PS7 Dunkleberger

Davis Dunkleberger

March 2023

1 Initial Guesses

	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	0.0	12.0	18.0	
tenure	259	0	6.0	5.5	0.0	3.8	25.9	
age	13	0	39.2	3.1	34.0	39.0	46.0	

The log of wages are missing at a 25% rate. I think it is likely to be missing at random (MAR). I used datasummary_skim() to create that table.

2 Model Summary

	Complete Cases	Mean Imputation	Predicted Values	Multiple Imputation
(Intercept)	0.534***	0.708***	0.534***	0.649***
	(0.146)	(0.116)	(0.112)	(0.144)
hgc	0.062***	0.050***	0.062***	0.057***
	(0.005)	(0.004)	(0.004)	(0.005)
collegenot college grad	0.145***	0.168***	0.145***	0.088**
	(0.034)	(0.026)	(0.025)	(0.032)
tenure	0.050***	0.038***	0.050***	0.044***
	(0.005)	(0.004)	(0.004)	(0.006)
I(tenure^2)	−0.002***	−0.001***	−0.002***	−0.001**

It appears the complete cases and predicted values are the same formula. That could be an error in my code or it could be that the formulas for those regressions are similar. There seems to be a significance between the models and what the actual value is. These models that I created do not seem to be close to the true value of 0.093. I do not think these models are the most accurate or true. The estimates are very close to each other. There might not be too much of a difference between the estimations in those estimates.

3 Final Project

I have not started thinking in earnest about what I am going to do with the final project. I do know I want to use the hockeyR package for the data. I might do something with goal numbers based on what I encountered in PS6. Finding a relationship or predicting goals for the next year could be interesting. I have not given it much thought beyond that honestly.