Fall 2015 Statistics and Practice

**Term Project**

**Soongsil University**

**School of Computer Science and Engineering**

|  |  |
| --- | --- |
| **Team Members** | **Student ID** |
| **김동현** | **20122343** |
| **나영현** | **20122367** |
| **박영준** | **20122379** |

Table of Contents

[**Analysis Planning** 3](#_Toc438491023)

[**Analysis Setting** 5](#_Toc438491024)

[**Analysis 1** 7](#_Toc438491025)

[**Test A:** 7](#_Toc438491026)

[**Test B:** 10](#_Toc438491027)

[**Test C** 15](#_Toc438491028)

[**Test D** 18](#_Toc438491029)

[**Analysis Outcome** 27](#_Toc438491030)

[**Analysis 2** 29](#_Toc438491031)

[**Test A** 29](#_Toc438491032)

[**Test B** 32](#_Toc438491033)

[**Test C** 36](#_Toc438491034)

[**Analysis Outcome** 46](#_Toc438491035)

[**Analysis 3** 48](#_Toc438491036)

[**Test A:** 48](#_Toc438491037)

[**Test B** 51](#_Toc438491038)

[**Test C** 55](#_Toc438491039)

[**Analysis Outcome** 57](#_Toc438491040)

[**Analysis Extra** 58](#_Toc438491041)

[**Test A** 58](#_Toc438491042)

[**Test B** 61](#_Toc438491043)

[**Analysis Outcome** 65](#_Toc438491044)

[**Final Project** 66](#_Toc438491045)

# **Analysis Planning**

|  |  |
| --- | --- |
| Variable Name | Related Questions |
| Privacy Score | 10,11,12,14,15,33,34 |
| Awareness Score | 5,18,19,24,27,29,30,36 |
| Utilization Score | 5,6,20,44 |
| Satisfaction Score | 13,21,31,32,35 |

|  |  |  |
| --- | --- | --- |
| Test No. | Analysis Statement | Hypothesis |
| 1 | To study the privacy concerns in loyalty program between male and female students | About the same |
| 2. | To study the privacy concerns in loyalty program between age group. | Older people more concern about privacy |
| 3. | To study the privacy concerns in loyalty program whether have experience Data leakage | People who have experience Data leakage are more concern about privacy |
| 4. | To study the privacy concerns in loyalty program between gender, age and whether they have experienced data leakage. | Male and Female about same. older people more concern about privacy. And people who have experienced data leakage are more concern about privacy |
| 5. | To study the Utilization concerns in loyalty program between male and female | Female use loyalty program more effectively than male. |
| 6. | To study the Utilization concerns in loyalty program according to sign purpose | All means About same. |
| 7. | To study the Utilization concerns in loyalty program between male and female according to sign purpose | Female’s Utilization is higher than Male’s Utilization and Sign purpose doesn’t affect Utilization |
| 8. | To study the awareness concerns in loyalty program between male and female | About same. |
| 9. | To study the awareness concerns in loyalty program according to age | Older people more concern about privacy. |
| 10 | To study the awareness concerns in loyalty program between age and gender. | Young people more concern about privacy than older people. Man and Female is about same |
| Extra | To study the satisfaction concerns in loyalty program whether have experienced data leakage. | People who experienced data leakage are more unsatisfied with loyalty program |
| To study the satisfaction concerns In loyalty program between male and female gained control awareness to reduce error. | Use Awareness Score for covariate reduce error, and experiment male, female’s satisfaction relationship |

# **Analysis Setting**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Analysis No. | Analysis Statement | Test Name | Variable Types | Hypothesis | Assumption Tests |
| 1 | To study the privacy concerns in loyalty program between male and female students | Independent T-Test | - Privacy Score (Dependent)  - Gender (Independent) |  | 1. Normality Test  2. Homogeneity Variance |
| To study the privacy concerns in loyalty program between age group. | One-Way ANOVA | - Privacy Score (Dependent)  - Age (Independent) | All the means are same  At least two means are different | 1. Normality Test  2. Homogeneity Variance |
| To study the privacy concerns in loyalty program whether have experienced data leakage | Independent T-Test | - Privacy Score (Dependent)  - Experience Data Leakage  (Independent) |  | 1. Normality Test  2. Homogeneity Variance |
| To study the privacy concerns in loyalty program between gender, age and whether they have experienced damage. | Three-Way ANOVA | - Privacy Score (Dependent)  - Gender (Independent)  - Age (Independent)  - Experience Data Leakage  (Independent) | All interaction means are the same  At least two interaction means are different | 1. Know Interaction means |
| 2 | To study the Utilization concerns in loyalty program between male and female. | Independent T-Test | - Utilization Score (Dependent)  - Gender (Independent) |  | 1. Normality Test  2. Homogeneity Variance |
| To study the Utilization concerns in loyalty program according to sign purpose. | One-Way ANOVA | - Utilization Score (Dependent)  - Sign Purpose (Independent) | All the means are same  At least two means are different | 1. Normality Test  2. Homogeneity Variance |
| To study the Utilization concerns in loyalty program between male and female according to sign purpose. | Two-Way ANOVA | - Utilization Score (Dependent)  - Gender (Independent)  - Sign Purpose (Independent) | All interaction means are the same  At least two interaction means are different | 1. Know Interaction means |
| 3 | To study the awareness concerns in loyalty program between male and female | Independent T-Test | - Awareness Score (Dependent)  - Gender (Independent) |  | 1. Normality Test  2. Homogeneity Variance |
| To study the awareness concerns in loyalty program according to age. | One-Way ANOVA | - Awareness Score (Dependent)  - Age (Independent) | All the means are same  At least two means are different | 1. Normality Test  2. Homogeneity Variance |
| To study the awareness concerns in loyalty program between age and gender. | Two-Way ANOVA | - Awareness Score (Dependent)  - Age (Independent)  - Gender (Independent) | All interaction means are the same  At least two interaction means are different | 1. Know Interaction means |
| Extra | To study the satisfaction concerns in loyalty program whether have experienced data leakage. | Independent T-Test | - Satisfaction Score (Dependent)  - Experience Data Leakage  (Independent) |  | 1. Normality Test  2. Homogeneity Variance |
| To study the satisfaction concerns  In loyalty program between male and female gained control awareness to reduce error. | ANCOVA | - Satisfaction Score (Dependent)  - Gender (Independent)  - Awareness Score (Covariate) | All adjusted means are same  At least two adjusted means are different | 1. Independence of the independent variable and covariate.  2. Regression Homogeneity |

# **Analysis 1**

## **Test A:**

|  |
| --- |
| Analysis Statement: To study the privacy concerns in loyalty program between male and female. |
| Assumption Test(s):      It’s an approximately normal because there is a symmetric line  Leven’s Test   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Independent Samples Test** | | | | | | | | | | | |  | | **Levene's Test for Equality of Variances** | | **t-test for Equality of Means** | | | | | | | | **F** | **Sig.** | **t** | **df** | **Sig. (2-tailed)** | **Mean Difference** | **Std. Error Difference** | **95% Confidence Interval of the Difference** | | | **Lower** | **Upper** | | **privacy** | **Equal variances assumed** | **4.575** | **.033** | **2.520** | **2794** | **.012** | **.382** | **.152** | **.085** | **.679** | | **Equal variances not assumed** |  |  | **2.534** | **2721.520** | **.011** | **.382** | **.151** | **.086** | **.678** |   : Homogeneity of variance can be assumed  : Homogeneity of variance cannot be assumed  F(2794) = 4.575 , sig = 0.033 (< 0.05)  This test is significant  Reject  Conclusion : Homogeneity of variance cannot be assumed  So, We do T-Test, when Equal Variance not Assumed |
| Analysis Results:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Group Statistics** | | | | | | |  | **Gender** | **N** | **Mean** | **Std. Deviation** | **Std. Error Mean** | | **privacy** | **male** | **1546** | **23.28** | **4.072** | **.104** | | **female** | **1250** | **22.90** | **3.874** | **.110** |   Independent T-Test:   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Independent Samples Test** | | | | | | | | | | | |  | | **Levene's Test for Equality of Variances** | | **t-test for Equality of Means** | | | | | | | | **F** | **Sig.** | **t** | **df** | **Sig. (2-tailed)** | **Mean Difference** | **Std. Error Difference** | **95% Confidence Interval of the Difference** | | | **Lower** | **Upper** | | **privacy** | **Equal variances assumed** | **4.575** | **.033** | **2.520** | **2794** | **.012** | **.382** | **.152** | **.085** | **.679** | | **Equal variances not assumed** |  |  | **2.534** | **2721.520** | **.011** | **.382** | **.151** | **.086** | **.678** |   :  :  Since Sig = 0.011 (< 0.05)  This test is significant  Reject  Conclusion: The different between two means is significant / are different. |
| Conclusion(s):  Male’s mean of privacy score is 23.28, SE = 0.104 and Female’s mean of privacy score is 22.90 SE =0.110, So, Male more concerns about privacy than female. |

## **Test B:**

|  |
| --- |
| Analysis Statement: To study the privacy concerns in loyalty program between age group |
| Assumption Test(s):    It’s an approximately normal because there is a symmetric line   |  |  |  |  | | --- | --- | --- | --- | | **Test of Homogeneity of Variances** | | | | | privacy | | | | | Levene Statistic | df1 | df2 | Sig. | | .346a | 4 | 2790 | .847 | | a. Groups with only one case are ignored in computing the test of homogeneity of variance for privacy. | | | |   Leven’s Test  : Homogeneity of variance can be assumed  : Homogeneity of variance cannot be assumed  F(4, 2790) = 0.346 , sig = 0.847 (> 0.05)  This is non-significant  Accept  Conclusion : Homogeneity of variance can be assumed |
| Analysis Results:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **ANOVA** | | | | | | | privacy | | | | | | |  | Sum of Squares | df | Mean Square | F | Sig. | | Between Groups | 380.880 | 4 | 95.220 | 6.028 | .000 | | Within Groups | 44072.920 | 2790 | 15.797 |  |  | | Total | 44453.800 | 2794 |  |  |  |   : All the means are same.  : At least two means are different.  F(4,2790) = 6.028, sig = 0.000(< 0.05)  This test is significant  Reject  Conclusion : At least two means are different  Post-Hoc Test   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Multiple Comparisons** | | | | | | | | Dependent Variable: privacy | | | | | | | | Tukey HSD | | | | | | | | (I) 나이 | (J) 나이 | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | | | Lower Bound | Upper Bound | | teenager | twenty | -1.24\* | .285 | .000 | -2.02 | -.46 | | thirty | -1.57\* | .365 | .000 | -2.57 | -.57 | | fourty | -.73 | .403 | .373 | -1.83 | .37 | | more than fifty | -1.29\* | .437 | .026 | -248 | -.10 | | twenty | teenager | 1.24\* | .285 | .000 | .46 | 2.02 | | thirty | -.33 | .260 | .707 | -1.04 | .38 | | fourty | .51 | .312 | .470 | -.34 | 1.36 | | more than fifty | -.05 | .355 | 1.000 | -1.02 | .92 | | thirty | teenager | 1.57\* | .365 | .000 | .57 | 2.57 | | twenty | .33 | .260 | .707 | -.38 | 1.04 | | fourty | .84 | .386 | .186 | -.21 | 1.90 | | more than fifty | .28 | .421 | .965 | -.87 | 1.43 | | fourty | teenager | .73 | .403 | .373 | -.37 | 1.83 | | twenty | -.51 | .312 | .470 | -1.36 | .34 | | thirty | -.84 | .386 | .186 | -1.90 | .21 | | more than fifty | -.56 | .455 | .728 | -1.81 | .68 | | more than fifty | teenager | 1.29\* | .437 | .026 | .10 | 2.48 | | twenty | .05 | .355 | 1.000 | -.92 | 1.02 | | thirty | -.28 | .421 | .965 | -1.43 | .87 | | fourty | .56 | .455 | .728 | -.68 | 1.81 | | Based on observed means.  The error term is Mean Square(Error) = 15.797. | | | | | | | | \*. The mean difference is significant at the 0.05 level. | | | | | | |   Comparison 1:  Since sig = 0.000(< 0.05)  This test is Significant  Reject  Conclusion  Comparison 2:  Since sig = 0.000(< 0.05)  This test is Significant  Reject  Conclusion  Comparison 3:  Since sig = 0.373(> 0.05)  This test is non-Significant  Accept  Conclusion  Comparison 4:  Since sig = 0.026(< 0.05)  This test is Significant  Reject  Conclusion  Total conclusion:   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Descriptives** | | | | | | | | | | privacy | | | | | | | | | |  | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum | | Lower Bound | Upper Bound | | teenager | 215 | 21.96 | 4.077 | .278 | 21.41 | 22.51 | 14 | 34 | | twenty | 200 | 23.20 | 3.936 | .088 | 23.03 | 23.37 | 11 | 35 | | thirty | 265 | 23.53 | 4.056 | .249 | 23.04 | 24.02 | 11 | 33 | | fourty | 177 | 22.69 | 4.133 | .311 | 22.08 | 23.30 | 14 | 35 | | more than fifty | 134 | 23.25 | 4.013 | .347 | 22.57 | 23.94 | 14 | 33 | | Total | 2795 | 23.11 | 3.989 | .075 | 22.96 | 23.26 | 11 | 35 | |
| Conclusion(s):  In conclusion, teenager(M = 21.96, SE = 0.278) is worst in Privacy, forty(M = 22. 69, SE = 0.311) are about same, twenty(M = 23.20, SE = 0.088) is better than forty, more than fifty(M = 23.25, SE = 0.347) is better than twenty, and best Score is thirty(M = 23.53, SE = 0.249), It means comparative young adult is better privacy concern in loyalty program |

## **Test C**

|  |
| --- |
| Analysis Statement: To study the privacy concerns in loyalty program whether they have experienced data leakage. |
| Assumption Test(s):    It’s an approximately normal because there is a symmetric line  Equal Variance   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Independent Samples Test** | | | | | | | | | | | |  | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | | | Lower | Upper | | privacy | Equal variances assumed | .943 | .332 | -.124 | 2790 | .901 | -.036 | .292 | -.608 | .536 | | Equal variances not assumed |  |  | -.118 | 229.935 | .906 | -.036 | .306 | -.639 | .566 |   Leven’s Test  : Homogenity variance can be assumed  : Homogenity variance can not be assumed  F(2790) = 0.943 , sig = 0.332 (> 0.05)  This is non-significant  Accept  Conclusion: Homogeneity variance can be assumed. |
| Analysis Results:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Group Statistics** | | | | | | |  | 멤버십 카드로 인해 피해를 받은 적이 있습니까? | N | Mean | Std. Deviation | Std. Error Mean | | privacy | yes | 202 | 23.07 | 4.203 | .296 | | no | 2590 | 23.11 | 3.975 | .078 |   Independent T-Test   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Independent Samples Test** | | | | | | | | | | | |  | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | | | Lower | Upper | | privacy | Equal variances assumed | .943 | .332 | -.124 | 2790 | .901 | -.036 | .292 | -.608 | .536 | | Equal variances not assumed |  |  | -.118 | 229.935 | .906 | -.036 | .306 | -.639 | .566 |   :  :  Since Sig = 0.901 (> 0.05)  This test is non-significant  Accept  Conclusion : The different between two means is non-significant / |
| Conclusion(s):  In conclusion, experience of data leakage is not important to affect privacy Score.  Answer Yes (M = 23.07, SE = 0.296) and No (M = 23.11, SE = 0.078) are about same. |

## **Test D**

|  |
| --- |
| Analysis Statement: To study the privacy concerns in loyalty program between gender, age and whether they have experienced data leakage. |
| Assumption Test(s): No Assumption test |
| Analysis Results:   |  |  |  |  | | --- | --- | --- | --- | | **Levene's Test of Equality of Error Variancesa** | | | | | Dependent Variable: privacy | | | | | F | df1 | df2 | Sig. | | .924 | 19 | 2772 | .553 | | Tests the null hypothesis that the error variance of the dependent variable is equal across groups. | | | | |  | | | |   a. Design: Intercept + Q1\_Gender + Q2\_Age + Q25\_Damage\_because\_of\_cards + Q1\_Gender \* Q2\_Age + Q1\_Gender \* Q25\_Damage\_because\_of\_cards + Q2\_Age \* Q25\_Damage\_because\_of\_cards + Q1\_Gender \* Q2\_Age \* Q25\_Damage\_because\_of\_cards  Leven’s Test  : Homogeneity of variance can be assumed  : Homogeneity of variance cannot be assumed  F(19, 2772) = 0.924, sig = 0.553 (> 0.05)  This is non-significant  Accept  Conclusion : Homogeneity of variance can be assumed   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Tests of Between-Subjects Effects** | | | | | | | Dependent Variable: privacy | | | | | | | Source | TypeIII Sum of Squares | df | Mean Square | F | Sig. | | Corrected Model | 771.118a | 19 | 40.585 | 2.576 | .000 | | Intercept | 184747.035 | 1 | 184747.035 | 11725.020 | .000 | | Q1\_Gender | 49.602 | 1 | 49.602 | 3.148 | .076 | | Q2\_Age | 135.797 | 4 | 33.949 | 2.155 | .072 | | Q25\_Damage\_because\_of\_cards | 2.253 | 1 | 2.253 | .143 | .705 | | Q1\_Gender \* Q2\_Age | 72.373 | 4 | 18.093 | 1.148 | .332 | | Q1\_Gender \* Q25\_Damage\_because\_of\_cards | 63.499 | 1 | 63.499 | 4.030 | .045 | | Q2\_Age \* Q25\_Damage\_because\_of\_cards | 113.819 | 4 | 28.455 | 1.806 | .125 | | Q1\_Gender \* Q2\_Age \* Q25\_Damage\_because\_of\_cards | 70.105 | 4 | 17.526 | 1.112 | .349 | | Error | 43677.432 | 2772 | 15.757 |  |  | | Total | 1535295.000 | 2792 |  |  |  | | Corrected Total | 44448.550 | 2791 |  |  |  | | a. R Squared = .017 (Adjusted R Squared = .011) | | | | | |   1)Main Effect of Gender   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **1. 성별** | | | | | | Dependent Variable: privacy | | | | | | 성별 | Mean | Std. Error | 95% Confidence Inteval | | | Lower Bound | Upper Bound | | male | 23.220 | .295 | 22.642 | 23.798 | | female | 22.471 | .302 | 21.880 | 23.063 |   :  :  F(1, 2772) = 3.148, sig=0.076(>0.05)  This test is non-significant    The test failed to reveal a main effect of Gender  2)Main Effect of Age   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **2. 나이** | | | | | | Dependent Variable: privacy | | | | | | 나이 | Mean | Std. Error | 95% Confidence Interval | | | Lower Bound | Upper Bound | | teenager | 21.840 | .472 | 20.914 | 22.765 | | twenty | 23.244 | .190 | 22.872 | 23.617 | | thirty | 23.099 | .358 | 22.398 | 23.801 | | fourty | 23.443 | .603 | 22.260 | 24.625 | | more than fifty | 22.602 | .602 | 21.422 | 23.783 |   : All the means are same.  : At least two means are different.  F(4, 2772) = 2.155, sig=0.072(>0.05)  This test is non-significant.  Accept  The test failed to revealed a main effect of Age  3)Main Effect of Experience Data Leakage   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **3. 멤버십 카드로 인해 피해를 받은 적이 있습니까?** | | | | | | Dependent Variable: privacy | | | | | | 멤버십 카드로 인해 피해를 받은 적이 있습니까? | Mean | Std. Error | 95% Confidence Interval | | | Lower Bound | Upper Bound | | yes | 22.766 | .403 | 21.977 | 23.555 | | no | 22.926 | .126 | 22.678 | 23.173 |   :  :  F(1, 2772) = 0.143, sig=0.705(>0.05)  This test is non-significant  Accept  The test failed to reveal a main effect of Experience Data Leakage  4)Interaction Effect of Gender and Age   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4. 성별 \* 나이** | | | | | | | Dependent Variable: privacy | | | | | | | 성별 | 나이 | Mean | Std. Error | 95% Confidence Interval | | | Lower Bound | Upper Bound | | male | teenager | 22.010 | .536 | 20.960 | 23.061 | | twenty | 23.654 | .246 | 23.171 | 24.137 | | thirty | 23.413 | .480 | 22.472 | 24.355 | | forty | 23.051 | .924 | 21.238 | 24.864 | | more than fifty | 23.972 | .860 | 22.284 | 25.659 | | female | teenager | 21.669 | .777 | 20.145 | 23.193 | | twenty | 22.835 | .290 | 22.267 | 23.403 | | thirty | 22.786 | .530 | 21.746 | 23.826 | | forty | 23.834 | .775 | 22.316 | 25.353 | | more than fifty | 21.233 | .842 | 19.582 | 22.884 |   : All the interaction means are same.  : At least two interaction means are different.  F(4, 2772) = 1.148, sig=0.332(>0.05)  This test is non-significant  Accept  The test failed to reveal an interaction effect of Gender an Age.    There is an interaction effect of Age and Gender  5)Interaction Effect of Gender and Experience Data Leakage   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **5. 성별 \* 멤버십 카드로 인해 피해를 받은 적이 있습니까?** | | | | | | | Dependent Variable: privacy | | | | | | | 성별 | 멤버십 카드로 인해 피해를 받은 적이 있습니까? | Mean | Std. Error | 95% Confidence Interval | | | Lower Bound | Upper Bound | | male | yes | 23.564 | .558 | 22.470 | 24.658 | | no | 22.876 | .191 | 22.501 | 23.251 | | female | yes | 21.968 | .581 | 20.830 | 23.107 | | no | 22.975 | .165 | 22.650 | 23.299 |   F(1, 2772) = 4.030, sig=0.045(<0.05)  This test is significant  Reject  The test revealed an interaction effect of Gender and Experience Data Leakage.    There is an interaction effect of Age and Experience Data Leakage  6)Interaction Effect of Age and Experience Data Leakage   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **6. 나이 \* 멤버십 카드로 인해 피해를 받은 적이 있습니까?** | | | | | | | Dependent Variable: privacy | | | | | | | 나이 | 멤버십 카드로 인해 피해를 받은 적이 있습니까? | Mean | Std. Error | 95% Confidence Interval | | | Lower Bound | Upper Bound | | teenager | yes | 21.705 | .899 | 19.942 | 23.469 | | no | 21.974 | .286 | 21.412 | 22.536 | | twenty | yes | 23.344 | .369 | 22.621 | 24.067 | | no | 23.145 | .093 | 22.963 | 23.328 | | thirty | yes | 22.519 | .666 | 21.213 | 23.824 | | no | 23.680 | .262 | 23.166 | 24.195 | | forty | yes | 24.429 | 1.162 | 22.150 | 26.707 | | no | 22.457 | .322 | 21.824 | 23.089 | | more than fifty | yes | 21.833 | 1.146 | 19.586 | 24.080 | | no | 23.372 | .369 | 22.648 | 24.096 |   : All the interaction means are same.  : At least two interaction means are different.  F(4, 2772) = 1.806, sig=0.125(>0.05)  This test is non-significant  Accept  The test failed to reveal an interaction effect of Age and Experience Data Leakage.    There is an interaction effect of Age and Experience Data Leakage.  7)Interaction Effect of Gender, Age and Experience Data Leakage   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **7. 성별 \* 나이 \* 멤버십 카드로 인해 피해를 받은 적이 있습니까?** | | | | | | | | Dependent Variable: privacy | | | | | | | | 성별 | 나이 | 멤버십 카드로 인해 피해를 받은 적이 있습니까? | Mean | Std. Error | 95% Confidence Interval | | | Lower Bound | Upper Bound | | male | teenager | yes | 22.125 | .992 | 20.179 | 24.071 | | no | 21.896 | .405 | 21.101 | 22.690 | | twenty | yes | 23.928 | .478 | 22.991 | 24.865 | | no | 23.381 | .119 | 23.147 | 23.615 | | thirty | yes | 23.100 | .888 | 21.360 | 24.840 | | no | 23.726 | .367 | 23.007 | 24.446 | | forty | yes | 24.000 | 1.775 | 20.519 | 27.481 | | no | 22.102 | .517 | 21.088 | 23.115 | | more than fifty | yes | 24.667 | 1.621 | 21.489 | 27.844 | | no | 23.277 | .579 | 22.141 | 24.412 | | female | teenager | yes | 21.286 | 1.500 | 18.344 | 24.228 | | no | 22.052 | .405 | 21.258 | 22.846 | | twenty | yes | 22.760 | .561 | 21.659 | 23.861 | | no | 22.910 | .143 | 22.630 | 23.189 | | thirty | yes | 21.937 | .992 | 19.992 | 23.883 | | no | 23.634 | .375 | 22.898 | 24.369 | | forty | yes | 24.857 | 1.500 | 21.915 | 27.799 | | no | 22.811 | .386 | 22.055 | 23.567 | | more than fifty | yes | 19.000 | 1.621 | 15.822 | 22.178 | | no | 23.467 | .458 | 22.568 | 24.365 |   : All the interaction means are same.  : At least two interaction means are different.  F(4, 2772) = 1.112, sig=0.349(>0.05)  This test is non-significant  Accept  The test failed to reveal an interaction effect of Age, Gender and Experience Data Leakage. |
| Conclusion(s):  All the interaction means are about same. |

## **Analysis Outcome**

|  |
| --- |
| Outcome |
| Visual Representation and Outcome Descriptions:  Many people think that privacy data is an important to them. So, we can easily suppose that who have experienced privacy data leakage are more concern about their privacy.      Above graphs show that people who have experienced the data leakage and not how much think about their privacy. Y-axis is a score of privacy. When the score is low, it means that privacy is not important to them. And high means important. In the first graph, the blue is the people who experienced data leakage. But the graph length is about same to green. It means the experience of data leakage does not big influence to thinking of privacy.    And this graph shows that the gender also does not affect thinking of privacy. The blue(male) and green(female) are about same each other. In conclusion, people’s idea of concerning about their privacy data does not change for experience of data leakage and gender. Maybe another important thing is existed somewhere. |

# **Analysis 2**

## **Test A**

|  |
| --- |
| Analysis Statement: To study the utilization concerns in loyalty program between male and female. |
| Assumption Test(s):  1) Normality test for each gender.  Our sample is fairly large, outliers are a much more pressing concerns than normality.      2) Homogeneity of Variance   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Independent Samples Test | | | | | | | | | | | |  | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | | | Lower | Upper | | utilization | Equal variances assumed | 18.414 | .000 | -12.613 | 2807 | .000 | -1.887 | .150 | -2.181 | -1.594 | | Equal variances not assumed |  |  | -12.422 | 2496.113 | .000 | -1.887 | .152 | -2.185 | -1.589 |   Leven’s Test  : Homogeneity of variance can be assumed  : Homogeneity of variance cannot be assumed  F(2807) = 18.414 , sig = 0.000 (< 0.05)  This is significant  Reject  Conclusion : Homogeneity variance cannot be assumed  So, We do T-Test, when Equal Variance not Assumed |
| Analysis Results:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Group Statistics | | | | | | |  | 성별 | N | Mean | Std. Deviation | Std. Eror Mean | | utilization | male | 1550 | 12.30 | 3.670 | .093 | | female | 1259 | 14.19 | 4.258 | .120 |   Independent T-Test   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Independent Samples Test | | | | | | | | | | | |  | | Levene's Test for Equality of Variances | | t-test for Equaliyof Means | | | | | | | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | | | Lower | Upper | | utilization | Equal variances assumed | 18.414 | .000 | -12.613 | 2807 | .000 | -1.887 | .150 | -2.181 | -1.594 | | Equal variances not assumed |  |  | -12.422 | 2496.113 | .000 | -1.887 | .152 | -2.185 | -1.589 |   :  :  Since Sig = 0.000(< 0.05)  This test is significant  Reject  Conclusion : The different between two means is significant / |
| Conclusion(s):  Female (M=14.19, SE=4.258) use loyalty program more effectively than Male (M=12.30, SE=3.670). |

## **Test B**

|  |
| --- |
| Analysis Statement: To study the utilization concerns in loyalty program according to sign purpose. |
| Assumption Test(s):    1) Normality test for each gender.  Our sample is fairly large, outliers are a much more pressing concerns than normality.      2) Homogeneity of variance   |  |  |  |  | | --- | --- | --- | --- | | **Test of Homogeneity of Variances** | | | | | utilization | | | | | Levene Statistic | df1 | df2 | Sig. | | 1.294 | 3 | 2795 | .275 |   : Homogeneity of variance can be assumed  : Homogeneity of variance cannot be assumed  F(3,2795) = 1.294, sig = 0.275(> 0.05)  This is Non-Significant  Accept  Conclusion : Homogeneity of Variance can be Assumed |
| Analysis Results:  One-Way ANOVA   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **ANOVA** | | | | | | | utilization | | | | | | |  | Sum of Squares | df | Mean Square | F | Sig. | | Between Groups | 391.632 | 3 | 130.544 | 8.005 | .000 | | Within Groups | 45582.663 | 2795 | 16.309 |  |  | | Total | 45974.294 | 2798 |  |  |  |   : All the means are the same  : At least two means are different  F(3,2795) = 8.005, sig = 0.000(< 0.05)  This test is significant  Reject  Conclusion : At least two means are different  Plane Contrast Test   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Contrast Coefficients** | | | | | | Contrast | 멤버십 카드를 가입한 주요 이유는 무엇입니까? | | | | | culture | food | transportation | travel | | 1 | 1 | 1 | 1 | -3 | | 2 | 1 | 1 | -2 | 0 | | 3 | 1 | -1 | 0 | 0 |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | | **Contrast Tests** | | | | | | | | |  |  | Contrast | Value of Contrast | Std. Error | t | df | Sig. (2-tailed) | | utilization | Assume equal variances | 1 | -.20 | 1.387 | -.147 | 2795 | .883 | | 2 | -1.02 | .797 | -1.279 | 2795 | .201 | | 3 | -.75 | .159 | -4.730 | 2795 | .000 | | Does not assume equal variances | 1 | -.20 | 1.478 | -.138 | 101.095 | .890 | | 2 | -1.02 | .880 | -1.158 | 113.118 | .249 | | 3 | -.75 | .157 | -4.768 | 2561.601 | .000 |   Contrast 3:  Since sig = 0.000(<0.05)  This test is significant  Reject  Conclusion :  Contrast 2:  Since sig = 0.201(>0.05)  This test is non-significant  Accept  Conclusion :  Contrast 1:  Since sig = 0.883(>0.05)  This test is non-significant  Accept  Conclusion : |
| Conclusion(s):  This Test Show that when sign purpose is culture, Worst Satisfaction (M = 12.73, SE = 0.114). On the other hand, food(M = 13.48, SE = 0.108), transportation(M = 13.62, SE = 0.108) and travel(M = 13.35, SE = 0.468) are about same means. |

## **Test C**

|  |
| --- |
| Analysis Statement:2 |
| Assumption Test(s): No Assumption Test |
| Analysis Results:   |  |  |  |  | | --- | --- | --- | --- | | **Levene's Test of Equality of Error Variancesa** | | | | | Dependent Variable: utilization | | | | | F | df1 | df2 | Sig. | | 3.478 | 7 | 2791 | .001 | | Tests the null hypothesis that the error variance of the dependent variable is equal across groups. | | | | |  | | | |   a. Design: Intercept + Q1\_Gender + Q8\_Kinds\_of\_membership + Q1\_Gender \* Q8\_Kinds\_of\_membership  : Homogeneity of variance can be assumed  : Homogeneity of variance cannot be assumed  F(7,2791) = 3.478, sig=0.001(<0.05)  This test is significant.  Reject  Conclusion :   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Tests of Between-Subjects Effects** | | | | | | | Dependent Variable: utilization | | | | | | | Source | Type III Sum of Squares | df | Mean Square | F | Sig. | | Corected Model | 2784.805a | 7 | 397.829 | 25.709 | .000 | | Intercept | 114752.735 | 1 | 114752.735 | 7415.575 | .000 | | Q1\_Gender | 981.743 | 1 | 981.743 | 63.442 | .000 | | Q8\_Kinds\_of\_membership | 287.012 | 3 | 95.671 | 6.182 | .000 | | Q1\_Gender \* Q8\_Kinds\_of\_membership | 126.726 | 3 | 42.242 | 2.730 | .042 | | Error | 43189.489 | 2791 | 15.475 |  |  | | Total | 530725.000 | 2799 |  |  |  | | Corrected Total | 45974.294 | 2798 |  |  |  | | a. R Squared = .061 (Adjusted R Squared = .058) | | | | | |   1)Main Effect of Type of Gender   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **1. 성별** | | | | | | Dependent Variable: utilization | | | | | | 성별 | Mean | Std. Error | 95% Confidence Interval | | | Lower Bound | Upper Bound | | male | 12.378 | .182 | 12.021 | 12.734 | | female | 14.901 | .259 | 14.392 | 15.409 |   :  :  F(1,2791) = 63.442, sig=0.000(<0.05)  This test is significant.  Reject  The test reveal a main effect of Gender  Main effect of Purpose   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **2. 멤버십 카드를 가입한 주요 이유는 무엇입니까?** | | | | | | Dependent Variable: utilization | | | | | | 멤버십 카드를 가입한 주요 이유는 무엇입니까? | Mean | Std. Error | 95% Confidence Interval | | | Lower Bound | Upper Bound | | culture | 12.952 | .116 | 12.724 | 13.180 | | food | 13.448 | .105 | 13.242 | 13.654 | | transportation | 14.344 | .412 | 13.537 | 15.151 | | ravel | 13.812 | .455 | 12.920 | 14.705 |   : All the means are same.  : At least two means are different.  F(3, 2791) = 2.730, sig=0.042(<0.05)  This test is significant.  Reject  The test revealed a main effect of Method  POST HOC   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Multiple Comparisons** | | | | | | | | Dependent Variable: utilization | | | | | | | | Games-Howell | | | | | | | | (I) 멤버십 카드를 가입한 주요 이유는 무엇입니까? | (J) 멤버십 카드를 가입한 주요 이유는 무엇입니까? | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | | | Lower Bound | Uper Bound | | culture | food | -.75\* | .157 | .000 | -1.15 | -.35 | | transportation | -.88 | .448 | .203 | -2.05 | .28 | | travel | -.61 | .482 | .583 | -1.87 | .65 | | food | culture | .75\* | .157 | .000 | .35 | 1.15 | | transportation | -.13 | .446 | .990 | -1.30 | 1.03 | | travel | .14 | .480 | .992 | -1.12 | 1.39 | | transportation | culture | .88 | .448 | .203 | -.28 | 2.05 | | food | .13 | .446 | .990 | -1.03 | 1.30 | | travel | .27 | .637 | .974 | -1.38 | 1.92 | | travel | culture | .61 | .482 | .583 | -.65 | 1.87 | | food | -.14 | .480 | .992 | -1.39 | 1.12 | | transportation | -.27 | .637 | .974 | -1.92 | 1.38 | | Based on observed means.  The error term is Mean Square(Error) = 16.309. | | | | | | | | \*. The mean difference is significant at the .05 level. | | | | | | |   Comparison 1:  Since sig=0.000(>0.05)  This test is significant      Comparison 2:    Since sig=0.203(>0.05)  This test is non-significant      Comparison 3:    Since sig=0.583(>0.05)  This test is non-significant      Total conclusion:  3)Interaction Effect of Gender and Sign Purpose   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **3. 성별 \* 멤버십 카드를 가입한 주요 이유는 무엇입니까?** | | | | | | | Dependent Variable: utilization | | | | | | | 성별 | 멤버십 카드를 가입한 주요 이유는 무엇입니까? | Mean | Std. Error | 95% Confidence Iterval | | | Lower Bound | Upper Bound | | male | culture | 12.003 | .144 | 11.720 | 12.286 | | food | 12.651 | .152 | 12.353 | 12.949 | | transportation | 12.446 | .457 | 11.549 | 13.343 | | travel | 12.411 | .526 | 11.380 | 13.441 | | female | culture | 13.901 | .183 | 13.543 | 14.259 | | food | 14.245 | .145 | 13.960 | 14.530 | | transportation | 16.242 | .685 | 14.900 | 17.585 | | travel | 15.214 | .743 | 13.757 | 16.672 |   : All the interaction means are same.  : At least two interaction means are different.  F(3,2791) = 2.730, sig=0.042(<0.05)  This test is significant  Reject  The test revealed an interaction effect of Gender and Sign Purpose.    There is an interaction effect of Gender and Sign Purpose   |  |  |  |  | | --- | --- | --- | --- | | **Between-Subjects Factors** | | | | |  | | Value Label | N | | 성별 | 0 | male | 1543 | | 1 | female | 1256 | | 멤버십 카드를 가입한 주요 이유는 무엇입니까? | 0 | culture | 1207 | | 1 | food | 1401 | | 2 | transportation | 107 | | 3 | travel | 84 |  |  |  | | --- | --- | | Group | Interaction | | 1 | Gender = 0, Purpose = 0 | | 2 | Gender = 0, Purpose = 1 | | 3 | Gender = 0, Purpose = 2 | | 4 | Gender = 0, Purpose = 3 | | 5 | Gender = 1, Purpose = 0 | | 6 | Gender = 1, Purpose = 1 | | 7 | Gender = 1, Purpose = 2 | | 8 | Gender = 1, Purpose = 3 |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Multiple Comparisons** | | | | | | | | Dependent Variable: utilization | | | | | | | | Games-Howell | | | | | | | | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | | | Lower Bound | Upper Bound | | 1 | 2 | -.64\* | .195 | .024 | -1.23 | -.05 | | 3 | -.44 | .460 | .978 | -1.87 | .99 | | 4 | -.41 | .516 | .993 | -2.03 | 1.21 | | 5 | -1.88\* | .239 | .000 | -2.61 | -1.15 | | 6 | -2.24\* | .202 | .000 | -2.85 | -1.63 | | 7 | -4.24\* | .849 | .000 | -6.98 | -1.50 | | 8 | -3.21\* | .910 | .027 | -6.19 | -.24 | | 2 | 1 | .64\* | .195 | .024 | .05 | 1.23 | | 3 | .19 | .464 | 1.000 | -1.25 | 1.64 | | 4 | .23 | .520 | 1.000 | -1.40 | 1.86 | | 5 | -1.24\* | .248 | .000 | -2.00 | -.49 | | 6 | -1.60\* | .211 | .000 | -2.25 | -.96 | | 7 | -3.60\* | .852 | .004 | -6.35 | -.85 | | 8 | -2.57 | .912 | .129 | -5.55 | .41 | | 3 | 1 | .44 | .460 | .978 | -.99 | 1.87 | | 2 | -.19 | .464 | 1.000 | -1.64 | 1.25 | | 4 | .04 | .666 | 1.000 | -2.02 | 2.09 | | 5 | -1.44 | .484 | .070 | -2.94 | .06 | | 6 | -1.80\* | .467 | .005 | -3.25 | -.35 | | 7 | -3.80\* | .948 | .005 | -6.79 | -.80 | | 8 | -2.77 | 1.002 | .134 | -5.97 | .43 | | 4 | 1 | .41 | .516 | .993 | -1.21 | 2.03 | | 2 | -.23 | .520 | 1.000 | -1.86 | 1.40 | | 3 | -.04 | .666 | 1.000 | -2.09 | 2.02 | | 5 | -1.47 | .538 | .128 | -3.15 | .21 | | 6 | -1.83\* | .522 | .017 | -3.47 | -.20 | | 7 | -3.83\* | .977 | .006 | -6.91 | -.75 | | 8 | -2.80 | 1.029 | .143 | -6.08 | .47 | | 5 | 1 | 1.88\* | .239 | .000 | 1.15 | 2.61 | | 2 | 1.24\* | .248 | .000 | .49 | 2.00 | | 3 | 1.44 | .484 | .070 | -.06 | 2.94 | | 4 | 1.47 | .538 | .128 | -.21 | 3.15 | | 6 | -.36 | .253 | .844 | -1.13 | .41 | | 7 | -2.36 | .863 | .146 | -5.13 | .42 | | 8 | -1.33 | .922 | .830 | -4.33 | 1.67 | | 6 | 1 | 2.24\* | .202 | .000 | 1.63 | 2.85 | | 2 | 1.60\* | .211 | .000 | .96 | 2.25 | | 3 | 1.80\* | .467 | .005 | .35 | 3.25 | | 4 | 1.83\* | .522 | .017 | .20 | 3.47 | | 5 | .36 | .253 | .844 | -.41 | 1.13 | | 7 | -2.00 | .853 | .302 | -4.75 | .76 | | 8 | -.97 | .913 | .960 | -3.95 | 2.01 | | 7 | 1 | 4.24\* | .849 | .000 | 1.50 | 6.98 | | 2 | 3.60\* | .852 | .004 | .85 | 6.35 | | 3 | 3.80\* | .948 | .005 | .80 | 6.79 | | 4 | 3.83\* | .977 | .006 | .75 | 6.91 | | 5 | 2.36 | .863 | .146 | -.42 | 5.13 | | 6 | 2.00 | .853 | .302 | -.76 | 4.75 | | 8 | 1.03 | 1.231 | .990 | -2.84 | 4.90 | | 8 | 1 | 3.21\* | .910 | .027 | .24 | 6.19 | | 2 | 2.57 | .912 | .129 | -.41 | 5.55 | | 3 | 2.77 | 1.002 | .134 | -.43 | 5.97 | | 4 | 2.80 | 1.029 | .143 | -.47 | 6.08 | | 5 | 1.33 | .922 | .830 | -1.67 | 4.33 | | 6 | .97 | .913 | .960 | -2.01 | 3.95 | | 7 | -1.03 | 1.231 | .990 | -4.90 | 2.84 | | Based on observed means.  The error term is Mean Square(Error) = 15.502. | | | | | | | | \*. The mean difference is significant at the 0.05 level. | | | | | | |   Comparison 1(group 1 & 2):    Since sig=0.024(<0.05)  This test is significant    Comparison 2(group 3 & 4):    Since sig=1.000(>0.05)  This test is non-significant    Comparison 3(group 5 & 6):    Since sig=0.844(>0.05)  This test is non-significant    Comparison 4(group 7 & 8):    Since sig=0.990(>0.05)  This test is non-significant |
| Conclusion(s):    = = |

## **Analysis Outcome**

|  |
| --- |
| Outcome |
| Visual Representation and Outcome Descriptions:    This analysis is about how people use their membership program effectively for each gender and sign purpose of the membership program. Commonly, Female is more detailed than male. And this trend is showed also our analysis.  As you can see above graph, female’s stick is more high than male’s. This stick is the utilization score. High score means that they utilize well and low score means they are not good at utilizing membership program. So, in our graph female is good at utilizing their membership program.    In this graph, there are many groups. Groups are divided by gender and sign purpose of membership. Culture (book, movie theater), Food (restaurant, cafe), Transportation (subway, bus), Travel (hotel).  So, each group means  Group1 = male, choose culture  Group2 = male, choose food  Group3 = male, choose transportation  Group4 = male, choose travel  Group5 = female, choose culture  Group6 = female, choose food  Group7 = female, choose transportation  Group8 = female, choose travel  As you can see, male’s stick is shorter than female and the longest stick is female choosing transportation. Also all female graph is longer than male. It means female is good at utilizing their membership program and transportation is best purpose of membership program. I think that this survey is done by young person like students, so the main purpose of membership program is seem to be a transportation. |

# **Analysis 3**

## **Test A:**

|  |
| --- |
| Analysis Statement: to study the awareness concerns in loyalty program between male and female. |
| Assumption Test(s):  1) Normality test for each gender.  Our sample is fairly large, outliers are a much more pressing concerns than normality.  The lowest score is 17 and the highest is 40.  2) Homogeneity of variance test   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Independent Samples Test** | | | | | | | | | | | |  | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | | | F | Sig. | T | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | | | Lower | Upper | | awareness | Equal variances assumed | 15.248 | .000 | -1.169 | 2609 | .242 | -.222 | .190 | -.594 | .150 | | Equal variances not assumed |  |  | -1.185 | 2578.708 | .236 | -.222 | .187 | -.590 | .145 |   : Homogeneity of variance can be assumed.  : Homogeneity of variance cannot be assumed.  F(2609)=15.248, sig = 0.000(<0.05)  This test is significant.  Reject  Conclusion: Homogeneity of variance cannot be assumed. |
| Analysis Results:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Group Statistics** | | | | | | |  | Gender | N | Mean | Std. Deviation | Std. Error Mean | | awareness | male | 1452 | 23.13 | 5.061 | .133 | | female | 1159 | 23.35 | 4.501 | .132 |   Independent T-Test   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Independent Samples Test** | | | | | | | | | | | |  | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | | | Lower | Upper | | awareness | Equal variances assumed | 15.248 | .000 | -1.169 | 2609 | .242 | -.222 | .190 | -.594 | .150 | | Equal variances not assumed |  |  | -1.185 | 2578.708 | .236 | -.222 | .187 | -.590 | .145 |   :  :  Since sig = 0.236(>0.05)  This test is non-significant.  Accept  Conclusion: The difference between two means is non-significant. |
| Conclusion(s):  The awareness between Male (M=23.13, SE=5.061) and Female (M=23.35, SE=4.501) is about the same. |

## **Test B**

|  |
| --- |
| Analysis Statement: To study the awareness concerns in loyalty program according to age |
| Assumption Test(s):  1) Normality test for each age group.  Our sample is fairly large, outliers are a much more pressing concerns than normality.    There is no outliers.  2) Homogeneity of variance test   |  |  |  |  | | --- | --- | --- | --- | | **Test of Homogeneity of Variances** | | | | | awareness | | | | | Levene Statistic | df1 | df2 | Sig. | | 4.727 | 4 | 2577 | .001 |   : Homogeneity of variance can be assumed.  : Homogeneity of variance cannot be assumed.  F(4,2577)=4.727, sig = 0.001(<0.05)  This test is significant.  Reject  Conclusion: Homogeneity of variance cannot be assumed. |
| Analysis Results:  Welch Test   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Robust Tests of Equality of Means** | | | | | | awareness | | | | | |  | Statistica | df1 | df2 | Sig. | | Welch | 18.692 | 4 | 379.563 | .000 | | Brown-Forsythe | 16.834 | 4 | 656.890 | .000 | | a. Asymptotically F distributed. | | | | |   : All the means are the same..  : At least two means are different.  F(4,379.563)=18.692, sig = 0.000(<0.05)  This test is significant.  Reject  Conclusion: At least two means are different.  Post-Hoc test   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Multiple Comparisons** | | | | | | | | Dependent Variable: awareness | | | | | | | | Games-Howell | | | | | | | | (I) 나이 | (J) 나이 | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | | | Lower Bound | Upper Bound | | teenager | twenty | 1.945\* | .340 | .000 | 1.01 | 2.88 | | thirty | .181 | .412 | .992 | -.95 | 1.31 | | fourty | .671 | .493 | .653 | -.68 | 2.02 | | more than fifty | .568 | .588 | .870 | -1.05 | 2.19 | | twenty | teenager | -1.945\* | .340 | .000 | -2.88 | -1.01 | | thirty | -1.765\* | .275 | .000 | -2.52 | -1.01 | | fourty | -1.275\* | .385 | .010 | -2.34 | -.21 | | more than fifty | -1.377 | .502 | .053 | -2.77 | .01 | | thirty | teenager | -.181 | .412 | .992 | -1.31 | .95 | | twenty | 1.765\* | .275 | .000 | 1.01 | 2.52 | | fourty | .490 | .450 | .812 | -.74 | 1.72 | | more than fifty | .387 | .553 | .956 | -1.14 | 1.91 | | fourty | teenager | -.671 | .493 | .653 | -2.02 | .68 | | twenty | 1.275\* | .385 | .010 | .21 | 2.34 | | thirty | -.490 | .450 | .812 | -1.72 | .74 | | more than fifty | -.103 | .615 | 1.000 | -1.79 | 1.59 | | more than fifty | teenager | -.568 | .588 | .870 | -2.19 | 1.05 | | twenty | 1.377 | .502 | .053 | -.01 | 2.77 | | thirty | -.387 | .553 | .956 | -1.91 | 1.14 | | fourty | .103 | .615 | 1.000 | -1.59 | 1.79 | | \*. The mean difference is significant at the 0.05 level. | | | | | | |   Comparison: 1  =    Since sig=0.000(<0.05)  This test is significant.  Conclusion:  Comparison: 2  =    Since sig = 0.992(>0.05)  This test is significant.  Conclusion:  Comparison: 3  =    Since sig= 0.493(>0.05)  This test is non-significant  Conclusion:  Comparison: 4  =    Since sig= 0.588(>0.05)  This test is non-significant  Conclusion:   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Descriptives** | | | | | | | | | | awareness | | | | | | | | | |  | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum | | Lower Bound | Upper Bound | | teenager | 203 | 24.54 | 4.617 | .324 | 23.90 | 25.18 | 17 | 36 | | twenty | 1849 | 22.60 | 4.482 | .104 | 22.39 | 22.80 | 17 | 36 | | thirty | 252 | 24.36 | 4.032 | .254 | 23.86 | 24.86 | 17 | 36 | | Forty | 163 | 23.87 | 4.736 | .371 | 23.14 | 24.60 | 17 | 35 | | more than fifty | 115 | 23.97 | 5.262 | .491 | 23.00 | 24.95 | 17 | 36 | | Total | 2582 | 23.06 | 4.564 | .090 | 22.89 | 23.24 | 17 | 36 | |
| Conclusion(s): =  The awareness of loyalty program in teenager(M=24.54, SE=4.617) and thirty group(M=24.36, SE=4.032) is about the same, but significantly different from twenty(M=22.60, SE=4.482), forty(M=23.97, SE=4.736), fifty group(M=23.97, SE=5.262), F(4, 2577)= 1489.093, p < 0.05 |

## **Test C**

|  |
| --- |
| Analysis Statement: To study the awareness concerns in loyalty program between age and gender. |
| Assumption Test(s): No |
| Analysis Results:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Tests of Between-Subjects Effects** | | | | | | | Dependent Variable: awareness | | | | | | | Source | Type III Sum of Squares | df | Mean Square | F | Sig. | | Corrected Model | 1570.795a | 9 | 174.533 | 8.600 | .000 | | Intercept | 571674.856 | 1 | 571674.856 | 28168.235 | .000 | | Q2\_Age | 1427.281 | 4 | 356.820 | 17.582 | .000 | | Q1\_Gender | 1.326 | 1 | 1.326 | .065 | .798 | | Q2\_Age \* Q1\_Gender | 59.808 | 4 | 14.952 | .737 | .567 | | Error | 52198.788 | 2572 | 20.295 |  |  | | Total | 1427202.000 | 2582 |  |  |  | | Corrected Total | 53769.583 | 2581 |  |  |  | | a. R Squared = .029 (Adjusted R Squared = .026) | | | | | |   1)main effect of type of age   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **1. 나이** | | | | | | Dependent Variable: awareness | | | | | | 나이 | Mean | Std. Error | 95% Confidence Interval | | | Lower Bound | Upper Bound | | teenager | 24.550 | .316 | 23.929 | 25.170 | | twenty | 22.620 | .106 | 22.412 | 22.829 | | thirty | 24.370 | .284 | 23.813 | 24.927 | | fourty | 23.825 | .366 | 23.108 | 24.542 | | more than fifty | 24.059 | .425 | 23.226 | 24.892 |   : All the means are same  : At least two means are different  F(4,2572)=17.582, sig=0.000(<0.05)  This test is significant  Reject  The test revealed a main effect of method  Post-Hoc test is done in Test B  2)Main effect of gender   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **2. 성별** | | | | | | Dependent Variable: awareness | | | | | | 성별 | Mean | Std. Error | 95% Confidence Interval | | | Lower Bound | Upper Bound | | male | 23.848 | .212 | 23.433 | 24.263 | | female | 23.921 | .190 | 23.548 | 24.294 |   :  :  Since sig = 0.789(>0.05)  This test is non-significant.  Accept  Conclusion: The test failed to reveal a main effect of gender.  3)Interaction effect of age and gender   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **3. 나이 \* 성별** | | | | | | | Dependent Variable: awareness | | | | | | | 나이 | 성별 | Mean | Std. Error | 95% Confidence Interval | | | Lower Bound | Upper Bound | | teenager | male | 24.324 | .440 | 23.462 | 25.186 | | female | 24.776 | .455 | 23.883 | 25.668 | | twenty | male | 22.482 | .137 | 22.214 | 22.751 | | female | 22.758 | .163 | 22.439 | 23.078 | | Thirty | male | 24.153 | .394 | 23.381 | 24.924 | | female | 24.587 | .410 | 23.784 | 25.390 | | Fourty | male | 23.650 | .582 | 22.510 | 24.790 | | female | 24.000 | .444 | 23.130 | 24.870 | | more than fifty | male | 24.633 | .644 | 23.371 | 25.895 | | female | 23.485 | .555 | 22.397 | 24.572 |   : All the interaction means are the same.  : At least two interaction means are different.  F(4,2572)=0.737, sig=0.567(>0.05)  This test is non-significant  Accept  The test failed to reveal an interaction effect. |
| Conclusion(s):  As a result, all the interaction means are the same. |

## **Analysis Outcome**

|  |
| --- |
| Outcome |
| Visual Representation and Outcome Descriptions:  Do you think that older people know more thing than other? They memorize everything they see? Commonly, older people know more than other. Because they lived more. This idea is also correct in membership program? I to put forth the conclusion first, it is wrong.    In this graph, Y-axis is the score that how much they know about membership program’s operation system such as, retention period of privacy data. High score means that they know many and low means few. As you can see, each age group no big difference. And twenty group is slightly lower than other groups.  As a result, age does not big influence to awareness about membership program. |

# **Analysis Extra**

## **Test A**

|  |
| --- |
| Analysis Statement: To study the satisfaction concerns in loyalty program whether they have experienced data leakage. |
| Assumption Test(s):  1) Normality test    It’s an approximately normal because there is a symmetric line  2)Homogeneity of variance   |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Independent Samples Test** | | | | | | | | | | | |  | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | | | Lower | Upper | | satisfaction | Equal variances assumed | 12.262 | .000 | .882 | 2781 | .378 | .278 | .315 | -.340 | .896 | | Equal variances not assumed |  |  | .762 | 225.076 | .447 | .278 | .365 | -.441 | .997 |   Leven’s Test  : Homogeneity of variance can be assumed  : Homogeneity of variance cannot be assumed  F(2781) = 12.262 , sig = 0.000 (< 0.05)  This is significant  Reject  Conclusion : Homogeneity of variance cannot be assumed  So, We do T-Test, when Equal Variance not Assumed |
| Analysis Results:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Group Statistics** | | | | | | |  | 멤버십 카드로 인해 피해를 받은 적이 있습니까? | N | Mean | Std. Deviation | Std. Error Mean | | satisfaction | yes | 203 | 17.78 | 5.061 | .355 | | no | 2580 | 17.50 | 4.258 | .084 |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Independent Samples Test** | | | | | | | | | | | |  | | Levene's Test for Equality of Variances | | t-test for Eqality of Means | | | | | | | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | | | Lower | Upper | | satisfaction | Equal variances assumed | 12.262 | .000 | .882 | 2781 | .378 | .278 | .315 | -.340 | .896 | | Equal variances not assumed |  |  | .762 | 225.076 | .447 | .278 | .365 | -.441 | .997 |   Independent T-Test  :  :  Since Sig = 0.447 (> 0.05)  This test is non-significant  Accept  Conclusion : The different between two means is non-significant / |
| Conclusion(s):  The experience of data leakage doesn’t affect Satisfaction. |

## **Test B**

|  |
| --- |
| Analysis Statement: |
| Assumption Test(s):  1) the Independence of the independent variable and covariate   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Tests of Between-Subjects Effects** | | | | | | | Dependent Variable: awareness | | | | | | | Source | Type III Sum of Squares | df | Mean Square | F | Sig. | | Corrected Model | .041a | 1 | .041 | .002 | .967 | | Intercept | 1448819.630 | 1 | 1448819.630 | 60091.284 | .000 | | Q1\_Gender | .041 | 1 | .041 | .002 | .967 | | Error | 67123.110 | 2784 | 24.110 |  |  | | Total | 1530677.000 | 2786 |  |  |  | | Corrected Total | 67123.150 | 2785 |  |  |  | | a. R Squared = .000 (Adjusted R Squared = .000) | | | | | |   : The covariate was roughly equal across levels of the independent variable  : The covariate was roughly not equal across levels of the independent variable  F(1,2784) = 0.002, sig=0.967(>0.05)  This test is non-significant  Accept  Conclusion : It is appropriate to use Awareness as a covariate in the analysis  2) Regression Homogeneity   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Tests of Between-Subjects Effects** | | | | | | | Dependent Variable: satisfaction | | | | | | | Source | Type III Sum of Squares | df | Mean Square | F | Sig. | | Corrected Model | 1469.127a | 3 | 489.709 | 26.982 | .000 | | Intercept | 25093.859 | 1 | 25093.859 | 1382.605 | .000 | | Q1\_Gender | .047 | 1 | .047 | .003 | .959 | | awareness | 1046.981 | 1 | 1046.981 | 57.686 | .000 | | Q1\_Gender \* awareness | 13.069 | 1 | 13.069 | .720 | .396 | | Error | 50492.439 | 2782 | 18.150 |  |  | | Total | 906960.000 | 2786 |  |  |  | | Corrected Total | 51961.566 | 2785 |  |  |  | | a. R Squared = .028 (Adjusted R Squared = .027) | | | | | |   : Homogeneity of regression slopes can be assumed.  : Homogeneity of r cannot be assumed.  F(1, 2782) = 0.720, sig=0.396(>0.05)  This test is non-significant  Accept  Conclusion: Homogeneity of regression slopes can be assumed. |
| Analysis Results:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Tests of Between-Subjects Effects** | | | | | | | Dependent Variable: satisfaction | | | | | | | Source | Type III Sum of Squares | df | Mean Square | F | Sig. | | Corrected Model | 1456.058a | 2 | 728.029 | 40.116 | .000 | | Intercept | 25717.188 | 1 | 25717.188 | 1417.092 | .000 | | Q1\_Gender | 328.796 | 1 | 328.796 | 18.118 | .000 | | awareness | 1128.208 | 1 | 1128.208 | 62.168 | .000 | | Error | 50505.509 | 2783 | 18.148 |  |  | | Total | 906960.000 | 2786 |  |  |  | | Corrected Total | 51961.566 | 2785 |  |  |  | | a. R Squared = .028 (Adjusted R Squared = .027) | | | | | |   :  :  F(1, 2783) = 18.118, sig=0.000(<0.05)  This test is significant  Reject  Conclusion : :  Original Data   |  |  |  |  | | --- | --- | --- | --- | | **Descriptive Statistics** | | | | | Dependent Variable: satisfaction | | | | | 성별 | Mean | Std. Deviation | N | | male | 17.83 | 4.382 | 1533 | | female | 17.14 | 4.213 | 1253 | | Total | 17.52 | 4.319 | 2786 |   Adjusted Data   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Estimates** | | | | | | Dependent Variable: satisfaction | | | | | | 성별 | Mean | Std. Error | 95% Confidence Interval | | | Lower Bound | Upper Bound | | male | 17.829a | .109 | 17.616 | 18.042 | | female | 17.138a | .120 | 16.902 | 17.374 | | a. Covariates appearing in the model are evaluated at the following values: awareness = 22.92. | | | | | |
| Conclusion(s):  In ANCOVA test we use Awareness Score with Covariate, and we Adjust male’s mean higher than female’s mean. So, male is more satisfied than female |

## **Analysis Outcome**

|  |
| --- |
| Outcome |
| Visual Representation and Outcome Descriptions:  Commonly, we think that people who have experienced privacy data leakage are unsatisfied with the company. But oddly, this idea is wrong, according to our analysis.  In this graph, Y-axis is the score that how much they are satisfied with membership program. High score means that they are pleased with the membership, and low is not. And X-axis is whether they have experienced personal data leakage. As you can see, there is no difference between two scores, it means that people’s satisfaction does not influenced by the experience of data leakage. It sounds like strange, it is the result of our analysis.  Thank you. |

# **Final Project**

In this project, Fist, we experienced many difficult. First, In three way ANOVA, in first analyze, interaction effect is almost non-significant, So, we unintentionally solve problem easily. We intend 2 \* 5 \* 2 ANOVA but, it is almost non-significant, only one process is significant.

Second, we presume normal this project, but it is unfortunately No! some Scores is Normal, some Scores is not Normal, So we recall remember that the Professor said, “Big Data don’t have to check normality. Big data are stroked nerves only outlier” So, we check outlier and control, delete outlier. Each Scores.

Third, we assume people who have experience Data Leakage have less Satisfaction Score. but it is wrong assume!! The test is invalid without knowing that, assume is wrong!! We shocked this result. Received harm clearly… Satisfaction is same! I don’t understand this result.

Fourth, This test is concentrated young, male student. So, result is concentrated too. I try add accurate result. But it is difficult.

Finally, Thank you for this year to teach us. I stroked nerves only coding or computer science, but this lecture let know us how to administrate big data, and data base. This field is so so so promising field. As far as I know is. So I thank you for awake us in some field like this one.

I expect meeting again in lecture, we want to stay hold relationship familiarly.

**Thank you**