

Introduction to Comparing Subsets of Data

So far, we've completed the first and second parts of analyzing our data: summarizing the data and drilling down into the data. Now, it's time for the third and final part: comparing subsets of the data.

When we summarized the data, we completed a big-picture analysis. When we drilled down into the data, we completed a targeted analysis. Now, we'll complete an analysis by grouping, or **aggregating**, the data.



CONNECT THE DOTS

In the previous activity, you used `loc` and `iloc` to drill down into subsets of data. Can you think of any subsets that might prove interesting to compare to one another?

When we aggregate the data, we do so to be able to compare the aggregations, which are subsets of the data. Pandas has functions that we can use to accomplish both the aggregation and the comparisons. Let's learn how to put those functions to use.

CHALLENGE CONNECTION

In this week's Challenge, you'll compare grades and schools by using the student dataset.

Now that you've received an introduction to comparing subsets of data, you'll next learn how to aggregate the data.

© 2020 - 2022 Trilogy Education Services, a 2U, Inc. brand. All Rights Reserved.