Pandas:

Pandas is a popular open-source data manipulation and analysis library for Python.

It provides easy-to-use data structures and functions needed to work with structured data,
such as spreadsheets or SQL tables, and time series data.

The primary data structures in pandas are:

DataFrame: A two-dimensional table of data with rows and columns. It can be thought of as a spreadsheet or SQL table. Each column can have a different data type (e.g., integers, strings, floating-point numbers).

Creating a Dataframe:

```
import pandas as pd
# Creating a DataFrame
data = {'Name': ['Alice', 'Bob', 'Charlie'],
        'Age': [25, 30, 35],
        'City': ['New York', 'San Francisco', 'Los Angeles']}
df = pd.DataFrame(data)

# Displaying the DataFrame
print(df)
   Name  Age          City
0  Alice   25       New York
1    Bob   30  San Francisco
2 Charlie  35    Los Angeles
```

Series: A one-dimensional labeled array capable of holding any data type. It can be thought of as a single column in a DataFrame.

```
import pandas as pd

# Create a series from a list
data_list = [1, 2, 3, 4, 5]
series_from_list = pd.Series(data_list)
print("Series from list:")
print(series_from_list)
```

Loading data:

```
import pandas as pd
df = pd.read_csv('pokemon_data.csv')
```

Reading Data using Pandas:

```
# print(df.head())

# print(df.head(10))

# print(df.tail())

# print(df.tail(15))
```

#Specific value

```
print(df.iloc[2,1])
```

```
print(df.columns)

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

Specific columns data:

```
print(df[['Name', 'Type 1', 'HP']])
```

```
## Read Each Row
#print(df.iloc[0:4])#4 records
```

Filtering the data:

```
#df.loc[df['Type 1'] == "Grass"]
```

Sorting Values:

Ascending:

```
df.sort_values(by=['Type 1'])
```

Multiple columns

```
df.sort_values(by=['Type 1', 'HP'])
```

Descending:

```
df.sort_values(by='Type 1', ascending=False)
```

Putting NAs at first

```
df.sort_values(by='Type 2', ascending=False, na_position='first')
```

Making changes to the data:

```
df['Total'] = df['HP'] + df['Attack'] + df['Defense'] + df['Sp. Atk'] + df['Sp. Def'] + df['Speed']
```

or

```
df['Total'] = df.iloc[:, 4:10].sum(axis=1)
```

df.head(5)

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

df.head()

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

Advanced Filtering data:

```
new_df = df.loc[(df['Type 1'] == 'Grass') & (df['Type 2'] == 'Poison') & (df['HP'] > 70)]
```

new_df

	Name	Type 1	Type 2	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	Generation	Legendary	Total
3	Venusaur	Grass	Poison	80	82	83	100	100	80	1	False	525
3	VenusaurMega Venusaur	Grass	Poison	80	100	123	122	120	80	1	False	625
45	Vileplume	Grass	Poison	75	80	85	110	90	50	1	False	490
71	Victreebel	Grass	Poison	80	105	65	100	70	70	1	False	490
591	Amoonguss	Grass	Poison	114	85	70	85	80	30	5	False	464

Saving/Exporting the Data:

```
df.to_csv('modified.csv', index=False)
```

Conditinal Changes:

```
df.loc[df['Total'] > 500, ['Generation','Legendary']] = ['Test 1', 'Test 2']
```

Aggregate Statistics (Groupby):

```
df.groupby(['Type 1']).count()['count']
```

```
df.groupby(['Type 1', 'Type 2']).count()['count']
```