Effects of Educational Television

In this exercise we're going to look at the effect of a educational television program The Electric Company that ran from 1971-77 on children's reading scores. We will investigate what reading gains, if any, were made by the 1st through 4th grade classes as part of a randomized experiment.

This exercise is based on:

> Joan G. Cooney (1976) The Electric Company: Television and Reading, 1971-1980: A Mid-Experiment Appraisal. Children's Television Network Report.

The data comes from a two location trial in which treatment was randomized at the level of school classes.[1] Each class was either treated (to watch the program) or control (to not watch the program). The outcome of interest is the score on a reading test administered at the end of each year called `post.score`. Note that these are distinct classes from all four years. The variables are:

Name	Description
`pair`	The index of the treated and control pair (ignored here).
`city`	The city: Fresno ("F") or Youngstown ("Y")
`grade`	Grade (1 through 4)
`supp`	Whether the program replaced ("R") or supplemented ("S") a reading activity
`treatment`	"T" if the class was treated, "C" otherwise (randomized)
`pre.score`	Class reading score *before* treatment, at the beginning of the school year
`post.score`	Class reading score at the end of the school year

Question 1

Read the data into an data frame named `electric`. What sort of variable has R assumed `grade` is? How will it be treated in a linear model? Under what circumstances would that be reasonable or unreasonable?

Make a new variable from `grade` that is a factor. How will a linear model treat this new variable? **Hint:** You may find that `summary` illuminates the new data set.

Finally, overwrite the existing treatment variable so that it is numerical: 1 when the class is treated and 0 when not.

Question 2

Let's now consider the effect of treatment. First, fit a linear model that predict `post.score` with just treatment. Then fit a model uses your factor version of `grade` as well as treatment.

Summarise both models in terms of how much of the variance in `post.score` they "explain" and the median size of their errors.

Now, consider each model's treatment coefficient. Are the estimates of this coefficient *different* in the two models? Why do you think that is?

[1] Classes were paired, but we will ignore that in the analysis

Question 3

(Optional). In the previous question we saw that the models agreed about the coefficient estimate.
This is a very rare thing in observational data, but it happens in when experimenters have carefully arranged features of the experiment to be "balanced" with respect to treatment. For example, the experimental design of this study is to have equal number of classes in treatment and in control within each grade. This makes the treatment indicator and grade indicators independent and therefore uncorrelated.

To investigate this further, first compute the correlation between grade and treatment assignment, and then make a table of these two variables. How does the table structure explain the correlation?

Compare this to the correlation of `post.score` and `treatment`.

Compute the average `post.score` for each grade. How does this this explain the correlation between post.score and treatment?

Why would it be helpful to "balance" variables like grade with respect to treatment in this way? **Hint:** think about our strategies for causal inference.

Question 4

Now make another model that uses the factor version of `grade` and `pre.score` (the reading score before the year begins) to predict `post.score`. Is this model better? If so, in what ways?

Question 5

Now let's consider the effect of treatment *within* each grade. We can use the `lm` function's `subset` argument to fit the model on just a subset of all the rows in the data set. For example, we can fit a model of the relationship of `post.score` to `treatment` and `pre.score` just in grade 2 like this:

```
mod <- lm(post.score ~ treatment + pre.score, data = electric,
          subset = grade == 2)
```

This is equivalent to

```
electric_grade2 <- filter(electric, grade == 2)
mod <- lm(post.score ~ treatment + pre.score,
          data = electric_grade2)
```

but a bit shorter to type.

Fit a linear model predicting `post.score` using treatment and `pre.score` for each grade. You can use either of the strategies above. There are now *four* treatment effects. How do they differ as grade increases?

Are these ATEs? If so, for which *population* are they ATEs for? What do we call ATEs for specific values of pre-treatment variables?

Question 6

Now let's try to learn about separate grade effects in a single model. One way to do this is to *interact* treatment with grade. Here's a general modeling principle:

> If you think the *effect* of variable A varies according to the *values* of variable B, then you should think of *adding an interaction* between A and B in your model

Reminder: In the `lm` formula interface this amounts to adding an `A:B` term. For example, if A and B interact to predict Y then the formula would be

```
Y ~ A + B + A:B
```

which would fit the model

$$Y_i = \beta_0 + A_i\beta_A + B_i\beta_B + (A_i \times B_i)\beta_{AB} + \epsilon_i$$

An alternative syntax to fit this model is `A*B`. So to fit the model above using this notation the formula is

```
Y ~ A * B
```

Since we always want to have `A` and `B` if we have an `A:B` term, this more compact notation makes sure we don't forget any of them. But they are equivalent.

Fit a model of all the grades that includes `pre.score`, `treatment`, and the factor version of grade, interacted with `treatment`. How would you construct grade-specific treatment effects from these coefficients?

Question 7

Constructing treatment effects from coefficients can be tricky. Let's take a different approach by creating some representative classes and plotting the difference treatment makes. First create 8 fictional classes: 4 treated and 4 untreated, each with an appropriate value of pre-score (for realism you can use the average `pre.score` in each grade, or for simplicity pick a single `pre.score`). Get predictions from the most recent model for these classes, and plot them. Provide a brief substantive interpretation of the results.