

# Introduction to Python

Define a variable:

```
In [1]: a = 1  
a
```

```
Out[1]: 1
```

What type of object is this?

```
In [2]: type(a) # Integer
```

```
Out[2]: int
```

```
In [3]: type(1.2)
```

```
Out[3]: float
```

```
In [4]: firm = 'Microsoft'
```

What type of object is this?

```
In [5]: type(firm) # string
```

```
Out[5]: str
```

Add integers:

```
In [6]: b = 2  
a + b
```

```
Out[6]: 3
```

Add "firm" and "b":

```
In [7]: firm + b
```

```
-----  
-----  
TypeError                                Traceback (most recent call last)  
  <ipython-input-7-da48f3104579> in <module>  
    ----> 1 firm + b  
  
TypeError: can only concatenate str (not "int") to str
```

We need to convert `b=2` to a string if we want to combine (concatenate) it with 'Microsoft':

```
In [8]: str(b)  # convert b to string (this creates a new object, b is still integer 2)
```

```
Out[8]: '2'
```

```
In [9]: firm + str(b)
```

```
Out[9]: 'Microsoft2'
```

What is firm now?

```
In [10]: firm
```

```
Out[10]: 'Microsoft'
```

If we want to change the content of "firm", we need to reassign it:

```
In [11]: firm = firm + str(b)  
         firm
```

```
Out[11]: 'Microsoft2'
```

Get character 'M' of Microsoft:

```
In [12]: firm = 'Microsoft'  
  
         firm[0]  # counting starts with 0
```

```
Out[12]: 'M'
```

```
In [13]: firm[1]
```

```
Out[13]: 'i'
```

```
In [14]: firm[-1] # last character
```

```
Out[14]: 't'
```

Find the position of 'c' in Microsoft:

```
In [15]: firm.find('c')
```

```
Out[15]: 2
```

Find 'o':

```
In [16]: firm.find('o') # first o
```

```
Out[16]: 4
```

Reverse find:

```
In [17]: firm.rfind('o') # last o
```

```
Out[17]: 6
```

"type(x)" is a function that returns the type of x.

".find('x')" is a string method (a function attached to strings).

Use a string method for integers:

```
In [18]: a.find(1) # Does not work -> find() can only be attached to strings
```

```
-----  
----  
AttributeError                                Traceback (most recent call 1  
ast)  
<ipython-input-18-002873b81352> in <module>  
----> 1 a.find(1) # Does not work -> find() can only be attached to s  
trings  
  
AttributeError: 'int' object has no attribute 'find'
```

Substring of firm:

```
In [19]: firm[1:4] # characters 1,2,3 (4 not included)
```

```
Out[19]: 'icr'
```

Don't specify first index:

```
In [20]: firm[:4] # from the beginning until 4, 4 not included
```

```
Out[20]: 'Micr'
```

Don't specify last index:

```
In [21]: firm[3:] # from 3 until the end
```

```
Out[21]: 'rosoft'
```

Same, counting from the back:

```
In [22]: firm[-6:]
```

```
Out[22]: 'rosoft'
```

Let's use Python to access information on the web.

Start by importing the Python request library:

```
In [23]: import requests
```

Get WSJ website:

```
In [24]: requests.get('https://www.wsj.com') # 403: forbidden
```

```
Out[24]: <Response [403]>
```

Get the SEC ticker symbols file:

```
In [25]: requests.get('https://www.sec.gov/files/company_tickers.json') # 200: o  
k
```

```
Out[25]: <Response [200]>
```

Get text from website:

```
In [26]: symbols = requests.get('https://www.sec.gov/files/company_tickers.json')  
.text # this is a string
```

First 200 characters of symbols:

```
In [27]: symbols[:200]
```

```
Out[27]: '{"0":{"cik_str":320193,"ticker":"AAPL","title":"Apple Inc."},"1":{"cik_str":789019,"ticker":"MSFT","title":"MICROSOFT CORP"},"2":{"cik_str":1018724,"ticker":"AMZN","title":"AMAZON COM INC"},"3":{"cik_'
```

Find index of Microsoft:

```
In [28]: i = symbols.find('MICROSOFT')
i
```

```
Out[28]: 108
```

Check this:

```
In [29]: symbols[i]
```

```
Out[29]: 'M'
```

What is the index of the closing bracket } for Microsoft?

Find closing bracket:

```
In [30]: symbols.find('}')
```

```
Out[30]: 59
```

But this is the first closing bracket in symbols, so it belongs to Apple.

Get string starting at position i:

```
In [31]: symbols[i:]          [:100] # show first 100 characters
```

```
Out[31]: 'MICROSOFT CORP"},"2":{"cik_str":1018724,"ticker":"AMZN","title":"AMAZO
N COM INC"},"3":{"cik_str":165'
```

```
In [32]: symbols[i:].find('}') # counting from the position of 'M'
```

```
Out[32]: 15
```

```
In [33]: last = i + symbols[i:].find('}') # 108 + 15
last
```

```
Out[33]: 123
```

```
In [34]: symbols[last] # check result
```

```
Out[34]: '}'
```

Find the index of the opening bracket:

```
In [35]: symbols.rfind('{') # rfind: last { in entire string
```

```
Out[35]: 860785
```

```
In [36]: symbols[:i] # everything until 108 ('MICROSOFT')
```

```
Out[36]: '{"0":{"cik_str":320193,"ticker":"AAPL","title":"Apple Inc."},"1":{"cik_str":789019,"ticker":"MSFT","title":""'}
```

```
In [37]: first = symbols[:i].rfind('{')
first
```

```
Out[37]: 65
```

Get entire entry:

```
In [38]: symbols[first:last+1]
```

```
Out[38]: '{"cik_str":789019,"ticker":"MSFT","title":"MICROSOFT CORP"}'
```

Find Walmart:

```
In [39]: i = symbols.find('Walmart')
first = symbols[:i].rfind('{')
last = i + symbols[i:].find('}')
symbols[first:last+1]
```

```
Out[39]: '{"cik_str":104169,"ticker":"WMT","title":"Walmart Inc."}'
```

Problem: some firms are in all caps.

Solution: convert entire string to lower case and then search with lower case letters.

Example:

```
In [40]: symbols.lower()[:100] # show first 100, note the lower case
```

```
Out[40]: '{"0":{"cik_str":320193,"ticker":"aapl","title":"apple inc."},"1":{"cik_str":789019,"ticker":"msft","title":""'}
```

```
In [41]: symbols.lower().find('microsoft') # now case does not matter!
```

```
Out[41]: 108
```

Entire code:

```
In [42]: i = symbols.lower().find('walmart')
first = symbols[:i].rfind('{')
last = i + symbols[i:].find('}')
symbols[first:last+1]
```

```
Out[42]: '{"cik_str":104169,"ticker":"WMT","title":"Walmart Inc."}'
```

Let's define a search function that runs these 4 lines so that we don't have to repeat this code everytime we look for a firm.

We can define our own custom functions like this:

```
In [43]: def add_this(x,y): # This function takes 2 arguments
        result = x + y
        return result
```

Run the function:

```
In [44]: add_this(2,3)
```

```
Out[44]: 5
```

```
In [45]: z = add_this('a','b')
z
```

```
Out[45]: 'ab'
```

Our search function:

```
In [46]: def find_firm(name): # 1 input argument
        i = symbols.lower().find(name) # find the input
        first = symbols[:i].rfind('{')
        last = i + symbols[i:].find('}')
        return symbols[first:last+1]
```

```
In [47]: find_firm('tesla')
```

```
Out[47]: '{"cik_str":1318605,"ticker":"TSLA","title":"Tesla, Inc."}'
```

Search by ticker symbol also works:

```
In [48]: find_firm('pton') # we can search by anything inside the {...}
```

```
Out[48]: '{"cik_str":1639825,"ticker":"PTON","title":"PELOTON INTERACTIVE, IN C."}'
```

Nice!