day-3-pandas-dataframe-part-1

January 2, 2023

```python
[95]: import pandas as pd
      import numpy as np
```

1 Creating a DataFrame

creating dataframe using list

```python
[96]: student_data = [
          [100,50,80],
          [90,70,7],
          [120,100,14],
          [78,50,81]
      ]
      pd.DataFrame(student_data,columns=["iq","marks","package"])
```

```
[96]:     iq  marks  package
      0  100     50       80
      1   90     70        7
      2  120    100       14
      3   78     50       81
```

Creating a dataframe using dict

```python
[200]: student = {
           "Name" : ["rohit","niddy","pushpa","ankit","piyush","vishal"],
           "id" :       [100,90,120,80,0,0],
           "marks" :   [55,66,93,78,0,0],
           "package" : [14,11,13,15,0,0]
       }
       stu=pd.DataFrame(student)
```

```python
[201]: stu
```

```
[201]:      Name   id  marks  package
       0   rohit  100     55       14
       1   niddy   90     66       11
       2  pushpa  120     93       13
```

```
3    ankit    80    78    15
4    piyush    0     0     0
5    vishal    0     0     0
```

Read CSV file

```python
[99]: movie=pd.read_csv("/content/movies.csv")
```

```python
[100]: movie
```

```
[100]:                                      title_x       imdb_id  \
       0                    Uri: The Surgical Strike    tt8291224
       1                                Battalion 609   tt9472208
       2        The Accidental Prime Minister (film)   tt6986710
       3                              Why Cheat India   tt8108208
       4                              Evening Shadows   tt6028796
       ...                                        ...         ...
       1624                    Tera Mera Saath Rahen    tt0301250
       1625                    Yeh Zindagi Ka Safar    tt0298607
       1626                         Sabse Bada Sukh   tt0069204
       1627                                    Daaka  tt10833860
       1628                                 Humsafar   tt2403201

                                             poster_path  \
       0      https://upload.wikimedia.org/wikipedia/en/thum…
       1                                                  NaN
       2      https://upload.wikimedia.org/wikipedia/en/thum…
       3      https://upload.wikimedia.org/wikipedia/en/thum…
       4                                                  NaN
       ...                                                ...
       1624   https://upload.wikimedia.org/wikipedia/en/2/2b…
       1625   https://upload.wikimedia.org/wikipedia/en/thum…
       1626                                               NaN
       1627   https://upload.wikimedia.org/wikipedia/en/thum…
       1628   https://upload.wikimedia.org/wikipedia/en/thum…

                                              wiki_link  \
       0      https://en.wikipedia.org/wiki/Uri:_The_Surgica…
       1              https://en.wikipedia.org/wiki/Battalion_609
       2      https://en.wikipedia.org/wiki/The_Accidental_P…
       3          https://en.wikipedia.org/wiki/Why_Cheat_India
       4          https://en.wikipedia.org/wiki/Evening_Shadows
       ...                                                ...
       1624   https://en.wikipedia.org/wiki/Tera_Mera_Saath_…
       1625   https://en.wikipedia.org/wiki/Yeh_Zindagi_Ka_S…
       1626       https://en.wikipedia.org/wiki/Sabse_Bada_Sukh
       1627                 https://en.wikipedia.org/wiki/Daaka
```

```
1628                 https://en.wikipedia.org/wiki/Humsafar

                              title_y              original_title  is_adult  \
0            Uri: The Surgical Strike      Uri: The Surgical Strike         0
1                       Battalion 609                 Battalion 609         0
2          The Accidental Prime Minister   The Accidental Prime Minister   0
3                       Why Cheat India               Why Cheat India      0
4                       Evening Shadows               Evening Shadows      0
...                               ...                           ...      ...
1624               Tera Mera Saath Rahen         Tera Mera Saath Rahen     0
1625               Yeh Zindagi Ka Safar          Yeh Zindagi Ka Safar      0
1626                      Sabse Bada Sukh               Sabse Bada Sukh    0
1627                              Daaka                         Daaka      0
1628                            Humsafar                      Humsafar     0

      year_of_release runtime          genres  imdb_rating  imdb_votes  \
0                2019     138  Action|Drama|War          8.4       35112
1                2019     131              War          4.1          73
2                2019     112  Biography|Drama          6.1        5549
3                2019     121      Crime|Drama          6.0        1891
4                2018     102            Drama          7.3         280
...               ...     ...              ...          ...         ...
1624             2001     148            Drama          4.9         278
1625             2001     146            Drama          3.0         133
1626             2018      \N     Comedy|Drama          6.1          13
1627             2019     136           Action          7.4          38
1628             2011      35    Drama|Romance          9.0        2968

                                            story  \
0        Divided over five chapters  the film chronicle…
1        The story revolves around a cricket match betw…
2        Based on the memoir by Indian policy analyst S…
3        The movie focuses on existing malpractices in …
4        While gay rights and marriage equality has bee…
...                                            …
1624  Raj Dixit lives with his younger brother  Rahu…
1625  Hindi pop-star  Sarina Devan  lives a wealthy …
1626  Village born Lalloo re-locates to Bombay  and …
1627  Shinda tries robbing a bank so he can be wealt…
1628  Sara and Ashar are childhood friends who share…

                                          summary tagline  \
0        Indian army special forces execute a covert op…     NaN
1        The story of Battalion 609 revolves around a c…     NaN
2        Explores Manmohan Singh's tenure as the Prime …     NaN
3        The movie focuses on existing malpractices in …     NaN
4        Under the 'Evening Shadows'  truth often plays…     NaN
```

```
...                                                      ...       ...
1624  A man is torn between his handicapped brother …     NaN
1625  A singer finds out she was adopted when the ed…     NaN
1626  Village born Lalloo re-locates to Bombay  and …     NaN
1627  Shinda tries robbing a bank so he can be wealt…     NaN
1628  Ashar and Khirad are forced to get married due…     NaN

                                                   actors  \
0        Vicky Kaushal|Paresh Rawal|Mohit Raina|Yami Ga…
1        Vicky Ahuja|Shoaib Ibrahim|Shrikant Kamat|Elen…
2        Anupam Kher|Akshaye Khanna|Aahana Kumra|Atul S…
3        Emraan Hashmi|Shreya Dhanwanthary|Snighdadeep …
4        Mona Ambegaonkar|Ananth Narayan Mahadevan|Deva…
...                                                    ...
1624     Ajay Devgn|Sonali Bendre|Namrata Shirodkar|Pre…
1625     Ameesha Patel|Jimmy Sheirgill|Nafisa Ali|Gulsh…
1626     Vijay Arora|Asrani|Rajni Bala|Kumud Damle|Utpa…
1627                             Gippy Grewal|Zareen Khan|
1628                                          Fawad Khan|

           wins_nominations              release_date
0                   4 wins    11 January 2019 (USA)
1                      NaN   11 January 2019 (India)
2                      NaN     11 January 2019 (USA)
3                      NaN      18 January 2019 (USA)
4      17 wins & 1 nomination   11 January 2019 (India)
...                      ...                       ...
1624                   NaN    7 November 2001 (India)
1625                   NaN   16 November 2001 (India)
1626                   NaN                        NaN
1627                   NaN     1 November 2019 (USA)
1628                   NaN       TV Series (2011-2012)

[1629 rows x 18 columns]
```

```python
ipl = pd.read_csv("/content/ipl-matches.csv")
ipl
```

```
          ID       City         Date   Season  MatchNumber  \
0    1312200   Ahmedabad   2022-05-29     2022        Final
1    1312199   Ahmedabad   2022-05-27     2022  Qualifier 2
2    1312198     Kolkata   2022-05-25     2022    Eliminator
3    1312197     Kolkata   2022-05-24     2022  Qualifier 1
4    1304116      Mumbai   2022-05-22     2022           70
..       ...         ...          ...      ...          ...
945   335986     Kolkata   2008-04-20  2007/08            4
946   335985      Mumbai   2008-04-20  2007/08            5
```

4

```
947     335984        Delhi  2008-04-19  2007/08                  3
948     335983   Chandigarh  2008-04-19  2007/08                  2
949     335982    Bangalore  2008-04-18  2007/08                  1

                            Team1                         Team2  \
0                 Rajasthan Royals                Gujarat Titans
1      Royal Challengers Bangalore              Rajasthan Royals
2      Royal Challengers Bangalore          Lucknow Super Giants
3                 Rajasthan Royals                Gujarat Titans
4               Sunrisers Hyderabad                  Punjab Kings
..                              …                             …
945           Kolkata Knight Riders               Deccan Chargers
946                  Mumbai Indians   Royal Challengers Bangalore
947                 Delhi Daredevils              Rajasthan Royals
948                 Kings XI Punjab           Chennai Super Kings
949     Royal Challengers Bangalore         Kolkata Knight Riders

                                      Venue                     TossWinner  \
0                Narendra Modi Stadium, Ahmedabad          Rajasthan Royals
1                Narendra Modi Stadium, Ahmedabad          Rajasthan Royals
2                            Eden Gardens, Kolkata      Lucknow Super Giants
3                            Eden Gardens, Kolkata            Gujarat Titans
4                       Wankhede Stadium, Mumbai       Sunrisers Hyderabad
..                                         …                             …
945                              Eden Gardens               Deccan Chargers
946                         Wankhede Stadium                Mumbai Indians
947                         Feroz Shah Kotla            Rajasthan Royals
948   Punjab Cricket Association Stadium, Mohali       Chennai Super Kings
949               M Chinnaswamy Stadium   Royal Challengers Bangalore

    TossDecision SuperOver                WinningTeam    WonBy  Margin  \
0            bat         N             Gujarat Titans  Wickets     7.0
1          field         N           Rajasthan Royals  Wickets     7.0
2          field         N  Royal Challengers Bangalore    Runs    14.0
3          field         N             Gujarat Titans  Wickets     7.0
4            bat         N               Punjab Kings  Wickets     5.0
..           …         …                        …        …       …
945          bat         N       Kolkata Knight Riders  Wickets     5.0
946          bat         N  Royal Challengers Bangalore  Wickets     5.0
947          bat         N           Delhi Daredevils  Wickets     9.0
948          bat         N        Chennai Super Kings     Runs    33.0
949        field         N       Kolkata Knight Riders     Runs   140.0

    method Player_of_Match                              Team1Players  \
0      NaN       HH Pandya   ['YBK Jaiswal', 'JC Buttler', 'SV Samson', 'D …
1      NaN       JC Buttler  ['V Kohli', 'F du Plessis', 'RM Patidar', 'GJ …
2      NaN       RM Patidar  ['V Kohli', 'F du Plessis', 'RM Patidar', 'GJ …
```

```
3       NaN      DA Miller    ['YBK Jaiswal', 'JC Buttler', 'SV Samson', 'D …
4       NaN   Harpreet Brar   ['PK Garg', 'Abhishek Sharma', 'RA Tripathi', …
..       …          …                                                      …
945     NaN      DJ Hussey    ['WP Saha', 'BB McCullum', 'RT Ponting', 'SC G…
946     NaN      MV Boucher   ['L Ronchi', 'ST Jayasuriya', 'DJ Thornely', '…
947     NaN      MF Maharoof  ['G Gambhir', 'V Sehwag', 'S Dhawan', 'MK Tiwa…
948     NaN      MEK Hussey   ['K Goel', 'JR Hopes', 'KC Sangakkara', 'Yuvra…
949     NaN      BB McCullum  ['R Dravid', 'W Jaffer', 'V Kohli', 'JH Kallis…

                                        Team2Players        Umpire1  \
0     ['WP Saha', 'Shubman Gill', 'MS Wade', 'HH Pan…    CB Gaffaney
1     ['YBK Jaiswal', 'JC Buttler', 'SV Samson', 'D …    CB Gaffaney
2     ['Q de Kock', 'KL Rahul', 'M Vohra', 'DJ Hooda…  J Madanagopal
3     ['WP Saha', 'Shubman Gill', 'MS Wade', 'HH Pan…    BNJ Oxenford
4     ['JM Bairstow', 'S Dhawan', 'M Shahrukh Khan',…    AK Chaudhary
..                                                 …              …
945   ['AC Gilchrist', 'Y Venugopal Rao', 'VVS Laxma…      BF Bowden
946   ['S Chanderpaul', 'R Dravid', 'LRPL Taylor', '…       SJ Davis
947   ['T Kohli', 'YK Pathan', 'SR Watson', 'M Kaif'…      Aleem Dar
948   ['PA Patel', 'ML Hayden', 'MEK Hussey', 'MS Dh…      MR Benson
949   ['SC Ganguly', 'BB McCullum', 'RT Ponting', 'D…      Asad Rauf

            Umpire2
0        Nitin Menon
1        Nitin Menon
2           MA Gough
3          VK Sharma
4      NA Patwardhan
..               …
945      K Hariharan
946        DJ Harper
947  GA Pratapkumar
948       SL Shastri
949       RE Koertzen

[950 rows x 20 columns]
```

2 DataFrame Method

Shape,dtype,index,column,values,head,tail,sample,info,isnull,duplicated,rename

```
[102]: movie.shape  #shape
```

```
[102]: (1629, 18)
```

```
[103]: ipl.shape
```

```
[103]: (950, 20)
```

```
[104]: movie.dtypes
```

```
[104]: title_x              object
       imdb_id              object
       poster_path          object
       wiki_link            object
       title_y              object
       original_title       object
       is_adult              int64
       year_of_release       int64
       runtime              object
       genres               object
       imdb_rating         float64
       imdb_votes            int64
       story                object
       summary              object
       tagline              object
       actors               object
       wins_nominations     object
       release_date         object
       dtype: object
```

```
[105]: ipl.dtypes
```

```
[105]: ID                    int64
       City                 object
       Date                 object
       Season               object
       MatchNumber          object
       Team1                object
       Team2                object
       Venue                object
       TossWinner           object
       TossDecision         object
       SuperOver            object
       WinningTeam          object
       WonBy                object
       Margin              float64
       method               object
       Player_of_Match      object
       Team1Players         object
       Team2Players         object
       Umpire1              object
       Umpire2              object
       dtype: object
```

```
[106]: movie.index    #index
```

```
[106]: RangeIndex(start=0, stop=1629, step=1)
```

```
[107]: ipl.index
```

```
[107]: RangeIndex(start=0, stop=950, step=1)
```

```
[108]: #column
       movie.columns
```

```
[108]: Index(['title_x', 'imdb_id', 'poster_path', 'wiki_link', 'title_y',
              'original_title', 'is_adult', 'year_of_release', 'runtime', 'genres',
              'imdb_rating', 'imdb_votes', 'story', 'summary', 'tagline', 'actors',
              'wins_nominations', 'release_date'],
             dtype='object')
```

```
[109]: ipl.columns
```

```
[109]: Index(['ID', 'City', 'Date', 'Season', 'MatchNumber', 'Team1', 'Team2',
              'Venue', 'TossWinner', 'TossDecision', 'SuperOver', 'WinningTeam',
              'WonBy', 'Margin', 'method', 'Player_of_Match', 'Team1Players',
              'Team2Players', 'Umpire1', 'Umpire2'],
             dtype='object')
```

```
[110]: stu.columns
```

```
[110]: Index(['id', 'marks', 'package'], dtype='object')
```

```
[111]: #values
       stu.values
```

```
[111]: array([[100,  55,  14],
              [ 90,  66,  11],
              [120,  93,  13],
              [ 80,  78,  15],
              [  0,   0,   0],
              [  0,   0,   0]])
```

```
[112]: ipl.values
```

```
[112]: array([[1312200, 'Ahmedabad', '2022-05-29', …,
              "['WP Saha', 'Shubman Gill', 'MS Wade', 'HH Pandya', 'DA Miller', 'R
       Tewatia', 'Rashid Khan', 'R Sai Kishore', 'LH Ferguson', 'Yash Dayal', 'Mohammed
       Shami']",
              'CB Gaffaney', 'Nitin Menon'],
              [1312199, 'Ahmedabad', '2022-05-27', …,
```

```
         "['YBK Jaiswal', 'JC Buttler', 'SV Samson', 'D Padikkal', 'SO Hetmyer',
  'R Parag', 'R Ashwin', 'TA Boult', 'YS Chahal', 'M Prasidh Krishna', 'OC
  McCoy']",
         'CB Gaffaney', 'Nitin Menon'],
        [1312198, 'Kolkata', '2022-05-25', …,
         "['Q de Kock', 'KL Rahul', 'M Vohra', 'DJ Hooda', 'MP Stoinis', 'E
  Lewis', 'KH Pandya', 'PVD Chameera', 'Mohsin Khan', 'Avesh Khan', 'Ravi
  Bishnoi']",
         'J Madanagopal', 'MA Gough'],
        …,
        [335984, 'Delhi', '2008-04-19', …,
         "['T Kohli', 'YK Pathan', 'SR Watson', 'M Kaif', 'DS Lehmann', 'RA
  Jadeja', 'M Rawat', 'D Salunkhe', 'SK Warne', 'SK Trivedi', 'MM Patel']",
         'Aleem Dar', 'GA Pratapkumar'],
        [335983, 'Chandigarh', '2008-04-19', …,
         "['PA Patel', 'ML Hayden', 'MEK Hussey', 'MS Dhoni', 'SK Raina', 'JDP
  Oram', 'S Badrinath', 'Joginder Sharma', 'P Amarnath', 'MS Gony', 'M
  Muralitharan']",
         'MR Benson', 'SL Shastri'],
        [335982, 'Bangalore', '2008-04-18', …,
         "['SC Ganguly', 'BB McCullum', 'RT Ponting', 'DJ Hussey', 'Mohammad
  Hafeez', 'LR Shukla', 'WP Saha', 'AB Agarkar', 'AB Dinda', 'M Kartik', 'I
  Sharma']",
         'Asad Rauf', 'RE Koertzen']], dtype=object)
```

[113]: `movie.head(2) #head`

[113]:
```
                        title_x     imdb_id  \
  0  Uri: The Surgical Strike  tt8291224
  1            Battalion 609  tt9472208


                                         poster_path  \
  0  https://upload.wikimedia.org/wikipedia/en/thum…
  1                                              NaN


                                           wiki_link  \
  0  https://en.wikipedia.org/wiki/Uri:_The_Surgica…
  1        https://en.wikipedia.org/wiki/Battalion_609


                   title_y            original_title  is_adult  \
  0  Uri: The Surgical Strike  Uri: The Surgical Strike         0
  1            Battalion 609            Battalion 609         0

    year_of_release runtime            genres  imdb_rating  imdb_votes  \
  0            2019     138  Action|Drama|War          8.4       35112
  1            2019     131               War          4.1          73
```

```
                                      story  \
0  Divided over five chapters  the film chronicle…
1  The story revolves around a cricket match betw…


                                    summary tagline  \
0  Indian army special forces execute a covert op…     NaN
1  The story of Battalion 609 revolves around a c…     NaN


                                     actors wins_nominations  \
0  Vicky Kaushal|Paresh Rawal|Mohit Raina|Yami Ga…          4 wins
1  Vicky Ahuja|Shoaib Ibrahim|Shrikant Kamat|Elen…             NaN


            release_date
0    11 January 2019 (USA)
1  11 January 2019 (India)
```

[114]: `movie.tail(2) #tail`

[114]:
```
          title_x       imdb_id                                   poster_path  \
1627       Daaka  tt10833860  https://upload.wikimedia.org/wikipedia/en/thum…
1628    Humsafar   tt2403201  https://upload.wikimedia.org/wikipedia/en/thum…


                                  wiki_link   title_y original_title  \
1627       https://en.wikipedia.org/wiki/Daaka     Daaka          Daaka
1628  https://en.wikipedia.org/wiki/Humsafar  Humsafar       Humsafar


       is_adult  year_of_release runtime          genres  imdb_rating  \
1627          0             2019     136          Action          7.4
1628          0             2011      35   Drama|Romance          9.0


       imdb_votes                                            story  \
1627           38  Shinda tries robbing a bank so he can be wealt…
1628         2968  Sara and Ashar are childhood friends who share…


                                    summary tagline  \
1627  Shinda tries robbing a bank so he can be wealt…     NaN
1628  Ashar and Khirad are forced to get married due…     NaN


                         actors wins_nominations          release_date
1627  Gippy Grewal|Zareen Khan|              NaN  1 November 2019 (USA)
1628              Fawad Khan|              NaN     TV Series (2011-2012)
```

[115]: `movie.sample(3)  #Randomly give the data from the dataset`

[115]:
```
              title_x     imdb_id  \
1405    88 Antop Hill  tt0349058
1057      Apna Asmaan  tt1073097
```

```
349    Moh Maya Money   tt5343678

                                                           poster_path  \
1405  https://upload.wikimedia.org/wikipedia/en/thum…
1057  https://upload.wikimedia.org/wikipedia/en/thum…
349   https://upload.wikimedia.org/wikipedia/en/thum…

                                    wiki_link         title_y  \
1405   https://en.wikipedia.org/wiki/88_Antop_Hill    88 Antop Hill
1057       https://en.wikipedia.org/wiki/Apna_Asmaan      Apna Asmaan
349   https://en.wikipedia.org/wiki/Moh_Maya_Money  Moh Maya Money

      original_title  is_adult  year_of_release runtime            genres  \
1405    88 Antop Hill         0             2003     119    Mystery|Thriller
1057      Apna Asmaan         0             2007     110          Drama|Sci-Fi
349   Moh Maya Money         0             2016     108  Crime|Drama|Thriller

      imdb_rating  imdb_votes  \
1405          4.6          59
1057          6.2         144
349           5.7         344

                                                      story  \
1405  Antara Shelar lives an upper middle-class life…
1057  Ravi and Padmini share a strained relationship…
349   A young real-estate broker tries to pull off a…

                                                    summary  \
1405  A prank goes horribly wrong when the victim en…
1057  In order to make their autistic child a genius…
349   A young real-estate broker tries to pull off a…

                                                    tagline  \
1405                                                    NaN
1057  How far would you go to make your child a genius?
349                                                     NaN

                                          actors wins_nominations  \
1405  Atul Kulkarni|Rahul Dev|Suchitra Pillai|Shweta…              NaN
1057  Irrfan Khan|Shobana|Dhruv Piyush Panjuani|Raja…              NaN
349   Neha Dhupia|Ranvir Shorey|Devendra Chowhan|Vid…              NaN

              release_date
1405       27 June 2003 (India)
1057  7 September 2007 (India)
349   25 November 2016 (India)
```

```
[116]: movie.info()   #info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1629 entries, 0 to 1628
Data columns (total 18 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   title_x           1629 non-null   object
 1   imdb_id           1629 non-null   object
 2   poster_path       1526 non-null   object
 3   wiki_link         1629 non-null   object
 4   title_y           1629 non-null   object
 5   original_title    1629 non-null   object
 6   is_adult          1629 non-null   int64
 7   year_of_release   1629 non-null   int64
 8   runtime           1629 non-null   object
 9   genres            1629 non-null   object
 10  imdb_rating       1629 non-null   float64
 11  imdb_votes        1629 non-null   int64
 12  story             1609 non-null   object
 13  summary           1629 non-null   object
 14  tagline           557 non-null    object
 15  actors            1624 non-null   object
 16  wins_nominations  707 non-null    object
 17  release_date      1522 non-null   object
dtypes: float64(1), int64(3), object(14)
memory usage: 229.2+ KB
```

```
[117]: ipl.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 950 entries, 0 to 949
Data columns (total 20 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   ID            950 non-null    int64
 1   City          899 non-null    object
 2   Date          950 non-null    object
 3   Season        950 non-null    object
 4   MatchNumber   950 non-null    object
 5   Team1         950 non-null    object
 6   Team2         950 non-null    object
 7   Venue         950 non-null    object
 8   TossWinner    950 non-null    object
 9   TossDecision  950 non-null    object
 10  SuperOver     946 non-null    object
 11  WinningTeam   946 non-null    object
 12  WonBy         950 non-null    object
```

```
13   Margin           932 non-null    float64
14   method            19 non-null    object
15   Player_of_Match  946 non-null    object
16   Team1Players     950 non-null    object
17   Team2Players     950 non-null    object
18   Umpire1          950 non-null    object
19   Umpire2          950 non-null    object
dtypes: float64(1), int64(1), object(18)
memory usage: 148.6+ KB
```

[118]: `#describe`
`movie.describe()` *#this function give the mathematicla summary about the data*

[118]:

	is_adult	year_of_release	imdb_rating	imdb_votes
count	1629.0	1629.000000	1629.000000	1629.000000
mean	0.0	2010.263966	5.557459	5384.263352
std	0.0	5.381542	1.567609	14552.103231
min	0.0	2001.000000	0.000000	0.000000
25%	0.0	2005.000000	4.400000	233.000000
50%	0.0	2011.000000	5.600000	1000.000000
75%	0.0	2015.000000	6.800000	4287.000000
max	0.0	2019.000000	9.400000	310481.000000

[119]: `#isnull`
`movie.isnull()`

[119]:

	title_x	imdb_id	poster_path	wiki_link	title_y	original_title \
0	False	False	False	False	False	False
1	False	False	True	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	True	False	False	False
...	...	...	...	...	...	...
1624	False	False	False	False	False	False
1625	False	False	False	False	False	False
1626	False	False	True	False	False	False
1627	False	False	False	False	False	False
1628	False	False	False	False	False	False

	is_adult	year_of_release	runtime	genres	imdb_rating	imdb_votes \
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	False	False	False	False
...	...	...	...	...	...	...
1624	False	False	False	False	False	False

```
1625     False              False     False     False          False          False
1626     False              False     False     False          False          False
1627     False              False     False     False          False          False
1628     False              False     False     False          False          False
```

```
        story   summary   tagline   actors   wins_nominations   release_date
0       False    False      True    False                False          False
1       False    False      True    False                 True          False
2       False    False      True    False                 True          False
3       False    False      True    False                 True          False
4       False    False      True    False                False          False
...       ...      ...       ...      ...                  ...            ...
1624    False    False      True    False                 True          False
1625    False    False      True    False                 True          False
1626    False    False      True    False                 True           True
1627    False    False      True    False                 True          False
1628    False    False      True    False                 True          False
```

[1629 rows x 18 columns]

[120]: `movie.isnull().sum()`

[120]:
```
title_x              0
imdb_id              0
poster_path        103
wiki_link            0
title_y              0
original_title       0
is_adult             0
year_of_release      0
runtime              0
genres               0
imdb_rating          0
imdb_votes           0
story               20
summary              0
tagline           1072
actors               5
wins_nominations   922
release_date       107
dtype: int64
```

[121]:
```
#duplicated
movie.duplicated()
```

[121]:
```
0        False
1        False
```

14

```
2        False
3        False
4        False
         …
1624     False
1625     False
1626     False
1627     False
1628     False
Length: 1629, dtype: bool
```

[122]: `movie.duplicated().sum()`

[122]: 0

[123]: `stu`

[123]:
```
    id  marks  package
0  100     55       14
1   90     66       11
2  120     93       13
3   80     78       15
4    0      0        0
5    0      0        0
```

[124]: `stu.duplicated()`

[124]:
```
0    False
1    False
2    False
3    False
4    False
5     True
dtype: bool
```

[125]: `stu.duplicated().sum()`

[125]: 1

[126]:
```
#rename
stu
```

[126]:
```
    id  marks  package
0  100     55       14
1   90     66       11
2  120     93       13
3   80     78       15
```

```
4     0        0           0
5     0        0           0
```

[127]: `stu.rename(columns={"marks":"Percentage","package":"LPA"}) #the rename function`␣
 ↪`is not permanant changes in the data set`
                                                                  `#if you want to`␣
 ↪`change permanant so you can use "inplace = True" parameter`

[127]:
```
     id  Percentage  LPA
0   100          55   14
1    90          66   11
2   120          93   13
3    80          78   15
4     0           0    0
5     0           0    0
```

[128]: `stu.rename(columns = {"marks":"percentage","package":"LPA"},inplace=True)` ␣
 ↪`#inplace = True for permanent changes in the dataframe`

[129]: `stu`

[129]:
```
     id  percentage  LPA
0   100          55   14
1    90          66   11
2   120          93   13
3    80          78   15
4     0           0    0
5     0           0    0
```

[130]: `stu.rename({"percentage":"Percentage"})`

[130]:
```
     id  percentage  LPA
0   100          55   14
1    90          66   11
2   120          93   13
3    80          78   15
4     0           0    0
5     0           0    0
```

3   Mathematical Function on the dataframe

[131]: `student=pd.DataFrame(student)`
`student`

[131]:
```
     id  marks  package
0   100     55       14
```

```
1    90      66          11
2   120      93          13
3    80      78          15
4     0       0           0
5     0       0           0
```

[133]: `student.sum() # sum, the default sum is sum of the column-wise`

[133]:
```
id         390
marks      292
package     53
dtype: int64
```

[135]: `student.sum(axis=1) #axis is parameter, that sum of the row-wise`

[135]:
```
0    169
1    167
2    226
3    173
4      0
5      0
dtype: int64
```

[136]: `student.mean()   #mean`

[136]:
```
id         65.000000
marks      48.666667
package     8.833333
dtype: float64
```

[138]: `student.mean(axis=1) # axis parameter mean of row-wise`

[138]:
```
0    56.333333
1    55.666667
2    75.333333
3    57.666667
4     0.000000
5     0.000000
dtype: float64
```

[144]:
```
#median
student.median()
```

[144]:
```
id         85.0
marks      60.5
package    12.0
dtype: float64
```

```
[146]: student.std()    #std
```

```
[146]: id          52.057660
       marks       39.757599
       package      6.968979
       dtype: float64
```

4   Selecting a Column from the DataFrame

```
[149]: movie['imdb_id']
```

```
[149]: 0          tt8291224
       1          tt9472208
       2          tt6986710
       3          tt8108208
       4          tt6028796
                    …
       1624       tt0301250
       1625       tt0298607
       1626       tt0069204
       1627      tt10833860
       1628       tt2403201
       Name: imdb_id, Length: 1629, dtype: object
```

```
[151]: ipl['Venue']
```

```
[151]: 0              Narendra Modi Stadium, Ahmedabad
       1              Narendra Modi Stadium, Ahmedabad
       2                       Eden Gardens, Kolkata
       3                       Eden Gardens, Kolkata
       4                    Wankhede Stadium, Mumbai
                                  …
       945                              Eden Gardens
       946                          Wankhede Stadium
       947                          Feroz Shah Kotla
       948   Punjab Cricket Association Stadium, Mohali
       949                     M Chinnaswamy Stadium
       Name: Venue, Length: 950, dtype: object
```

```
[153]: student['package']
```

```
[153]: 0    14
       1    11
       2    13
       3    15
       4     0
```

```
5      0
Name: package, dtype: int64
```

Selecting a multipule column

```
[159]: movie[['title_y','title_x','original_title']]
```

```
[159]:                        title_y                         title_x  \
       0            Uri: The Surgical Strike         Uri: The Surgical Strike
       1                       Battalion 609                   Battalion 609
       2         The Accidental Prime Minister   The Accidental Prime Minister (film)
       3                      Why Cheat India                 Why Cheat India
       4                      Evening Shadows                 Evening Shadows
       ...                              ...                             ...
       1624            Tera Mera Saath Rahen           Tera Mera Saath Rahen
       1625            Yeh Zindagi Ka Safar           Yeh Zindagi Ka Safar
       1626                 Sabse Bada Sukh                 Sabse Bada Sukh
       1627                            Daaka                           Daaka
       1628                         Humsafar                        Humsafar

                          original_title
       0            Uri: The Surgical Strike
       1                       Battalion 609
       2         The Accidental Prime Minister
       3                      Why Cheat India
       4                      Evening Shadows
       ...                              ...
       1624            Tera Mera Saath Rahen
       1625            Yeh Zindagi Ka Safar
       1626                 Sabse Bada Sukh
       1627                            Daaka
       1628                         Humsafar

       [1629 rows x 3 columns]
```

```
[162]: type(movie[['title_y','title_x','original_title']]) #when we fatch the multiple␣
        ↪column then it comes "DataFrame" type
```

```
[162]: pandas.core.frame.DataFrame
```

```
[163]: type(ipl['Venue'])   #when we fatch the single column then it comes "Series" type
```

```
[163]: pandas.core.series.Series
```

```
[168]: ipl[['Team1',"WinningTeam"]]
```

```
[168]:                         Team1                    WinningTeam
     0             Rajasthan Royals                 Gujarat Titans
     1    Royal Challengers Bangalore              Rajasthan Royals
     2    Royal Challengers Bangalore  Royal Challengers Bangalore
     3             Rajasthan Royals                 Gujarat Titans
     4           Sunrisers Hyderabad                   Punjab Kings
     ..                         …                             …
     945        Kolkata Knight Riders         Kolkata Knight Riders
     946              Mumbai Indians   Royal Challengers Bangalore
     947             Delhi Daredevils              Delhi Daredevils
     948              Kings XI Punjab          Chennai Super Kings
     949  Royal Challengers Bangalore        Kolkata Knight Riders

     [950 rows x 2 columns]
```

5 How to fecth Row in the DataFrame

iloc fuction work on index postion

loc fnction work on index label when we provide

```
[179]: movie.iloc[0]
```

```
[179]: title_x                              Uri: The Surgical Strike
       imdb_id                                          tt8291224
       poster_path        https://upload.wikimedia.org/wikipedia/en/thum…
       wiki_link          https://en.wikipedia.org/wiki/Uri:_The_Surgica…
       title_y                              Uri: The Surgical Strike
       original_title                       Uri: The Surgical Strike
       is_adult                                                    0
       year_of_release                                          2019
       runtime                                                   138
       genres                                       Action|Drama|War
       imdb_rating                                               8.4
       imdb_votes                                              35112
       story              Divided over five chapters  the film chronicle…
       summary            Indian army special forces execute a covert op…
       tagline                                                   NaN
       actors             Vicky Kaushal|Paresh Rawal|Mohit Raina|Yami Ga…
       wins_nominations                                       4 wins
       release_date                         11 January 2019 (USA)
       Name: 0, dtype: object
```

```
[181]: type(movie.iloc[2])
```

```
[181]: pandas.core.series.Series
```

```
[183]: movie.iloc[1:10:2]
```

```
[183]:              title_x      imdb_id  \
       1        Battalion 609  tt9472208
       3      Why Cheat India  tt8108208
       5          Soni (film)  tt6078866
       7           Bombairiya  tt4971258
       9      Thackeray (film)  tt7777196


                                        poster_path  \
       1                                        NaN
       3  https://upload.wikimedia.org/wikipedia/en/thum…
       5  https://upload.wikimedia.org/wikipedia/en/thum…
       7  https://upload.wikimedia.org/wikipedia/en/thum…
       9  https://upload.wikimedia.org/wikipedia/en/thum…


                                        wiki_link          title_y  \
       1      https://en.wikipedia.org/wiki/Battalion_609    Battalion 609
       3    https://en.wikipedia.org/wiki/Why_Cheat_India  Why Cheat India
       5        https://en.wikipedia.org/wiki/Soni_(film)             Soni
       7          https://en.wikipedia.org/wiki/Bombairiya       Bombairiya
       9  https://en.wikipedia.org/wiki/Thackeray_(film)        Thackeray


          original_title  is_adult  year_of_release runtime              genres  \
       1   Battalion 609         0             2019     131                 War
       3  Why Cheat India        0             2019     121          Crime|Drama
       5            Soni         0             2018      97                Drama
       7      Bombairiya         0             2019     104  Comedy|Crime|Drama
       9       Thackeray         0             2019     120      Biography|Drama


          imdb_rating  imdb_votes                                          story  \
       1          4.1          73  The story revolves around a cricket match betw…
       3          6.0        1891  The movie focuses on existing malpractices in …
       5          7.2        1595  Soni  a young policewoman in Delhi  and her su…
       7          4.3         295  It follows the story of Meghna  who gets embro…
       9          5.1        2301  Balasaheb Thackrey works as a cartoonist for a…


                                           summary  \
       1  The story of Battalion 609 revolves around a c…
       3  The movie focuses on existing malpractices in …
       5  While fighting crimes against women in Delhi  …
       7  It follows the story of Meghna  who gets embro…
       9  Biographical account of Shiv Sena Supremo  Bal…


                                           tagline  \
       1                                        NaN
       3                                        NaN
```

```
5                                                   NaN
7   They didn't mean to change the world.
9                                                   NaN


                                            actors       wins_nominations  \
1   Vicky Ahuja|Shoaib Ibrahim|Shrikant Kamat|Elen…                  NaN
3   Emraan Hashmi|Shreya Dhanwanthary|Snighdadeep …                  NaN
5   Geetika Vidya Ohlyan|Saloni Batra|Vikas Shukla…  3 wins & 5 nominations
7   Radhika Apte|Akshay Oberoi|Siddhanth Kapoor|Ra…                  NaN
9   Nawazuddin Siddiqui|Amrita Rao|Abdul Quadir Am…                  NaN


             release_date
1   11 January 2019 (India)
3     18 January 2019 (USA)
5     18 January 2019 (USA)
7   18 January 2019 (India)
9   25 January 2019 (India)
```

[184]: `movie.iloc[:3]`

```
[184]:                           title_x    imdb_id  \
       0         Uri: The Surgical Strike  tt8291224
       1                    Battalion 609  tt9472208
       2   The Accidental Prime Minister (film)  tt6986710


                                        poster_path  \
       0   https://upload.wikimedia.org/wikipedia/en/thum…
       1                                           NaN
       2   https://upload.wikimedia.org/wikipedia/en/thum…


                                          wiki_link  \
       0   https://en.wikipedia.org/wiki/Uri:_The_Surgica…
       1       https://en.wikipedia.org/wiki/Battalion_609
       2   https://en.wikipedia.org/wiki/The_Accidental_P…


                            title_y                original_title  is_adult  \
       0      Uri: The Surgical Strike      Uri: The Surgical Strike         0
       1                 Battalion 609                 Battalion 609         0
       2   The Accidental Prime Minister  The Accidental Prime Minister         0

           year_of_release runtime          genres  imdb_rating  imdb_votes  \
       0             2019     138  Action|Drama|War          8.4       35112
       1             2019     131               War          4.1          73
       2             2019     112  Biography|Drama           6.1        5549


                                              story  \
       0   Divided over five chapters  the film chronicle…
```

```
1  The story revolves around a cricket match betw…
2  Based on the memoir by Indian policy analyst S…

                                         summary tagline  \
0  Indian army special forces execute a covert op…    NaN
1  The story of Battalion 609 revolves around a c…    NaN
2  Explores Manmohan Singh's tenure as the Prime …    NaN

                                          actors wins_nominations  \
0  Vicky Kaushal|Paresh Rawal|Mohit Raina|Yami Ga…          4 wins
1  Vicky Ahuja|Shoaib Ibrahim|Shrikant Kamat|Elen…             NaN
2  Anupam Kher|Akshaye Khanna|Aahana Kumra|Atul S…             NaN

            release_date
0    11 January 2019 (USA)
1  11 January 2019 (India)
2    11 January 2019 (USA)
```

[186]: `movie.iloc[[0,1,4,6,8]]`

[186]:
```
                            title_x      imdb_id  \
0          Uri: The Surgical Strike  tt8291224
1                    Battalion 609  tt9472208
4                   Evening Shadows  tt6028796
6                     Fraud Saiyaan  tt5013008
8  Manikarnika: The Queen of Jhansi  tt6903440

                                         poster_path  \
0  https://upload.wikimedia.org/wikipedia/en/thum…
1                                                NaN
4                                                NaN
6  https://upload.wikimedia.org/wikipedia/en/thum…
8  https://upload.wikimedia.org/wikipedia/en/thum…

                                          wiki_link  \
0  https://en.wikipedia.org/wiki/Uri:_The_Surgica…
1        https://en.wikipedia.org/wiki/Battalion_609
4      https://en.wikipedia.org/wiki/Evening_Shadows
6        https://en.wikipedia.org/wiki/Fraud_Saiyaan
8  https://en.wikipedia.org/wiki/Manikarnika:_The…

                            title_y                   original_title  \
0          Uri: The Surgical Strike          Uri: The Surgical Strike
1                    Battalion 609                    Battalion 609
4                   Evening Shadows                   Evening Shadows
6                     Fraud Saiyaan                     Fraud Saiyyan
8  Manikarnika: The Queen of Jhansi  Manikarnika: The Queen of Jhansi
```

```
   is_adult  year_of_release runtime                    genres  imdb_rating  \
0         0             2019     138           Action|Drama|War          8.4
1         0             2019     131                        War          4.1
4         0             2018     102                      Drama          7.3
6         0             2019     109               Comedy|Drama          4.2
8         0             2019     148    Action|Biography|Drama          6.5

   imdb_votes                                           story  \
0       35112  Divided over five chapters  the film chronicle…
1          73  The story revolves around a cricket match betw…
4         280  While gay rights and marriage equality has bee…
6         504  Fraud Saiyyan is the story of a con artist in …
8        7361  Manikarnika born in Varanasi when Dixt a minis…

                                        summary tagline  \
0  Indian army special forces execute a covert op…     NaN
1  The story of Battalion 609 revolves around a c…     NaN
4  Under the 'Evening Shadows'  truth often plays…     NaN
6  Fraud Saiyyan is the story of a con artist in …     NaN
8  Story of Rani Lakshmibai  one of the leading f…     NaN

                                        actors         wins_nominations  \
0  Vicky Kaushal|Paresh Rawal|Mohit Raina|Yami Ga…                4 wins
1  Vicky Ahuja|Shoaib Ibrahim|Shrikant Kamat|Elen…                   NaN
4  Mona Ambegaonkar|Ananth Narayan Mahadevan|Deva…  17 wins & 1 nomination
6  Arshad Warsi|Saurabh Shukla|Flora Saini|Sara L…                   NaN
8  Kangana Ranaut|Rimi Sen|Atul Kulkarni|Nalneesh…                   NaN

             release_date
0    11 January 2019 (USA)
1  11 January 2019 (India)
4  11 January 2019 (India)
6  18 January 2019 (India)
8     25 January 2019 (USA)
```

loc function

```
[202]:  stu
```

```
[202]:      Name   id  marks  package
        0   rohit  100     55       14
        1   niddy   90     66       11
        2  pushpa  120     93       13
        3   ankit   80     78       15
        4  piyush    0      0        0
        5  vishal    0      0        0
```

```
[207]: stu.set_index("Name",inplace=True) #in this we set the name index by using␣
       ↪"set_index" method
```

```
[208]: stu
```

```
[208]:         id  marks  package
       Name
       rohit   100     55       14
       niddy    90     66       11
       pushpa  120     93       13
       ankit    80     78       15
       piyush    0      0        0
       vishal    0      0        0
```

```
[209]: stu.loc['rohit']   #Selecting a perticular row
```

```
[209]: id         100
       marks       55
       package     14
       Name: rohit, dtype: int64
```

```
[212]: stu.loc[['pushpa','niddy']]   #fetch multiple row
```

```
[212]:         id  marks  package
       Name
       pushpa  120     93       13
       niddy    90     66       11
```

```
[214]: movie.loc[3:4]
```

```
[214]:             title_x     imdb_id  \
       3  Why Cheat India  tt8108208
       4  Evening Shadows  tt6028796

                                             poster_path  \
       3  https://upload.wikimedia.org/wikipedia/en/thum…
       4                                             NaN

                                            wiki_link         title_y  \
       3  https://en.wikipedia.org/wiki/Why_Cheat_India  Why Cheat India
       4  https://en.wikipedia.org/wiki/Evening_Shadows  Evening Shadows

           original_title  is_adult  year_of_release runtime      genres  \
       3  Why Cheat India         0             2019     121  Crime|Drama
       4  Evening Shadows         0             2018     102        Drama

          imdb_rating  imdb_votes                                         story  \
```

```
3           6.0          1891   The movie focuses on existing malpractices in …
4           7.3           280   While gay rights and marriage equality has bee…

                                    summary tagline  \
3  The movie focuses on existing malpractices in …      NaN
4  Under the 'Evening Shadows'   truth often plays…      NaN


                                    actors         wins_nominations  \
3  Emraan Hashmi|Shreya Dhanwanthary|Snighdadeep …                      NaN
4  Mona Ambegaonkar|Ananth Narayan Mahadevan|Deva…  17 wins & 1 nomination

           release_date
3     18 January 2019 (USA)
4  11 January 2019 (India)
```

[215]:
```
stu
```

[215]:
```
         id  marks  package
Name
rohit   100     55       14
niddy    90     66       11
pushpa  120     93       13
ankit    80     78       15
piyush    0      0        0
vishal    0      0        0
```

[216]:
```
stu.loc['rohit':"piyush":2]
```

[216]:
```
         id  marks  package
Name
rohit   100     55       14
pushpa  120     93       13
piyush    0      0        0
```

[219]:
```
stu.iloc[0]     #we can also use the iloc function in this dataframe
```

[219]:
```
id         100
marks       55
package     14
Name: rohit, dtype: int64
```

6   SElecting Both Row And Column

[225]:
```
movie.iloc[0:3,0:3]
```

```
[225]:                          title_x      imdb_id  \
       0             Uri: The Surgical Strike  tt8291224
       1                       Battalion 609  tt9472208
       2  The Accidental Prime Minister (film)  tt6986710


                                               poster_path
       0  https://upload.wikimedia.org/wikipedia/en/thum…
       1                                              NaN
       2  https://upload.wikimedia.org/wikipedia/en/thum…
```

```
[229]: movie.loc[0:2,"title_x":"poster_path"]   #[Row(Start:End),Column(Start:End)] loc␣
       ↪function work with labels
```

```
[229]:                          title_x      imdb_id  \
       0             Uri: The Surgical Strike  tt8291224
       1                       Battalion 609  tt9472208
       2  The Accidental Prime Minister (film)  tt6986710


                                               poster_path
       0  https://upload.wikimedia.org/wikipedia/en/thum…
       1                                              NaN
       2  https://upload.wikimedia.org/wikipedia/en/thum…
```

7 Filtering a DataFrame

Some Question on the filternig data frame 1. find all the final winner

```
[244]: ipl.head(2)
```

```
[244]:          ID       City        Date  Season   MatchNumber  \
       0  1312200  Ahmedabad  2022-05-29    2022         Final
       1  1312199  Ahmedabad  2022-05-27    2022   Qualifier 2


                            Team1               Team2  \
       0            Rajasthan Royals      Gujarat Titans
       1  Royal Challengers Bangalore  Rajasthan Royals


                                   Venue        TossWinner TossDecision SuperOver  \
       0  Narendra Modi Stadium, Ahmedabad  Rajasthan Royals          bat         N
       1  Narendra Modi Stadium, Ahmedabad  Rajasthan Royals        field         N


              WinningTeam     WonBy  Margin method Player_of_Match  \
       0    Gujarat Titans  Wickets     7.0    NaN       HH Pandya
       1  Rajasthan Royals  Wickets     7.0    NaN       JC Buttler


                                          Team1Players  \
       0  ['YBK Jaiswal', 'JC Buttler', 'SV Samson', 'D …
```

```
1  ['V Kohli', 'F du Plessis', 'RM Patidar', 'GJ …
```

```
                                        Team2Players        Umpire1      Umpire2
0  ['WP Saha', 'Shubman Gill', 'MS Wade', 'HH Pan…  CB Gaffaney  Nitin Menon
1  ['YBK Jaiswal', 'JC Buttler', 'SV Samson', 'D …  CB Gaffaney  Nitin Menon
```

[256]: `#find all the final winner`
`match=ipl['MatchNumber']=='Final'`
`new_df=ipl[match]`
`new_df[['Season','WinningTeam']].value_counts()`

[256]:
```
Season   WinningTeam
2007/08  Rajasthan Royals       1
2009     Deccan Chargers        1
2009/10  Chennai Super Kings    1
2011     Chennai Super Kings    1
2012     Kolkata Knight Riders  1
2013     Mumbai Indians         1
2014     Kolkata Knight Riders  1
2015     Mumbai Indians         1
2016     Sunrisers Hyderabad    1
2017     Mumbai Indians         1
2018     Chennai Super Kings    1
2019     Mumbai Indians         1
2020/21  Mumbai Indians         1
2021     Chennai Super Kings    1
2022     Gujarat Titans         1
dtype: int64
```

[266]: `from pandas._libs.hashtable import value_count`
`ipl[ipl['MatchNumber']=='Final'][["Season","WinningTeam"]]`

[266]:
```
       Season           WinningTeam
0        2022         Gujarat Titans
74       2021    Chennai Super Kings
134   2020/21         Mumbai Indians
194      2019         Mumbai Indians
254      2018    Chennai Super Kings
314      2017         Mumbai Indians
373      2016    Sunrisers Hyderabad
433      2015         Mumbai Indians
492      2014  Kolkata Knight Riders
552      2013         Mumbai Indians
628      2012  Kolkata Knight Riders
702      2011    Chennai Super Kings
775   2009/10    Chennai Super Kings
835      2009        Deccan Chargers
```

```
    892   2007/08        Rajasthan Royals
```

```
[272]: # how many super over finishes have occured
       ipl[ipl['SuperOver'] == "Y"].shape[0]
```

[272]: 14

```
[291]: # how many matches has csk won in kolkata
       ipl[(ipl["City"]== "Kolkata") & (ipl['WinningTeam']=='Chennai Super Kings')].
        ↪shape[0]
```

[291]: 5

```
[298]: #toss winner is match winner in percentage
       ipl[ipl["TossWinner"] == ipl['WinningTeam']].shape[0]/ipl.shape[0]*100
```

[298]: 51.473684210526315

```
[309]: # movie with raing than 8 and votes > 10000
       movie[(movie["imdb_rating"] > 8) & (movie["imdb_votes"] > 10000)].shape[0]
```

[309]: 43

Action movie with rating higher than 7.5

```
[315]: #Action movie with rating higher than 7.5
       mov1 = movie['genres'].str.split('|').apply(lambda x: 'Action' in x)
       mov2 = movie['imdb_rating'] > 7.5
       movie[mov1 & mov2].shape
```

[315]: (33, 18)

```
[316]: mov1 = movie['genres'].str.contains('Action')
       mov2 = movie['imdb_rating'] > 7.5
       movie[mov1 & mov2].shape
```

[316]: (33, 18)

8 Add New Column

```
[317]: movie["Camera"] = "Imax"
       movie
```

```
[317]:                              title_x     imdb_id  \
       0                Uri: The Surgical Strike    tt8291224
       1                           Battalion 609    tt9472208
       2       The Accidental Prime Minister (film)    tt6986710
```

```
3                           Why Cheat India    tt8108208
4                           Evening Shadows    tt6028796
…                                       …              …
1624               Tera Mera Saath Rahen    tt0301250
1625               Yeh Zindagi Ka Safar    tt0298607
1626                     Sabse Bada Sukh    tt0069204
1627                               Daaka  tt10833860
1628                             Humsafar    tt2403201


                                              poster_path  \
0     https://upload.wikimedia.org/wikipedia/en/thum…
1                                                   NaN
2     https://upload.wikimedia.org/wikipedia/en/thum…
3     https://upload.wikimedia.org/wikipedia/en/thum…
4                                                   NaN
…                                                     …
1624  https://upload.wikimedia.org/wikipedia/en/2/2b…
1625  https://upload.wikimedia.org/wikipedia/en/thum…
1626                                                NaN
1627  https://upload.wikimedia.org/wikipedia/en/thum…
1628  https://upload.wikimedia.org/wikipedia/en/thum…


                                                wiki_link  \
0     https://en.wikipedia.org/wiki/Uri:_The_Surgica…
1           https://en.wikipedia.org/wiki/Battalion_609
2     https://en.wikipedia.org/wiki/The_Accidental_P…
3          https://en.wikipedia.org/wiki/Why_Cheat_India
4          https://en.wikipedia.org/wiki/Evening_Shadows
…                                                       …
1624  https://en.wikipedia.org/wiki/Tera_Mera_Saath_…
1625  https://en.wikipedia.org/wiki/Yeh_Zindagi_Ka_S…
1626      https://en.wikipedia.org/wiki/Sabse_Bada_Sukh
1627                https://en.wikipedia.org/wiki/Daaka
1628              https://en.wikipedia.org/wiki/Humsafar


                            title_y                  original_title  is_adult  \
0            Uri: The Surgical Strike      Uri: The Surgical Strike         0
1                       Battalion 609                 Battalion 609         0
2     The Accidental Prime Minister  The Accidental Prime Minister         0
3                     Why Cheat India               Why Cheat India         0
4                     Evening Shadows               Evening Shadows         0
…                                 …                             …          …
1624            Tera Mera Saath Rahen         Tera Mera Saath Rahen         0
1625             Yeh Zindagi Ka Safar          Yeh Zindagi Ka Safar         0
1626                   Sabse Bada Sukh               Sabse Bada Sukh         0
1627                             Daaka                         Daaka         0
1628                           Humsafar                     Humsafar         0
```

```
      year_of_release runtime            genres  imdb_rating  imdb_votes  \
0                2019     138   Action|Drama|War          8.4       35112
1                2019     131                War          4.1          73
2                2019     112    Biography|Drama          6.1        5549
3                2019     121        Crime|Drama          6.0        1891
4                2018     102              Drama          7.3         280
...               ...     ...                ...          ...         ...
1624             2001     148              Drama          4.9         278
1625             2001     146              Drama          3.0         133
1626             2018      \N       Comedy|Drama          6.1          13
1627             2019     136             Action          7.4          38
1628             2011      35      Drama|Romance          9.0        2968

                                                  story  \
0        Divided over five chapters  the film chronicle…
1        The story revolves around a cricket match betw…
2        Based on the memoir by Indian policy analyst S…
3        The movie focuses on existing malpractices in …
4        While gay rights and marriage equality has bee…
...                                                   …
1624     Raj Dixit lives with his younger brother  Rahu…
1625     Hindi pop-star  Sarina Devan  lives a wealthy …
1626     Village born Lalloo re-locates to Bombay  and …
1627     Shinda tries robbing a bank so he can be wealt…
1628     Sara and Ashar are childhood friends who share…

                                                summary tagline  \
0        Indian army special forces execute a covert op…     NaN
1        The story of Battalion 609 revolves around a c…     NaN
2        Explores Manmohan Singh's tenure as the Prime …     NaN
3        The movie focuses on existing malpractices in …     NaN
4        Under the 'Evening Shadows'  truth often plays…     NaN
...                                                   …       …
1624     A man is torn between his handicapped brother …     NaN
1625     A singer finds out she was adopted when the ed…     NaN
1626     Village born Lalloo re-locates to Bombay  and …     NaN
1627     Shinda tries robbing a bank so he can be wealt…     NaN
1628     Ashar and Khirad are forced to get married due…     NaN

                                                 actors  \
0        Vicky Kaushal|Paresh Rawal|Mohit Raina|Yami Ga…
1        Vicky Ahuja|Shoaib Ibrahim|Shrikant Kamat|Elen…
2        Anupam Kher|Akshaye Khanna|Aahana Kumra|Atul S…
3        Emraan Hashmi|Shreya Dhanwanthary|Snighdadeep …
4        Mona Ambegaonkar|Ananth Narayan Mahadevan|Deva…
...                                                   …
```

```
1624   Ajay Devgn|Sonali Bendre|Namrata Shirodkar|Pre…
1625   Ameesha Patel|Jimmy Sheirgill|Nafisa Ali|Gulsh…
1626   Vijay Arora|Asrani|Rajni Bala|Kumud Damle|Utpa…
1627                              Gippy Grewal|Zareen Khan|
1628                                        Fawad Khan|

          wins_nominations            release_date Camera
0                  4 wins     11 January 2019 (USA)   Imax
1                     NaN   11 January 2019 (India)   Imax
2                     NaN     11 January 2019 (USA)   Imax
3                     NaN     18 January 2019 (USA)   Imax
4     17 wins & 1 nomination  11 January 2019 (India)  Imax
…                     …                        …      …
1624                  NaN    7 November 2001 (India)   Imax
1625                  NaN   16 November 2001 (India)   Imax
1626                  NaN                        NaN   Imax
1627                  NaN     1 November 2019 (USA)   Imax
1628                  NaN      TV Series (2011-2012)   Imax

[1629 rows x 19 columns]
```

[319]: ` movie.info() `

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1629 entries, 0 to 1628
Data columns (total 19 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   title_x           1629 non-null   object
 1   imdb_id           1629 non-null   object
 2   poster_path       1526 non-null   object
 3   wiki_link         1629 non-null   object
 4   title_y           1629 non-null   object
 5   original_title    1629 non-null   object
 6   is_adult          1629 non-null   int64
 7   year_of_release   1629 non-null   int64
 8   runtime           1629 non-null   object
 9   genres            1629 non-null   object
 10  imdb_rating       1629 non-null   float64
 11  imdb_votes        1629 non-null   int64
 12  story             1609 non-null   object
 13  summary           1629 non-null   object
 14  tagline           557 non-null    object
 15  actors            1624 non-null   object
 16  wins_nominations  707 non-null    object
 17  release_date      1522 non-null   object
 18  Camera            1629 non-null   object
```

```
dtypes: float64(1), int64(3), object(15)
memory usage: 241.9+ KB
```

[326]:
```
movie1=movie.dropna()
movie1
```

[326]:

	title_x	imdb_id
11	Gully Boy	tt2395469
34	Yeh Hai India	tt5525846
37	Article 15 (film)	tt10324144
87	Aiyaary	tt6774212
96	Raid (2018 film)	tt7363076
…	…	…
1600	Kasoor	tt0272688
1601	Maya (2001 film)	tt0299027
1607	Nayak (2001 Hindi film)	tt0291376
1621	Tum Bin	tt0290326
1623	Zubeidaa	tt0255713

	poster_path
11	https://upload.wikimedia.org/wikipedia/en/thum…
34	https://upload.wikimedia.org/wikipedia/en/thum…
37	https://upload.wikimedia.org/wikipedia/en/thum…
87	https://upload.wikimedia.org/wikipedia/en/thum…
96	https://upload.wikimedia.org/wikipedia/en/thum…
…	…
1600	https://upload.wikimedia.org/wikipedia/en/thum…
1601	https://upload.wikimedia.org/wikipedia/en/thum…
1607	https://upload.wikimedia.org/wikipedia/en/thum…
1621	https://upload.wikimedia.org/wikipedia/en/thum…
1623	https://upload.wikimedia.org/wikipedia/en/thum…

	wiki_link
11	https://en.wikipedia.org/wiki/Gully_Boy
34	https://en.wikipedia.org/wiki/Yeh_Hai_India
37	https://en.wikipedia.org/wiki/Article_15_(film)
87	https://en.wikipedia.org/wiki/Aiyaary
96	https://en.wikipedia.org/wiki/Raid_(2018_film)
…	…
1600	https://en.wikipedia.org/wiki/Kasoor
1601	https://en.wikipedia.org/wiki/Maya_(2001_film)
1607	https://en.wikipedia.org/wiki/Nayak_(2001_Hind…
1621	https://en.wikipedia.org/wiki/Tum_Bin
1623	https://en.wikipedia.org/wiki/Zubeidaa

	title_y	original_title
11	Gully Boy	Gully Boy

```
34                       Yeh Hai India                       Yeh Hai India
37                          Article 15                          Article 15
87                             Aiyaary                             Aiyaary
96                                Raid                                Raid
…                                   …                                   …
1600                              Kasoor                              Kasoor
1601                                Maya                                Maya
1607                 Nayak: The Real Hero                 Nayak: The Real Hero
1621   Tum Bin…: Love Will Find a Way  Tum Bin…: Love Will Find a Way
1623                             Zubeidaa                             Zubeidaa


       is_adult  year_of_release  runtime                    genres  imdb_rating  \
11            0             2019      153              Drama|Music          8.2
34            0             2017      128     Action|Adventure|Drama          5.7
37            0             2019      130               Crime|Drama          8.3
87            0             2018      157            Action|Thriller          5.2
96            0             2018      122         Action|Crime|Drama          7.4
…             …                …        …                         …            …
1600          0             2001      151     Drama|Mystery|Thriller          6.0
1601          0             2001      105                     Drama          7.0
1607          0             2001      187      Action|Drama|Thriller          7.8
1621          0             2001      158             Drama|Romance          7.5
1623          0             2001      153  Biography|Drama|History          6.2


       imdb_votes                                              story  \
11          22440  Gully Boy is a film about a 22-year-old boy "M…
34            169  Yeh Hai India  follows the story of a 25 years…
37          13417  In the rural heartlands of India  an upright p…
87           3538  General Gurinder Singh comes with a proposal t…
96          13159  Set in the 80s in Uttar Pradesh  India  Raid i…
…               …                                                 …
1600          801  Shekar Saxena is arrested and charged with the…
1601          277  For 12 year-old Maya  life is very simple in h…
1607        12522  Employed as a camera-man at a popular televisi…
1621         2879  Shekar and Amar meet at a party where Amar ask…
1623         1384  The film begins with Riyaz (Rajat Kapoor)  Zub…


                                                 summary  \
11     A coming-of-age story based on the lives of st…
34     Yeh Hai India  follows the story of a 25 years…
37     In the rural heartlands of India  an upright p…
87     After finding out about an illegal arms deal  …
96     A fearless income tax officer raids the mansio…
…                                                      …
1600   Shekhar appoints Simran  a lawyer with an impe…
1601   MAYA is a coming of age story set against an '…
1607   A man accepts a challenge by the chief ministe…
```

```
1621  Shekhar accidentally kills his associate Amar …
1623  Zubeidaa  an aspiring Muslim actress  marries …


                                              tagline  \
11                                  Apna Time Aayega!
34                              A Film for Every Indian
37              Farq Bahut Kar Liya| Ab Farq Laayenge.
87                                The Ultimate Trickery
96                    Heroes don't always come in uniform
…                                                    …
1600                            A dangerous passion…
1601  A child's innocence shattered by unreal tradit…
1607                                    Fight the power
1621                                Love will find a way
1623                            The Story of a Princess


                                               actors  \
11    Ranveer Singh|Alia Bhatt|Siddhant Chaturvedi|V…
34     Gavie Chahal|Mohan Agashe|Mohan Joshi|Lom Harsh|
37    Ayushmann Khurrana|Nassar|Manoj Pahwa|Kumud Mi…
87    Sidharth Malhotra|Manoj Bajpayee|Rakul Preet S…
96    Ajay Devgn|Saurabh Shukla|Ileana D'Cruz|Amit S…
…                                                    …
1600  Divya Dutta|Sucheta Paushe|Lisa Ray|Apoorva Ag…
1601  Anant Nag|Mita Vasisht|Nitya Shetty|Nikhil Yad…
1607  Anil Kapoor|Rani Mukerji|Amrish Puri|Johnny Le…
1621  Priyanshu Chatterjee|Himanshu Malik|Sandali Si…
1623  Karisma Kapoor|Rekha|Manoj Bajpayee|Rajit Kapo…


              wins_nominations              release_date Camera
11        6 wins & 3 nominations    14 February 2019 (USA)   Imax
34        2 wins & 1 nomination      24 May 2019 (India)    Imax
37                        1 win      28 June 2019 (USA)     Imax
87                   1 nomination   16 February 2018 (USA)   Imax
96        2 wins & 3 nominations    16 March 2018 (India)   Imax
…                              …                        …      …
1600              2 nominations     2 February 2001 (India)   Imax
1601    2 wins & 4 nominations   16 January 2002 (Belgium)   Imax
1607              2 nominations    7 September 2001 (India)   Imax
1621              6 nominations        13 July 2001 (India)   Imax
1623  3 wins & 13 nominations    19 January 2001 (India)   Imax


[298 rows x 19 columns]
```

[328]: `movie['Lead Actore']=movie1['actors'].str.split('|').apply(lambda x : x[0])`

[329]: `movie.head(2)`

```
[329]:                      title_x      imdb_id  \
        0  Uri: The Surgical Strike  tt8291224
        1             Battalion 609  tt9472208


                                        poster_path  \
        0  https://upload.wikimedia.org/wikipedia/en/thum…
        1                                        NaN


                                          wiki_link  \
        0  https://en.wikipedia.org/wiki/Uri:_The_Surgica…
        1       https://en.wikipedia.org/wiki/Battalion_609


                         title_y           original_title  is_adult  \
        0  Uri: The Surgical Strike  Uri: The Surgical Strike         0
        1             Battalion 609             Battalion 609         0


           year_of_release runtime          genres  imdb_rating  imdb_votes  \
        0             2019     138  Action|Drama|War          8.4       35112
        1             2019     131               War          4.1          73


                                          story  \
        0  Divided over five chapters  the film chronicle…
        1  The story revolves around a cricket match betw…


                                        summary tagline  \
        0  Indian army special forces execute a covert op…     NaN
        1  The story of Battalion 609 revolves around a c…     NaN


                                         actors wins_nominations  \
        0  Vicky Kaushal|Paresh Rawal|Mohit Raina|Yami Ga…          4 wins
        1  Vicky Ahuja|Shoaib Ibrahim|Shrikant Kamat|Elen…             NaN


                  release_date Camera Lead Actore
        0    11 January 2019 (USA)   Imax          NaN
        1  11 January 2019 (India)   Imax          NaN
```

Some DataFrame Function

```
[332]:  #astype:-- change the data type
        ipl['ID']=ipl['ID'].astype('int32')
```

```
[333]:  ipl.info()


        <class 'pandas.core.frame.DataFrame'>
        RangeIndex: 950 entries, 0 to 949
        Data columns (total 20 columns):
         #   Column              Non-Null Count  Dtype
        ---  ------              --------------  -----
```

```
0   ID               950 non-null   int32
1   City             899 non-null   object
2   Date             950 non-null   object
3   Season           950 non-null   object
4   MatchNumber      950 non-null   object
5   Team1            950 non-null   object
6   Team2            950 non-null   object
7   Venue            950 non-null   object
8   TossWinner       950 non-null   object
9   TossDecision     950 non-null   object
10  SuperOver        946 non-null   object
11  WinningTeam      946 non-null   object
12  WonBy            950 non-null   object
13  Margin           932 non-null   float64
14  method           19 non-null    object
15  Player_of_Match  946 non-null   object
16  Team1Players     950 non-null   object
17  Team2Players     950 non-null   object
18  Umpire1          950 non-null   object
19  Umpire2          950 non-null   object
dtypes: float64(1), int32(1), object(18)
memory usage: 144.9+ KB
```

[335]: `ipl['Season']=ipl['Season'].astype('category')`

[336]: `ipl.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 950 entries, 0 to 949
Data columns (total 20 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   ID               950 non-null    int32
 1   City             899 non-null    object
 2   Date             950 non-null    object
 3   Season           950 non-null    category
 4   MatchNumber      950 non-null    object
 5   Team1            950 non-null    object
 6   Team2            950 non-null    object
 7   Venue            950 non-null    object
 8   TossWinner       950 non-null    object
 9   TossDecision     950 non-null    object
 10  SuperOver        946 non-null    object
 11  WinningTeam      946 non-null    object
 12  WonBy            950 non-null    object
 13  Margin           932 non-null    float64
 14  method           19 non-null     object
 15  Player_of_Match  946 non-null    object
```

```
16   Team1Players      950 non-null     object
17   Team2Players      950 non-null     object
18   Umpire1           950 non-null     object
19   Umpire2           950 non-null     object
dtypes: category(1), float64(1), int32(1), object(17)
memory usage: 139.0+ KB
```

[338]: `ipl['Team1']=ipl["Team1"].astype('category')`

[339]: `ipl['Team2'] = ipl["Team2"].astype('category')`

[340]: `ipl.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 950 entries, 0 to 949
Data columns (total 20 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   ID              950 non-null    int32
 1   City            899 non-null    object
 2   Date            950 non-null    object
 3   Season          950 non-null    category
 4   MatchNumber     950 non-null    object
 5   Team1           950 non-null    category
 6   Team2           950 non-null    category
 7   Venue           950 non-null    object
 8   TossWinner      950 non-null    object
 9   TossDecision    950 non-null    object
 10  SuperOver       946 non-null    object
 11  WinningTeam     946 non-null    object
 12  WonBy           950 non-null    object
 13  Margin          932 non-null    float64
 14  method          19 non-null     object
 15  Player_of_Match 946 non-null    object
 16  Team1Players    950 non-null    object
 17  Team2Players    950 non-null    object
 18  Umpire1         950 non-null    object
 19  Umpire2         950 non-null    object
dtypes: category(3), float64(1), int32(1), object(15)
memory usage: 127.4+ KB
```

[ ]: