Gotta catch 'em all !

The dataset which I'll use includes information about Pokemons.

An interesting set of questions answered using python to get a basic understanding of pandas and data visualization libraries.

1)importing all important libraries

For eg, "import numpy as np"

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
import warnings
warnings.filterwarnings("ignore")
```

2) Read the csv file and assign it to a variable .

```
#your code here
data = pd.read_excel("Pokemon.xlsx")
data
```

	#	Name	Type 1	Type 2	Total	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	Generation	Legendary
0	1	Bulbasaur	Grass	Poison	318	45	49	49	65	65	45	1	False
1	2	Ivysaur	Grass	Poison	405	60	62	63	80	80	60	1	False
2	3	Venusaur	Grass	Poison	525	80	82	83	100	100	80	1	False
3	3	VenusaurMega Venusaur	Grass	Poison	625	80	100	123	122	120	80	1	False
4	4	Charmander	Fire	NaN	309	39	52	43	60	50	65	1	False
...	...	...	...	...	...	...	...	...	...	...	...	...	...
795	719	Diancie	Rock	Fairy	600	50	100	150	100	150	50	6	True
796	719	DiancieMega Diancie	Rock	Fairy	700	50	160	110	160	110	110	6	True
797	720	HoopaHoopa Confined	Psychic	Ghost	600	80	110	60	150	130	70	6	True

3) Display shape of dataframe

Expected Output - (800, 13)

```
#your code here
data.shape
```

```
(800, 13)
```

4) Print all columns of dataframe

Return an array containing names of all the columns.

```
#your code here
data.columns
```

```
Index(['#', 'Name', 'Type 1', 'Type 2', 'Total', 'HP', 'Attack', 'Defense',
       'Sp. Atk', 'Sp. Def', 'Speed', 'Generation', 'Legendary'],
      dtype='object')
```

5) Remove the column '#' and update the dataframe.

```
#your code here
data=data.drop(['#'],axis=1)
```

6) Set the 'Name' column as the index of dataframe

```
data = data.set_index('Name')
data.head()
```

Name	Type 1	Type 2	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	Total	Generation	Legendary
Bulbasaur	Grass	Poison	45	49	49	65	65	45	318	1	False
Ivysaur	Grass	Poison	60	62	63	80	80	60	405	1	False
Venusaur	Grass	Poison	80	82	83	100	100	80	525	1	False
VenusaurMega Venusaur	Grass	Poison	80	100	123	122	120	80	625	1	False
Charmander	Fire	NaN	39	52	43	60	50	65	309	1	False

7) Print a list of all the unique Type-1 powers

```
#your code here
print(list(data['Type 1'].unique()))
```

```
['Grass', 'Fire', 'Water', 'Bug', 'Normal', 'Poison', 'Electric', 'Ground', 'Fairy', 'Fighting', 'Psychic', 'Rock', 'Ghost', 'Ice', 'Dragon', 'Dark', 'Steel', 'Flying']
```

8) Create a column which contains the Type 1 and Type 2 abilities of pokemons, seperated with a '+" sign. Also, display the no. of pokemons that have type-1 power as 'Psychic' and type 2 power as 'Flying' using this new column.

```
#Create a column with name 'Type 1 + 2', which contains the Type 1 and Type 2 abilities of pokemons, seperated with a '+'' sign
data['Type 1 + 2'] = data['Type 1'] + " + " + data['Type 2']

#The number of pokemons with Type 1 ability as 'Psychic' and Type 2 ability as 'Flying' are:
data['Type 1 + 2'].value_counts()['Psychic + Flying']
```

```
6
```

Questions about the dataset

1. How many pokemons have 'Mega' in their name?

```
(data.index.str.contains('Mega')).sum()
```

```
49
```

2. What is the standard deviation of Sp. Def. in the dataset ?

```
#your code here
np.round(data['Sp. Def'].describe()['std'],3)
```

```
27.829
```

3. What percentage (upto 3 decimal places) of pokemons are legendary ?

```
#your code here
(((data['Legendary']==True).sum())/(data.shape[0]))*100
```

```
8.125
```

4. Name the pokemon(s) with Maximum Defense.

```
#your code here
print("\nMax Defence:\n",data["Defense"].sort_values(ascending= False).head(5))
```

```
Max Defence:
 Name
Shuckle                230
SteelixMega Steelix    230
AggronMega Aggron      230
Steelix                200
Regirock               200
Name: Defense, dtype: int64
```

The top 3 pokemons with highest defense are SteelixMega Steelix, Shuckle, AggrnMega Aggron

5. Which poison pokemon has the strongest attack ?

```
#your code here
data.loc[((data['Type 1']== 'Poison') | (data['Type 2']  == 'Poison')),'Attack'].sort_values(ascending=False).head()
```

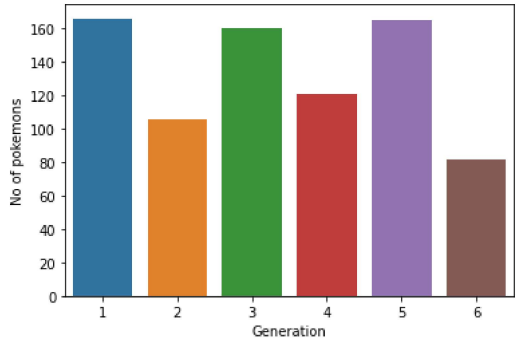
```
Name
BeedrillMega Beedrill    150
Toxicroak                106
Muk                      105
Victreebel               105
Nidoking                 102
Name: Attack, dtype: int64
```

BeadrillMega Beedrill is the poison pokemon with the strongest attack.

6. Using seaborn make different types of plots, observe the trend and answer the questions given in the form.

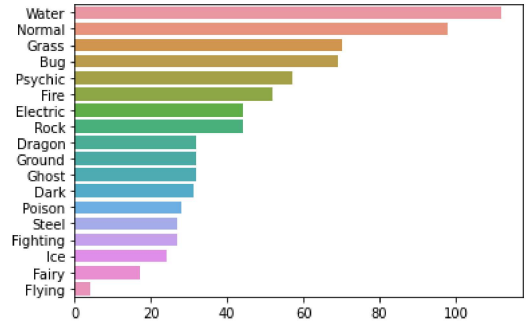
```
#your code here
a=data.Generation.value_counts()
ab=sns.barplot(a.index,a.values)
ab.set_xlabel("Generation")
ab.set_ylabel("No of pokemons")
```

```
Text(0, 0.5, 'No of pokemons')
```




```
#your code here
b= data["Type 1"].value_counts()
sns.barplot(b.values,b.index)
```

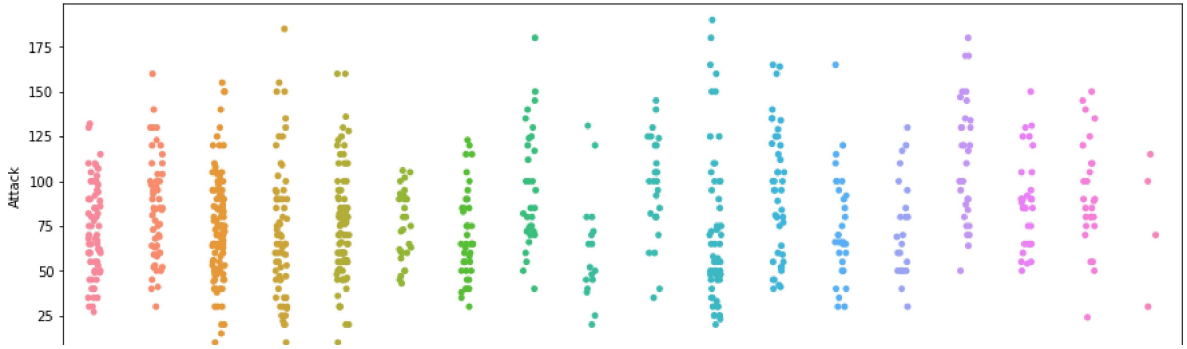
```
<matplotlib.axes._subplots.AxesSubplot at 0x7faacf39fc10>
```



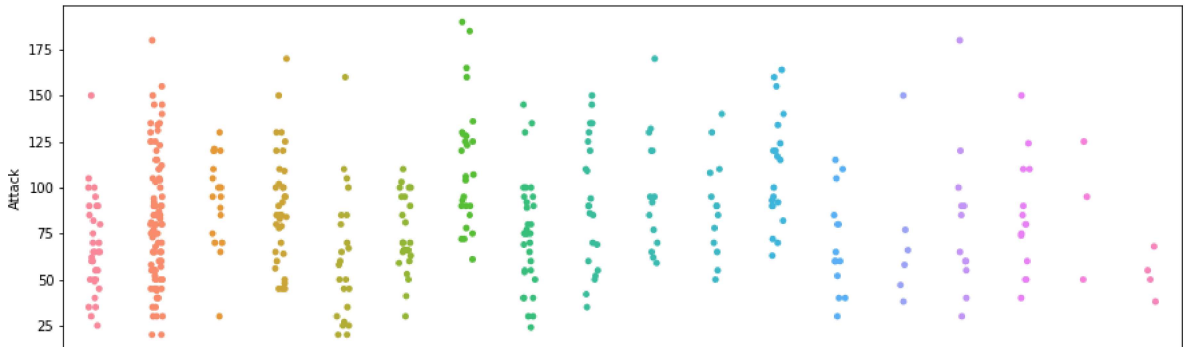

```
c= data["Type 2"].value_counts()
sns.barplot(c.values,c.index)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7faacf2ff610>
```


```
#your code here
fig, ax = plt.subplots(figsize=(15, 5))
ax = sns.stripplot(x='Type 1', y='Attack', data=data)
```

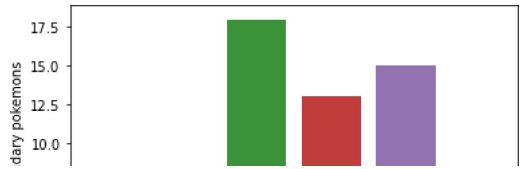



```
fig, ay = plt.subplots(figsize=(15, 5))
ay = sns.stripplot(x='Type 2', y='Attack', data=data)
```




```
#your code here
legendary = data[data['Legendary']==True]
x= legendary.Generation.value_counts()
ax= sns.barplot(x.index,x.values)
ax.set_xlabel("Generation")
ax.set_ylabel("No of Legendary pokemons")
```

```
Text(0, 0.5, 'No of Legendary pokemons')
```




```
data[(data['Type 1']=='Psychic')&(data['Type 2']=='Flying')]
```

Name	Type 1	Type 2	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	Total	Generation	Legendary	Type 1 + 2
Natu	Psychic	Flying	40	50	45	70	45	70	320	2	False	Psychic + Flying
Xatu	Psychic	Flying	65	75	70	95	70	95	470	2	False	Psychic + Flying
Lugia	Psychic	Flying	106	90	130	90	154	110	680	2	True	Psychic + Flying
Woobat	Psychic	Flying	55	45	43	55	43	72	313	5	False	Psychic + Flying
Swoobat	Psychic	Flying	67	57	55	77	55	114	425	5	False	Psychic + Flying
Sigilyph	Psychic	Flying	72	58	80	103	80	97	490	5	False	Psychic + Flying

7. Which is the second fastest non-legendary 'Ghost' type pokemon from 4th generation ?

```
data.loc[(((data['Type 1']== 'Ghost') | (data['Type 2']  == 'Ghost')) & (data['Legendary']==False) & (data['Generation']==4)),'Speed'].sort_values(ascending=False).head()
```

```
Name
Froslass     110
Mismagius    105
Rotom         91
Drifblim      80
Drifloon      70
Name: Speed, dtype: int64
```

8. How many non-legendary pokemons have stronger defence but weaker attack than Charizard?

```
#your code here
data[data.index == 'Charizard']
data[(data.Attack<84)&(data.Defense>78)&(data.Legendary==False)].shape[0]
```

```
116
```

9. Which pokemon has the highest combined value of Attack and Sp. Atk ?

```
#your code here
data['Combined'] = data['Attack'] + data['Sp. Atk']
data.Combined.sort_values(ascending=False).head()
```

```
Name
DeoxysAttack Forme       360
RayquazaMega Rayquaza    360
MewtwoMega Mewtwo X      344
MewtwoMega Mewtwo Y      344
KyogrePrimal Kyogre      330
Name: Combined, dtype: int64
```

RayquazaMega Rayquaza and DeoxysAttack Forme are the pokemons with highest combined value of Attacka and Sp. Atk

10. Which type of legendary pokemons are the most common?

```
#your code here
print(legendary['Type 1'].value_counts().head() , legendary['Type 2'].value_counts().head())
```

```
Psychic     14
Dragon      12
Fire         5
Electric     4
Water        4
Name: Type 1, dtype: int64 Flying      13
Psychic      5
Fighting     4
Dragon       4
Fire         3
Name: Type 2, dtype: int64
```

11. How many generation-3, non-legendary pokemons have higher HP than the weakest generation-6, legendary pokemon?

```
#your code here
data[(data.Legendary==True)&(data.Generation ==6)].HP.min()
```

```
50
```

```
data[(data.Legendary==False)&(data.Generation==3)&(data.HP>50)].shape[0]
```

```
95
```

12. Print out the third slowest pokemon(s) in the dataset.

```
#your code here
data.Speed.sort_values().head(15)
```

```
Name
Munchlax       5
Shuckle        5
Ferroseed     10
Bonsly        10
Trapinch      10
Silcoon       15
Foongus       15
Wooper        15
Slowpoke      15
Cleffa        15
Igglybuff     15
Roggenrola    15
Cascoon       15
Pineco        15
Escavalier    20
Name: Speed, dtype: int64
```

Third slowest speed is 15 . There are 9 pokemonns with this speed, namely :

- Silcoon
- Foongus
- Wooper
- Slowpoke
- Cleffa
- Igglybuff
- Roggenrola
- Cascoon
- Pineco

13. Which pokemon type has the highest average HP?

```
#your code here
((data.groupby('Type 2').HP.mean())+(data.groupby('Type 1').HP.mean()))/2
```

```
Type 2
Bug         55.108696
Dark        71.178226
Dragon      82.739583
Electric    73.981061
Fairy       69.210997
```

```
Fighting    74.656695
Fire        70.576923
Flying      71.070876
Ghost       61.790179
Grass       64.955714
Ground      75.504911
Ice         81.000000
Normal      70.387755
Poison      63.007353
Psychic     71.421850
Rock        66.717532
Steel       64.929293
Water       67.388393
Name: HP, dtype: float64
```

'Dragon' type has the highest average HP

✓ 0s completed at 5:43 PM