# 4.12.1 Create Bins for School Size

**Creating** the spending bins for the scores by school spending per student was a bit of a challenge, but you did a great job. In fact, Maria is so impressed that she now wants you to group the same averages and percentages by the school size.

Previously, we grouped the scores by school spending per student. Now we'll group the scores by school size. For this DataFrame we'll create three bins: small, medium, and large. To determine these bins, we can look at the `per_school_counts`, or total students, from each school from the `per_school_summary_df` DataFrame.

## REWIND

---

We retrieved the `per_school_counts` using the following code:

Looking at the `per_school_counts` Series, we can see that the highest student population is 4,976. Therefore, the upper edge will be 5,000. Because we have only three bins—small, medium, and large—here's how we can group the school sizes:

- "Small" schools have fewer than 1,000 students
- "Medium" schools have 1,000–1,999 students
- "Large" schools have 2,000–5,000 students

In the new DataFrame, calculate the average math and reading scores, math and reading passing percentages, and overall passing percentage for each bin, as shown in the following image:

To establish the bins and group names, add the following code to a new cell and run the cell.

```
# Establish the bins.
size_bins = [0, 1000, 2000, 5000]
group_names = ["Small (<1000)", "Medium (1000-2000)", "Large (2000-5000)"]
```

Now we can "cut" the `per_school_summary_df` DataFrame using the `cut` function to get the three sizing bins.