Follow-along instructions: Explore linear regression with Python

Accessing and utilizing resources in this section

While watching the video that follows this reading, you may find it helpful to track the instructor's progress by following along in your own Jupyter notebook. To do so, open the annotated follow-along guide for the video. The content in this notebook is identical to the content shown in this lesson's instructional video. In addition to that content, you'll find additional information throughout the notebook. That information is provided to explain the purpose of each concept covered, why the code is written in a certain way, and tips for running the code.

Steps to complete:

1. Read this page of instructions.
2. Open the Annotated follow-along guide: Explore linear regression with Python, which contains a version of the same notebook the instructor will use in the video.
3. Follow along with the instructor as they go over the code in the notebook.
4. Learn from the instructor and practice running the code in your notebook.

Data dictionary

This lesson uses a dataset called `penguins`. It represents data about three penguin species. This dataset is frequently used when teaching regression fundamentals. For more information about the dataset, refer to the Palmer penguins introduction.

The dataset contains:

344 rows – each row is a unique penguin

7 columns:

Column name	Type	Description
species	str	Penguin species
island	str	Island name in the Palmer Archipelago
bill_length_mm	float	Length of penguin's bill in millimeters
bill_depth_mm	float	Depth of penguin's bill in millimeters
flipper_length_mm	float	Length of penguin's flipper in millimeters
body_mass_g	float	Penguin's body mass in grams
sex	string	Penguin's sex