Data camp 笔记1

String is immutable

s = “abcd”

s[0] = “y” report error

s = “y”+s[1:4] is allowd

Python list

第一位的index 是0

最后一位可以表示为-1

list[1:n] 提取包括index为1 末尾到n-1的元素

已有list 名为 area

area\_copy = area

area\_copy[5] = 1

也会改变area中原来的取值

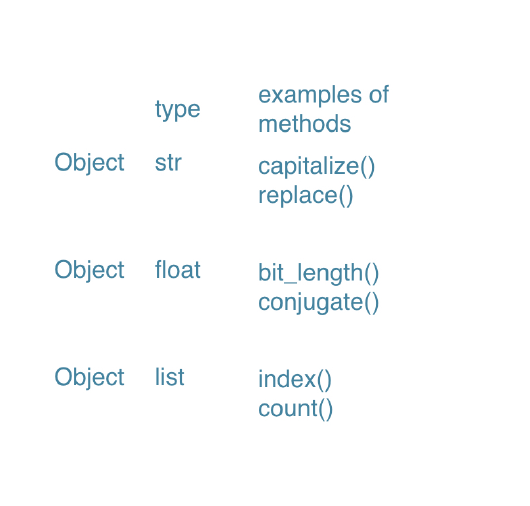
正确赋值是 area\_copy = area[:]

areas.append(“x”) 在area的list后面加上x

areas.count(“x”) 计算areas的list里面有几个x

areas.count（“x“）给出x在area里的有几个

areas.index(“x”) 给出x在area里面的index



调用numpy的package

这样的话使用numpy里面的function array用法是numpy.array

import numpy as np

调用numpy的package 名字可以命名为np

这样的话使用numpy里面的function array用法是np.array

也可以只调用arrary

from numpy import array

然后直接array（）

create a new array

np.array([1, 2,3])

注意一定要加括号

注意调用package以后里面的dataset是直接使用的 不用再加包的名字

例如在numpy里面的dataset height\_in 可以直接使用 type(height\_in)

多dimension 的array有点像R里面matrix的用法

如果是提取某一行某一列的 直接是x[1, 3] 如果x是array

而x如果是list 则是x[1][3]

想知道x的dimension

命令是x.shape 前提是已经调用了numpy package

作图

调用 matplotlib.pyplot 的package

import matplotlib.pyplot as plt

line plot

plt.plot(x,y)

scatter plot

plt.scatter(x,y)

plt.hist(x, bins= )

改变x轴为log

plt.xscale(‘log’)

显示图

plt.show()

清除

plt.clf()

改axis title 名称

plt.xlabel()

plt.ylabel()

plt.title()

换x y 轴上某个坐标的lab

用tick命令 即y轴上对应的0,1,2 被对应换成 one two three

plt.yticks([0,1,2], ["one","two","three"])

plot.scatter里面的s 是size

alpha 透明度

c 是颜色

加文字注释 前面两个是坐标

# Additional customizations

plt.text(1550, 71, 'India')

plt.text(5700, 80, 'China')

图像有网格的话是grid

# Add grid() call

plt.grid(True)

# Definition of countries and capital

countries = ['spain', 'france', 'germany', 'norway']

capitals = ['madrid', 'paris', 'berlin', 'oslo']

# Get index of 'germany': ind\_ger

ind\_ger = countries.index('germany')

# Use ind\_ger to print out capital of Germany

print(capitals[ind\_ger])

get index commands to get the index of a specific element

xxxx.index()