**PSY 653 Module 1: Missing Data Techniques and Meta-Analyses**

**Jan 29, 2020**

**Missing Data Techniques**

*Part 1: In-Class Demo*

Follow along as we talk through the steps of using Multiple Imputations by Chained Equations (MICE) to handle missing data.

1. Create a new R notebook
2. Install and load the mice package. Also load the following libraries: olsrr and tidyverse.
3. Read in the datafile mice\_data1.csv
4. Conduct a simple linear regression in which X1 is regressed on X2. This uses pairwise deletion of missing values by default
5. Using the mice() function, impute the dataset 5 times
6. Using the with() function, perform a simple linear regression on the five imputed versions of the dataset in which X1 is regressed on X2
   1. Examine how the model estimates vary across imputed versions of the dataset
7. Using the pool() function, calculate the pooled model estimates for the linear model
   1. How does the pooled regression estimate differ from that of the original model that used pairwise deletion?

*Part 2: Try It Yourself*

1. Read in the datafile called mice\_data2.csv
2. Conduct a simple linear regression in which X1 is regressed on X2, X3, and X4. This uses pairwise deletion of missing values by default
3. Using the mice() function, impute the dataset 5 times
4. Using the with() function, perform a simple linear regression on the five imputed versions of the dataset in which X1 is regressed on X2, X3, and X4
   1. Examine how the model estimates vary across imputed versions of the dataset
5. Using the pool() function, calculate the pooled model estimates for the linear model
   1. How does the pooled regression estimate differ from that of the original model that used pairwise deletion?

**Meta-Analyses Techniques**

The file “studies.csv” includes the outcomes of 20 studies where a particular test was correlated with a standard measure of academic achievement. The file includes the correlation, N, and a study designator (S1, S2..)

*Part 1:* In class Demo

Follow along as we talk through the steps of using the meta package to run a meta analysis on 20 different study correlation effect sizes.

1. Create a new R notebook
2. Install and load the meta package. Also load the following libraries: tidyverse.
3. Read in the datafile studies.csv
4. Use the metacor function to perform a meta-analysis across the 20 studies. Save it to an object
5. View your newly created object.
6. What conclusions do you reach?
7. Use the forest function to create a forest plot of your studies
   1. What does the forest plot tell you?
8. Use the funnel function to create a funnel plot of your studies
   1. What does the funnel plot tell you?

*Part 2: Try It Yourself*

1. studies2.csv contains the same results with the following 5 studies added

\_\_\_\_\_\_\_\_\_\_

r N

.04 400

-.14 2000

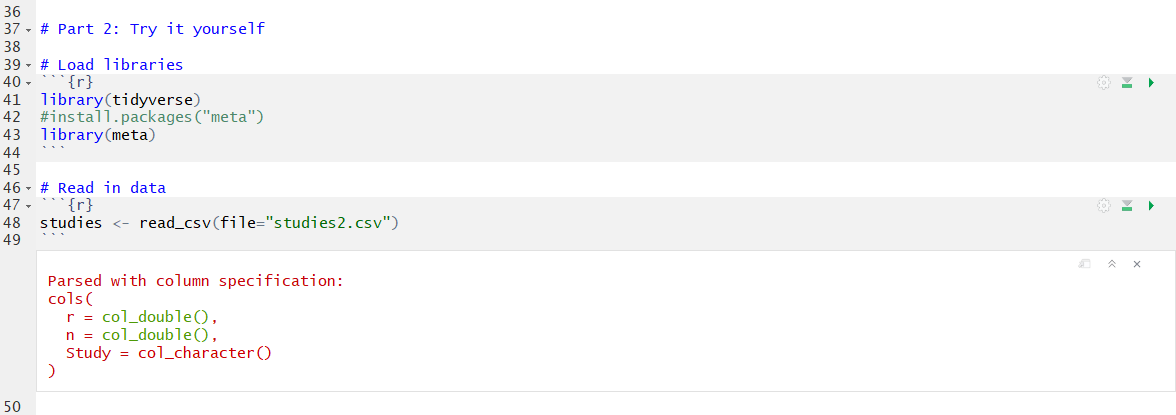
.22 310

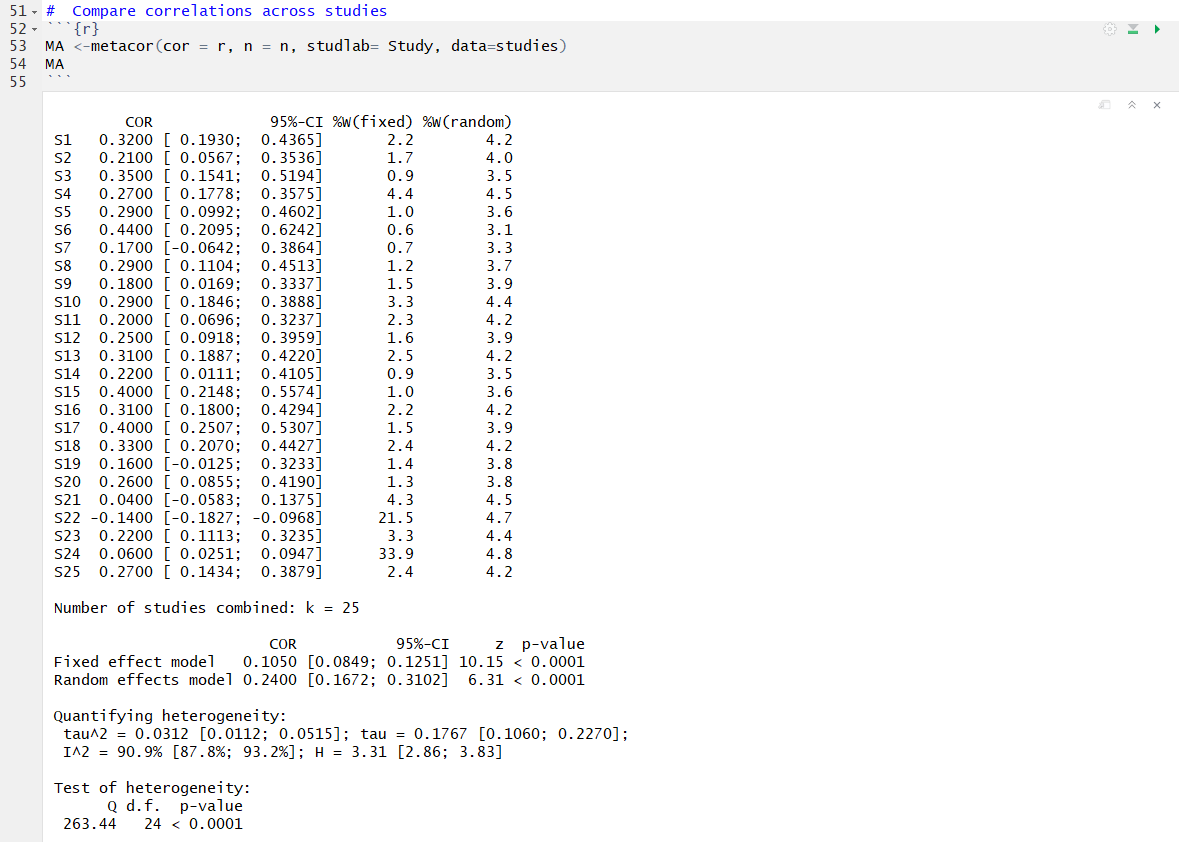
.06 3150

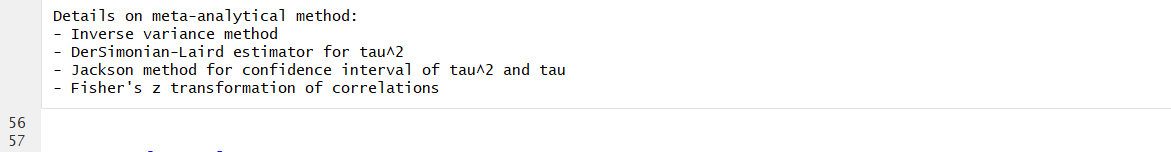
.27 222

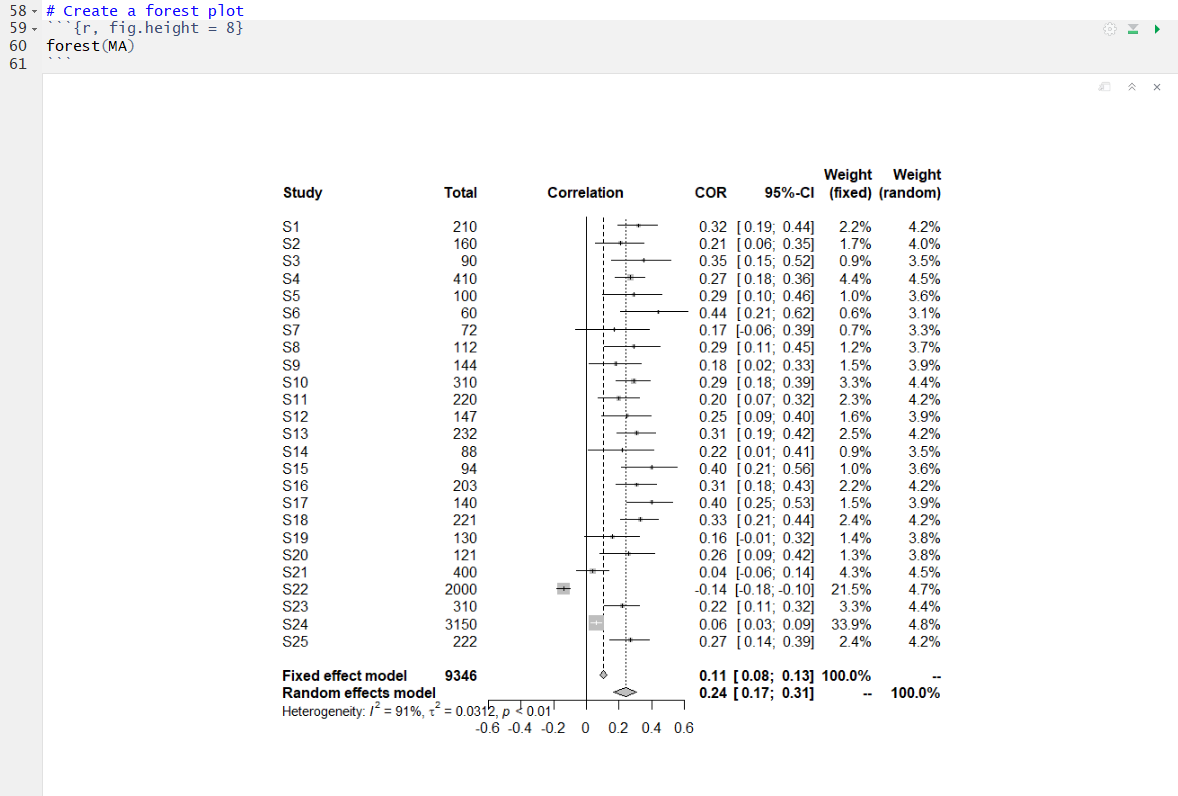
\_\_\_\_\_\_\_\_\_\_

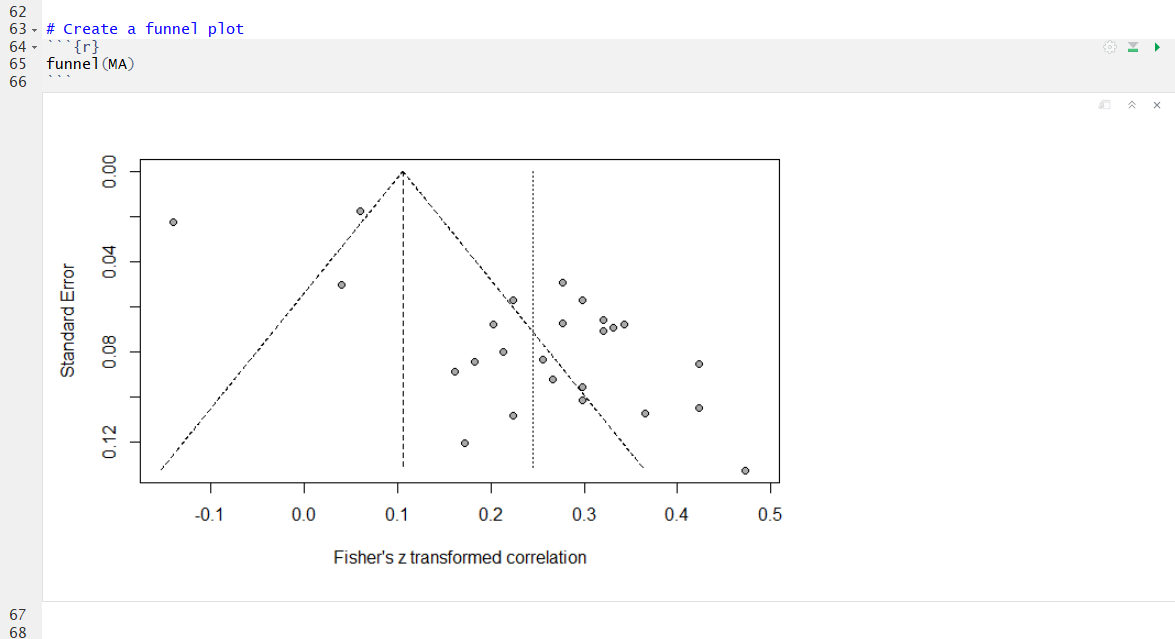
1. Using the new studies2.csv dataset, run a meta analyses and see how your results change.











* 1. What different conclusions do you make? Why?

With the new set of studies included, we reach different conclusions. Now, we see a large amount of heterogeneity across studies, this is shown by our large I2 value (90.9%). The I2 value is the percentage of true heterogeneity seen across studies. Ideally, we would like our I2 to be at least below 30%. Our forest plot shows that the last 5 studies have large samples and are skewing the effect size drastically. Thus, indicating that the last 5 studies added may be measuring a different effect and we should interpret the results with caution. Lastly, the funnel plot indicates a possibility of publication bias due to its skewness to the right. We would like this chart to ideally be symmetrical. If we were to report this, we would need to consider removing the studies that may be measuring a different effect and more investigation would be needed.