

D

APPENDIX D: SUMMARY OF STATISTICAL TERMS

Analysis of Variance (ANOVA) is a statistical technique used to measure and test differences between the mean scores of three or more groups of respondents responding to questions which require interval or ratio data.

Categorical data sets contain pieces of information (i.e. values or observations) that can be sorted into discrete groups. An example of categorical data is whether a respondent is female (group 1) or male (group 2).

Chi-square tests are used to measure and test differences between proportions of respondents in different groups in how they have answered questions which generate categorical or ordinal data.

Interval data are pieces of information measured by asking respondents to rank concepts on a numerical scale with equal intervals, and where the scale does not have a natural zero. Examples of interval data are those gathered in questions which ask respondents to rate on a scale of 1 to 5 how bad they would feel if they were to experience certain alcohol-related consequences.

Kendall's tau-b is a measure of correlation for ordinal or ranked variables; or, in other words, is a form of significance testing for testing the relationship between ordinal data generated by various groups of respondents. In this report, it is used to measure correlation between waves of research and ordinal data. Where data is only available for one wave, Kendall's tau-b is used to measure correlation between age (across three ordinal categories, 15-17 years, 18-21 years and 22-25 years) and ordinal data. Significant relationships are reported where appropriate.

Mean is the average score given to scores derived through gathering interval or ratio data. Means are derived by adding all scores given by respondents (excluding those who did not answer), then dividing them by the total number of respondents who answered.

Ordinal data are pieces of information measured by asking respondents to respond to questions which ask for rankings which have a logical order, but are not separated by equal intervals. Examples of ordinal data are those gathered in questions which ask whether respondents strongly disagree, disagree just a little, agree just a little or strongly agree with a statement.

Ratio data sets contain pieces of information that lie along a continuous scale which has a natural zero and where differences and ratios are interpretable. Examples of ratio data are those gathered in questions asking respondents to indicate how many glasses or bottles of different types of alcohol were consumed the last time they drank alcohol.

Statistical significance is the term used to describe differences between groups where the likelihood of those differences occurring purely by chance is low. In most cases, statistical significance is reported when the likelihood of differences occurring as a result of chance is less than 5% (p for probability <0.050), and cases in which the likelihood are lower (e.g., either lower than 1%, $p < 0.010$ or lower than 0.1%, $p < 0.001$) are reported where appropriate.