pandas Series

Presenter: Steve Baskauf
vanderbi.lt/codegraf

Jean & Alexander Heard
LIBRARIES

CodeGraf landing page

- vanderbi.lt/codegraf

What is Pandas ?

Jean & Alexander Heard
LIBRARIES

Pandas library for Python

- Open source library for data analysis
- Name from "panel data" (econometrics term)
- Built on NumPy
- Creates a DataFrame object
- Easy import from CSV and Excel

Import statement

- Conventional format:

```
import pandas as pd
```

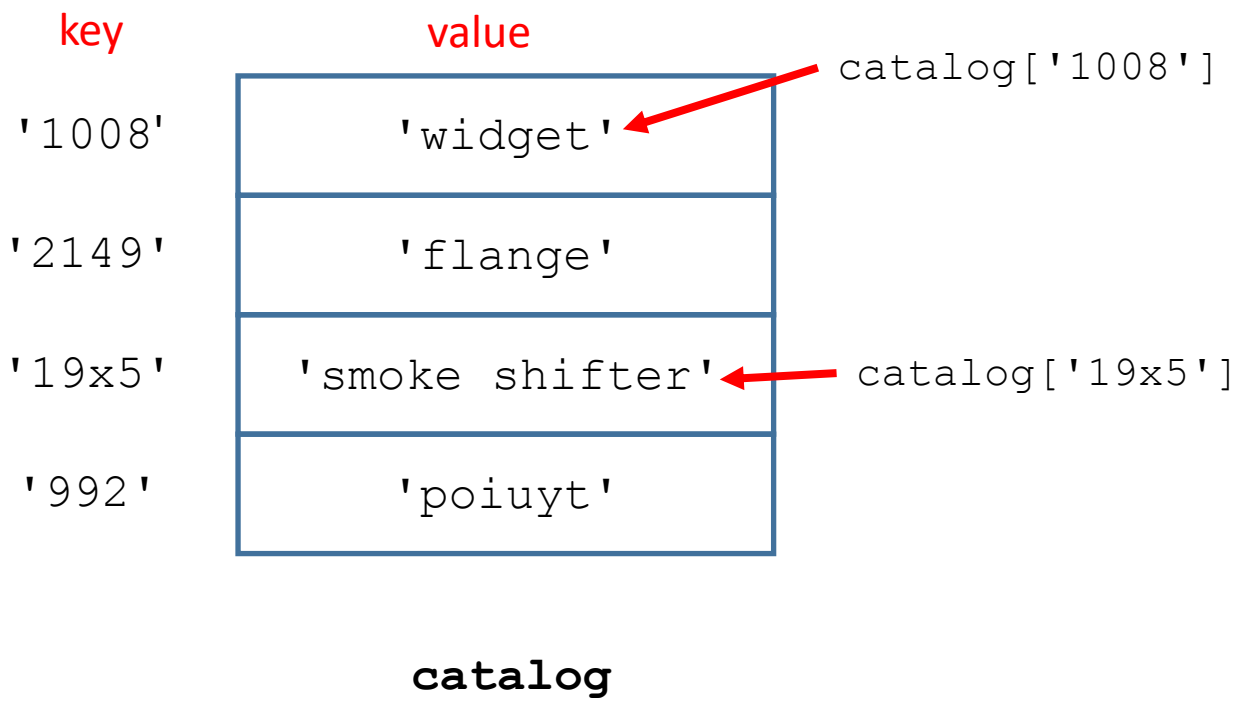
- Use this to make your code compatible with everyone else's code

One dimensional data structures

Jean & Alexander Heard
LIBRARIES

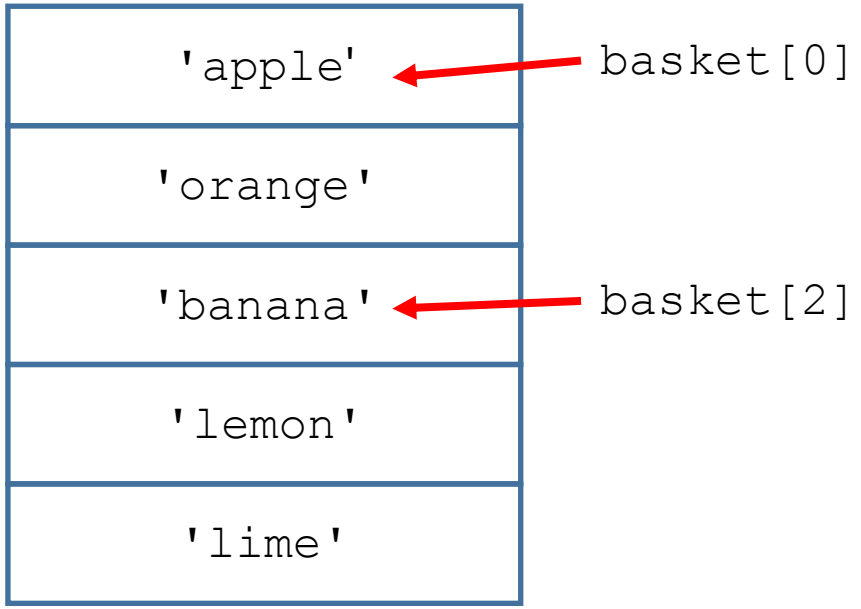
Dictionaries

- Values are referenced by keys.



Lists

- Items are referenced by an integer index (0-based)



For more videos like this, visit the CodeGraf landing page

vanderbi.lt/codegraf

pandas Series

Jean & Alexander Heard
LIBRARIES

Series

- A specific one-dimensional data structure with **named elements**.
- Series items can be referenced by either their integer index or string label.
- Series are built by instantiating a **`pd.Series()`** object.

- pandas Series are built from NumPy arrays + an index (**integers** and **labels**).

- Series can be used in vectorized operations just like NumPy arrays.

For more videos like this, visit the CodeGraf landing page

vanderbi.lt/codegraf

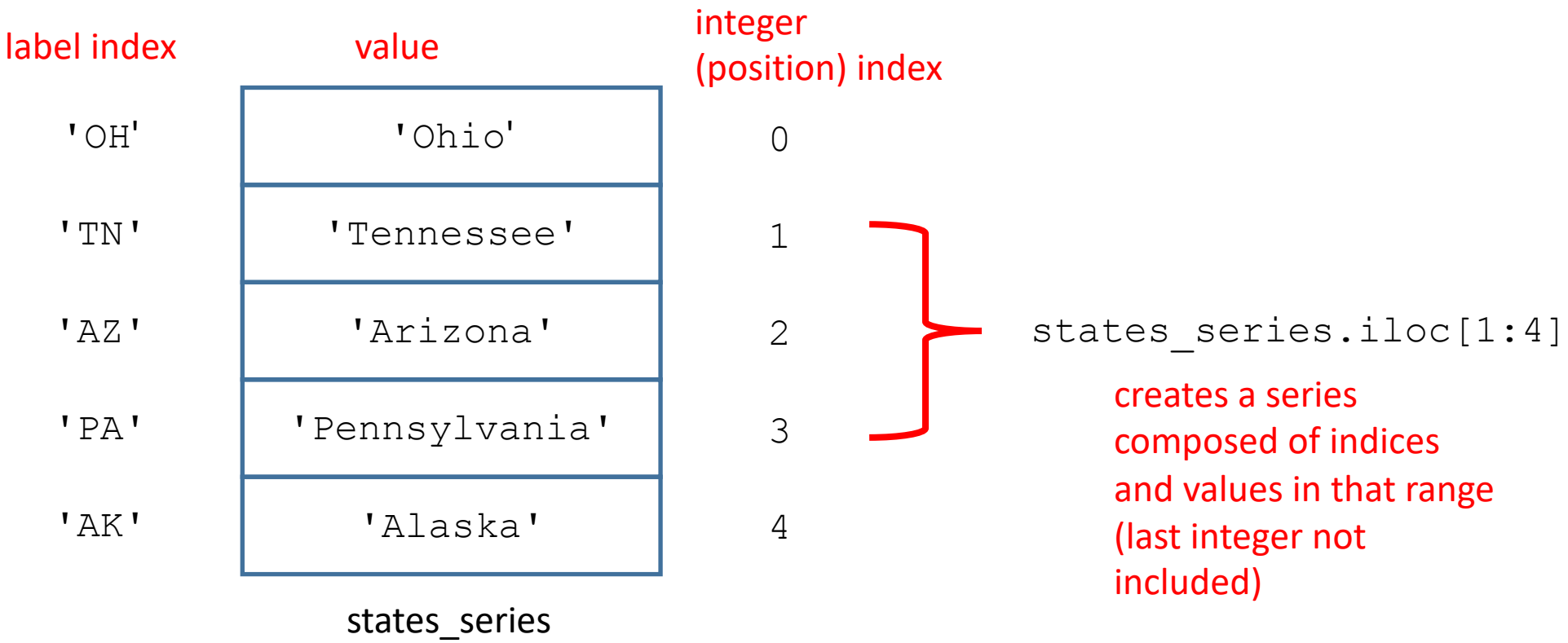
Slicing a Series

DISC DIGITAL SCHOLARSHIP AND COMMUNICATIONS

Jean & Alexander Heard LIBRARIES

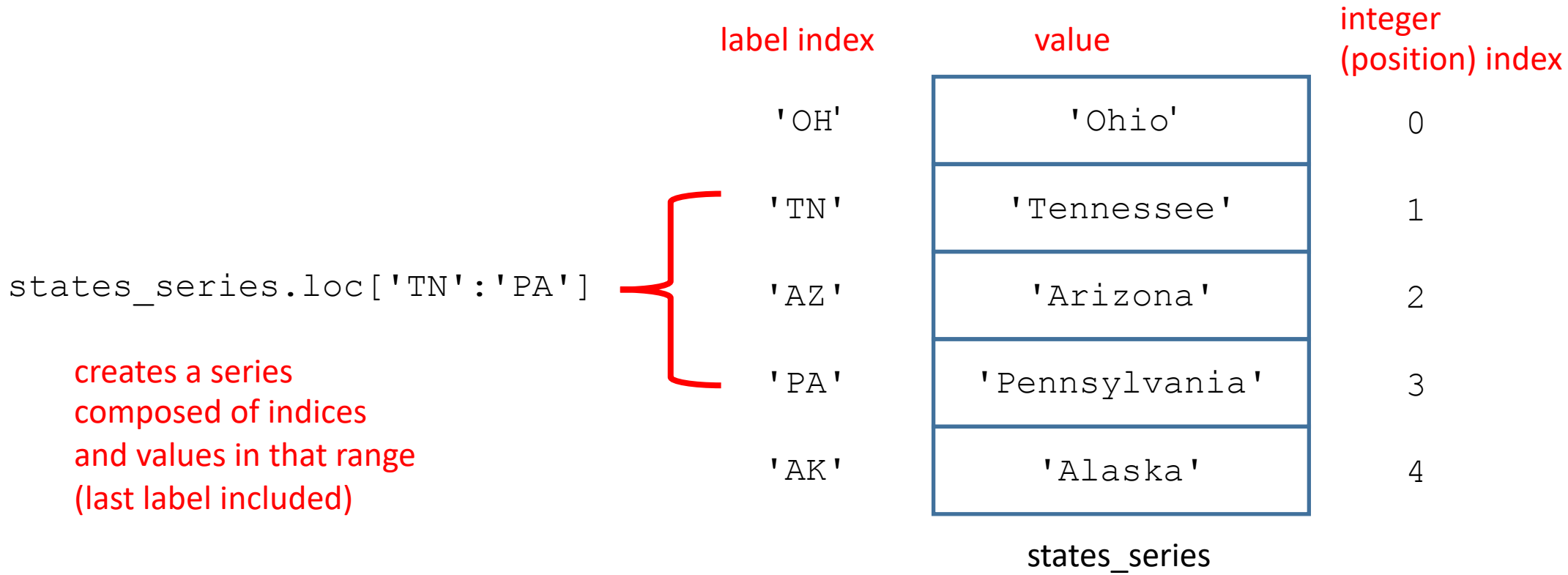
Slicing a Series by integer index: `.iloc[]`

- When slicing, a range is given instead of a single integer
- End of range is one less than ending number
- Slicing generates another Series



Slicing a Series by label index : `.loc[]`

- End of range includes final label.
- Slicing generates another Series



For more videos like this, visit the CodeGraf landing page

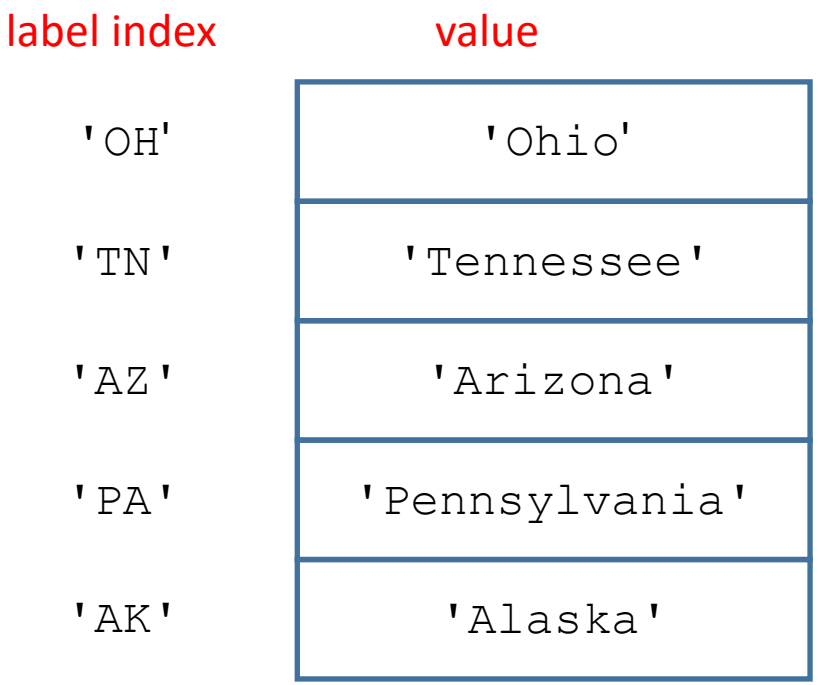
vanderbi.lt/codegraf

Slicing a Series by condition

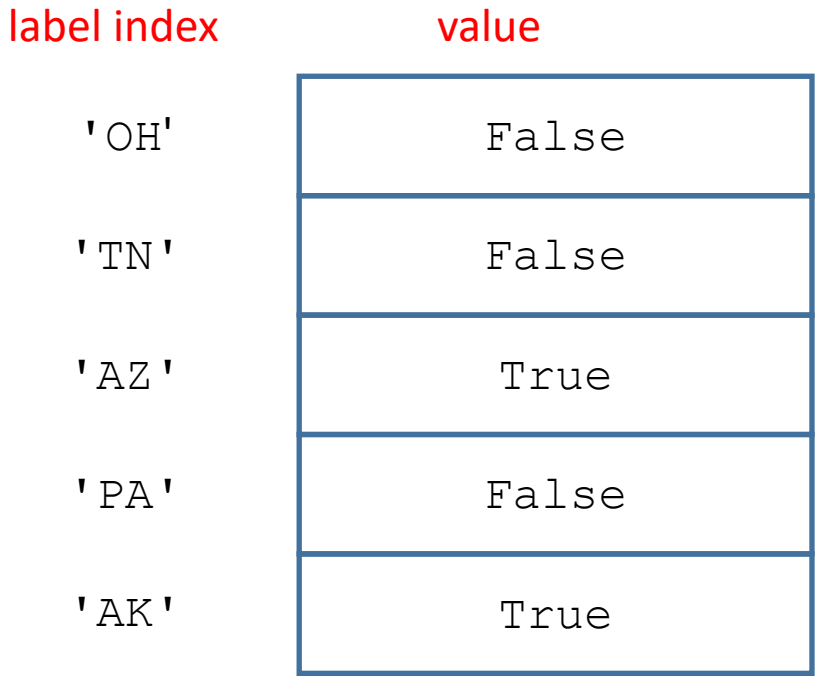
Jean & Alexander Heard
LIBRARIES

Result of a vectorized boolean operation

- The operation generates another Series with corresponding label indices.
- The values of the resulting Series are the evaluation for each item in the original Series.



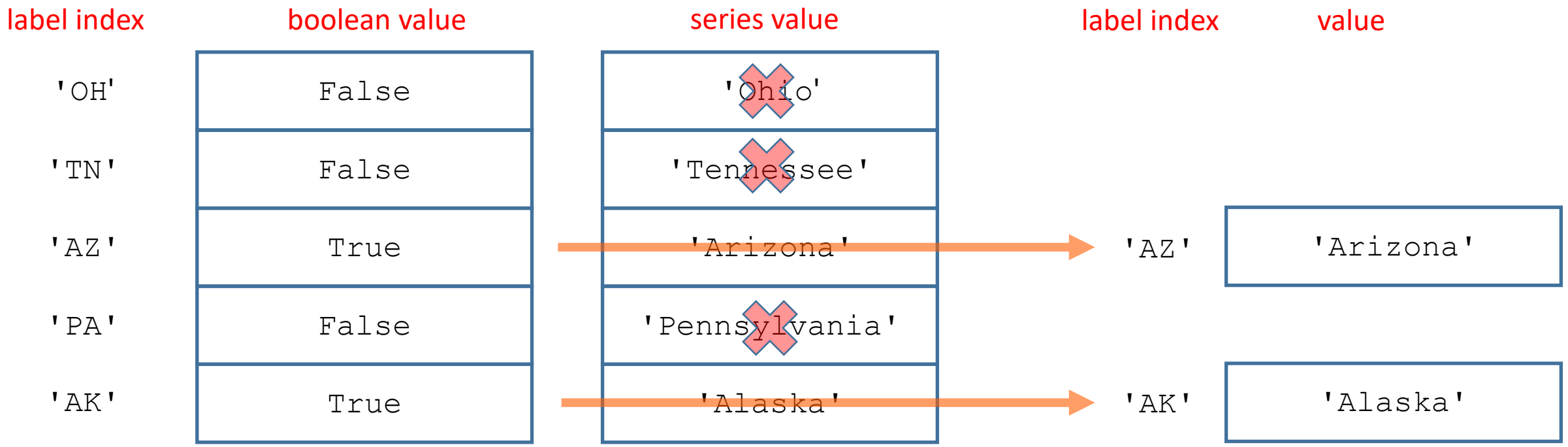
states_series



states_series.str.startswith('A')

Slicing using `.loc[]` with boolean values

- Series values are added to the slice if the corresponding boolean is **`True`**.
- The operation generates another Series with corresponding label indices.




```
states_series.str.startswith('A')
```

```
states_series.loc[states_series.str.startswith('A')]
```

For more videos like this, visit the CodeGraf landing page

vanderbi.lt/codegraf

Slices vs. copies

DISC DIGITAL SCHOLARSHIP AND COMMUNICATIONS

Jean & Alexander Heard LIBRARIES

For more videos like this, visit the CodeGraf landing page

vanderbi.lt/codegraf

Making changes permanent

DISC DIGITAL SCHOLARSHIP AND COMMUNICATIONS

Jean & Alexander Heard LIBRARIES

For more videos like this, visit the CodeGraf landing page

vanderbi.lt/codegraf