

Here are some general comments for HW3. No one was penalized for anything related to these comments. These are just general thoughts. In some instances you were explicitly asked NOT to do something (like perform a thorough exploratory data analysis), but I included some comments regardless.

- In general, answer your question first and then show the R output second. It is just more convenient for the reader that way. I
- With exploratory data analysis (and residual analysis), tell me not only what you see but why its important. Are certain patterns in the data indicative of a larger problem? If so, what is it? How do these results inform your future analysis plan?
- Dig deeper and run EDA for each group or any combination of variables that make you suspicious. Do this for residuals too. In general, if you see any grouping variables (race, state etc.) you should examine them closely. They can be a great source of heteroskedasticity at the very least, and at the very worst you might see evidence that the data distributions are so different from one another that it makes you doubt that they should belong in the same regression.
- Interpret square term in general terms. What does the quadratic term tell us? It can be an indication of increasing returns, decreasing returns, or even reversal in directions. With respect to diagnosing whether or not there is a trend reversal, you have to do more than examine the coefficients. Some algebra or predicted values are needed.
 - Even if the quadratic term is statistically insignificant, you need to incorporate it into your answer as to the overall rate of return of experience on education.
- How to interpret race? Tricky because people cannot change from being white to being AA. Keep this in mind when thinking about the logic of causality. It might be easier to keep things correlational here. As in, “these results suggest that African American workers are paid more on average when starting a job, but their income drops relative to non-African Americans as experience increases”
- Work on having presentable tables, but this is not vital for this class.
- Rescaling comment: Sometimes it matters and sometimes it does not. In this case, it does matter, but sometimes it does not.

Here are some thoughts and comments from HW2 that I think apply to all of the submissions. These are more like guidelines and tips on how to think more deeply about a given problem.

Comments on HW 2

- Examining the distribution of your DV can inform which type of analysis you can run. Here, there is a spike at 100 and observations have an upper limit. This tells us two things. One, we might want to use a model that accounts for bounded observations (out of the scope of this class); two, there might be two different processes at play here. The first is the decision to participate at 100 or not and the second is how much to contribute if a person chooses not to match at 100.
- Why is scale location chart different than others?