ANA-522-OL1 Spring 2022
Mod02 Week04 Lab: pandas
Due: Friday February 4th at midnight

Explore Pokemon Dataset

The complete Pokemon Dataset is available on Kaggle.com

This dataset contains information on all 802 Pokemon from all Seven Generations of Pokemon. The information contained in this dataset include Base Stats, Performance against Other Types, Height, Weight, Classification, Egg Steps, Experience Points, Abilities, etc. The information was scraped from http://serebii.net/

We are going to use pandas and some NumPy functions to explore the dataset and extract information by manipulating data points from featured attributes.

[ ]:
```
import pandas as pd
import numpy as np
```

[ ]:
```
# pandas has the read_csv function to get the pokemon dataset
# More details can be found in Chapter 6 of the text by Wes McKinney.
#
# You can always access to the dataset file using the absolute path:
# '/home/ANA522/mod02/pokemon.csv'
# on the ANA-522 JupyterLab server
#
# If you are running your Notebook locally on your own computer (not using␣
 ↪ANA-522 JupyterLab server),
# you need to have a local copy of pokemon.cvs to be in the same folder with␣
 ↪your Notebook file, and do the following statement instead:
# pokemon = pd.read_csv('pokemon.csv', sep = ',')

pokemon = pd.read_csv('/home/ANA522/mod02/pokemon.csv', sep = ',')
```

1 Q01: Find a pandas function to overview the Pokemon dataset.

	against_bug	against_dark	against_dragon	against_electric	against_fairy	against_fight	against_fire	against_flying	against_ghost	against_grass	...	height_m	hp	percentage_
count	801.000000	801.000000	801.000000	801.000000	801.000000	801.000000	801.000000	801.000000	801.000000	801.000000	...	781.000000	801.000000	703.00
mean	0.996255	1.057116	0.968789	1.073970	1.068976	1.065543	1.135456	1.192884	0.985019	1.034020	...	1.163892	68.958801	55.1
std	0.597248	0.438142	0.353058	0.654962	0.522167	0.717251	0.691853	0.604488	0.558256	0.788896	...	1.080326	26.576015	20.2
min	0.250000	0.250000	0.000000	0.000000	0.250000	0.000000	0.250000	0.250000	0.000000	0.250000	...	0.100000	1.000000	0.00
25%	0.500000	1.000000	1.000000	0.500000	1.000000	0.500000	0.500000	1.000000	1.000000	0.500000	...	0.600000	50.000000	50.00
50%	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	...	1.000000	65.000000	50.00
75%	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	2.000000	1.000000	1.000000	1.000000	...	1.500000	80.000000	50.00
max	4.000000	4.000000	2.000000	4.000000	4.000000	4.000000	4.000000	4.000000	4.000000	4.000000	...	14.500000	255.000000	100.00

8 rows × 34 columns

2 Q02: Find a pandas attribute to return the dimension of the Pokemon dataset.

(801, 41)

3 Q03: Find a pandas attribute to return all attribute names of the Pokemon Dataset.

```
Index(['abilities', 'against_bug', 'against_dark', 'against_dragon',
       'against_electric', 'against_fairy', 'against_fight', 'against_fire',
       'against_flying', 'against_ghost', 'against_grass', 'against_ground',
       'against_ice', 'against_normal', 'against_poison', 'against_psychic',
       'against_rock', 'against_steel', 'against_water', 'attack',
       'base_egg_steps', 'base_happiness', 'base_total', 'capture_rate',
       'classfication', 'defense', 'experience_growth', 'height_m', 'hp',
       'japanese_name', 'name', 'percentage_male', 'pokedex_number',
       'sp_attack', 'sp_defense', 'speed', 'type1', 'type2', 'weight_kg',
       'generation', 'is_legendary'],
      dtype='object')
```

4 Q04: Find a pandas attribute to retrieve the data type of each column in Pokemon Dataset.

```
abilities            object
against_bug         float64
against_dark        float64
against_dragon      float64
against_electric    float64
against_fairy       float64
against_fight       float64
against_fire        float64
against_flying      float64
against_ghost       float64
against_grass       float64
against_ground      float64
against_ice         float64
against_normal      float64
against_poison      float64
against_psychic     float64
against_rock        float64
against_steel       float64
against_water       float64
attack                int64
base_egg_steps        int64
base_happiness        int64
base_total            int64
capture_rate         object
classfication        object
defense               int64
experience_growth     int64
height_m            float64
hp                    int64
japanese_name        object
name                 object
percentage_male     float64
pokedex_number        int64
sp_attack             int64
sp_defense            int64
speed                 int64
type1                object
type2                object
weight_kg           float64
generation            int64
is_legendary          int64
dtype: object
```

5 Q05: Display the whole Pokemon Dataset.

	abilities	against_bug	against_dark	against_dragon	against_electric	against_fairy	against_fight	against_fire	against_flying	against_ghost	...	percentage_male	pokedex_number	sp_attack	sp_defense	speed	type1	type2	weight_kg	generation
0	['Overgrow', 'Chlorophyll']	1.00	1.0	1.0	0.5	0.5	0.5	2.0	2.0	1.0	...	88.1	1	65	65	45	grass	poison	6.9	1
1	['Overgrow', 'Chlorophyll']	1.00	1.0	1.0	0.5	0.5	0.5	2.0	2.0	1.0	...	88.1	2	80	80	60	grass	poison	13.0	1
2	['Overgrow', 'Chlorophyll']	1.00	1.0	1.0	0.5	0.5	0.5	2.0	2.0	1.0	...	88.1	3	122	120	80	grass	poison	100.0	1
3	['Blaze', 'Solar Power']	0.50	1.0	1.0	1.0	0.5	1.0	0.5	1.0	1.0	...	88.1	4	60	50	65	fire	NaN	8.5	1
4	['Blaze', 'Solar Power']	0.50	1.0	1.0	1.0	0.5	1.0	0.5	1.0	1.0	...	88.1	5	80	65	80	fire	NaN	19.0	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
796	['Beast Boost']	0.25	1.0	0.5	2.0	0.5	1.0	2.0	0.5	1.0	...	NaN	797	107	101	61	steel	flying	999.9	7
797	['Beast Boost']	1.00	1.0	0.5	0.5	0.5	2.0	4.0	1.0	1.0	...	NaN	798	59	31	109	grass	steel	0.1	7
798	['Beast Boost']	2.00	0.5	2.0	0.5	4.0	2.0	0.5	1.0	0.5	...	NaN	799	97	53	43	dark	dragon	888.0	7
799	['Prism Armor']	2.00	2.0	1.0	1.0	1.0	0.5	1.0	1.0	2.0	...	NaN	800	127	89	79	psychic	NaN	230.0	7
800	['Soul-Heart']	0.25	0.5	0.0	1.0	0.5	1.0	2.0	0.5	1.0	...	NaN	801	130	115	65	steel	fairy	80.5	7

801 rows × 41 columns

6 Q06: Find a pandas attribute to return the Pokemon Dataset as a NumPy array.

```
array([["['Overgrow', 'Chlorophyll']", 1.0, 1.0, ..., 6.9, 1, 0],
       ["['Overgrow', 'Chlorophyll']", 1.0, 1.0, ..., 13.0, 1, 0],
       ["['Overgrow', 'Chlorophyll']", 1.0, 1.0, ..., 100.0, 1, 0],
       ...,
       ["['Beast Boost']", 2.0, 0.5, ..., 888.0, 7, 1],
       ["['Prism Armor']", 2.0, 2.0, ..., 230.0, 7, 1],
       ["['Soul-Heart']", 0.25, 0.5, ..., 80.5, 7, 1]], dtype=object)
```

7 Q07: Display the first 10 records in the Pokemon Dataset.

	abilities	against_bug	against_dark	against_dragon	against_electric	against_fairy	against_fight	against_fire	against_flying	against_ghost	...	percentage_male	pokedex_number	sp_attack	sp_defense	speed	type1	type2	weight_kg
0	['Overgrow', 'Chlorophyll']	1.00	1.0	1.0	0.5	0.5	0.5	2.0	2.0	1.0	...	88.1	1	65	65	45	grass	poison	6.9
1	['Overgrow', 'Chlorophyll']	1.00	1.0	1.0	0.5	0.5	0.5	2.0	2.0	1.0	...	88.1	2	80	80	60	grass	poison	13.0
2	['Overgrow', 'Chlorophyll']	1.00	1.0	1.0	0.5	0.5	0.5	2.0	2.0	1.0	...	88.1	3	122	120	80	grass	poison	100.0
3	['Blaze', 'Solar Power']	0.50	1.0	1.0	1.0	0.5	1.0	0.5	1.0	1.0	...	88.1	4	60	50	65	fire	NaN	8.5
4	['Blaze', 'Solar Power']	0.50	1.0	1.0	1.0	0.5	1.0	0.5	1.0	1.0	...	88.1	5	80	65	80	fire	NaN	19.0
5	['Blaze', 'Solar Power']	0.25	1.0	1.0	2.0	0.5	0.5	0.5	1.0	1.0	...	88.1	6	159	115	100	fire	flying	90.5
6	['Torrent', 'Rain Dish']	1.00	1.0	1.0	2.0	1.0	1.0	0.5	1.0	1.0	...	88.1	7	50	64	43	water	NaN	9.0
7	['Torrent', 'Rain Dish']	1.00	1.0	1.0	2.0	1.0	1.0	0.5	1.0	1.0	...	88.1	8	65	80	58	water	NaN	22.5
8	['Torrent', 'Rain Dish']	1.00	1.0	1.0	2.0	1.0	1.0	0.5	1.0	1.0	...	88.1	9	135	115	78	water	NaN	85.5
9	['Shield Dust', 'Run Away']	1.00	1.0	1.0	1.0	1.0	0.5	2.0	2.0	1.0	...	50.0	10	20	20	45	bug	NaN	2.9

10 rows × 41 columns

8 Q08: Display all records in the Pokemon Dataset with name, type1, and type2 attributes only.

	name	type1	type2
0	Bulbasaur	grass	poison
1	Ivysaur	grass	poison
2	Venusaur	grass	poison
3	Charmander	fire	NaN
4	Charmeleon	fire	NaN
...	...	...	...
796	Celesteela	steel	flying
797	Kartana	grass	steel
798	Guzzlord	dark	dragon
799	Necrozma	psychic	NaN
800	Magearna	steel	fairy

801 rows × 3 columns

9 Q09: Display all Pokemon whose height is greater than 9 meters with only the name and height attributes.

	name	height_m
207	Steelix	9.2
320	Wailord	14.5
796	Celesteela	9.2

10 Q10: Display the index number 207 Pokemon with its name, type1, and type2 attributes.

```
name     Steelix
type1      steel
type2     ground
Name: 207, dtype: object
```

11 Q11: Display the index number 207 Pokemon with attributes of column indices 1,2,3,4, and 0 in order using iloc()

```
against_bug                                                  0.5
against_dark                                                 1.0
against_dragon                                               0.5
against_electric                                             0.0
abilities           ['Rock Head', 'Sturdy', 'Sheer Force']
Name: 207, dtype: object
```

12 Q12: Display index number 207 and 208 Pokemon with attributes of column index 1,2,3,4, and 0 in order using iloc()

	against_bug	against_dark	against_dragon	against_electric	abilities
207	0.5	1.0	0.5	0.0	['Rock Head', 'Sturdy', 'Sheer Force']
208	0.5	0.5	0.0	1.0	['Intimidate', 'Run Away', 'Rattled']

13 Q13: Display all of the index 207 and 208 Pokemon whose weight_kg is greater than or equal to 400kg with attributes of column index 1,2,3,4, and 0 in order using iloc()

	against_bug	against_dark	against_dragon	against_electric	abilities
207	0.5	1.0	0.5	0.0	['Rock Head', 'Sturdy', 'Sheer Force']

14 Q14: Display all records of the Pokemon Dataset sorted by attribute speed.

	abilities	against_bug	against_dark	against_dragon	against_electric	against_fairy	against_fight	against_fire	against_flying	against_ghost	...	percentage_male	pokedex_number	sp
212	['Sturdy', 'Gluttony', 'Contrary']	1.0	1.0	1.0	1.0	1.0	1.00	1.0	1.0	1.0	...	50.0	213	
445	['Pickup', 'Thick Fat', 'Gluttony']	1.0	1.0	1.0	1.0	1.0	2.00	1.0	1.0	0.0	...	88.1	446	
770	['Innards Out', 'Unaware']	1.0	1.0	1.0	2.0	1.0	1.00	0.5	1.0	1.0	...	50.0	771	
596	['Iron Barbs']	1.0	1.0	0.5	0.5	0.5	2.00	4.0	1.0	1.0	...	50.0	597	
437	['Sturdy', 'Rock Head', 'Rattled']	1.0	1.0	1.0	1.0	1.0	2.00	0.5	0.5	1.0	...	50.0	438	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
141	['Rock Head', 'Pressure', 'Unnerve']	0.5	1.0	1.0	2.0	1.0	1.00	0.5	0.5	1.0	...	88.1	142	
100	['Soundproof', 'Static', 'Aftermath']	1.0	1.0	1.0	0.5	1.0	1.00	1.0	0.5	1.0	...	NaN	101	
794	['Beast Boost']	0.5	0.5	1.0	1.0	2.0	0.50	2.0	4.0	1.0	...	NaN	795	
290	['Speed Boost', 'Infiltrator']	0.5	1.0	1.0	2.0	1.0	0.25	2.0	2.0	1.0	...	50.0	291	
385	['Pressure']	2.0	2.0	1.0	1.0	1.0	0.50	1.0	1.0	2.0	...	NaN	386	

801 rows × 41 columns

15 Q15: Display all records of the Pokemon dataset sorted by attribute generation then by speed.

	abilities	against_bug	against_dark	against_dragon	against_electric	against_fairy	against_fight	against_fire	against_flying	against_ghost	...	percentage_male	pokedex_number	sp_attack	sp_defense	sp
78	['Oblivious', 'Own Tempo', 'Regenerator']	2.00	2.0	1.0	2.0	1.00	0.50	0.5	1.0	2.0	...	50.0	79	40	40	
38	['Cute Charm', 'Competitive', 'Friend Guard']	0.50	0.5	0.0	1.0	1.00	1.00	1.0	1.0	0.0	...	24.6	39	45	25	
73	['Rock Head', 'Sturdy', 'Sand Veil', 'Magnet P...	1.00	1.0	1.0	0.0	1.00	2.00	0.5	0.5	1.0	...	50.0	74	30	30	
45	['Effect Spore', 'Dry Skin', 'Damp']	2.00	1.0	1.0	0.5	1.00	0.50	4.0	4.0	1.0	...	50.0	46	45	55	
87	['Stench', 'Sticky Hold', 'Poison Touch', 'Poi...	0.50	1.0	1.0	1.0	0.50	0.50	1.0	1.0	1.0	...	50.0	88	40	50	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
757	['Corrosion', 'Oblivious']	0.25	1.0	1.0	1.0	0.25	0.50	0.5	1.0	1.0	...	0.0	758	111	60	
773	['Shields Down']	0.50	1.0	1.0	2.0	1.00	1.00	0.5	0.5	1.0	...	NaN	774	100	60	
742	['Honey Gather', 'Shield Dust', 'Sweet Veil']	0.50	0.5	0.0	1.0	1.00	0.25	2.0	2.0	1.0	...	50.0	743	95	70	
784	['Electric Surge', 'Telepathy']	0.50	0.5	0.0	0.5	1.00	0.50	1.0	0.5	1.0	...	NaN	785	95	75	
794	['Beast Boost']	0.50	0.5	1.0	1.0	2.00	0.50	2.0	4.0	1.0	...	NaN	795	137	37	

801 rows × 41 columns

16 Q16: How many generations are there in the Pokemon Dataset? What are they?

```
There are 7 in total. They are: 1 2 3 4 5 6 7 generations
```

17 Q17: How many Pokemon are there for each generation?

```
There are 7 generations in total
1    151
2    100
3    135
4    107
5    156
6     72
7     80
Name: generation, dtype: int64
```

18 Q18: List all speed values of Pokemon in asscending order.

```
There are 113 in total. They are:
5 10 15 20 22 23 24 25 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
46 47 48 49 50 51 52 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73
74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 95 96 97 98 99 100 101
102 103 104 105 106 108 109 110 112 113 114 115 116 117 118 120 121 123 124 125 126
127 128 130 132 135 140 145 150 151 160 180
```

19 Q19: What is the minimum speed of all Pokemon in the Dataset?

5

20 Q20: Display all Pokemon whose speed is minimum speed with all attributes.

	abilities	against_bug	against_dark	against_dragon	against_electric	against_fairy	against_fight	against_fire	against_flying	against_ghost	...	percentage_male	pokedex_number	sp_attack	sp_defense	speed	type1	type2	weight_kg	generation	is_legendary
212	['Sturdy', 'Gluttony', 'Contrary']	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	...	50.0	213	10	230	5	bug	rock	20.5	2	0
445	['Pickup', 'Thick Fat', 'Gluttony']	1.0	1.0	1.0	1.0	1.0	2.0	1.0	1.0	0.0	...	88.1	446	40	85	5	normal	NaN	105.0	4	0
770	['Innards Out', 'Unaware']	1.0	1.0	1.0	2.0	1.0	1.0	0.5	1.0	1.0	...	50.0	771	30	130	5	water	NaN	1.2	7	0

3 rows × 41 columns

21 Q21: Convert all Pokemon attak numbers into floats. You can create a lambda function, and apply the function to convert all attack values to the float data type. Print the converted attack column as output.

```
0       49.0
1       62.0
2      100.0
3       52.0
4       64.0
       ...
796    101.0
797    181.0
798    101.0
799    107.0
800     95.0
Name: attack, Length: 801, dtype: float64
```

22 Q22: What is the average and total weight in kg of all the Pokemon's combined. Present the average and the total value in two decimal points

```
The average weight of Pokemon is 61.38
Total wieight is 47936.30 kgs.
```

23 Q23: How many Pokemon whose type1 is grass?

```
There are 78 Pokemon whose type1 is grass.
```

24 Q24: How many Pokemon have a NaN value in their type2 attribute?

```
There are 384 Pokemon whose type2 value is NaN.
```

25 Q25: What are the top10 numbers of Pokemon in type1?

```
water       114
normal      105
grass        78
bug          72
psychic      53
fire         52
rock         45
electric     39
ground       32
poison       32
Name: type1, dtype: int64
```