Pandas agg

WHAT YOU'LL LEARN

- How to use the Pandas agg method

- How to aggregate your data
  - calculate the mean
  - median
  - sum
  - etc

Pandas agg overview

The agg method "aggregates" your data

IN:
```
supercars.agg('mean')
```
← Specify an aggregation function
(like mean, etc)

OUT:
```
year               1994.520913
horsepower          304.655894
torque              301.180838
seconds_0_60          3.965491
engine_size         220.011316
top_speed           150.194515
weight                1.476125
dtype: float64
```
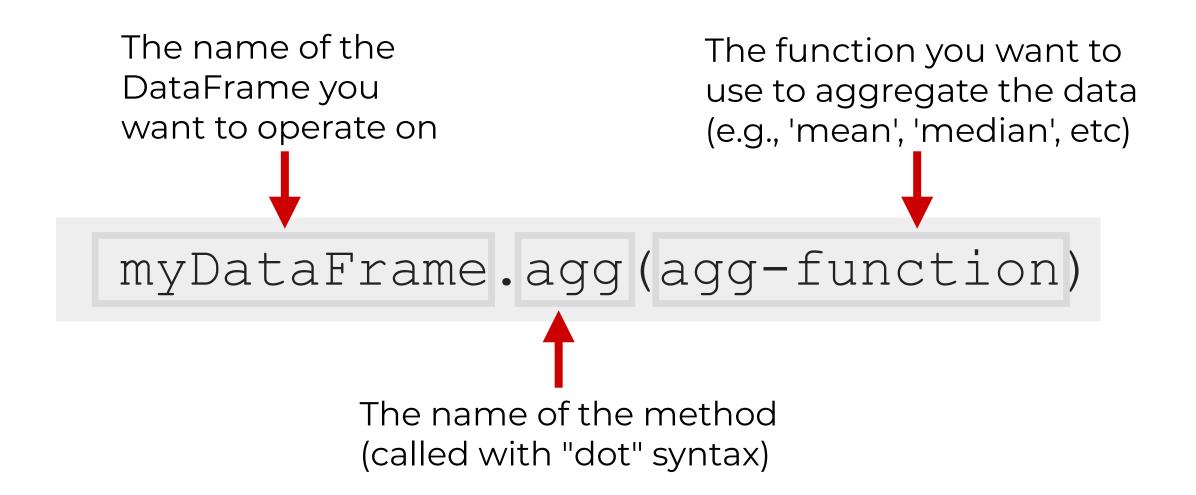← Pandas agg will calculate the summary statistic(s) that you specify on the variable(s) that you specify

Pandas agg syntax

Syntax: aggregate all variables

The name of the
DataFrame you
want to operate on

The function you want to
use to aggregate the data
(e.g., 'mean', 'median', etc)

```
myDataFrame.agg(agg-function)
```

The name of the method
(called with "dot" syntax)

Calling the function this way will apply the aggregation
function to all numeric variables

Syntax: aggregate one variable

The name of the
DataFrame you
want to operate on

The function you want to
use to aggregate the data
(e.g., 'mean', 'median', etc)

```
myDataFrame['var-name'].agg(agg-function)
```

The name of the variable you
want to aggregate
(called with "bracket" syntax)

Calling the function this way will apply the aggregation
function to a single variable only

SYNTAX: AGGREGATE SELECT VARIABLES

```
myDataFrame[['var-name']].agg(agg-function)
```

You can also provide a list of variable names to aggregate select variables

SYNTAX: USE AGGREGATION FUNCTIONS

You can also provide a list of multiple aggregation functions (e.g., `['mean', 'median']`)

```
myDataFrame['var-name'].agg(agg-functions)
```

This will calculate the specified aggregation functions for the variable(s)

Parameters of Pandas agg

THE PARAMETERS OF PANDAS AGG

Parameter	What it does	notes
`agg-function`	The aggregation function you want to apply to your data	Can be single function or multiple functions in a list
`axis`	The "direction" (i.e., axis) you want to perform the calculation	Default is 0 We almost never use this parameter

THE OUTPUT OF PANDAS AGG

The output of agg depends on
how the method is called…

If the input is a …	And agg is called with …	agg will return a …
Series	a single function	scalar
DataFrame	a single function	Series
DataFrame	several functions	DataFrame

Recap

RECAP OF WHAT WE LEARNED

- The Pandas `agg` method aggregates your data

- You can use agg to calculate a single statistic, or multiple statistics

- **Next Steps**: watch the code walkthrough for examples