Scientific Programming in Python — Exercise Sheet 5

(submission: 13/05/2017 8pm)

Task 1: Installing Pandas and Matplotlib

OS X/Linux: through pip. For matplotlib, Linux users will also need to install **python-tk** through the package manager.

Windows: `http://www.lfd.uci.edu/~gohlke/pythonlibs/`

Task 2: == Special Submission Task(s) ==

On stud.ip, you'll find **Pokemon.csv** which we'll be using as our dataset for this homework.

Create a new folder in your repo called **pandas_exercises**. In there create a file called **pokemon.py**. You'll be creating a few functions that take in a data frame (already populated with the Pokemon data) and return a data frame containing whatever that function is looking for.

```
def first_gen(df):
    '''find the names of all Pokemon from generation 1'''

def highest_hp(df):
    '''find the name(s) of the Pokemon with the highest HP'''

def mean_attack_by_type(df):
    '''find the mean attack power of each type (just use Type 1), result will
    contain just type and attack columns
    '''
```

```
def high_defense(df):
    '''find just the Name and Defense rating of Pokemon that have an
    above average (>) Defense.
    '''

def deduplicated(df):
    '''some Pokemon are in the list multiple times with different names.
    Find the names of the pokemon but without these duplicates. You
    can use the # column to tell if they're the same pokemon. Example:
    "Venusaur" has # 3 and there is a "VenusaurMega Venusaur" also
    with # 3, the result should have just "Venusaur"
    '''
```

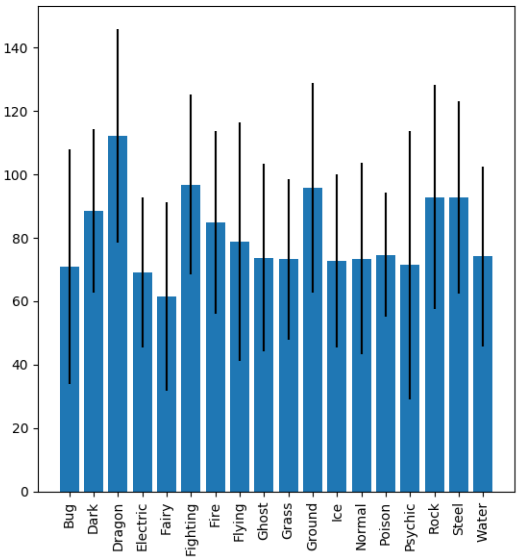
Task 3: Matplotlib

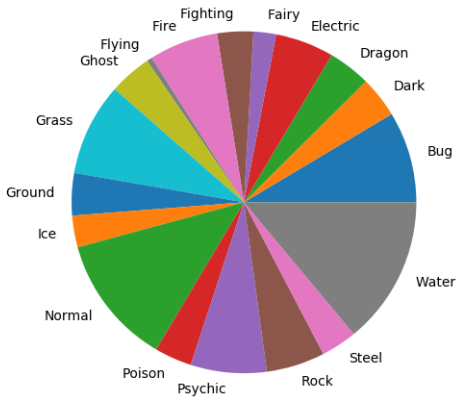
NOTE: Since this exercise is not graded, you do not have to put it on the github repository. If you do, please put all optional exercises somewhere into a different folder, so they do not clash with our evaluation!

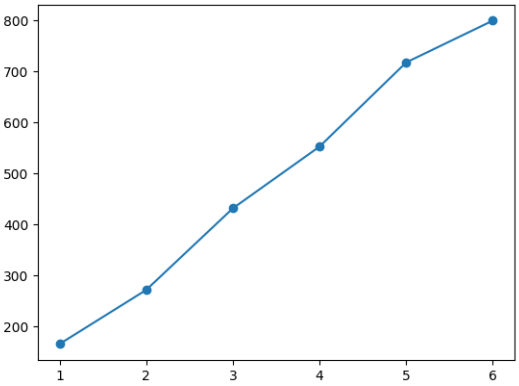
To get comfortable with matplotlib you'll make a few simple graphs based off of the Pokemon data. We've included examples of what they'll roughly look like. You should take some time to play around with the options that matplotlib provides to make them look nicer.

While most of our matplotlib tasks aren't graded since they'd be very difficult to evaluate automatically, we will have one task later on that we look at manually which will involve matplotlib so you should take the time now to familiarize yourself with it some if you haven't already.

1. Make a bar graph based off of the mean attack by type data frame. Add some error bars for the standard deviation too.
2. Make a pie chart of the number of Pokemon of each type (again, just Type 1).
3. Make a line graph to show how the number of Pokemon in total has increased across the generations.


140
120
100
80
60
40
20
0
Bug
Dark
Dragon
Electric
Fairy
Fighting
Fire
Flying
Ghost
Grass
Ground
Ice
Normal
Poison
Psychic
Rock
Steel
Water


Fighting
Fire
Fairy
Electric
Flying
Ghost
Dragon
Dark
Grass
Bug
Ground
Ice
Water
Normal
Steel
Poison
Rock
Psychic


800
700
600
500
400
300
200
1
2
3
4
5
6