Data Transformations in Python

What are Transformations?

- Code that changes the shape of your data

Transpose

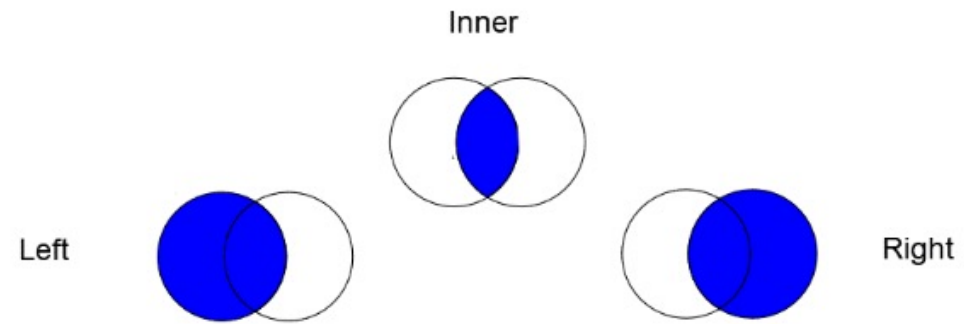**dataFrame.T**

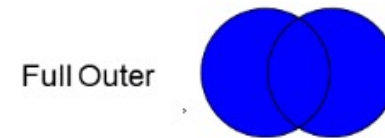| Student	Name	Grade
1	Terry	A
2	Tina	B
3	Tom	A

→

Student	1	2	3
Name	Terry	Tina	Tom
Grade	A	B	A

Types of Merges

- Like JOIN in SQL

Merge

NewDataFrame = pd.merge(data1, data2, on=ID)

Student	Name	Grade
1	Terry	A
2	Tina	B
3	Tom	A

+

Student	Class
1	Stats
2	R
3	Python

=

Student	Name	Grade	Class
1	Terry	A	Stats
2	Tina	B	R
3	Tom	A	Python

Append – Add Columns

- You don't need an id to merge on – just adds them on the left

- Be careful! Be certain your data's in the same order!

pd.concat([data1, data2], axis=1)

Name	Grade
Terry	A
Tina	B
Tom	A

+

Class
Stats
R
Python

=

Name	Grade	Class
Terry	A	Stats
Tina	B	R
Tom	A	Python

Append – Add Rows

- **Must have the same name and number of columns**
pd.concat([data1, data2])

Student	Name	Grade
1	Terry	A
2	Tina	B
3	Tom	A

+

Student	Name	Grade
4	Gina	A
5	Reggie	B
6	Rufus	C

=

Student	Name	Grade
1	Terry	A
2	Tina	B
3	Tom	A
4	Gina	A
5	Reggie	B
6	Rufus	C