

Introduction to Summarizing Data

So far, you've completed the first two phases of the three-phase data analysis process: collecting the data and preparing the data. Now, it's time for the third phase: analyzing the data.

Analyzing the data might seem like an open-ended and intimidating task. For example, how do you determine the most impactful product line that your company sells? Or when presented with a dataset of customer information, like their locations, purchase histories, and ad exposures, how do you find insights into how to compare the demographics from a marketing perspective? Regarding data analysis, the starting and ending points might seem impossible to find. But with a few key techniques, data analysis becomes more manageable and approachable.

We'll break down this phase into three parts: summarizing the data, drilling down into the data, and comparing subsets of the data. This lesson focuses on summarizing the data. Later lessons will focus on drilling down into the data and comparing subsets of the data.

In this lesson, you'll specifically learn how to analyze data at a high level, including the use of visualizations and summary statistics. And, you'll learn how to do a targeted analysis with the help of Pandas functions. By the end of this lesson, you'll have the tools that you need to develop a game plan for a data analysis.

CHALLENGE CONNECTION

In this week's Challenge, you'll use Pandas to do both high-level and date-focused analyses of school data.

Now that you've received an introduction to summarizing data, you'll next learn about doing a big-picture analysis.

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