

cricket analysis

July 22, 2024

```
[4]: # import libraries
import pandas as pd
```

```
[14]: import json
```

MATCH SUMMARY

```
[8]: # load the match results json file
with open(r'C:
↪\Users\ritik\Downloads\data-analytics-project-for-beginners\t20_json_files\t20_json_files\t
↪json') as f:
    data=json.load(f)
```

```
[9]: data
```

```
[9]: [{'matchSummary': [{'team1': 'Namibia',
    'team2': 'Sri Lanka',
    'winner': 'Namibia',
    'margin': '55 runs',
    'ground': 'Geelong',
    'matchDate': 'Oct 16, 2022',
    'scorecard': 'T20I # 1823'},
    {'team1': 'Netherlands',
    'team2': 'U.A.E.',
    'winner': 'Netherlands',
    'margin': '3 wickets',
    'ground': 'Geelong',
    'matchDate': 'Oct 16, 2022',
    'scorecard': 'T20I # 1825'},
    {'team1': 'Scotland',
    'team2': 'West Indies',
    'winner': 'Scotland',
    'margin': '42 runs',
    'ground': 'Hobart',
    'matchDate': 'Oct 17, 2022',
    'scorecard': 'T20I # 1826'},
    {'team1': 'Ireland',
    'team2': 'Zimbabwe',
```

```

    'winner': 'Zimbabwe',
    'margin': '31 runs',
    'ground': 'Hobart',
    'matchDate': 'Oct 17, 2022',
    'scorecard': 'T20I # 1828'},
    {'team1': 'Namibia',
     'team2': 'Netherlands',
     'winner': 'Netherlands',
     'margin': '5 wickets',
     'ground': 'Geelong',
     'matchDate': 'Oct 18, 2022',
     'scorecard': 'T20I # 1830'},
    {'team1': 'Sri Lanka',
     'team2': 'U.A.E.',
     'winner': 'Sri Lanka',
     'margin': '79 runs',
     'ground': 'Geelong',
     'matchDate': 'Oct 18, 2022',
     'scorecard': 'T20I # 1832'},
    {'team1': 'Ireland',
     'team2': 'Scotland',
     'winner': 'Ireland',
     'margin': '6 wickets',
     'ground': 'Hobart',
     'matchDate': 'Oct 19, 2022',
     'scorecard': 'T20I # 1833'},
    {'team1': 'West Indies',
     'team2': 'Zimbabwe',
     'winner': 'West Indies',
     'margin': '31 runs',
     'ground': 'Hobart',
     'matchDate': 'Oct 19, 2022',
     'scorecard': 'T20I # 1834'},
    {'team1': 'Netherlands',
     'team2': 'Sri Lanka',
     'winner': 'Sri Lanka',
     'margin': '16 runs',
     'ground': 'Geelong',
     'matchDate': 'Oct 20, 2022',
     'scorecard': 'T20I # 1835'},
    {'team1': 'Namibia',
     'team2': 'U.A.E.',
     'winner': 'U.A.E.',
     'margin': '7 runs',
     'ground': 'Geelong',
     'matchDate': 'Oct 20, 2022',
     'scorecard': 'T20I # 1836'},

```

```

{'team1': 'Ireland',
 'team2': 'West Indies',
 'winner': 'Ireland',
 'margin': '9 wickets',
 'ground': 'Hobart',
 'matchDate': 'Oct 21, 2022',
 'scorecard': 'T20I # 1837'},
{'team1': 'Scotland',
 'team2': 'Zimbabwe',
 'winner': 'Zimbabwe',
 'margin': '5 wickets',
 'ground': 'Hobart',
 'matchDate': 'Oct 21, 2022',
 'scorecard': 'T20I # 1838'},
{'team1': 'Australia',
 'team2': 'New Zealand',
 'winner': 'New Zealand',
 'margin': '89 runs',
 'ground': 'Sydney',
 'matchDate': 'Oct 22, 2022',
 'scorecard': 'T20I # 1839'},
{'team1': 'Afghanistan',
 'team2': 'England',
 'winner': 'England',
 'margin': '5 wickets',
 'ground': 'Perth',
 'matchDate': 'Oct 22, 2022',
 'scorecard': 'T20I # 1840'},
{'team1': 'Ireland',
 'team2': 'Sri Lanka',
 'winner': 'Sri Lanka',
 'margin': '9 wickets',
 'ground': 'Hobart',
 'matchDate': 'Oct 23, 2022',
 'scorecard': 'T20I # 1841'},
{'team1': 'India',
 'team2': 'Pakistan',
 'winner': 'India',
 'margin': '4 wickets',
 'ground': 'Melbourne',
 'matchDate': 'Oct 23, 2022',
 'scorecard': 'T20I # 1842'},
{'team1': 'Bangladesh',
 'team2': 'Netherlands',
 'winner': 'Bangladesh',
 'margin': '9 runs',
 'ground': 'Hobart',

```

```

    'matchDate': 'Oct 24, 2022',
    'scorecard': 'T20I # 1843'},
{'team1': 'South Africa',
 'team2': 'Zimbabwe',
 'winner': 'no result',
 'margin': '',
 'ground': 'Hobart',
 'matchDate': 'Oct 24, 2022',
 'scorecard': 'T20I # 1844'},
{'team1': 'Australia',
 'team2': 'Sri Lanka',
 'winner': 'Australia',
 'margin': '7 wickets',
 'ground': 'Perth',
 'matchDate': 'Oct 25, 2022',
 'scorecard': 'T20I # 1845'},
{'team1': 'England',
 'team2': 'Ireland',
 'winner': 'Ireland',
 'margin': '5 runs',
 'ground': 'Melbourne',
 'matchDate': 'Oct 26, 2022',
 'scorecard': 'T20I # 1846'},
{'team1': 'Afghanistan',
 'team2': 'New Zealand',
 'winner': 'abandoned',
 'margin': '',
 'ground': 'Melbourne',
 'matchDate': 'Oct 26, 2022',
 'scorecard': 'T20I # 1846a'},
{'team1': 'Bangladesh',
 'team2': 'South Africa',
 'winner': 'South Africa',
 'margin': '104 runs',
 'ground': 'Sydney',
 'matchDate': 'Oct 27, 2022',
 'scorecard': 'T20I # 1847'},
{'team1': 'India',
 'team2': 'Netherlands',
 'winner': 'India',
 'margin': '56 runs',
 'ground': 'Sydney',
 'matchDate': 'Oct 27, 2022',
 'scorecard': 'T20I # 1848'},
{'team1': 'Pakistan',
 'team2': 'Zimbabwe',
 'winner': 'Zimbabwe',

```

```

    'margin': '1 run',
    'ground': 'Perth',
    'matchDate': 'Oct 27, 2022',
    'scorecard': 'T20I # 1849'},
{'team1': 'Afghanistan',
 'team2': 'Ireland',
 'winner': 'abandoned',
 'margin': '',
 'ground': 'Melbourne',
 'matchDate': 'Oct 28, 2022',
 'scorecard': 'T20I # 1849a'},
{'team1': 'Australia',
 'team2': 'England',
 'winner': 'abandoned',
 'margin': '',
 'ground': 'Melbourne',
 'matchDate': 'Oct 28, 2022',
 'scorecard': 'T20I # 1849b'},
{'team1': 'New Zealand',
 'team2': 'Sri Lanka',
 'winner': 'New Zealand',
 'margin': '65 runs',
 'ground': 'Sydney',
 'matchDate': 'Oct 29, 2022',
 'scorecard': 'T20I # 1850'},
{'team1': 'Bangladesh',
 'team2': 'Zimbabwe',
 'winner': 'Bangladesh',
 'margin': '3 runs',
 'ground': 'Brisbane',
 'matchDate': 'Oct 30, 2022',
 'scorecard': 'T20I # 1851'},
{'team1': 'Netherlands',
 'team2': 'Pakistan',
 'winner': 'Pakistan',
 'margin': '6 wickets',
 'ground': 'Perth',
 'matchDate': 'Oct 30, 2022',
 'scorecard': 'T20I # 1852'},
{'team1': 'India',
 'team2': 'South Africa',
 'winner': 'South Africa',
 'margin': '5 wickets',
 'ground': 'Perth',
 'matchDate': 'Oct 30, 2022',
 'scorecard': 'T20I # 1853'},
{'team1': 'Australia',

```

```

    'team2': 'Ireland',
    'winner': 'Australia',
    'margin': '42 runs',
    'ground': 'Brisbane',
    'matchDate': 'Oct 31, 2022',
    'scorecard': 'T20I # 1855'},
    {'team1': 'Afghanistan',
     'team2': 'Sri Lanka',
     'winner': 'Sri Lanka',
     'margin': '6 wickets',
     'ground': 'Brisbane',
     'matchDate': 'Nov 1, 2022',
     'scorecard': 'T20I # 1856'},
    {'team1': 'England',
     'team2': 'New Zealand',
     'winner': 'England',
     'margin': '20 runs',
     'ground': 'Brisbane',
     'matchDate': 'Nov 1, 2022',
     'scorecard': 'T20I # 1858'},
    {'team1': 'Netherlands',
     'team2': 'Zimbabwe',
     'winner': 'Netherlands',
     'margin': '5 wickets',
     'ground': 'Adelaide',
     'matchDate': 'Nov 2, 2022',
     'scorecard': 'T20I # 1859'},
    {'team1': 'Bangladesh',
     'team2': 'India',
     'winner': 'India',
     'margin': '5 runs',
     'ground': 'Adelaide',
     'matchDate': 'Nov 2, 2022',
     'scorecard': 'T20I # 1860'},
    {'team1': 'Pakistan',
     'team2': 'South Africa',
     'winner': 'Pakistan',
     'margin': '33 runs',
     'ground': 'Sydney',
     'matchDate': 'Nov 3, 2022',
     'scorecard': 'T20I # 1861'},
    {'team1': 'Ireland',
     'team2': 'New Zealand',
     'winner': 'New Zealand',
     'margin': '35 runs',
     'ground': 'Adelaide',
     'matchDate': 'Nov 4, 2022',

```

```

    'scorecard': 'T20I # 1862'},
{'team1': 'Australia',
 'team2': 'Afghanistan',
 'winner': 'Australia',
 'margin': '4 runs',
 'ground': 'Adelaide',
 'matchDate': 'Nov 4, 2022',
 'scorecard': 'T20I # 1864'},
{'team1': 'England',
 'team2': 'Sri Lanka',
 'winner': 'England',
 'margin': '4 wickets',
 'ground': 'Sydney',
 'matchDate': 'Nov 5, 2022',
 'scorecard': 'T20I # 1867'},
{'team1': 'Netherlands',
 'team2': 'South Africa',
 'winner': 'Netherlands',
 'margin': '13 runs',
 'ground': 'Adelaide',
 'matchDate': 'Nov 6, 2022',
 'scorecard': 'T20I # 1871'},
{'team1': 'Bangladesh',
 'team2': 'Pakistan',
 'winner': 'Pakistan',
 'margin': '5 wickets',
 'ground': 'Adelaide',
 'matchDate': 'Nov 6, 2022',
 'scorecard': 'T20I # 1872'},
{'team1': 'India',
 'team2': 'Zimbabwe',
 'winner': 'India',
 'margin': '71 runs',
 'ground': 'Melbourne',
 'matchDate': 'Nov 6, 2022',
 'scorecard': 'T20I # 1873'},
{'team1': 'New Zealand',
 'team2': 'Pakistan',
 'winner': 'Pakistan',
 'margin': '7 wickets',
 'ground': 'Sydney',
 'matchDate': 'Nov 9, 2022',
 'scorecard': 'T20I # 1877'},
{'team1': 'England',
 'team2': 'India',
 'winner': 'England',
 'margin': '10 wickets',

```

```

    'ground': 'Adelaide',
    'matchDate': 'Nov 10, 2022',
    'scorecard': 'T20I # 1878'},
{'team1': 'England',
 'team2': 'Pakistan',
 'winner': 'England',
 'margin': '5 wickets',
 'ground': 'Melbourne',
 'matchDate': 'Nov 13, 2022',
 'scorecard': 'T20I # 1879'}}]]

```

```

[10]: #Converts data from json file to a table format using pandas DataFrame
df_matches = pd.DataFrame(data[0]['matchSummary'])
df_matches.head(3)

```

```

[10]:
   team1 team2 winner margin ground matchDate \
0  Namibia  Sri Lanka  Namibia  55 runs  Geelong  Oct 16, 2022
1  Netherlands  U.A.E.  Netherlands  3 wickets  Geelong  Oct 16, 2022
2  Scotland  West Indies  Scotland  42 runs  Hobart  Oct 17, 2022

   scorecard
0  T20I # 1823
1  T20I # 1825
2  T20I # 1826

```

```

[11]: df_matches.shape

```

```

[11]: (45, 7)

```

```

[12]: df_matches.rename({'scorecard': 'Match ID'}, axis =1, inplace = 1) #rename
      ↪scorecard to Match ID
df_matches.head(10)

```

```

[12]:
   team1 team2 winner margin ground matchDate \
0  Namibia  Sri Lanka  Namibia  55 runs  Geelong  Oct 16, 2022
1  Netherlands  U.A.E.  Netherlands  3 wickets  Geelong  Oct 16, 2022
2  Scotland  West Indies  Scotland  42 runs  Hobart  Oct 17, 2022
3  Ireland  Zimbabwe  Zimbabwe  31 runs  Hobart  Oct 17, 2022
4  Namibia  Netherlands  Netherlands  5 wickets  Geelong  Oct 18, 2022
5  Sri Lanka  U.A.E.  Sri Lanka  79 runs  Geelong  Oct 18, 2022
6  Ireland  Scotland  Ireland  6 wickets  Hobart  Oct 19, 2022
7  West Indies  Zimbabwe  West Indies  31 runs  Hobart  Oct 19, 2022
8  Netherlands  Sri Lanka  Sri Lanka  16 runs  Geelong  Oct 20, 2022
9  Namibia  U.A.E.  U.A.E.  7 runs  Geelong  Oct 20, 2022

   Match ID
0  T20I # 1823

```



```

1 T20I # 1825
2 T20I # 1826
3 T20I # 1828
4 T20I # 1830
5 T20I # 1832
6 T20I # 1833
7 T20I # 1834
8 T20I # 1835
9 T20I # 1836

```

```

[22]: # transfer the dataframe to match summary csv file
df_matches.to_csv(r'C:
↪\Users\ritik\Downloads\data-analytics-project-for-beginners\t20_csv_files\t20_csv_files\dim
↪csv',index=True)

```

```

[23]: match_ids_dict = {}

for index, row in df_matches.iterrows():
    key1 = row['team1'] + ' Vs ' + row['team2']
    key2 = row['team2'] + ' Vs ' + row['team1']
    match_ids_dict[key1] = row['Match ID']
    match_ids_dict[key2] = row['Match ID']

match_ids_dict

```

```

[23]: {'Namibia Vs Sri Lanka': 'T20I # 1823',
'Sri Lanka Vs Namibia': 'T20I # 1823',
'Netherlands Vs U.A.E.': 'T20I # 1825',
'U.A.E. Vs Netherlands': 'T20I # 1825',
'Scotland Vs West Indies': 'T20I # 1826',
'West Indies Vs Scotland': 'T20I # 1826',
'Ireland Vs Zimbabwe': 'T20I # 1828',
'Zimbabwe Vs Ireland': 'T20I # 1828',
'Namibia Vs Netherlands': 'T20I # 1830',
'Netherlands Vs Namibia': 'T20I # 1830',
'Sri Lanka Vs U.A.E.': 'T20I # 1832',
'U.A.E. Vs Sri Lanka': 'T20I # 1832',
'Ireland Vs Scotland': 'T20I # 1833',
'Scotland Vs Ireland': 'T20I # 1833',
'West Indies Vs Zimbabwe': 'T20I # 1834',
'Zimbabwe Vs West Indies': 'T20I # 1834',
'Netherlands Vs Sri Lanka': 'T20I # 1835',
'Sri Lanka Vs Netherlands': 'T20I # 1835',
'Namibia Vs U.A.E.': 'T20I # 1836',
'U.A.E. Vs Namibia': 'T20I # 1836',
'Ireland Vs West Indies': 'T20I # 1837',
'West Indies Vs Ireland': 'T20I # 1837',

```

'Scotland Vs Zimbabwe': 'T20I # 1838',
'Zimbabwe Vs Scotland': 'T20I # 1838',
'Australia Vs New Zealand': 'T20I # 1839',
'New Zealand Vs Australia': 'T20I # 1839',
'Afghanistan Vs England': 'T20I # 1840',
'England Vs Afghanistan': 'T20I # 1840',
'Ireland Vs Sri Lanka': 'T20I # 1841',
'Sri Lanka Vs Ireland': 'T20I # 1841',
'India Vs Pakistan': 'T20I # 1842',
'Pakistan Vs India': 'T20I # 1842',
'Bangladesh Vs Netherlands': 'T20I # 1843',
'Netherlands Vs Bangladesh': 'T20I # 1843',
'South Africa Vs Zimbabwe': 'T20I # 1844',
'Zimbabwe Vs South Africa': 'T20I # 1844',
'Australia Vs Sri Lanka': 'T20I # 1845',
'Sri Lanka Vs Australia': 'T20I # 1845',
'England Vs Ireland': 'T20I # 1846',
'Ireland Vs England': 'T20I # 1846',
'Afghanistan Vs New Zealand': 'T20I # 1846a',
'New Zealand Vs Afghanistan': 'T20I # 1846a',
'Bangladesh Vs South Africa': 'T20I # 1847',
'South Africa Vs Bangladesh': 'T20I # 1847',
'India Vs Netherlands': 'T20I # 1848',
'Netherlands Vs India': 'T20I # 1848',
'Pakistan Vs Zimbabwe': 'T20I # 1849',
'Zimbabwe Vs Pakistan': 'T20I # 1849',
'Afghanistan Vs Ireland': 'T20I # 1849a',
'Ireland Vs Afghanistan': 'T20I # 1849a',
'Australia Vs England': 'T20I # 1849b',
'England Vs Australia': 'T20I # 1849b',
'New Zealand Vs Sri Lanka': 'T20I # 1850',
'Sri Lanka Vs New Zealand': 'T20I # 1850',
'Bangladesh Vs Zimbabwe': 'T20I # 1851',
'Zimbabwe Vs Bangladesh': 'T20I # 1851',
'Netherlands Vs Pakistan': 'T20I # 1852',
'Pakistan Vs Netherlands': 'T20I # 1852',
'India Vs South Africa': 'T20I # 1853',
'South Africa Vs India': 'T20I # 1853',
'Australia Vs Ireland': 'T20I # 1855',
'Ireland Vs Australia': 'T20I # 1855',
'Afghanistan Vs Sri Lanka': 'T20I # 1856',
'Sri Lanka Vs Afghanistan': 'T20I # 1856',
'England Vs New Zealand': 'T20I # 1858',
'New Zealand Vs England': 'T20I # 1858',
'Netherlands Vs Zimbabwe': 'T20I # 1859',
'Zimbabwe Vs Netherlands': 'T20I # 1859',
'Bangladesh Vs India': 'T20I # 1860',

```
'India Vs Bangladesh': 'T20I # 1860',
'Pakistan Vs South Africa': 'T20I # 1861',
'South Africa Vs Pakistan': 'T20I # 1861',
'Ireland Vs New Zealand': 'T20I # 1862',
'New Zealand Vs Ireland': 'T20I # 1862',
'Australia Vs Afghanistan': 'T20I # 1864',
'Afghanistan Vs Australia': 'T20I # 1864',
'England Vs Sri Lanka': 'T20I # 1867',
'Sri Lanka Vs England': 'T20I # 1867',
'Netherlands Vs South Africa': 'T20I # 1871',
'South Africa Vs Netherlands': 'T20I # 1871',
'Bangladesh Vs Pakistan': 'T20I # 1872',
'Pakistan Vs Bangladesh': 'T20I # 1872',
'India Vs Zimbabwe': 'T20I # 1873',
'Zimbabwe Vs India': 'T20I # 1873',
'New Zealand Vs Pakistan': 'T20I # 1877',
'Pakistan Vs New Zealand': 'T20I # 1877',
'England Vs India': 'T20I # 1878',
'India Vs England': 'T20I # 1878',
'England Vs Pakistan': 'T20I # 1879',
'Pakistan Vs England': 'T20I # 1879'}
```

BATTING SUMMARY

```
[25]: with open(r'C:
↳\Users\ritik\Downloads\data-analytics-project-for-beginners\t20_json_files\t20_json_files\t
↳json') as f:
    data1=json.load(f)

    # loads all batting summary of all matches into one single table by joining
↳batting summary of every match using extend
    all_records = []
    for rec in data1:
        all_records.extend(rec['battingSummary'])

df_bat = pd.DataFrame(all_records)
df_bat.head()
```

```
[25]:
```

	match	team	Innings	battingPos	batsmanName	\
0	Namibia Vs Sri Lanka	Namibia		1	Michael van Lingen	
1	Namibia Vs Sri Lanka	Namibia		2	Divan la Cock	
2	Namibia Vs Sri Lanka	Namibia		3	Jan Nicol Loftie-Eaton	
3	Namibia Vs Sri Lanka	Namibia		4	Stephan Baard	
4	Namibia Vs Sri Lanka	Namibia		5	Gerhard Erasmus(c)	

		dismissal	runs	balls	4s	6s	SR
0	c Pramod Madushan	b Chameera	3	6	0	0	50.00

1	c	Shanaka	b	Pramod Madushan	9	9	1	0	100.00
2	c	â€	Mendis	b Karunaratne	20	12	1	2	166.66
3	c	DM de Silva	b	Pramod Madushan	26	24	2	0	108.33
4	c	Gunathilaka	b	PWH de Silva	20	24	0	0	83.33

```
[18]: # Adds column 'Out/Not out'
df_bat["Out/Not_out"] = df_bat.dismissal.apply(lambda x: 'Out' if len(x)>0 else
↳ 'Not out')
# remove the dismissal column
df_bat.drop(columns = ['dismissal'], inplace = True)
df_bat.head(10)
```

```
[18]:
```

	match	team	Innings	battingPos	batsmanName	runs	\
0	Namibia Vs Sri Lanka	Namibia		1	Michael van Lingen	3	
1	Namibia Vs Sri Lanka	Namibia		2	Divan la Cock	9	
2	Namibia Vs Sri Lanka	Namibia		3	Jan Nicol Loftie-Eaton	20	
3	Namibia Vs Sri Lanka	Namibia		4	Stephan Baard	26	
4	Namibia Vs Sri Lanka	Namibia		5	Gerhard Erasmus(c)	20	
5	Namibia Vs Sri Lanka	Namibia		6	Jan Frylinck	44	
6	Namibia Vs Sri Lanka	Namibia		7	David Wiese	0	
7	Namibia Vs Sri Lanka	Namibia		8	JJ Smit	31	
8	Namibia Vs Sri Lanka	Sri Lanka		1	Pathum Nissanka	9	
9	Namibia Vs Sri Lanka	Sri Lanka		2	Kusal Mendisâ€	6	

	balls	4s	6s	SR	Out/Not_out
0	6	0	0	50.00	Out
1	9	1	0	100.00	Out
2	12	1	2	166.66	Out
3	24	2	0	108.33	Out
4	24	0	0	83.33	Out
5	28	4	0	157.14	Out
6	1	0	0	0.00	Out
7	16	2	2	193.75	Not out
8	10	1	0	90.00	Out
9	6	0	0	100.00	Out

```
[19]: df_bat["batsmanName"] = df_bat['batsmanName'].replace('â€', '')
df_bat["batsmanName"] = df_bat['batsmanName'].replace('\xa0', '')
df_bat.head(10)
```

```
[19]:
```

	match	team	Innings	battingPos	batsmanName	runs	\
0	Namibia Vs Sri Lanka	Namibia		1	Michael van Lingen	3	
1	Namibia Vs Sri Lanka	Namibia		2	Divan la Cock	9	
2	Namibia Vs Sri Lanka	Namibia		3	Jan Nicol Loftie-Eaton	20	
3	Namibia Vs Sri Lanka	Namibia		4	Stephan Baard	26	
4	Namibia Vs Sri Lanka	Namibia		5	Gerhard Erasmus(c)	20	
5	Namibia Vs Sri Lanka	Namibia		6	Jan Frylinck	44	

6	Namibia Vs Sri Lanka	Namibia	7	David Wiese	0
7	Namibia Vs Sri Lanka	Namibia	8	JJ Smit	31
8	Namibia Vs Sri Lanka	Sri Lanka	1	Pathum Nissanka	9
9	Namibia Vs Sri Lanka	Sri Lanka	2	Kusal Mendis	6

	balls	4s	6s	SR	Out/Not_out
0	6	0	0	50.00	Out
1	9	1	0	100.00	Out
2	12	1	2	166.66	Out
3	24	2	0	108.33	Out
4	24	0	0	83.33	Out
5	28	4	0	157.14	Out
6	1	0	0	0.00	Out
7	16	2	2	193.75	Not out
8	10	1	0	90.00	Out
9	6	0	0	100.00	Out

```
[24]: # adds match ID to batting table by linking 'Match ID' column to 'match' column
      ↪ using map feature
df_bat['Match ID'] = df_bat['match'].map(match_ids_dict)
df_bat.head(10)
```

	match	team	Innings	battingPos	batsmanName	runs	\
0	Namibia Vs Sri Lanka	Namibia		1	Michael van Lingen	3	
1	Namibia Vs Sri Lanka	Namibia		2	Divan la Cock	9	
2	Namibia Vs Sri Lanka	Namibia		3	Jan Nicol Loftie-Eaton	20	
3	Namibia Vs Sri Lanka	Namibia		4	Stephan Baard	26	
4	Namibia Vs Sri Lanka	Namibia		5	Gerhard Erasmus(c)	20	
5	Namibia Vs Sri Lanka	Namibia		6	Jan Frylinck	44	
6	Namibia Vs Sri Lanka	Namibia		7	David Wiese	0	
7	Namibia Vs Sri Lanka	Namibia		8	JJ Smit	31	
8	Namibia Vs Sri Lanka	Sri Lanka		1	Pathum Nissanka	9	
9	Namibia Vs Sri Lanka	Sri Lanka		2	Kusal Mendis	6	

	balls	4s	6s	SR	Out/Not_out	Match ID
0	6	0	0	50.00	Out	T20I # 1823
1	9	1	0	100.00	Out	T20I # 1823
2	12	1	2	166.66	Out	T20I # 1823
3	24	2	0	108.33	Out	T20I # 1823
4	24	0	0	83.33	Out	T20I # 1823
5	28	4	0	157.14	Out	T20I # 1823
6	1	0	0	0.00	Out	T20I # 1823
7	16	2	2	193.75	Not out	T20I # 1823
8	10	1	0	90.00	Out	T20I # 1823
9	6	0	0	100.00	Out	T20I # 1823

```
[26]: # transfer the dataframe to batting summary csv file
df_bat.to_csv(r'C:
↳\Users\ritik\Downloads\data-analytics-project-for-beginners\t20_csv_files\t20_csv_files\fac
↳csv',index=True)
```

BOWLING SUMMARY

```
[27]: # load the bowling summary json file
with open(r'C:
↳\Users\ritik\Downloads\data-analytics-project-for-beginners\t20_json_files\t20_json_files\t
↳json') as f:
    data2=json.load(f)
    # loads all bowling summary of all matches into one single table by
    ↳joining bowling summary of every match using extend
    bowl_records=[]
    for rec in data2:
        bowl_records.extend(rec['bowlingSummary'])

df_bowl=pd.DataFrame(bowl_records)
df_bowl
```

```
[27]:
```

	match	bowlingTeam	bowlerName	overs	\
0	Namibia Vs Sri Lanka	Sri Lanka	Maheesh Theekshana	4	
1	Namibia Vs Sri Lanka	Sri Lanka	Dushmantha Chameera	4	
2	Namibia Vs Sri Lanka	Sri Lanka	Pramod Madushan	4	
3	Namibia Vs Sri Lanka	Sri Lanka	Chamika Karunaratne	4	
4	Namibia Vs Sri Lanka	Sri Lanka	Wanindu Hasaranga de Silva	4	
..	
495	Pakistan Vs England	Pakistan	Naseem Shah	4	
496	Pakistan Vs England	Pakistan	Haris Rauf	4	
497	Pakistan Vs England	Pakistan	Shadab Khan	4	
498	Pakistan Vs England	Pakistan	Mohammad Wasim	4	
499	Pakistan Vs England	Pakistan	Iftikhar Ahmed	0.5	

	maiden	runs	wickets	economy	0s	4s	6s	wides	noBalls
0	0	23	1	5.75	7	0	0	2	0
1	0	39	1	9.75	6	3	1	2	0
2	0	37	2	9.25	6	3	1	0	0
3	0	36	1	9.00	7	3	1	1	0
4	0	27	1	6.75	8	1	1	0	0
..
495	0	30	0	7.50	15	3	1	1	0
496	0	23	2	5.75	13	3	0	1	0
497	0	20	1	5.00	10	1	0	0	0
498	0	38	1	9.50	5	5	0	2	0
499	0	13	0	15.60	0	1	1	0	0

[500 rows x 13 columns]

```
[28]: # adds match ID to bowling table by linking 'Match ID' column to 'match' column
      ↪ using map feature
df_bowl['Match ID']= df_bowl['match'].map(match_ids_dict)
df_bowl
```

```
[28]:
```

	match	bowlingTeam	bowlerName	overs	\
0	Namibia Vs Sri Lanka	Sri Lanka	Maheesh Theekshana	4	
1	Namibia Vs Sri Lanka	Sri Lanka	Dushmantha Chameera	4	
2	Namibia Vs Sri Lanka	Sri Lanka	Pramod Madushan	4	
3	Namibia Vs Sri Lanka	Sri Lanka	Chamika Karunaratne	4	
4	Namibia Vs Sri Lanka	Sri Lanka	Wanindu Hasaranga de Silva	4	
..	
495	Pakistan Vs England	Pakistan	Naseem Shah	4	
496	Pakistan Vs England	Pakistan	Haris Rauf	4	
497	Pakistan Vs England	Pakistan	Shadab Khan	4	
498	Pakistan Vs England	Pakistan	Mohammad Wasim	4	
499	Pakistan Vs England	Pakistan	Iftikhar Ahmed	0.5	

	maiden	runs	wickets	economy	0s	4s	6s	wides	noBalls	Match ID
0	0	23	1	5.75	7	0	0	2	0	T20I # 1823
1	0	39	1	9.75	6	3	1	2	0	T20I # 1823
2	0	37	2	9.25	6	3	1	0	0	T20I # 1823
3	0	36	1	9.00	7	3	1	1	0	T20I # 1823
4	0	27	1	6.75	8	1	1	0	0	T20I # 1823
..	
495	0	30	0	7.50	15	3	1	1	0	T20I # 1879
496	0	23	2	5.75	13	3	0	1	0	T20I # 1879
497	0	20	1	5.00	10	1	0	0	0	T20I # 1879
498	0	38	1	9.50	5	5	0	2	0	T20I # 1879
499	0	13	0	15.60	0	1	1	0	0	T20I # 1879

[500 rows x 14 columns]

```
[29]: # transfer into bowling summary csv file
df_bowl.to_csv(r'C:
      ↪ \Users\ritik\Downloads\data-analytics-project-for-beginners\t20_csv_files\t20_csv_files\fac
      ↪ csv')
```

Player Info

```
[31]: # load the json file
with open(r'C:
      ↪ \Users\ritik\Downloads\data-analytics-project-for-beginners\t20_json_files\t20_json_files\t
      ↪ json') as f:
    data3=json.load(f)
```

```
df_players_info=pd.DataFrame(data3)
df_players_info.head(10)
```

```
[31]:
```

	name	team	battingStyle	\
0	Michael van Lingen	Namibia	Left hand Bat	
1	Divan la Cock	Namibia	Right hand Bat	
2	Jan Nicol Loftie-Eaton	Namibia	Left hand Bat	
3	Stephan Baard	Namibia	Right hand Bat	
4	Gerhard Erasmus(c)	Namibia	Right hand Bat	
5	Jan Frylinck	Namibia	Left hand Bat	
6	David Wiese	Namibia	Right hand Bat	
7	JJ Smit	Namibia	Right hand Bat	
8	Pathum Nissanka	Sri Lanka	Right hand Bat	
9	Kusal Mendis	Sri Lanka	Right hand Bat	

	bowlingStyle	playingRole	\
0	Left arm Medium	Bowling Allrounder	
1	Legbreak	Opening Batter	
2	Right arm Medium, Legbreak	Batter	
3	Right arm Medium fast	Batter	
4	Right arm Offbreak	Allrounder	
5	Left arm Fast medium	Allrounder	
6	Right arm Medium fast	Allrounder	
7	Left arm Medium fast	Bowling Allrounder	
8		Top order Batter	
9	Legbreak	Wicketkeeper Batter	

	description
0	
1	
2	
3	
4	
5	
6	David Wiese joined a marked outflow of South A...
7	
8	
9	Blessed with a compact technique, an aggressiv...

```
[32]: df_players_info['name'] = df_players_info['name'].apply(lambda x: x.
    ↪replace('â€', ''))
df_players_info['name'] = df_players_info['name'].apply(lambda x: x.
    ↪replace('†', ''))
df_players_info['name'] = df_players_info['name'].apply(lambda x: x.
    ↪replace('\xa0', ''))
df_players_info.head(10)
```



```
[32]:
```

	name	team	battingStyle	\
0	Michael van Lingen	Namibia	Left hand Bat	
1	Divan la Cock	Namibia	Right hand Bat	
2	Jan Nicol Loftie-Eaton	Namibia	Left hand Bat	
3	Stephan Baard	Namibia	Right hand Bat	
4	Gerhard Erasmus(c)	Namibia	Right hand Bat	
5	Jan Frylinck	Namibia	Left hand Bat	
6	David Wiese	Namibia	Right hand Bat	
7	JJ Smit	Namibia	Right hand Bat	
8	Pathum Nissanka	Sri Lanka	Right hand Bat	
9	Kusal Mendis	Sri Lanka	Right hand Bat	

	bowlingStyle	playingRole	\
0	Left arm Medium	Bowling Allrounder	
1	Legbreak	Opening Batter	
2	Right arm Medium, Legbreak	Batter	
3	Right arm Medium fast	Batter	
4	Right arm Offbreak	Allrounder	
5	Left arm Fast medium	Allrounder	
6	Right arm Medium fast	Allrounder	
7	Left arm Medium fast	Bowling Allrounder	
8		Top order Batter	
9	Legbreak	Wicketkeeper Batter	

	description
0	
1	
2	
3	
4	
5	
6	David Wiese joined a marked outflow of South A...
7	
8	
9	Blessed with a compact technique, an aggressiv...

```
[37]: df_players_info.to_csv(r'C:\Users\ritik\Downloads\data-analytics-project-for-beginners\t20_csv_files\t20_csv_files\dim
↪ csv')
```

```
[ ]:
```