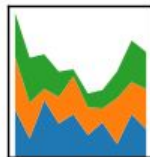
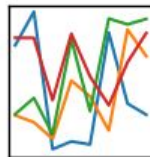


Combining DataFrames

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



Combining DataFrames

Two Main Ways:

1. Concatenating
2. Merging

Concatenating two DataFrames:

`pd.concat([<df1>, <df2>, <df3>])`

pass a **list** of dataframes to concatenate

df1

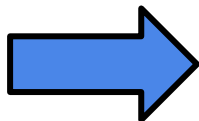
	A	B	C	D
0	A0	B0	C0	D0
1	A1	B1	C1	D1
2	A2	B2	C2	D2
3	A3	B3	C3	D3

df2

	A	B	C	D
4	A4	B4	C4	D4
5	A5	B5	C5	D5
6	A6	B6	C6	D6
7	A7	B7	C7	D7

df3

	A	B	C	D
8	A8	B8	C8	D8
9	A9	B9	C9	D9
10	A10	B10	C10	D10
11	A11	B11	C11	D11



Result

	A	B	C	D
0	A0	B0	C0	D0
1	A1	B1	C1	D1
2	A2	B2	C2	D2
3	A3	B3	C3	D3
4	A4	B4	C4	D4
5	A5	B5	C5	D5
6	A6	B6	C6	D6
7	A7	B7	C7	D7
8	A8	B8	C8	D8
9	A9	B9	C9	D9
10	A10	B10	C10	D10
11	A11	B11	C11	D11

- Same columns
- Like pasting them together

Merging two DataFrames:

`pd.merge(<df1>, <df2>, on = <col or list of cols to join on>, how = <join_type>)`

(or numerous other ways using the various possible arguments)

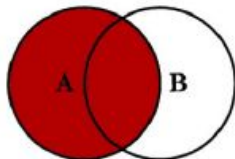
left				right				Result					
	key	A	B		key	C	D		key	A	B	C	D
0	K0	A0	B0	0	K0	C0	D0	0	K0	A0	B0	C0	D0
1	K1	A1	B1	1	K1	C1	D1	1	K1	A1	B1	C1	D1
2	K2	A2	B2	2	K2	C2	D2	2	K2	A2	B2	C2	D2
3	K3	A3	B3	3	K3	C3	D3	3	K3	A3	B3	C3	D3

- Need one or more “key” columns to join on
- Pastes matching rows together along the key column(s)

pandas merge types

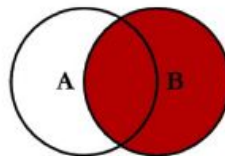
how = 'left'

Keeps all rows from the left table and only the matching rows from the right table.



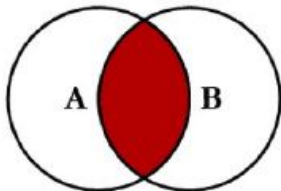
how = 'right'

Keeps all rows from the right table and only the matching rows from the left table.



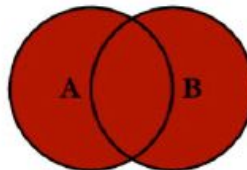
how = 'inner'

Keeps only rows that have a match in both tables.



how = 'outer'

Keeps all rows from both tables, whether they match on the specified key or not.



Question: Concatenate or Merge?

Interested in studying traffic accidents in Nashville, you download the 2017 and 2018 traffic accidents datasets from data.nashville.gov.


After reading in the two years as separate DataFrames, you wish to combine into a single DataFrame for analysis.


Traffic Accidents (2018) Police Dataset	
Details about traffic accidents reported to Metro Nashville Police Department in calendar year 2018. More	Created January 8, 2018 Views 32,161
Tags street, road, weather, fatality, motor, and 11 more API Docs	

Traffic Accidents (2017) Police Dataset	
Details about traffic accidents reported to Metro Nashville Police Department in calendar year 2017. More	Created October 24, 2017 Views 976
Tags street, road, weather, fatality, motor, and 11 more API Docs	

Question: Concatenate or Merge?

You are interested in studying the types of service requests that occur by council district. After your analysis, you want to incorporate information about the council member for each district.

hubNashville (311) Service Requests Public Services		 Dataset
Details of service requests to hubNashville, Metro Nashville government's comprehensive customer service system. Residents or visitors can connect with a Metro representative to request More		Updated September 5, 2020
Tags permit , recycling , litter , garbage , trash , and 21 more		Views 5,116
API Docs		

Metropolitan Council Members Metro Government		 Dataset
Names and contact information for each elected Member of the Metropolitan Council.		Updated October 7, 2019
Tags charter , at-large , district , representative , elected , and 8 more		Views 372
API Docs		