Feature extraction of Anime data using Pandas

In this Project we are going to Perform the below tasks

1. make a new column for episode count
2. make a new column for time stamp
3. which anime has the highest score
4. give me top 5 highest scoring anime
5. which anime has the highest episode count
6. animes with top 5 episode count
7. which is the longest running anime
8. Top 5 longest running Animes

```
In [63]: ## Importing pandas & numpy libraries
         import numpy as np
         import pandas as pd
```

```
In [64]: ## Loading the data from CSV file
         df = pd.read_csv(r"C:\Users\balus\OneDrive\Desktop\Data Science prep\Python libraries- Shrey\Pandas\anime.csv")
```

```
In [65]: df
```

Out[65]:

	Rank	Title	Score
0	1	Fullmetal Alchemist: BrotherhoodTV (64 eps)Apr...	9.10
1	2	Steins;GateTV (24 eps)Apr 2011 - Sep 20112,473...	9.07
2	3	Bleach: Sennen Kessen-henTV (13 eps)Oct 2022 -...	9.06
3	4	Gintama°TV (51 eps)Apr 2015 - Mar 2016605,113 ...	9.06
4	5	Shingeki no Kyojin Season 3 Part 2TV (10 eps)A...	9.05
5	6	Gintama'TV (51 eps)Apr 2011 - Mar 2012534,105 ...	9.04
6	7	Gintama: The FinalMovie (1 eps)Jan 2021 - Jan ...	9.04
7	8	Hunter x Hunter TV (148 eps)Oct 2011 - Sep 201...	9.04
8	9	Kaguya-sama wa Kokurasetai: Ultra RomanticTV (...	9.04
9	10	Gintama': EnchousenTV (13 eps)Oct 2012 - Mar 2...	9.03
10	11	Shingeki no Kyojin: The Final Season - Kankets...	9.03
11	12	Ginga Eiyuu DensetsuOVA (110 eps)Jan 1988 - Ma...	9.02
12	13	Bleach: Sennen Kessen-hen - Ketsubetsu-tanTV (...	8.99
13	14	Fruits Basket: The FinalTV (13 eps)Apr 2021 - ...	8.99
14	15	Gintama.TV (12 eps)Jan 2017 - Mar 2017302,232 ...	8.98
15	16	GintamaTV (201 eps)Apr 2006 - Mar 20101,034,41...	8.94
16	17	Koe no KatachiMovie (1 eps)Sep 2016 - Sep 2016...	8.94
17	18	3-gatsu no Lion 2nd SeasonTV (22 eps)Oct 2017 ...	8.93
18	19	Clannad: After StoryTV (24 eps)Oct 2008 - Mar ...	8.93
19	20	Code Geass: Hangyaku no Lelouch R2TV (25 eps)A...	8.91
20	21	Gintama Movie 2: Kanketsu-hen - Yorozuya yo Ei...	8.91
21	22	Violet Evergarden MovieMovie (1 eps)Sep 2020 -...	8.89
22	23	Owarimonogatari 2nd SeasonTV (7 eps)Aug 2017 -...	8.88
23	24	Gintama.: Shirogane no Tamashii-hen - Kouhan-s...	8.88
24	25	MonsterTV (74 eps)Apr 2004 - Sep 20051,041,081...	8.87
25	26	Kimi no Na wa.Movie (1 eps)Aug 2016 - Aug 2016...	8.84
26	27	Kaguya-sama wa Kokurasetai: First Kiss wa Owar...	8.84
27	28	Bocchi the Rock!TV (12 eps)Oct 2022 - Dec 2022...	8.83
28	29	The First Slam DunkMovie (1 eps)Dec 2022 - Dec...	8.83
29	30	Vinland Saga Season 2TV (24 eps)Jan 2023 - Jun...	8.82
30	31	Jujutsu Kaisen 2nd SeasonTV (23 eps)Jul 2023 -...	8.82
31	32	Gintama.: Shirogane no Tamashii-henTV (12 eps)...	8.81
32	33	Kingdom 3rd SeasonTV (26 eps)Apr 2020 - Oct 20...	8.81

33	34	Mob Psycho 100 IITV (13 eps)Jan 2019 - Apr 201...	8.80
34	35	Shingeki no Kyojin: The Final SeasonTV (16 eps...	8.80
35	36	Kimetsu no Yaiba: Yuukaku-henTV (11 eps)Dec 20...	8.79
36	37	Kizumonogatari III: Reiketsu-henMovie (1 eps)J...	8.79
37	38	"Oshi no Ko"TV (11 eps)Apr 2023 - Jun 2023645,...	8.79
38	39	Haikyuu!! Karasuno Koukou vs. Shiratorizawa Ga...	8.78
39	40	Sen to Chihiro no KamikakushiMovie (1 eps)Jul ...	8.78
40	41	Monogatari Series: Second SeasonTV (26 eps)Jul...	8.77
41	42	Shingeki no Kyojin: The Final Season Part 2TV ...	8.77
42	43	Hajime no IppoTV (75 eps)Oct 2000 - Mar 200255...	8.76
43	44	Cowboy BebopTV (26 eps)Apr 1998 - Apr 19991,79...	8.75
44	45	Kingdom 4th SeasonTV (26 eps)Apr 2022 - Oct 20...	8.75
45	46	Vinland SagaTV (24 eps)Jul 2019 - Dec 20191,40...	8.74
46	47	Mushishi Zoku Shou 2nd SeasonTV (10 eps)Oct 20...	8.73
47	48	Shouwa Genroku Rakugo Shinjuu: Sukeroku Futata...	8.72
48	49	Mob Psycho 100 IIITV (12 eps)Oct 2022 - Dec 20...	8.71
49	50	Rurouni Kenshin: Meiji Kenkaku Romantan - Tsui...	8.71

1. make a new column for episode count

```
In [66]:   ## Function to extract episodes
           def extract_episodes(txt):
               check = False
               data = ""
               for i in txt:
                   if i == ")":
                       check = False
                       return data
                   if check == True:
                       data = data + i
                   if i == '(':
                       check = True
```

```
In [67]:   ## applying function on Title colmn
           df["Episodes"] = df["Title"].apply(extract_episodes)
```

```
In [68]:   df["Episodes"]
```

```
Out[68]:  0        64 eps
          1        24 eps
          2        13 eps
          3        51 eps
          4        10 eps
          5        51 eps
          6         1 eps
          7       148 eps
          8        13 eps
          9        13 eps
          10        2 eps
          11      110 eps
          12       13 eps
          13       13 eps
          14       12 eps
          15      201 eps
          16        1 eps
          17       22 eps
          18       24 eps
          19       25 eps
          20        1 eps
          21        1 eps
          22        7 eps
          23       14 eps
          24       74 eps
          25        1 eps
          26        1 eps
          27       12 eps
          28        1 eps
          29       24 eps
          30       23 eps
          31       12 eps
          32       26 eps
          33       13 eps
          34       16 eps
          35       11 eps
          36        1 eps
          37       11 eps
          38       10 eps
          39        1 eps
          40       26 eps
          41       12 eps
          42       75 eps
          43       26 eps
          44       26 eps
          45       24 eps
          46       10 eps
          47       12 eps
          48       12 eps
          49        4 eps
          Name: Episodes, dtype: object
```

```python
In [69]:  ## removing "eps" from Episodes column
          df['Episodes']= df['Episodes'].str.replace(" eps","")
```

```python
In [70]:  ## changing the column type to Int
          df['Episodes'] = df['Episodes'].astype(int)
```

```python
In [71]:  df["Episodes"]
```

```
Out[71]:  0      64
          1      24
          2      13
          3      51
          4      10
          5      51
          6       1
          7     148
          8      13
          9      13
          10      2
          11    110
          12     13
          13     13
          14     12
          15    201
          16      1
          17     22
          18     24
          19     25
          20      1
          21      1
          22      7
          23     14
          24     74
          25      1
          26      1
          27     12
          28      1
          29     24
          30     23
          31     12
          32     26
          33     13
          34     16
          35     11
          36      1
          37     11
          38     10
          39      1
          40     26
          41     12
          42     75
          43     26
          44     26
          45     24
          46     10
          47     12
          48     12
          49      4
          Name: Episodes, dtype: int64
```

In [72]: `df`

Out[72]:

	Rank	Title	Score	Episodes
0	1	Fullmetal Alchemist: BrotherhoodTV (64 eps)Apr...	9.10	64
1	2	Steins;GateTV (24 eps)Apr 2011 - Sep 20112,473...	9.07	24
2	3	Bleach: Sennen Kessen-henTV (13 eps)Oct 2022 -...	9.06	13
3	4	Gintama°TV (51 eps)Apr 2015 - Mar 2016605,113 ...	9.06	51
4	5	Shingeki no Kyojin Season 3 Part 2TV (10 eps)A...	9.05	10
5	6	Gintama'TV (51 eps)Apr 2011 - Mar 2012534,105 ...	9.04	51
6	7	Gintama: The FinalMovie (1 eps)Jan 2021 - Jan ...	9.04	1
7	8	Hunter x Hunter TV (148 eps)Oct 2011 - Sep 201...	9.04	148
8	9	Kaguya-sama wa Kokurasetai: Ultra RomanticTV (...	9.04	13
9	10	Gintama': EnchousenTV (13 eps)Oct 2012 - Mar 2...	9.03	13
10	11	Shingeki no Kyojin: The Final Season - Kankets...	9.03	2
11	12	Ginga Eiyuu DensetsuOVA (110 eps)Jan 1988 - Ma...	9.02	110
12	13	Bleach: Sennen Kessen-hen - Ketsubetsu-tanTV (...	8.99	13
13	14	Fruits Basket: The FinalTV (13 eps)Apr 2021 - ...	8.99	13
14	15	Gintama.TV (12 eps)Jan 2017 - Mar 2017302,232 ...	8.98	12
15	16	GintamaTV (201 eps)Apr 2006 - Mar 20101,034,41...	8.94	201
16	17	Koe no KatachiMovie (1 eps)Sep 2016 - Sep 2016...	8.94	1

17	18	3-gatsu no Lion 2nd SeasonTV (22 eps)Oct 2017 ...	8.93	22
18	19	Clannad: After StoryTV (24 eps)Oct 2008 - Mar ...	8.93	24
19	20	Code Geass: Hangyaku no Lelouch R2TV (25 eps)A...	8.91	25
20	21	Gintama Movie 2: Kanketsu-hen - Yorozuya yo Ei...	8.91	1
21	22	Violet Evergarden MovieMovie (1 eps)Sep 2020 -...	8.89	1
22	23	Owarimonogatari 2nd SeasonTV (7 eps)Aug 2017 -...	8.88	7
23	24	Gintama.: Shirogane no Tamashii-hen - Kouhan-s...	8.88	14
24	25	MonsterTV (74 eps)Apr 2004 - Sep 20051,041,081...	8.87	74
25	26	Kimi no Na wa.Movie (1 eps)Aug 2016 - Aug 2016...	8.84	1
26	27	Kaguya-sama wa Kokurasetai: First Kiss wa Owar...	8.84	1
27	28	Bocchi the Rock!TV (12 eps)Oct 2022 - Dec 2022...	8.83	12
28	29	The First Slam DunkMovie (1 eps)Dec 2022 - Dec...	8.83	1
29	30	Vinland Saga Season 2TV (24 eps)Jan 2023 - Jun...	8.82	24
30	31	Jujutsu Kaisen 2nd SeasonTV (23 eps)Jul 2023 -...	8.82	23
31	32	Gintama.: Shirogane no Tamashii-henTV (12 eps)...	8.81	12
32	33	Kingdom 3rd SeasonTV (26 eps)Apr 2020 - Oct 20...	8.81	26
33	34	Mob Psycho 100 IITV (13 eps)Jan 2019 - Apr 201...	8.80	13
34	35	Shingeki no Kyojin: The Final SeasonTV (16 eps...	8.80	16
35	36	Kimetsu no Yaiba: Yuukaku-henTV (11 eps)Dec 20...	8.79	11
36	37	Kizumonogatari III: Reiketsu-henMovie (1 eps)J...	8.79	1
37	38	"Oshi no Ko"TV (11 eps)Apr 2023 - Jun 2023645,...	8.79	11
38	39	Haikyuu!! Karasuno Koukou vs. Shiratorizawa Ga...	8.78	10
39	40	Sen to Chihiro no KamikakushiMovie (1 eps)Jul ...	8.78	1
40	41	Monogatari Series: Second SeasonTV (26 eps)Jul...	8.77	26
41	42	Shingeki no Kyojin: The Final Season Part 2TV ...	8.77	12
42	43	Hajime no IppoTV (75 eps)Oct 2000 - Mar 200255...	8.76	75
43	44	Cowboy BebopTV (26 eps)Apr 1998 - Apr 19991,79...	8.75	26
44	45	Kingdom 4th SeasonTV (26 eps)Apr 2022 - Oct 20...	8.75	26
45	46	Vinland SagaTV (24 eps)Jul 2019 - Dec 20191,40...	8.74	24
46	47	Mushishi Zoku Shou 2nd SeasonTV (10 eps)Oct 20...	8.73	10
47	48	Shouwa Genroku Rakugo Shinjuu: Sukeroku Futata...	8.72	12
48	49	Mob Psycho 100 IIITV (12 eps)Oct 2022 - Dec 20...	8.71	12
49	50	Rurouni Kenshin: Meiji Kenkaku Romantan - Tsui...	8.71	4

2. make a new column for time stamp

```python
In [73]: ## Function to extract timestamp
         def extraction_time(txt):
             check = False
             data = ""
             for i in range(len(txt)):
                 if txt[i] == ')':
                     for j in range(i+1,i+20):
                         data += txt[j]

                     return data
```

```python
In [74]: ## Creating a new Total time column by applying the function to title column
         df['Total Time'] = df['Title'].apply(extraction_time)
```

```python
In [75]: df.head()
```

	Rank	Title	Score	Episodes	Total Time
0	1	Fullmetal Alchemist: BrotherhoodTV (64 eps)Apr...	9.10	64	Apr 2009 - Jul 2010
1	2	Steins;GateTV (24 eps)Apr 2011 - Sep 20112,473...	9.07	24	Apr 2011 - Sep 2011
2	3	Bleach: Sennen Kessen-henTV (13 eps)Oct 2022 -...	9.06	13	Oct 2022 - Dec 2022
3	4	Gintama°TV (51 eps)Apr 2015 - Mar 2016605,113 ...	9.06	51	Apr 2015 - Mar 2016
4	5	Shingeki no Kyojin Season 3 Part 2TV (10 eps)A...	9.05	10	Apr 2019 - Jul 2019

In [76]:
```python
## Extracting the months using Total time column
from dateutil.relativedelta import relativedelta
from datetime import datetime

def calculate_total_months(period):
    try:
        start_str, end_str = period.split(' - ')
        start_date = datetime.strptime(start_str, '%b %Y')
        end_date = datetime.strptime(end_str, '%b %Y')
        r = relativedelta(end_date, start_date)
        return r.years * 12 + r.months + 1  # +1 to include the starting month
    except:
        return None

df['Months'] = df['Total Time'].apply(calculate_total_months)
```

3. which anime has the highest score

In [77]:
```python
df[df["Score"]==df["Score"].max()]["Title"]
```

Out[77]:
```
0    Fullmetal Alchemist: BrotherhoodTV (64 eps)Apr...
Name: Title, dtype: object
```

4. give me top 5 highest scoring anime

In [78]:
```python
df["Title"].head()
```

Out[78]:
```
0    Fullmetal Alchemist: BrotherhoodTV (64 eps)Apr...
1    Steins;GateTV (24 eps)Apr 2011 - Sep 20112,473...
2    Bleach: Sennen Kessen-henTV (13 eps)Oct 2022 -...
3    Gintama°TV (51 eps)Apr 2015 - Mar 2016605,113 ...
4    Shingeki no Kyojin Season 3 Part 2TV (10 eps)A...
Name: Title, dtype: object
```

5. which anime has the highest episode count

In [79]:
```python
df[df["Episodes"]==df["Episodes"].max()]["Title"]
```

Out[79]:
```
15    GintamaTV (201 eps)Apr 2006 - Mar 20101,034,41...
Name: Title, dtype: object
```

6. animes with top 5 episode count

In [ ]:
```python
df.sort_values(by="Episodes",ascending=False)["Title"].head() ## First method of extracrting top 5
```

Out[ ]:
```
15    GintamaTV (201 eps)Apr 2006 - Mar 20101,034,41...
7     Hunter x Hunter TV (148 eps)Oct 2011 - Sep 201...
11    Ginga Eiyuu DensetsuOVA (110 eps)Jan 1988 - Ma...
42    Hajime no IppoTV (75 eps)Oct 2000 - Mar 200255...
24    MonsterTV (74 eps)Apr 2004 - Sep 20051,041,081...
Name: Title, dtype: object
```

7. which is the longest running anime

In [81]:
```python
df[df["Months"]==df["Months"].max()]["Title"]
```

Out[81]:
```
11    Ginga Eiyuu DensetsuOVA (110 eps)Jan 1988 - Ma...
Name: Title, dtype: object
```

8. Top 5 longest running Animes

In [ ]:
```python
df.nlargest(5,columns="Months")["Title"] ## Second method of extracrting top 5
```

11   Ginga Eiyuu DensetsuOVA (110 eps)Jan 1988 - Ma...
        15   GintamaTV (201 eps)Apr 2006 - Mar 20101,034,41...
        7   Hunter x Hunter TV (148 eps)Oct 2011 - Sep 201...
        32   Kingdom 3rd SeasonTV (26 eps)Apr 2020 - Oct 20...
        24   MonsterTV (74 eps)Apr 2004 - Sep 20051,041,081...
        Name: Title, dtype: object

In [ ]:   ## Exporting data to csv
          df.to_csv("Anime_with_new_features.csv")