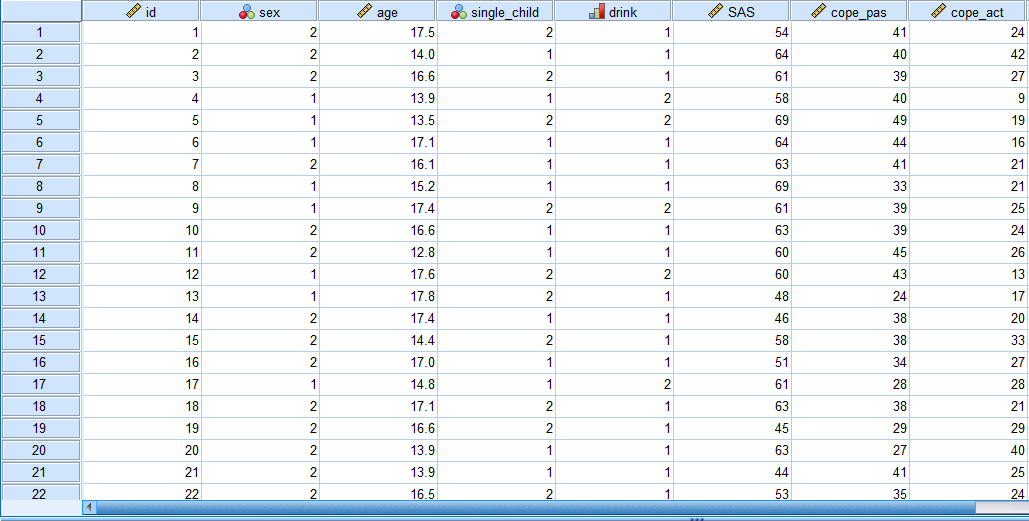
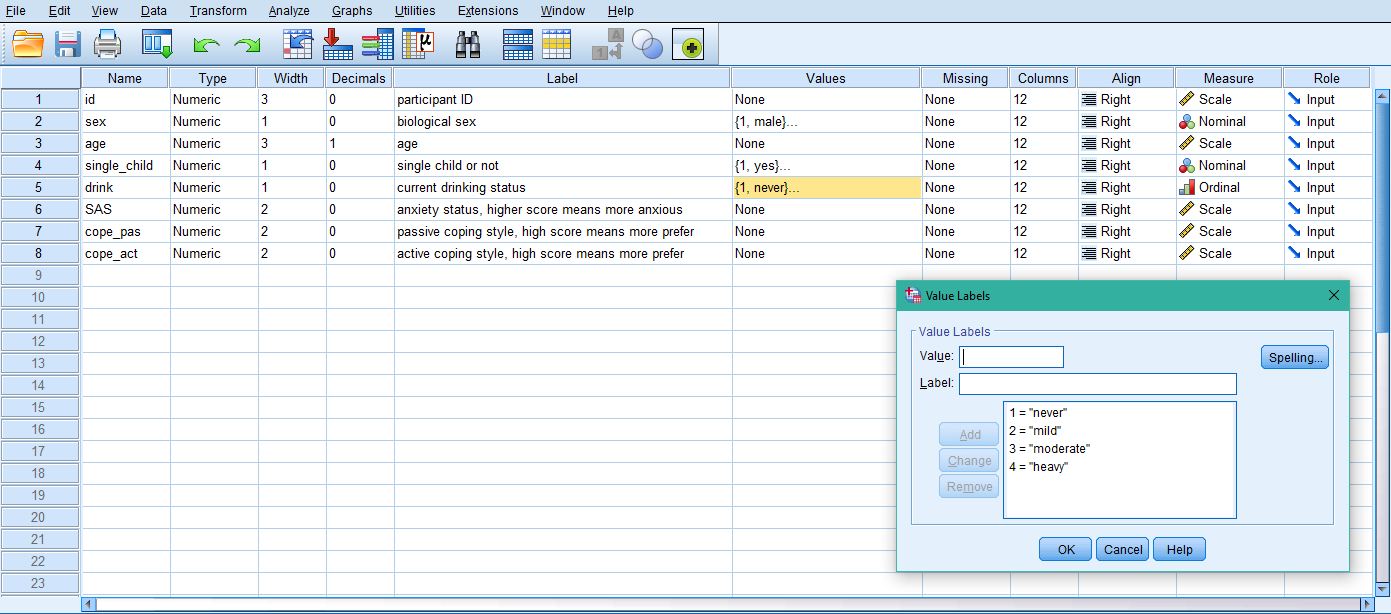
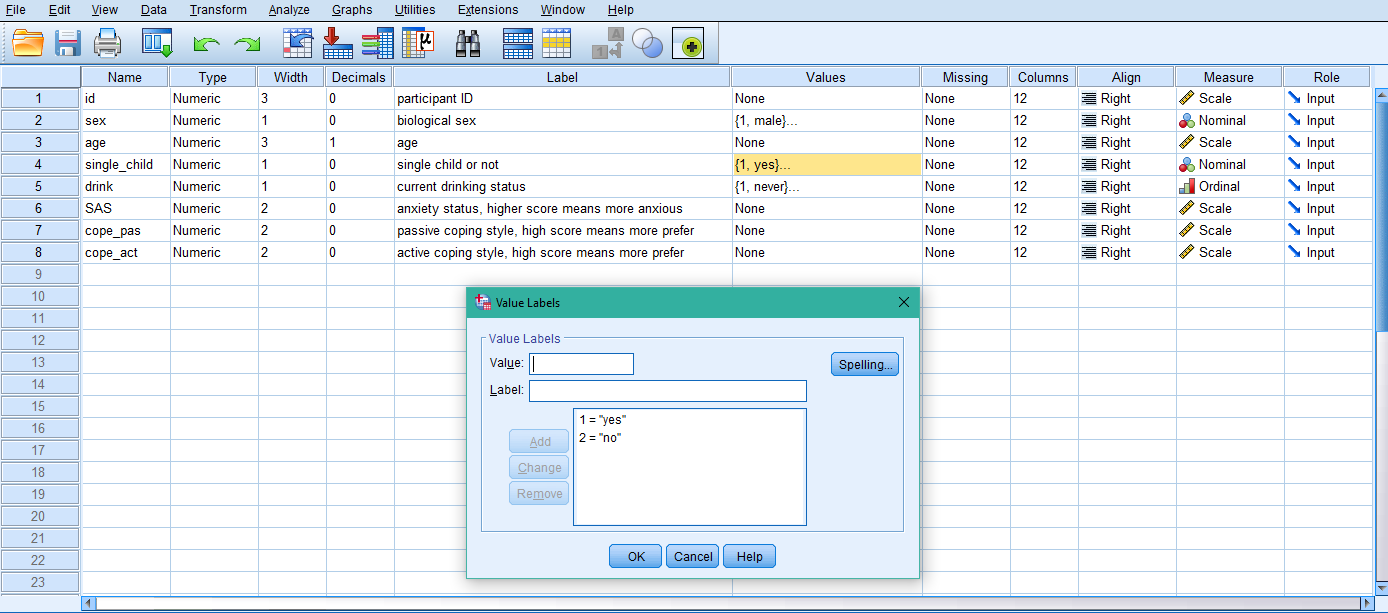
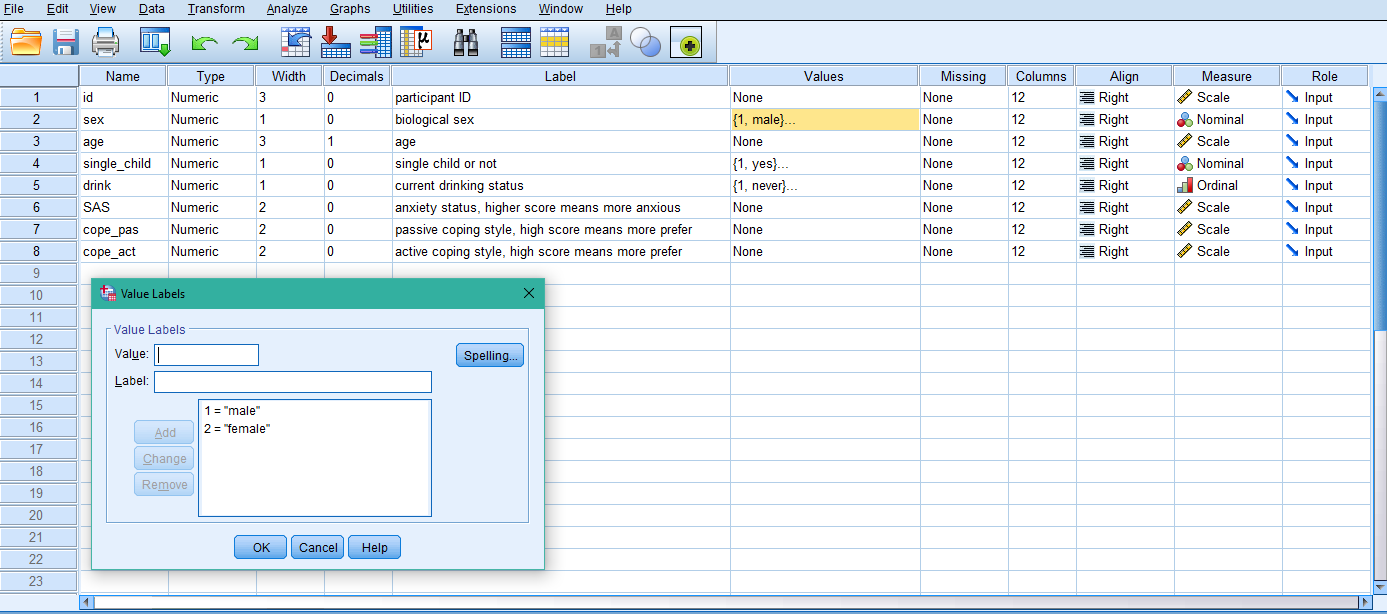
Justin Harding

450221916

**INFO2150 Case Study**

1. **SPSS**

Below are screenshots of the SPSS file created for this case study. The variable names, types, decimal places, width, variable labels and value labels match those outlined by the original excel file.



1. **Variable Types**

For our dataset there are many different data types. The first variable “ID” is simply a label variable. The two nominal variables are “sex” and “single\_child”, since they do not have any quantitative value. These are also classified as dichotomous variables, since there are only two distinct values. The ordinal variable in this set is “drink” since there is an order between values, however the difference between them is not known. Age is a continuous variable since it has an infinite number of possible values. SAS, cope\_pas and cope\_act are scale variables, since there is a meaningful metric and the distances can be calculated.

1. **Relationship between sex and anxiety of young people**

**Research Hypothesis:** There is association between sex and the anxiety status of young people.

**Null Hypothesis:** There is no association between sex and the anxiety status of young people.

The statistical test I will use when analyzing this data will be a chi-squared test, since I am testing whether two variables are related or not.

Below is the crosstabulation of SAS and sex, with the expected count.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **anxiety status, higher score means more anxious \* biological sex Crosstabulation** | | | | | |
|  | | | biological sex | | Total |
| male | female |
| anxiety status, higher score means more anxious | 25 | Count | 0 | 1 | 1 |
| Expected Count | .4 | .6 | 1.0 |
| 26 | Count | 1 | 3 | 4 |
| Expected Count | 1.8 | 2.2 | 4.0 |
| 28 | Count | 1 | 2 | 3 |
| Expected Count | 1.3 | 1.7 | 3.0 |
| 29 | Count | 5 | 5 | 10 |
| Expected Count | 4.4 | 5.6 | 10.0 |
| 30 | Count | 2 | 8 | 10 |
| Expected Count | 4.4 | 5.6 | 10.0 |
| 31 | Count | 3 | 2 | 5 |
| Expected Count | 2.2 | 2.8 | 5.0 |
| 33 | Count | 4 | 5 | 9 |
| Expected Count | 4.0 | 5.0 | 9.0 |
| 34 | Count | 9 | 10 | 19 |
| Expected Count | 8.4 | 10.6 | 19.0 |
| 35 | Count | 5 | 8 | 13 |
| Expected Count | 5.8 | 7.2 | 13.0 |
| 36 | Count | 14 | 16 | 30 |
| Expected Count | 13.3 | 16.7 | 30.0 |
| 38 | Count | 9 | 13 | 22 |
| Expected Count | 9.8 | 12.2 | 22.0 |
| 39 | Count | 13 | 8 | 21 |
| Expected Count | 9.3 | 11.7 | 21.0 |
| 40 | Count | 17 | 12 | 29 |
| Expected Count | 12.9 | 16.1 | 29.0 |
| 41 | Count | 8 | 16 | 24 |
| Expected Count | 10.6 | 13.4 | 24.0 |
| 43 | Count | 12 | 9 | 21 |
| Expected Count | 9.3 | 11.7 | 21.0 |
| 44 | Count | 26 | 25 | 51 |
| Expected Count | 22.6 | 28.4 | 51.0 |
| 45 | Count | 5 | 10 | 15 |
| Expected Count | 6.6 | 8.4 | 15.0 |
| 46 | Count | 3 | 5 | 8 |
| Expected Count | 3.5 | 4.5 | 8.0 |
| 48 | Count | 5 | 7 | 12 |
| Expected Count | 5.3 | 6.7 | 12.0 |
| 49 | Count | 4 | 2 | 6 |
| Expected Count | 2.7 | 3.3 | 6.0 |
| 50 | Count | 2 | 5 | 7 |
| Expected Count | 3.1 | 3.9 | 7.0 |
| 51 | Count | 2 | 10 | 12 |
| Expected Count | 5.3 | 6.7 | 12.0 |
| 53 | Count | 1 | 1 | 2 |
| Expected Count | .9 | 1.1 | 2.0 |
| 54 | Count | 1 | 4 | 5 |
| Expected Count | 2.2 | 2.8 | 5.0 |
| 55 | Count | 2 | 3 | 5 |
| Expected Count | 2.2 | 2.8 | 5.0 |
| 56 | Count | 2 | 4 | 6 |
| Expected Count | 2.7 | 3.3 | 6.0 |
| 58 | Count | 1 | 1 | 2 |
| Expected Count | .9 | 1.1 | 2.0 |
| 59 | Count | 1 | 2 | 3 |
| Expected Count | 1.3 | 1.7 | 3.0 |
| 60 | Count | 1 | 1 | 2 |
| Expected Count | .9 | 1.1 | 2.0 |
| 61 | Count | 2 | 3 | 5 |
| Expected Count | 2.2 | 2.8 | 5.0 |
| 63 | Count | 0 | 4 | 4 |
| Expected Count | 1.8 | 2.2 | 4.0 |
| 64 | Count | 1 | 1 | 2 |
| Expected Count | .9 | 1.1 | 2.0 |
| 69 | Count | 2 | 0 | 2 |
| Expected Count | .9 | 1.1 | 2.0 |
| Total | | Count | 164 | 206 | 370 |
| Expected Count | 164.0 | 206.0 | 370.0 |

Next is the Chi-Square Tests results.

|  |  |  |  |
| --- | --- | --- | --- |
| **Chi-Square Tests** | | | |
|  | Value | df | Asymptotic Significance (2-sided) |
| Pearson Chi-Square | 27.591a | 32 | .689 |
| Likelihood Ratio | 31.051 | 32 | .514 |
| Linear-by-Linear Association | .338 | 1 | .561 |
| N of Valid Cases | 370 |  |  |
| a. 39 cells (59.1%) have expected count less than 5. The minimum expected count is .44. | | | |

Our p-value is 0.05. The assumption that less than 20% have expected count less than 5 has been violated. Therefore, we cannot use the Pearson Chi-Square to determine if our null hypothesis is proven or not. Instead, we will use the Likelihood Ratio. The asymptotic significance is .514, which is much greater than our p-value. Therefore, our null hypothesis is accepted, and we have proven that there is no association between sex and young people.

1. **Passive or active coping styles**

**Research Hypothesis:** Young people do not use active and passive coping styles equally.

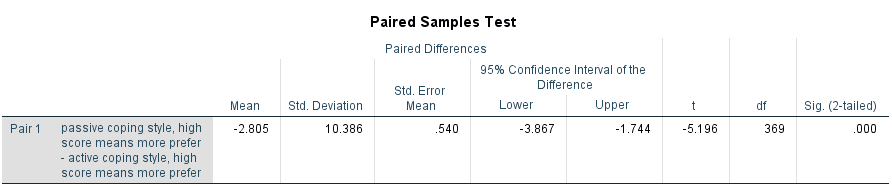
**Null Hypothesis:** Young people use active and passive coping styles equally.

**Alternative Hypothesis 1:** Young people use active coping styles more than passive coping styles.

**Alternative Hypothesis 2:** Young people use passive coping styles more than active coping styles.

The statistical test I will use is the paired sample t-test, since we must compare means of the data, and we are interested in comparing the differences between each person’s preferred coping style.

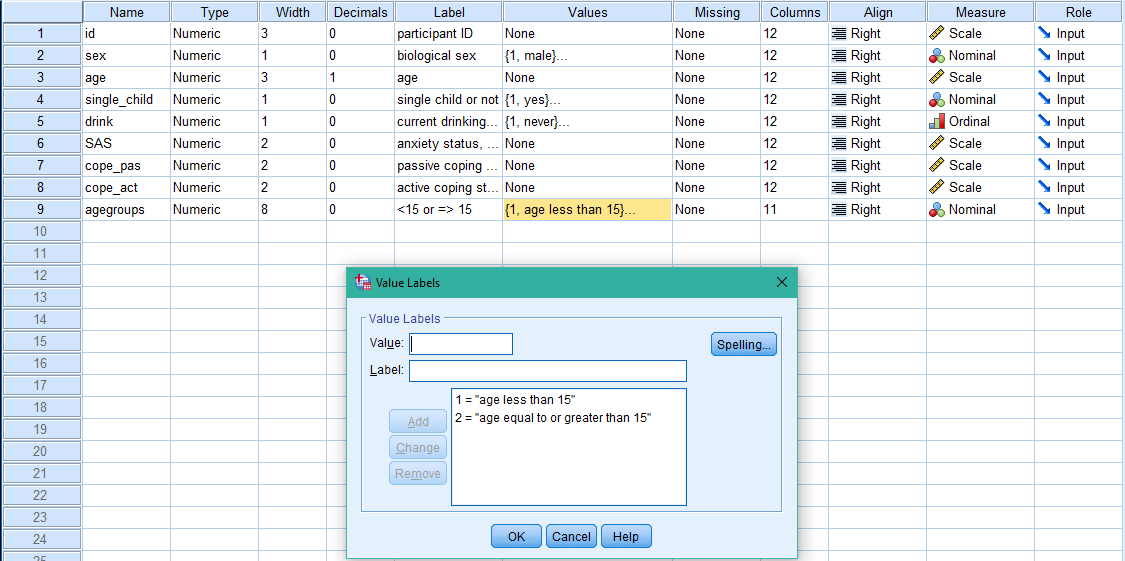
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Paired Samples Statistics** | | | | | |
|  | | Mean | N | Std. Deviation | Std. Error Mean |
| Pair 1 | passive coping style, high score means more prefer | 24.76 | 370 | 8.489 | .441 |
| active coping style, high score means more prefer | 27.57 | 370 | 8.292 | .431 |

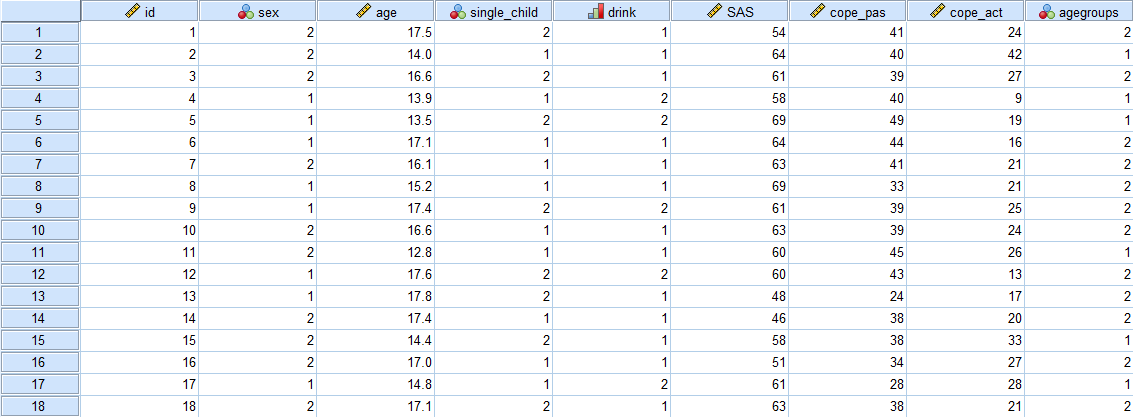


The .000 Sig is smaller than the p-value of 0.05. This indicates that the null hypothesis is rejected, and an alternative hypothesis is accepted. The negative t value indicates that an active coping style is used more than a passive coping style. The negative mean difference and the higher mean of active coping style shows that it is more preferred as well.

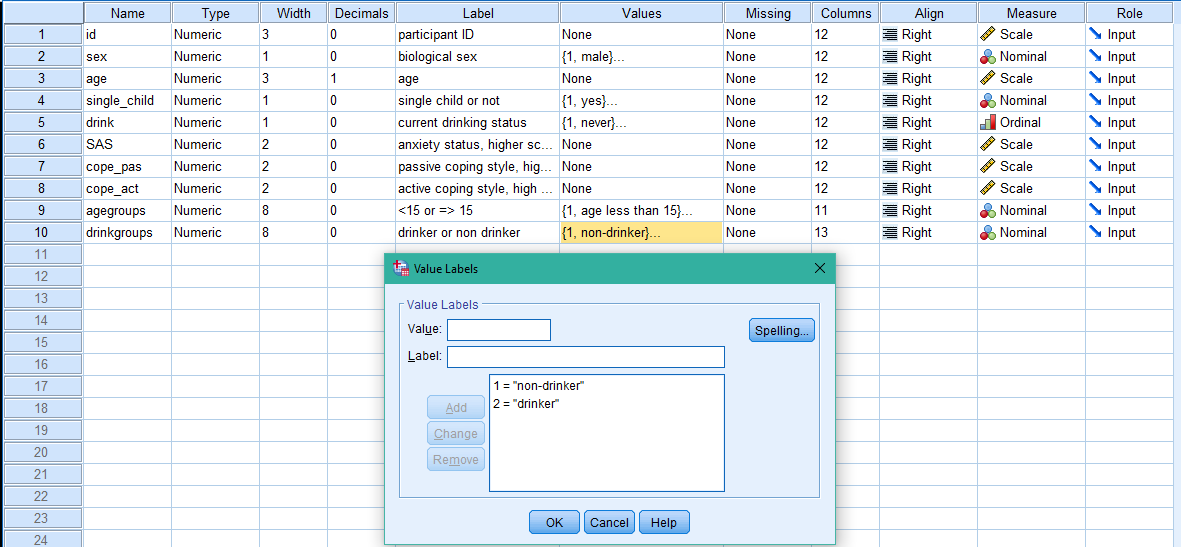
1. **Age and drinking**

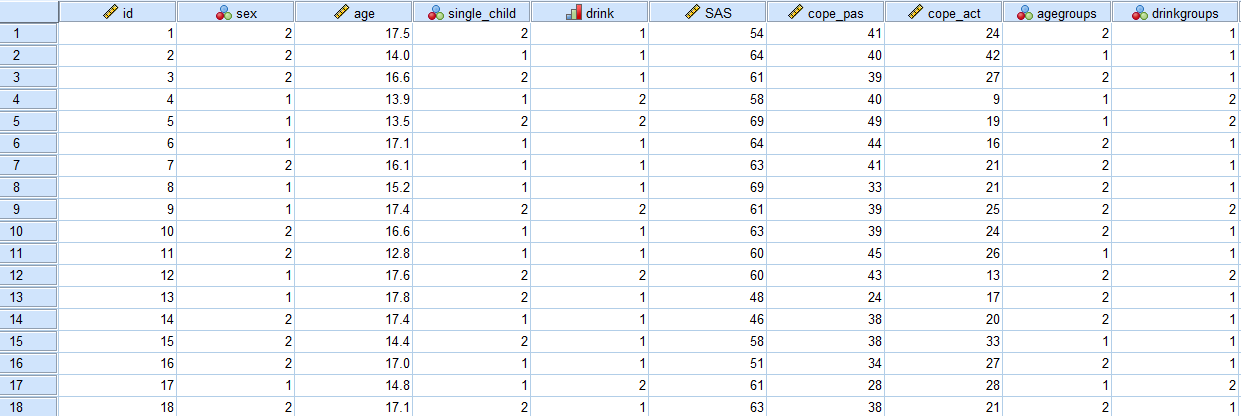
The screenshots below show the new variable agegroups which groups ages into two values.





Next, a new variable was created to change the ordinal groups for drinking from never, mild, moderate or heavy to drinker or non-drinker. It was assumed that mild, moderate and heavy are all classified as drinker and only never is classified as non-drinker.



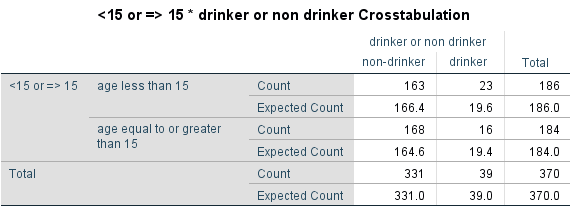


**Research Hypothesis:** There is association between age and drinking in young people.

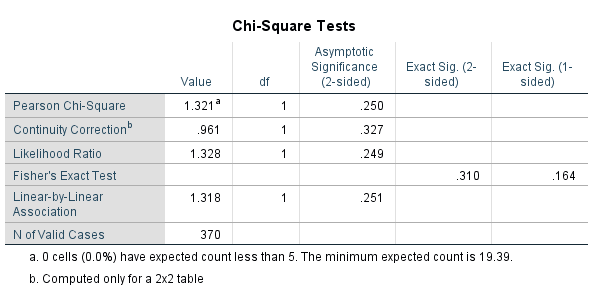
**Null Hypothesis:** There is no association between age and drinking in young people.

The statistical test I will use when analyzing this data will be a chi-squared test, since I am testing whether two variables are related or not.

Below is the crosstabulation of agegroup and drinkgroup, with the expected count.



Next is the chi-square test results



Our p-value is 0.05. The assumption that less than 20% have expected count less than 5 has been not been violated. Therefore, we can use the Pearson Chi-Square. The asymptotic significance is .250, which is much greater than our p-value. Therefore, our null hypothesis is accepted, and we have proven that there is no association between age and drinking in young people.