# ANCOVA

**Description:**

ANCOVA is a type of analysis of variance that allows you to first control for variable(s) that change the relationship of the independent variable with the dependent variable.

IV

DV

CV

Usually the section marked with an arrow would be considered error. But since we used the control variable (covariate, see below), it’s now accounted for, which means there is less error in the DV (always a good thing).

**Definitions/Abbreviations:**

* CV – covariate. This variable changes with the dependent variable and is used to adjust the dependent variable to create equal groups or eliminate nuisance information. CVs can be either a continuous variable (such as age) or a dichotomous variable (such as gender). CVs are used to adjust the values of the dependent variables, usually to create larger group differences on your independent variables. This variable, if actually related to the DV, helps control for error variance.
* IV – independent variable. This variable *has* to be a dichotomous variable. You can put people into groups based on any category (gender, handedness) or your experimental manipulation (instructions versus no instructions).
* DV – dependent variable. The dependent variable *needs* to be a continuous variable or another type of analysis might work better (see log regression). Your dependent variable should be the measurement you took in your study or what information you are expecting to see changed over groups.

**Relation to other designs:**

* ANCOVA is a type of ANOVA with control/covariates. However, ANOVA is mathematically related to regression – so it’s fair to say that ANCOVA is half-way between a traditional regression and ANOVA.
* First, the variance between the DV and CV is analyzed, looking at the overlap in the two (or more) variables. Then, the DV is adjusted to not include the CV variance (i.e. that variance is taken out). This step is very similar to regression, especially if the CV is continuous.
* Second, a regular ANOVA is analyzed on the adjusted means.
* **Therefore:** 
  + Between subjects designs:
    - Still get and interpret Levene’s test for homogeneity
    - Post hoc test: Independent t or BN ANCOVA on adjusted means
    - Post hoc correction: Tukey/Bonferroni
  + Repeated measures designs
    - Still get and interpret Mauchly’s test for Sphericity
    - Post hoc test: Dependent t or RM ANCOVA on adjusted means
    - Post hoc correction: Tukey/Bonferroni
  + Mixed designs
    - Get both Levene’s and Mauchly’s test
    - Post hoc test: depends on which IV you choose to split/analyze

Chart of ANCOVA Analysis:

|  |  |  |  |
| --- | --- | --- | --- |
|  | ANCOVA | | |
|  | Main Effect Between | Main Effect Repeated Measures | Interaction |
| If levels > 2  And significant | Independent t-test  Tukey correction | Dependent t-test  Tukey correction | SPLIT one IV column  Independent t-test OR  Dependent t-test  Tukey correction |
| If levels = 2 | Interpret means | Interpret means |

If the interaction is significant, often people ignore any analyses with the main effects:

* This procedure reduces Type 1 error because you are running less post hoc tests.
* You are interested in the interaction anyway, so why only interpret one variable at a time?
* Also, be sure to follow up with the correct test type – do not do dependent t on the between subjects factor.

**New to this section:**

* We will have to get the *adjusted means* to calculate effect size and for the post hoc tests.
* We will have to use a different type of post hoc correction – with ANCOVA, you have to control for the CV, and therefore, we cannot use pairwise.t.test, which would be calculated on the non-adjusted means.
* Tukey is another type of post hoc correction, which corrects on the *mean difference needed*, rather than the *p-value needed (alpha)*. These different places to correct have the same overall effect of controlling for Type 1 error. Tukey is less stringent than Bonferroni – but arguably, Bonferroni is easier to run.

**Notes on Assumptions:**

* Accuracy: same as regular ANOVA.
* Missing: same as regular ANOVA – mostly you cannot fill in because the percent of missing data would be too high.
* Outliers:
  + For between subjects ANOVA, we have been using z-scores to check for outliers because we would need more than one column of data to be able to use Mahalanobis. With ANCOVA, we have at least two columns – one for the DV and one for the CV. Therefore, you would be making sure that people do not have a strange combination of answers on the CV and DV.
* Additivity:
  + If you have ONE CV: not necessary.
  + If you have TWO OR MORE CVs: You want to check your CVs by using a correlation to make sure they do not overlap too much. If they overlap a great deal, then you want to use only one of them or combine them. Look for variables with *r* > .70. Highly correlated CVs cause you to lose power in this type of analysis.
* Linearity:
  + Linearity between the CV and DV is a very important issue. If you are using a regression to adjust the values of the DV with the CV, then there needs to be a linear relationship. You can check for this value using our fake regression procedure and looking at the QQ Plot.
* Normality:
  + You also want the CV and DV combination to be normally distributed. You can check for multivariate normality by running a fake regression and looking at the histogram of the residuals.
* Homogeneity:
  + The variance of the groups from your IV need to be equal across both the DV and CV. You can check this information with a residual plot from your fake regression (you do not want raining or an unequal spread of the dots around 0). You can also use Levene’s test of homogeneity – you *do not* want p < .001.
  + You will also check for Sphericity for repeated measures designs – you *do not* want p < .001.
* Homoscedasticity:
  + The spread of the scores on the CV need to be equal around all the values of the DV. You have to check this assumption by looking at a residual plot from a fake regression. You do not want a megaphone shape.

# Complete Example – All Between

With comments for Repeated Measures

The researcher has measured participant’s attitudes toward the use of drugs to see if there are group differences in attitudes. They also measured physical health symptoms to control for people who report more symptoms will have more positive attitudes towards drugs (because they are sicker). Finally, they grouped people by employment status and religion to analyze group differences and interactions on attitudes.

**Datafile:** bn ancova.csv

**IVs:**

* Employment – employed versus unemployed.
* Religion – none, Jewish, Protestant, Catholic.

**CV:**

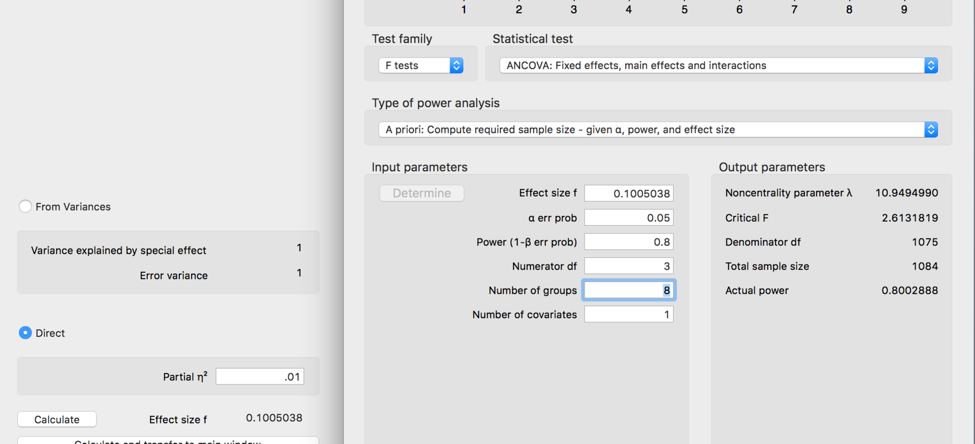
* Phyheal – physical health, higher scores indicate more health symptoms (problems).

**DV:**

* Attdrug – attitudes towards drugs, higher scores indicate more positive attitudes.

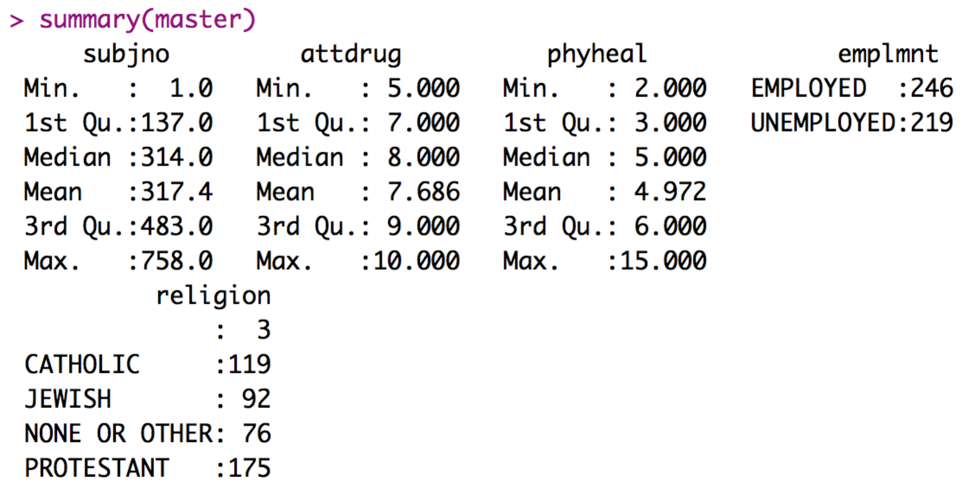
**Power:**

1. Open Gpower!
   1. Test family: F-test
   2. Statistical Test: ANCOVA: Fixed effects, main effects, and interactions
   3. Estimate an effect size: click determine 🡪 click direct 🡪 use eta square sizes you think might be accurate, remember small, medium, and large estimates from the notes.
   4. Alpha = .05
   5. Power (1-beta .20) = .80
   6. Numerator df:
      1. Main effects: Levels – 1
      2. Interaction: (Levels-1)\*(Levels-1)
2. Number of groups
   1. Main effects: Number of Levels
   2. Interactions: Number of Conditions
3. Number of covariates = number of CVs.
4. Let’s estimate the following:
   1. Small effect size (eta = .01)
   2. Interaction test:
      1. Numerator df = (2-1)\*(4-1) = 3
      2. Number of groups: 2\*4 = 8
   3. Number of covariates = 1
5. Let’s hope it’s not a small effect, as our power analysis says we need 1084 people!

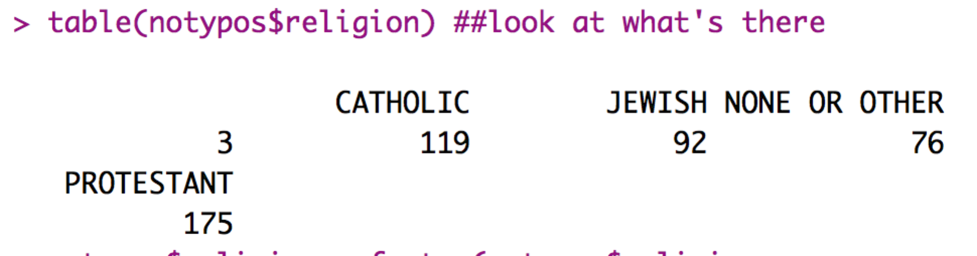


**Assumptions:**

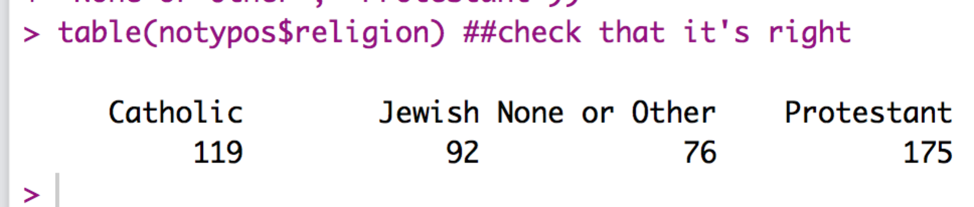
1. Accuracy:
   1. Use the summary(*dataset name*) function to get the basic information for the data.
   2. Let’s check out minimum and maximum:
      1. This data should be factored and not go below zero.
      2. The min and max are ok, but we should clean up the labels on the religion variable because there are some weird blank ones.
         1. Remember, that you can use the factor command to delete levels, just do NOT include it when you have the levels command.



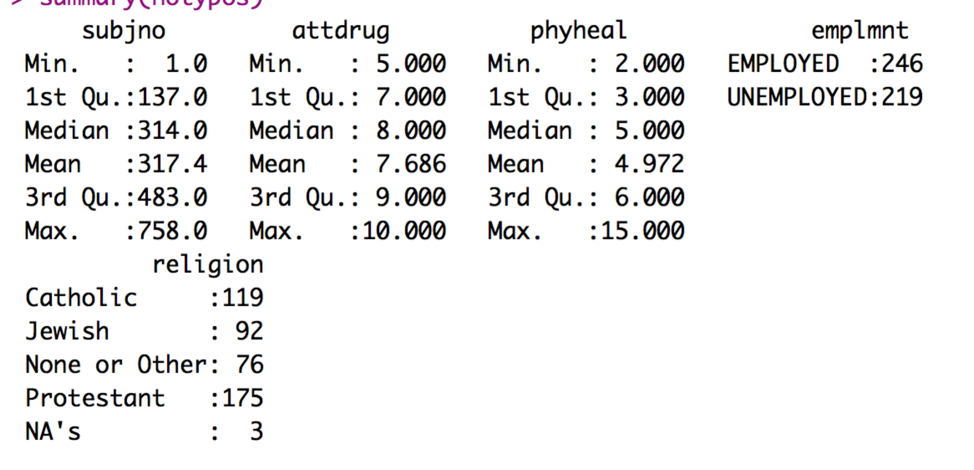
Before factoring:



After factoring:



1. Missing:
   1. With the summary function, I can also see that I have missing data.
   2. Even if there was missing data, remember that any missing data ends up being more than 5% for each participant in an ANOVA. Therefore, they should normally get excluded.

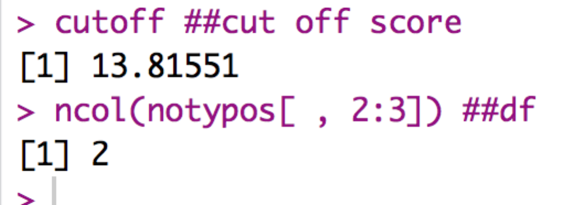


1. Outliers:
   1. We have both DV and CV(s), we have several columns to work with, which means we can use Mahalanobis values. We want to use this format for data screening because it accounts for the fact that people have more than one variable to screen.
   2. BUT: don’t forget that you cannot use the factored columns in Mahalanobis.
   3. Create the Mahalanobis values:
      1. mahal = mahalanobis(*dataset*,

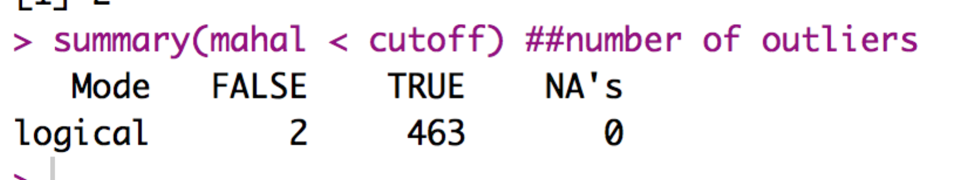
colMeans(*dataset*, na.rm = T),

cov(*dataset*, use = “pairwise.complete.obs”))

* 1. Create the cut off score:
     1. cutoff = qchisq(1-.001, ncol(*dataset*))
  2. Remember you can use:
     1. cutoff to get the cutoff score
     2. ncol(*dataset*) to get the *df*

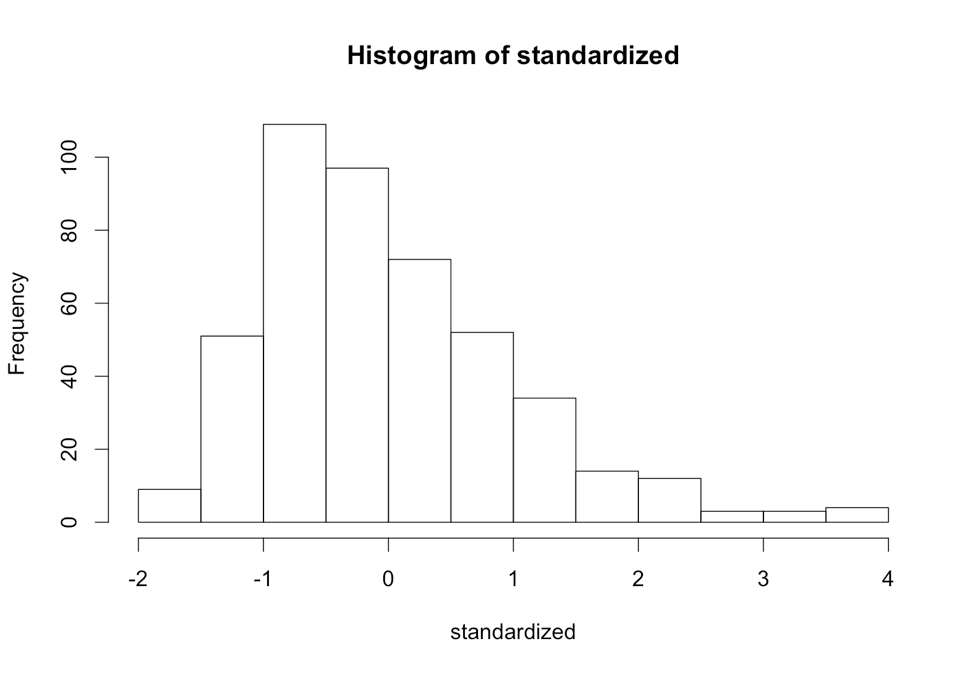


* 1. See how many outliers you have:
     1. summary(mahal < cutoff)

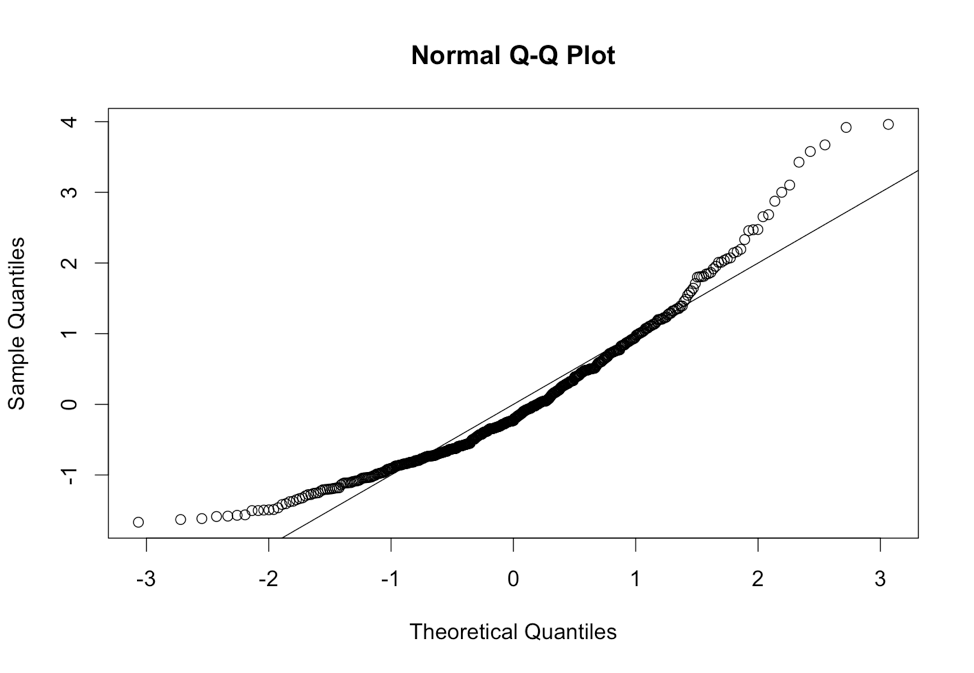


* + 1. Remember FALSE is bad.
    2. I have two outliers!
  1. Exclude outliers:
     1. noout = subset(*dataset*, mahal < cutoff)

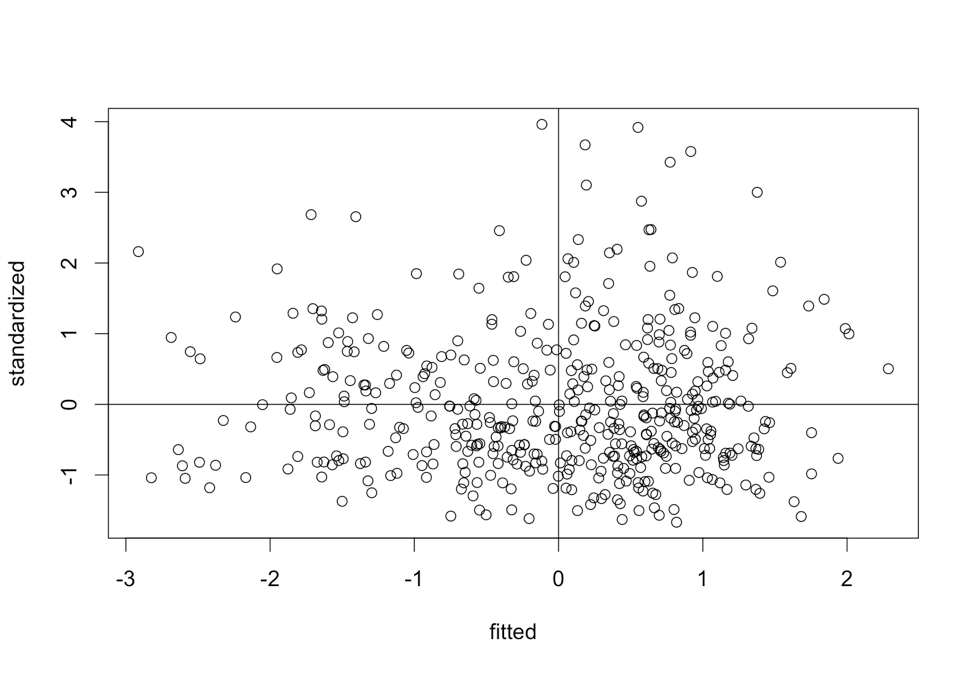
1. Additivity
   1. We do have to worry about correlations with repeated measures and ANCOVA, but not quite in the same way we talked about it for overall data screening.
   2. In general, you *want* the various repeated measurements (if you have them, we don’t in this example) to be highly correlated – it will give you more power if they are correlated and less if they are not.
      1. However, they cannot be perfectly correlated or the ANOVA will not run.
      2. Mainly we are checking that we don’t get any 1s other than the diagonal in our symbols chart. So, basically, the rule is the *r* < .999.
   3. For CVs, we do not want them to correlated over .70 – that means that they would overlap by 49%, which means that you are wasting your time using both CVs.
   4. Get the correlations:
      1. correl = cor(*dataset*, use = “pairwise.complete.obs”)
   5. Get the symbols chart:
      1. symnum(correl)
   6. Look for 1s NOT on the diagonal for repeated measures, and , + or B for CVs.
2. Set up the rest of the assumptions:
   1. Make a random variable:
      1. random = rchisq(nrow(*dataset*), 7)
   2. Run a fake regression:
      1. fake = lm(random~., data = *dataset*)
   3. Create the standardized residuals:
      1. standardized = rstudent(fake)
   4. Create the fitted values:
      1. fitted = scale(fake$fitted.values)
3. Normality:
   1. hist(standardized)
   2. Most of the data is between -2 and 2 and is centered over 0 – but there definitely is a skew to the distribution, even after taking out outliers.
   3. Because we have more than 30 people, we do not have to worry because of the central limit theorem.



* 1. Linearity:
     1. qqnorm(standardized)
     2. abline(0,1)
     3. This graph is ok … most of the data between -2 and 2 is on the line, remember it’s hard to predict the values out on the tails.



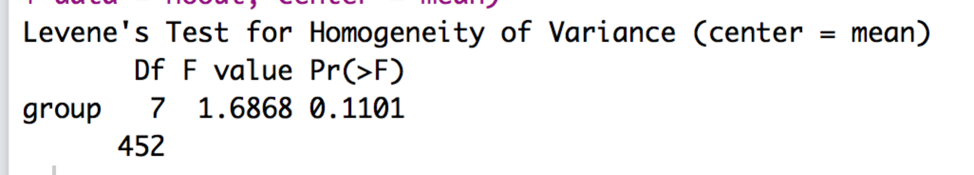
* 1. Homogeneity and Homoscedasticity
     1. plot(fitted,standardized)
     2. abline(0,0)
     3. abline(v = 0)
     4. Homogeneity: remember you want the data centered evenly around zero.
        1. Not so much here on the vertical axis. The horizontal axis is ok, but the vertical runs -1 to 4. We should check Levene’s test to make sure everything is ok.
     5. Homoscedasticity: the data seems fairly evenly spread across the graph, so I would say it’s ok. You have to check this part because of the linear regression part of the CV and DV.



* 1. Homogeneity: Take 2 Levene’s Test
     1. Levene’s is a test for homogeneity between groups, so it looks to see if the variances are equal across your IV levels.
     2. It is notoriously **oversensitive**, but can be a good place to start if you want to check a real number, rather than this scatterplot.
     3. With large sample sizes, it is often significant (remembering the big important rule, p<.001), and with large sample sizes it matters less. Ergo, if you have big *n* in each group, then don’t worry about it so much.
     4. You will have to run the ANOVA to get Levene’s Test, see below.
     5. **HOW TO GET IT:**
        1. Because we are not using ezANOVA for ANCOVA, we have to get Levene’s separately.
        2. Load the car library.
           1. library(car)
        3. Run the Levene’s test with only the between subjects factors:
           1. leveneTest(*dv column* ~ *iv column*\**iv column*,

data = *dataset,* center = mean)

* + - * 1. If you only have one between subjects factor, just take out the \**iv column* half. If you have more than two, just add more \**iv columns.*
        2. Use center = mean because we are analyzing means in ANOVA.
      1. Woohoo – it’s not significant because p > .001.



* 1. Homogeneity: Take 3 Mauchly’s Test for Sphericity (if you have repeated measures, which we do not in this example).

1. Mauchly’s is a test for homogeneity between repeated measures (so to speak), which is called sphericity.
2. The assumption is considered *compound symmetry*:
   * + 1. The correlations between all the levels are equal.
       2. The variance of the difference scores between each level combination is the same.
     1. It is almost impossible to meet this assumption:
        1. Generally, you are examining if there are differences in levels.
        2. They are often taking the same thing over and over.
        3. So, the variances often get much smaller or larger across the levels.
        4. It’s such a problem, people often ignore sphericity.
     2. Something important!
        1. IF there are only two levels of an IV, you will NOT get Mauchly’s test. Why? It’s hard to compare the correlation and variance of the difference scores if there are only two levels (because there is only one correlation and one difference score).
     3. **HOW TO GET IT:**
        1. I’d suggest running an ezANOVA with just the repeated measures to determine if you need to correct for sphericity – it’s not exactly easy for ANCOVA.
        2. library(ez)
        3. ezANOVA(data = *dataframe,*

dv = *dv column,*

wid = *participant number column,*

within = .(*repeated measures columns*),

type = 3)

**Running the ANOVA:**

1. Remember, if you have repeated measures switch from WIDE to LONG format.
   1. Long format for repeated measures means that each level + participant get their own row … so that there is one column for the IV and one column for the DV.
   2. Install / load the reshape library (NOT reshape2).
   3. library(reshape).
   4. Melt the data (run all these lines):

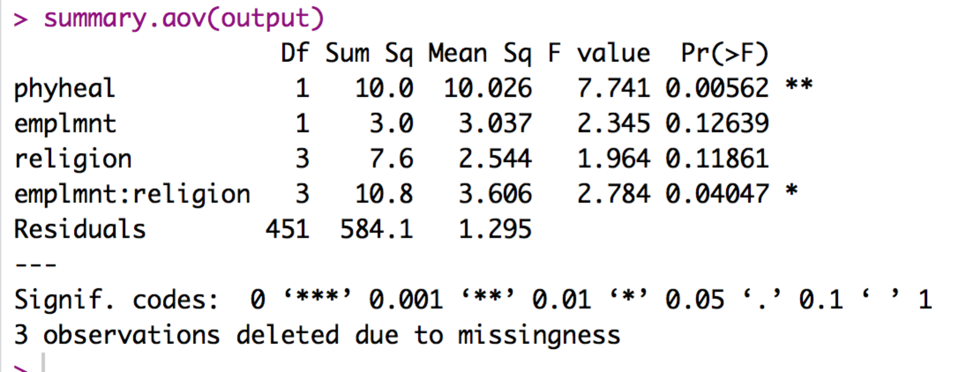
longdata = melt(*dataset*,

id = c("*partno*", “*between subjects column*”),

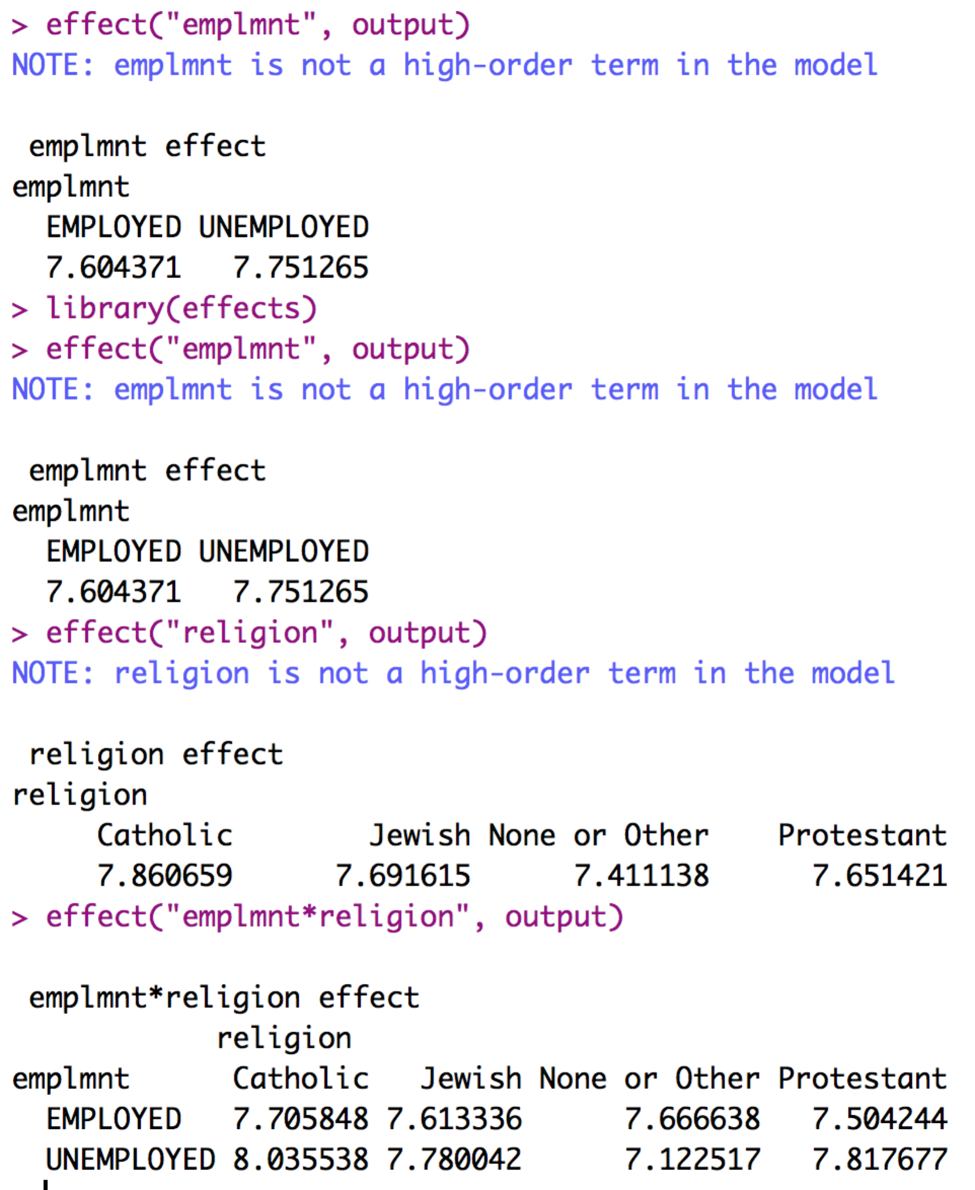
measured = c("*level column*", " *level column* ", " *level column*"))

* 1. Be sure to put your between subjects group column in the id section – you do not want to melt that data (plus it will give you an error because it’s not repeated across subject number).
  2. I’d suggest relabeling the column names since variable and value are not that helpful – change *column* out here to the new names.
     1. colnames(*dataset*) = c(“*column”, “column”, “column”*)

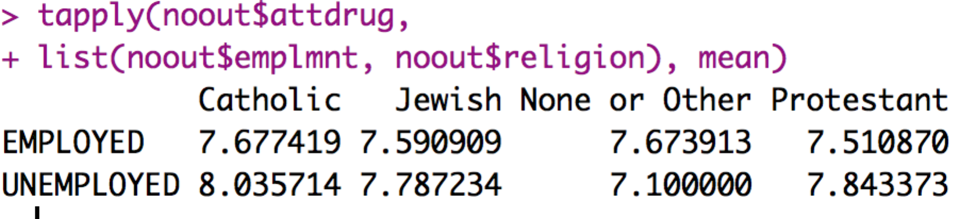
1. Run the ANCOVA using the lm() function.
   1. output = lm(*DV* ~ *IV\*IV* + *CV*, data = *dataset*)
   2. summary.aov(output)
   3. What’s going on?
      1. We’ve been using the lm function for our fake regressions, but now we are going to use it because ANCOVA is half-regression/half-ANOVA.
         1. (ok, so ANOVA is regression, but most people interpret it in traditional ANOVA fashion).
      2. Therefore, we are going to create a formula for our DV…it is ~ (predicted) by the IVs and the CV. We multiply the IVs to get interactions, but just add the CV to control for it.
      3. If you multiply a continuous CV and categorical IVs, you might as well skip to the regression section on moderation, as that is what you are doing instead of ANCOVA.
      4. We are going to treat CVs as control variables, so only control for them. (hence, the adding and not multiplying).
      5. summary.aov gives you the output in ANOVA style.



1. Interpret the output:
   1. In this output
      1. DF = df but the model df is next to the name of the variable, and the error df is next to the word residuals. So, we will use 451 as the second number for each of our DFs.
      2. Use F value as F and pr (>F) as p.
      3. We will use Sum Sq for effect size calculation.
   2. Check the omnibus (overall) test for your CV:
      1. *F*(1, 451) = 7.74, *p* = .006 … wait what happened to eta? ☹
      2. Calculate partial eta squared: Sum Sq for the variable / (Sum Sq for the Residuals + Sum sq for the variable)
      3. For phyheal = 10.0 / (584.1+10.0) = .02
      4. *F*(1, 451) = 7.74, *p* = .006, ηp2 = .02
      5. What does a significant CV mean?
         1. It means that the CV is significant related to the DV. Most people say that means it’s a significant *adjustor* of the DV.
         2. You can run a correlation to figure out the direction and magnitude of the relationship.
         3. cor(*dataset*$*dv*, *dataset*$*iv*)
         4. The correlation is *r* = .12, which means it a small positive relationship. As physical health symptoms increase, the attitude toward drugs also gets more positive (increases).
         5. Now, we are going to interpret the ANOVA main effects *controlling* for this relationship.
   3. Check the Omnibus (overall) test for your IVs:
      1. We have three of them! What happened?!
         1. You will get one *F* test for each IV and then also the interaction.
         2. You will interpret each one separately.
         3. Remember, if the interaction is significant, only do post hocs for the interaction.
         4. You will use the same eta formula for these effects as well.
      2. Write that up:
         1. Employment: *F*(1, 451) = 2.35, *p* = .12, ηp2 = .01.
         2. Religion: *F*(3, 451) = 1.96, *p* = .12, ηp2 = .01.
         3. Interaction: *F*(3, 451) = 2.78, *p* = .04, ηp2 = .02.
2. Post Hoc Interpretation/Plan:
   1. Oh boy. We have to calculate post hocs on the adjusted (controlled) means. That’s a bit of trouble.
   2. To get the adjusted means:
      1. Load the effects library.
      2. Use the following code:
         1. effect(“*IV column*”, output)
         2. IV column must be in quotes.
         3. Output is whatever you save the lm() as, which we called output earlier.
         4. You can do “*IV\*IV”* for the interaction.



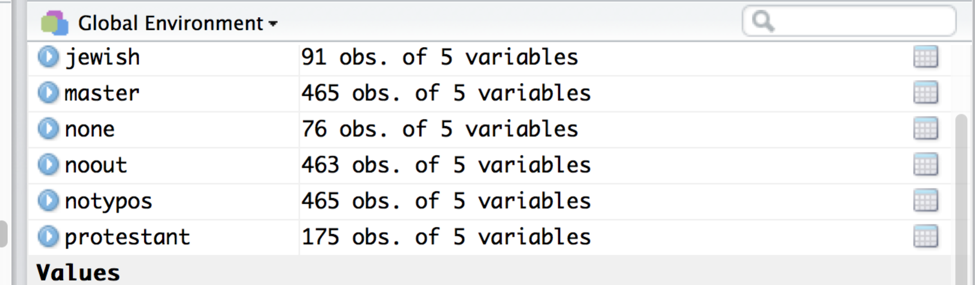
* 1. Just as quick proof the means are different run tapply with the mean to see the unadjusted means.
     1. You will see only a small difference because our correlation between CV and DV is pretty small.



* 1. To get the SDs and N, we can use tapply. We are (more than likely) slightly overestimating the SD, but that’s ok, as overestimating the error is better than underestimating
     1. tapply(*dataset$DV*, list(*dataset*$*IV, dataset$IV*), sd)
     2. tapply(*dataset$DV*, list(*dataset*$*IV, dataset$IV*), length)
     3. Remember, you can take out one of the IVs to just get main effects.
  2. None of the main effects are significant, so we are going to just say that employment and religion do not have any individual effect.
  3. Interaction: is significant, so let’s analyze that!
     1. We have 2X4 ANOVA – so we have eight boxes to consider:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | Religion | | | |
|  |  | None | Catholic | Jewish | Protestant |
| Employment | Unemployed |  |  |  |  |
| Employed |  |  |  |  |

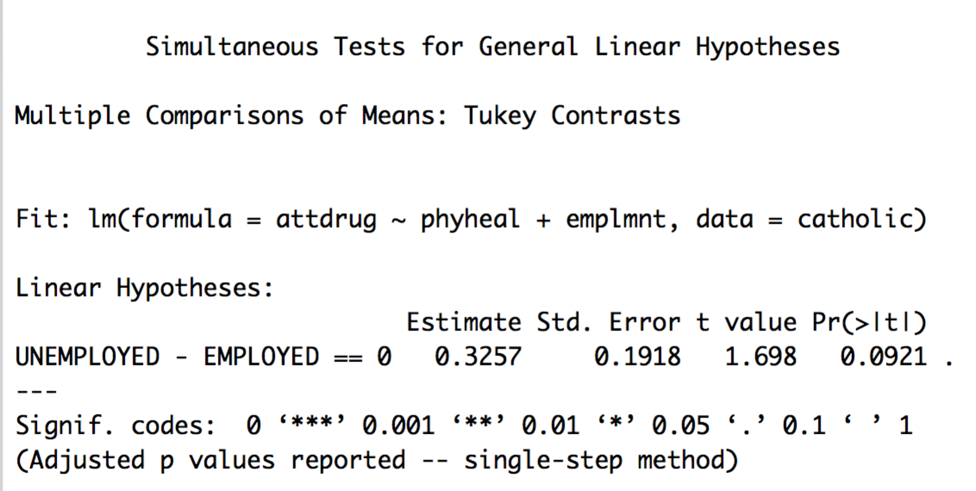
* + 1. How do we want to compare them?
    2. The rule usually is across or down but not both.
       1. Which way?
          1. First, pick based on a hypothesis.
          2. Second, go with the lesser number of tests to save type 1 error rate.
       2. So, we are going to split on religion, so we only get four tests: for each religion, we are going to compare employed versus unemployed.
    3. Once you pick a direction, you will need to SPLIT the dataset into chunks to analyze each piece separately.
       1. Use the subset function!
       2. Be sure to look at your N values to make sure they changed – it should total up to the full number of rows in the last dataset.



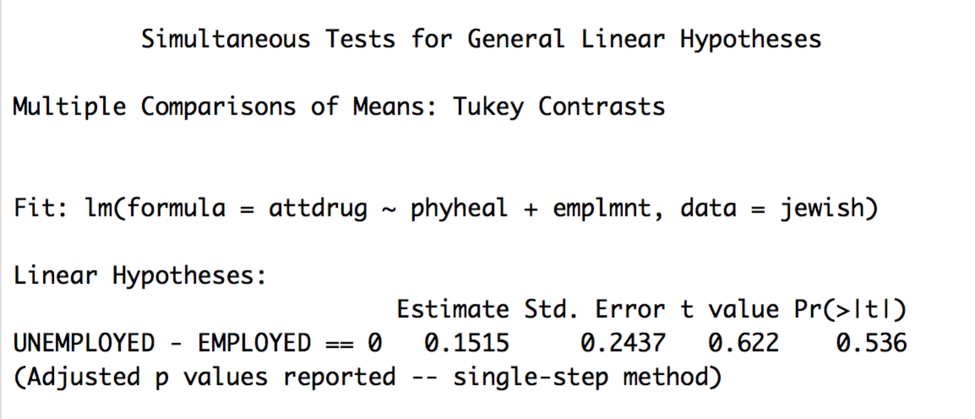
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mean 1 | Mean 2 | P-value | Explain? | Effect size |
| CATHOLIC  UN  M = 8.04  SD = 1.09  N = 56 | CATHOLIC  EMP  M = 7.71  SD = 0.99  N = 62 |  |  |  |
| JEWISH  UN  M = 7.78  SD = 1.21  N = 47 | JEWISH  EMP  M = 7.61  SD = 1.11  N = 44 |  |  |  |
| PROTESTANT  UN  M = 7.82  SD = 1.19  N = 83 | PROTESTANT  EMP  M = 7.50  SD = 1.08  N = 92 |  |  |  |
| NONE  UN  M = 7.12  SD = 1.18  N = 30 | NONE  EMP  M = 7.67  SD = 1.35  N = 46 |  |  |  |

* 1. For the post hoc test and correction:
     1. Load the multcomp library.
     2. For interactions, run little mini ANCOVAs on each new dataset.
        1. Be sure to use the new dataset! You will have to do the same thing for each dataset that you just created.
        2. Also, you will have to take out the variable you split on – or it will give you a grumpy error.
        3. Save each of the outputs.
     3. Run the glht() function for each output.
        1. post\_output = glht(*ancova output*, linfct = mcp(*IV analyzing,* = “Tukey”)
        2. IV name should not be in quotes.
        3. summary(post\_output)
     4. Use these corrected p values to fill in your table.

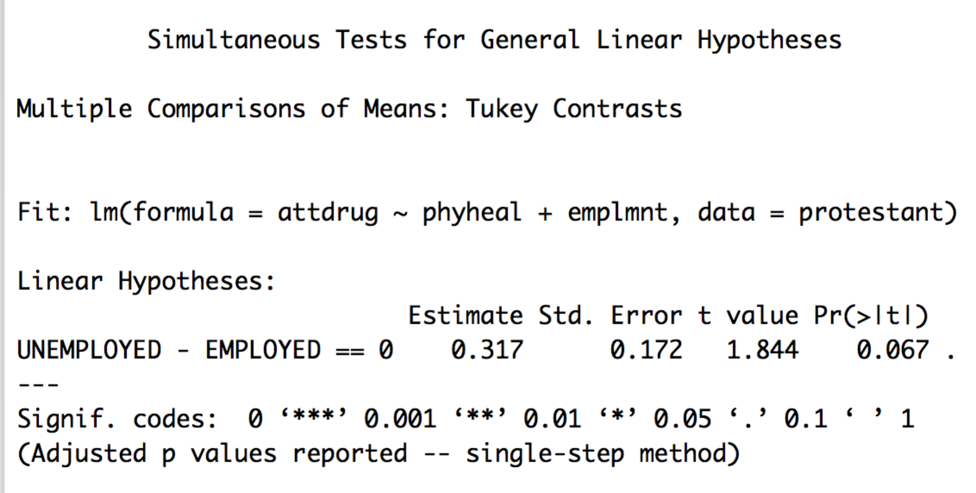
Catholic:



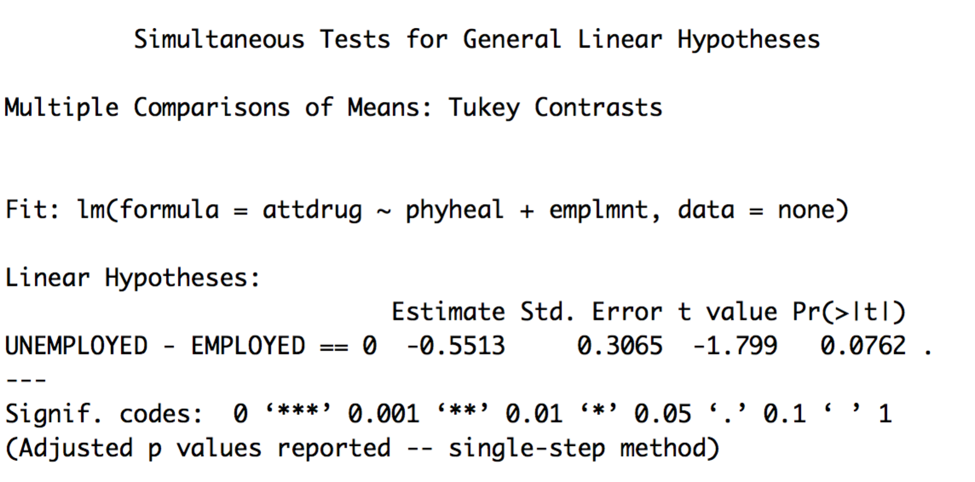
Jewish:



Protestant:



None:



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mean 1 | Mean 2 | P-value | Explain? | Effect size |
| CATHOLIC  UN  M = 8.04  SD = 1.09  N = 56 | CATHOLIC  EMP  M = 7.71  SD = 0.99  N = 62 | .09 | Not significant,  Marginal un > employed |  |
| JEWISH  UN  M = 7.78  SD = 1.21  N = 47 | JEWISH  EMP  M = 7.61  SD = 1.11  N = 44 | .54 | Not significant |  |
| PROTESTANT  UN  M = 7.82  SD = 1.19  N = 83 | PROTESTANT  EMP  M = 7.50  SD = 1.08  N = 92 | .06 | Not significant, marginal un > employ |  |
| NONE  UN  M = 7.12  SD = 1.18  N = 30 | NONE  EMP  M = 7.67  SD = 1.35  N = 46 | .07 | Not significant, marginal employ > un |  |

* 1. Note: So, what happened? Why didn’t we get any significant post hocs if we got a significant interaction?
  2. Couple reasons:
     1. If the interaction is small, sometimes the correction factor (aka Tukey, Bonferroni), will wipe out the small differences that existed in your omnibus ANCOVA.
     2. The differences in means existed in the other direction: so if I had run analyses comparing each religion for employed groups and then compared religions for unemployed groups, I would have found significant differences in:
        1. No differences for employed
        2. Unemployed: none v catholic, none v protestant
     3. The differences in means are in a strange combination you wouldn’t test.
        1. For example, it doesn’t make a whole lot of sense to compare none\_employed groups to jewish\_unemployed groups.
  3. What would you do? Generally, you might reconsider the analysis you present / think about what that means for your hypothesis test (do you have enough people? Is this result practically important?). For practical importance, look at effect sizes and directions – for Catholic um > employed but for None it was the reverse, so that’s what is probably driving the interaction (because they are in opposite directions).
  4. You can use MOTE to calculate the effect sizes by loading the MOTE library.
     1. library(MOTE)
  5. We will use d.ind.t for these calculations because it matches between subjects.
     1. d.ind.t(m1 = #, m2 = #, sd1 = #, sd2 = #, n1 = #, n2 = # a = .05)
  6. Make sure each M, SD, and N look correct.
  7. Enter *d* only into your table.
  8. You can make *d* values positive or negative – I tend to report them as always positive because the negative just indicates that you subtracted the smaller mean first, not anything about the actual effect size.
     1. These are all pretty small effects (which matches eta square sizes as well).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mean 1 | Mean 2 | P-value | Explain? | Effect size |
| CATHOLIC  UN  M = 8.04  SD = 1.09  N = 56 | CATHOLIC  EMP  M = 7.71  SD = 0.99  N = 62 | .09 | Not significant,  Marginal un > employed | .32 |
| JEWISH  UN  M = 7.78  SD = 1.21  N = 47 | JEWISH  EMP  M = 7.61  SD = 1.11  N = 44 | .54 | Not significant | .15 |
| PROTESTANT  UN  M = 7.82  SD = 1.19  N = 83 | PROTESTANT  EMP  M = 7.50  SD = 1.08  N = 92 | .06 | Not significant,  Marginal un > employed | .28 |
| NONE  UN  M = 7.12  SD = 1.18  N = 30 | NONE  EMP  M = 7.67  SD = 1.35  N = 46 | .07 | Not significant, marginal employ > un | .43 |

**Graphs:**

1. The best type of chart for anything analyzing group means is a bar chart with error bars.
2. We are going to use ggplot2 to build all our graphs.
   1. The package works like a transparency machine – you build layers and add them to the graph. You will really want to learn to stack your code, so that it’s easy to troubleshoot any problems you have.
3. Load the ggplot2 library.
   1. library(ggplot2).
4. We are going to clean up the gray background, the nondiscriminate axes, and the tiny type that always happens with plots.
   1. Separate from the graph code, run this code exactly:

cleanup = theme(panel.grid.major = element\_blank(),

panel.grid.minor = element\_blank(),

panel.background = element\_blank(),

axis.line = element\_line(colour = "black"),

legend.key = element\_rect(fill = "white"),

text = element\_text(size = 15))

* 1. This code saves a whole bunch of settings as cleanup, which then we can add to our graph.

1. Create a blank graph with the right variables.
   1. X = IV, Y = DV.
   2. bargraph = ggplot(*datasetname,* aes(*Xcolumn, Ycolumn,* fill = *IVcolumn*))
   3. Note: fill has to be a factored variable. This variable will be put into a legend.
2. Which one should be the legend versus X axis?
   1. I put my split variable for interactions on the X axis, so the post hoc tests match the bars that are paired together.
3. Add things to the plot:

bargraph +

stat\_summary(fun.y = mean,

geom = "bar") +

stat\_summary(fun.data = mean\_cl\_normal,

geom = "errorbar",

position = position\_dodge(width = 0.90),

width = 0.2)

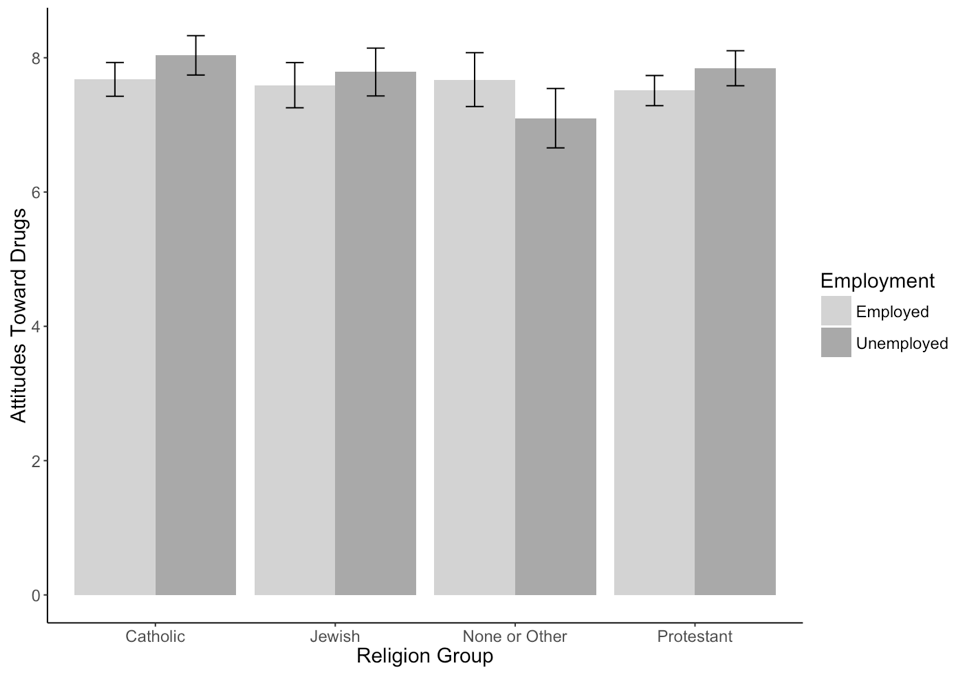
* 1. Please note:
     1. That code above stays exactly the same, but remember that “” doesn’t copy correctly sometimes.
     2. What does it do?
        1. The first stat\_summary adds the bars to the graph by graphing the mean for each group.
        2. The second stat\_summary adds the error bars of the confidence interval (approximately 2\*SE). These bars help you see how much the variance is spread around each group

1. Label X and Y:
   1. xlab(“Text that you want”) + ylab(“Text that you want”) will fix the axes labels.
2. Next issue – the bad looking legend and colors:
   1. Notice in the first line we created the graph, we used the word FILL.
   2. We can do scale\_fill\_manual to fix that problem. The name part will change the overall label, and you can use labels if you want to fix the level labels.
   3. You can also make it black / gray / white / green / purple by using the values command.
   4. scale\_fill\_manual(name = c(“*Name of IV*”),

labels = c(“*level”, “level” ,…*),

values = c(“*color*”, *“color”, …*))

1. You could also change the Y axis length if you had specific limits.
   1. coord\_cartesian(ylim = c(#,#))
   2. Where # are the actual lower and upper limits.



*Figure 1*. Adjusted interaction means for religion and employment on attitudes towards drugs.

**Results**

Participants were measured on their attitudes toward drug use and physical health symptoms to examine the interaction between employment and religion on attitudes. The data were screened for assumptions, and there were two multivariate outliers were present (using Mahalanobis distance), which were removed for the final analysis. Therefore, all participants were used for this analysis. Linearity, homogeneity (Levene’s *F*(7, 452) = 1.67, *p* = .11), homoscedasticity, and normality were all met.

A between subjects ANCOVA was used to analyze the interaction between employment and religion on attitudes towards the use of drugs after controlling for physical health symptoms. Physical health was a significant adjustor of attitudes towards the use of drugs, *F*(1, 451) = 7.74, *p* = .006, *ηp2*= .02. Physical health was positively related to attitudes of the use of drugs (*r* = .12), which indicated that as more symptoms are listed more positive attitudes towards drugs exist. After controlling for physical health, the main effect of employment was not significant, *F*(1, 451) = 2.35, *p* = .12, *ηp2* = .01, indicating that employed participants (*M* = 7.62, *SD* = 1.11) had the same attitudes as unemployed participants (*M* = 7.70, *SD* = 1.20). The main effect of religion was not significant, *F*(3, 451) = 1.96, *p* = .12, *ηp2* = .01, with None or other groups reporting low attitudes (*M* = 7.40, *SD* = 1.13), followed by Protestant groups (*M* = 7.66, *SD* = 1.09), Jewish (*M* = 7.71, *SD* = 1.12), and finally Catholic participants (*M* = 7.86, *SD* = 1.11).

The interaction between employment and religion was significant, *F*(3, 451) = 2.78, *p* = .04, *ηp2* = .02, which is shown in Figure 1. Independent t-tests with a Tukey correction were used to examine the adjusted means by employment type. Catholic participants showed a more positive attitude when unemployed, *p* = .09, *d* = 0.32, while none or other participants showed a more positive attitude when employed, *p* = .07, *d* = 0.43. The other two comparisons for Jewish, *p* = .54, *d* = 0.15, and Protestant, *p* = .62, *d* = 0.28, where not significant.