Data Wrangling in Python

Add Columns

dataFrame['newColumnName'] = value

Can do conditional things if you need to, but we'll talk about that later!

Rename Columns

dataFrame.**rename(columns={'oldName' : 'newname},
inplace=True)**

You use a key-value pair dictionary format!

inplace=True ensures that the changes gets made
permanently!

Combine Columns

• Must be the same data type to combine this way

```
dataFrame['newColumnName'] =
dataFrame['columnName'] + "separator" +
dataFrame['columnName'].map(str)
```

Separate Columns

newDataFrame **=** dataFrame['columnName']**.str.split(**'separator', **expand=True)**

Then add the columns into the dataFrame again:

newDataFrame **= pd.concat([dataFrame1, dataFrame2], axis =1)**

axis=1 specifies columns instead of rows

Subset Data

- Drop Columns

dataFrame.drop([columnName], axis=1, inplace = True)

- Keep Columns

newDataFrame = dataFrame[['column1', 'column2']]

- Choose Rows

newDataFrame = dataFrame[:index row to stop before]