empirical_proj_2

Lindsey Greenhill

4/9/2022

Question 1

The fundamental problem of causal inference is that you cannot see the outcome for a child who both was in a small class and not in a small class. In other words, it is impossible to measure the dual outcomes when a child both is in a small class and not in a small class, because the child can only either be in a small class or not in a small class.

Question 2

The following variables have missing values:

towncode: 318 missingmath: 58 missingverb: 466 missingses_index: 318 missing

• boy: 1037 missing

##	student_id		towncode		schlcode		class_id	
##	Min. :	1	Min.	: 166	Min.	:11005	Min.	:110051
##	1st Qu.:1	3600	1st Qu.	:26104	1st Qu.	:31049	1st Qu.	:310491
##	Median :2	7412	Median	:62000	Median	:41284	Median	:412841
##	Mean :2	7395	Mean	:50434	Mean	:39776	Mean	:397762
##	3rd Qu.:4	1158	3rd Qu.	:79004	3rd Qu.	:51247	3rd Qu.	:512471
##	Max. :5	4860	Max.	:98004	Max.	:61365	Max.	:613651
##			NA's	:318				
##	school_enrollmen		nt class_size		math		verb	
##	Min. :	9.00	Min.	: 7.00	Min.	: 1.0	Min.	: 1.00
##	1st Qu.:	57.00	1st Q	u.:28.00	1st Q	u.:23.0	1st Qu	.:25.00
##	Median :	75.00	Media	n :32.00	Media	n:27.0	Median	:27.00
##	Mean :	80.41	Mean	:31.58	Mean	:25.3	Mean	:25.97
##	3rd Qu.:105.00		3rd Qu.:36.00		3rd Qu.:29.0		3rd Qu.:29.00	
##	Max. :1	88.00	Max.	:44.00	${\tt Max.}$:30.0	Max.	:30.00
##					NA's	:58	NA's	:466
##	ses_index		boy		born_isr		religious	
##	Min. :	0.00	Min.	:0.0000	Min.	:0.0000	Min.	:0.0000
##	1st Qu.: 4.00		1st Qu.:0.0000		1st Qu.:1.0000		1st Qu.:0.0000	
##	Median :	9.00	Median	:1.0000	Median	:1.0000	Media	n:0.0000
##	Mean :1	3.17	Mean	:0.5078	Mean	:0.9608	Mean	:0.2242
##	3rd Qu.:1	8.00	3rd Qu.	:1.0000	3rd Qu	1.:1.0000	3rd Q	u.:0.0000
##	Max. :7	6.00	Max.	:1.0000	Max.	:1.0000	Max.	:1.0000
##	NA's :3	18	NA's	:1037				