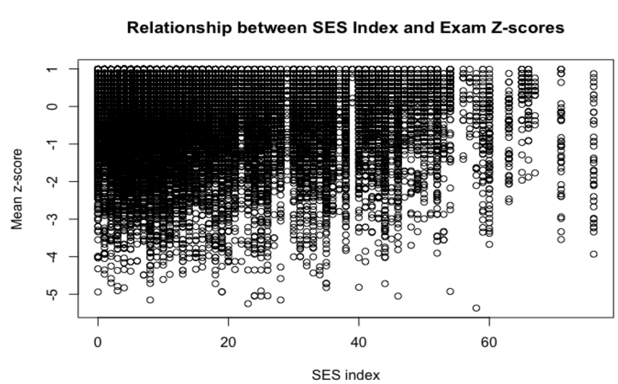
1. I remember my elementary-school classes having about 30 students on average. This is pretty consistent with the California state average of 29.7 students in an elementary class (NCES).

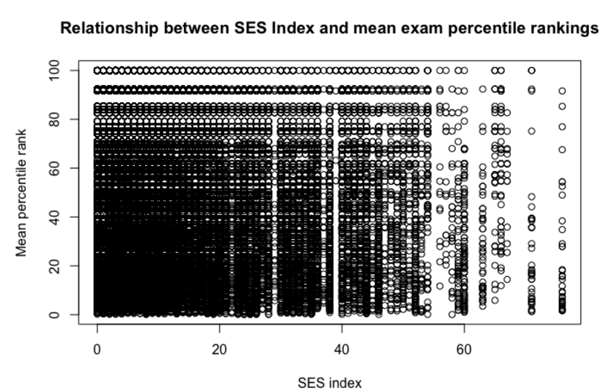
2. Summary statistics for math scores (third_grade$math), verbal scores (third_grade$verb), test Z-scores (third_grade$zscore), and test percentile ranks (third_grade$perc) are displayed below:

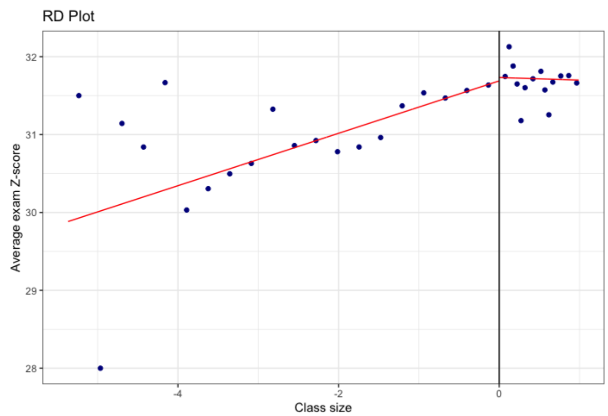
```
> summary(third_grade$math)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
    1.0    23.0    27.0    25.3    29.0    30.0    3668 
> summary(third_grade$verb)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
   1.00   25.00   27.00   25.97   29.00   30.00    4076 
```

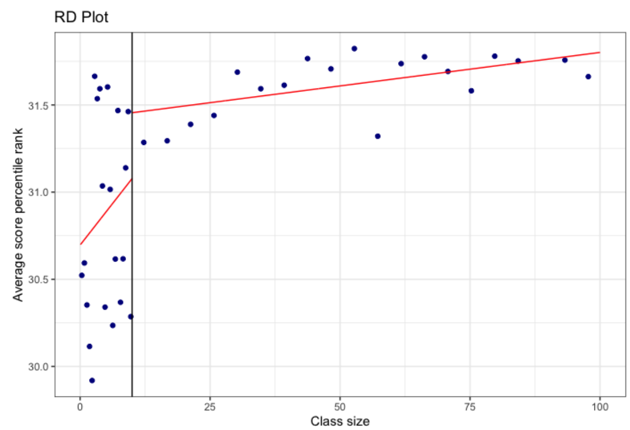
```
> summary(third_grade$zscore)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
 -5.368  -0.374   0.295   0.003   0.652   0.990    4134 
> summary(third_grade$perc)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
  0.078  31.499  54.652  53.815  76.952 100.000    4134 
```

3.





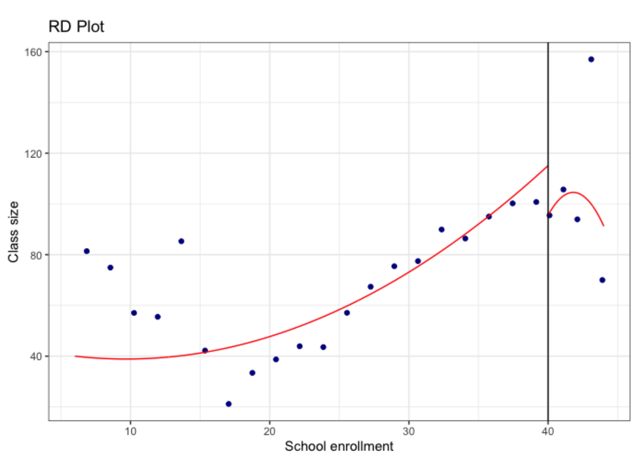


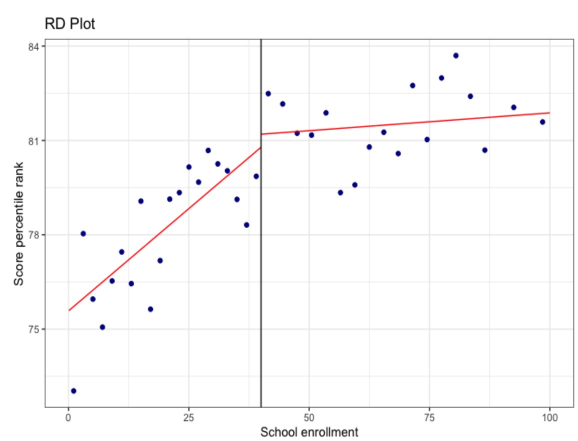


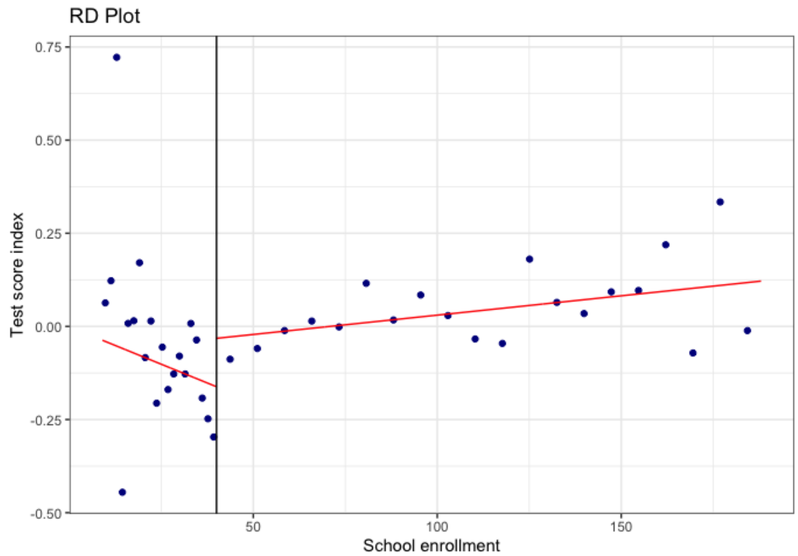
From the regression-discontinuity scatterplots above, we observe a stark difference in the slopes of the linear fits on either side of a class-size cutoff. So far, we cannot draw any conclusions about causation, but this difference in slope may suggest a causal effect on exam performance from class size.

4. Causal inference is fundamentally a missing-data problem due to the fact that we observe the outcomes of only one treatment per observation. The results of other treatments inherently cannot be observed and are therefore "missing data," and we must employ causal inference to infer what these missing data points are likely to be.

5.







6. The first two models had a prediction difference of -15.869. This indicates that the second model, which was fit above the cutoff point, predicts a class size that is about 16 students smaller than the prediction of the first model, which was fit below the cutoff, for a school with 40 students enrolled. The next two models had a prediction difference of 0.0647, indicating that the second model predicted an average exam z-score that was 0.0647 units higher than the first model's prediction for a 40-student school. The last two models had a prediction difference of 1.747, indicating the second model predicted an average exam ranking that was 1.747 percentile points higher than the first model's prediction for a 40-student school.

7.

```
> confint(mod1)
                       2.5 %     97.5 %
(Intercept)       16.7319743 17.1021915
school_enrollment  0.2294545  0.2358401
> confint(mod2)
                         2.5 %       97.5 %
(Intercept)       -0.235191647 -0.161520937
school_enrollment  0.002371705  0.003640943
> confint(mod3)
                        2.5 %      97.5 %
(Intercept)       48.05671730 50.27485820
school_enrollment  0.04669453  0.08490984
```

Because none of these confidence intervals for the school_enrollment coefficient contain 0, these results are statistically significant from 0 at the 0.05 confidence level.

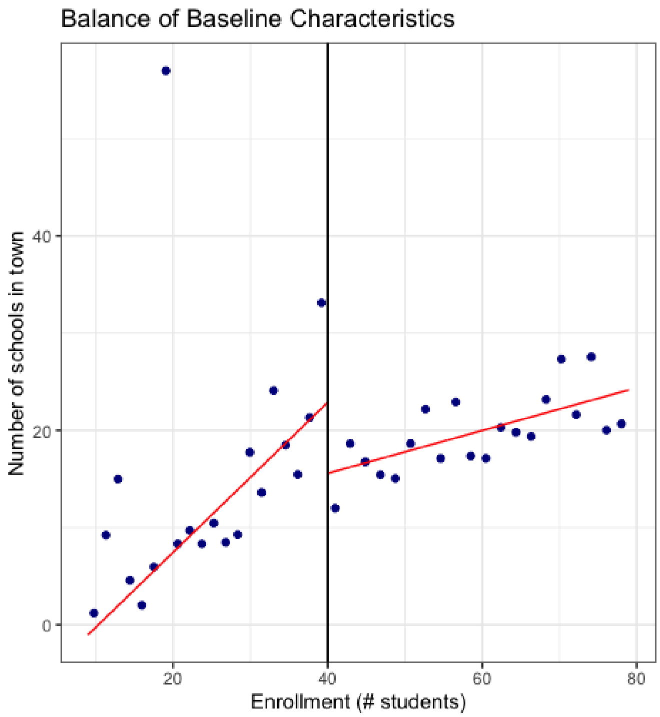
8. Assuming no more than 40 students per class and assuming no classes added until after reaching the maximum class size, the average class size would drop from 40 students to 20.5 students at the 40-student threshold. (Angrist and Lavy 535) This is not exactly observed in the data because schools can usually "afford to add extra classes before reaching the maximum class size." (542)

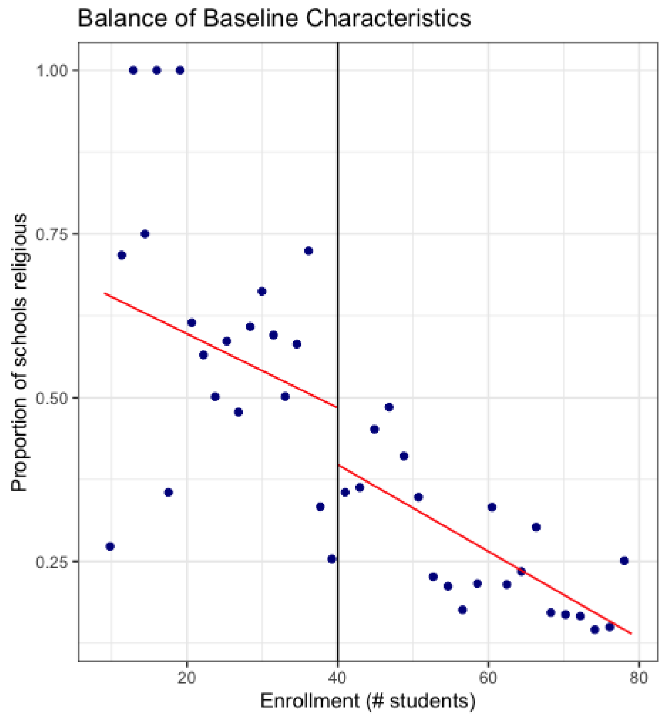
9. We expect a decrease of about 0.039 units in the average exam z-score and a decrease of about 0.958 percentiles in the average exam percentile ranking associated with a class-size decrease from 40 to 35 students.

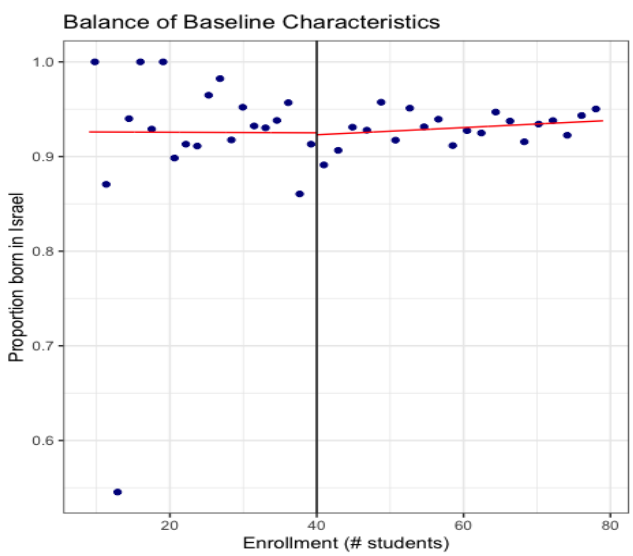
10. Yes, I would feel confident making the same prediction. This change from 20 to 15 students is an equivalent drop in the number of students to the scenario in question 9, and this drop in students from 20 to 15 is still in the range of the data used to fit the linear model from the previous question. Therefore, the results of that model are generalizable to this change in class size.

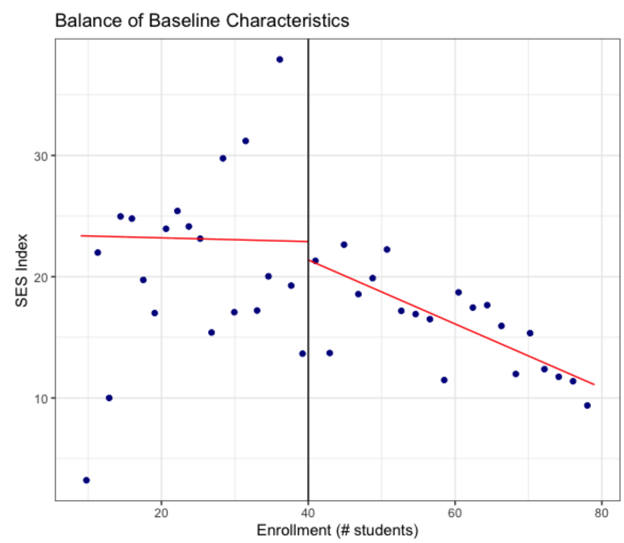
11. The Tennessee data's estimated change in test percentile ranking of 5.0026 is significantly larger than the estimate of 0.958 from this dataset. However, the confidence intervals generated using this dataset do not have an impact on the Tennessee results at all. The data used to build our confidence intervals represents a fundamentally different sample population - Israeli schools - from the Tennessee dataset.

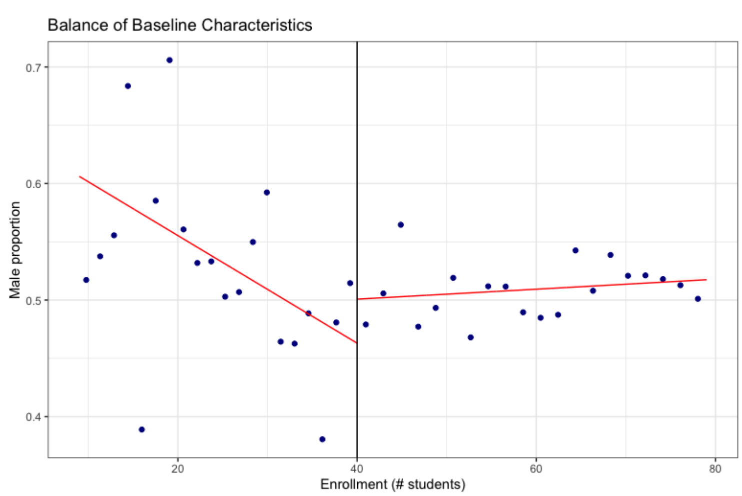
12.

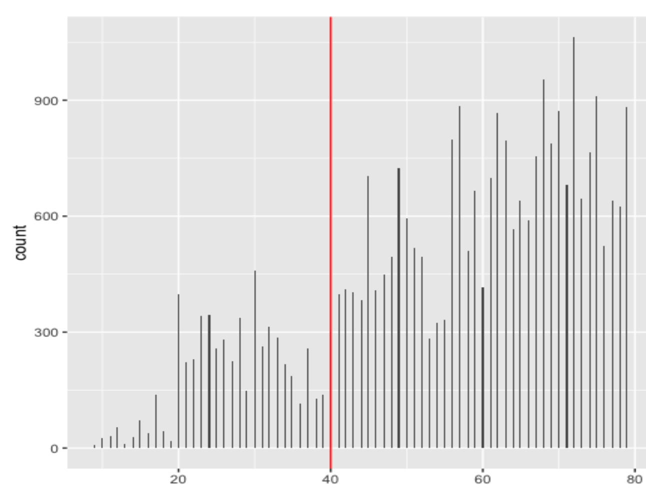

Balance of Baseline Characteristics
Number of schools in town
0
20
40
20
40
60
80
Enrollment (# students)


Balance of Baseline Characteristics
Proportion of schools religious
0.25
0.50
0.75
1.00
20
40
60
80
Enrollment (# students)


Balance of Baseline Characteristics
Proportion born in Israel
0.6
0.7
0.8
0.9
1.0
20
40
60
80
Enrollment (# students)


Balance of Baseline Characteristics
SES Index
10
20
30
20
40
60
80
Enrollment (# students)


Balance of Baseline Characteristics
Male proportion
0.4
0.5
0.6
0.7
20
40
60
80
Enrollment (# students)


count
0
300
600
900
20
40
60
80
school_enrollment

13. For a regression discontinuity analysis to be appropriate, the observations directly on either side of the discontinuity cutoff should be comparable. From the discontinuity plots above, we notice that the identification assumption appears to be satisfied. The baseline characteristics seem to evolve relatively smoothly across the 40-student enrollment threshold.

14. I disagree with this statement. The statement mistakenly concludes that the plot demonstrates a causal relationship between enrollment and test scores; however, from the plot, we only observe a positive association between enrollment and test scores. Correlation does not imply causation, so it would be erroneous to conclude a causal link based solely on this plot.

15. Based on the statistically significant results from this project and from the STAR experiment, it seems plausible that there may be a positive causal link between increased class size and test scores. However, I would not immediately conclude that being in a bigger class would have improved my progress in school. The results from these datasets are only generalizable to the populations they represent, and it is entirely possible that for someone like myself, who did not grow up in either Tennessee or Israel, these results would not apply. Nevertheless, the results of these two datasets confirm each other and may provide important insights into the effect of class size on academic performance.

Works Cited

Angrist, Joshua, and Victor Lavy. “Using Maimonides' Rule to Estimate the Effect of Class Size on Student Achievement.” *The Quarterly Journal of Economics*, May 1999, doi:10.3386/w5888.

“Schools and Staffing Survey (SASS).” *National Center for Education Statistics (NCES) Home Page, a Part of the U.S. Department of Education*, nces.ed.gov/surveys/sass/tables/sass0708_2009324_t1s_08.asp.