The most important library is SciPy for python: Pandas and matplot toolkits are from SciPy

Google trends

Data driven society

It is not limited to tech companies only

Data driven intelligence

Statistical knowledge, spceiliased and computing (hacking skills)

\_

Python is high level language

Interactive

Don’t need to compile like c and java

Def is used to define a function

You can assign a variable to a function! Normal…

We can assign a function to a variable

You can append a list [ ] but you cannot append a tuple ( )

You can iterate in a list or a tuple

If you have a list the index is [] e.g. x[1]

Tuples and lists can be concatenated by +

Lists only can be repeated using \* example [1]\*3 but not tuples

In can be used to see if a value is inside a tuple or a list

Slicing is with indexes

Example x[1:5]

String is a list of characters so you can use slicing directly

[-1] or minus is from the end

If empty it is from default beginning or end

Regular expressions is about finding a special pattern in a set

Which is used in text mining

The function split: how does it work:

stringYouWant.split(here the pattern you wanna split with e.g. ‘ ‘) for a space

this will return a list

then you choose the index of the element you want

Dictonary is a lbeled collection

It has a key

Which is called a map

It is with { }

The value and key is separated by :

You can check the value by its key

e.g. x[‘key’]

if x is a dixtionary

x.values() will give the value without key

x.items() will give everything

you can assign a tuple or a list to variables

tuple: ()

list: []

dictionary: {}

UNICODE (UTF) is way better than ASCII

Str() to convert to string

With a string ().format() you don’t need to change to string etc when you combine int with a string to print

Inside format is a dictionary

To convert a csv file to a list of dictionaries:

With open (file name) as csvfile:

Mpg = list(csv.dictReader(csvfile))

So the csvReader is the fuction

To get the dictionary keys use the keys method

dictionaryName[1].keys()

example to get the average for a certain key value in a list of dictionaries

sum(int(l['cyl']) for l in mpg)/ len(mpg)

Use `set` to return the unique values for the number of cylinders the cars in our dataset have.

cylinders = set(d['cyl'] for d in mpg)

Get back to the notebook for date and time

Python can be object oriented

Where you define a class

Map is used to iterate a function among iteratable objects

Lambda is anyn function without name

Syntax

Lambda list-of-arguments : function-to-be-done

Get more familiar with list comprehension

Where you convert a for loop to a one line

[Number for number in range (0,1000) if number % 2 ==0]

Will give a list!!

times\_tables = [i\*j for i in range(10) for j in range(10))]

Import numpy

To convert a list to a numpy array: np.array(list-name)

We can have a multi-dimensional array np.array (list-one, list-two)

Array name if it is m

m.shape will give array dimension

to create array with range: np.arange(0,30,2)

will give even numbers from 0 to 30

docker run -p 8888:8888 -v C:\Users\uceehgh\Documents\DS4A:/home/jovyan/work jupyter/scipy-notebook:17aba6048f44 --name DS4A

for number of elements instead of the step size use linspace not arrange

np.resize

np.reshape are the same

np.zeros, np.ones, np.eye(3), ni.diag()

np.repeat([array], how many times)

np.vstack or hstack to stack arrays horizontally otr vertically

the to power (^) is \*\* in numpy

dot product is

x.dot(y)

len(array) gives the number of rows

array.T for transpose

array.dtype to know the array type

array.astype(‘f’) to convert to float for example

A second `:` can be used to indicate step-size. `array[start:stop:stepsize]`

s[-5::-2]

To make a copy of an array: array.copy()