DSC 10, Spring 2018
Lecture 9

Functions

sites.google.com/eng.ucsd.edu/dsc-10-spring-2018

Credit: Anindita Adhikari and John DeNero

Defining Functions

Defining functions

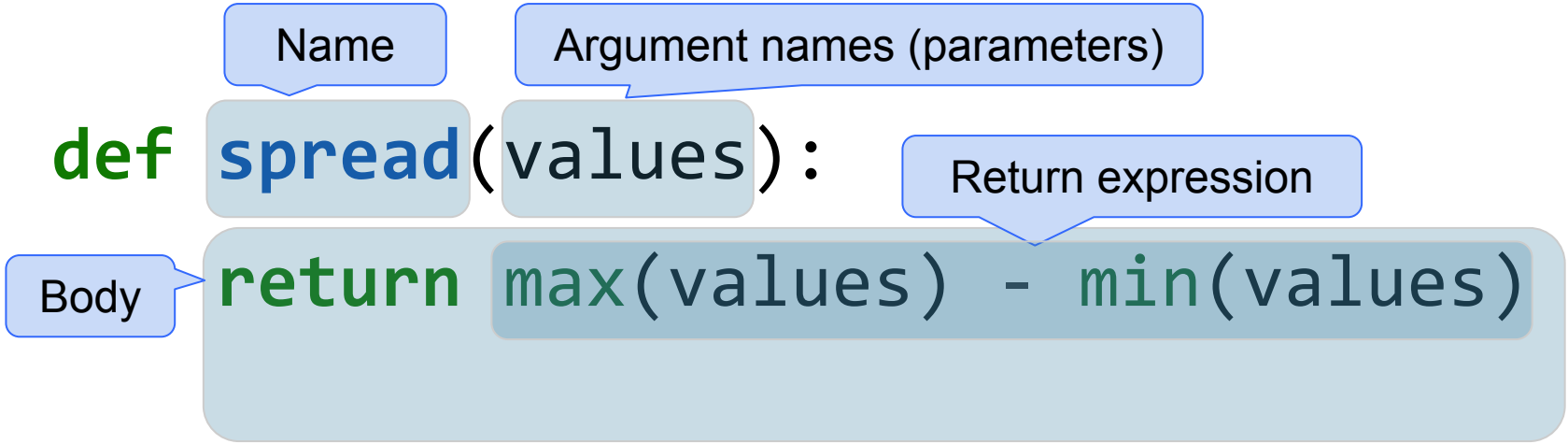
- You get to write your own functions.

- Useful when you want to do the same computation over and over again, and Python does not have a function for it.

(Demo)

Def Statements

User-defined functions give names to blocks of code



Discussion Question

```
def f(s):
    return np.round(s / sum(s) * 100, 2)
```

What does this function do?	What kind of input does it take?	What output will it give?
A. `Computes average`	`array of numbers`	`one number`
B. `Computes average`	`one number`	`one number`
C. `Computes percents`	`array of numbers`	`array of numbers`
D. `Computes percents`	`array of numbers`	`one number`
E. `None of the above`		

What's a reasonable name?

(Demo)

Apply

Apply

The `apply` method creates an array by calling a function on every element in input column(s)

- First argument: Function to apply
- Other argument(s): The input column(s)

```
table_name.apply(function_name, 'column_label')
```

Discussion Question

Input:

```
‘Data Science rocks!”
```

Output:

```
'Data Science rocks!: length is 19'
```

Define a function **str_len** that takes a string as a parameter and returns a new string that consists of:

- The given string
- A colon and a space
- “length is”
- The length of the string

Discussion Question

Input:

Title
Avatar
Star Wars: Episode I - The Phantom Menace
Star Wars
Star Wars: Episode III - Revenge of the Sith
Star Wars: Episode II - Attack of the Clones
Return of the Jedi

```
array(['Avatar: length is 6',
       'Star Wars: Episode I - The Phantom Menace: length is 41',
       'Star Wars: length is 9',
       'Star Wars: Episode III - Revenge of the Sith: length is 44',
       'Star Wars: Episode II - Attack of the Clones: length is 44',
       'Return of the Jedi: length is 18',
```

If the name of the table is ***top*** and the name of our function is ***str_len***, how do we find the length of each movie title?

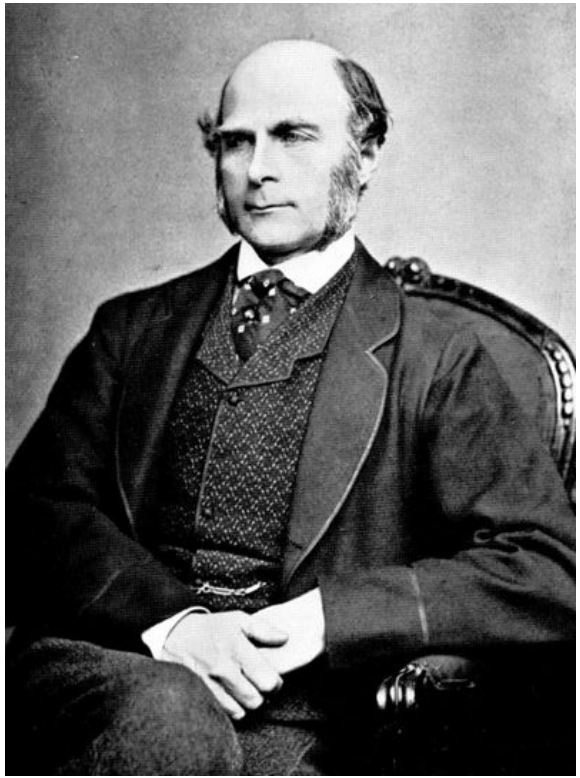
A. `top.apply(str_len(string), 'Title')`
B. `top.apply(str_len( ), 'Title')`
C. `top.apply(str_len, 'Title')`
D. `Title.apply(str_len, 'top')`
E. `Title.apply(str_len(), 'top')`

(Demo)

Example: Prediction

Sir Francis Galton

- 1822 - 1911 (knighted in 1909)
- A pioneer in making predictions
- Particular interest in heredity
- Charles Darwin's half-cousin



Apply with Multiple Arguments

Apply with multiple arguments

The `apply` method creates an array by calling a function on every element in one or more input columns

- First argument: Function to apply
- Other argument(s): The input column(s)

```
table_name.apply(one_arg_function, 'column_label')
table_name.apply(two_arg_function,
                 'column_label_for_first_arg',
                 'column_label_for_second_arg')
```

`apply` called with only a function applies it to each row

Extra Practice

```
def my_func():
    x = 10
    print("Value inside function:",x)

x = 20
my_func()
print("Value outside function:",x)
```

What is the value of x after this code is executed?

A. 10
B. 20
C. This code will not run because of an error.
D. None, x has no value since it is defined inside a function.

(Demo)

Extra Practice

Input:

	Title	Studio	Gross	Gross (Adjusted)	Year
	Avatar	Fox	760507625	846120800	2009
	Star Wars: Episode I - The Phantom Menace	Fox	474544677	785715000	1999
	Star Wars	Fox	460998007	1549640500	1977

Output:

	Title	Studio	Gross	Gross (Adjusted)	Year	Difference
	Avatar	Fox	760507625	846120800	2009	85613175
	Star Wars: Episode I - The Phantom Menace	Fox	474544677	785715000	1999	311170323
	Star Wars	Fox	460998007	1549640500	1977	1088642493
	Star Wars: Episode III - Revenge of the Sith	Fox	380270577	516123900	2005	135853323

Discuss how you would create the output table.

Way 1) Defining a function and using **`apply`**.

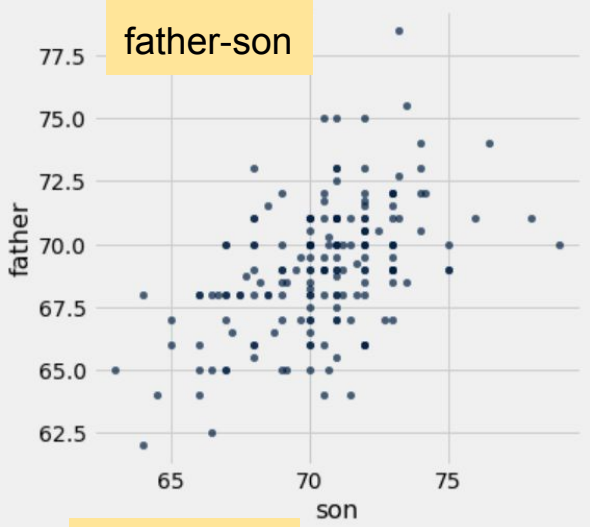
Way 2) Without defining a function or using **`apply`**.

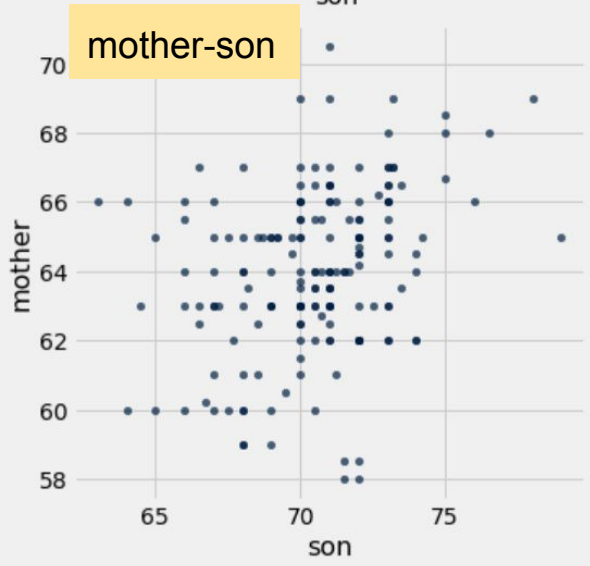
(Demo)

Father or Mother?

Is a son's height more influenced by his father's height or his mother's height?

A. Father, because difference between father and son height is smaller than difference between mother and son height.
B. Mother, because there is more variability in mother's heights than father's heights.
C. Father, because the points on the father-son plot more strongly resemble a line than those on the mother-son plot.
D. Father, because the points on the father-son plot form a steeper curve than the those on the mother-son plot.





Discussion Question

This histogram describes a **year** of daily temperatures in degrees F
(horizontal: temperature (degrees F); vertical: percent per degree F)

Try to answer these questions:

- What proportion of days had a high temp in the range 60-69?
- What proportion had a low of 45 or more?
- What proportion of days had a difference of more than 20 degrees between their high and low temperatures?