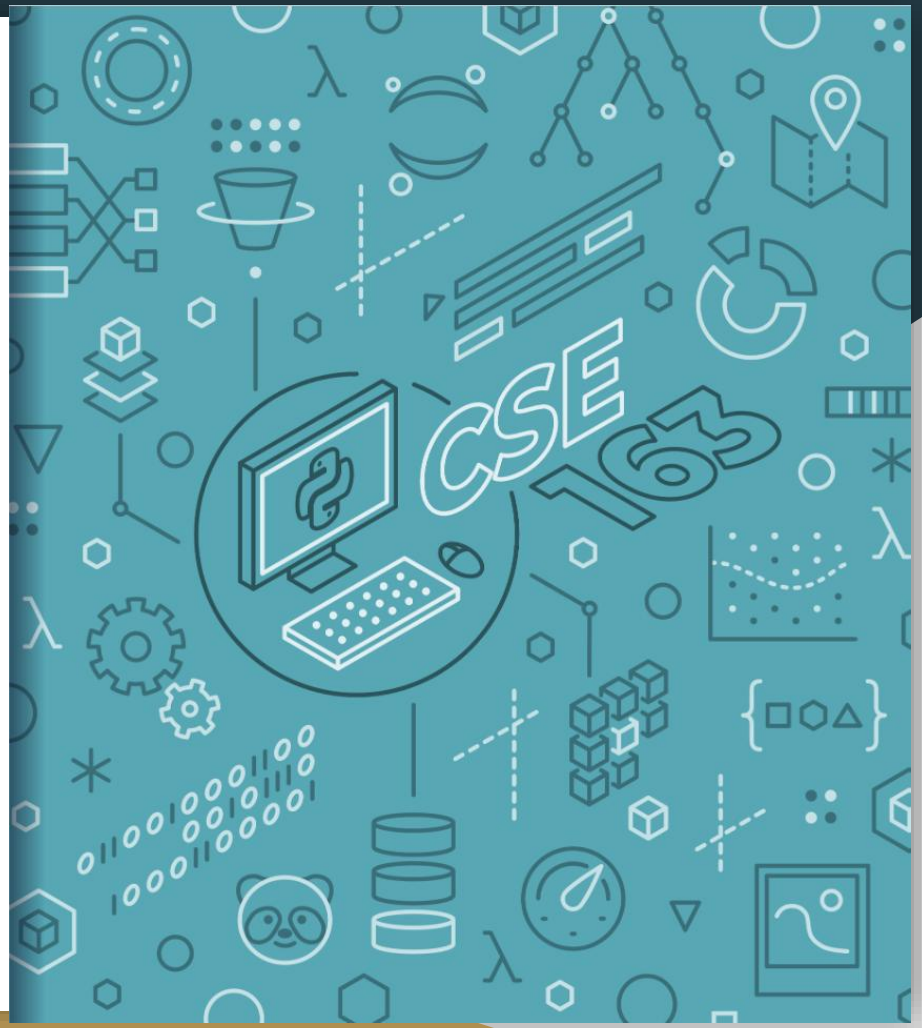
CSE 163
Pandas: Groupby & Indexing

Adrian Salguero
Spring 2026

💭 Icebreaker (discuss with neighbors): ‘
Favorite sunny day activity? Add to our Slido!

slido.com
#cse163



Announcements

- **Take Home Assessment 2: Networks** due Thursday, April 23rd at 11:59pm!
- **Lesson 10 Canvas Quiz** due tonight at 11:59pm!
- **Reading Assignment 3** due Monday, April 27th at 11:59pm!
- **Project Part 1** now avaialble, due Friday May 1st at 11:59pm on Gradescope!
  - Choose one of the following: Project (group of max 3) or Portfolio (individual)
  - You cannot switch once you've selected your option.
  - If you are doing a Project, only one person need submit on Gradescope but make sure your partners are added to your submission!

DataFrame

- One of the basic data types from pandas is a ***DataFrame!***

Columns

Index (row)

	id	year	month	day	latitude	longitude	name	magnitude
0	nc72666881	2016	7	27	37.672333	-121.619000	California	1.43
1	us20006i0y	2016	7	27	21.514600	94.572100	Burma	4.90
2	nc72666891	2016	7	27	37.576500	-118.859167	California	0.06

Groupby Demo

	col1	col2
0	A	1
1	B	2
2	C	3
3	A	4
4	C	5

Groupby Demo

	col1	col2
0	A	1
1	B	2
2	C	3
3	A	4
4	C	5

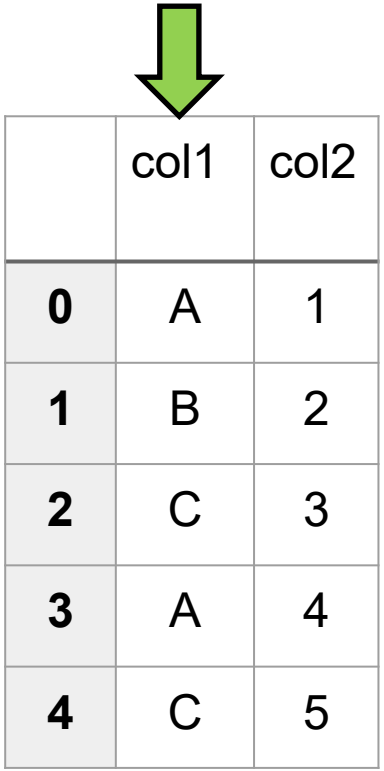
```
result = data.groupby('col1')
```

A	1
A	4

B	2

C	3
C	5

Groupby Demo



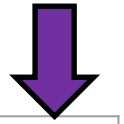
	col1	col2
0	A	1
1	B	2
2	C	3
3	A	4
4	C	5

```
result = data.groupby('col1')
```

A	1
A	4
B	2
C	3
C	5

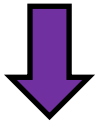
A *Groupby* DataFrame

Groupby Demo



```
result = data.groupby('col1')['col2']
```

	col1	col2
0	A	1
1	B	2
2	C	3
3	A	4
4	C	5



	col2
A	1
A	4
B	2
C	3
C	5

Groupby Demo

```
result = data.groupby('col1')['col2'].sum()
```

	col1	col2
0	A	1
1	B	2
2	C	3
3	A	4
4	C	5

col2

A	1
A	4
B	2
C	3
C	5

`.sum()`

`.sum()`

`.sum()`

Groupby Demo

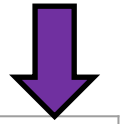
```
result = data.groupby('col1')['col2'].sum()
```

	col1	col2
0	A	1
1	B	2
2	C	3
3	A	4
4	C	5

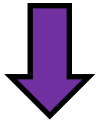
col2

A	5
B	2
C	8

Groupby Demo



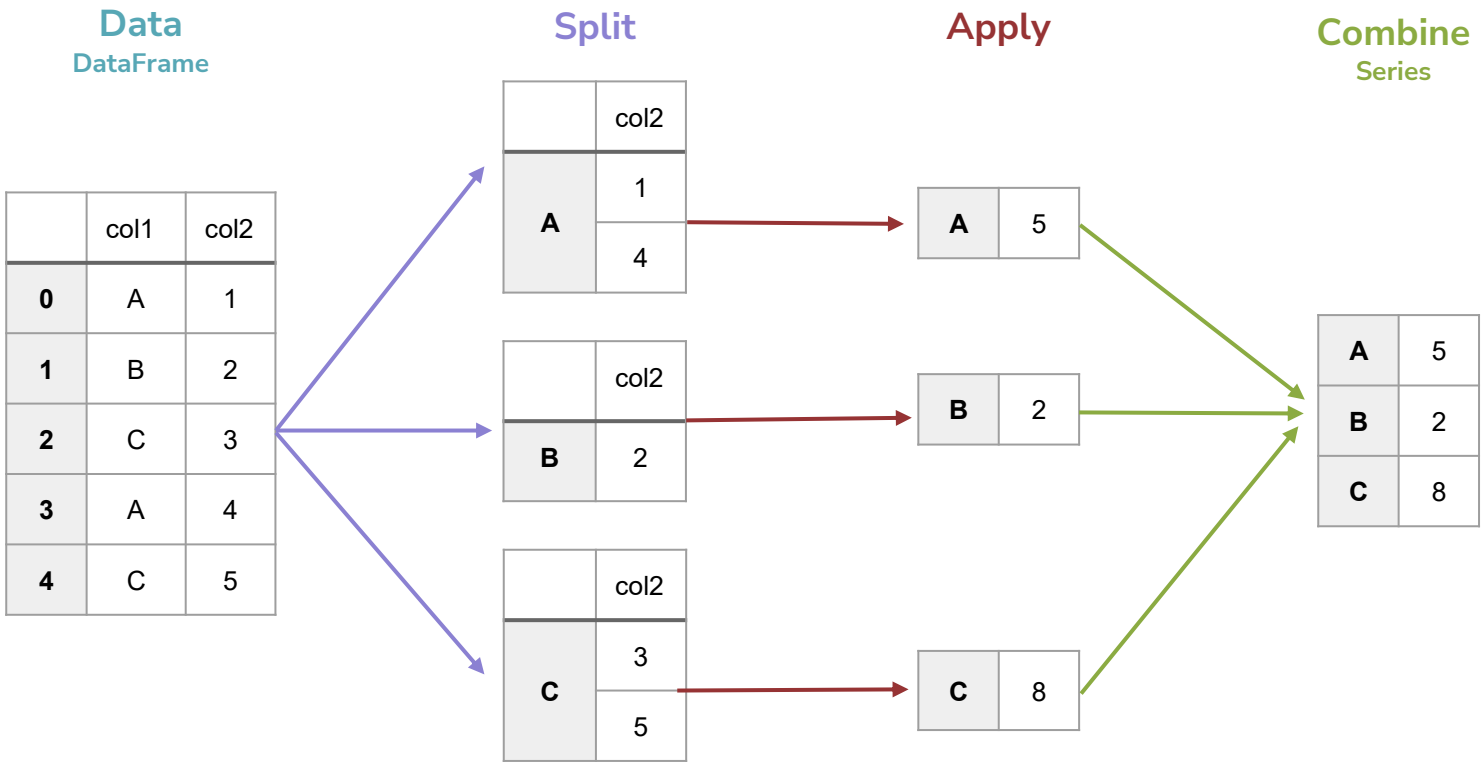
```
result = data.groupby('col1')['col2'].sum()
```



	col1	col2
0	A	1
1	B	2
2	C	3
3	A	4
4	C	5

A	5
B	2
C	8

```
result = data.groupby('col1')['col2'].sum()
```



Apply

- We have shown how to filter and group data, but what if you want to transform/modify your data
- Change numerical data with operators: + , -, /, *, abs(), min(), max(), etc.
- You can change Strings using built-in methods (e.g. len(), upper(), etc.) or even define your own custom function that works on Strings.

```
data['name'].str.len()
data['name'].str.upper()
data['name'].apply(len)     #Function as parameter
data['name'].apply(my_function) #Custom my_function
```