```
#colab link :
https://colab.research.google.com/drive/1CF6M1hPpnVwjhJmhpcgFs2VzRuVe4
i6u
#Name: Atharva Kangralkar
#Branch: CS(AIML) - A
#Roll no : 54
#Batch : 2
#LAB ASSIGNMENT 1: PANDAS FUNCTION

#1 What are Pandas?
#pandas is a Python package providing fast, flexible, and expressive
data structures designed to make working with “relational” or
“labeled” data both easy and intuitive.
#It aims to be the fundamental high-level building block for doing
practical, real-world data analysis in Python.
#Additionally, it has the broader goal of becoming the most powerful
and flexible open source data analysis/manipulation tool available in
any language.

#2 Write details of IPL match dataset. What is your insight about the
data?
#It contains 17 columns and 636 rows.
#The dataset contains information about match outcomes, cities, and
teams.
#A trend shows that Mumbai Indians have the most wins
#Seasons between 2008 to 2017 have been covered

#3 How to import excel data in a notebook?
# Returns a DataFrame
#pd.read_excel("path_to_file.xls", sheet_name="Sheet1")

#4 Import two libraries (pandas, numpy)
import pandas as pd
import numpy as np

#5 Read the dataset and open the file
df = pd.read_csv("/content/ipl_dataset.csv")
df
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 636,\n  \"fields\": [\
n    {\n      \"column\": \"id\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 183,\n        \"min\": 1,\n
\"max\": 636,\n        \"num_unique_values\": 636,\n
\"samples\": [\n          275,\n          632,\n          83\
n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"season\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 2,\n        \"min\": 2008,\n        \"max\": 2017,\n
\"num_unique_values\": 10,\n        \"samples\": [\n          2015,\n
2008,\n          2012\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
```

```
\"city\",\n      \"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 30,\n        \"samples\": [\n
\"Ranchi\",\n          \"Centurion\",\n          \"Dharamsala\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"date\",\n      \"properties\":
{\n        \"dtype\": \"object\",\n        \"num_unique_values\":
450,\n        \"samples\": [\n          \"2016-04-14\",\n
\"2016-05-21\",\n          \"2009-05-19\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"team1\",\n      \"properties\": {\
n        \"dtype\": \"category\",\n        \"num_unique_values\": 14,\
n        \"samples\": [\n          \"Rajasthan Royals\",\n
\"Kochi Tuskers Kerala\",\n          \"Sunrisers Hyderabad\"\
n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"team2\",\n      \"properties\": {\n        \"dtype\": \"category\",\
n        \"num_unique_values\": 14,\n        \"samples\": [\n
\"Chennai Super Kings\",\n          \"Pune Warriors\",\n
\"Royal Challengers Bangalore\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"toss_winner\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 14,\n        \"samples\": [\n
\"Rajasthan Royals\",\n          \"Kochi Tuskers Kerala\",\n
\"Royal Challengers Bangalore\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"toss_decision\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"bat\",\
n          \"field\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"result\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 3,\n        \"samples\":
[\n          \"normal\",\n          \"tie\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"dl_applied\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
0,\n        \"min\": 0,\n        \"max\": 1,\n
\"num_unique_values\": 2,\n        \"samples\": [\n          1,\n
0\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"winner\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 14,\n
\"samples\": [\n          \"Rajasthan Royals\",\n          \"Pune
Warriors\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"win_by_runs\",\n      \"properties\": {\n        \"dtype\":
\"number\",\n        \"std\": 23,\n        \"min\": 0,\n
\"max\": 146,\n        \"num_unique_values\": 86,\n
```

\"samples\": [\n   60,\n   35\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"win_by_wickets\",\n
\"properties\": {\n   \"dtype\": \"number\",\n   \"std\":
3,\n   \"min\": 0,\n   \"max\": 10,\n
\"num_unique_values\": 11,\n   \"samples\": [\n   4,\n
0\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"player_of_match\",\n   \"properties\": {\n   \"dtype\":
\"category\",\n   \"num_unique_values\": 201,\n
\"samples\": [\n   \"DP Nannes\",\n   \"BA Stokes\"\n
],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n
}\n   },\n   {\n   \"column\": \"venue\",\n   \"properties\":
{\n   \"dtype\": \"category\",\n   \"num_unique_values\":
35,\n   \"samples\": [\n   \"Himachal Pradesh Cricket
Association Stadium\",\n   \"Dr DY Patil Sports Academy\"\n
],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n
}\n   },\n   {\n   \"column\": \"umpire1\",\n
\"properties\": {\n   \"dtype\": \"category\",\n
\"num_unique_values\": 44,\n   \"samples\": [\n   \"C
Shamshuddin\",\n   \"S Asnani\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"umpire2\",\n   \"properties\":
{\n   \"dtype\": \"category\",\n   \"num_unique_values\":
45,\n   \"samples\": [\n   \"K Srinath\",\n   \"SD
Ranade\"\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"umpire3\",\n   \"properties\": {\n   \"dtype\": \"number\",\
n   \"std\": null,\n   \"min\": null,\n   \"max\":
null,\n   \"num_unique_values\": 0,\n   \"samples\": [],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   }\n]\n}","type":"dataframe","variable_name":"df"}

```
#6 Show the type of data frame.
df.dtypes
```

```
id                  int64
season              int64
city               object
date               object
team1              object
team2              object
toss_winner        object
toss_decision      object
result             object
dl_applied          int64
winner             object
win_by_runs         int64
win_by_wickets      int64
player_of_match    object
```

```
venue                object
umpire1              object
umpire2              object
umpire3             float64
dtype: object
```

```
#7 Show the First 2 lines or last 2 lines of the dataset ?
# df.head(2)
df.tail(2)
```


```
{"repr_error":"0","type":"dataframe"}
```


```
#8 What is the shape of the dataset ?
df.shape
```

```
(636, 18)
```

```
#9 What is the total preview of the dataset ?
df.head(10)
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 636,\n  \"fields\": [\
n    {\n      \"column\": \"id\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 183,\n        \"min\": 1,\n
\"max\": 636,\n        \"num_unique_values\": 636,\n
\"samples\": [\n          275,\n          632,\n          83\
n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"season\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 2,\n        \"min\": 2008,\n        \"max\": 2017,\n
\"num_unique_values\": 10,\n        \"samples\": [\n          2015,\n
2008,\n          2012\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"city\",\n      \"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 30,\n        \"samples\": [\n
\"Ranchi\",\n          \"Centurion\",\n          \"Dharamsala\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"date\",\n      \"properties\":
{\n        \"dtype\": \"object\",\n        \"num_unique_values\":
450,\n        \"samples\": [\n          \"2016-04-14\",\n
\"2016-05-21\",\n          \"2009-05-19\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"team1\",\n      \"properties\": {\
n        \"dtype\": \"category\",\n        \"num_unique_values\": 14,\
n        \"samples\": [\n          \"Rajasthan Royals\",\n
\"Kochi Tuskers Kerala\",\n          \"Sunrisers Hyderabad\"\
n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"team2\",\n      \"properties\": {\n        \"dtype\": \"category\",\
n        \"num_unique_values\": 14,\n        \"samples\": [\n
\"Chennai Super Kings\",\n          \"Pune Warriors\",\n
\"Royal Challengers Bangalore\"\n        ],\n
```

```
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"toss_winner\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 14,\n        \"samples\": [\n
\"Rajasthan Royals\",\n          \"Kochi Tuskers Kerala\",\n
\"Royal Challengers Bangalore\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"toss_decision\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"bat\",\
n          \"field\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"result\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 3,\n        \"samples\":
[\n          \"normal\",\n          \"tie\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"dl_applied\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
0,\n        \"min\": 0,\n        \"max\": 1,\n
\"num_unique_values\": 2,\n        \"samples\": [\n          1,\n
0\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"winner\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 14,\n
\"samples\": [\n          \"Rajasthan Royals\",\n          \"Pune
Warriors\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"win_by_runs\",\n      \"properties\": {\n        \"dtype\":
\"number\",\n        \"std\": 23,\n        \"min\": 0,\n
\"max\": 146,\n        \"num_unique_values\": 86,\n
\"samples\": [\n          60,\n          35\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"win_by_wickets\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
3,\n        \"min\": 0,\n        \"max\": 10,\n
\"num_unique_values\": 11,\n        \"samples\": [\n          4,\n
0\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"player_of_match\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 201,\n
\"samples\": [\n          \"DP Nannes\",\n          \"BA Stokes\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"venue\",\n      \"properties\":
{\n        \"dtype\": \"category\",\n        \"num_unique_values\":
35,\n        \"samples\": [\n          \"Himachal Pradesh Cricket
Association Stadium\",\n          \"Dr DY Patil Sports Academy\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"umpire1\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
```

\"num_unique_values\": 44,\n   \"samples\": [\n   \"C Shamshuddin\",\n   \"S Asnani\"\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"umpire2\",\n   \"properties\": {\n   \"dtype\": \"category\",\n   \"num_unique_values\": 45,\n   \"samples\": [\n   \"K Srinath\",\n   \"SD Ranade\"\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"umpire3\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": null,\n   \"min\": null,\n   \"max\": null,\n   \"num_unique_values\": 0,\n   \"samples\": [],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   }\n]\n}","type":"dataframe","variable_name":"df"}

```
#10 Show the preview of the numerical data column?
df.select_dtypes(include='number')
```

{"summary":"{\n \"name\": \"df\",\n \"rows\": 636,\n \"fields\": [\n   {\n   \"column\": \"id\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": 183,\n   \"min\": 1,\n   \"max\": 636,\n   \"num_unique_values\": 636,\n   \"samples\": [\n   275,\n   632,\n   83\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"season\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": 2,\n   \"min\": 2008,\n   \"max\": 2017,\n   \"num_unique_values\": 10,\n   \"samples\": [\n   2015,\n   2008,\n   2012\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"dl_applied\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": 0,\n   \"min\": 0,\n   \"max\": 1,\n   \"num_unique_values\": 2,\n   \"samples\": [\n   1,\n   0\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"win_by_runs\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": 23,\n   \"min\": 0,\n   \"max\": 146,\n   \"num_unique_values\": 86,\n   \"samples\": [\n   60,\n   35\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"win_by_wickets\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": 3,\n   \"min\": 0,\n   \"max\": 10,\n   \"num_unique_values\": 11,\n   \"samples\": [\n   4,\n   0\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"umpire3\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": null,\n   \"min\": null,\n   \"max\": null,\n   \"num_unique_values\": 0,\n   \"samples\": [],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   }\n]\n}","type":"dataframe"}

```
#12 Find out one column from the dataset?
df['season']
```

```
0      2017
1      2017
2      2017
3      2017
4      2017
       ...
631    2016
632    2016
633    2016
634    2016
635    2016
Name: season, Length: 636, dtype: int64
```

```
#13 Find out type of data types of above data Q.12
df['season'].dtypes
```

```
dtype('int64')
```

```
#14 Find out shape of above data Q.12
df['season'].shape
```

```
(636,)
```

```
#15 Find out multiple columns from the dataframe?
df[['season','city' ,'result']]
```


```
{"summary":"{\n  \"name\": \"df[['season','city' ,'result']]\",\n  \"rows\": 636,\n  \"fields\": [\n    {\n      \"column\": \"season\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 2,\n        \"min\": 2008,\n        \"max\": 2017,\n        \"num_unique_values\": 10,\n        \"samples\": [\n          2015,\n          2008,\n          2012\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"city\",\n      \"properties\": {\n        \"dtype\": \"category\",\n        \"num_unique_values\": 30,\n        \"samples\": [\n          \"Ranchi\",\n          \"Centurion\",\n          \"Dharamsala\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"result\",\n      \"properties\": {\n        \"dtype\": \"category\",\n        \"num_unique_values\": 3,\n        \"samples\": [\n          \"normal\",\n          \"tie\",\n          \"no result\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    }\n  ]\n}","type":"dataframe"}
```


```
#16 Find out the type of above data of Q.15?
df[['season','city' ,'result']].dtypes
```

```
season     int64
city      object
result    object
dtype: object
```

```
#17 Find out the shape of the above data of Q.15?
df[['season','city' ,'result']].shape
```

```
(636, 3)
```

```
#18 Find out the fifth row of the dataframe?
df.iloc[4]
```

```
id                                              5
season                                       2017
city                                      Bangalore
date                                     2017-04-08
team1               Royal Challengers Bangalore
team2                           Delhi Daredevils
toss_winner         Royal Challengers Bangalore
toss_decision                               bat
result                                   normal
dl_applied                                    0
winner              Royal Challengers Bangalore
win_by_runs                                  15
win_by_wickets                                0
player_of_match                        KM Jadhav
venue                       M Chinnaswamy Stadium
umpire1                                     NaN
umpire2                                     NaN
umpire3                                     NaN
Name: 4, dtype: object
```

```
#19  Find out the type of above data of Q.18?
df.iloc[5].dtypes
```

```
dtype('O')
```

```
#20  Find out multiple rows (24   to 29) from the dataset?
df.iloc[23:29]
```

```
{"repr_error":"0","type":"dataframe"}
```

```
#21  Find out data type of above Q.20
df.iloc[24:29].dtypes
```

```
id                   int64
season               int64
city                object
date                object
team1               object
```

```
team2                object
toss_winner          object
toss_decision        object
result               object
dl_applied            int64
winner               object
win_by_runs           int64
win_by_wickets        int64
player_of_match      object
venue                object
umpire1              object
umpire2              object
umpire3             float64
dtype: object
```

```
#22  Find out alternate matches data from dataframe?
df.iloc[::2]
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 318,\n  \"fields\": [\
n    {\n      \"column\": \"id\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 183,\n        \"min\": 1,\n
\"max\": 635,\n        \"num_unique_values\": 318,\n
\"samples\": [\n          147,\n          555,\n          51\
n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"season\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 2,\n        \"min\": 2008,\n        \"max\": 2017,\n
\"num_unique_values\": 10,\n        \"samples\": [\n          2015,\n
2008,\n          2012\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"city\",\n      \"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 27,\n        \"samples\": [\n
\"Kanpur\",\n          \"Port Elizabeth\",\n          \"Jaipur\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"date\",\n      \"properties\":
{\n        \"dtype\": \"object\",\n        \"num_unique_values\":
313,\n        \"samples\": [\n          \"2008-05-11\",\n
\"2017-04-17\",\n          \"2013-04-27\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"team1\",\n      \"properties\": {\
n        \"dtype\": \"category\",\n        \"num_unique_values\": 14,\
n        \"samples\": [\n          \"Deccan Chargers\",\n
\"Kochi Tuskers Kerala\",\n          \"Sunrisers Hyderabad\"\
n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"team2\",\n      \"properties\": {\n        \"dtype\": \"category\",\
n        \"num_unique_values\": 14,\n        \"samples\": [\n
\"Chennai Super Kings\",\n          \"Pune Warriors\",\n
\"Royal Challengers Bangalore\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
```

```
n    },\n    {\n      \"column\": \"toss_winner\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 14,\n        \"samples\": [\n          \"Deccan
Chargers\",\n          \"Kochi Tuskers Kerala\",\n          \"Royal
Challengers Bangalore\"\n        ],\n        \"semantic_type\": \"\",\
n        \"description\": \"\"\n      }\n    },\n    {\n
\"column\": \"toss_decision\",\n      \"properties\": {\n
\"dtype\": \"category\",\n        \"num_unique_values\": 2,\n
\"samples\": [\n          \"bat\",\n          \"field\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"result\",\n      \"properties\":
{\n        \"dtype\": \"category\",\n        \"num_unique_values\":
3,\n        \"samples\": [\n          \"normal\",\n          \"no
result\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"dl_applied\",\n      \"properties\": {\n        \"dtype\":
\"number\",\n        \"std\": 0,\n        \"min\": 0,\n
\"max\": 1,\n        \"num_unique_values\": 2,\n        \"samples\":
[\n          1,\n          0\n        ],\n        \"semantic_type\":
\"\",\n        \"description\": \"\"\n      }\n    },\n    {\n
\"column\": \"winner\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 14,\n
\"samples\": [\n          \"Rajasthan Royals\",\n          \"Pune
Warriors\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"win_by_runs\",\n      \"properties\": {\n        \"dtype\":
\"number\",\n        \"std\": 23,\n        \"min\": 0,\n
\"max\": 138,\n        \"num_unique_values\": 65,\n
\"samples\": [\n          48,\n          32\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"win_by_wickets\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
3,\n        \"min\": 0,\n        \"max\": 10,\n
\"num_unique_values\": 10,\n        \"samples\": [\n          9,\n
10\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"player_of_match\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 139,\n
\"samples\": [\n          \"Q de Kock\",\n          \"MK Pandey\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"venue\",\n      \"properties\":
{\n        \"dtype\": \"category\",\n        \"num_unique_values\":
32,\n        \"samples\": [\n          \"Sharjah Cricket Stadium\",\n
\"St George's Park\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"umpire1\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 42,\n
\"samples\": [\n          \"SS Hazare\",\n          \"Aleem Dar\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
```

}\n   },\n   {\n   \"column\": \"umpire2\",\n
\"properties\": {\n   \"dtype\": \"category\",\n
\"num_unique_values\": 40,\n   \"samples\": [\n   \"IL
Howell\",\n   \"K Hariharan\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"umpire3\",\n   \"properties\":
{\n   \"dtype\": \"number\",\n   \"std\": null,\n
\"min\": null,\n   \"max\": null,\n   \"num_unique_values\":
0,\n   \"samples\": [],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   }\n]\n}","type":"dataframe"}

```
#23  Find out rows 2, 8, 10 from dataframe?
df.iloc[[2,8,10]]
```

{"repr_error":"0","type":"dataframe"}

```
#24  Find out column (5,6,10) from the dataframe using iloc ?
df.iloc[:,[5,6,10]]
```

{"summary":"{\n \"name\": \"df\",\n \"rows\": 636,\n \"fields\": [\
n   {\n   \"column\": \"team2\",\n   \"properties\": {\n
\"dtype\": \"category\",\n   \"num_unique_values\": 14,\n
\"samples\": [\n   \"Chennai Super Kings\",\n   \"Pune
Warriors\",\n   \"Royal Challengers Bangalore\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"toss_winner\",\n
\"properties\": {\n   \"dtype\": \"category\",\n
\"num_unique_values\": 14,\n   \"samples\": [\n
\"Rajasthan Royals\",\n   \"Kochi Tuskers Kerala\",\n
\"Royal Challengers Bangalore\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"winner\",\n   \"properties\":
{\n   \"dtype\": \"category\",\n   \"num_unique_values\":
14,\n   \"samples\": [\n   \"Rajasthan Royals\",\n
\"Pune Warriors\",\n   \"Sunrisers Hyderabad\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   }\n]\n}","type":"dataframe"}

```
#25  Find out matches that take place in Pune city?
df[df['city'] == 'Pune']
```

{"repr_error":"0","type":"dataframe"}

```
#26  What is the data type of Q.25 output?
df[df['city'] == 'Pune'].dtypes
```

```
id                   int64
season               int64
city                object
date                object
team1               object
```

```
team2                object
toss_winner          object
toss_decision        object
result               object
dl_applied            int64
winner               object
win_by_runs           int64
win_by_wickets        int64
player_of_match      object
venue                object
umpire1              object
umpire2              object
umpire3             float64
dtype: object
```

```
#27  Find out the shape of Q.25 output data?
df[df['city'] == 'Pune'].shape
```

```
(32, 18)
```

```
#28  Find out 2013 seasons matches from the dataset?
df[df['season'] == 2013]
```

```
{"repr_error":"0","type":"dataframe"}
```

```
#29  Find out matches played in Pune city after 2012?
df[(df['city'] == 'Pune') & (df['season'] > 2012)]
```

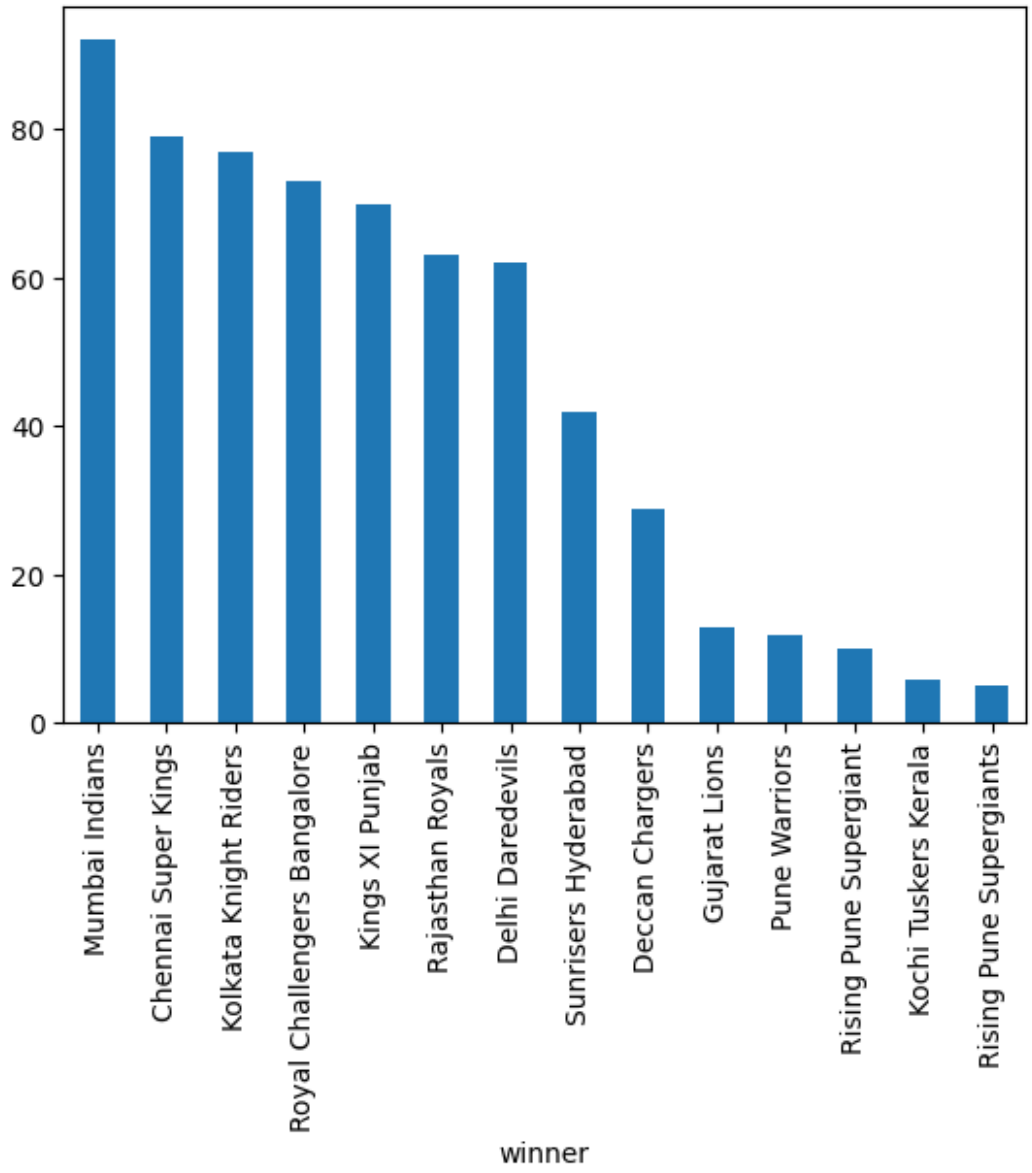
```
{"repr_error":"0","type":"dataframe"}
```

```
#30  Find out how many matches are won by an individual team?
#df['winner'].value_counts()
df.groupby(['winner']).size()
```

```
winner
Chennai Super Kings             79
Deccan Chargers                 29
Delhi Daredevils                62
Gujarat Lions                   13
Kings XI Punjab                 70
Kochi Tuskers Kerala             6
Kolkata Knight Riders           77
Mumbai Indians                  92
Pune Warriors                   12
Rajasthan Royals                63
Rising Pune Supergiant          10
Rising Pune Supergiants          5
Royal Challengers Bangalore     73
Sunrisers Hyderabad             42
dtype: int64
```

```
#31-A Plot graph to show how many matches won by an individual team?
df['winner'].value_counts().plot(kind='bar')
```

```
<Axes: xlabel='winner'>
```




```
#31-B Show the relation between data frame columns. Use a match
dataset.
df.corr(numeric_only = True)
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 6,\n  \"fields\": [\n
{\n      \"column\": \"id\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 0.4446674454811242,\n
\"min\": -0.015509761243645356,\n        \"max\": 1.0,\n
```

\"num_unique_values\": 5,\n   \"samples\": [\n
0.47108676169661295,\n   -0.015509761243645356,\n
0.024280698612303364\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"season\",\n   \"properties\": {\n   \"dtype\": \"number\",\n
\"std\": 0.44643738507547415,\n   \"min\": -
0.016814675872375222,\n   \"max\": 1.0,\n
\"num_unique_values\": 5,\n   \"samples\": [\n   1.0,\n
-0.0007080041173234443,\n   0.004170478073391533\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"dl_applied\",\n
\"properties\": {\n   \"dtype\": \"number\",\n   \"std\":
0.4467904251290839,\n   \"min\": -0.011640193653903199,\n
\"max\": 1.0,\n   \"num_unique_values\": 5,\n   \"samples\":
[\n   0.004170478073391533,\n   -0.011640193653903199,\n
1.0\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"win_by_runs\",\n   \"properties\": {\n   \"dtype\":
\"number\",\n   \"std\": 0.5675469278041096,\n   \"min\": -
0.5651806422603185,\n   \"max\": 1.0,\n
\"num_unique_values\": 5,\n   \"samples\": [\n   -
0.016814675872375222,\n   -0.5651806422603185,\n   -
0.010892615458697906\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"win_by_wickets\",\n   \"properties\": {\n   \"dtype\":
\"number\",\n   \"std\": 0.5671600723004355,\n   \"min\": -
0.5651806422603185,\n   \"max\": 1.0,\n
\"num_unique_values\": 5,\n   \"samples\": [\n   -
0.0007080041173234443,\n   1.0,\n   -
0.011640193653903199\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"umpire3\",\n   \"properties\": {\n   \"dtype\": \"number\",\
n   \"std\": null,\n   \"min\": null,\n   \"max\":
null,\n   \"num_unique_values\": 0,\n   \"samples\": [],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   }\n]\n}","type":"dataframe"}

```
#32  Show the name of the season as year and dl_applied as dl. Use a
match dataset.
df.rename({'season': 'year', 'dl_applied': 'dl'}, inplace=True,
axis=1)

#33  Show id as index in data frame.
df.set_index('id', inplace=True)

#34  Show id as column of data frame instead of index.
df.reset_index(drop='True', inplace=True)
```

```
#35  Find out how many IPL matches won by an individual IPL team. Also show the data in data frame format.
df.groupby(['winner']).size()
```

```
winner
Chennai Super Kings             79
Deccan Chargers                 29
Delhi Daredevils                62
Gujarat Lions                   13
Kings XI Punjab                 70
Kochi Tuskers Kerala             6
Kolkata Knight Riders           77
Mumbai Indians                  92
Pune Warriors                   12
Rajasthan Royals                63
Rising Pune Supergiant          10
Rising Pune Supergiants          5
Royal Challengers Bangalore     73
Sunrisers Hyderabad             42
dtype: int64
```

```
#Assignment Part - 2

#1   Find out how many matches won by each individual team.
df.groupby(['winner']).size()
```

```
winner
Chennai Super Kings             79
Deccan Chargers                 29
Delhi Daredevils                62
Gujarat Lions                   13
Kings XI Punjab                 70
Kochi Tuskers Kerala             6
Kolkata Knight Riders           77
Mumbai Indians                  92
Pune Warriors                   12
Rajasthan Royals                63
Rising Pune Supergiant          10
Rising Pune Supergiants          5
Royal Challengers Bangalore     73
Sunrisers Hyderabad             42
dtype: int64
```

```
#2   What is the type of output of Q.1?
df.groupby(['winner']).size().dtypes
```

```
dtype('int64')
```

```
#3   Find out the index of Q.2 data.
matches_won = df['winner'].value_counts()
matches_won.index
```

```
Index(['Mumbai Indians', 'Chennai Super Kings', 'Kolkata Knight
Riders',
       'Royal Challengers Bangalore', 'Kings XI Punjab', 'Rajasthan
Royals',
       'Delhi Daredevils', 'Sunrisers Hyderabad', 'Deccan Chargers',
       'Gujarat Lions', 'Pune Warriors', 'Rising Pune Supergiant',
       'Kochi Tuskers Kerala', 'Rising Pune Supergiants'],
      dtype='object', name='winner')
```

```
#4   Find out values of Q.2 output data.
matches_won.values
```

```
array([92, 79, 77, 73, 70, 63, 62, 42, 29, 13, 12, 10,  6,  5])
```

```
#5   Find out the first 3 rows of Q.2 output data.
matches_won.head(3)
```

```
winner
Mumbai Indians            92
Chennai Super Kings       79
Kolkata Knight Riders     77
Name: count, dtype: int64
```

```
#6   Find out last 3 rows of Q.2 output data
matches_won.tail(3)
```

```
winner
Rising Pune Supergiant     10
Kochi Tuskers Kerala        6
Rising Pune Supergiants     5
Name: count, dtype: int64
```

```
#7   Find out Pune Warriors won how many matches?
matches_won.get('Pune Warriors', 0)
```

```
12
```

```
#8   Find out total matches played by Pune Warriors.
df[(df['team1'] == 'Pune Warriors') | (df['team2'] == 'Pune
Warriors')].shape[0]
```

```
46
```

```
#9   Find out the top 4 teams playing maximum matches in IPL?
team1_matches = df['team1'].value_counts()
team2_matches = df['team2'].value_counts()
total_matches = team1_matches.add(team2_matches,
fill_value=0).sort_values(ascending=False)
top_4_teams = total_matches.head(4)
top_4_teams
```

```
Mumbai Indians                  157
Royal Challengers Bangalore     152
Kings XI Punjab                 148
Kolkata Knight Riders           148
Name: count, dtype: int64
```

```
#10  Arrange date column data in ascending order and city in
descending order.
df.sort_values(by=['date', 'city'], ascending=[True, False])
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 636,\n  \"fields\": [\
n    {\n      \"column\": \"year\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 2,\n        \"min\": 2008,\n
\"max\": 2017,\n        \"num_unique_values\": 10,\n
\"samples\": [\n          2016,\n          2009,\n          2013\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"city\",\n      \"properties\":
{\n        \"dtype\": \"category\",\n        \"num_unique_values\":
30,\n        \"samples\": [\n          \"Sharjah\",\n
\"Bloemfontein\",\n          \"Pune\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"date\",\n      \"properties\": {\n
\"dtype\": \"object\",\n        \"num_unique_values\": 450,\n
\"samples\": [\n          \"2017-04-05\",\n          \"2017-05-13\",\n
\"2010-04-18\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"team1\",\n      \"properties\": {\n        \"dtype\": \"category\",\
n        \"num_unique_values\": 14,\n        \"samples\": [\n
\"Pune Warriors\",\n          \"Rising Pune Supergiants\",\n
\"Kolkata Knight Riders\"\n        ],\n        \"semantic_type\":
\"\",\n        \"description\": \"\"\n      }\n    },\n    {\n
\"column\": \"team2\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 14,\n
\"samples\": [\n          \"Kochi Tuskers Kerala\",\n
\"Rising Pune Supergiants\",\n          \"Royal Challengers
Bangalore\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"toss_winner\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 14,\n
\"samples\": [\n          \"Pune Warriors\",\n          \"Gujarat
Lions\",\n          \"Royal Challengers Bangalore\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"toss_decision\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"bat\",\
n          \"field\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"result\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 3,\n        \"samples\":
[\n          \"normal\",\n          \"tie\"\n        ],\n
```

\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"dl\",\n   \"properties\": {\n
\"dtype\": \"number\",\n   \"std\": 0,\n   \"min\": 0,\n
\"max\": 1,\n   \"num_unique_values\": 2,\n   \"samples\":
[\n   1,\n   0\n   ],\n   \"semantic_type\":
\"\",\n   \"description\": \"\"\n   }\n   },\n   {\n
\"column\": \"winner\",\n   \"properties\": {\n   \"dtype\":
\"category\",\n   \"num_unique_values\": 14,\n
\"samples\": [\n   \"Kochi Tuskers Kerala\",\n
\"Rising Pune Supergiants\"\n   ],\n   \"semantic_type\":
\"\",\n   \"description\": \"\"\n   }\n   },\n   {\n
\"column\": \"win_by_runs\",\n   \"properties\": {\n
\"dtype\": \"number\",\n   \"std\": 23,\n   \"min\": 0,\n
\"max\": 146,\n   \"num_unique_values\": 86,\n
\"samples\": [\n   93,\n   140\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"win_by_wickets\",\n
\"properties\": {\n   \"dtype\": \"number\",\n   \"std\":
3,\n   \"min\": 0,\n   \"max\": 10,\n
\"num_unique_values\": 11,\n   \"samples\": [\n   7,\n
0\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"player_of_match\",\n   \"properties\": {\n   \"dtype\":
\"category\",\n   \"num_unique_values\": 201,\n
\"samples\": [\n   \"DE Bollinger\",\n   \"GD McGrath\"\
n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"venue\",\n   \"properties\": {\n   \"dtype\": \"category\",\
n   \"num_unique_values\": 35,\n   \"samples\": [\n
\"Shaheed Veer Narayan Singh International Stadium\",\n
\"Buffalo Park\"\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"umpire1\",\n   \"properties\": {\n   \"dtype\":
\"category\",\n   \"num_unique_values\": 44,\n
\"samples\": [\n   \"Nitin Menon\",\n   \"JD Cloete\"\n
],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n
}\n   },\n   {\n   \"column\": \"umpire2\",\n
\"properties\": {\n   \"dtype\": \"category\",\n
\"num_unique_values\": 45,\n   \"samples\": [\n   \"K
Srinivasan\",\n   \"VA Kulkarni\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"umpire3\",\n   \"properties\":
{\n   \"dtype\": \"number\",\n   \"std\": null,\n
\"min\": null,\n   \"max\": null,\n   \"num_unique_values\":
0,\n   \"samples\": [],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   }\n]\n}","type":"dataframe"}

```
#11  Find the data which consists of a unique city in the data frame.
df['city'].unique()
```

```
array(['Hyderabad', 'Pune', 'Rajkot', 'Indore', 'Bangalore', 'Mumbai',
       'Kolkata', 'Delhi', 'Chandigarh', 'Kanpur', 'Jaipur',
'Chennai',
       'Cape Town', 'Port Elizabeth', 'Durban', 'Centurion',
       'East London', 'Johannesburg', 'Kimberley', 'Bloemfontein',
       'Ahmedabad', 'Cuttack', 'Nagpur', 'Dharamsala', 'Kochi',
       'Visakhapatnam', 'Raipur', 'Ranchi', 'Abu Dhabi', 'Sharjah',
nan],
      dtype=object)
```

```
#12  Find out the IPL team winner of each season
# Get the last match of each season (which would be the final match)
last_matches = df.sort_values('date').groupby('year').last()

# Create a simple list/table of winners by season
last_matches[['winner']]
```


```
{"summary":"{\n  \"name\": \"last_matches[['winner']]\",\n  \"rows\":
10,\n  \"fields\": [\n    {\n      \"column\": \"year\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
3,\n        \"min\": 2008,\n        \"max\": 2017,\n
\"num_unique_values\": 10,\n        \"samples\": [\n          2016,\n
2009,\n          2013\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"winner\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 6,\n        \"samples\": [\n
\"Rajasthan Royals\",\n          \"Deccan Chargers\",\n
\"Sunrisers Hyderabad\"\n        ],\n        \"semantic_type\": \"\",\
n        \"description\": \"\"\n      }\n    }\n  ]\
n}","type":"dataframe"}
```