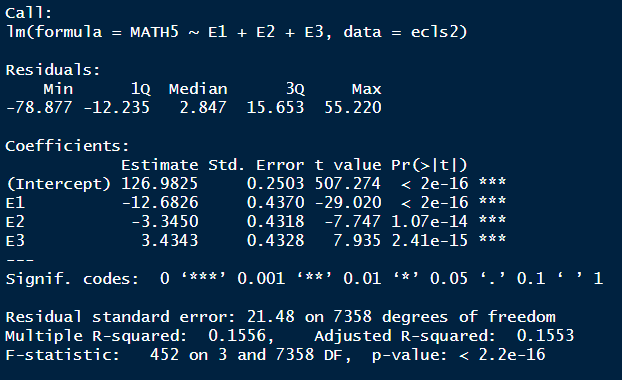
Hey All,

Thank you for being receptive yesterday as we went through the lab. I know that it is a lot and I wish that we could have a bit more time to make sure that you understand the content and get good practice in R, as well. I received some feedback about some things which I would like to address! My high school teacher personality is coming out.

1. **Pipelines:** Remember that pipelines are NOT a necessary tool for you to use, however you can use them to make adjustments to your code more readily. If you do not want to, we can take a piece of code and just redo it without pipelines. For example, we had the code that was associated with filtering out the “NA” rows (missing data) using the “type” column.   
     
   (…, data = Prestige %>% filter(is.na(type) == F))  
     
   This takes the Prestige data set and runs it through another function (filter) to “remove” the rows that do not have a type designation. It is row-wise deletion. We can instead do this with nested codes.  
     
   (…, data = filter(.data = Prestige, is.na(Prestige$type) == F))  
     
   This does the same thing and may be more understandable to you! If you wish to read more about pipelines, you can read it [here](https://www.datacamp.com/community/tutorials/pipe-r-tutorial). {<https://www.datacamp.com/community/tutorials/pipe-r-tutorial>}

1. **Dummy versus Deviation Coding:** In theory, this is fairly easy to understand; however you may be lost on how to interpret the analyses. While dummy coding is more reasonable (comparing each group to a pre-determined reference group), deviation coding is more practical and is often adopted in other software (SPSS). Instead of comparing to a reference group, each group is compared to the **grand mean**.   
     
   Why would you want that? If you think about the fact that all statistical analyses are trying to use a sample to speak to the overall population, then the grand mean is an estimate of the overall average of EVERYONE that is pertinent to your data sample. For example, in the ecls analysis, the grand mean would be an estimate of ALL STUDENTS’ fifth grade math score, that took the particular exam for that year. It’s a pretty important measure!  
     
   The deviation codes are comparing each group to the overall mean. Let’s look at the example from class really quickly.

Here, you can see the grand mean (overall estimate of 5th grade math score of everyone in the population that is represented by this sample) is 126.98 points. The group with the lowest SES category is 12.69 points lower than the overall average of everyone. The group 3 is 3.43 points higher than everyone and the SES group 4 has an average score of 126.98 – (-12.68) – (-3.345) – 3.434 = 139.58 points on the 5th grade math.

See how this helps and let me know if you need anything else!