Ex-11.5.1

Note: I have given the terms in the dataset more sensible names to read at a glance, for myself. In the source of this document there is a `subset` call, followed by a `colnames` call which should act as a key for this change when looking into my work. I have chosen to work with grade 3 arbitrarily.

1. Suppose that we are interested in using this dataset to evaluate the impact of class size on test scores (we have data on reading and math). Based on the data description, would we be able to estimate a causal effect of class size on these test scores? Explain why or why not.

Tentatively, sort-of. Looking at the data, we have a stand-in for class size, not actually class size. This isn't as good as just directly having the exact number of people in the class. We also have lots of other terms which may be confounders, or would otherwise interfere with isolating the causal effect if we couldn't control for them.

The other thing we would need to do is evaluate an ADG showing which terms we should include and which not.

If we were not worried about the causal effect, and only wished to know the correlation, then we would include pretty much every term. The question is, does the ADG change that? Let's find out.

The terms I am most worried about are ethnicity. Ethnicity, unfortunately, is still a major determination of your lot in life in this country. Lots of racist people, lots of racist institutions. So, I think ethnicity causally effects a lot of things. This will be reflected in the graph. Generally, ethnicity will determine some part of location and opportunity. `teacher_highest_degree` may be lower, if they were not given good opportunities in collage because of their ethnicity; `school_system` may be different as some ethnicities are more likely to live in low-income areas; etc. . .

I believe that `school_system`, `school_type`, and `school` may have some determination on `star_class_size`. The population of a given area will drastically change the sizes of each class.

However, it is *also* the case that test scores are effected by the quality of schooling at those schools, and even which teachers are teaching at those schools.

I propose including `school_system`, `school_type`, and `school` along with `star_class_size`, as a minimal model for isolating the causal effect of `star_class_size` on test scores.

There are no coliders in this model, and also no descendants of `star_class_size`. The only value I do not believe has any correlation with on scores would be `sex`, but it wouldn't hurt to include it anyway, as it's possible that old-fashioned societal pressures may pressure people to do better or worse depending on their sex. So the kitchen-sink approach would not be unfounded either, as opposed to the minimal model.
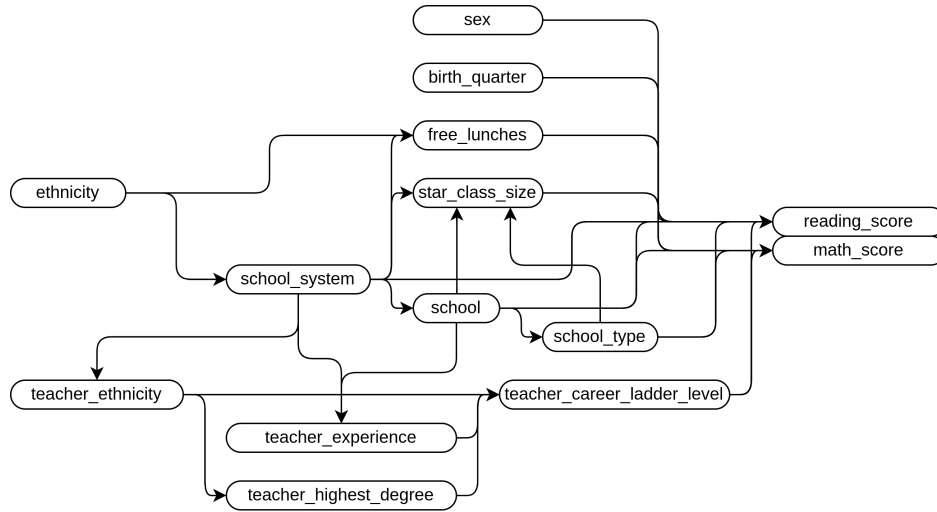
(graph appears on next page →)

Figure 1: an acyclic directed graph representing the ideas above.

2. Using `star_class_size` as the explanatory variable, use linear regression to estimate the effect of class size on test scores. Interpret the results.

I believe you intend for me to *not* use either my minimal model, or the kitchen sink model, as you use the word "variable" as a singular. So the model I will use is: $score \sim \beta_0 + \beta_1 star\_class\_size$

Table 1:

	Dependent variable:	
	math_score	reading_score
	(1)	(2)
star__class__sizeregular+aide	0.951	−0.100
	(2.021)	(1.879)
star__class__sizesmall	4.816**	5.611***
	(1.968)	(1.830)
Constant	624.731***	622.957***
	(1.537)	(1.429)
Observations	2,678	2,678
R^2	0.003	0.006
Adjusted R^2	0.002	0.005
Residual Std. Error (df = 2675)	40.097	37.290
F Statistic (df = 2; 2675)	3.740**	7.412***

Note:	*p<0.1; **p<0.05; ***p<0.01

Test scores in both reading and math noticeably and significantly increase given a smaller class size. Adding a teaching assistant does not produce statistically significant results to analyze one way or the other.

3. Produce a plot that shows the effect of class size on test scores

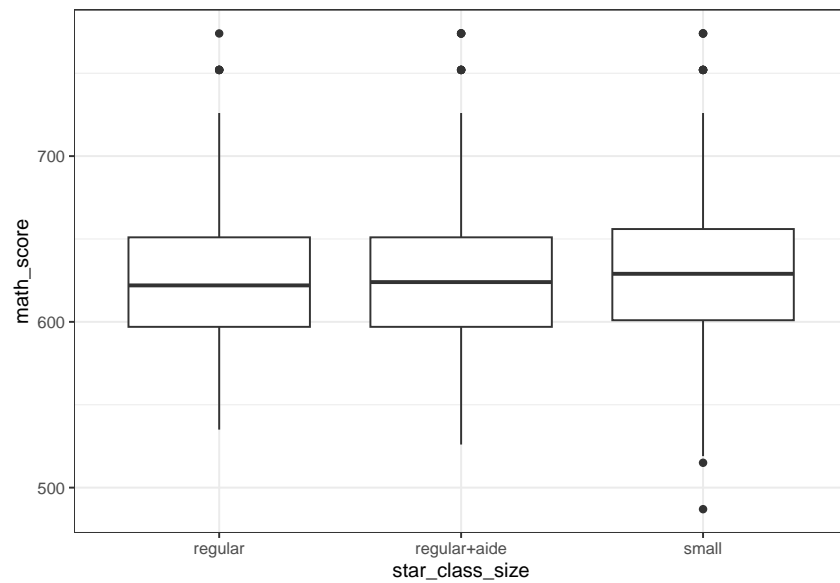Below are two box-plots showing test scores based on class size:

Figure 2: a box-plot showing the relationship between class size and math test scores.
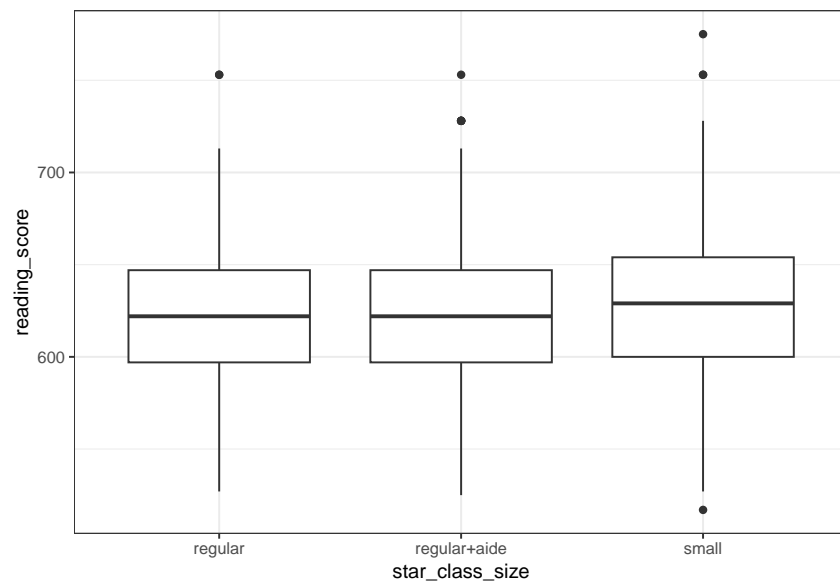

Figure 3: a box-plot showing the relationship between class size and reading test scores.

You'll notice in both plots a small, yet noticeable increase in both the median and the first and third quartiles when looking at the smaller class size vs the normal one.

4. Discuss why you might want to include the variable `birth_quarter` as a control variable. Explain what problem it is solving. Make reference to your answer in part 1.

`birth_quarter` may contribute to test scores for a number of reasons. It could represent students who were held back, students who are younger and thus less experienced (admittedly by a small margin) etc... It also has no effect on `star_class_size`. While including it in the model wouldn't remove any bias, it would increase accuracy.

5. Based on your model in part 2, create a plot that could show heteroskedasticity. Re-estimate your regression in part 2 using heteroskedasticity-robust standard errors and explain whether they change any conclusions.

So, it's at this point I really wish we had a continuous variable for class size, and not the categorical one. However, we work with what we have.

I believe the box plots above appropriately show that if there may in fact be some heteroskedasticity. The box is (admittedly very slightly) wider for smaller classes.

I'll be switching to the minimal causality model from question 1.

The following are `coeftests` for both models:

```
##
## t test of coefficients:
##
##                              Estimate Std. Error  t value  Pr(>|t|)
## (Intercept)                 604.92605    2.34563 257.8951 < 2.2e-16 ***
## star_class_sizeregular+aide   0.75809    1.90730   0.3975  0.691054
## star_class_sizesmall          5.50337    1.87567   2.9341  0.003374 **
## school_system                 2.01018    0.29776   6.7511 1.794e-11 ***
## school_typerural             20.62351    2.26906   9.0890 < 2.2e-16 ***
## school_typesuburban          22.00201    2.54106   8.6586 < 2.2e-16 ***
## school_typeurban             16.55154    3.18614   5.1949 2.203e-07 ***
## school                       -0.89738    0.14309  -6.2713 4.162e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## t test of coefficients:
##
##                              Estimate Std. Error  t value  Pr(>|t|)
## (Intercept)                 602.70497    2.25263 267.5566 < 2.2e-16 ***
## star_class_sizeregular+aide  -0.35020    1.78001  -0.1967 0.8440441
## star_class_sizesmall          6.10690    1.76869   3.4528 0.0005635 ***
## school_system                 1.50502    0.27416   5.4896 4.408e-08 ***
## school_typerural             19.32542    2.20145   8.7785 < 2.2e-16 ***
## school_typesuburban          21.42045    2.38649   8.9757 < 2.2e-16 ***
## school_typeurban             17.69536    3.14923   5.6190 2.120e-08 ***
## school                       -0.62870    0.13136  -4.7859 1.795e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The conclusions should not change, as the findings are still statistically significant.

6. Is it possible to estimate clustered standard errors using this dataset? Explain why or why not. If you can, explain why you might want to, and which variable(s) you would use.

There are several clusters which could be used. Same `school`, `school_type`, `school_system`, `ethnicity`, etc... Any or all of these categorical variables may exhibit some clustering, and could be used to estimate clustered standard errors.

7. Test the following hypotheses:

    a. Class size has no effect on test scores.

The f-statistic for mathematics, restricting only the `star_class_size` variable, is 11.0096615. For reading, it is 19.466645. (Derived from the tables below.)

The critical value for math is 3.844944, and for reading it is 3.844944. Those values are both well below the f-statistics, and thus we can reject the null.

tables showing the information from which this was all derived on following pages $\rightarrow$

    b. The effect of class size on test scores depends on the highest degree earned of the teacher.

New model: $score \sim \beta_0 + \beta_1 star\_class\_size + \beta_2 teacher\_highest\_degree + \beta_3 star\_class\_size\,teacher\_highest\_degree$

The f-statistic for mathematics, restricting the interaction between `star_class_size` and `teacher_highest_degree`, is 3.903546. For reading, it is 2.3613068. (Derived from the tables below.)

The critical value for math is 2.9990985, and for reading it is 2.9990985. For mathematics, the f-statistic is above the critical value, and so we may want to reject the null, and the opposite is true of reading.

However. In both mathematics and reading several coefficients lose statistical significance, so the interpretation of that model is less concrete.

tables showing the information from which this was all derived on following pages $\rightarrow$

    c. The effect of class size on test scores depends on the years of the teacher's total teaching experience.

New model: $score \sim \beta_0 + \beta_1 star\_class\_size + \beta_2 teacher\_experience + \beta_3 star\_class\_size\,teacher\_experience$

The f-statistic for mathematics, restricting the interaction between `star_class_size` and `teacher_highest_degree`, is 4.8632981. For reading, it is 2.1881984. (Derived from the tables below.)

The critical value for math is 2.9990935, and for reading it is 2.9990935. For mathematics, the f-statistic is above the critical value, and so we may want to reject the null, and the opposite is true of reading.

However. In both mathematics and reading several coefficients lose statistical significance, so the interpretation of that model is less concrete.

This is the same situation as 7.b.

tables showing the information from which this was all derived on following pages $\rightarrow$

Table 2:

	math_score	
	Dependent variable:	
	(1)	(2)
star_class_sizeregular+aide	0.758	
	(1.958)	
star_class_sizesmall	5.503***	
	(1.909)	
school_system	2.010***	1.997***
	(0.299)	(0.300)
school_typerural	20.624***	20.356***
	(2.381)	(2.384)
school_typesuburban	22.002***	22.062***
	(2.597)	(2.601)
school_typeurban	16.552***	16.811***
	(3.359)	(3.364)
school	−0.897***	−0.890***
	(0.143)	(0.143)
Constant	604.926***	607.467***
	(2.508)	(2.166)
Observations	2,678	2,678
R^2	0.067	0.064
Adjusted R^2	0.065	0.062
Residual Std. Error	38.813 (df = 2670)	38.878 (df = 2672)
F Statistic	27.565*** (df = 7; 2670)	36.266*** (df = 5; 2672)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3:

	reading_score	
	Dependent variable:	
	(1)	(2)
star_class_sizeregular+aide	−0.350	
	(1.823)	
star_class_sizesmall	6.107***	
	(1.778)	
school_system	1.505***	1.487***
	(0.279)	(0.280)
school_typerural	19.325***	19.017***
	(2.218)	(2.223)
school_typesuburban	21.420***	21.503***
	(2.418)	(2.426)
school_typeurban	17.695***	18.015***
	(3.129)	(3.138)
school	−0.629***	−0.620***
	(0.133)	(0.134)
Constant	602.705***	605.146***
	(2.335)	(2.021)
Observations	2,678	2,678
R^2	0.067	0.061
Adjusted R^2	0.065	0.059
Residual Std. Error	36.146 (df = 2670)	36.264 (df = 2672)
F Statistic	27.531*** (df = 7; 2670)	34.424*** (df = 5; 2672)

Note: *p<0.1; **p<0.05; ***p<0.01

	Dependent variable:	
	math_score	
	(1)	(2)
star_class_sizeregular+aide	−40.643**	0.742
	(17.707)	(2.021)
star_class_sizesmall	−7.226	4.550**
	(15.717)	(1.968)
teacher_highest_degreebachelor	−58.002***	−44.332***
	(19.905)	(12.133)
teacher_highest_degreemaster	−50.561**	−40.777***
	(19.948)	(12.157)
teacher_highest_degreespecialist	−33.583**	−35.701**
	(16.677)	(13.909)
star_class_sizeregular+aide:teacher_highest_degreebachelor	43.835**	
	(17.917)	
star_class_sizesmall:teacher_highest_degreebachelor	14.419	
	(15.924)	
star_class_sizeregular+aide:teacher_highest_degreemaster	39.306**	
	(17.967)	
star_class_sizesmall:teacher_highest_degreemaster	8.500	
	(16.027)	
star_class_sizeregular+aide:teacher_highest_degreespecialist		
star_class_sizesmall:teacher_highest_degreespecialist		
Constant	679.226***	667.450***
	(19.802)	(12.214)
Observations	2,678	2,678
R^2	0.013	0.010
Adjusted R^2	0.009	0.008
Residual Std. Error	39.952 (df = 2668)	39.981 (df = 2672)
F Statistic	3.779*** (df = 9; 2668)	5.234*** (df = 5; 2672)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5:

	reading_score	
	Dependent variable:	
	(1)	(2)
star_class_sizeregular+aide	0.554	0.118
	(16.488)	(1.881)
star_class_sizesmall	28.095*	5.283***
	(14.635)	(1.831)
teacher_highest_degreebachelor	−14.881	−38.106***
	(18.535)	(11.291)
teacher_highest_degreemaster	−16.497	−39.138***
	(18.575)	(11.314)
teacher_highest_degreespecialist	−12.879	−27.537**
	(15.529)	(12.944)
star_class_sizeregular+aide:teacher_highest_degreebachelor	−2.022	
	(16.684)	
star_class_sizesmall:teacher_highest_degreebachelor	−22.627	
	(14.828)	
star_class_sizeregular+aide:teacher_highest_degreemaster	1.062	
	(16.730)	
star_class_sizesmall:teacher_highest_degreemaster	−24.058	
	(14.923)	
star_class_sizeregular+aide:teacher_highest_degreespecialist		
star_class_sizesmall:teacher_highest_degreespecialist		
Constant	638.450***	661.263***
	(18.439)	(11.367)
Observations	2,678	2,678
R^2	0.013	0.011
Adjusted R^2	0.009	0.009
Residual Std. Error	37.202 (df = 2668)	37.207 (df = 2672)
F Statistic	3.847*** (df = 9; 2668)	5.978*** (df = 5; 2672)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 6:

	math_score	
	Dependent variable:	
	(1)	(2)
star_class_sizeregular+aide	8.920**	0.945
	(3.858)	(2.021)
star_class_sizesmall	4.594	4.833**
	(3.702)	(1.969)
teacher_experience	0.225	0.032
	(0.184)	(0.091)
star_class_sizeregular+aide:teacher_experience	−0.578**	
	(0.239)	
star_class_sizesmall:teacher_experience	0.026	
	(0.233)	
Constant	621.659***	624.297***
	(2.939)	(1.978)
Observations	2,678	2,678
R^2	0.006	0.003
Adjusted R^2	0.005	0.002
Residual Std. Error	40.046 (df = 2672)	40.104 (df = 2674)
F Statistic	3.469*** (df = 5; 2672)	2.533* (df = 3; 2674)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 7:

	Dependent variable:	
	reading_score	
	(1)	(2)
star_class_sizeregular+aide	2.961	−0.121
	(3.590)	(1.879)
star_class_sizesmall	3.217	5.669***
	(3.445)	(1.830)
teacher_experience	0.112	0.109
	(0.171)	(0.085)
star_class_sizeregular+aide:teacher_experience	−0.223	
	(0.223)	
star_class_sizesmall:teacher_experience	0.187	
	(0.217)	
Constant	621.424***	621.468***
	(2.735)	(1.839)
Observations	2,678	2,678
R^2	0.008	0.006
Adjusted R^2	0.006	0.005
Residual Std. Error	37.269 (df = 2672)	37.285 (df = 2674)
F Statistic	4.174*** (df = 5; 2672)	5.493*** (df = 3; 2674)

Note: *p<0.1; **p<0.05; ***p<0.01

8. Students' assignment to the class size groups was not held constant across the four years. Write a line or two of code that shows that this was the case.

```
star_filtered <- star |> filter(
  star_class_size!=star_class_size_k |
    star_class_size!=star_class_size_f |
    star_class_size!=star_class_size_s |
    star_class_size_k!=star_class_size_f |
    star_class_size_k!=star_class_size_s |
    star_class_size_f!=star_class_size_s)

num_changed <- nrow(star_filtered)
```

It can be shown that 1232 people changed class size.

9. What is the effect of class size in kindergarten on test scores in grade 1? How much of this effect can be attributable to the effect of kindergarten test scores on grade 1 test scores?

New model: $score_f \sim \beta_0 + \beta_1 star\_class\_size\_k + \beta_2 score\_k + \beta_3 star\_class\_size\_k \cdot score\_k$

The effect of k score on first grade score is much smaller than the effect of k class size on first grade score, if we allow for interaction between the two variables. Otherwise, the effects of either are about the same.

Table 8:

	math_score_f	
	Dependent variable:	
	(1)	(2)
star_class_size_kregular+aide	−12.072	0.262
	(17.843)	(1.559)
star_class_size_ksmall	17.934	4.646***
	(17.654)	(1.595)
math_score_k	0.494***	0.492***
	(0.025)	(0.015)
star_class_size_kregular+aide:math_score_k	0.025	
	(0.036)	
star_class_size_ksmall:math_score_k	−0.026	
	(0.035)	
Constant	297.047***	297.627***
	(12.261)	(7.315)
Observations	2,678	2,678
R^2	0.309	0.308
Adjusted R^2	0.308	0.308
Residual Std. Error	33.358 (df = 2672)	33.359 (df = 2674)
F Statistic	238.963*** (df = 5; 2672)	397.600*** (df = 3; 2674)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 9:

	reading_score_f	
	Dependent variable:	
	(1)	(2)
star_class_size_kregular+aide	66.931**	0.174
	(28.950)	(2.034)
star_class_size_ksmall	62.599**	3.500*
	(29.546)	(2.080)
reading_score_k	1.014***	0.920***
	(0.046)	(0.027)
star_class_size_kregular+aide:reading_score_k	−0.151**	
	(0.065)	
star_class_size_ksmall:reading_score_k	−0.133**	
	(0.066)	
Constant	87.682***	129.529***
	(20.573)	(11.975)
Observations	2,678	2,678
R^2	0.311	0.310
Adjusted R^2	0.310	0.309
Residual Std. Error	43.490 (df = 2672)	43.525 (df = 2674)
F Statistic	241.752*** (df = 5; 2672)	400.165*** (df = 3; 2674)

Note: *p<0.1; **p<0.05; ***p<0.01