**Assignment #5**

**20175404 Rocku Oh**

You will analyze the scarcity dataset. Produce the tables or output followed by an explanation. You will first run a MANOVA with the dependent variables being q2, q3, q4, and q7 and the independent variable being status. Then you will run a MANCOVA with the dependent variables being q2, q3, q4, and q7 and the main independent variable being status controlling for company, product, use, need, and shipping. For the MANCOVA, you are at least supposed to evaluate the interaction effects of status with each control variable.

**Question 1**

Assess whether the data is multivariate normally distributed and meets basic assumptions.

 



Table 1

*Test of equality of covariance matrices across dependent variables (4 questions)*

|  |  |  |  |
| --- | --- | --- | --- |
| Coefficients | dof | Value | Prob > value |
| Modified LR chi2 |  | 16.06631 |  |
| Box F | 20 | 0.80 | 0.7176 |
| Box chi2 | 20 | 2.9417 | 0.7176 |

Table 2

*Correlation matrix of the dependent variables (4 questions)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Q2 | Q3 | Q4 | Q7 |
| Q2 (disappointment for the product status) |  |  |  |  |
| Q3 (disappointment for the website) | 0.3584 |  |  |  |
| Q4 (belief of 'bad luck' when product was not available) | 0.1816 | -0.0677 |  |  |
| Q7 (likelihood of re-visit in the future) | 0.0746 | -0.3291 | 0.1983 |  |

As shown in the Table 1, Box’s *M* is not statistically significant with the level of 0.05, it means the dependent variables satisfy assumption that the covariance matrices are homogeneity.

Dependent variables are not highly correlated based on the correlation values of dependent variable pairs (see Table 2).

**Question 2**

What hypotheses will your one-way MANOVA be testing?

A MANOVA or multivariate analysis of variance is a way to test the hypothesis that one or more independent variables, or factors, have an effect on a set of two or more dependent variables. The null hypothesis is that there are no difference between group means for a particular combination of dependent variables.

* H0 : μ2 = μ3 = μ4 = μ7
* H1 :μik≠μjk for at least one i≠j and at least one variable k

**Question 3**

Run a one-way MANOVA. Report and discuss Wilk’s Lambda?

As shown in the Table 3, a one-way MANOVA revealed a significant multivariate main effect for product status, Wilks’ λ = .965, F-statistic = 5.01, p <. 001 (significant level of 0.05). Thus we can reject hull hypothesis, which means there are statistically significant difference between group means for 4 questions.

Table 3

*MANOVA result for the 4 questions on product status*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Source | Test | statistic | df | F(df1 | df 2) | F | Prob > F |
| status | Wilks’ lambda | 0.9649 | 2 | 8 | 2222 | 5.01 | 0 |
|  | Hotelling trace | 0.0362 |  | 8 | 2220 | 5.03 | 0 |
|  | Pillai’s trace | 0.0353 |  | 8 | 2224 | 4.99 | 0 |
|  | Roy’s largest root | 0.0308 |  | 4 | 1112 | 8.56 | 0 |

*Note.* The number of total sample size and residual is 1116 and 1114 respectively.

Table 4

*ANOVA results of product status on the 4 questions*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Source | Dependent variable | Sum of Squares | df | Mean Square | F | Prob > F |
| status | Q2 | 14.1889 | 2 | 7.0945 | 4.65 | 0.0097 |
|  | Q3 | 36.043763 | 2 | 18.021882 | 12.03 | 0 |
|  | Q4 | 18.13272 | 2 | 9.06636 | 5.63 | 0.0037 |
|  | Q7 | 3.544239 | 2 | 1.7721195 | 1.56 | 0.2116 |

**Question 4**

What specific conclusion can you draw from the MANOVA results and the univariate ANOVA follow-up tests?

The Table 4 reported the result of the ANOVA tests on the four dependent variables, Q2 (disappointment of product status), Q3 (disappointment of the website), Q4 (belief of 'bad luck' when product was not available), and Q7 (likelihood of re-visit in the future).

Given the significance of the overall test, the univariate main effects were examined. Significant univariate main effects for product status were obtained for Q2, *F* (2, 1114) = 4.65, *p* =.0097; Q3, *F* (2, 1114) =12.03, *p* <.001; and Q4, *F* (2, 1114) =5.63, *p* =.0037 with the confidence level of 0.0125 (=0.05/4).

**Question 5**

What specific conclusions can you draw from the pairwise comparison?

Significant pairwise differences were obtained in question of disappointment for the product status between the ‘sold out’ and both the ‘Unavailable’ and ‘Out-of-stock’. Since we are doing 3 significant tests looking at the pairwise tests comparing disappointment level by product status, we need to use smaller confidence level again to protect against inflated alpha error, so let’s divide the .05 by 3 and set .017 as our error level.

As shown in the Table 5, the ‘Out-of-stock’ and ‘Sold out’ are significantly different in disappointment level of customers when the product is not available to buy (significance level of .05). ‘Out-of-stock’ give 0.28 point more disappointment emotion to customer as compared to the ‘Sold out’ status.

In the case of the question 3 which is how the customers are disappointed with the website, the difference between the ‘Sold out’ and the ‘Unavailable’ and the ‘Out-of-stock’ and the ‘Sold out’ are statistically significant in disappointment of website (significance level of .05). ‘Out-of-stock’ arouse 0.31 point more disappointed emotion to the customers as compared to the ‘Sold out’ status. ‘Sold out’ give 0.42 point less disappointed emotion to the customers as comparing with the ‘Unavailable’ status.

In the case of the question 4 which is how the customers believe it was simply 'bad luck' that the product was out-of-stock/sold out/unavailable, the difference between the ‘Sold out’ and the ‘Unavailable’ are statistically significant (significance level of .05). The customers are 0.3 point level more likely to believe that the status of the product is simply bad luck compared to the ‘Unavailable’.

A question 7 (How likely are you to visit this website in the future?) is not statistically influenced by the variable of current product status.

Table 5

*Pairwise comparison for group mean of dependent variable on product status*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Dependent Variable | (I) status | (J) status | Mean Difference  (I-J) | Standard Error | t-statistic | Prob > F | 95% CI |
| Q2 | Sold out | Unavailable | -0.1484 | 0.0897 | -1.65 | 0.098 | -0.3245 0.0277 |
|  | Out-of-stock | Unavailable | 0.1291 | 0.0907 | 1.42 | 0.155 | -0.0488 0.3070 |
|  | Out-of-stock | Sold out | 0.2776 | 0.0911 | 3.05 | **0.002** | 0.0988 0.4563 |
|  |  |  |  |  |  |  |  |
| Q3 | Sold out | Unavailable | -0.4218 | 0.0890 | -4.74 | **0.000** | -0.5963 -0.2472 |
|  | Out-of-stock | Unavailable | -0.1102 | 0.0899 | -1.23 | 0.220 | -0.2866 0.0661 |
|  | Out-of-stock | Sold out | 0.3116 | 0.0903 | 3.45 | **0.001** | 0.1344 0.4887 |
|  |  |  |  |  |  |  |  |
| Q4 | Sold out | Unavailable | 0.3092 | 0.0922 | 3.35 | **0.001** | 0.1283 0.4902 |
|  | Out-of-stock | Unavailable | 0.1659 | 0.0932 | 1.78 | 0.075 | -0.0170 0.3487 |
|  | Out-of-stock | Sold out | -0.1433 | 0.0936 | -1.53 | 0.126 | -0.3270 0.0403 |
|  |  |  |  |  |  |  |  |
| Q7 | Sold out | Unavailable | 0.1217 | 0.0776 | 1.57 | 0.117 | -0.0306 0.2739 |
|  | Out-of-stock | Unavailable | 0.0051 | 0.0784 | 0.07 | 0.948 | -0.1487 0.1590 |
|  | Out-of-stock | Sold out | -0.1166 | 0.0788 | -1.48 | 0.139 | -0.2711 0.0380 |

**Question 6**

Run the MANCOVA, univariate ANCOVA, and pairwise comparisons. Just provide the tables. You do not have to provide the interpretation.

Table 6

*MANOVA result (Wilk’s lambda) for the 4 questions on independent variables (status, company, product, use, need, and shipping)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Source | statistic | df | F(df1 | df 2) | F | Prob > F |
| status | 0.9936 | 2 | 8 | 2192 | 0.88 | 0.531 |
| company | 0.9415 | 1 | 4 | 1096 | 17.03 | **<0.001** |
| product | 0.9937 | 1 | 4 | 1096 | 1.74 | 0.1381 |
| use | 0.9878 | 1 | 4 | 1096 | 3.39 | 0.0091 |
| need | 0.9935 | 1 | 4 | 1096 | 1.81 | 0.1252 |
| shipping | 0.9936 | 1 | 4 | 1096 | 1.77 | 0.1330 |
| status \* company | 0.986 | 2 | 8 | 2192 | 1.94 | 0.0507 |
| status \* product | 0.9953 | 2 | 8 | 2192 | 0.64 | 0.7444 |
| status \* use | 0.9953 | 2 | 8 | 2192 | 0.64 | 0.7411 |
| status \* need | 0.9926 | 2 | 8 | 2192 | 1.02 | 0.4168 |
| status \* shipping | 0.9943 | 2 | 8 | 2192 | 0.79 | 0.6150 |

*Note.* The number of total sample size and residual is 1116 and 1099 respectively. The significant level is 0.008 (=0.05/6). The only significant difference between group means is the case of the company variable.

Table 7

*ANOVA results of different independent variables status and company with interaction term on the 4 questions*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Source | Dependent variable | Sum of Squares | df | Mean Square | F | Prob > F |
| company | Q2 | 3.3436 | 1 | 3.3436 | 2.2 | 0.1385 |
|  | Q3 | 17.7652 | 1 | 17.7652 | 11.79 | **0.0006** |
|  | Q4 | 3.5374 | 1 | 3.5374 | 2.19 | 0.1388 |
|  | Q7 | 60.8229 | 1 | 60.8229 | 55.86 | **< 0.001** |
| company\*status | Q2 | 14.8184 | 2 | 7.4092 | 4.87 | **0.0078** |
|  | Q3 | 9.9226 | 2 | 4.9613 | 3.29 | 0.0375 |
|  | Q4 | 13.5953 | 2 | 6.7977 | 4.22 | 0.015 |
|  | Q7 | 0.0325 | 2 | 0.0162 | 0.01 | 0.9852 |

*Note.* The number of total sample size and residual is 1116 and 1099 respectively. The significant level is 0.0125 (=0.05/4). The only significant difference between group means is the case of the company variable.