 0.232.

In terms of familiarity, this concept is also significantly related to brand attitude. The effect of familiarity is positive ($\beta = .336$, $p < .001$). The 95% confidence interval is between 0.240 and 0.432.

Conclusion

The major conclusions that can be drawn based on the analyses conducted earlier are the following:

- H1 – There is a positive association between Transportation and Sympathy and Empathy: hypothesis confirmed, since the test was statistically significant;
- H2 - There is a positive association between Retrospective Reflection and Sympathy and Empathy: hypothesis confirmed, since the test was statistically significant;
- H3a – There is a positive influence of Transportation on Brand Attitude, while holding Involvement and Familiarity constant: confirmed, since the test was statistically significant;
- H3b - There is a positive influence of Involvement on Brand Attitude, while holding Transportation and Familiarity constant: confirmed, since the test was statistically significant;
- H3c: There is a positive influence of Involvement on Brand Attitude, while holding Transportation and Familiarity constant: confirmed, since the test was statistically significant.