Pivot and Loc's

21 October 2023 22:05

PIVOT

In Python, when working with data analysis and manipulation, you can use the pivot function to reshape your data. The pivot function allows you to transform your data from a long format to a wide format, or vice versa, based on the values in the DataFrame. The pivot function is available in various libraries like Pandas, which is commonly used for data manipulation and analysis. Here is an example of how to use the pivot function in Pandas:

```
import pandas as pd

# Create a sample DataFrame
data = {'date': ['2023-10-01', '2023-10-01', '2023-10-02', '2023-10-02'],
        'category': ['A', 'B', 'A', 'B'],
        'value': [1, 2, 3, 4]}
df = pd.DataFrame(data)

# Pivot the DataFrame
pivot_df = df.pivot(index='date', columns='category', values='value')
print(pivot_df)
```

This will output:

```
category     A    B
date
2023-10-01  1.0  2.0
2023-10-02  3.0  4.0
```

In the context of Pandas, both loc and iloc are used for accessing data in a DataFrame, but they are used for different types of indexing. Here's the difference between them:

1. loc: It is label-based indexing. When you use loc, you can access a group of rows and columns by labels or a boolean array. This means you need to specify the row labels and column labels. For example:

   Python:
   ```
   df.loc[row_label, column_label]
   ```

2. iloc: It is integer-based indexing. When you use iloc, you can access a group of rows and columns by integer position. This means you need to specify the integer location for rows and columns. For example:

   python
   ```
   df.iloc[row_index, column_index]
   ```

Here's a simple example to illustrate the difference:

```
import pandas as pd

data = {'A': [1, 2, 3, 4, 5], 'B': [10, 20, 30, 40, 50]}
df = pd.DataFrame(data, index=['a', 'b', 'c', 'd', 'e'])
```

```
# Using loc
print(df.loc['a', 'A'])  # Output: 1

# Using iloc
print(df.iloc[0, 0])  # Output: 1
```

In this example, `loc['a', 'A']` and `iloc[0, 0]` both access the first element of the DataFrame, but using different methods based on label and integer position, respectively.