4.6.4

Activity: Aggregating Answers

Now that Maria has gotten your targeted analysis, she's identified some questions that she wants you to answer by aggregating and comparing the data. Specifically, she wants to discover if the test scores differ among grades or among schools.

In this activity, you'll use `groupby` with other Pandas functions to answer complicated questions about the dataset.

Files

Download the following files to help you get started:

Aggregating Answers files (https://2u-data-curriculum-team.s3.amazonaws.com/dataviz-online/v2/module_4/4-6-Student_Data_Starter_Code.zip)

sing the Jupyter Notebook file in the `Unsolved` folder, navigate to the "Step 5" section to write your code. Alternatively, you can use the Jupyter Notebook file that you created in the previous lesson to write your code.

Instructions

1. Aggregate the data to get the average math and average reading scores for every grade, as the following code shows:

```
avg_student_scores_by_grade = student_df.groupby(by='grade').mean()
avg_student_scores_by_grade.loc[:, ["math_score", "reading_score"]]
```

2. Find the lowest math and lowest reading scores in every school, as the following code shows:

```
min_scores_by_school = student_df.groupby(by='school_name').min()
min_scores_by_school.loc[:, ["math_score", "reading_score"]]
```

3. Sort the data from the previous step to show the lowest math scores first, as the following code shows:

```
min_scores_by_school = student_df.groupby(by='school_name').min()
min_scores_by_school.loc[:, ["math_score", "reading_score"]].sort_values("math_score")
```

4. Aggregate the data by school name and then by grade, and then calculate the average reading score and the average math score, as the following code shows:

```
avg_by_school = student_df.groupby(['school_name', 'grade']).mean()
avg_by_school.loc[:, ["reading_score", "math_score"]]
```

Solution

How did you do?

You can refer to the solution file in the `Solved` folder, which is in the zipped folder that you downloaded for this activity.

What's Next?

The analysis of the school data is complete! Now that you've learned how to use the foundational Pandas tools, it's time to apply your skills in this week's Challenge.