# Follow-along instructions: Compute descriptive statistics 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 <u>Annotated follow-along guide: Compute descriptive statistics with Python</u> ⬈, 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

In this lesson's video, your notebook will include the dataset **education_districtwise.csv**. This dataset represents a list of school districts in an anonymous country. The data includes district and state names, total population, and the literacy rate.

The dataset contains:

680 rows – each row is a different school district

7 columns

Column name	Type	Description
DISTNAME	str	The names of an anonymous country's school districts
STATNAME	str	The names of an anonymous country's states where school districts are located
BLOCKS	int64	The number of blocks in the school district. Blocks are the smallest organizational structure in the education system of the anonymous country.
VILLAGES	int64	column shows how many villages are in each district
CLUSTERS	int64	The number of clusters in the school district. Clusters are the second smallest organizational structure in the education system of the anonymous country.
TOTPOPULAT	int64	column shows the population for each district
OVERALL_LI	int64	shows the literacy rate for each district

Mark as completed