

## Analyzing Data with Pandas

3 min

Catherine wants to know if Codecademy learners are finding what they need. She needs to analyze data from a survey that was administered to a subset of visitors.

Catherine selects the survey data from the SQL [database](#) and downloads it onto her computer. She saves the data as a [CSV](#) file (extension **.csv**), which stands for Comma-Separated Values. CSV is a text-only spreadsheet format that lets us store and explore data.

To analyze the survey data, Catherine will use **pandas**. Python is a programming language, and pandas is a special set of commands in Python that lets us analyze spreadsheet data. Pandas can do a lot of the things that SQL can do, but it's also backed by the power of Python, so we can easily transition from analyzing our data with pandas to visualizing it using other Python tools.

### Instructions

1. Checkpoint 1 Passed

1.

Catherine wrote some code to analyze her data in **script.py**, the file in the center of the screen.

If you press "Run" now, it will generate errors because she is missing this line of code:

```
import pandas as pd
```

Paste that code into the top of **script.py** and then press "Run".

You'll see the results of Catherine's analysis in the right panel.

### **script.py**

```
import codecademylib3_seaborn
```

```
# Paste code here:
```

```
import pandas as pd
```

```
# Load data
```

```
df = pd.read_csv('page_visits.csv')
```

```
# Display data
```

```
print(df.head())
```

	first_name	last_name	gender	age	visit_date	website_goal
0	James	Mcguire	male	50.0	2017-10-26	Find a job
1	John	Vaughan	male	35.0	2017-11-09	Pay a bill
2	Robert	Dillon	male	33.0	2017-11-28	Pay a bill
3	Michael	Cameron	male	40.0	2017-09-27	When is trash day?
4	William	Yang	male	38.0	2017-11-01	Contact City Council