ibimbing

Pokemon Data Analysis

PORTFOLIO

Fariz Rifky Berliano

Table of Contents

Tools

These are the tools used to analyze and visualize data.

Questions to Answer

From the data analysis and data visualization that has been done, we can answer the following questions:

1. How are pokemon ***generations*** distributed?

2. How are ***legendary*** and ***non-legendary*** pokemon distributed?

3. Who is the ***top 5 strongest*** pokemon by total stats?

Data Understanding

This section contains information about the dataset/dataframe that will be used for analysis and visualization, the data used is "pokemon.csv" obtained from:
https://bit.ly/data-pokemon-dsf

Pokemon DataFrame

	#	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
798	720	HoopaHoopa Unbound	Psychic	Dark	680	80	160	60	170	130	80	6	True
799	721	Volcanion	Fire	Water	600	80	110	120	130	90	70	6	True

Data Understanding

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 800 entries, 0 to 799
Data columns (total 13 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   #           800 non-null    int64
 1   Name        800 non-null    object
 2   Type 1      800 non-null    object
 3   Type 2      414 non-null    object
 4   Total       800 non-null    int64
 5   HP          800 non-null    int64
 6   Attack      800 non-null    int64
 7   Defense     800 non-null    int64
 8   Sp. Atk     800 non-null    int64
 9   Sp. Def     800 non-null    int64
 10  Speed       800 non-null    int64
 11  Generation  800 non-null    int64
 12  Legendary   800 non-null    object
dtypes: int64(9), object(4)
memory usage: 81.4+ KB
```

Summary of DataFrame Structure

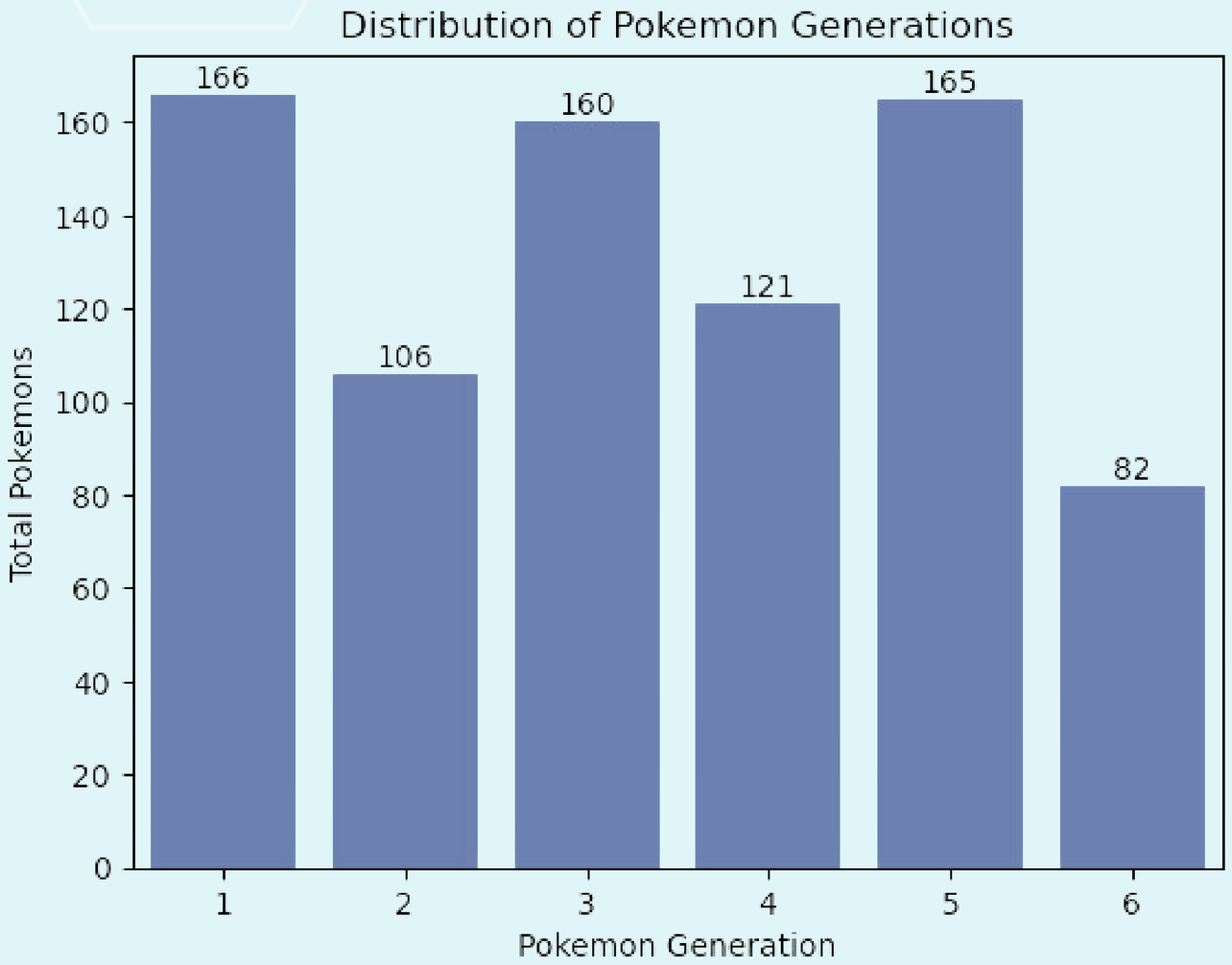
From summary on the side, it can be seen that there are 800 pokemon available(rows) in the data and has 13 columns with 2 data types, namely object/string and int64.
We also can see in the "Type 2" column there are some empty data because there are pokemon that only have 1 type.

Descriptive Statistics of DataFrame

	#	Total	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	Generation
count	800.000000	800.00000	800.000000	800.000000	800.000000	800.000000	800.000000	800.000000	800.00000
mean	362.813750	435.10250	69.258750	79.001250	73.842500	72.820000	71.902500	68.277500	3.32375
std	208.343798	119.96304	25.534669	32.457366	31.183501	32.722294	27.828916	29.060474	1.66129
min	1.000000	180.00000	1.000000	5.000000	5.000000	10.000000	20.000000	5.000000	1.00000
25%	184.750000	330.00000	50.000000	55.000000	50.000000	49.750000	50.000000	45.000000	2.00000
50%	364.500000	450.00000	65.000000	75.000000	70.000000	65.000000	70.000000	65.000000	3.00000
75%	539.250000	515.00000	80.000000	100.000000	90.000000	95.000000	90.000000	90.000000	5.00000
max	721.000000	780.00000	255.000000	190.000000	230.000000	194.000000	230.000000	180.000000	6.00000

The image above is the descriptive statistics of pokemon dataframe.
These descriptive statistics include a summary of key values such as count (number of non-empty rows), average (mean), standard deviation, minimum value, quartiles, and maximum value of the overall pokemon stats.

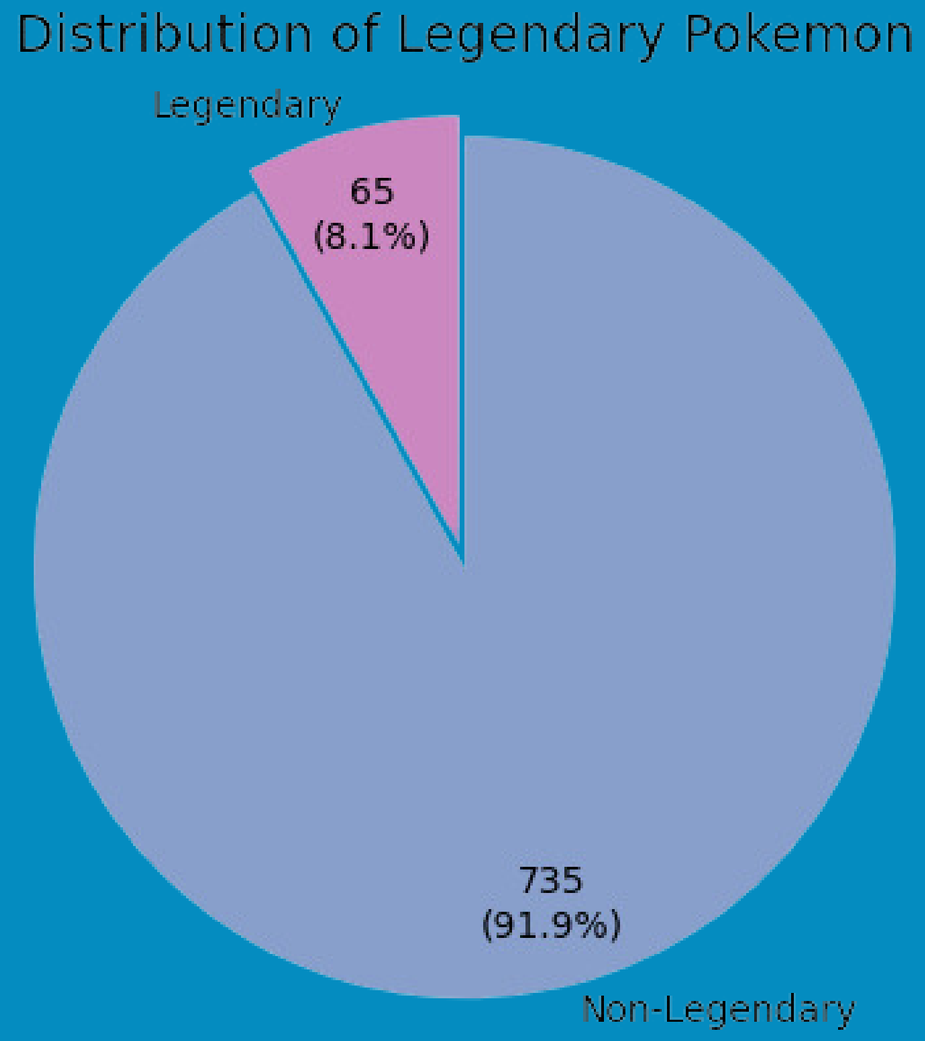
Answer of Questions



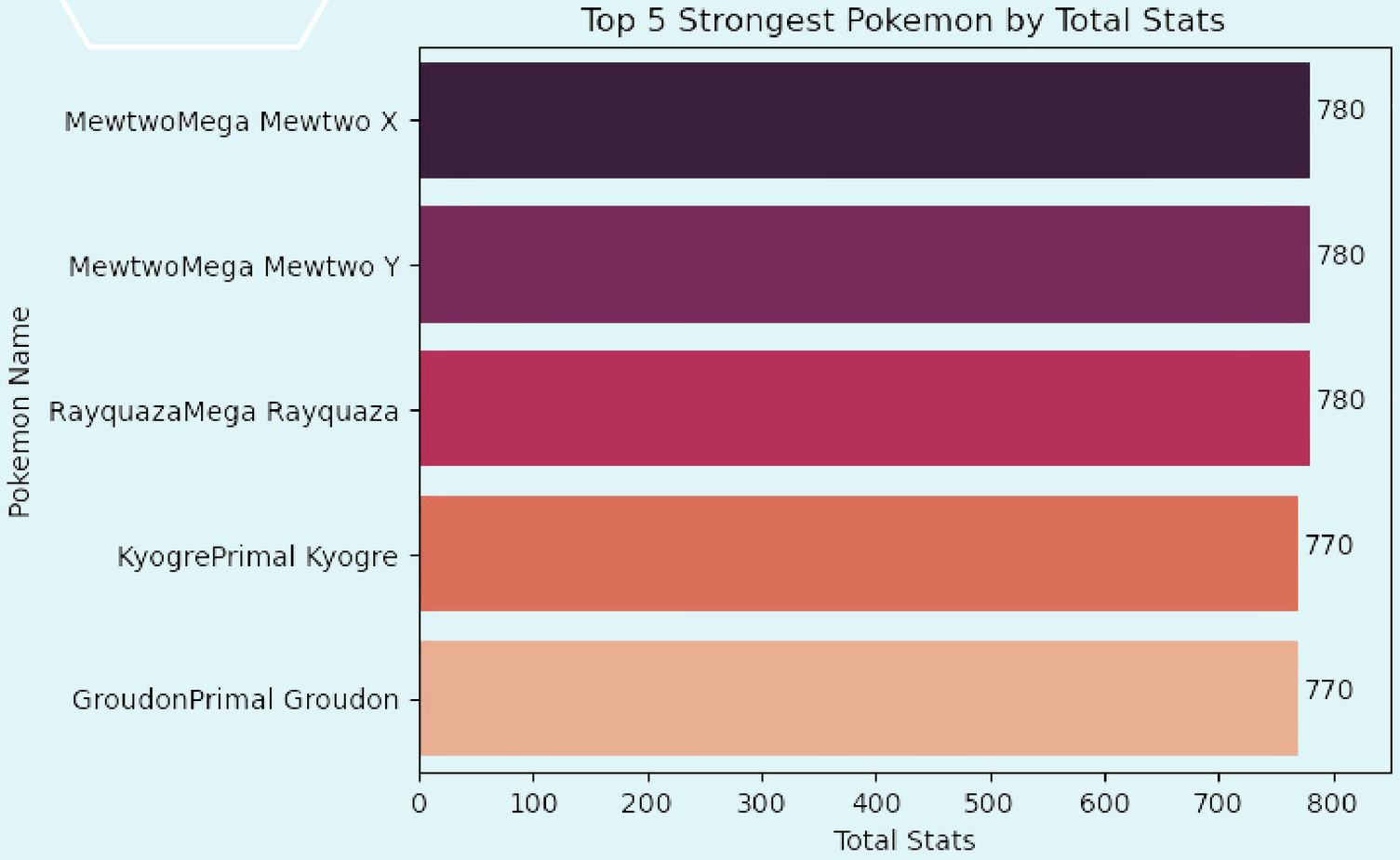
1. How are pokemon ***generations*** distributed?
=> From the graph bar on the side, we can conclude that ***generation 1*** pokemon has the most total pokemons (166 pokemons), with a difference of only 1 pokemon compared to ***generation 5*** (165 pokemons).

Answer of Questions

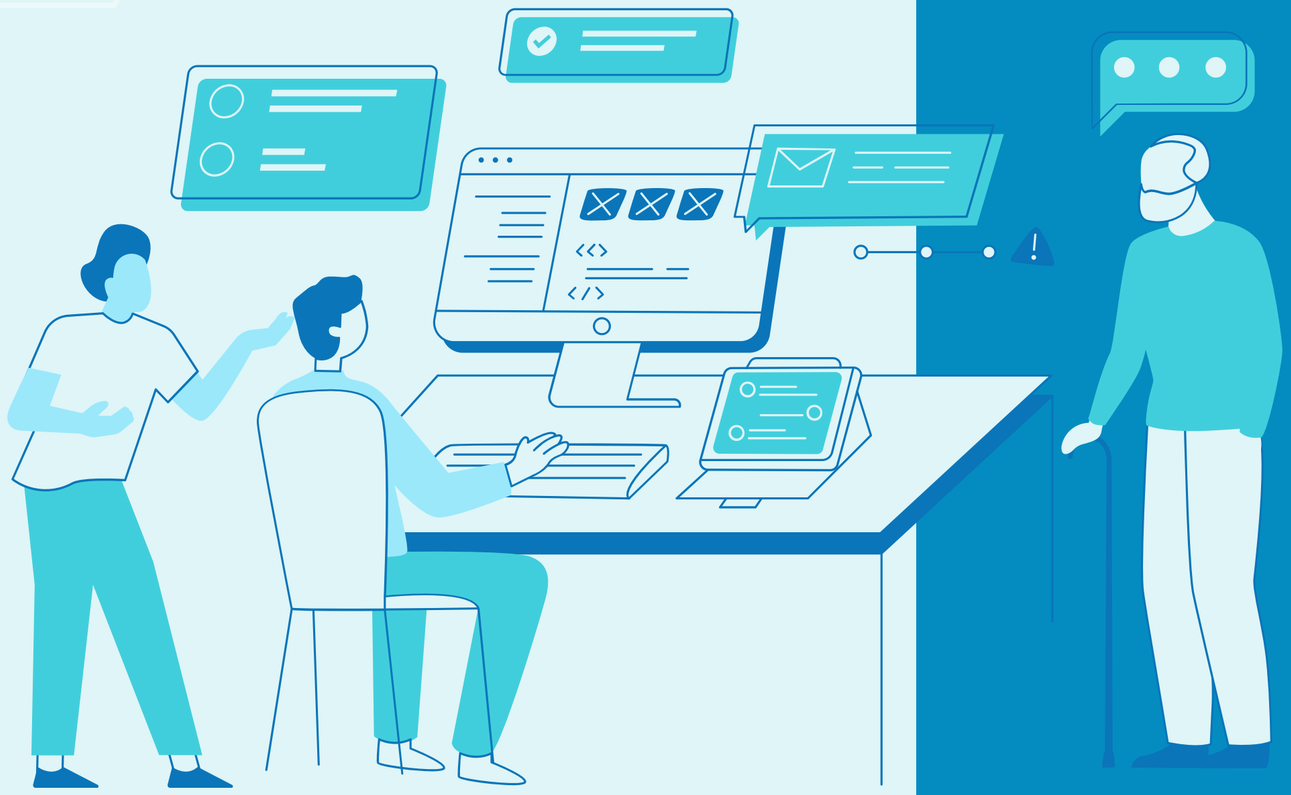
2. How are ***legendary*** and ***non-legendary*** pokemon distributed?
=> From the following pie chart, the distribution of legendary pokemon and non-legendary is **very far**.
We can conclude that legendary pokemon are **very rare** and **difficult to find** in the game/movie.



Answer of Questions



3. Who is the ***top 5 strongest*** pokemon by total stats?
=> We can see that the 5 strongest pokemon based on their total stats are as follows (graph on the side). The differences in their total stats are **very close** and **some are even the same**.



THANK YOU

farizrifkyberliano@gmail.com

https://github.com/ifarbie

Let’s Connect on LinkedIn!