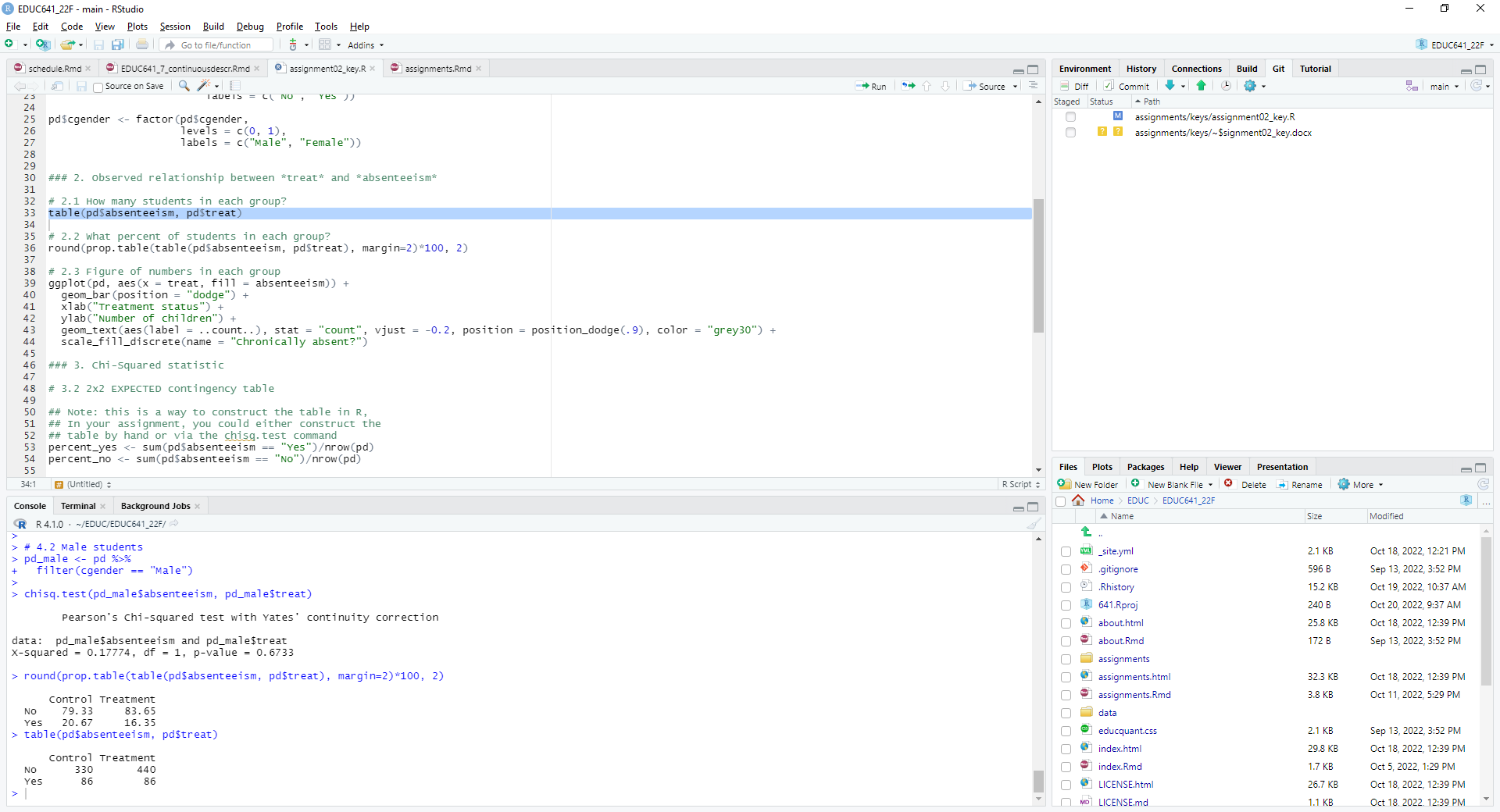
EDUC 641 Assignment 02 Key

1. Read in the dataset

#### 2. Observed relationship between treat and absenteeism

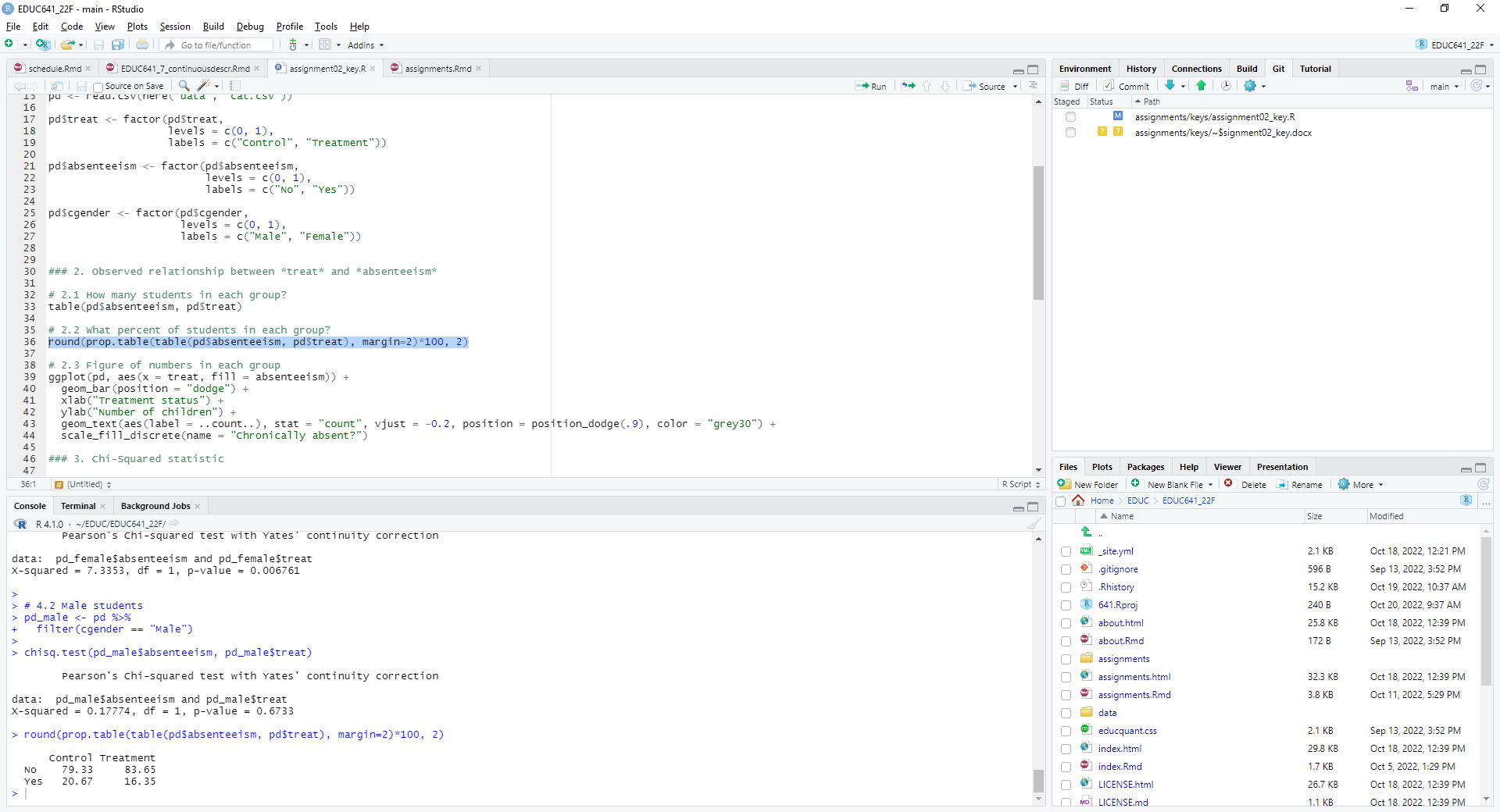
2.1. How many students in the treatment group were chronically absent and not chronically absent, respectively? How many students in the control group were chronically absent/not chronically absent? Write a 1-2 sentence response and create a 2x2 contingency table (frequencies in cell) to demonstrate your answer.

86 students in the treatment group were chronically absent and 440 were not chronically absent. In the control group, 86 students were also chronically absent, but 330 were not.

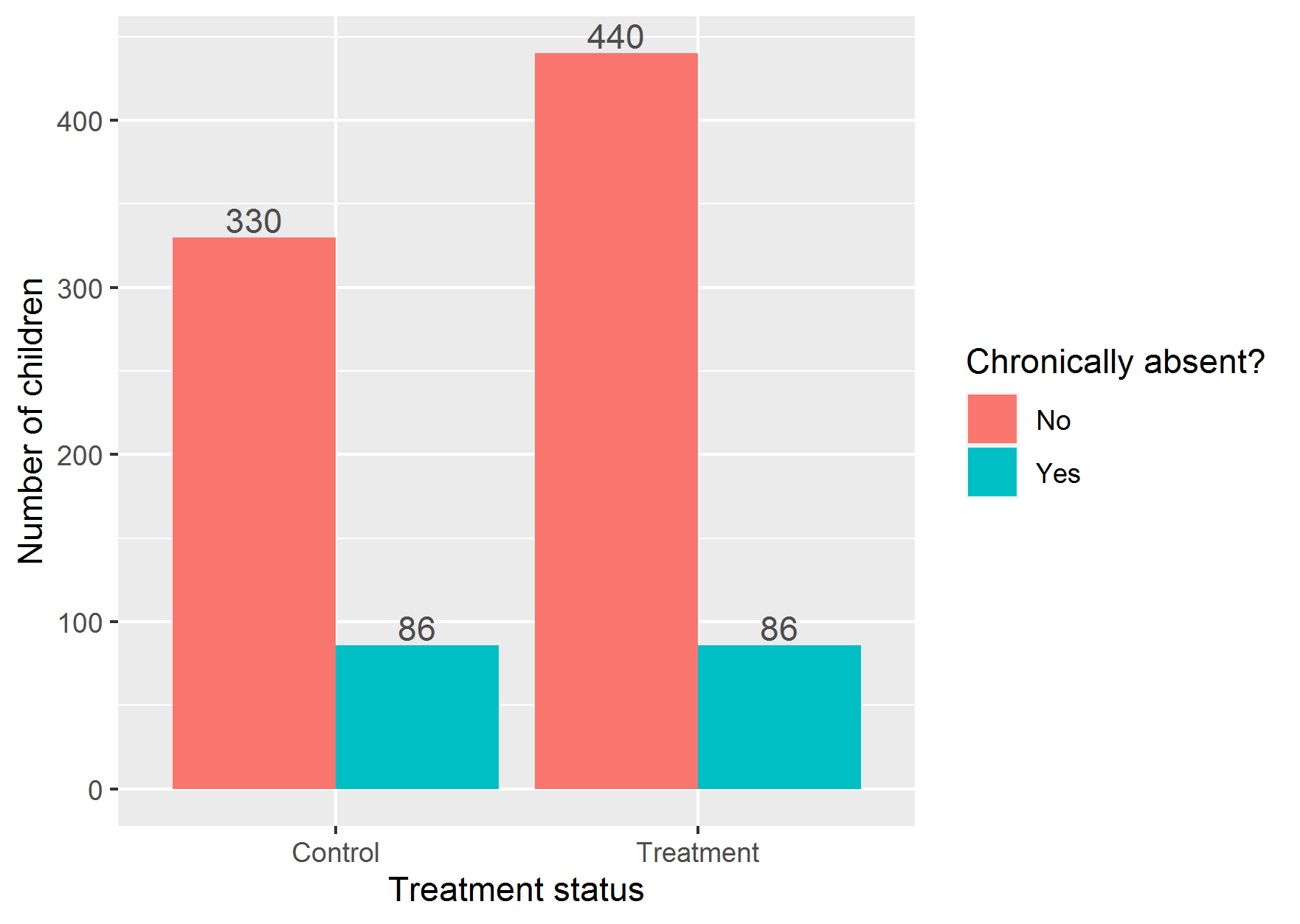


2.2. What percent of students in the treatment group were chronically absent and not chronically absent, respectively? How about in the control group? Write a 1-2 sentence response and create a 2x2 table (percentages in cell) to demonstrate your answer.

16.4 percent of students in the treatment group were chronically absent and 83.7 percent were not chronically. In the control group, 20.7 percent of students were chronically absent and 79.3 percent were not.



2.3. Create one figure to visualize the numbers of chronically absent/not chronically absent students in the treatment and control groups. Make sure to label the x- and y-axes.



2.4. Based on your observations from 2.1 to 2.3, do you think there is a relationship between absenteeism and treat? Why or why not?

The table in 2.2. shows that there were 16 percent students in treatment group who were chronically absent, which is 5 percentage points less than the 21 percent who were absent in the control group. As the figure from 2.3 shows, this primarily comes from the larger number of students who were not chronically absent in the treatment group. However, we would need to investigate this further to determine whether these differences are due to idiosyncratic sampling from the two underlying populations.

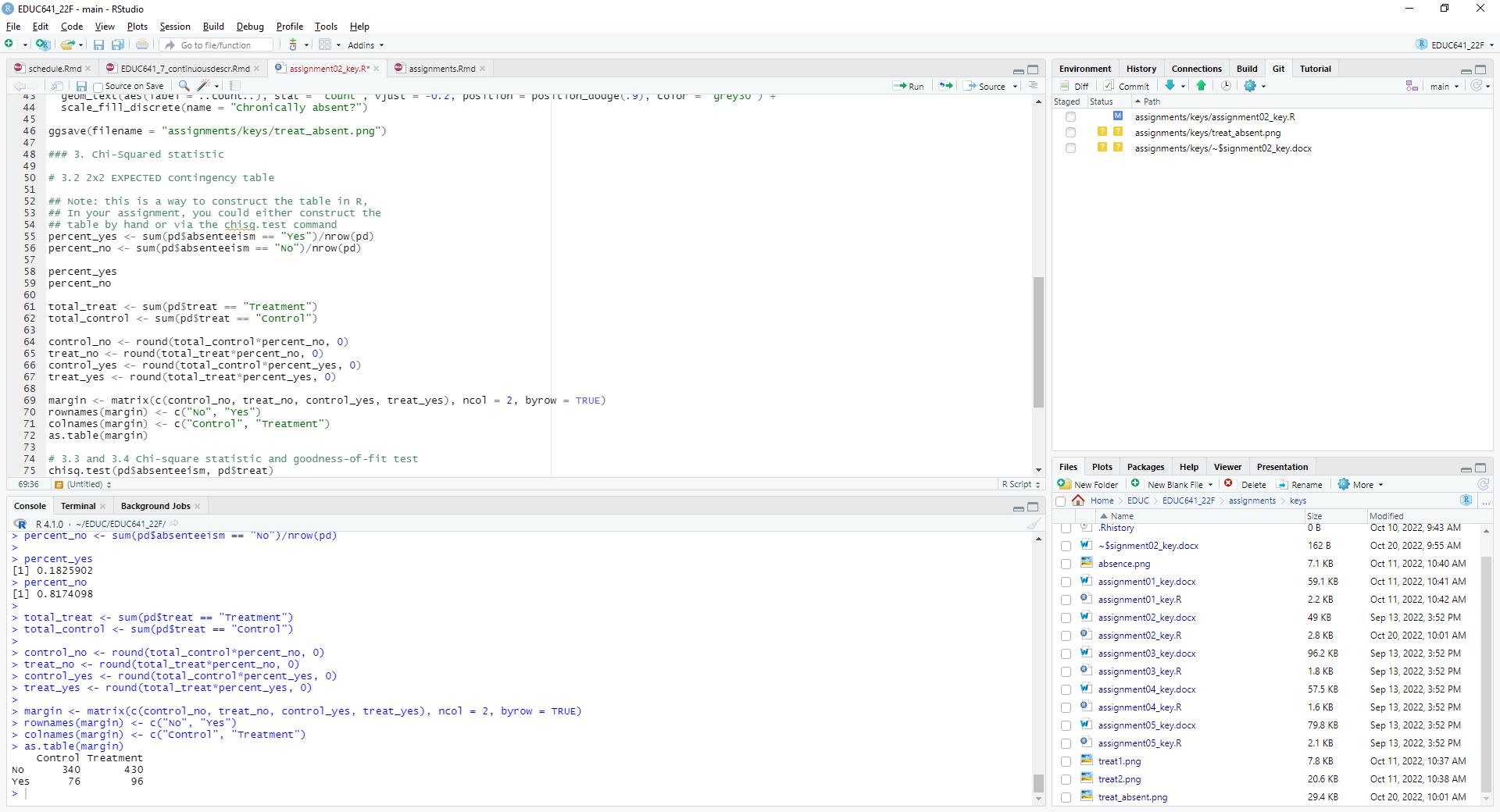
#### 3. Chi-Square goodness-of-fit test of categorical association

3.1. Our first research question of interest is whether teacher participation in professional development decreased rates of student chronic absenteeism. State your null hypothesis for this research question?

There is no relationship between teacher participation in professional development and the rate of student chronic absenteeism, on average in the population.

3.2. If there were **NO** relationship between absenteeism and treat, on average in the population, how many students would you **expect** to be chronically absent/not chronically absent in each of the treatment/control groups? Re-create the table from question 2.1 to show these expected values and interpret the values in this new table in 2-3 sentences.

Overall, about 18 percent of students were chronically absent and 82 percent were not. If there were no relationship between *absenteeism* and *treat* we would expect the proportion of chronically absent students *in the treatment group* to be 18 percent as well, and we would expect the same for students in the control group. In particular, we would expect 340 students in the control group to not be chronically absent and 76 to be chronically absent. We would expect 430 students in the treatment group not to be chronically absent and 96 to be chronically absent.



3.3. Calculate the Chi-squared statistic **by hand** using the tables you generated in 2.1 and 3.2. In one sentence, state what this Chi-squared statistic represents.

The Chi-Squared value of 2.884 measures the net discrepancy between the observed and expected frequencies.

*Note: you will get a slightly different Chi-square value if you calculate this by hand than if you calculate this in R, due to rounding the expected values to whole numbers and due to the “Yates’ continuity correction” that R applies. It will not make any difference to your inference. If you want to test it out, calculate the Chi-square by hand without rounding and then specify the chisq.test in R with the option: correct = FALSE, and you will get the exact same answers.*

3.4. Perform a Chi-squared goodness-of-fit test in R to examine the relationship between absenteeism and treat. Write 3-4 sentences to interpret your results and answer the research question: Were students less likely to be chronically absent if their teachers participated the consultancy PD intervention?

We conducted a Chi-squared goodness-of-fit test with an alpha-threshold of 0.05, and we failed to reject our null hypothesis. In the population from which the sample for this study was drawn, we did not detect any statistically significant relationship, on average, between student chronic absenteeism and whether their teachers were randomly assigned to participate in a professional development consultancy (*χ2* = 2.627, *p* = 0.105).

#### 4. Sub-sample comparisons

4.1. We are also interested in whether absenteeism and treat are related for female students. Perform a Chi-squared goodness-of-fit test in R to investigate the relationship between absenteeism and treat for female students. Write 2-3 sentences to state your null hypothesis, interpret your results, and answer the research question: Were female students less likely to be chronically absent if their teachers participated in the consultancy PD intervention?

We conducted a Chi-squared goodness-of-fit test with an alpha-threshold of 0.05, and we reject the null hypothesis that any differences in our data were due to sampling idiosyncrasy. In the population of female students in the context in which the study was conducted, we found a statistically significant relationship, on average, between student chronic absenteeism and whether female students’ teachers were randomly assigned to participate in a professional development consultancy (*χ2* = 7.335, *p* = 0.007).

4.2. Do the same for male students.

We conducted a Chi-squared goodness-of-fit test with an alpha-threshold of 0.05, and we failed to reject our null hypothesis. In the population of male students in the setting in which the study was conducted, we did not detect any statistically significant relationship, on average, between male student chronic absenteeism and whether their teachers were randomly assigned to participate in a professional development consultancy (*χ2* = 0.178, *p* = 0.673).