

Driving Right (1) in Python

The screenshot shows a web browser window with the URL `campus.datacamp.com/courses/intermediate-python/logic-control-flow-and-filtering/ex=15`. The page is titled "Driving right (1)" and is part of an "Exercise" section. It contains instructions for a task involving the 'cars' dataset. The instructions are as follows:

- Remember that `cars` dataset, containing the cars per 1000 people (`cars_per_cap`) and whether people drive right (`drives_right`) for different countries (`country`). The code that imports this data in CSV format into Python as a DataFrame is included in the script.
- In the video, you saw a step-by-step approach to filter observations from a DataFrame based on boolean arrays. Let's start simple and try to find all observations in `cars` where `drives_right` is `True`.
- `drives_right` is a boolean column, so you'll have to extract it as a Series and then use this boolean Series to select observations from `cars`.

Below the instructions, there are "Instructions" and "Take Hint (-30 XP)" buttons. To the right, there is a code editor with the following Python code:

```
1 # Import cars data
2 import pandas as pd
3 cars = pd.read_csv('cars.csv', index_col = 0)
4
5 # Extract drives_right column as Series: dr
6
7
8 # Use dr to subset cars: sel
9
10
11 # Print sel
12
```

At the bottom of the code editor, there are buttons for "Run Code" and "Submit Answer". Below the code editor is a "Python Shell" window with the prompt `In [1]:`.

****Question:****

Remember that `cars` dataset contains the cars per 1000 people (`cars_per_cap`) and whether people drive right (`drives_right`) for different countries (`country`). The code that imports this data in CSV format into Python as a DataFrame is included in the script.

1. Extract the `drives_right` column as a Pandas Series and store it as `dr`.
2. Use `dr`, a boolean Series, to subset the `cars` DataFrame. Store the resulting selection in `sel`.
3. Print `sel`, and assert that `drives_right` is `True` for all observations.

****Answer:****

Here is the Python code that solves the problem:

```
# Import cars data
import pandas as pd
```

```
# Read the dataset
cars = pd.read_csv('cars.csv', index_col=0)
```

```
# Extract drives_right column as Series: dr
```

```
dr = cars['drives_right']

# Use dr to subset cars: sel
sel = cars[dr]

# Print sel
print(sel)
```

****Explanation:****

1. ****Import Pandas****: The Pandas library is imported to handle tabular data in a DataFrame.
2. ****Read the dataset****: The `pd.read_csv()` function reads the dataset from a CSV file into a Pandas DataFrame. The `index_col=0` parameter sets the first column as the index.
3. ****Extract `drives_right` column****: The `drives_right` column is extracted as a Pandas Series and stored in `dr`.
4. ****Subset the DataFrame****: The `cars[dr]` expression uses the boolean Series `dr` to filter the rows of the DataFrame where `drives_right` is `True`. The result is stored in `sel`.
5. ****Print the result****: The `print(sel)` statement displays the subsetted DataFrame.