Abstract

Cricket is sport enjoyed across India. We have extracted data for the cricket's newest format t20 and fit it in our conceptual database.

Data Description

The dataset for all three methods has been gathered from the espncricinfo and Kaggle websites. The file from Kaggle is t20_matches.csv. It contains 26 columns and 6417 entries. Below is the link for the dataset:

https://www.kaggle.com/imrankhan17/t20matches (https://www.kaggle.com/imrankhan17/t20matches)

We have also used the same dataset to create a different file which we have used for API data extraction.

```
In [280]:   #Below are the columns in t20_cricket

            #*   match_id = shows the match id for that particular match
            #*   series_id = shows the series id for that particular series
            #*   match_details = shows the match details
            #*   result = shows the end result of the match
            #*   scores = shows the final scores for both teams
            #*   date = shows the date of the day when the game was played
            #*   venue = shows the venue of the match
            #*   round = shows the round of the tournament
            #*   home = shows the home team
            #*   away = shows the away team
            #*   winner = shows the winning team
            #*   win_by_runs = shows the margin of winning by number of runs
            #*   win_by_wickets = shows the margin of winning by number of wickets
            #*   balls_remaining = shows the balls remaining after the end of the game
            #*   innings1 = shows the team batting first
            #*   innings1_runs = shows the first inning score
            #*   innings1_wickets = shows the first inning wickets
            #*   innings1_overs_batted = shows the first innings overs bowled/batted
            #*   innings1_overs = shows the total number of overs in the second innings
            #*   innings2 = shows the team batting second
            #*   innings2_runs = shows the second inning score
            #*   innings2_wickets = shows the second inning wickets
            #*   innings2_overs_batted = shows the second innings overs bowled/batted
            #*   innings2_overs = shows the total number of overs in the second innings
            #*   D/L_method = shows wether duckworth lewis method was used
            #*   target = shows the target for the match
```

Report

We searched for data for cricket and zeroed in on the t20 format as it is the latest format. For the CSV file and Web API scraping we have created a schema which show the match ID's and Match description against it. This will give the viewer an idea about the teams that played the match and where and when the match was played. Additionally the viewer will also get an idea about the series which the match was a part of.

We have deleted the columns which contained null entries from the csv file for Audit Validity. We have taken the match ID's from the CSV file and created another file named india.csv whcih we used in conjuction with API and put the extracted data in the schema.

We used the espncricinfo site for webscraping and used it extract data of t20 playing india batsmen. We used BeautifulSoup library for the same.

Importing Libraries

```
In [1]: import pandas as pd
        import os
```

Read From CSV File

```
In [2]: cricket = pd.read_csv('t20_matches.csv')
```

Output Of Data

In [3]: `cricket.head()`

Out[3]:

	match_id	series_id	match_details	result	scores	date	ven
0	304744	124121	North Group:Durham v Nottinghamshire at Cheste...	Durham won by 6 wickets (with 5 balls remaining)	Nottinghamshire 157/7 (20/20 ov); Durham 160/4...	13/06/2003	Chester-Stre
1	129610	124121	South Group:Hampshire v Sussex at Southampton-...	Hampshire won by 5 runs	Hampshire 153 (19.4/20 ov); Sussex 148/7 (20/2...	13/06/2003	Southampt
2	129597	124121	Midlands/Wales/West Group:Somerset v Warwicksh...	Warwickshire won by 19 runs	Warwickshire 188/7 (20/20 ov); Somerset 169 (2...	13/06/2003	Taunt
3	304745	124121	South Group:Surrey v Middlesex at The Oval- Ju...	Surrey won by 4 wickets (with 4 balls remaining)	Middlesex 155 (20/20 ov); Surrey 158/6 (19.2/2...	13/06/2003	The Ov
4	304746	124121	Midlands/Wales/West Group:Worcestershire v Nor...	Worcestershire won by 1 wicket (with 2 balls r...	Northamptonshire 150/9 (20/20 ov); Worcestersh...	13/06/2003	Worcest

5 rows × 26 columns

Checking Shape of Data

In [4]: `cricket.shape`

Out[4]: `(6417, 26)`

In [5]: `cricket.columns`

Out[5]: `Index(['match_id', 'series_id', 'match_details', 'result', 'scores', 'date',`
`        'venue', 'round', 'home', 'away', 'winner', 'win_by_runs',`
`        'win_by_wickets', 'balls_remaining', 'innings1', 'innings1_runs',`
`        'innings1_wickets', 'innings1_overs_batted', 'innings1_overs',`
`        'innings2', 'innings2_runs', 'innings2_wickets',`
`        'innings2_overs_batted', 'innings2_overs', 'D/L_method', 'target'],`
`        dtype='object')`

Checking Info in Every Column

In [6]: `cricket.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6417 entries, 0 to 6416
Data columns (total 26 columns):
match_id                6417 non-null int64
series_id               6417 non-null int64
match_details           6417 non-null object
result                  6403 non-null object
scores                  6151 non-null object
date                    6417 non-null object
venue                   6417 non-null object
round                   5143 non-null object
home                    6417 non-null object
away                    6417 non-null object
winner                  6403 non-null object
win_by_runs             2931 non-null float64
win_by_wickets          3004 non-null float64
balls_remaining         2969 non-null float64
innings1                6151 non-null object
innings1_runs           6151 non-null float64
innings1_wickets        6151 non-null float64
innings1_overs_batted   6151 non-null float64
innings1_overs          6151 non-null float64
innings2                6070 non-null object
innings2_runs           6070 non-null float64
innings2_wickets        6070 non-null float64
innings2_overs_batted   6070 non-null float64
innings2_overs          6070 non-null float64
D/L_method              233 non-null float64
target                  231 non-null float64
dtypes: float64(13), int64(2), object(11)
memory usage: 1.3+ MB
```

Checking for Null values

```
In [7]:  cricket.isnull().sum()
```

```
Out[7]:  match_id                 0
         series_id                0
         match_details            0
         result                  14
         scores                 266
         date                     0
         venue                    0
         round                 1274
         home                     0
         away                     0
         winner                  14
         win_by_runs           3486
         win_by_wickets        3413
         balls_remaining       3448
         innings1               266
         innings1_runs          266
         innings1_wickets       266
         innings1_overs_batted  266
         innings1_overs         266
         innings2               347
         innings2_runs          347
         innings2_wickets       347
         innings2_overs_batted  347
         innings2_overs         347
         D/L_method            6184
         target                6186
         dtype: int64
```

```
In [8]:  cricket.drop(['target','D/L_method','balls_remaining','win_by_wickets','win_by_r
         uns','round'],axis=1, inplace=True)
```

In [9]: `cricket.head()`

Out[9]:

	match_id	series_id	match_details	result	scores	date	venue
0	304744	124121	North Group:Durham v Nottinghamshire at Cheste...	Durham won by 6 wickets (with 5 balls remaining)	Nottinghamshire 157/7 (20/20 ov); Durham 160/4...	13/06/2003	Chester-Stre
1	129610	124121	South Group:Hampshire v Sussex at Southampton-...	Hampshire won by 5 runs	Hampshire 153 (19.4/20 ov); Sussex 148/7 (20/2...	13/06/2003	Southampto
2	129597	124121	Midlands/Wales/West Group:Somerset v Warwicksh...	Warwickshire won by 19 runs	Warwickshire 188/7 (20/20 ov); Somerset 169 (2...	13/06/2003	Taunt
3	304745	124121	South Group:Surrey v Middlesex at The Oval- Ju...	Surrey won by 4 wickets (with 4 balls remaining)	Middlesex 155 (20/20 ov); Surrey 158/6 (19.2/2...	13/06/2003	The Ov
4	304746	124121	Midlands/Wales/West Group:Worcestershire v Nor...	Worcestershire won by 1 wicket (with 2 balls r...	Northamptonshire 150/9 (20/20 ov); Worcestersh...	13/06/2003	Worcest

Using Web API

In [111]: `INDmatches = pd.read_csv('india.csv')`

In [112]: `INDmatches.head()`

Out[112]:

	MATCH_ID
0	287859
1	287862
2	287865
3	287879
4	297800

Creating a list for Match Description

In [113]: `Description_list=[]`

Creating a list for Match ID

In [114]: `Matchid_list=[]`

In [115]: `Matchid_list = INDmatches.MATCH_ID`

Using Web API to gather match description data

In [116]:
```python
from espncricinfo.match import Match
for matchId in INDmatches.MATCH_ID:
    try:
        m = Match(matchId)
        print(m.description)
        Description_list.append(m.description)
    except:
        Description_list.append("Error")
        print("error fetching data for match with id: "+str(matchId))
```

ICC World Twenty20, 7th Match, Group D: India v Scotland at Durban, Sep 13, 2007

ICC World Twenty20, 10th Match, Group D: India v Pakistan at Durban, Sep 14, 2007

ICC World Twenty20, 13th Match, Group E: India v New Zealand at Johannesburg, Sep 16, 2007

ICC World Twenty20, Final: India v Pakistan at Johannesburg, Sep 24, 2007

Australia tour of India, Only T20I: India v Australia at Mumbai (BS), Oct 20, 2007

ICC World Twenty20, 12th Match, Group A: India v Ireland at Nottingham, Jun 10, 2009

ICC World Twenty20, 16th Match, Group E: India v West Indies at Lord's, Jun 12, 2009

ICC World Twenty20, 24th Match, Group E: India v South Africa at Nottingham, Jun 16, 2009

Sri Lanka tour of India, 1st T20I: India v Sri Lanka at Nagpur, Dec 9, 2009

Sri Lanka tour of India, 2nd T20I: India v Sri Lanka at Mohali, Dec 12, 2009

ICC World Twenty20, 5th Match, Group C: India v South Africa at Gros Islet, May 2, 2010

ICC World Twenty20, 23rd Match, Group F: India v Sri Lanka at Gros Islet, May 11, 2010

England tour of India, Only T20I: India v England at Kolkata, Oct 29, 2011

error fetching data for match with id: 565819

New Zealand tour of India, 2nd T20I: India v New Zealand at Chennai, Sep 11, 2012

ICC World Twenty20, 20th Match, Group 2: India v Pakistan at Colombo (RPS), Sep 30, 2012

ICC World Twenty20, 24th Match, Group 2: India v South Africa at Colombo (RPS), Oct 2, 2012

England tour of India, 1st T20I: India v England at Pune, Dec 20, 2012

England tour of India, 2nd T20I: India v England at Mumbai, Dec 22, 2012

Pakistan tour of India, 1st T20I: India v Pakistan at Bengaluru, Dec 25, 2012

Pakistan tour of India, 2nd T20I: India v Pakistan at Ahmedabad, Dec 28, 2012

Australia tour of India, Only T20I: India v Australia at Rajkot, Oct 10, 2013

World T20, 13th Match, Group 2: India v Pakistan at Dhaka, Mar 21, 2014

World T20, 17th Match, Group 2: India v West Indies at Dhaka, Mar 23, 2014

World T20, 2nd Semi-Final: India v South Africa at Dhaka, Apr 4, 2014

World T20, Final: India v Sri Lanka at Dhaka, Apr 6, 2014

South Africa tour of India, 1st T20I: India v South Africa at Dharamsala, Oct 2, 2015

South Africa tour of India, 2nd T20I: India v South Africa at Cuttack, Oct 5, 2015

error fetching data for match with id: 903591

Sri Lanka tour of India and Bangladesh, 1st T20I: India v Sri Lanka at Pune, Feb 9, 2016

Sri Lanka tour of India and Bangladesh, 2nd T20I: India v Sri Lanka at Ranchi, Feb 12, 2016

Sri Lanka tour of India and Bangladesh, 3rd T20I: India v Sri Lanka at Visakhapatnam, Feb 14, 2016

Asia Cup, 4th Match: India v Pakistan at Dhaka, Feb 27, 2016

Asia Cup, 7th Match: India v Sri Lanka at Dhaka, Mar 1, 2016

Asia Cup, 9th Match: India v United Arab Emirates at Dhaka, Mar 3, 2016

World T20, 13th Match, Super 10 Group 2: India v New Zealand at Nagpur, Mar 15, 2016

World T20, 19th Match, Super 10 Group 2: India v Pakistan at Kolkata, Mar 19, 2016

World T20, 25th Match, Super 10 Group 2: India v Bangladesh at Bengaluru, Mar 23, 2016

World T20, 31st Match, Super 10 Group 2: India v Australia at Mohali, Mar 27, 2
016
World T20, 2nd Semi-Final: India v West Indies at Mumbai, Mar 31, 2016
England tour of India, 1st T20I: India v England at Kanpur, Jan 26, 2017
England tour of India, 2nd T20I: India v England at Nagpur, Jan 29, 2017
England tour of India, 3rd T20I: India v England at Bengaluru, Feb 1, 2017
India tour of South Africa, Only T20I: South Africa v India at Johannesburg, De
c 1, 2006
ICC World Twenty20, 21st Match, Group E: England v India at Durban, Sep 19, 200
7
ICC World Twenty20, 24th Match, Group E: South Africa v India at Durban, Sep 2
0, 2007
ICC World Twenty20, 2nd Semi Final: Australia v India at Durban, Sep 22, 2007
India tour of Australia, Only T20I: Australia v India at Melbourne, Feb 1, 2008
India tour of Sri Lanka, Only T20I: Sri Lanka v India at Colombo (RPS), Feb 10,
2009
India tour of New Zealand, 1st T20I: New Zealand v India at Christchurch, Feb 2
5, 2009
India tour of New Zealand, 2nd T20I: New Zealand v India at Wellington, Feb 27,
2009
ICC World Twenty20, 4th Match, Group A: Bangladesh v India at Nottingham, Jun
6, 2009
ICC World Twenty20, 20th Match, Group E: England v India at Lord's, Jun 14, 200
9
ICC World Twenty20, 3rd Match, Group C: Afghanistan v India at Gros Islet, May
1, 2010
ICC World Twenty20, 15th Match, Group F: Australia v India at Bridgetown, May
7, 2010
ICC World Twenty20, 19th Match, Group F: West Indies v India at Bridgetown, May
9, 2010
India tour of Zimbabwe, 1st T20I: Zimbabwe v India at Harare, Jun 12, 2010
India tour of Zimbabwe, 2nd T20I: Zimbabwe v India at Harare, Jun 13, 2010
India tour of South Africa, Only T20I: South Africa v India at Durban, Jan 9, 2
011
India tour of West Indies, Only T20I: West Indies v India at Port of Spain, Jun
4, 2011
India tour of England, Only T20I: England v India at Manchester, Aug 31, 2011
India tour of Australia, 1st T20I: Australia v India at Sydney, Feb 1, 2012
India tour of Australia, 2nd T20I: Australia v India at Melbourne, Feb 3, 2012
India tour of South Africa, Only T20I: South Africa v India at Johannesburg, Ma
r 30, 2012
India tour of Sri Lanka, Only T20I: Sri Lanka v India at Pallekele, Aug 7, 2012
ICC World Twenty20, 3rd Match, Group A: Afghanistan v India at Colombo (RPS), S
ep 19, 2012
ICC World Twenty20, 10th Match, Group A: England v India at Colombo (RPS), Sep
23, 2012
ICC World Twenty20, 16th Match, Group 2: Australia v India at Colombo (RPS), Se
p 28, 2012
World T20, 24th Match, Group 2: Bangladesh v India at Dhaka, Mar 28, 2014
World T20, 28th Match, Group 2: Australia v India at Dhaka, Mar 30, 2014
India tour of England, Only T20I: England v India at Birmingham, Sep 7, 2014
India tour of Zimbabwe, 1st T20I: Zimbabwe v India at Harare, Jul 17, 2015
India tour of Zimbabwe, 2nd T20I: Zimbabwe v India at Harare, Jul 19, 2015
India tour of Australia, 1st T20I: Australia v India at Adelaide, Jan 26, 2016
India tour of Australia, 2nd T20I: Australia v India at Melbourne, Jan 29, 2016
India tour of Australia, 3rd T20I: Australia v India at Sydney, Jan 31, 2016
Asia Cup, 1st Match: Bangladesh v India at Dhaka, Feb 24, 2016
Asia Cup, Final: Bangladesh v India at Dhaka, Mar 6, 2016

India tour of Zimbabwe, 1st T20I: Zimbabwe v India at Harare, Jun 18, 2016
India tour of Zimbabwe, 2nd T20I: Zimbabwe v India at Harare, Jun 20, 2016
India tour of Zimbabwe, 3rd T20I: Zimbabwe v India at Harare, Jun 22, 2016
India tour of West Indies and United States of America, 1st T20I: India v West
Indies at Lauderhill, Aug 27, 2016
India tour of West Indies and United States of America, 2nd T20I: India v West
Indies at Lauderhill, Aug 28, 2016

Creating a table for ID and Description

```
In [120]: Match_List = pd.DataFrame(Matchid_list)
          Match_List['Description'] = Description_list
```

In [122]: Match_List

Out[122]:

	MATCH_ID	Description
0	287859	ICC World Twenty20, 7th Match, Group D: India ...
1	287862	ICC World Twenty20, 10th Match, Group D: India...
2	287865	ICC World Twenty20, 13th Match, Group E: India...
3	287879	ICC World Twenty20, Final: India v Pakistan at...
4	297800	Australia tour of India, Only T20I: India v Au...
5	356001	ICC World Twenty20, 12th Match, Group A: India...
6	356006	ICC World Twenty20, 16th Match, Group E: India...
7	356014	ICC World Twenty20, 24th Match, Group E: India...
8	430884	Sri Lanka tour of India, 1st T20I: India v Sri...
9	430885	Sri Lanka tour of India, 2nd T20I: India v Sri...
10	412682	ICC World Twenty20, 5th Match, Group C: India ...
11	412699	ICC World Twenty20, 23rd Match, Group F: India...
12	521217	England tour of India, Only T20I: India v Engl...
13	565819	Error
14	565820	New Zealand tour of India, 2nd T20I: India v N...
15	533291	ICC World Twenty20, 20th Match, Group 2: India...
16	533295	ICC World Twenty20, 24th Match, Group 2: India...
17	565810	England tour of India, 1st T20I: India v Engla...
18	565811	England tour of India, 2nd T20I: India v Engla...
19	589306	Pakistan tour of India, 1st T20I: India v Paki...
20	589307	Pakistan tour of India, 2nd T20I: India v Paki...
21	647247	Australia tour of India, Only T20I: India v Au...
22	682921	World T20, 13th Match, Group 2: India v Pakist...
23	682929	World T20, 17th Match, Group 2: India v West I...
24	682963	World T20, 2nd Semi-Final: India v South Afric...
25	682965	World T20, Final: India v Sri Lanka at Dhaka, ...
26	903587	South Africa tour of India, 1st T20I: India v ...
27	903589	South Africa tour of India, 2nd T20I: India v ...
28	903591	Error
29	963697	Sri Lanka tour of India and Bangladesh, 1st T2...
...	...	...
53	412679	ICC World Twenty20, 3rd Match, Group C: Afghan...
54	412691	ICC World Twenty20, 15th Match, Group F: Austr...
55	412695	ICC World Twenty20, 19th Match, Group F: West ...
56	452153	India tour of Zimbabwe, 1st T20I: Zimbabwe v I...
57	452154	India tour of Zimbabwe, 2nd T20I: Zimbabwe v I...

	MATCH_ID	Description
58	463149	India tour of South Africa, Only T20I: South A...
59	489220	India tour of West Indies, Only T20I: West Ind...
60	474476	India tour of England, Only T20I: England v In...
61	518954	India tour of Australia, 1st T20I: Australia v...
62	518955	India tour of Australia, 2nd T20I: Australia v...
63	556252	India tour of South Africa, Only T20I: South A...
64	564786	India tour of Sri Lanka, Only T20I: Sri Lanka ...
65	533274	ICC World Twenty20, 3rd Match, Group A: Afghan...
66	533281	ICC World Twenty20, 10th Match, Group A: Engla...
67	533287	ICC World Twenty20, 16th Match, Group 2: Austr...
68	682943	World T20, 24th Match, Group 2: Bangladesh v I...
69	682951	World T20, 28th Match, Group 2: Australia v In...
70	667731	India tour of England, Only T20I: England v In...
71	885969	India tour of Zimbabwe, 1st T20I: Zimbabwe v I...
72	885971	India tour of Zimbabwe, 2nd T20I: Zimbabwe v I...
73	895817	India tour of Australia, 1st T20I: Australia v...
74	895819	India tour of Australia, 2nd T20I: Australia v...
75	895821	India tour of Australia, 3rd T20I: Australia v...
76	966745	Asia Cup, 1st Match: Bangladesh v India at Dha...
77	966765	Asia Cup, Final: Bangladesh v India at Dhaka, ...
78	1007655	India tour of Zimbabwe, 1st T20I: Zimbabwe v I...
79	1007657	India tour of Zimbabwe, 2nd T20I: Zimbabwe v I...
80	1007659	India tour of Zimbabwe, 3rd T20I: Zimbabwe v I...
81	1041615	India tour of West Indies and United States of...
82	1041617	India tour of West Indies and United States of...

83 rows × 2 columns

WEB SCRAPING USING PYTHON

```
In [265]: import bs4
```

```
In [266]: from urllib.request import urlopen as uReq
```

```
In [267]: from bs4 import BeautifulSoup as soup
```

```
In [268]: my_url = 'http://stats.espncricinfo.com/ci/engine/stats/index.html?class=3;filter=advanced;orderby=runs;team=6;template=results;type=batting'
```

Grabbing the Webpage

```
In [269]:  uSite = uReq(my_url)
```

```
In [270]:  page = uSite.read()
```

```
In [271]:  uSite.close()
```

HTML parsing

```
In [272]:  page_cricketers = soup(page, "html.parser")
```

```
In [273]:  Boxes = page_cricketers.findAll("tr", {"class":"data1"})
```

```
In [274]:  Box= Boxes[0]
```

```
In [275]: filename = "T20info.csv"
          f = open(filename,"w")

          headers = "Name, Matches, Runs, Centuries, Half Centuries\n"

          f.write(headers)

          for Box in Boxes:
              Name = Box.td.a.text
              Matches = Box.findAll('td')[2].text
              Runs = Box.findAll('td')[5].b.text
              Centuries = Box.findAll('td')[10].text
              Half_Centuries = Box.findAll('td')[11].text


              print("Name: " + Name)
              print("Matches: " + Matches)
              print("Runs: " + Runs)
              print("Centuries: " + Centuries)
              print("Half_Centuries: " + Half_Centuries)

              f.write(Name + "," +Matches + "," +Runs + "," +Centuries + "," +Half_Centuri
          es +"\n")
          f.close()
```

Name: RG Sharma
Matches: 90
Runs: 2237
Centuries: 4
Half_Centuries: 15
Name: V Kohli
Matches: 65
Runs: 2167
Centuries: 0
Half_Centuries: 19
Name: SK Raina
Matches: 78
Runs: 1605
Centuries: 1
Half_Centuries: 5
Name: MS Dhoni
Matches: 93
Runs: 1487
Centuries: 0
Half_Centuries: 2
Name: S Dhawan
Matches: 46
Runs: 1232
Centuries: 0
Half_Centuries: 9
Name: Yuvraj Singh
Matches: 58
Runs: 1177
Centuries: 0
Half_Centuries: 8
Name: G Gambhir
Matches: 37
Runs: 932
Centuries: 0
Half_Centuries: 7
Name: KL Rahul
Matches: 25
Runs: 782
Centuries: 2
Half_Centuries: 4
Name: MK Pandey
Matches: 28
Runs: 538
Centuries: 0
Half_Centuries: 2
Name: V Sehwag
Matches: 19
Runs: 394
Centuries: 0
Half_Centuries: 2
Name: AM Rahane
Matches: 20
Runs: 375
Centuries: 0
Half_Centuries: 1
Name: KD Karthik
Matches: 26
Runs: 352

Centuries: 0
Half_Centuries: 0
Name: HH Pandya
Matches: 35
Runs: 271
Centuries: 0
Half_Centuries: 0
Name: RV Uthappa
Matches: 13
Runs: 249
Centuries: 0
Half_Centuries: 1
Name: YK Pathan
Matches: 22
Runs: 236
Centuries: 0
Half_Centuries: 0
Name: IK Pathan
Matches: 24
Runs: 172
Centuries: 0
Half_Centuries: 0
Name: M Vijay
Matches: 9
Runs: 169
Centuries: 0
Half_Centuries: 0
Name: RR Pant
Matches: 10
Runs: 157
Centuries: 0
Half_Centuries: 1
Name: R Ashwin
Matches: 46
Runs: 123
Centuries: 0
Half_Centuries: 0
Name: KM Jadhav
Matches: 9
Runs: 122
Centuries: 0
Half_Centuries: 1
Name: RA Jadeja
Matches: 40
Runs: 116
Centuries: 0
Half_Centuries: 0
Name: Harbhajan Singh
Matches: 28
Runs: 108
Centuries: 0
Half_Centuries: 0
Name: Mandeep Singh
Matches: 3
Runs: 87
Centuries: 0
Half_Centuries: 1
Name: SS Iyer

Matches: 6
Runs: 83
Centuries: 0
Half_Centuries: 0
Name: AR Patel
Matches: 11
Runs: 68
Centuries: 0
Half_Centuries: 0
Name: S Badrinath
Matches: 1
Runs: 43
Centuries: 0
Half_Centuries: 0
Name: AT Rayudu
Matches: 6
Runs: 42
Centuries: 0
Half_Centuries: 0
Name: D Mongia
Matches: 1
Runs: 38
Centuries: 0
Half_Centuries: 0
Name: PA Patel
Matches: 2
Runs: 36
Centuries: 0
Half_Centuries: 0
Name: STR Binny
Matches: 3
Runs: 35
Centuries: 0
Half_Centuries: 0
Name: R Dravid
Matches: 1
Runs: 31
Centuries: 0
Half_Centuries: 0
Name: A Nehra
Matches: 27
Runs: 28
Centuries: 0
Half_Centuries: 0
Name: KH Pandya
Matches: 6
Runs: 23
Centuries: 0
Half_Centuries: 0
Name: AB Dinda
Matches: 9
Runs: 22
Centuries: 0
Half_Centuries: 0
Name: B Kumar
Matches: 34
Runs: 22
Centuries: 0

Half_Centuries: 0
Name: Kuldeep Yadav
Matches: 17
Runs: 20
Centuries: 0
Half_Centuries: 0
Name: S Sreesanth
Matches: 10
Runs: 20
Centuries: 0
Half_Centuries: 0
Name: SV Samson
Matches: 1
Runs: 19
Centuries: 0
Half_Centuries: 0
Name: V Shankar
Matches: 5
Runs: 17
Centuries: 0
Half_Centuries: 0
Name: AB Agarkar
Matches: 4
Runs: 15
Centuries: 0
Half_Centuries: 0
Name: MK Tiwary
Matches: 3
Runs: 15
Centuries: 0
Half_Centuries: 0
Name: Z Khan
Matches: 17
Runs: 13
Centuries: 0
Half_Centuries: 0
Name: NV Ojha
Matches: 2
Runs: 12
Centuries: 0
Half_Centuries: 0
Name: PP Ojha
Matches: 6
Runs: 10
Centuries: 0
Half_Centuries: 0
Name: SR Tendulkar
Matches: 1
Runs: 10
Centuries: 0
Half_Centuries: 0
Name: JJ Bumrah
Matches: 40
Runs: 8
Centuries: 0
Half_Centuries: 0
Name: I Sharma
Matches: 14

```
Runs: 8
Centuries: 0
Half_Centuries: 0
Name: P Kumar
Matches: 10
Runs: 7
Centuries: 0
Half_Centuries: 0
Name: Parvez Rasool
Matches: 1
Runs: 5
Centuries: 0
Half_Centuries: 0
Name: YS Chahal
Matches: 27
Runs: 3
Centuries: 0
Half_Centuries: 0
```

Conclusion

We have thus created a databse for t20 cricket matches by using csv file and Web API scraper and created one for t20 indian cricket batsmen with python web scraping.

Citation

https://www.youtube.com/watch?v=XQgXKtPSzUI&t=961s (https://www.youtube.com/watch?v=XQgXKtPSzUI&t=961s)

https://github.com/nikbearbrown/INFO_6210/tree/master/Movie_DB_Example (https://github.com/nikbearbrown/INFO_6210/tree/master/Movie_DB_Example)

License

Percentage Contribution

Amey Babar = 50%

responsible for writing half of the code and data search. Major part in python web scraping

Aniruddha Joshi = 50%

responsible for writing half of the code and data search. Major part in Web API scraping and CSV file scraping.