This module will discuss how to implement three statistical tests in R: a two-sample t-test, analysis of variance, and a chi-square test of independence.

**Lines 7 to 24** provide a reminder about how to set your working directory and read in data. Variable and value labels for the items that are used in the demonstration follow on **lines 28 to 66**. A dataframe that only includes the variables of interest is generated at **line 76**. Data cleaning procedures follow on **lines 70 to 135**.

[SCROLL TO 117]

An example of a nested ifelse statement can be found at **lines 117 to 126**. Recall that ifelse statements have three parts: the condition to test in the data, the value to supply if the condition is TRUE, and the value to supply if the condition is FALSE. A nested ifelse statement uses another ifelse statement as the value to supply if the condition is FALSE. An alternate approach to this recoding task is presented at **lines** **128-132.**

**Lines 139 to 166** demonstrate how toconduct a t-test comparing mean weight between subjects who endorse lifetime high blood pressure and those who do not endorse lifetime high blood pressure. The t test function generates an output object that we name TtestOutput. **Lines 157 to 166** demonstrate how to access the results of the test, which are stored as elements of the output object.

**Lines 170 to 203** demonstrate how to conduct an ANOVA comparing mean weight betweensubjects who endorse lifetime high blood pressure and take medication, subjects who endorse lifetime high blood pressure and do not take medication, and those who do not endorse lifetime high blood pressure. ANOVA results can be extracted using the summary() function. A posthoc tukey test is conducted at **line 187** to identify the pairwise comparisons where mean differences are statistically significant.

**Lines 207 to 226** demonstrate how to conduct a chi-square test of independence to examine if categorical BMI and lifetime high blood pressure are statistically independent from each other. **Lines 218 to 226** demonstrate how to access the results of the test, which are stored as elements of the output object.

**Lines 230 to 378** provide a few examples of how to generate tables that summarize the results of the t-test, ANOVA, and chi-square test. Custom tables are generated for the t-test and ANOVA results by extracting the relevant results, collating them in a dataframe, and exporting the dataframe with stargazer.

[SCROLL TO 351]

sjPlot has a convenient function for generating formatted crosstabs to accompany chi-square hypothesis test results. Note that the amount of information displayed in the table can be adjusted by setting the arguments at **lines 360-362 and 364** to FALSE.

**Lines 382 to 435** demonstrate how to visualize results using a selection of functions from sjPlot. These specialized functions can be quicker to set up than the code in ggplot, but are somewhat less flexible. Box plots are used to present group means for the t-test and ANOVA.

[SCROLL TO 408]

A line plot is used to display the proportion of subjects in each category of BMI, separated by lifetime high blood pressure status. This information can also be displayed as a bar plot by commenting out **line 412**, and removing the hashtag on **line 413**.