

## Pandas Fundamentals

- Pandas is like a toolbox to use for data scientists.

### 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[]`

Name	Indexing Pattern
loc	<code>name.loc[row_label, col_label]</code>
iloc	<code>name.iloc[row_index, col_index]</code>

### Selection

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

```
df.loc[0:1, ["A","B"]]
```

```
df.iloc[ : , len(df.columns))
```

### Filtering

Selecting values of a dataset where certain conditions are true.

A popular pattern is `df[condition]`

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

### Combining Dataframes

Three techniques:

Concatenate: Naively combines along an axis

Merge: combine through shared column

Join: Combine using shared indices