|  |
| --- |
| **import** numpy **as** np **import** pandas **as** pd  **import** matplotlib.pyplot **as** plt **import** seaborn **as** sns  df\_data**=**pd**.**read\_csv('/content/anime9.csv') df\_data |

**Name: Jeevan Ramesh Jadhav**

**Class: FY BSC CS A**

**Roll No.: 2149**

**Course Roll No.: 03**

**Task on Data cleaning and data visualization**

In [ ]:

Out[ ]: **English**

**MAL\_ID Name Score Genres Type Episodes Aired Premiered**

**name**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 1 | Cowboy Bebop | 8.78 | Action,  Adventure,  Comedy,  Drama,  Sci-Fi,  Space | Cowboy Bebop | TV | 26 | Apr  3,  1998 to Apr  24, 1999 | Spring 1998 |
| **1** | 5 | Cowboy Bebop:  Tengoku no Tobira | 8.39 | NaN | NaN | NaN | 1 | Sep  1,  2001 | Unknown |
| **2** | 6 | Trigun | 8.24 | Action,  Sci-Fi,  Adventure,  Comedy,  Drama, Shounen | Trigun | TV | 26 | Apr  1,  1998 to Sep  30, 1998 | Spring 1998 |
| **3** | 7 | Witch  Hunter Robin | 7.27 | NaN | NaN | TV | 26 | Jul 2,  2002 to Dec  24, 2002 | Summer 2002 |
| **4** | 8 | Bouken Ou Beet | 6.98 | NaN | NaN | NaN | 52 | Sep  30,  2004 to Sep  29, 2005 | Fall 2004 |
| **5** | 15 | Eyeshield 21 | 7.95 | Action,  Sports,  Comedy, Shounen | NaN | TV | 145 | Apr  6,  2005 to Mar  19, 2008 | Spring 2005 |
| **6** | 16 | Hachimitsu to Clover | 8.06 | Comedy,  Drama,  Josei,  Romance,  Slice of Life | Honey and Clover | NaN | 24 | Apr  15,  2005 to Sep  27, 2005 | Spring 2005 |
| **7** | 17 | Hungry Heart:  Wild  Striker | 7.59 | Slice of  Life,  Comedy,  Sports, Shounen | NaN | NaN | 52 | Sep  11,  2002 to Sep | Fall 2002 |

10,

2003 Apr

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Initial D  **8** 18 Fourth Stage | 8.15 | Action,  Cars,  Sports,  Drama, Seinen | NaN | TV | 24 | 17,  2004 to Feb  18, 2006 | Spring 2004 |
| **9** 19 Monster | 8.76 | NaN | NaN | TV | 74 | Apr  7,  2004 to Sep  28, 2005 | Spring 2004 |
| **10** 20 Naruto | 7.91 | Action,  Adventure,  Comedy,  Super Power, | Naruto | TV | 220 | Oct  3,  2002 to Feb | Fall 2002 |
| df\_data**.**isnull()**.**sum() |  |  |  |  |  |  |  |
| MAL\_ID 0  Name 0  Score 0  Genres 4  English name 7  Type 4  Episodes 0  Aired 0  Premiered 0  Producers 7  Licensors 7  Studios 5 Source 5 dtype: int64 |  |  |  |  |  |  |  |
| df\_data**.**describe() |  |  |  |  |  |  |  |

In [ ]:

Out[ ]:

In [ ]:

Out[ ]: **MAL\_ID Score Episodes**

|  |  |  |  |
| --- | --- | --- | --- |
| **count** | 11.000000 | 11.000000 | 11.000000 |
| **mean** | 12.000000 | 8.007273 | 60.909091 |
| **std** | 6.678323 | 0.563313 | 65.330628 |
| **min** | 1.000000 | 6.980000 | 1.000000 |
| **25%** | 6.500000 | 7.750000 | 25.000000 |
| **50%** | 15.000000 | 8.060000 | 26.000000 |
| **75%** | 17.500000 | 8.315000 | 63.000000 |
| **max** | 20.000000 | 8.780000 | 220.000000 |

|  |
| --- |
| df\_data**.**isna()**.**sum() |

In [ ]:

Out[ ]: MAL\_ID 0 Name 0

Score 0

Genres 4

English name 7

Type 4

Episodes 0

Aired 0

Premiered 0

Producers 7

Licensors 7

Studios 5 Source 5 dtype: int64

|  |
| --- |
| df\_data |

In [ ]:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 1 | Cowboy Bebop | 8.78 | Action,  Adventure,  Comedy,  Drama,  Sci-Fi,  Space | Cowboy Bebop | TV | 26 | Apr  3,  1998 to Apr  24, 1999 | Spring 1998 |
| **1** | 5 | Cowboy Bebop:  Tengoku no Tobira | 8.39 | NaN | NaN | NaN | 1 | Sep  1,  2001 | Unknown |
| **2** | 6 | Trigun | 8.24 | Action,  Sci-Fi,  Adventure,  Comedy,  Drama, Shounen | Trigun | TV | 26 | Apr  1,  1998 to Sep  30, 1998 | Spring 1998 |
| **3** | 7 | Witch  Hunter Robin | 7.27 | NaN | NaN | TV | 26 | Jul 2,  2002 to Dec  24, 2002 | Summer 2002 |
| **4** | 8 | Bouken Ou Beet | 6.98 | NaN | NaN | NaN | 52 | Sep  30,  2004 to Sep  29, 2005 | Fall 2004 |
| **5** | 15 | Eyeshield 21 | 7.95 | Action,  Sports,  Comedy, Shounen | NaN | TV | 145 | Apr  6,  2005 to Mar  19, 2008 | Spring 2005 |
| **6** | 16 | Hachimitsu to Clover | 8.06 | Comedy,  Drama,  Josei,  Romance,  Slice of Life | Honey and Clover | NaN | 24 | Apr  15,  2005 to Sep  27, 2005 | Spring 2005 |
| **7** | 17 | Hungry Heart:  Wild  Striker | 7.59 | Slice of  Life,  Comedy,  Sports, Shounen | NaN | NaN | 52 | Sep  11,  2002 to Sep | Fall 2002 |

10,

2003

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **8** | 18 | Initial D  Fourth Stage | 8.15 | Action,  Cars,  Sports,  Drama, Seinen | NaN | TV | 24 | Apr  17,  2004 to Feb  18, 2006 | Spring 2004 |
| **9** | 19 | Monster | 8.76 | NaN | NaN | TV | 74 | Apr  7,  2004 to Sep  28, 2005 | Spring 2004 |
| **10** | 20 | Naruto | 7.91 | Action,  Adventure,  Comedy, Super | Naruto | TV | 220 | Oct  3,  2002 to Feb | Fall 2002 |

Power,

|  |
| --- |
| df\_data**=**df\_data**.**drop(['Licensors'],axis**=**1) df\_data |

In [ ]:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 1 | Cowboy Bebop | 8.78 | Action,  Adventure,  Comedy,  Drama,  Sci-Fi,  Space | Cowboy Bebop | TV | 26 | Apr  3,  1998 to Apr  24, 1999 | Spring 1998 |
| **1** | 5 | Cowboy Bebop:  Tengoku no Tobira | 8.39 | NaN | NaN | NaN | 1 | Sep  1,  2001 | Unknown |
| **2** | 6 | Trigun | 8.24 | Action,  Sci-Fi,  Adventure,  Comedy,  Drama, Shounen | Trigun | TV | 26 | Apr  1,  1998 to Sep  30, 1998 | Spring 1998 |
| **3** | 7 | Witch  Hunter Robin | 7.27 | NaN | NaN | TV | 26 | Jul 2,  2002 to Dec  24, 2002 | Summer 2002 |
| **4** | 8 | Bouken Ou Beet | 6.98 | NaN | NaN | NaN | 52 | Sep  30,  2004 to Sep  29, 2005 | Fall 2004 |
| **5** | 15 | Eyeshield 21 | 7.95 | Action,  Sports,  Comedy, Shounen | NaN | TV | 145 | Apr  6,  2005 to Mar  19, 2008 | Spring 2005 |
| **6** | 16 | Hachimitsu to Clover | 8.06 | Comedy,  Drama,  Josei,  Romance,  Slice of Life | Honey and Clover | NaN | 24 | Apr  15,  2005 to Sep  27, 2005 | Spring 2005 |
| **7** | 17 | Hungry Heart:  Wild  Striker | 7.59 | Slice of  Life,  Comedy,  Sports, Shounen | NaN | NaN | 52 | Sep  11,  2002 to Sep | Fall 2002 |

10,

2003 Apr

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Initial D  **8** 18 Fourth Stage | 8.15 | Action,  Cars,  Sports,  Drama, Seinen | NaN | TV | 24 | 17,  2004 to Feb  18, 2006 | Spring 2004 |
| **9** 19 Monster | 8.76 | NaN | NaN | TV | 74 | Apr  7,  2004 to Sep  28, 2005 | Spring 2004 |
| **10** 20 Naruto | 7.91 | Action,  Adventure,  Comedy,  Super Power, | Naruto | TV | 220 | Oct  3,  2002 to Feb | Fall 2002 |
| df\_data['Producers']**.**value\_counts() | | |  |  |  |  |  |
| Victor Entertainment 1  TV Tokyo, Dentsu 1  TV Tokyo, Nihon Ad Systems, TV Tokyo Music, Shueisha 1  OB Planning, Studio Jack 1 Name: Producers, dtype: int64 | | | | |  |  |  |
| b**=**df\_data['Producers']**.**fillna('TV Tokyo') b | | | | |  |  |  |
| 1. TV Tokyo 2. TV Tokyo 3. Victor Entertainment 4. TV Tokyo 5. TV Tokyo, Dentsu5 TV Tokyo, Nihon Ad Systems, TV Tokyo Music, Sh... 6. TV Tokyo 7. TV Tokyo 8. OB Planning, Studio Jack 9. TV Tokyo 10. TV TokyoName: Producers, dtype: object | | | | |  |  |  |
| df\_data['English name']**.**fillna('One Piece') | | | | |  |  |  |

In [ ]:

Out[ ]:

In [ ]:

Out[ ]:

In [ ]:

1. Cowboy Bebop
2. One Piece
3. Trigun
4. One Piece
5. One Piece
6. One Piece
7. Honey and Clover
8. One Piece
9. One Piece
10. One Piece
11. Naruto

Name: English name, dtype: object

In [ ]: df\_data**.**value\_counts()

Out[ ]: MAL\_ID Name Score Genres Englis h name Type Episodes Aired Premiered Producers Studios Source

6 Trigun 8.24 Action, Sci-Fi, Adventure, Comedy, Drama, Shounen Trigun

TV 26 Apr 1, 1998 to Sep 30, 1998 Spring 1998 Victor Entertainment Madhouse Manga 1 dtype: int64

In [ ]: df\_data**.**fillna(method**=**'bfill')

Out[ ]:

**Type Episodes Aired Premiered**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 1 | Cowboy Bebop | | 8.78 | Action,  Adventure,  Comedy,  Drama,  Sci-Fi,  Space | Cowboy Bebop | | TV | 26 | Apr  3,  1998 to Apr  24, 1999 | Spring 1998 |
| **1** | 5 | Cowboy Bebop:  Tengoku no Tobira | | 8.39 | Action,  Sci-Fi,  Adventure,  Comedy,  Drama, Shounen | Trigun | | TV | 1 | Sep  1,  2001 | Unknown |
| **2** | 6 | Trigun | | 8.24 | Action,  Sci-Fi,  Adventure,  Comedy,  Drama, Shounen | Trigun | | TV | 26 | Apr  1,  1998 to Sep  30, 1998 | Spring 1998 |
| **3** | 7 | Witch  Hunter Robin | | 7.27 | Action,  Sports,  Comedy, Shounen | Honey and Clover | | TV | 26 | Jul 2,  2002 to Dec  24, 2002 | Summer 2002 |
| **4** | 8 | Bouken Ou Beet | | 6.98 | Action,  Sports,  Comedy, Shounen | Honey and Clover | | TV | 52 | Sep  30,  2004 to Sep  29, 2005 | Fall 2004 |
| **5** | 15 | Eyeshield 21 | | 7.95 | Action,  Sports,  Comedy, Shounen | Honey and Clover | | TV | 145 | Apr  6,  2005 to Mar  19, 2008 | Spring 2005 |
| **6** | 16 | Hachimitsu to Clover | | 8.06 | Comedy,  Drama,  Josei,  Romance,  Slice of Life | Honey and Clover | | TV | 24 | Apr  15,  2005 to Sep  27, 2005 | Spring 2005 |
| **7** | 17 | Hungry Heart: Wild | | 7.59 | Slice of  Life,  Comedy, | Naruto | | TV | 52 | Sep  11, 2002 | Fall 2002 |
|  | |  | |  |  |  | |  | **Type** | **Episodes** | **Aired** | **Premiered** |
|  | |  | | Striker |  | Sports, Shounen | |  |  |  | to Sep  10, 2003 |  |
| **8** | | 18 | | Initial D  Fourth Stage | 8.15 | Action,  Cars,  Sports,  Drama, Seinen | | Naruto | TV | 24 | Apr  17,  2004 to Feb  18, 2006 | Spring 2004 |
| **9** | | 19 | | Monster | 8.76 | Action,  Adventure,  Comedy,  Super  Power, Martia... | | Naruto | TV | 74 | Apr  7,  2004 to Sep  28, 2005 | Spring 2004 |
| **10** | | 20 | | Naruto | 7.91 | Action,  Adventure,  Comedy,  Super Power, | | Naruto | TV | 220 | Oct  3,  2002 to Feb | Fall 2002 |

|  |
| --- |
| df\_data**.**fillna('BAKA!!') |

In [ ]:

Out[ ]:

**Type Episodes Aired Premiered**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 1 | Cowboy Bebop | 8.78 | Action,  Adventure,  Comedy,  Drama,  Sci-Fi,  Space | Cowboy Bebop | TV | 26 | Apr  3,  1998 to Apr  24, 1999 | Spring 1998 |
| **1** | 5 | Cowboy Bebop:  Tengoku no Tobira | 8.39 | BAKA!! | BAKA!! | BAKA!! | 1 | Sep  1,  2001 | Unknown |
| **2** | 6 | Trigun | 8.24 | Action,  Sci-Fi,  Adventure,  Comedy,  Drama, Shounen | Trigun | TV | 26 | Apr  1,  1998 to Sep  30, 1998 | Spring 1998 |
| **3** | 7 | Witch  Hunter Robin | 7.27 | BAKA!! | BAKA!! | TV | 26 | Jul 2,  2002 to Dec  24, 2002 | Summer 2002 |
| **4** | 8 | Bouken Ou Beet | 6.98 | BAKA!! | BAKA!! | BAKA!! | 52 | Sep  30,  2004 to Sep  29, 2005 | Fall 2004 |
| **5** | 15 | Eyeshield 21 | 7.95 | Action,  Sports,  Comedy, Shounen | BAKA!! | TV | 145 | Apr  6,  2005 to Mar  19, 2008 | Spring 2005 |
| **6** | 16 | Hachimitsu to Clover | 8.06 | Comedy,  Drama,  Josei,  Romance,  Slice of Life | Honey and Clover | BAKA!! | 24 | Apr  15,  2005 to Sep  27, 2005 | Spring 2005 |
| **7** | 17 | Hungry Heart:  Wild  Striker | 7.59 | Slice of  Life,  Comedy,  Sports, Shounen | BAKA!! | BAKA!! | 52 | Sep  11,  2002 to Sep | Fall 2002 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  | **Type** | **Episodes** | **Aired**  10, 2003 | **Premiered** |
| **8** | 18 | Initial D  Fourth Stage | 8.15 | Action,  Cars,  Sports,  Drama, Seinen | BAKA!! | TV | 24 | Apr  17,  2004 to Feb  18, 2006 | Spring 2004 |
| **9** | 19 | Monster | 8.76 | BAKA!! | BAKA!! | TV | 74 | Apr  7,  2004 to Sep  28, 2005 | Spring 2004 |
| **10** | 20 | Naruto | 7.91 | Action,  Adventure,  Comedy, Super | Naruto | TV | 220 | Oct  3,  2002 to Feb | Fall 2002 |

Power,

|  |
| --- |
| df\_data**.**duplicated() |

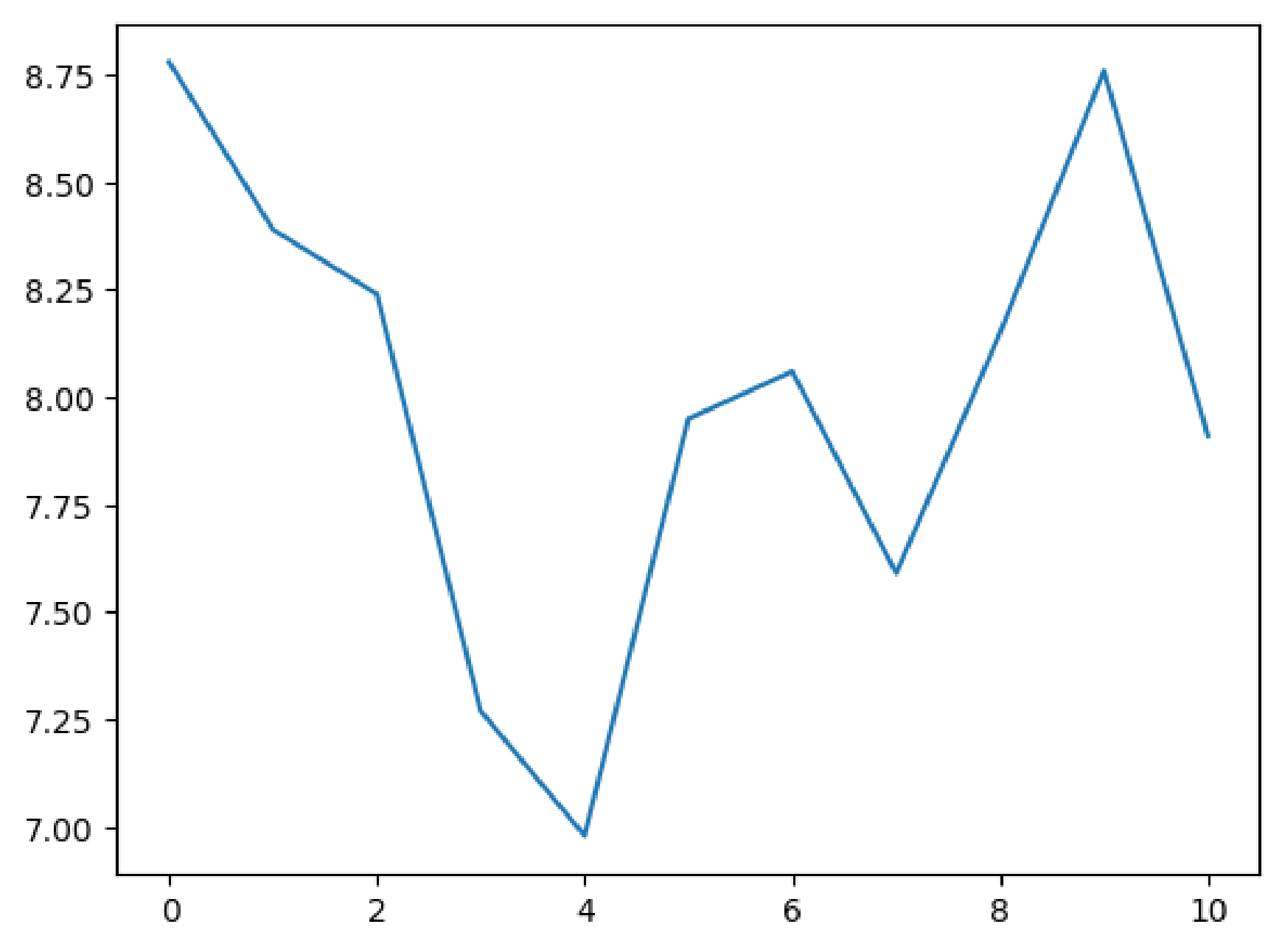
In [ ]:

Out[ ]: 0 False 1 False

1. False
2. False
3. False
4. False
5. False
6. False
7. False
8. False10 False dtype: bool

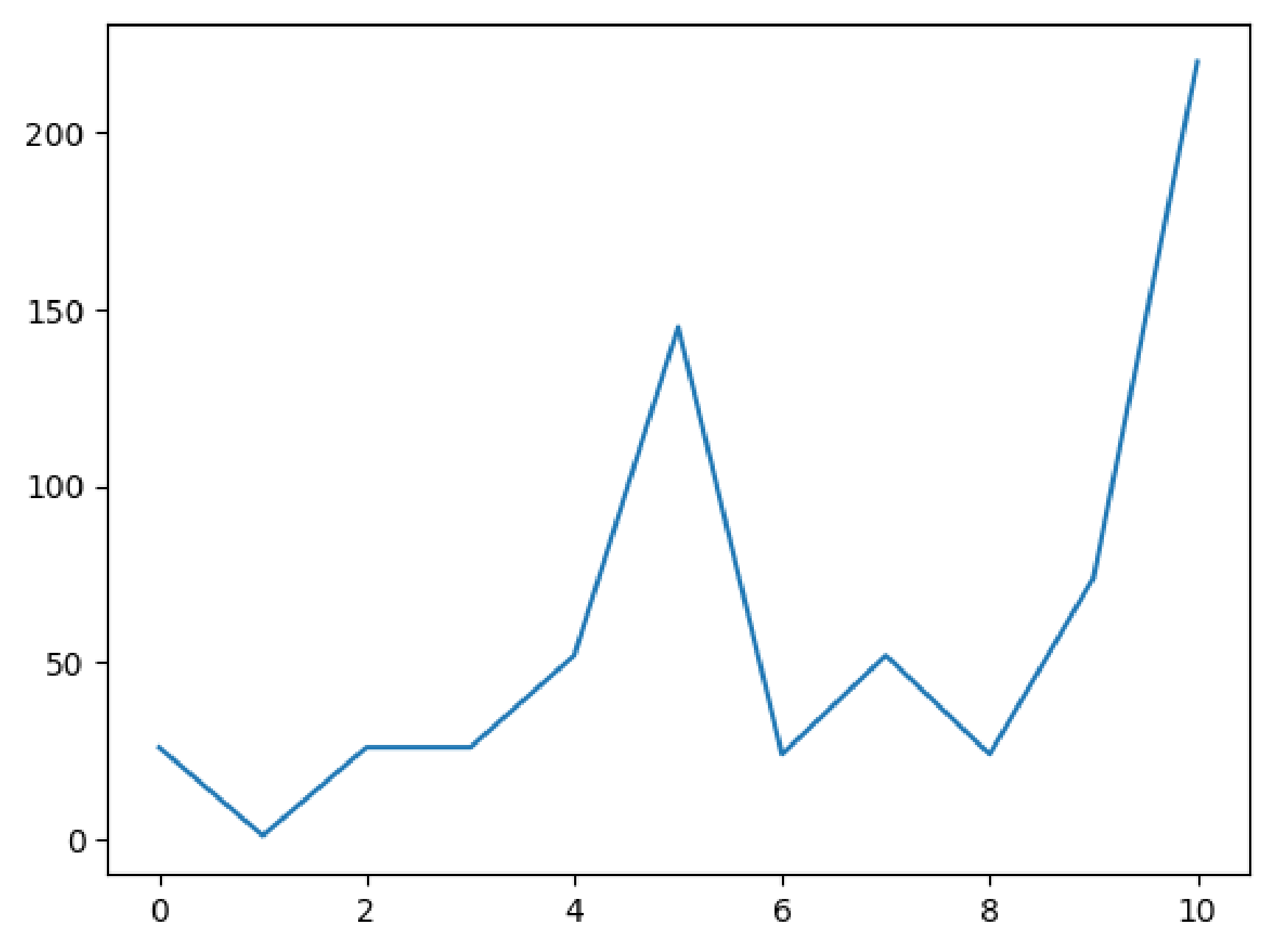
|  |
| --- |
| plt**.**plot('Score',data**=**df\_data) plt**.**show() |

In [ ]:



|  |
| --- |
| plt**.**plot('Episodes',data**=**df\_data) plt**.**show() |

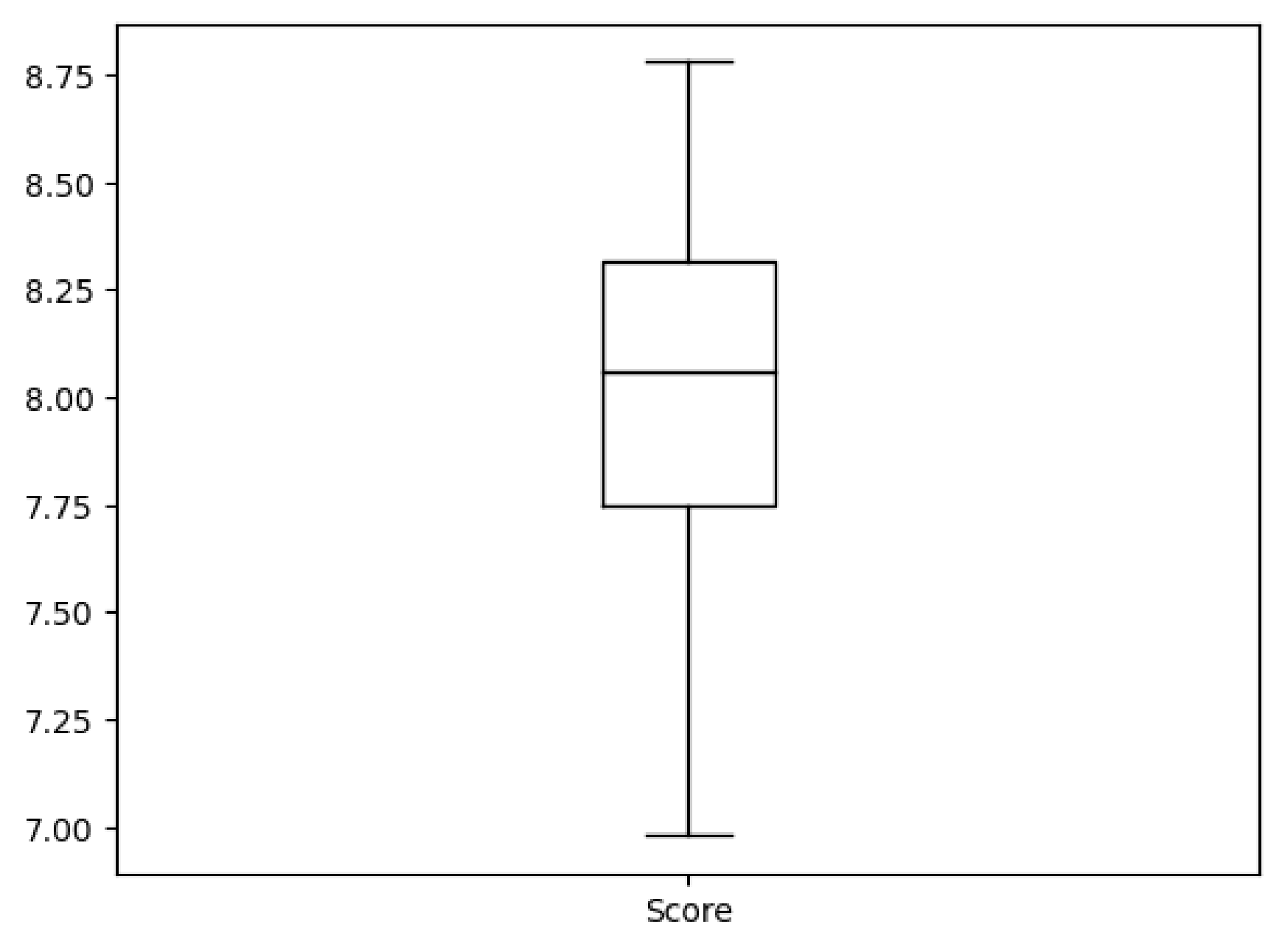
In [ ]:



|  |
| --- |
| df\_data**.**boxplot(column**=**['Score'], grid**=False**, color**=**'black') |

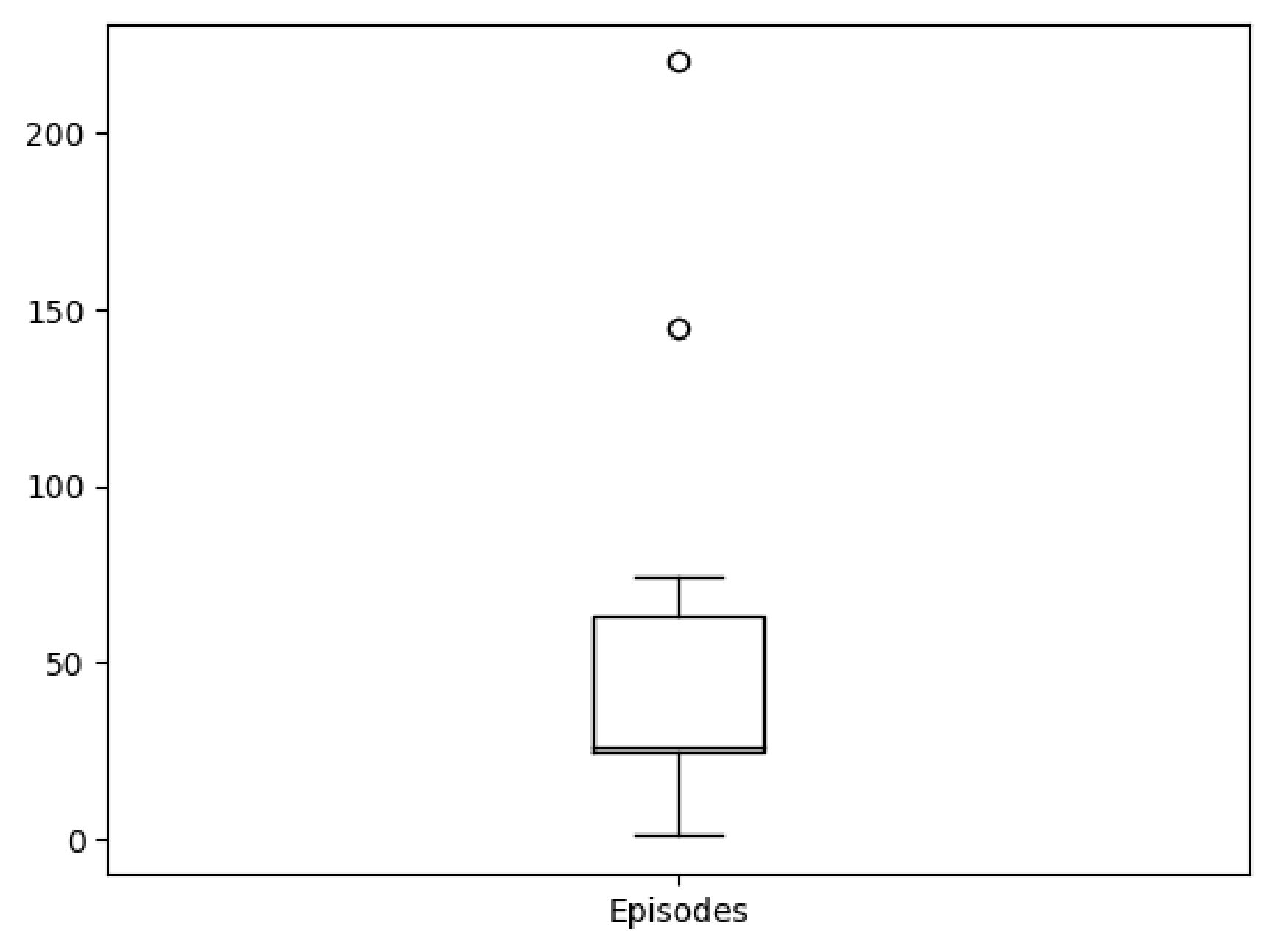
In [ ]:

Out[ ]: <Axes: >



|  |
| --- |
| df\_data**.**boxplot(column**=**['Episodes'], grid**=False**, color**=**'black') |

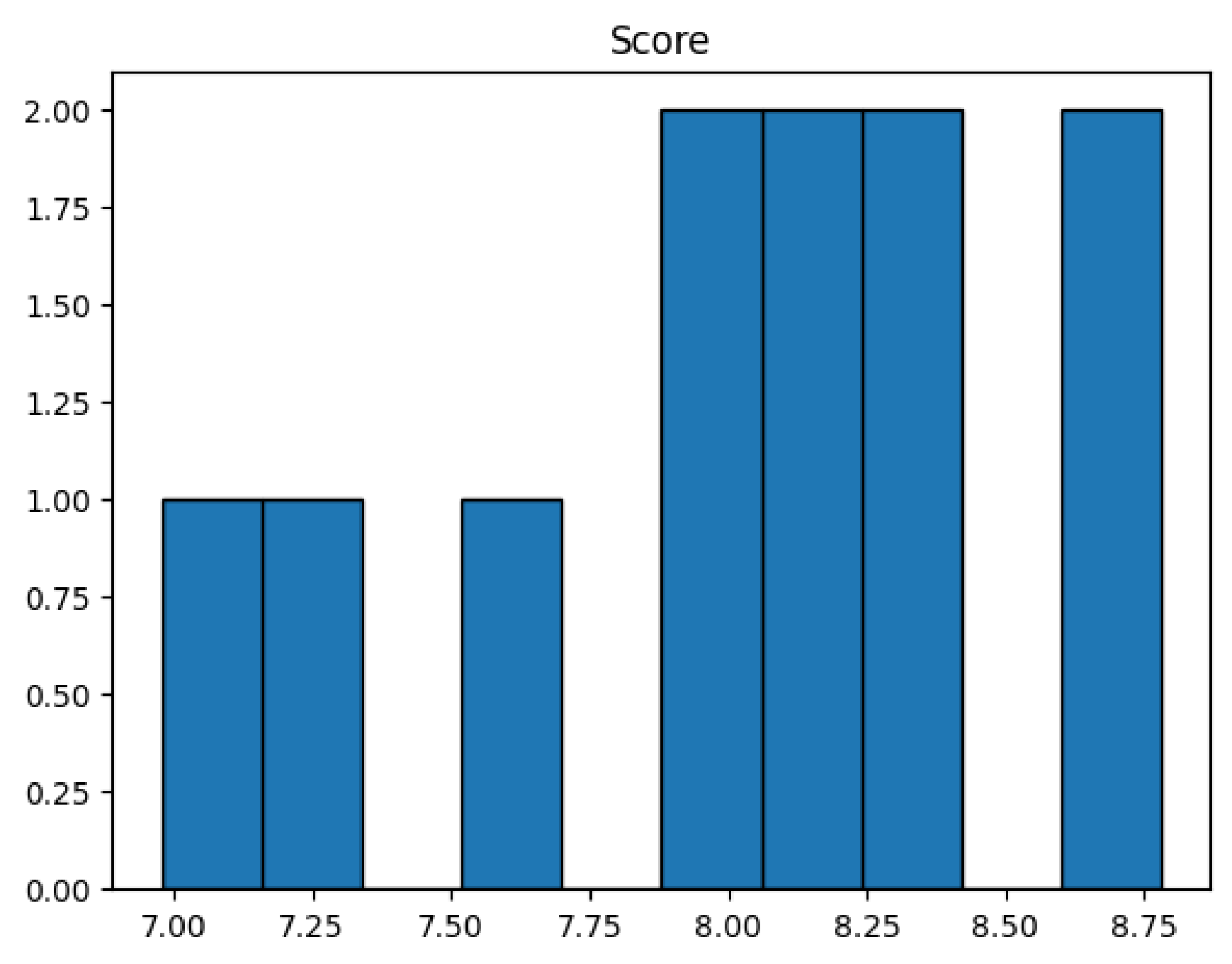
In [ ]: Out[ ]: <Axes: >



|  |
| --- |
| df\_data**.**hist(column**=**'Score', grid**=False**, edgecolor**=**'black') |

In [ ]:

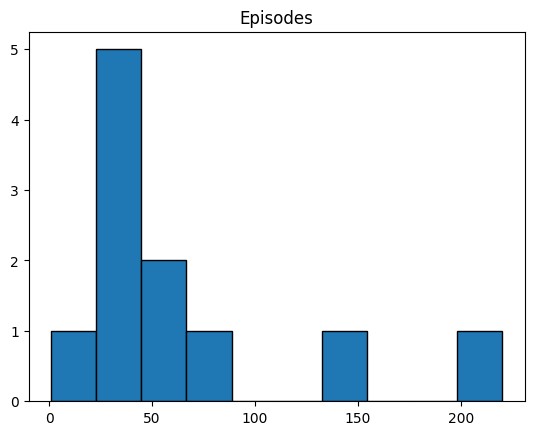
Out[ ]: array([[<Axes: title={'center': 'Score'}>]], dtype=object)



|  |
| --- |
| df\_data**.**hist(column**=**'Episodes', grid**=False**, edgecolor**=**'black') |

In [ ]:

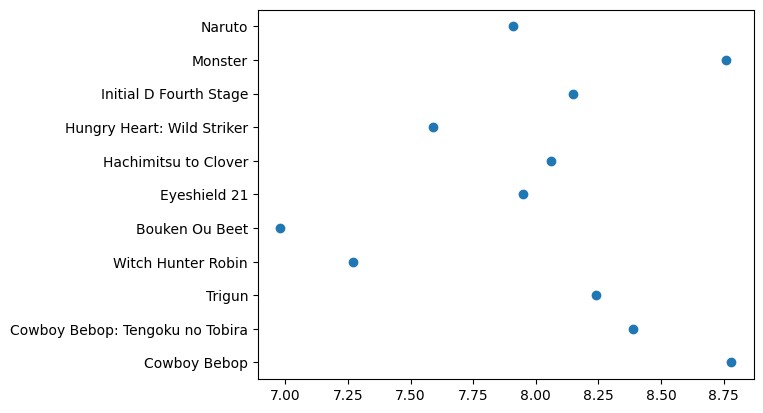
Out[ ]: array([[<Axes: title={'center': 'Episodes'}>]], dtype=object)



|  |
| --- |
| **import** matplotlib.pyplot **as** plt  x**=**('Name') y**=**('Score') plt**.**scatter('Score','Name',data**=**df\_data) plt**.**show |

In [ ]:

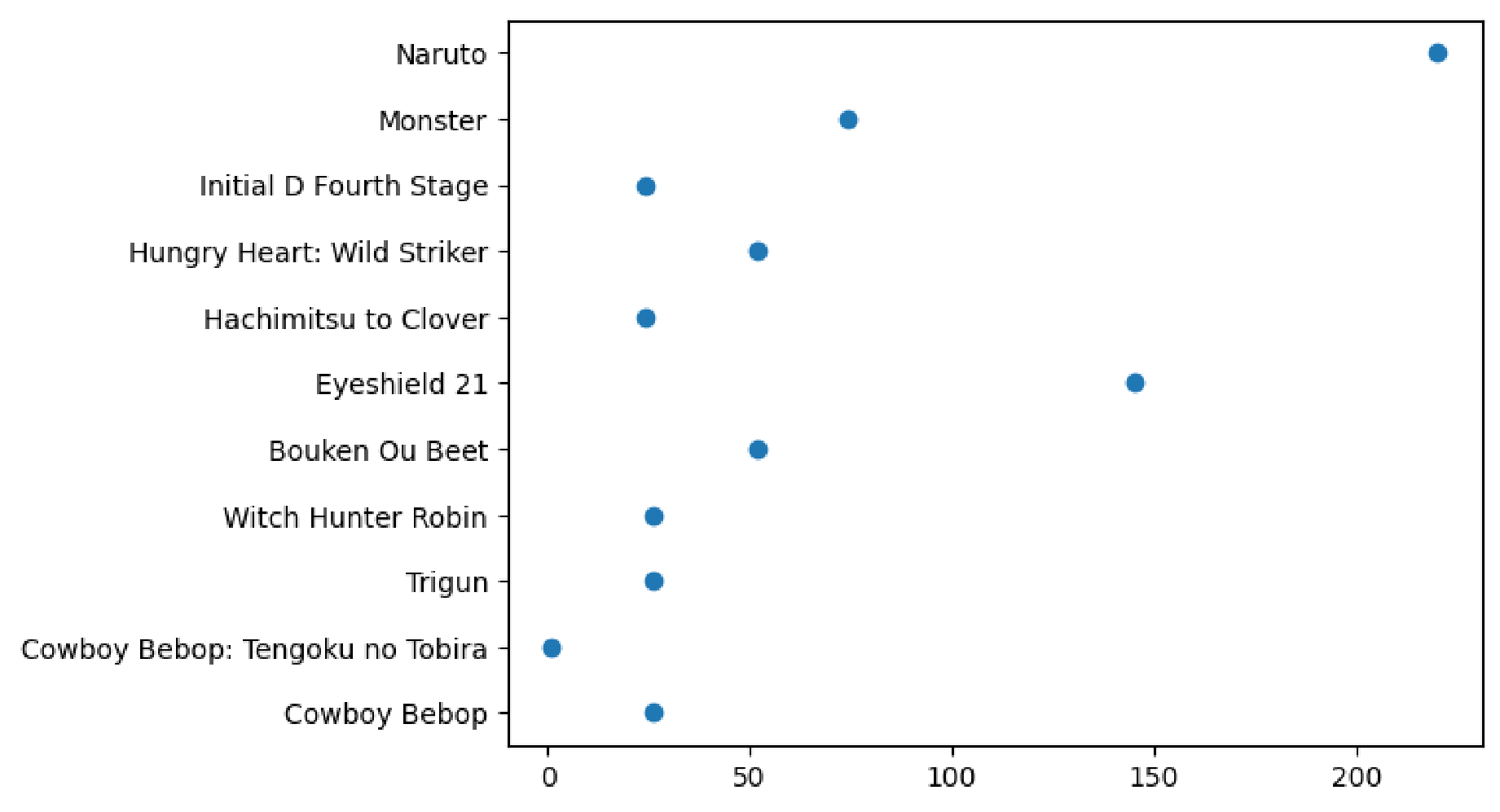
Out[ ]: <function matplotlib.pyplot.show(close=None, block=None)>



|  |
| --- |
| plt**.**scatter('Episodes','Name',data**=**df\_data) plt**.**show |

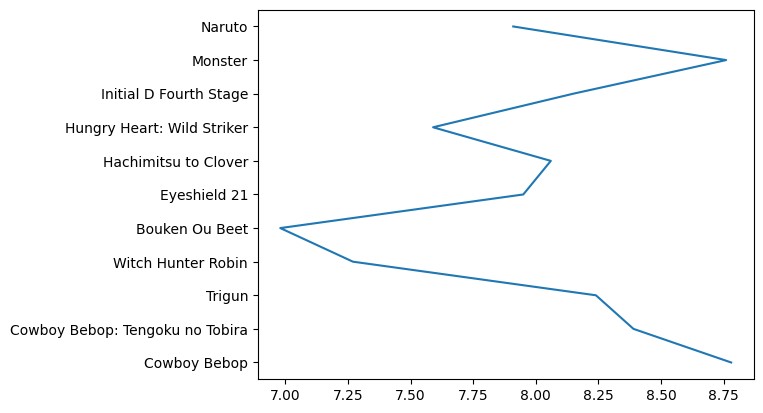
In [ ]:

Out[ ]: <function matplotlib.pyplot.show(close=None, block=None)>



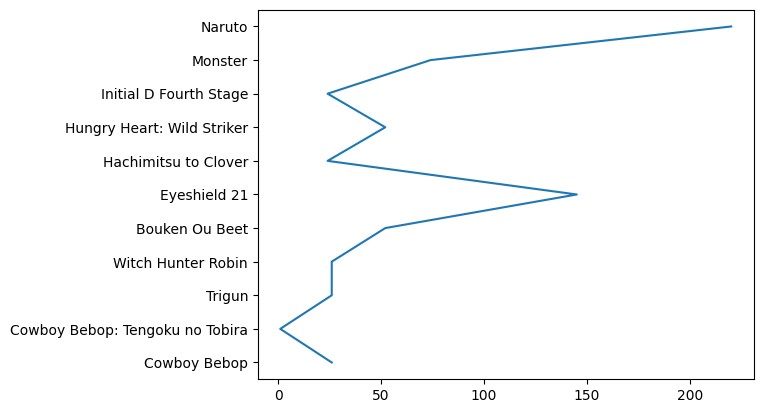
|  |
| --- |
| plt**.**plot('Score','Name',data**=**df\_data) plt**.**show() |

In [ ]:



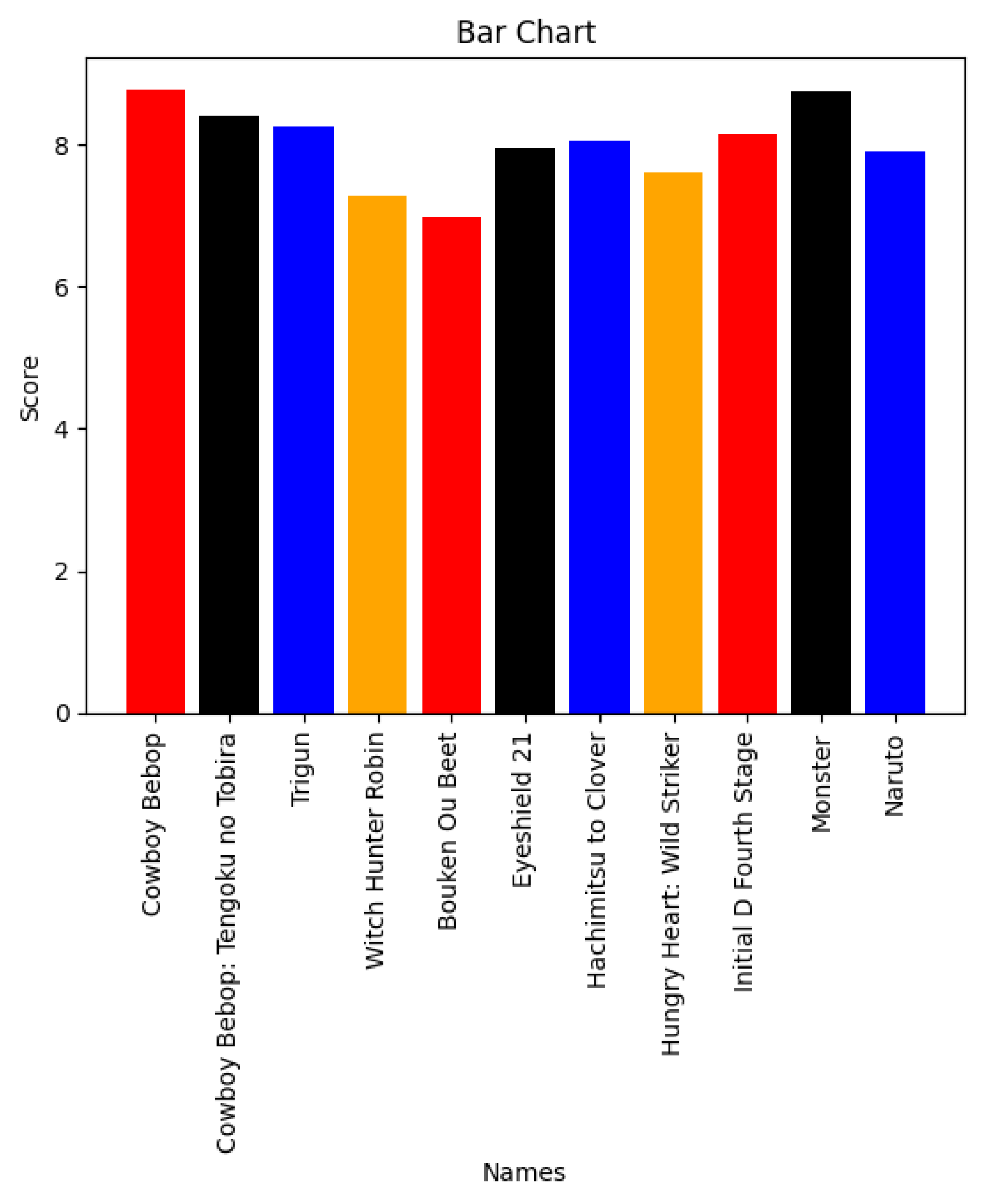
|  |
| --- |
| plt**.**plot('Episodes','Name',data**=**df\_data) plt**.**show() |

In [ ]:



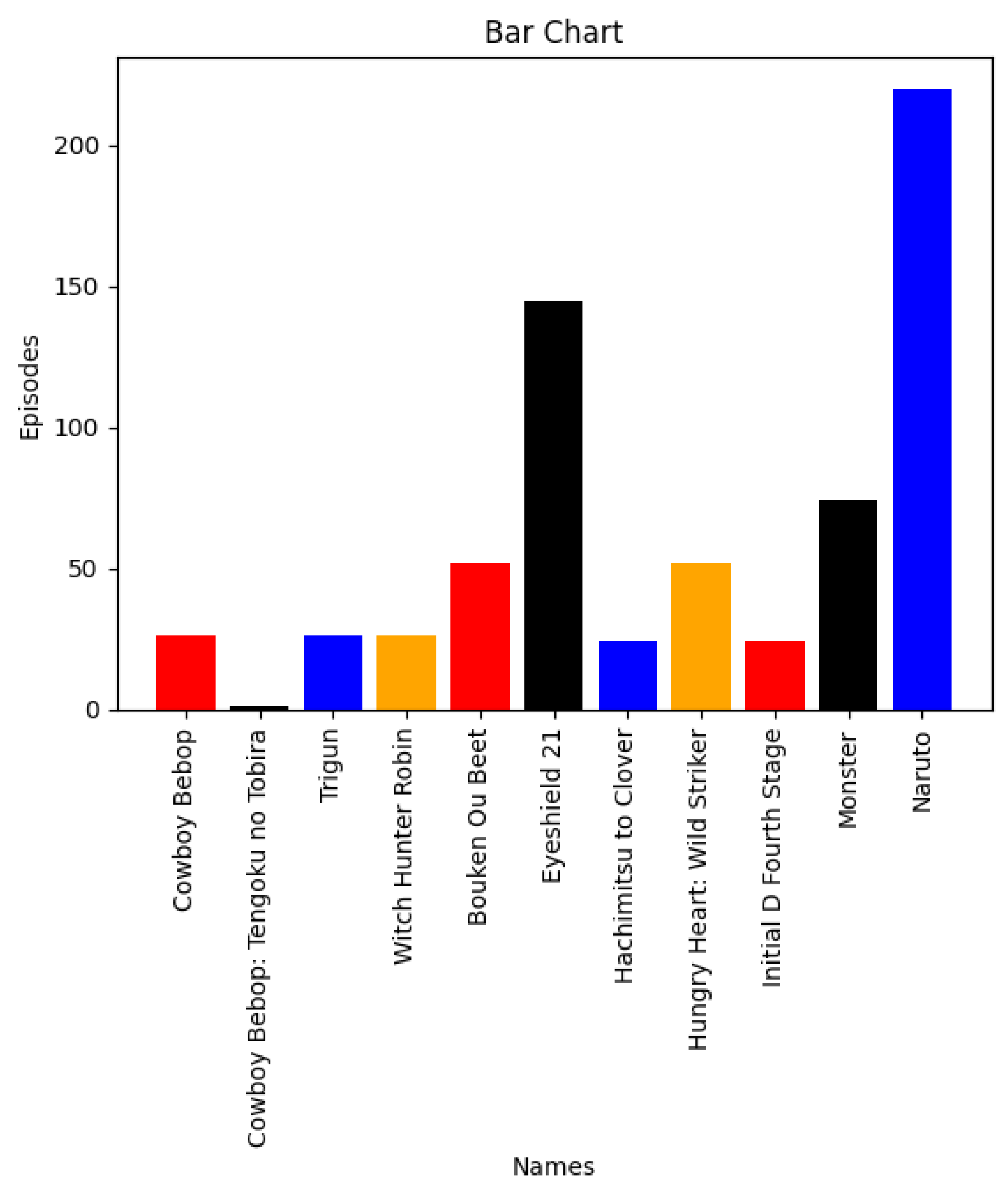
|  |
| --- |
| colors**=**['r','k','b','orange'] plt**.**bar('Name','Score',color**=**colors,data**=**df\_data) plt**.**title('Bar Chart') plt**.**xlabel('Names') plt**.**ylabel('Score') plt**.**xticks(rotation **=** 90) plt**.**show() |

In [ ]:



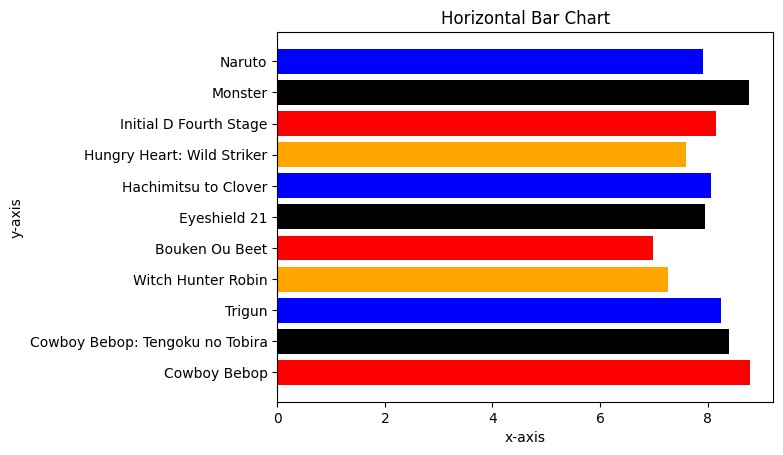
|  |
| --- |
| colors**=**['r','k','b','orange'] plt**.**bar('Name','Episodes',color**=**colors,data**=**df\_data) plt**.**title('Bar Chart') plt**.**xlabel('Names') plt**.**ylabel('Episodes') plt**.**xticks(rotation **=** 90) plt**.**show() |

In [ ]:



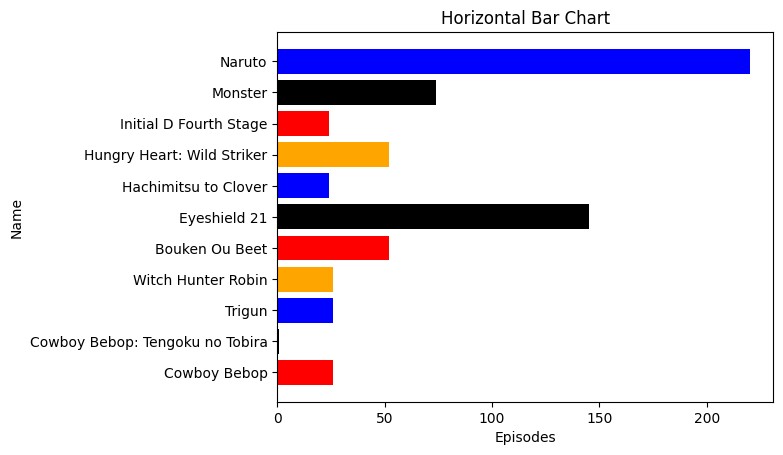
|  |
| --- |
| colors**=**['r','k','b','orange'] plt**.**barh('Name','Score',color**=**colors,data**=**df\_data) plt**.**title('Horizontal Bar Chart') plt**.**xlabel('Score') plt**.**ylabel('Name') plt**.**show() |

In [ ]:



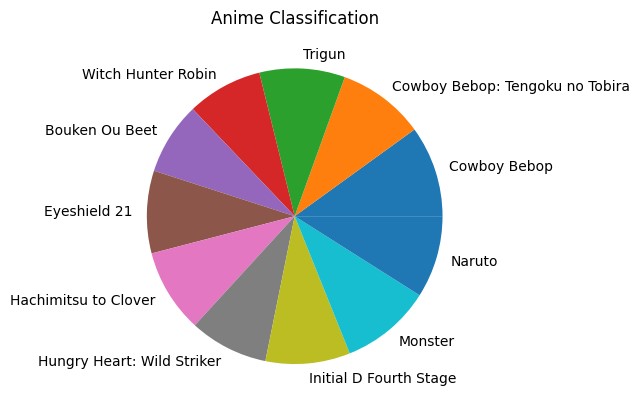
|  |
| --- |
| colors**=**['r','k','b','orange'] plt**.**barh('Name','Episodes',color**=**colors,data**=**df\_data) plt**.**title('Horizontal Bar Chart') plt**.**xlabel('Episodes') plt**.**ylabel('Name') plt**.**show() |

In [ ]:



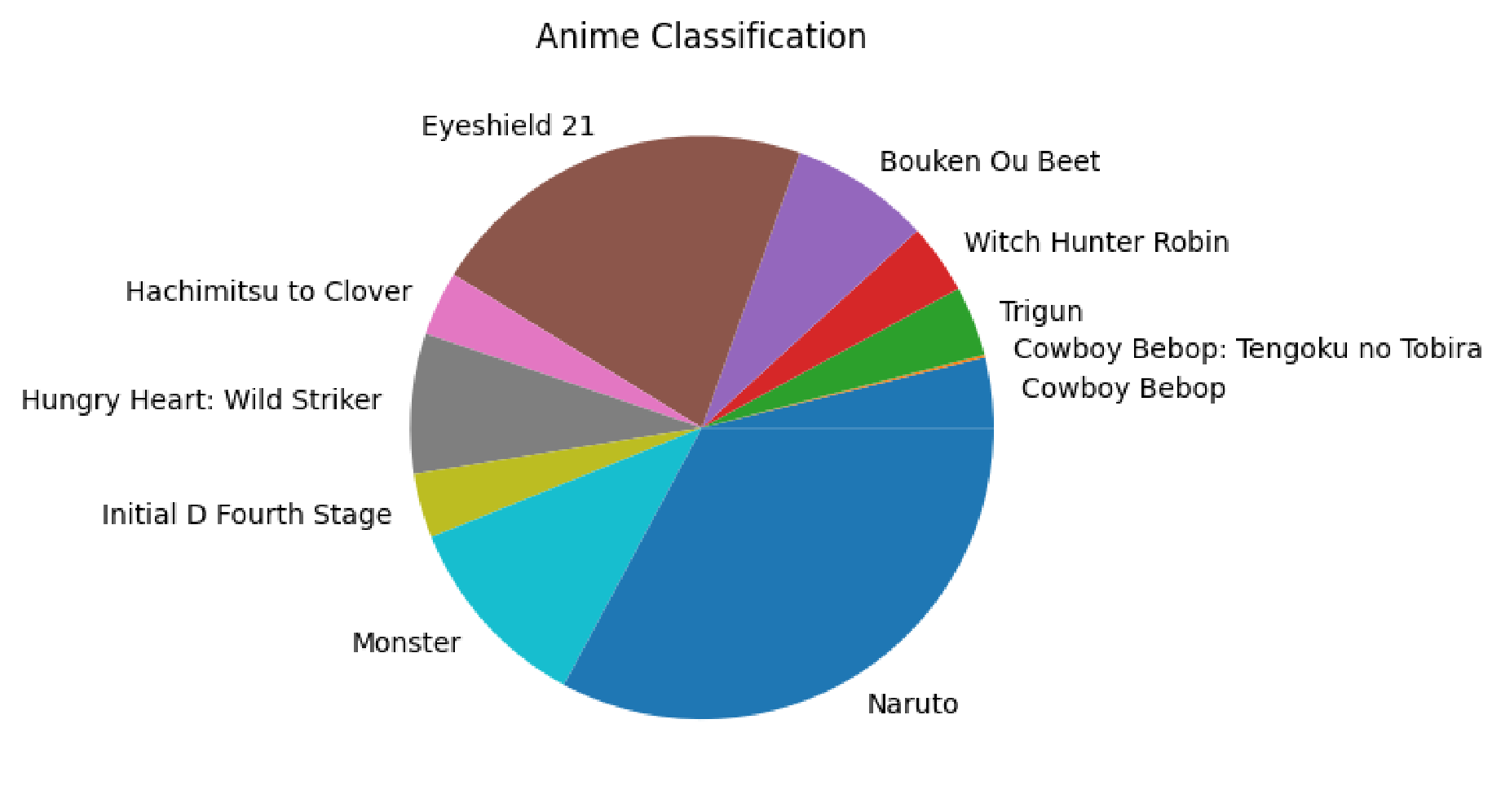
|  |
| --- |
| plt**.**pie('Score',labels**=**'Name',data**=**df\_data) plt**.**title('Anime Classification') plt**.**show() |

In [ ]:



|  |
| --- |
| plt**.**pie('Episodes',labels**=**'Name',data**=**df\_data) plt**.**title('Anime Classification') plt**.**show() |

In [ ]:



Conclusion: In the given Data, we have a total number of 11 Anime which may consist of anime or a movie, in which the relative score of those anime varies from a range of 7.00-8.75 and when we take episodes then we could see that 9 outoff 11 anime have their episodes in the range of 1-52 but there are outliers which have there number of episodes out of the range.This a study of the univariate analysis, but when we do a bivariate analysis then we come to know that as the particular anime like'Naruto' have a huge amount of episodes in it in the comparison of other animes which affects its overall Score, which makes us to understand that when the episodes are more then it affects the overall Score of the anime and when anime has less number of episodes then it may not affect the score of it.

When we compare the episodes and score of the anime then we understand that for the animes which have a small time duration then we can invest up into the anime called Cowboy bebop have a great number of Rating Score with the increasing numbers of episodes,but when we look on the other anime named hungry heart it's score decreases with it's increasing episodes. If we look upon the log time duration Anime then we can invest up into the anime called Naruto as with thye increasing of episodes it overall mean rating Score doesn't fall sharply,but when we consider the same type of other anime eyeshield 21 then we can see that it has a less amount of episodes but its score is decreasing with it's increasing number of episodes.