

STUDY GUIDE

PLOTTING WITH PANDAS

Key Terms and Definitions

- » **Pandas:** The Python library for organizing, cleaning, and otherwise manipulating your data, which a different approach to visualization. Instead of creating the plot and then inserting the data, it creates the graph directly from the data itself.
- » **Exploratory Data Analysis (EDA):** The process of investigating data visually as a primary step in understanding the data set.
- » **.plot():** The main method in Pandas for creating a visualization.
 - Alternately, you can use the particular method of the type of chart you're trying to build to go directly to that module, then use parameters unique to that type of chart (e.g., `.bar()` or `.scatter()`).
- » **kind=:** The argument within `.plot()` that allows you to tell Pandas what type of chart to create. Choices include bar, hist, box, kde (or density), area, scatter, hexbin, and pie.

Guiding Questions

1. Why is EDA an important early step in the data science workflow?
2. What are some key insights you should get from your EDA process?

Additional Resources

1. [Using Pandas to Create Charts](#)
2. [DataCamp: Pandas Foundations](#)
 - » Specifically, look at Section 2: "Exploratory Data Analysis."

Technical Resources

1. [Install Anaconda](#) and [Verify Installation](#)
2. [Jupyter Notebook Practice](#)

