

Introduction to Data Analysis: The Three-Phase Process

Maria feels excited to have you join the team, but she also understands that onboarding can be challenging. So to help you prepare for your new tasks and responsibilities, she's organized training activities to get you up to speed.

Before we start collecting data, let's get an introduction to data analysis as a whole. Although no single best way exists for analyzing data, it typically follows a three-phase process:

1. **Collect the data:** In this phase, we collect the data for analysis by importing it into a structure that Python and Pandas can work on.
2. **Prepare the data:** In this phase, we clean the data to account for missing, incomplete, erroneous, and duplicated data. This phase also includes manipulating segments of the dataset to highlight the relationships among variables.
3. **Analyze the data:** This phase consists of an iterative process and varies depending on the goal of the analysis or research. Regardless of the specifics, the analysis involves testing a thesis, reviewing the results, and refining both until we reach a conclusion.

We'll take this three-phase approach in this module.

CHALLENGE CONNECTION

In this week's Challenge, you'll apply this three-phase approach as you further prepare and analyze the school data by using Pandas and JupyterLab.

This lesson focuses on the first phase: collect the data. In later lessons, you'll learn how to perform the second and third phases. Specifically, you'll learn how to prepare and analyze data to identify business opportunities and key insights. And, you'll do all this by using Pandas and Jupyter Notebook.

HIDE PRO TIP

Keep in mind that this three-phase process isn't a one-size-fits-all solution. In a future data analyst role, you might need to follow a specific process that differs from the collect-prepare-analyze approach. The key thing to remember is that a structured analysis process—regardless of the exact phases—helps ensure consistency and prevents missing critical steps.

Now that you've received an introduction to data analysis, you'll next receive a review of your data analysis tools.

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