HW2

May 19, 2017

1. Finish creating the following function that takes a list and returns the average value.

```
In [4]: def average(my_list):
            """
            This function computes the arithmetic average
            """
            total = 0
            for item in my_list:
                total += item

            return total/len(my_list)

        average([1,2,1,4,3,2,5,9])
```

```
Out[4]: 3.375
```

2. Using a Dictionary keep track of the count of numbers (or items) from a list

```
In [2]: def counts(my_list):
            """
            This function keeps track of the count of numbers from a list
            """
            counts = dict()
            for item in my_list:
                counts[item] = my_list.count(item)

            return counts

        counts([1,2,1,4,3,2,5,9])
```

```
Out[2]: {1: 2, 2: 2, 3: 1, 4: 1, 5: 1, 9: 1}
```

3. Using the counts() function and the .split() function, return a dictionary of most occuring words from the following paragraph. Bonus, remove punctuation from words

```
In [3]: paragraph_text = '''
        For a minute or two she stood looking at the house, and wondering what to d
        The Fish-Footman began by producing from under his arm a great letter, near
        Then they both bowed low, and their curls got entangled together.
```

```
Alice laughed so much at this, that she had to run back into the wood for f
Alice went timidly up to the door, and knocked.
'There's no sort of use in knocking,' said the Footman, 'and that for two r
'Please, then,' said Alice, 'how am I to get in?'
'There might be some sense in your knocking,' the Footman went on without a
'I shall sit here,' the Footman remarked, 'till tomorrow--'
At this moment the door of the house opened, and a large plate came skimmin

import string
punct = set(string.punctuation)
a = "".join(x for x in paragraph_text if x not in punct)
counts(a.split())
```

Out[3]:
```
{'Alice': 5,
 'An': 2,
 'And': 1,
 'At': 1,
 'But': 1,
 'Duchess': 2,
 'First': 1,
 'FishFootman': 2,
 'Footman': 3,
 'Footman's': 1,
 'For': 2,
 'FrogFootman': 1,
 'He': 1,
 'I': 3,
 'It': 1,
 'I'm': 1,
 'Queen': 2,
 'She': 1,
 'The': 2,
 'Then': 1,
 'a': 15,
 'about': 1,
 'against': 1,
 'all': 3,
 'aloud': 1,
 'am': 2,
 'and': 16,
 'another': 1,
 'answer': 1,
 'any': 1,
 'are': 2,
 'arm': 1,
 'as': 4,
 'at': 6,
 'attending': 1,
```

```
'back': 1,
'be': 2,
'because': 3,
'been': 1,
'began': 1,
'behind': 1,
'between': 1,
'both': 2,
'bowed': 1,
'broke': 1,
'broken': 1,
'by': 3,
'called': 1,
'came': 2,
'can’t': 1,
'certainly': 1,
'changing': 1,
'considered': 1,
'constant': 1,
'could': 2,
'crash': 1,
'crept': 1,
'croquet’': 2,
'curious': 1,
'curled': 1,
'curls': 1,
'decidedly': 1,
'dish': 1,
'do': 1,
'door': 6,
'entangled': 1,
'every': 1,
'extraordinary': 1,
'eyes': 2,
'face': 2,
'fear': 1,
'felt': 1,
'fish--and': 1,
'footman': 3,
'footmen': 1,
'for': 3,
'frog': 1,
'from': 2,
'get': 2,
'going': 1,
'gone': 1,
'got': 1,
'grazed': 1,
```

```
'great': 2,
'ground': 1,
'had': 4,
'hair': 1,
'handed': 1,
'have': 1,
'he': 5,
'head': 2,
'heads': 1,
'hear': 1,
'hearing': 1,
'help': 1,
'her': 2,
'here’': 1,
'herself': 1,
'him': 3,
'himself': 1,
'his': 5,
'house': 2,
'howling': 1,
'if': 2,
'in': 7,
'inside': 2,
'instance': 1,
'into': 3,
'invitation': 2,
'in’': 2,
'it': 2,
'it’': 1,
'judging': 1,
'just': 1,
'kettle': 1,
'knock': 1,
'knocked': 1,
'knocking’': 2,
'know': 1,
'know’': 1,
'knuckles': 1,
'large': 3,
'laughed': 1,
'let': 1,
'letter': 1,
'like': 1,
'listen': 1,
'little': 2,
'livery': 3,
'looking': 2,
'loudly': 1,
```

```
'low': 1,
'making': 1,
'might': 3,
'minute': 1,
'moment': 1,
'most': 1,
'much': 1,
'near': 1,
'nearly': 2,
'next': 2,
'no': 2,
'noise': 2,
'nose': 1,
'noticed': 1,
'now': 1,
'of': 9,
'on': 4,
'one': 2,
'only': 2,
'opened': 2,
'or': 2,
'order': 1,
'other': 2,
'otherwise': 1,
'out': 5,
'over': 2,
'peeped': 1,
'perhaps': 1,
'pieces': 2,
'plate': 1,
'play': 2,
'possibly': 1,
'powdered': 1,
'producing': 1,
'questions--How': 1,
'rapped': 1,
'rate': 1,
'reasons': 1,
'remarked': 1,
'repeated': 2,
'round': 1,
'run': 1,
'running': 1,
'said': 3,
'same': 2,
'saying': 1,
'secondly': 1,
'sense': 1,
```

```
'shall': 1,
'she': 6,
'side': 1,
'sit': 1,
'sitting': 1,
'skimming': 1,
'sky': 2,
'sneezing': 1,
'so': 2,
'solemn': 2,
'some': 1,
'sort': 1,
'speaking': 1,
'staring': 1,
'stood': 1,
'straight': 1,
'stupidly': 1,
'such': 1,
'suddenly': 1,
'that': 3,
'the': 32,
'their': 3,
'then': 1,
'then’': 1,
'there': 1,
'they': 1,
'they’re': 1,
'this': 4,
'thought': 1,
'time': 1,
'timidly': 1,
'to': 15,
'together': 1,
'tomorrow--’': 1,
'tone': 2,
'top': 1,
'trees': 1,
'two': 2,
'uncivil': 1,
'under': 1,
'up': 3,
'us': 1,
'use': 1,
'very': 2,
'was': 8,
'way': 1,
'we': 1,
'went': 2,
```

```
 'were': 1,
 'what': 2,
 'when': 2,
 'with': 2,
 'within--a': 1,
 'without': 1,
 'wondering': 1,
 'wood': 2,
 'wood--she': 1,
 'words': 1,
 'would': 1,
 'you': 5,
 'your': 1,
 'you’': 1,
 '‘But': 1,
 '‘For': 1,
 '‘From': 1,
 '‘I': 1,
 '‘Please': 1,
 '‘There': 1,
 '‘There’s': 1,
 '‘and': 1,
 '‘his': 1,
 '‘how': 1,
 '‘if': 1,
 '‘till': 1}
```

1.Read in a file and write each line from the file to a new file Title-ized

```
In [25]: Str1 = 'data science is great'
         Str1.upper()

Out[25]: 'DATA SCIENCE IS GREAT'
```

1. Given a list, find the average using a numpy function.

```
In [27]: import numpy as np
         simple_list  = [1,2,1,4,3,2,5,9]
         x = np.array(simple_list)
         x.mean()

Out[27]: 3.375
```

2. Given two lists of Heights and Weights of individual, calculate the BMI of those individuals, without writing a for-loop

```
In [32]: import numpy as np
         heights = np.array([174, 173, 173, 175, 171])
         weights = np.array([88, 83, 92, 74, 77])
         BMI = 703*weights/(weights*heights)
         BMI
```

```
Out[32]: array([ 4.04022989,  4.06358382,  4.06358382,  4.01714286,  4.11111111])
```

3. Create an array of length 20 filled with random values (between 0 to 1)

```
In [34]: import numpy as np

         x = np.random.uniform(0, 1,20)
         x

Out[34]: array([ 0.6713075 ,  0.71660997,  0.16459845,  0.48139437,  0.95194432,
                 0.97891152,  0.25273975,  0.39495007,  0.29857946,  0.41923733,
                 0.53982213,  0.75874212,  0.12938339,  0.4829185 ,  0.85905321,
                 0.42931614,  0.30386443,  0.2875226 ,  0.11127129,  0.18364928])
```

```
In [ ]: %%file hw_data.csv
        id,sex,weight,height
        1,M,190,77
        2,F,120,70
        3,F,110,68
        4,M,150,72
        5,O,120,66
        6,M,120,60
        7,F,140,70
```

```
In [1]: import pandas as pd
        df1=pd.read_csv("/Users/e_chertow/Documents/Ahmed/Documents/JhuMachineLearr
```

```
In [5]: print(df1)
        df1.dtypes
```

```
   id sex  weight  height
0   1   M     190      77
1   2   F     120      70
2   3   F     110      68
3   4   M     150      72
4   5   O     120      66
5   6   M     120      60
6   7   F     140      70


Out[5]: id          int64
        sex        object
        weight      int64
        height      int64
        dtype: object
```

2. Find the average weight

```
In [6]: df1["weight"].mean()
```

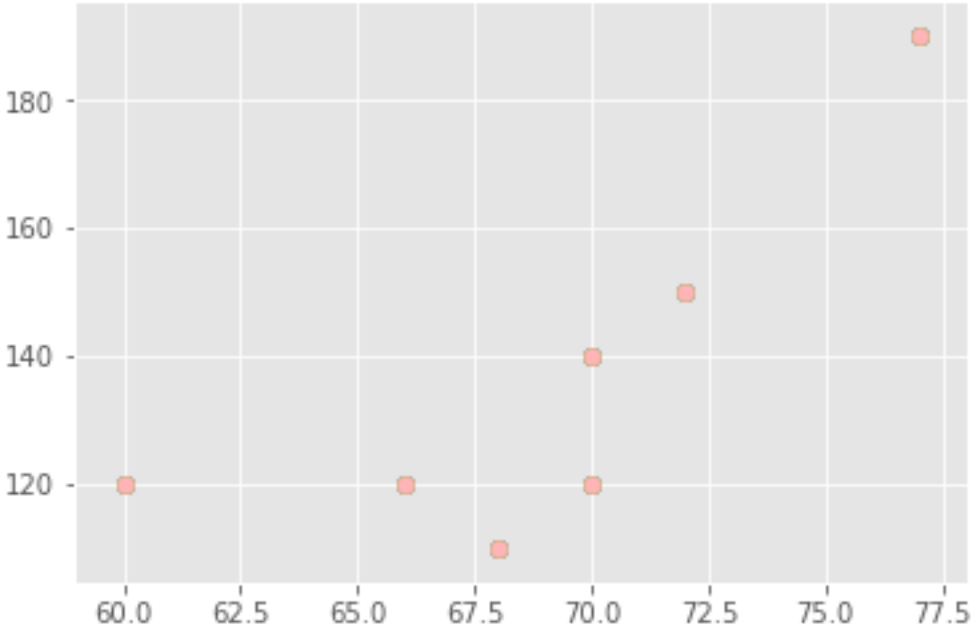
```
Out[6]: 135.71428571428572
```

Find the Value Counts on column sex

```
In [10]: df1['sex'].value_counts()

Out[10]: F    3
         M    3
         O    1
         Name: sex, dtype: int64
```

```
In [17]: import matplotlib.pyplot as plt
         plt.style.use('ggplot')
         plt.scatter(df1['height'],df1['weight'])
         plt.show()
```



Calculate BMI and save as a new column

```
In [18]: df1['BMI'] = 703*df1['weight']/(df1['weight']*df1['height'])

In [19]: print(df1)

   id sex  weight  height        BMI
0   1   M     190      77   9.129870
1   2   F     120      70  10.042857
2   3   F     110      68  10.338235
3   4   M     150      72   9.763889
```

```
4   5   O     120      66  10.651515
5   6   M     120      60  11.716667
6   7   F     140      70  10.042857
```

Save sheet as a new CSV file hw_dataB.csv

```
In [20]: df1.to_csv("/Users/e_chertow/Documents/Ahmed/Documents/JhuMachineLearning/
```

```
In [1]: !cat /Users/e_chertow/Documents/Ahmed/Documents/JhuMachineLearning/hw_dataB
```

```
,id,sex,weight,height,BMI
0,1,M,190,77,9.12987012987013
1,2,F,120,70,10.042857142857143
2,3,F,110,68,10.338235294117647
3,4,M,150,72,9.76388888888889
4,5,O,120,66,10.651515151515152
5,6,M,120,60,11.716666666666667
6,7,F,140,70,10.042857142857143
```

```
In [ ]:
```