Pandas:
A python library that can make analyzing data easier
Documentation: pandas - Python Data Analysis Library (pydata.org)

Dataframes:

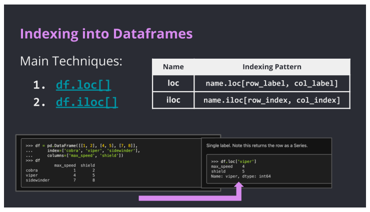
- A pandas object that is used to store a dataset
- Information is organized in rows and columns
- Dataframes simplify common operations, like sorting data

Series:

- A pandas object used to create dataframes
- Seen as a one dimensional list of data
    - Think of it as a single column in a dataframe

Indexing into Dataframes
Main Techniques:

1. df.loc[]
2. df.iloc[]



Indexing into Dataframes

Main Techniques:

1. `df.loc[]`
2. `df.iloc[]`

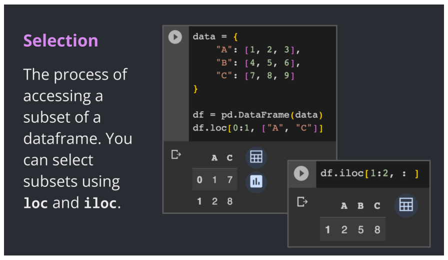
Name	Indexing Pattern
loc	name.loc[row_label, col_label]
iloc	name.iloc[row_index, col_index]

```
>>> df = pd.DataFrame([[1, 2], [4, 5], [7, 8]],
...      index=['cobra', 'viper', 'sidewinder'],
...      columns=['max_speed', 'shield'])
>>> df
            max_speed  shield
cobra               1       2
viper               4       5
sidewinder          7       8
```

Single label. Note this returns the row as a Series.

```
>>> df.loc['viper']
max_speed    4
shield       5
Name: viper, dtype: int64
```

9-19-2023



Selection

The process of accessing a subset of a dataframe. You can select subsets using **loc** and **iloc**.

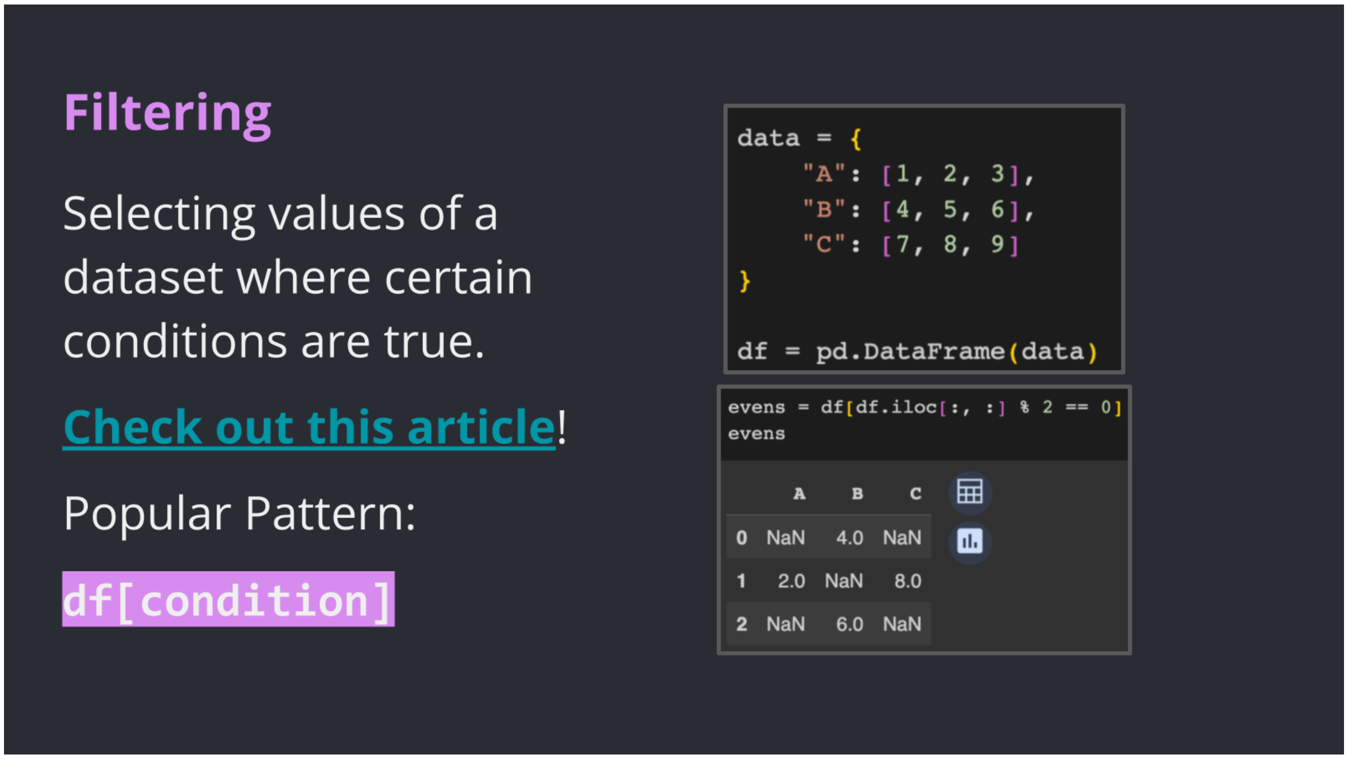
```
data = {
    "A": [1, 2, 3],
    "B": [4, 5, 6],
    "C": [7, 8, 9]
}

df = pd.DataFrame(data)
df.loc[0:1, ["A", "C"]]
```

```
   A  C
0  1  7
1  2  8
```

```
df.iloc[1:2, : ]
```

```
   A  B  C
1  2  5  8
```



Filtering

Selecting values of a dataset where certain conditions are true.

Check out this article!

Popular Pattern:

`df[condition]`

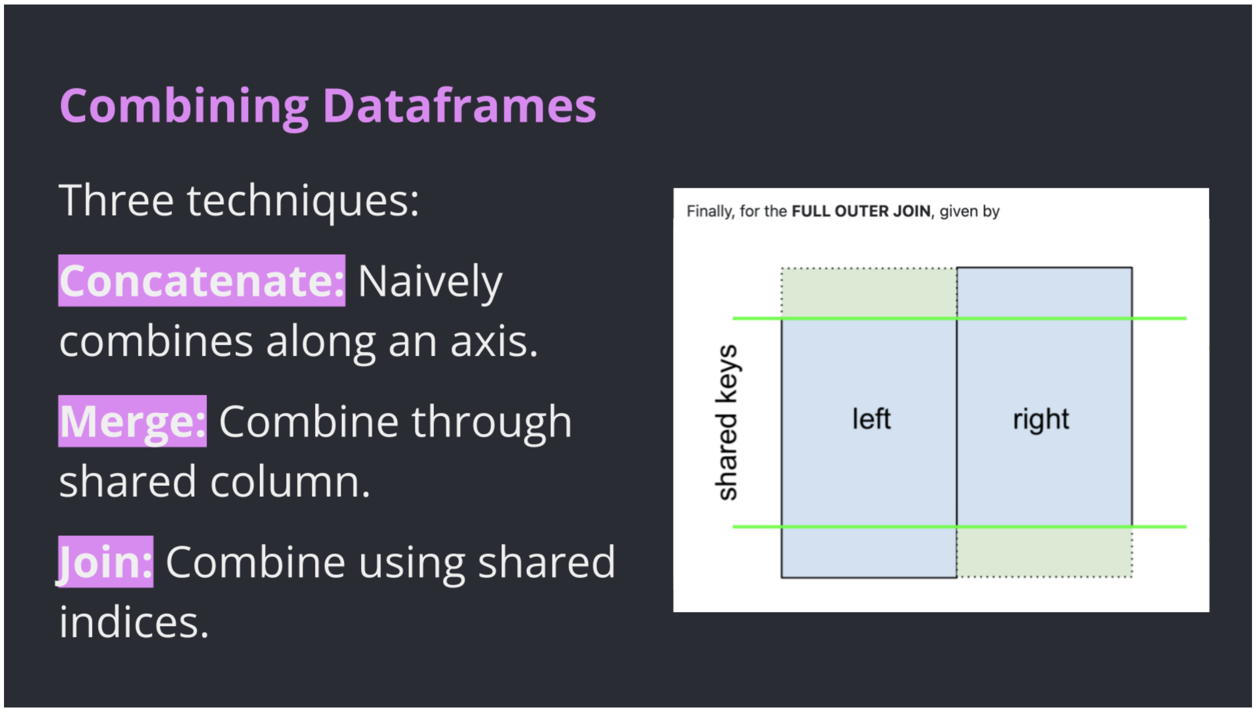
```
data = {
     "A": [1, 2, 3],
     "B": [4, 5, 6],
     "C": [7, 8, 9]
}

df = pd.DataFrame(data)
```

```
evens = df[df.iloc[:, :] % 2 == 0]
evens
```

	A	B	C
0	NaN	4.0	NaN
1	2.0	NaN	8.0
2	NaN	6.0	NaN

Check out this article!



Combining Dataframes

Three techniques:

Concatenate: Naively combines along an axis.

Merge: Combine through shared column.

Join: Combine using shared indices.

Finally, for the **FULL OUTER JOIN**, given by

shared keys
left
right