Partner Coding Practice 4

Your names

We will work with the `satgpa` data from the `openintro` package. In the following code chunk, load in the `openintro` package and the package we use for wrangling. Then take a look at the Help file of the `satgpa` dataset.

Question 1

We'd like to know if performance on the SAT and high school GPA are correlated with first year GPA. Create a visualization that effectively shows the relationship between a student's 1) average SAT percentile (averaged across math and verbal), 2) their high school GPA, and 3) their first year GPA. Interpret what you find.

Question 2

Using wrangling code, find the proportion of students who had better first year GPAs than high school GPAs.

Question 3

Create a new data frame called `satgpa2` that contains two new variables: one called `hs_gpa_cat` and another called `fy_gpa_cat`. Each of these variables should take the value "under" if the respective GPA is less than or equal to 3.0, and "over" if the respective GPA is greater than 3.0.

Question 4

Using wrangling code, obtain the probability of having a first year GPA greater than 3.0, conditioning on the different levels of high school GPA category. Wrangle until your displayed output has only two rows and two variables. *In choosing how to reduce the number of rows and columns, think about what is the most important information you need to present.*