

Examining the Dataset

The screenshot shows a web browser window with the URL `campus.datacamp.com/courses/analyzing-police-activity-with-pandas/preparing-the-data-for-analysis/ex-2`. The page is titled "Examining the Dataset" and contains the following instructions:

- Import pandas using the alias `pd`.
- Read the file `police.csv` into a DataFrame named `ri`.
- Examine the first 5 rows of the DataFrame (known as the 'head').
- Count the number of missing values in each column. Use `.isnull()` to check which DataFrame elements are missing, and then take the `.sum()` to count the number of True values in each column.

The code editor on the right contains the following Python code:

```
1 # Import the pandas library as pd
2 import pandas as pd
3
4 # Read 'police.csv' into a DataFrame named ri
5 ri = pd.read_csv('police.csv')
6
7 # Examine the head of the DataFrame
8 print(ri.head())
9
10 # Count the number of missing values in each column
11 print(ri.isnull().sum())
```

Task Description

1. Import pandas using the alias `pd`.
2. Read the file `'police.csv'` into a DataFrame named `ri`.
3. Examine the first 5 rows of the DataFrame (known as the `'head'`).
4. Count the number of missing values in each column using `.isnull()` to check which DataFrame elements are missing, and then take the `.sum()` to count the number of True values in each column.

Code Solution

```
# Import the pandas library as pd
import pandas as pd
```

```
# Read 'police.csv' into a DataFrame named ri
ri = pd.read_csv('police.csv')
```

```
# Examine the head of the DataFrame
print(ri.head())
```

```
# Count the number of missing values in each column
print(ri.isnull().sum())
```

Code Explanation

1. The line `'import pandas as pd'` imports the pandas library, which is essential for working with data in tabular format in Python.
2. The line `'ri = pd.read_csv('police.csv')'` reads the CSV file named `'police.csv'` into a pandas DataFrame called `'ri'`. This DataFrame is used to hold and manipulate the data.
3. The line `'print(ri.head())'` displays the first five rows of the DataFrame. This is a quick way to inspect the structure and content of the data.
4. The line `'print(ri.isnull().sum())'` calculates the number of missing values in each column. The `'isnull()'` method identifies missing data, and `'sum()'` adds up the missing values for each column.