

I was planning to go to Peru this week, things went awry. Going to see Godzilla vs Kong instead.

PS7 - Luke Denton

March 29, 2021

## Exercise 1

logwage is missing at 25 percent, and I think that this variable is missing not at random.

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
logwage	670	25	1.6	0.4	0.0	1.7	2.3
hgc	16	0	13.1	2.5	0	12.0	18
tenure	259	0	6.0	5.5	0.0	3.8	25.9
age	13	0	39.2	3.1	34	39.0	46

## Exercise 2

The iterated deletion of the na's for model 1 actually got closer to the true value of Beta 1 of hgc than Model 2 or 3, which I think is very interesting. Mean imputation was by far the worst method for calculating our Beta 1 value. Using predicted values got us close to the value we obtained for both model 1 and 4. The estimates in Model 3 and 4 were similar, although our multiple imputation model did a better job of predicting the true value of Beta 1.

	Model 1	Model 2	Model 3	Model 4
(Intercept)	0.534 (0.146)	0.708 (0.116)	0.563 (0.112)	0.495 (0.137)
hgc	0.062 (0.005)	0.050 (0.004)	0.059 (0.004)	0.064 (0.005)
collegenot college grad	0.145 (0.034)	0.169 (0.026)	0.178 (0.025)	0.145 (0.032)
tenure	0.050 (0.005)	0.038 (0.004)	0.047 (0.004)	0.050 (0.005)
tenuresq	-0.002 (0.000)	-0.001 (0.000)	-0.002 (0.000)	-0.002 (0.000)
age	0.000 (0.003)	0.000 (0.002)	0.000 (0.002)	0.001 (0.003)
marriedsingle	-0.022 (0.018)	-0.027 (0.014)	-0.028 (0.013)	-0.023 (0.019)
Num.Obs.	1669	2229	2229	2229
Num.Imp.				10
R2	0.208	0.146	0.223	0.233
R2 Adj.	0.206	0.144	0.221	0.231
AIC	1179.9	1093.8	957.6	
BIC	1223.2	1139.5	1003.2	
Log.Lik.	-581.936	-538.912	-470.776	
F	72.917	63.461	106.192	

## Exercise 3

We talked about this in class, but I haven't come up with a concrete idea for my project yet. I looked into the `gapmindr` package, but unfortunately it's not available under the current version of R. I want to do something international development related, ideally within the region of Latin America. The World Bank and UN Development Programme have great amounts of data, it just might be a bit tedious cleaning and converting from excel files to tibbles. I definitely want to include visualizations in my project this semester, as development reports typically include a lot of visualizations. But, for specific topics, I'm still undecided.