

# Multiple Linear Regression Report

## 8.1.1 Descriptive Statistics

Measures of central tendency and measures of spread:

|              | <b>continuance_purchase_intention</b> | <b>enjoyment</b> | <b>entertainment</b> | <b>education</b> | <b>flow</b> | <b>design_background</b> |
|--------------|---------------------------------------|------------------|----------------------|------------------|-------------|--------------------------|
| <b>count</b> | 309.000                               | 309.000          | 309.000              | 309.000          | 309.000     | 309.000                  |
| <b>mean</b>  | 4.099                                 | 4.188            | 4.228                | 4.286            | 4.125       | 4.453                    |
| <b>std</b>   | 0.780                                 | 0.808            | 0.758                | 0.689            | 0.697       | 0.535                    |
| <b>min</b>   | 1.000                                 | 1.000            | 1.500                | 1.250            | 1.200       | 1.500                    |
| <b>25%</b>   | 3.667                                 | 4.000            | 4.000                | 4.000            | 3.800       | 4.000                    |
| <b>50%</b>   | 4.333                                 | 4.000            | 4.500                | 4.250            | 4.200       | 4.667                    |
| <b>75%</b>   | 4.667                                 | 5.000            | 5.000                | 5.000            | 4.600       | 4.833                    |
| <b>max</b>   | 5.000                                 | 5.000            | 5.000                | 5.000            | 5.000       | 5.000                    |

Box and Whisker Plot:

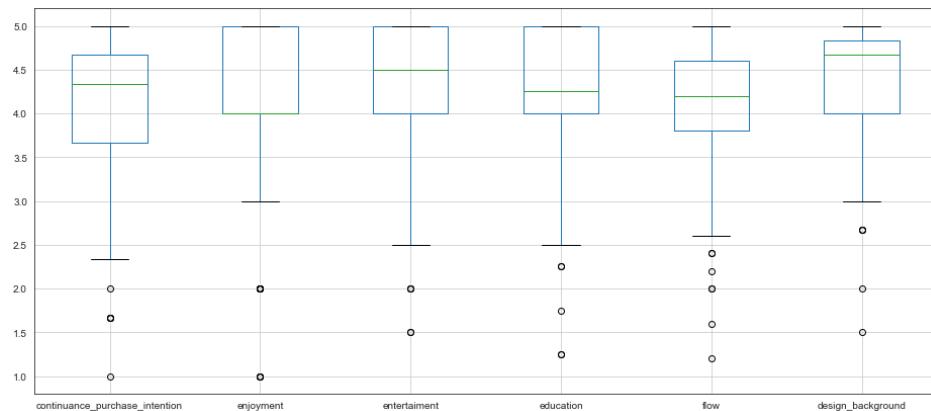


Figure 1 Box and Whisker Plot

Barchart of Continuance Purchase Intention

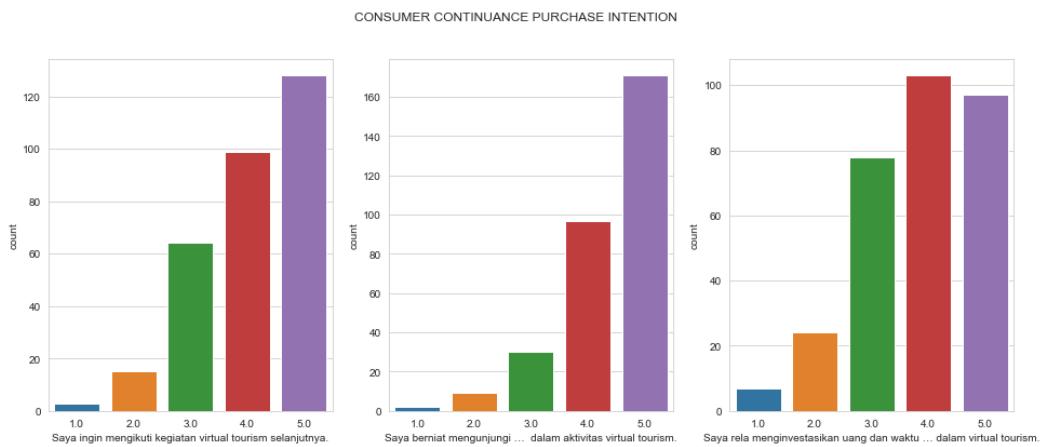
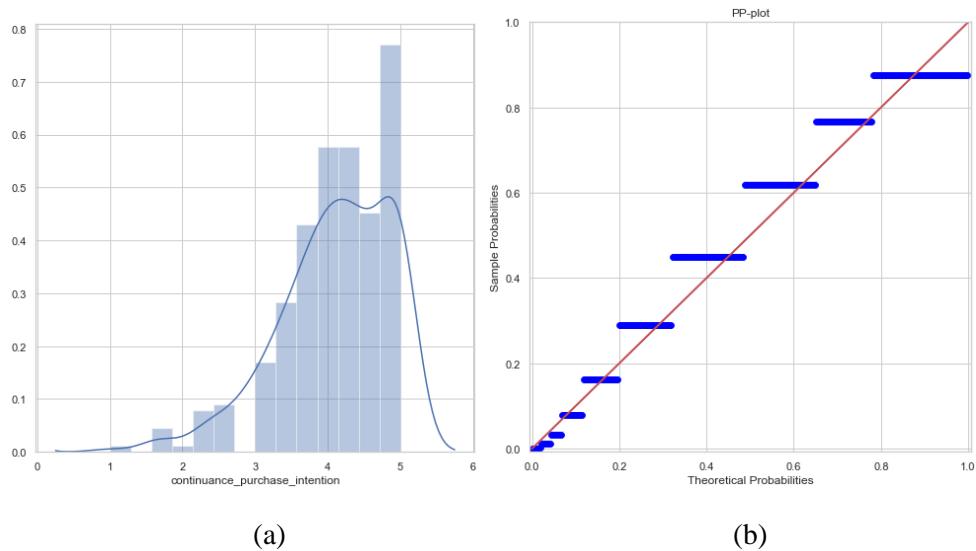


Figure 2 Barchart of Continuance Purchase Intention

### 8.1.2 Normal Distribution



*Figure 3 (a) Histogram plot with KDE of Continuance Purchase Intention. (b) Probability plot of Continuance Purchase Intention*

Skewness test was conducted to find the Z score of Continuance Purchase Intention. The result is shown in table 1.

*Table 1 Skewness Test Result*

|                |           |
|----------------|-----------|
| <b>Z Score</b> | -5.892    |
| <b>P value</b> | 3.797e-09 |

To find if the dataset has been drawn from a normality distributed population, Normality Test (D'Agostino and Pearson's) was conducted. . The result is shown in table 2.

*Table 2 D'Agostino and Pearson's Normality Test Result*

|                |           |
|----------------|-----------|
| <b>P value</b> | 1.442e-09 |
|----------------|-----------|

Since the P value is below the significance test value ( $\alpha = 0.05$ ), the null hypothesis can be rejected (Normal Distribution).

### 8.1.3 Pearson Correlation and Sigma 2 tailed

Pearson correlation test was conducted to find the correlation value and two-tailed p-value for all of the variables. The result can be shown in table 3.

*Table 3 Pearson Correlation Table*

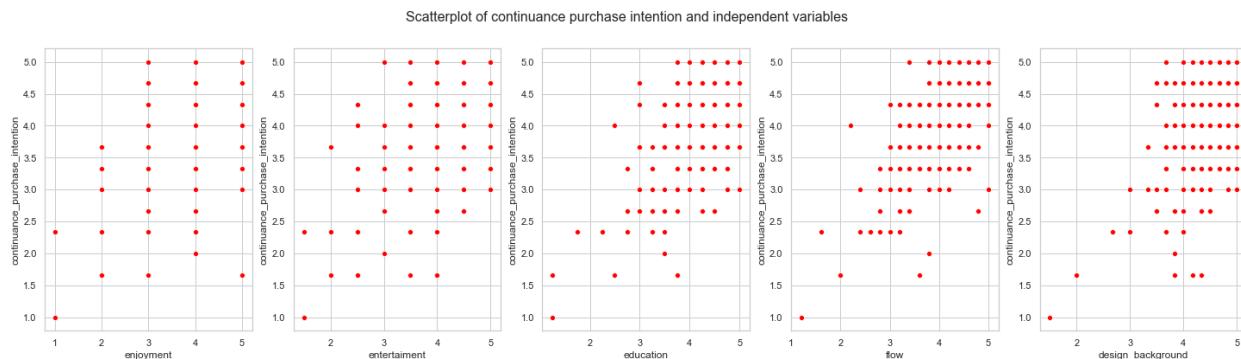
|                                       | Continuance Purchase Intention | Enjoyment | Entertainment | Education | Flow  | Design Background |
|---------------------------------------|--------------------------------|-----------|---------------|-----------|-------|-------------------|
| <b>Continuance Purchase Intention</b> | 1.000                          | 0.577     | 0.631         | 0.722     | 0.718 | 0.589             |
| <b>Enjoyment</b>                      | 0.577                          | 1.000     | 0.735         | 0.715     | 0.681 | 0.510             |
| <b>Entertainment</b>                  | 0.631                          | 0.735     | 1.000         | 0.715     | 0.745 | 0.536             |
| <b>Education</b>                      | 0.722                          | 0.715     | 0.715         | 1.000     | 0.778 | 0.675             |
| <b>Flow</b>                           | 0.718                          | 0.681     | 0.745         | 0.778     | 1.000 | 0.641             |
| <b>Design Background</b>              | 0.589                          | 0.510     | 0.536         | 0.675     | 0.641 | 1.000             |

The two tailed P value of the predictor against response is shown in table 4.

*Table 4 two tailed P value of the predictor against response*

| Predictor                | Response | Two-tailed P-value |
|--------------------------|----------|--------------------|
| <b>Enjoyment</b>         |          | 0.000              |
| <b>Entertainment</b>     |          | 0.000              |
| <b>Education</b>         |          | 0.000              |
| <b>Flow</b>              |          | 0.000              |
| <b>Design Background</b> |          | 0.000              |

All predictors have significant p-value (below 0.05). The highest pearson correlation between predictor and response was found in the Education – Continuance Purchase Intention (0,722), followed by Flow – Continuance Purchase Intention (0,718).



*Figure 4 Scatterplot of Continuance Purchase Intention and Independent Variables*

## 8.1.4 ANOVA

To analyze ANOVA, first the independent variables was fitted to multiple regression model. The model architecture is shown in the figure 5.

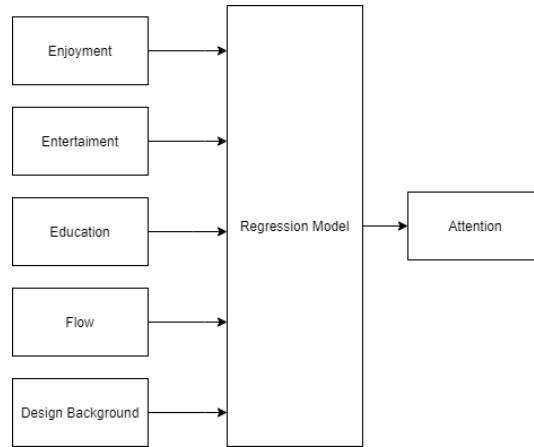


Figure 5 Multiple Regression Model

The result or summary of the regression is shown in the figure 6.

| OLS Regression Results |                                |                     |          |        |        |       |
|------------------------|--------------------------------|---------------------|----------|--------|--------|-------|
| Dep. Variable:         | continuance_purchase_intention | R-squared:          | 0.594    |        |        |       |
| Model:                 | OLS                            | Adj. R-squared:     | 0.587    |        |        |       |
| Method:                | Least Squares                  | F-statistic:        | 88.55    |        |        |       |
| Date:                  | Fri, 29 Jan 2021               | Prob (F-statistic): | 3.62e-57 |        |        |       |
| Time:                  | 15:38:19                       | Log-Likelihood:     | -221.86  |        |        |       |
| No. Observations:      | 309                            | AIC:                | 455.7    |        |        |       |
| Df Residuals:          | 303                            | BIC:                | 478.1    |        |        |       |
| Df Model:              | 5                              |                     |          |        |        |       |
| Covariance Type:       | nonrobust                      |                     |          |        |        |       |
| coef                   | std err                        | t                   | P> t     | [0.025 | 0.975] |       |
| Intercept              | -0.1099                        | 0.243               | -0.453   | 0.651  | -0.587 | 0.368 |
| enjoyment              | -0.0111                        | 0.057               | -0.194   | 0.846  | -0.124 | 0.101 |
| entertainment          | 0.1076                         | 0.065               | 1.664    | 0.097  | -0.020 | 0.235 |
| education              | 0.3778                         | 0.078               | 4.867    | 0.000  | 0.225  | 0.531 |
| flow                   | 0.3551                         | 0.075               | 4.757    | 0.000  | 0.208  | 0.502 |
| design_background      | 0.1609                         | 0.075               | 2.153    | 0.032  | 0.014  | 0.308 |
| Omnibus:               | 37.377                         | Durbin-Watson:      | 2.040    |        |        |       |
| Prob(Omnibus):         | 0.000                          | Jarque-Bera (JB):   | 69.844   |        |        |       |
| Skew:                  | -0.675                         | Prob(JB):           | 6.82e-16 |        |        |       |
| Kurtosis:              | 4.899                          | Cond. No.           | 83.9     |        |        |       |

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Figure 6 Regression Result

The F-test result for regression model is shown in the table 5.

*Table 5 F-test Result for Regression Model*

|                     |          |
|---------------------|----------|
| <b>F Statistics</b> | 88.55    |
| <b>P Value</b>      | 3.62e-57 |

Figure 6 print out the probability of occurrence of the F-statistic under the assumption that the null hypothesis is true. The p value shown by figure 6 and table 5 is 3.62e-57 which is smaller than alpha (0.05). Thus, there is much less than 5% chance that the F-statistic of 88.55 could have occurred by chance under the assumption of a valid null hypothesis. This means that the model can explain the variance in the dependent variable better than intercept-only model.

ANOVA per predictors was conducted and the result is shown in the table 6.

*Table 6 ANOVA per Predictors*

|                          | <b>Sum of Square</b> | <b>Degree of Freedom</b> | <b>F Score</b> | <b>P Value (PR (&gt; F))</b> |
|--------------------------|----------------------|--------------------------|----------------|------------------------------|
| <b>enjoyment</b>         | 62.273               | 62.273                   | 248.099        | 0.000                        |
| <b>entertainment</b>     | 17.382               | 17.382                   | 69.251         | 0.000                        |
| <b>education</b>         | 22.990               | 22.990                   | 91.595         | 0.000                        |
| <b>flow</b>              | 7.316                | 7.316                    | 29.149         | 0.000                        |
| <b>design_background</b> | 1.163                | 1.163                    | 4.635          | 0.032                        |
| <b>Residual</b>          | 76.053               | 0.251                    |                |                              |

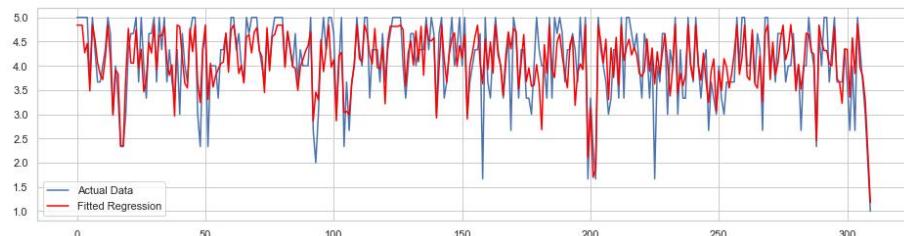
Table 6 shows that each predictor have significant value. It means that there is a statistically significant difference between the groups and their effects.

ANOVA one way between regression line and actual data (dependent variable) was also conducted, the result shown in table 7.

*Table 7 ANOVA one way between Regression Line and Dependent Variable*

|                             |           |
|-----------------------------|-----------|
| <b>F one-way statistics</b> | 2.842e-28 |
| <b>P value</b>              | 1.0       |

P value is higher than 0.05, that means there is no difference in the variance between regression line and the actual data (dependent variable) which means the regression line can explain the variance.



*Figure 7 Actual Data vs Fitted Regression*

### **8.1.5 Model Summary or Goodness of Fit**

Multi regression model's goodness of fit is explained by R squared and Adjusted R squared in the table 8.

*Table 8 R squared and Adjusted R squared of Regression Model*

|                           |       |
|---------------------------|-------|
| <b>R squared</b>          | 0.594 |
| <b>Adjusted R Squared</b> | 0.587 |

The R squared of the regression model is higher than 0.5, it means the multiple regression model is fit to predict the outcome in equal to R-squared value.

### **8.1.6 Table of Coefficient**

Multicollinearity in multiple regression model's predictors is shown in the table 9. Table 9 shows that there is high multicollinearity among the predictors. The highest multicollinearity value is owned by Education variable.

*Table 9 Multiple Regression Predictor's VIF*

| <b>Variables</b>  | <b>VIF</b> |
|-------------------|------------|
| Enjoyment         | 73.111     |
| Entertainment     | 94.165     |
| Education         | 139.771    |
| Flow              | 120.017    |
| Design Background | 83.504     |

The positive or negative correlation of the independent variable in regression coefficient along with its p value is shown in table 10.

*Table 10 Coefficient Correlation and P-values*

| <b>Variables</b>  | <b>Coefficient</b> | <b>P Value</b> |
|-------------------|--------------------|----------------|
| Enjoyment         | -0.011             | 0.846          |
| Entertainment     | 0.108              | 0.097          |
| Education         | 0.378              | 0.000          |
| Flow              | 0.356              | 0.000          |
| Design Background | 0.161              | 0.032          |

Table 10 concluded that enjoyment and entertainment are not statistically significant and this result favor that these variables may have no significant correlation with the continuance purchase intention as the p values exceed the alpha value (0.05). On the other side, education, flow, and design background are statistically significant and have positive correlation with continuance purchase intention as the p value falls below the alpha value (0.05).

### **8.1.7 Summary of each hypothesis to continuance purchase intention**

Terdapat 5 hipotesis dengan masing-masing hipotesis adalah korelasi antara variabel independen dengan variabel dependen. Dari tes yang telah dilakukan, didapatkan kesimpulan sebagai berikut:

- Variabel *Enjoyment* tidak memiliki hubungan langsung secara positif terhadap *continuance purchase intention*.
- Variabel *Entertainment* tidak memiliki hubungan langsung secara positif terhadap *continuance purchase intention*.
- Variabel *Education* memiliki hubungan langsung secara positif terhadap *continuance purchase intention*.
- Variabel *Flow* memiliki hubungan langsung secara positif terhadap *continuance purchase intention*
- Variabel *Design Background* memiliki hubungan langsung secara positif terhadap *continuance purchase intention*