**COMPUTER LITERACY ASSIGNMENT 3**

**SSCM 1303**

**LECTURER NAME:**

**DR NORHAIZA BT AHMAD**

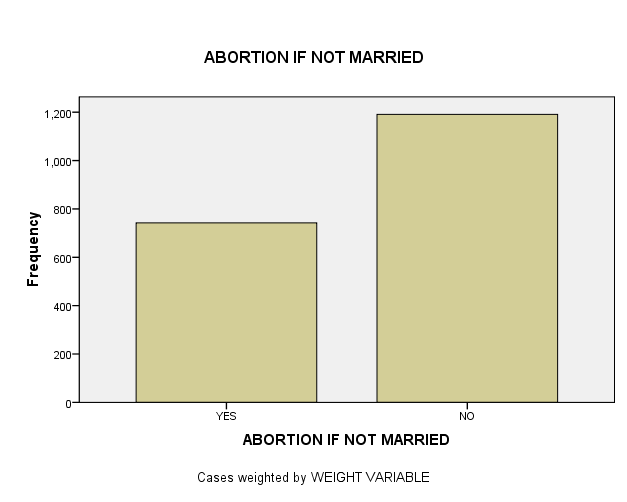
**GROUP MEMBER NAME:**

**SIA JIA YONG 911103-01-6097**

**OOI CHUN LIN 910131-08-5693**

**LAU MUN HING 910218-08-6023**

**Exercise 4 Question 1**



| **ABORTION IF NOT MARRIED** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  |  | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | YES | 742 | 16.4 | 38.4 | 38.4 |
| NO | 1191 | 26.4 | 61.6 | 100.0 |
| Total | 1933 | 42.8 | 100.0 |  |
| Missing | NAP | 2513 | 55.7 |  |  |
| DK | 53 | 1.2 |  |  |
| NA | 13 | .3 |  |  |
| Total | 2579 | 57.2 |  |  |
| Total | | 4512 | 100.0 |  |  |

We can see from the chart, the graphic is a bar chart with the categories at the bottom, the X-axis, and the frequency scale at the left, the Y-axis. The variable label ABORTION-IF NOT MARRIED is displayed at the top of the chart. We see from the frequency distribution that there are more “no” 26.4% answers than “yes”, 16.4% answers, when respondents were asked if a woman should able to get an abortion for not married. There are 55.7% is NAP(Not Appropriate), 1.2% is DK(Don’t know) and 0.3% is NA(Not Answered).

But for valid percent excluded missing value, we can get to know that 38.4% of respondent answered “yes” while 61.6% of respondent answered “no”.

**Exercise 6**

1.

| **Report** | | | |
| --- | --- | --- | --- |
| AGE OF RESPONDENT | | | |
| VOTE FOR KERRY, BUSH, NADER | Mean | N | Std. Deviation |
| KERRY | 47.58 | 1339 | 16.878 |
| BUSH | 48.93 | 1476 | 16.035 |
| OTHER | 46.95 | 41 | 12.371 |
| Total | 48.27 | 2856 | 16.400 |

The youngest mean age group is respondents who voted for other which is 46.95 whereas the oldest mean age group is respondents who voted for Bush which is 48.93.

2.

| **Group Statistics** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | RESPONDENT'S SEX | N | Mean | Std. Deviation | Std. Error Mean |
| TOTAL FAMILY INCOME | MALE | 1789 | 17.92 | 5.100 | .121 |
| FEMALE | 2016 | 16.83 | 5.540 | .123 |

| **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 | |
| TOTAL FAMILY INCOME | Equal variances assumed | 24.423 | .000 | 6.265 | 3804 | .000 | 1.086 | .173 | .746 | 1.426 |
| Equal variances not assumed |  |  | 6.296 | 3798.587 | .000 | 1.086 | .172 | .748 | 1.424 |

Man has a highest mean income. Yes, because the significance value is less than 0.0005 so for sure is more than 0.05.

3.

| **Group Statistics** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | BELIEF IN LIFE AFTER DEATH | N | Mean | Std. Deviation | Std. Error Mean |
| AGE OF RESPONDENT | YES | 2178 | 45.13 | 16.648 | .357 |
| NO | 453 | 46.40 | 17.390 | .817 |

| **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 |
| AGE OF RESPONDENT | Equal variances assumed | 5.497 | .019 | -1.471 | 2629 | .141 | -1.274 | .866 | -2.973 | .424 |
| Equal variances not assumed |  |  | -1.429 | 636.274 | .153 | -1.274 | .891 | -3.025 | .476 |

Respondents who don’t believe in life after death contribute highest mean age of respondent which is 46.40. No because the significance value more than 0.05.

4.

| **Paired Samples Statistics** | | | | | |
| --- | --- | --- | --- | --- | --- |
| Pair 1 |  | Mean | N | Std. Deviation | Std. Error Mean |
| R'S MOTHER S SOCIOECONOMIC INDEX | 44.774 | 1379 | 19.8009 | .5333 |
| R'S FATHER S SOCIOECONOMIC INDEX | 48.184 | 1379 | 19.1088 | .5146 |

| **Paired Samples Test** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Paired Differences | | | | | t | df | Sig.  (2-tailed) |
| Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | |
| Lower | Upper |
| Pair 1 | R'S MOTHER S SOCIOECONOMIC INDEX - R'S FATHER S SOCIOECONOMIC INDEX | -3.4098 | 20.9974 | .5655 | -4.5192 | -2.3005 | -6.030 | 1378 | .000 |

Respondent’s father’s socioeconomic index has the highest mean socioeconomic status. Yes because the significance value is lower than 0.0005 and for sure is lower than 0.05.

**EXERCISE 7**

1) The higher the age of the people, the higher the tvhours.

2)

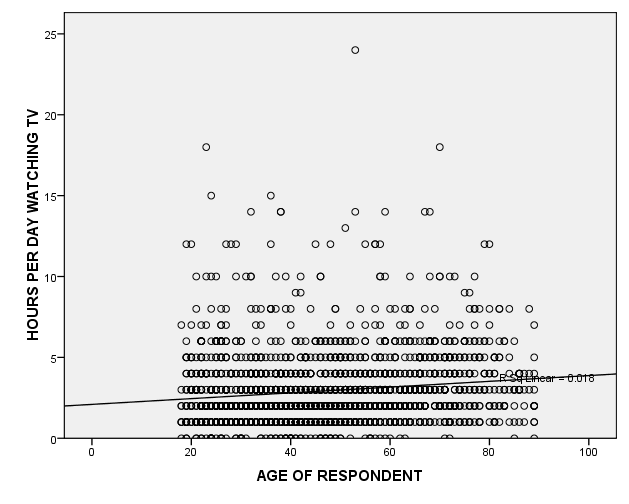
| **Correlations** | | | |
| --- | --- | --- | --- |
|  |  | AGE OF RESPONDENT | HOURS PER DAY WATCHING TV |
| AGE OF RESPONDENT | Pearson Correlation | 1 | .127\*\* |
| Sig. (2-tailed) |  | .000 |
| N | 4496 | 1982 |
| HOURS PER DAY WATCHING TV | Pearson Correlation | .127\*\* | 1 |
| Sig. (2-tailed) | .000 |  |
| N | 1982 | 1986 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | |  |

| **Correlations** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  | AGE OF RESPONDENT | HOURS PER DAY WATCHING TV |
| Spearman's rho | AGE OF RESPONDENT |  | Correlation Coefficient | 1.000 | .134\*\* |
|  | Sig. (2-tailed) | . | .000 |
|  | N | 4183 | 1845 |
| HOURS PER DAY WATCHING TV |  | Correlation Coefficient | .134\*\* | 1.000 |
|  | Sig. (2-tailed) | .000 | . |
|  | N | 1845 | 1847 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | |  |  |

Yes, my hypothesis was supported with the three things. No, the coefficient is not 0. My prediction of the hypothesized direction of the relationship between age of respondents and the hours per day watching tv is positive.Positive 0.134 correlation coefficients tell me there is a direct relationship between age and tvhours. When age of the respondents increases, the hours per day watching tv increases.

3) The strength of the relationship between age and tvhours is weak. The hours per day respondents have to relax might also influence the amount of television that people watch.

5)



| **Variables Entered/Removedb** | | | |
| --- | --- | --- | --- |
| Model | Variables Entered | Variables Removed | Method |
| 1 | AGE OF RESPONDENTa | . | Enter |
| a. All requested variables entered. | | |  |
| b. Dependent Variable: HOURS PER DAY WATCHING TV | | | |

| **Model Summary** | | | | |
| --- | --- | --- | --- | --- |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .127a | .016 | .016 | 2.131 |
| a. Predictors: (Constant), AGE OF RESPONDENT | | | | |

| **ANOVAb** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 148.212 | 1 | 148.212 | 32.634 | .000a |
| Residual | 8993.903 | 1980 | 4.542 |  |  |
| Total | 9142.115 | 1981 |  |  |  |
| a. Predictors: (Constant), AGE OF RESPONDENT | | | | |  |  |
| b. Dependent Variable: HOURS PER DAY WATCHING TV | | | | |  |  |

| **Coefficientsa** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | 2.047 | .137 |  | 14.917 | .000 |
| AGE OF RESPONDENT | .016 | .003 | .127 | 5.713 | .000 |
| a. Dependent Variable: HOURS PER DAY WATCHING TV | | | |  |  |  |

Y = a + bX

Y refers to the value of the hours per day watching tv, a is the Y-intercept (the point where the line crosses the Y-axis, b is the slope of the line which describes the relationship between the age of respondents and the hours per day watching tv, and X is the value of the age of respondents.

We know that the linear relationship between the age of respondents and the hours per day watching tvis not perfect. The correlation coefficient is 0.134, and the scatterplot showed a lot of cases that did not fall directly on the line. Thus, it is clear to us that knowing the age of respondents will not tell us the hours per day watching tv.We are only analysing a sample of cases and not the whole population to which we want to generalize our findings about the relationship between age of respondents and the hours per day watching tv. It is clear that there are some outliers in our findings.We draw a best fit line to show the relationship between the age of respondents and the hours per day watching hours. For these reasons, it is conventional to write the formula for the line as

Ŷ = a + bX + e, where e refers to error.

Since we do not know the actual value of errors that we make, we choose to ignore the e.

Ŷ = 2.047 + 0.016X

2.047= the Y-intercept is interpreted as the hours per day watching hours (our dependent, or Y variable),

0.016 = the slope of the linethe age of respondents (our independent, or X variable).

Since X refers to the value of the independent variable, and age of respondentsis our independent variable, all we have to do is enter 12 (age of respondents) into our equation as follows:

Ŷ = 2.047 + 0.016(12)

Ŷ = 2.239

We find that the age of respondents which is12 is associated with the hours per day watching tv is 2.239 hours. In sum, the error tells us about the distance from actual values of Y (the answers that the GSS survey respondents gave) and predicted values of Y (the ones you calculate based on the GSS respondent’s information in the “X” variable).

Error = the difference between a predicted value of Y for a given case and the actual value of Y for a given case (Ŷ-Y).