Chain Indexing

Context

In Pandas, `df["one"]["two"]` and `df.loc[:,("one","two")]` give the same result. The `df["one"]["two"]` is called **chain indexing.**

Problem

- Chain indexing executes as two separate operations. First, `df["one"]` retrieves a subset, and then `["two"]` operates on the result. This approach is slower than using `df.loc`, which performs the operation in one step.
- Also, assigning values with chain indexing may fail because Pandas does not guarantee whether `df["col"]` returns a view (affects the original data) or a copy (creates a new object).

Solution

Developers using Pandas **should avoid** using chain indexing but they should use **df.loc to access data.**

Existing Stage	**Effect**
Data importing and **Cleaning**.	Error-prone & Efficiency

Example

Python

```python
### Pandas
import pandas as pd
df = pd.DataFrame([[1,2,3],[4,5,6]])
col = 1
x = 0
- df[col][x] = 42

+ df.loc[x, col] = 42
```