

# Pandas\_Basics

June 5, 2025

Pandas Basics by Isha Borgaonkar

```
[132]: #Importing Pandas
import pandas as pd
```

```
[133]: #Read the dataset into Pandas DataFrame
df=pd.read_csv('IPLData.csv')
```

```
[134]: #printing the dataframe
df
```

```
[134]:
```

	match_id	inning	batting_team \
0	1	1	Sunrisers Hyderabad
1	1	1	Sunrisers Hyderabad
2	1	1	Sunrisers Hyderabad
3	1	1	Sunrisers Hyderabad
4	1	1	Sunrisers Hyderabad
...	...	...	...
150455	636	2	Royal Challengers Bangalore
150456	636	2	Royal Challengers Bangalore
150457	636	2	Royal Challengers Bangalore
150458	636	2	Royal Challengers Bangalore
150459	636	2	Royal Challengers Bangalore

  

	bowling_team	over	ball	batsman	non_striker \
0	Royal Challengers Bangalore	1	1	DA Warner	S Dhawan
1	Royal Challengers Bangalore	1	2	DA Warner	S Dhawan
2	Royal Challengers Bangalore	1	3	DA Warner	S Dhawan
3	Royal Challengers Bangalore	1	4	DA Warner	S Dhawan
4	Royal Challengers Bangalore	1	5	DA Warner	S Dhawan
...	...	...	...	...	...
150455	Sunrisers Hyderabad	20	2	Sachin Baby	CJ Jordan
150456	Sunrisers Hyderabad	20	3	Sachin Baby	CJ Jordan
150457	Sunrisers Hyderabad	20	4	Iqbal Abdulla	Sachin Baby
150458	Sunrisers Hyderabad	20	5	Sachin Baby	Iqbal Abdulla
150459	Sunrisers Hyderabad	20	6	Iqbal Abdulla	Sachin Baby

  

	bowler	is_super_over	...	bye_runs	legbye_runs	noball_runs \
--	--------	---------------	-----	----------	-------------	---------------

0	TS Mills	0	...	0	0	0
1	TS Mills	0	...	0	0	0
2	TS Mills	0	...	0	0	0
3	TS Mills	0	...	0	0	0
4	TS Mills	0	...	0	0	0
...	...	...	...	...	...	...
150455	B Kumar	0	...	0	0	0
150456	B Kumar	0	...	0	0	0
150457	B Kumar	0	...	0	1	0
150458	B Kumar	0	...	0	0	0
150459	B Kumar	0	...	0	0	0

	penalty_runs	batsman_runs	extra_runs	total_runs	player_dismissed	\
0	0	0	0	0		NaN
1	0	0	0	0		NaN
2	0	4	0	4		NaN
3	0	0	0	0		NaN
4	0	0	2	2		NaN
...	...	...	...	...	...	...
150455	0	2	0	2		NaN
150456	0	0	0	0	CJ Jordan	
150457	0	0	1	1		NaN
150458	0	1	0	1		NaN
150459	0	4	0	4		NaN

	dismissal_kind	fielder
0	NaN	NaN
1	NaN	NaN
2	NaN	NaN
3	NaN	NaN
4	NaN	NaN
...	...	...
150455	NaN	NaN
150456	run out	NV Ojha
150457	NaN	NaN
150458	NaN	NaN
150459	NaN	NaN

[150460 rows x 21 columns]

```
[135]: # View the first few rows of the dataset
df.head()
```

```
[135]: match_id  inning  batting_team  bowling_team  over  \
0         1        1  Sunrisers Hyderabad  Royal Challengers Bangalore  1
1         1        1  Sunrisers Hyderabad  Royal Challengers Bangalore  1
2         1        1  Sunrisers Hyderabad  Royal Challengers Bangalore  1
```

3	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1
4	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1

	ball	batsman	non_striker	bowler	is_super_over	...	bye_runs	\
0	1	DA Warner	S Dhawan	TS Mills	0	...	0	
1	2	DA Warner	S Dhawan	TS Mills	0	...	0	
2	3	DA Warner	S Dhawan	TS Mills	0	...	0	
3	4	DA Warner	S Dhawan	TS Mills	0	...	0	
4	5	DA Warner	S Dhawan	TS Mills	0	...	0	

	legbye_runs	noball_runs	penalty_runs	batsman_runs	extra_runs	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	4	0	
3	0	0	0	0	0	
4	0	0	0	0	2	

	total_runs	player_dismissed	dismissal_kind	fielder
0	0	NaN	NaN	NaN
1	0	NaN	NaN	NaN
2	4	NaN	NaN	NaN
3	0	NaN	NaN	NaN
4	2	NaN	NaN	NaN

[5 rows x 21 columns]

```
[137]: #printing first 10 rows in the datasets
df.head(10)
```

```
[137]: match_id  inning      batting_team      bowling_team  over  \
0          1          1  Sunrisers Hyderabad  Royal Challengers Bangalore  1
1          1          1  Sunrisers Hyderabad  Royal Challengers Bangalore  1
2          1          1  Sunrisers Hyderabad  Royal Challengers Bangalore  1
3          1          1  Sunrisers Hyderabad  Royal Challengers Bangalore  1
4          1          1  Sunrisers Hyderabad  Royal Challengers Bangalore  1
5          1          1  Sunrisers Hyderabad  Royal Challengers Bangalore  1
6          1          1  Sunrisers Hyderabad  Royal Challengers Bangalore  1
7          1          1  Sunrisers Hyderabad  Royal Challengers Bangalore  2
8          1          1  Sunrisers Hyderabad  Royal Challengers Bangalore  2
9          1          1  Sunrisers Hyderabad  Royal Challengers Bangalore  2
```

	ball	batsman	non_striker	bowler	is_super_over	...	bye_runs	\
0	1	DA Warner	S Dhawan	TS Mills	0	...	0	
1	2	DA Warner	S Dhawan	TS Mills	0	...	0	
2	3	DA Warner	S Dhawan	TS Mills	0	...	0	
3	4	DA Warner	S Dhawan	TS Mills	0	...	0	
4	5	DA Warner	S Dhawan	TS Mills	0	...	0	

5	6	S Dhawan	DA Warner	TS Mills	0	...	0
6	7	S Dhawan	DA Warner	TS Mills	0	...	0
7	1	S Dhawan	DA Warner	A Choudhary	0	...	0
8	2	DA Warner	S Dhawan	A Choudhary	0	...	0
9	3	DA Warner	S Dhawan	A Choudhary	0	...	0

	legbye_runs	noball_runs	penalty_runs	batsman_runs	extra_runs	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	4	0	
3	0	0	0	0	0	
4	0	0	0	0	2	
5	0	0	0	0	0	
6	1	0	0	0	1	
7	0	0	0	1	0	
8	0	0	0	4	0	
9	0	1	0	0	1	

	total_runs	player_dismissed	dismissal_kind	fielder
0	0	NaN	NaN	NaN
1	0	NaN	NaN	NaN
2	4	NaN	NaN	NaN
3	0	NaN	NaN	NaN
4	2	NaN	NaN	NaN
5	0	NaN	NaN	NaN
6	1	NaN	NaN	NaN
7	1	NaN	NaN	NaN
8	4	NaN	NaN	NaN
9	1	NaN	NaN	NaN

[10 rows x 21 columns]

```
[140]: #printing last 5 rows by tail function()
df.tail()
```

```
[140]:      match_id  inning      batting_team      bowling_team \
150455      636      2  Royal Challengers Bangalore  Sunrisers Hyderabad
150456      636      2  Royal Challengers Bangalore  Sunrisers Hyderabad
150457      636      2  Royal Challengers Bangalore  Sunrisers Hyderabad
150458      636      2  Royal Challengers Bangalore  Sunrisers Hyderabad
150459      636      2  Royal Challengers Bangalore  Sunrisers Hyderabad

      over  ball      batsman  non_striker  bowler  is_super_over  ...  \
150455   20    2    Sachin Baby    CJ Jordan  B Kumar           0  ...
150456   20    3    Sachin Baby    CJ Jordan  B Kumar           0  ...
150457   20    4  Iqbal Abdulla    Sachin Baby  B Kumar           0  ...
150458   20    5    Sachin Baby  Iqbal Abdulla  B Kumar           0  ...
```

150459	20	6	Iqbal Abdulla	Sachin Baby	B Kumar	0	...
--------	----	---	---------------	-------------	---------	---	-----

	bye_runs	legbye_runs	noball_runs	penalty_runs	batsman_runs	\
150455	0	0	0	0	2	
150456	0	0	0	0	0	
150457	0	1	0	0	0	
150458	0	0	0	0	1	
150459	0	0	0	0	4	

	extra_runs	total_runs	player_dismissed	dismissal_kind	fielder
150455	0	2	NaN	NaN	NaN
150456	0	0	CJ Jordan	run out	NV Ojha
150457	1	1	NaN	NaN	NaN
150458	0	1	NaN	NaN	NaN
150459	0	4	NaN	NaN	NaN

[5 rows x 21 columns]

```
[7]: #printing Last 10 rows by tail function
df.tail(10)
```

```
[7]:
```

	match_id	inning	batting_team	bowling_team	\
150450	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	
150451	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	
150452	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	
150453	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	
150454	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	
150455	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	
150456	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	
150457	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	
150458	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	
150459	636	2	Royal Challengers Bangalore	Sunrisers Hyderabad	

	over	ball	batsman	non_striker	bowler	\
150450	19	3	CJ Jordan	Sachin Baby	Mustafizur Rahman	
150451	19	4	CJ Jordan	Sachin Baby	Mustafizur Rahman	
150452	19	5	Sachin Baby	CJ Jordan	Mustafizur Rahman	
150453	19	6	Sachin Baby	CJ Jordan	Mustafizur Rahman	
150454	20	1	CJ Jordan	Sachin Baby	B Kumar	
150455	20	2	Sachin Baby	CJ Jordan	B Kumar	
150456	20	3	Sachin Baby	CJ Jordan	B Kumar	
150457	20	4	Iqbal Abdulla	Sachin Baby	B Kumar	
150458	20	5	Sachin Baby	Iqbal Abdulla	B Kumar	
150459	20	6	Iqbal Abdulla	Sachin Baby	B Kumar	

	is_super_over	...	bye_runs	legbye_runs	noball_runs	penalty_runs	\
150450		0	...	0	0	0	0

150451	0	...	0	1	0	0
150452	0	...	0	0	0	0
150453	0	...	0	0	0	0
150454	0	...	0	0	0	0
150455	0	...	0	0	0	0
150456	0	...	0	0	0	0
150457	0	...	0	1	0	0
150458	0	...	0	0	0	0
150459	0	...	0	0	0	0

  

	batsman_runs	extra_runs	total_runs	player_dismissed	dismissal_kind \
150450	0	0	0	NaN	NaN
150451	0	1	1	NaN	NaN
150452	2	0	2	NaN	NaN
150453	6	0	6	NaN	NaN
150454	1	0	1	NaN	NaN
150455	2	0	2	NaN	NaN
150456	0	0	0	CJ Jordan	run out
150457	0	1	1	NaN	NaN
150458	1	0	1	NaN	NaN
150459	4	0	4	NaN	NaN

  

	fielder
150450	NaN
150451	NaN
150452	NaN
150453	NaN
150454	NaN
150455	NaN
150456	NV Ojha
150457	NaN
150458	NaN
150459	NaN

[10 rows x 21 columns]

```
[141]: #Get a concise summary of the DataFrame
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150460 entries, 0 to 150459
Data columns (total 21 columns):
#   Column          Non-Null Count  Dtype
---  -
0   match_id        150460 non-null int64
1   inning          150460 non-null int64
2   batting_team    150460 non-null object
3   bowling_team    150460 non-null object
```

```

4   over                150460 non-null int64
5   ball                150460 non-null int64
6   batsman            150460 non-null object
7   non_striker        150460 non-null object
8   bowler             150460 non-null object
9   is_super_over      150460 non-null int64
10  wide_runs          150460 non-null int64
11  bye_runs           150460 non-null int64
12  legbye_runs        150460 non-null int64
13  noball_runs        150460 non-null int64
14  penalty_runs       150460 non-null int64
15  batsman_runs       150460 non-null int64
16  extra_runs         150460 non-null int64
17  total_runs         150460 non-null int64
18  player_dismissed   7438 non-null object
19  dismissal_kind     7438 non-null object
20  fielder            5369 non-null object

```

dtypes: int64(13), object(8)

memory usage: 24.1+ MB

```
[142]: # Statistical summary of numeric columns
df.describe()
```

```
[142]:
```

	match_id	inning	over	ball \
count	150460.000000	150460.000000	150460.000000	150460.000000
mean	318.281317	1.482188	10.142649	3.616483
std	182.955531	0.501768	5.674338	1.807698
min	1.000000	1.000000	1.000000	1.000000
25%	161.000000	1.000000	5.000000	2.000000
50%	319.000000	1.000000	10.000000	4.000000
75%	476.000000	2.000000	15.000000	5.000000
max	636.000000	4.000000	20.000000	9.000000

  

	is_super_over	wide_runs	bye_runs	legbye_runs \
count	150460.000000	150460.000000	150460.000000	150460.000000
mean	0.000538	0.037498	0.004885	0.022232
std	0.023196	0.257398	0.114234	0.200104
min	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000	0.000000
50%	0.000000	0.000000	0.000000	0.000000
75%	0.000000	0.000000	0.000000	0.000000
max	1.000000	5.000000	4.000000	5.000000

  

	noball_runs	penalty_runs	batsman_runs	extra_runs \
count	150460.000000	150460.000000	150460.000000	150460.000000
mean	0.004340	0.000066	1.222445	0.069022
std	0.072652	0.018229	1.594509	0.349667

min	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000	0.000000
50%	0.000000	0.000000	1.000000	0.000000
75%	0.000000	0.000000	1.000000	0.000000
max	5.000000	5.000000	6.000000	7.000000

	total_runs
count	150460.000000
mean	1.291466
std	1.583240
min	0.000000
25%	0.000000
50%	1.000000
75%	1.000000
max	7.000000

```
[10]: #enlisting data types from pandas
df.dtypes
```

```
[10]: match_id      int64
inning           int64
batting_team     object
bowling_team     object
over            int64
ball            int64
batsman          object
non_striker      object
bowler           object
is_super_over    int64
wide_runs        int64
bye_runs         int64
legbye_runs      int64
noball_runs      int64
penalty_runs     int64
batsman_runs     int64
extra_runs       int64
total_runs       int64
player_dismissed object
dismissal_kind   object
fielder          object
dtype: object
```

```
[143]: #Printing ktype of data frame
type(df)
```

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



```
[12]: #List all column names
df.columns
```

```
[12]: Index(['match_id', 'inning', 'batting_team', 'bowling_team', 'over', 'ball',
        'batsman', 'non_striker', 'bowler', 'is_super_over', 'wide_runs',
        'bye_runs', 'legbye_runs', 'noball_runs', 'penalty_runs',
        'batsman_runs', 'extra_runs', 'total_runs', 'player_dismissed',
        'dismissal_kind', 'fielder'],
        dtype='object')
```

```
[144]: #Read the dataset into a DataFrame
df1 = pd.read_csv('IPLData.csv')
```

```
[145]: #Printing the Dataset
df1
```

```
[145]:
```

	match_id	inning	batting_team \
0	1	1	Sunrisers Hyderabad
1	1	1	Sunrisers Hyderabad
2	1	1	Sunrisers Hyderabad
3	1	1	Sunrisers Hyderabad
4	1	1	Sunrisers Hyderabad
...	...	...	...
150455	636	2	Royal Challengers Bangalore
150456	636	2	Royal Challengers Bangalore
150457	636	2	Royal Challengers Bangalore
150458	636	2	Royal Challengers Bangalore
150459	636	2	Royal Challengers Bangalore

  

	bowling_team	over	ball	batsman	non_striker \
0	Royal Challengers Bangalore	1	1	DA Warner	S Dhawan
1	Royal Challengers Bangalore	1	2	DA Warner	S Dhawan
2	Royal Challengers Bangalore	1	3	DA Warner	S Dhawan
3	Royal Challengers Bangalore	1	4	DA Warner	S Dhawan
4	Royal Challengers Bangalore	1	5	DA Warner	S Dhawan
...	...	...	...	...	...
150455	Sunrisers Hyderabad	20	2	Sachin Baby	CJ Jordan
150456	Sunrisers Hyderabad	20	3	Sachin Baby	CJ Jordan
150457	Sunrisers Hyderabad	20	4	Iqbal Abdulla	Sachin Baby
150458	Sunrisers Hyderabad	20	5	Sachin Baby	Iqbal Abdulla
150459	Sunrisers Hyderabad	20	6	Iqbal Abdulla	Sachin Baby

  

	bowler	is_super_over	...	bye_runs	legbye_runs	noball_runs \
0	TS Mills	0	...	0	0	0
1	TS Mills	0	...	0	0	0
2	TS Mills	0	...	0	0	0
3	TS Mills	0	...	0	0	0

4	TS Mills	0	...	0	0	0
...	...	...	...	...	...	...
150455	B Kumar	0	...	0	0	0
150456	B Kumar	0	...	0	0	0
150457	B Kumar	0	...	0	1	0
150458	B Kumar	0	...	0	0	0
150459	B Kumar	0	...	0	0	0

  

	penalty_runs	batsman_runs	extra_runs	total_runs	player_dismissed	\
0	0	0	0	0		NaN
1	0	0	0	0		NaN
2	0	4	0	4		NaN
3	0	0	0	0		NaN
4	0	0	2	2		NaN
...	...	...	...	...	...	...
150455	0	2	0	2		NaN
150456	0	0	0	0	CJ Jordan	
150457	0	0	1	1		NaN
150458	0	1	0	1		NaN
150459	0	4	0	4		NaN

  

	dismissal_kind	fielder
0	NaN	NaN
1	NaN	NaN
2	NaN	NaN
3	NaN	NaN
4	NaN	NaN
...	...	...
150455	NaN	NaN
150456	run out	NV Ojha
150457	NaN	NaN
150458	NaN	NaN
150459	NaN	NaN

[150460 rows x 21 columns]

```
[147]: # Save the DataFrame `df1` to a CSV file named 'result.csv' in the current
        ↪ directory
        df1.to_csv("result.csv")
```

```
[148]: #Data frame to csv setting index and header as false
        df1.to_csv("result.csv",index=False,header=False)
```

```
[149]: #Separating the commas by hashtag
        df1.to_csv("result2.csv",index=False,header=False,sep = "#")
```

```
[150]: #Creating the result4.csv file only with batting team and bowling team column
df1.to_csv("result4.csv",columns=['batting_team','bowling_team'])
```

```
[151]: #only printing cols values of batting_team
df1['batting_team']
```

```
[151]: 0          Sunrisers Hyderabad
1          Sunrisers Hyderabad
2          Sunrisers Hyderabad
3          Sunrisers Hyderabad
4          Sunrisers Hyderabad
...
150455    Royal Challengers Bangalore
150456    Royal Challengers Bangalore
150457    Royal Challengers Bangalore
150458    Royal Challengers Bangalore
150459    Royal Challengers Bangalore
Name: batting_team, Length: 150460, dtype: object
```

```
[152]: # This prints the type of a single column (Series) - returns <class 'pandas.
↳core.series.Series'>
type(df1['batting_team'])
```

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

you have only one record in a Series, but you're treating it like it has many values, like a list."

```
[153]: # This prints the type of a column wrapped in double brackets - returns <class_
↳'pandas.core.frame.DataFrame'>
type(df1[['batting_team']])
```

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

```
[154]: # Select and display specific columns
df1[['match_id','inning','batting_team','bowling_team']]
```

```
[154]:      match_id  inning      batting_team \
0          1        1      Sunrisers Hyderabad
1          1        1      Sunrisers Hyderabad
2          1        1      Sunrisers Hyderabad
3          1        1      Sunrisers Hyderabad
4          1        1      Sunrisers Hyderabad
...
150455      636        2  Royal Challengers Bangalore
150456      636        2  Royal Challengers Bangalore
150457      636        2  Royal Challengers Bangalore
150458      636        2  Royal Challengers Bangalore
150459      636        2  Royal Challengers Bangalore
```

	bowling_team
0	Royal Challengers Bangalore
1	Royal Challengers Bangalore
2	Royal Challengers Bangalore
3	Royal Challengers Bangalore
4	Royal Challengers Bangalore
...	...
150455	Sunrisers Hyderabad
150456	Sunrisers Hyderabad
150457	Sunrisers Hyderabad
150458	Sunrisers Hyderabad
150459	Sunrisers Hyderabad

[150460 rows x 4 columns]

New Dataset

```
[155]: # Read all HTML tables from the given Basketball Reference player page into a
      ↪ list of DataFrames
df4 = pd.read_html("https://www.basketball-reference.com/players/i/ibakase01.
      ↪ html?utm_source=direct&utm_medium=Share&utm_campaign=ShareTool")
```

```
[156]: # Check and print the data type of df4
type(df4)
```

[156]: list

```
[157]: # Print the number of rows (length) in the dataset df1
len(df4)
```

[157]: 11

```
[158]: # Display the first HTML table (as a DataFrame) from the list of tables stored
      ↪ in df4
df4[0]
```

```
[158]:
```

	Season	Age \
0	2008-09	Did not play - other pro league
1	2009-10	20
2	2010-11	21
3	2011-12	22
4	2012-13	23
5	2013-14	24
6	2014-15	25
7	2015-16	26
8	2016-17	27
9	2016-17	27

10	2016-17	27
11	2017-18	28
12	2018-19	29
13	2019-20	30
14	2020-21	31
15	2021-22	32
16	2021-22	32
17	2021-22	32
18	2022-23	33
19	2023-24	Did not play - other pro league
20	NaN	NaN
21	14 Yrs	14 Yrs
22	NaN	NaN
23	OKC (7 Yrs)	OKC (7 Yrs)
24	TOR (4 Yrs)	TOR (4 Yrs)
25	LAC (2 Yrs)	LAC (2 Yrs)
26	MIL (2 Yrs)	MIL (2 Yrs)
27	ORL (1 Yr)	ORL (1 Yr)

	Team	Lg \
0	Did not play - other pro league	Did not play - other pro league
1	OKC	NBA
2	OKC	NBA
3	OKC	NBA
4	OKC	NBA
5	OKC	NBA
6	OKC	NBA
7	OKC	NBA
8	2TM	NBA
9	ORL	NBA
10	TOR	NBA
11	TOR	NBA
12	TOR	NBA
13	TOR	NBA
14	LAC	NBA
15	2TM	NBA
16	LAC	NBA
17	MIL	NBA
18	MIL	NBA
19	Did not play - other pro league	Did not play - other pro league
20	NaN	NaN
21	14 Yrs	14 Yrs
22	NaN	NaN
23	OKC (7 Yrs)	OKC (7 Yrs)
24	TOR (4 Yrs)	TOR (4 Yrs)
25	LAC (2 Yrs)	LAC (2 Yrs)
26	MIL (2 Yrs)	MIL (2 Yrs)

27	ORL (1 Yr)	ORL (1 Yr)
	Pos	G \
0	Did not play - other pro league	Did not play - other pro league
1	PF	73
2	PF	82
3	PF	66
4	PF	80
5	PF	81
6	PF	64
7	PF	78
8	PF	79
9	PF	56
10	PF	23
11	PF	76
12	C	74
13	C	55
14	C	41
15	C	54
16	C	35
17	C	19
18	C	16
19	Did not play - other pro league	Did not play - other pro league
20	Pos	G
21	NaN	919
22	NaN	NaN
23	NaN	524
24	NaN	228
25	NaN	76
26	NaN	35
27	NaN	56
	GS	MP \
0	Did not play - other pro league	Did not play - other pro league
1	0	18.1
2	44	27.0
3	66	27.2
4	80	31.1
5	81	32.9
6	64	33.1
7	78	32.1
8	79	30.7
9	56	30.5
10	23	31.0
11	76	27.5
12	51	27.2
13	27	27.0

14	39	23.3
15	12	16.2
16	10	15.4
17	2	17.8
18	0	11.6
19	Did not play - other pro league	Did not play - other pro league
20	GS	MP
21	697	27.3
22	NaN	NaN
23	413	28.8
24	177	27.6
25	49	19.6
26	2	15.0
27	56	30.5

	FG	FGA	...	\
0	Did not play - other pro league	Did not play - other pro league	...	
1	2.8	5.1	...	
2	4.1	7.5	...	
3	4.0	7.4	...	
4	5.6	9.7	...	
5	6.5	12.1	...	
6	5.8	12.3	...	
7	5.3	11.1	...	
8	6.0	12.4	...	
9	6.1	12.6	...	
10	5.6	12.1	...	
11	5.0	10.3	...	
12	6.3	11.9	...	
13	6.2	12.2	...	
14	4.5	8.9	...	
15	2.7	5.5	...	
16	2.7	5.5	...	
17	2.8	5.5	...	
18	1.6	3.4	...	
19	Did not play - other pro league	Did not play - other pro league	...	
20	FG	FGA	...	
21	5.0	9.7	...	
22	NaN	NaN	...	
23	4.9	9.3	...	
24	5.8	11.4	...	
25	3.7	7.3	...	
26	2.3	4.5	...	
27	6.1	12.6	...	

	ORB	DRB	...	\
0	Did not play - other pro league	Did not play - other pro league	...	

1	1.9	3.5
2	2.6	5.0
3	2.9	4.6
4	2.8	4.9
5	2.8	6.0
6	2.1	5.7
7	1.8	5.0
8	1.6	5.2
9	1.8	5.0
10	1.3	5.5
11	1.0	5.3
12	2.1	6.0
13	2.1	6.1
14	1.8	4.9
15	1.1	3.5
16	1.1	3.1
17	1.2	4.1
18	0.9	1.8
19	Did not play - other pro league	Did not play - other pro league
20	ORB	DRB
21	2.1	5.0
22	NaN	NaN
23	2.4	5.0
24	1.7	5.7
25	1.5	4.1
26	1.1	3.1
27	1.8	5.0

	TRB	AST \
0	Did not play - other pro league	Did not play - other pro league
1	5.4	0.1
2	7.6	0.3
3	7.5	0.4
4	7.7	0.5
5	8.8	1.0
6	7.8	0.9
7	6.8	0.8
8	6.8	0.9
9	6.8	1.1
10	6.8	0.7
11	6.3	0.8
12	8.1	1.3
13	8.2	1.4
14	6.7	1.8
15	4.6	0.9
16	4.3	1.0
17	5.3	0.7



18		2.8		0.3
19	Did not play - other pro league		Did not play - other pro league	
20		TRB		AST
21		7.1		0.8
22		NaN		NaN
23		7.4		0.6
24		7.4		1.1
25		5.6		1.4
26		4.1		0.5
27		6.8		1.1

		STL		BLK \
0	Did not play - other pro league		Did not play - other pro league	
1		0.3		1.3
2		0.4		2.4
3		0.5		3.7
4		0.4		3.0
5		0.5		2.7
6		0.5		2.4
7		0.5		1.9
8		0.5		1.6
9		0.6		1.6
10		0.3		1.4
11		0.4		1.3
12		0.4		1.4
13		0.5		0.8
14		0.2		1.1
15		0.2		0.6
16		0.2		0.7
17		0.2		0.4
18		0.1		0.4
19	Did not play - other pro league		Did not play - other pro league	
20		STL		BLK
21		0.4		1.9
22		NaN		NaN
23		0.4		2.5
24		0.4		1.2
25		0.2		1.0
26		0.2		0.4
27		0.6		1.6

		TOV		PF \
0	Did not play - other pro league		Did not play - other pro league	
1		0.9		2.7
2		0.9		3.3
3		1.2		2.7
4		1.6		2.7

5	1.5	2.9
6	1.5	3.0
7	1.4	2.7
8	1.3	2.7
9	1.2	2.5
10	1.7	3.3
11	1.2	2.8
12	1.5	2.9
13	2.0	2.8
14	1.1	1.9
15	0.8	1.9
16	0.9	1.6
17	0.5	2.4
18	0.7	1.4
19	Did not play - other pro league	Did not play - other pro league
20	TOV	PF
21	1.3	2.7
22	NaN	NaN
23	1.3	2.8
24	1.5	2.8
25	1.0	1.7
26	0.6	2.0
27	1.2	2.5

	PTS	Awards
0	Did not play - other pro league	Did not play - other pro league
1	6.3	NaN
2	9.9	DPOY-14
3	9.1	DPOY-2,DEF1
4	13.2	DPOY-3,DEF1
5	15.1	DPOY-4,DEF1
6	14.3	NaN
7	12.6	NaN
8	14.8	NaN
9	15.1	NaN
10	14.2	NaN
11	12.6	NaN
12	15.0	NaN
13	15.4	NaN
14	11.1	NaN
15	6.8	NaN
16	6.6	NaN
17	7.0	NaN
18	4.1	NaN
19	Did not play - other pro league	Did not play - other pro league
20	PTS	Awards
21	12.0	NaN

22	NaN	NaN
23	11.6	NaN
24	14.2	NaN
25	9.0	NaN
26	5.7	NaN
27	15.1	NaN

[28 rows x 31 columns]

```
[159]: # Save the fourth HTML table (index 3) from df4 to a CSV file named
        ↪ 'players_data.csv'
df4[3].to_csv("players_data.csv")
```

```
[160]: # Read all HTML tables from the specified Basketball Reference player page into
        ↪ a list of DataFrames and store it in df9
df9=pd.read_html(''https://www.basketball-reference.com/players/i/ibakase01.
        ↪html?utm_source=direct&utm_medium=Share&utm_campaign=ShareTool
        ''')
```

```
[161]: # Display the first HTML table (index 0) from the list of DataFrames stored in
        ↪ df9
df9[0]
```

```
[161]:
```

	Season	Age \
0	2008-09	Did not play - other pro league
1	2009-10	20
2	2010-11	21
3	2011-12	22
4	2012-13	23
5	2013-14	24
6	2014-15	25
7	2015-16	26
8	2016-17	27
9	2016-17	27
10	2016-17	27
11	2017-18	28
12	2018-19	29
13	2019-20	30
14	2020-21	31
15	2021-22	32
16	2021-22	32
17	2021-22	32
18	2022-23	33
19	2023-24	Did not play - other pro league
20	NaN	NaN
21	14 Yrs	14 Yrs
22	NaN	NaN

23	OKC (7 Yrs)	OKC (7 Yrs)
24	TOR (4 Yrs)	TOR (4 Yrs)
25	LAC (2 Yrs)	LAC (2 Yrs)
26	MIL (2 Yrs)	MIL (2 Yrs)
27	ORL (1 Yr)	ORL (1 Yr)

	Team	Lg \
0	Did not play - other pro league	Did not play - other pro league
1	OKC	NBA
2	OKC	NBA
3	OKC	NBA
4	OKC	NBA
5	OKC	NBA
6	OKC	NBA
7	OKC	NBA
8	2TM	NBA
9	ORL	NBA
10	TOR	NBA
11	TOR	NBA
12	TOR	NBA
13	TOR	NBA
14	LAC	NBA
15	2TM	NBA
16	LAC	NBA
17	MIL	NBA
18	MIL	NBA
19	Did not play - other pro league	Did not play - other pro league
20	NaN	NaN
21	14 Yrs	14 Yrs
22	NaN	NaN
23	OKC (7 Yrs)	OKC (7 Yrs)
24	TOR (4 Yrs)	TOR (4 Yrs)
25	LAC (2 Yrs)	LAC (2 Yrs)
26	MIL (2 Yrs)	MIL (2 Yrs)
27	ORL (1 Yr)	ORL (1 Yr)

	Pos	G \
0	Did not play - other pro league	Did not play - other pro league
1	PF	73
2	PF	82
3	PF	66
4	PF	80
5	PF	81
6	PF	64
7	PF	78
8	PF	79
9	PF	56

10	PF	23
11	PF	76
12	C	74
13	C	55
14	C	41
15	C	54
16	C	35
17	C	19
18	C	16
19	Did not play - other pro league	Did not play - other pro league
20	Pos	G
21	NaN	919
22	NaN	NaN
23	NaN	524
24	NaN	228
25	NaN	76
26	NaN	35
27	NaN	56

	GS	MP \
0	Did not play - other pro league	Did not play - other pro league
1	0	18.1
2	44	27.0
3	66	27.2
4	80	31.1
5	81	32.9
6	64	33.1
7	78	32.1
8	79	30.7
9	56	30.5
10	23	31.0
11	76	27.5
12	51	27.2
13	27	27.0
14	39	23.3
15	12	16.2
16	10	15.4
17	2	17.8
18	0	11.6
19	Did not play - other pro league	Did not play - other pro league
20	GS	MP
21	697	27.3
22	NaN	NaN
23	413	28.8
24	177	27.6
25	49	19.6
26	2	15.0

27		56	30.5
		FG	FGA ... \
0	Did not play - other pro league	Did not play - other pro league	...
1		2.8	5.1 ...
2		4.1	7.5 ...
3		4.0	7.4 ...
4		5.6	9.7 ...
5		6.5	12.1 ...
6		5.8	12.3 ...
7		5.3	11.1 ...
8		6.0	12.4 ...
9		6.1	12.6 ...
10		5.6	12.1 ...
11		5.0	10.3 ...
12		6.3	11.9 ...
13		6.2	12.2 ...
14		4.5	8.9 ...
15		2.7	5.5 ...
16		2.7	5.5 ...
17		2.8	5.5 ...
18		1.6	3.4 ...
19	Did not play - other pro league	Did not play - other pro league	...
20		FG	FGA ...
21		5.0	9.7 ...
22		NaN	NaN ...
23		4.9	9.3 ...
24		5.8	11.4 ...
25		3.7	7.3 ...
26		2.3	4.5 ...
27		6.1	12.6 ...
		ORB	DRB \
0	Did not play - other pro league	Did not play - other pro league	
1		1.9	3.5
2		2.6	5.0
3		2.9	4.6
4		2.8	4.9
5		2.8	6.0
6		2.1	5.7
7		1.8	5.0
8		1.6	5.2
9		1.8	5.0
10		1.3	5.5
11		1.0	5.3
12		2.1	6.0
13		2.1	6.1

14	1.8	4.9
15	1.1	3.5
16	1.1	3.1
17	1.2	4.1
18	0.9	1.8
19	Did not play - other pro league	Did not play - other pro league
20	ORB	DRB
21	2.1	5.0
22	NaN	NaN
23	2.4	5.0
24	1.7	5.7
25	1.5	4.1
26	1.1	3.1
27	1.8	5.0

	TRB	AST \
0	Did not play - other pro league	Did not play - other pro league
1	5.4	0.1
2	7.6	0.3
3	7.5	0.4
4	7.7	0.5
5	8.8	1.0
6	7.8	0.9
7	6.8	0.8
8	6.8	0.9
9	6.8	1.1
10	6.8	0.7
11	6.3	0.8
12	8.1	1.3
13	8.2	1.4
14	6.7	1.8
15	4.6	0.9
16	4.3	1.0
17	5.3	0.7
18	2.8	0.3
19	Did not play - other pro league	Did not play - other pro league
20	TRB	AST
21	7.1	0.8
22	NaN	NaN
23	7.4	0.6
24	7.4	1.1
25	5.6	1.4
26	4.1	0.5
27	6.8	1.1

	STL	BLK \
0	Did not play - other pro league	Did not play - other pro league

1	0.3	1.3
2	0.4	2.4
3	0.5	3.7
4	0.4	3.0
5	0.5	2.7
6	0.5	2.4
7	0.5	1.9
8	0.5	1.6
9	0.6	1.6
10	0.3	1.4
11	0.4	1.3
12	0.4	1.4
13	0.5	0.8
14	0.2	1.1
15	0.2	0.6
16	0.2	0.7
17	0.2	0.4
18	0.1	0.4
19	Did not play - other pro league	Did not play - other pro league
20	STL	BLK
21	0.4	1.9
22	NaN	NaN
23	0.4	2.5
24	0.4	1.2
25	0.2	1.0
26	0.2	0.4
27	0.6	1.6

	TOV	PF \
0	Did not play - other pro league	Did not play - other pro league
1	0.9	2.7
2	0.9	3.3
3	1.2	2.7
4	1.6	2.7
5	1.5	2.9
6	1.5	3.0
7	1.4	2.7
8	1.3	2.7
9	1.2	2.5
10	1.7	3.3
11	1.2	2.8
12	1.5	2.9
13	2.0	2.8
14	1.1	1.9
15	0.8	1.9
16	0.9	1.6
17	0.5	2.4



18		0.7		1.4
19	Did not play - other pro league		Did not play - other pro league	
20		TOV		PF
21		1.3		2.7
22		NaN		NaN
23		1.3		2.8
24		1.5		2.8
25		1.0		1.7
26		0.6		2.0
27		1.2		2.5

		PTS		Awards
0	Did not play - other pro league		Did not play - other pro league	
1		6.3		NaN
2		9.9		DPOY-14
3		9.1		DPOY-2,DEF1
4		13.2		DPOY-3,DEF1
5		15.1		DPOY-4,DEF1
6		14.3		NaN
7		12.6		NaN
8		14.8		NaN
9		15.1		NaN
10		14.2		NaN
11		12.6		NaN
12		15.0		NaN
13		15.4		NaN
14		11.1		NaN
15		6.8		NaN
16		6.6		NaN
17		7.0		NaN
18		4.1		NaN
19	Did not play - other pro league		Did not play - other pro league	
20		PTS		Awards
21		12.0		NaN
22		NaN		NaN
23		11.6		NaN
24		14.2		NaN
25		9.0		NaN
26		5.7		NaN
27		15.1		NaN

[28 rows x 31 columns]

```
[162]: # Read all HTML tables from the specified Basketball Reference player page into
        ↪ a list of DataFrames
html_data = pd.read_html("https://www.basketball-reference.com/players/i/
        ↪ ibakase01.html?utm_source=direct&utm_medium=Share&utm_campaign=ShareTool")
```

```
# Loop through each DataFrame in the list and save it as a separate CSV file
for i in range(len(html_data)):
    # Save the i-th table to a CSV file named 'html_data0.csv', 'html_data1.
    ↪ csv', etc., without including the index column
    html_data[i].to_csv(f"html_data{i}.csv", index=False)
```

Reading data from the github

```
[163]: # Read the dataset into a DataFrame
pd.read_csv("https://raw.githubusercontent.com/datasciencedojo/datasets/master/
    ↪ titanic.csv")
```

```
[163]:
```

	PassengerId	Survived	Pclass	\
0	1	0	3	
1	2	1	1	
2	3	1	3	
3	4	1	1	
4	5	0	3	
..	...	...	...	
886	887	0	2	
887	888	1	1	
888	889	0	3	
889	890	1	1	
890	891	0	3	

  

	Name	Sex	Age	SibSp	\
0	Braund, Mr. Owen Harris	male	22.0	1	
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	
2	Heikkinen, Miss. Laina	female	26.0	0	
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	
4	Allen, Mr. William Henry	male	35.0	0	
..	...	...	...	...	
886	Montvila, Rev. Juozas	male	27.0	0	
887	Graham, Miss. Margaret Edith	female	19.0	0	
888	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	
889	Behr, Mr. Karl Howell	male	26.0	0	
890	Dooley, Mr. Patrick	male	32.0	0	

  

	Parch	Ticket	Fare	Cabin	Embarked
0	0	A/5 21171	7.2500	NaN	S
1	0	PC 17599	71.2833	C85	C
2	0	STON/O2. 3101282	7.9250	NaN	S
3	0	113803	53.1000	C123	S
4	0	373450	8.0500	NaN	S
..	...	...	...	...	
886	0	211536	13.0000	NaN	S
887	0	112053	30.0000	B42	S

888	2	W./C. 6607	23.4500	NaN	S
889	0	111369	30.0000	C148	C
890	0	370376	7.7500	NaN	Q

[891 rows x 12 columns]

Creating the String/ Creating own data

```
[164]: # Define a multi-line string that looks like a JSON object
s = """{
"name" : "Isha",
"phone_no" : 45454545,
"comp": ["Google", "EY", "unilver", "verizon", "deloitte", "wipro"]
}
"""
```

Java Script Object Notation key value pair Loading the data and passing the string and printing that

```
[165]: import json
```

```
[166]: # Parse the JSON-formtted string `s` into a Python dictionary and store it in
↪ `s1`
s1 = json.loads(s)
```

```
[167]: # Display the contents of the dictionary parsed from the JSON string
s1
```

```
[167]: {'name': 'Isha',
'phone_no': 45454545,
'comp': ['Google', 'EY', 'unilver', 'verizon', 'deloitte', 'wipro']}
```

```
[168]: # Create a DataFrame using the 'comp' key from the dictionary s1 and select
↪ only the 'comp' column
pd.DataFrame(s1, columns=['comp'])
```

```
[168]:      comp
0    Google
1       EY
2    unilver
3    verizon
4    deloitte
5     wipro
```

```
[169]: # Convert the dictionary s1 into a DataFrame - Pandas will create columns from
↪ keys,
# and broadcast scalar values ('name' and 'phone_no') to match the length of
↪ the list under 'comp'
```

```
pd.DataFrame(s1)
```

```
[169]:
```

	name	phone_no	comp
0	Isha	45454545	Google
1	Isha	45454545	EY
2	Isha	45454545	unilver
3	Isha	45454545	verizon
4	Isha	45454545	deloitte
5	Isha	45454545	wipro

```
[170]: # Import the requests library to handle HTTP requests
import requests

# Send a GET request to the GitHub API to retrieve open issues from the pandas
↪ repository
res = requests.get('https://api.github.com/repos/pandas-dev/pandas/issues')

# Check the response status (200 means the request was successful)
print(res)

# Parse the JSON response content into a Python object (usually a list of
↪ dictionaries)
js = res.json()
```

```
<Response [200]>
```

```
[171]: # Dataset available in the format of list
print(js)
```

```
[{'url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61576',
'repository_url': 'https://api.github.com/repos/pandas-dev/pandas',
'labels_url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61576/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61576/comments',
'events_url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61576/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61576',
'id': 3122405296,
'node_id': 'PR_kwDOAA0YD86ZQ5dU',
'number': 61576,
'title': 'DataFrames Class update',
'user': {'login': 'SaraTamname',
'id': 116681400,
'node_id': 'U_kgDOBvRquA',
'avatar_url': 'https://avatars.githubusercontent.com/u/116681400?v=4',
'gravatar_id': '',
'url': 'https://api.github.com/users/SaraTamname',
'html_url': 'https://github.com/SaraTamname',
'followers_url': 'https://api.github.com/users/SaraTamname/followers',
'following_url': 'https://api.github.com/users/SaraTamname/following{/other_user}',
'gists_url': 'https://api.github.com/users/SaraTamname/gists{/gist_id}',
'starred_url': 'https://api.github.com/users/SaraTamname/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/SaraTamname/subscriptions',
'organizations_url': 'https://api.github.com/users/SaraTamname/orgs',
```

```

'repos_url': 'https://api.github.com/users/SaraTammame/repos', 'events_url':
'https://api.github.com/users/SaraTammame/events{/privacy}',
'received_events_url':
'https://api.github.com/users/SaraTammame/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False, 'labels': [], 'state': 'open',
'locked': False, 'assignee': None, 'assignees': [], 'milestone': None,
'comments': 0, 'created_at': '2025-06-05T19:32:17Z', 'updated_at':
'2025-06-05T19:32:17Z', 'closed_at': None, 'author_association': 'NONE', 'type':
None, 'active_lock_reason': None, 'draft': False, 'pull_request': {'url':
'https://api.github.com/repos/pandas-dev/pandas/pulls/61576', 'html_url':
'https://github.com/pandas-dev/pandas/pull/61576', 'diff_url':
'https://github.com/pandas-dev/pandas/pull/61576.diff', 'patch_url':
'https://github.com/pandas-dev/pandas/pull/61576.patch', 'merged_at': None},
'body': '- [ ] closes #xxxx (Replace xxxx with the GitHub issue number)\r\n- [ ]
[Tests added and passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#writing-tests) if fixing a bug
or adding a new feature\r\n- [ ] All [code checks
passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#pre-commit).\r\n- [ ] Added
[type annotations](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#type-hints) to new
arguments/methods/functions.\r\n- [ ] Added an entry in the latest
`doc/source/whatsnew/vX.X.X.rst` file if fixing a bug or adding a new
feature.\r\n', 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61576/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61576/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61574', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61574/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61574/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61574/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61574', 'id':
3122363438, 'node_id': 'I_kwDOAAOYD866G3wu', 'number': 61574, 'title': "BUG:
2.3.0 didn't publish wheels for musl-aarch64 (arm)", 'user': {'login':
'virajmehta', 'id': 9244420, 'node_id': 'MDQ6VXNlcjkyNDQOMjA=', 'avatar_url':
'https://avatars.githubusercontent.com/u/9244420?v=4', 'gravatar_id': '', 'url':
'https://api.github.com/users/virajmehta', 'html_url':
'https://github.com/virajmehta', 'followers_url':
'https://api.github.com/users/virajmehta/followers', 'following_url':
'https://api.github.com/users/virajmehta/following{/other_user}', 'gists_url':
'https://api.github.com/users/virajmehta/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/virajmehta/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/virajmehta/subscriptions',
'organizations_url': 'https://api.github.com/users/virajmehta/orgs',

```

```

'repos_url': 'https://api.github.com/users/virajmehta/repos', 'events_url':
'https://api.github.com/users/virajmehta/events{/privacy}',
'received_events_url':
'https://api.github.com/users/virajmehta/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{'id': 76811,
'node_id': 'MDU6TGFiZWw3NjgxMQ==', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Bug', 'name': 'Bug', 'color': 'e10c02', 'default': False,
'description': None}, {'id': 1954720290, 'node_id': 'MDU6TGFiZWw3OTUONzIwMjkw',
'url': 'https://api.github.com/repos/pandas-dev/pandas/labels/Needs%20Triage',
'name': 'Needs Triage', 'color': '0052cc', 'default': False, 'description':
'Issue that has not been reviewed by a pandas team member'}], 'state': 'open',
'locked': False, 'assignee': None, 'assignees': [], 'milestone': None,
'comments': 0, 'created_at': '2025-06-05T19:18:15Z', 'updated_at':
'2025-06-05T19:18:18Z', 'closed_at': None, 'author_association': 'NONE', 'type':
None, 'active_lock_reason': None, 'sub_issues_summary': {'total': 0,
'completed': 0, 'percent_completed': 0}, 'body': '### Pandas version checks\n\n-
[x] I have checked that this issue has not already been reported.\n\n- [x] I
have confirmed this bug exists on the [latest
version](https://pandas.pydata.org/docs/whatsnew/index.html) of pandas.\n\n- [x]
I have confirmed this bug exists on the [main branch](https://pandas.pydata.org/
docs/dev/getting_started/install.html#installing-the-development-version-of-
pandas) of pandas.\n\n\n### Reproducible Example\n\n``python\nhttps://pypi.org/
project/pandas/2.3.0/#files\nhttps://pypi.org/project/pandas/2.2.3/#files\n``\n
\n### Issue Description\n\nctrl-f for musl on the first page gives 15
results\nfor the second page gives 36 results.\nctrl on the 2.3.0 page for
"-cp313-cp313-musllinux_1_2_aarch64.whl" gives no results but there are results
on the 2.2.3 page.j\n\n### Expected Behavior\n\nwould like a wheel for musl /
arm64\n\n### Installed Versions\n\nnot relevant', 'closed_by': None,
'reactions': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61574/reactions', 'total_count': 1, '+1': 1, '-1': 0, 'laugh':
0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 0},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61574/timeline', 'performed_via_github_app': None,
'state_reason': None}, {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61573', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61573/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61573/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61573/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61573', 'id':
3122306508, 'node_id': 'I_kwDOAAOYD866Gp3M', 'number': 61573, 'title': 'WEB:
Test that our versions JSON is valid', 'user': {'login': 'datapythonista', 'id':
10058240, 'node_id': 'MDQ6VXNlcjEwMDU4MjQw', 'avatar_url':
'https://avatars.githubusercontent.com/u/10058240?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/datapythonista', 'html_url':
'https://github.com/datapythonista', 'followers_url':
'https://api.github.com/users/datapythonista/followers', 'following_url':

```

```

'https://api.github.com/users/datapythonista/following{/other_user}',
'gists_url': 'https://api.github.com/users/datapythonista/gists{/gist_id}',
'starred_url':
'https://api.github.com/users/datapythonista/starred{/owner}/{/repo}',
'subscriptions_url':
'https://api.github.com/users/datapythonista/subscriptions',
'organizations_url': 'https://api.github.com/users/datapythonista/orgs',
'repos_url': 'https://api.github.com/users/datapythonista/repos', 'events_url':
'https://api.github.com/users/datapythonista/events{/privacy}',
'received_events_url':
'https://api.github.com/users/datapythonista/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{ 'id': 717120670,
'node_id': 'MDU6TGFiZWw3MTcxMjA2NzA=', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/good%20first%20issue',
'name': 'good first issue', 'color': '0e8a16', 'default': True, 'description':
None}, { 'id': 1508144531, 'node_id': 'MDU6TGFiZWw3NTA4MTQONTMx', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Web', 'name': 'Web',
'color': '23ef97', 'default': False, 'description': 'pandas website'}], 'state':
'open', 'locked': False, 'assignee': None, 'assignees': [], 'milestone': None,
'comments': 0, 'created_at': '2025-06-05T18:59:22Z', 'updated_at':
'2025-06-05T18:59:22Z', 'closed_at': None, 'author_association': 'MEMBER',
'type': None, 'active_lock_reason': None, 'sub_issues_summary': { 'total': 0,
'completed': 0, 'percent_completed': 0}, 'body': 'See #61572\n\nI think this can
be as simple as loading the file with `json.load` when calling `pandas_web.py`.
This way, if the file is not valid JSON the CI should break. But we need to
double check that `json.load` fails if an extra comma is present.', 'closed_by':
None, 'reactions': { 'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61573/reactions', 'total_count': 0, '+1': 0, '-1': 0, 'laugh':
0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 0},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61573/timeline', 'performed_via_github_app': None,
'state_reason': None}, { 'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61570', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61570/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61570/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61570/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61570', 'id':
3122126379, 'node_id': 'I_kwDOAA0YD866F94r', 'number': 61570, 'title': 'BUG:
comparing strings of different dtypes errors', 'user': { 'login': 'a-reich',
'id': 73507369, 'node_id': 'MDQ6VXNlcjczNTA3MzY5', 'avatar_url':
'https://avatars.githubusercontent.com/u/73507369?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/a-reich', 'html_url':
'https://github.com/a-reich', 'followers_url':
'https://api.github.com/users/a-reich/followers', 'following_url':
'https://api.github.com/users/a-reich/following{/other_user}', 'gists_url':
'https://api.github.com/users/a-reich/gists{/gist_id}', 'starred_url':

```

```

'https://api.github.com/users/a-reich/starred{/owner}/{repo}',
'subscriptions_url': 'https://api.github.com/users/a-reich/subscriptions',
'organizations_url': 'https://api.github.com/users/a-reich/orgs', 'repos_url':
'https://api.github.com/users/a-reich/repos', 'events_url':
'https://api.github.com/users/a-reich/events{/privacy}', 'received_events_url':
'https://api.github.com/users/a-reich/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{ 'id': 76811,
'node_id': 'MDU6TGFiZWw3NjgXMQ==', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Bug', 'name': 'Bug', 'color': 'e10c02', 'default': False,
'description': None}, { 'id': 1954720290, 'node_id': 'MDU6TGFiZWw3OTUONzIwMjkw',
'url': 'https://api.github.com/repos/pandas-dev/pandas/labels/Needs%20Triage',
'name': 'Needs Triage', 'color': '0052cc', 'default': False, 'description':
'Issue that has not been reviewed by a pandas team member'}], 'state': 'open',
'locked': False, 'assignee': None, 'assignees': [], 'milestone': None,
'comments': 0, 'created_at': '2025-06-05T17:57:17Z', 'updated_at':
'2025-06-05T17:57:17Z', 'closed_at': None, 'author_association': 'NONE', 'type':
None, 'active_lock_reason': None, 'sub_issues_summary': { 'total': 0,
'completed': 0, 'percent_completed': 0}, 'body': '### Pandas version checks\n\n-
[x] I have checked that this issue has not already been reported.\n\n- [x] I
have confirmed this bug exists on the [latest
version](https://pandas.pydata.org/docs/whatsnew/index.html) of pandas.\n\n- [ ]
I have confirmed this bug exists on the [main branch](https://pandas.pydata.org/
docs/dev/getting_started/install.html#installing-the-development-version-of-
pandas) of pandas.\n\n\n### Reproducible Example\n\n```\npython\nimport pandas as
pd, numpy as np\narr1 = pd.array([],pd.StringDtype("pyarrow",
na_value=pd.NA))\narr2 = pd.array([], pd.StringDtype("python",
na_value=np.nan))\narr1 == arr2 # NotImplementedError: eq not implemented for
<class \'pandas.core.arrays.string_.StringArrayNumpySemantics\'>\n\n\n###
Issue Description\n\nThis appears to be the type of issue discussed in
https://github.com/pandas-dev/pandas/issues/60639. That issue was closed, but I
got an error when I tried running the above reproducer on the example given in
the [whatsnew](https://pandas.pydata.org/pandas-
docs/version/2.3.0/whatsnew/v2.3.0.html#notable-bug-fix1) for release 2.3.\n\nMy
understanding was that the issue was closed when https://github.com/pandas-
dev/pandas/pull/61138 was merged to main, but it\'s unclear if the fix was
successfully backported to the 2.3.x branch. I haven\'t had the time yet to try
when building pandas myself from main.\n\n\n### Expected Behavior\n\nComparisons
of string arrays/series with different dtypes should not error and the return
dtype should follow the behavior laid out in #60639 .\n\n\n### Installed
Versions\n\n<details>\n\nINSTALLED VERSIONS\n-----\ncommit
: 2cc37625532045f4ac55b27176454bbbc9baf213\npython
:
3.12.2\npython-bits : 64\nOS : Linux\nOS-release
: 6.6.87.1-microsoft-standard-WSL2\nVersion : #1 SMP
PREEMPT_DYNAMIC Mon Apr 21 17:08:54 UTC 2025\nmachine
:
x86_64\nprocessor : x86_64\nbyteorder : little\nLC_ALL
: None\nLANG : C.UTF-8\nLOCALE :
C.UTF-8\n\npandas : 2.3.0\nnumpy : 2.2.6\npytz
: 2025.2\ndateutil : 2.9.0.post0\npip
:

```



```

None\nCython                : None\nsphinx                : None\nIPython
: None\nadbc-driver-postgresql: None\nadbc-driver-sqlite      : None\nbs4
: None\nblosc                : None\nbottleneck            : None\ndataframe-
api-compat : None\nfastparquet          : None\nfsspec                :
None\nhtml5lib              : None\nhypothesis            : None\nngcsfs
: None\njinja2              : None\nlxml.etree            : None\nmatplotlib
: None\nnumba                : None\nnumexpr              : None\nodfpy
: None\nopenpyxl            : None\npandas_gbq           : None\npsycopg2
: None\npymysql              : None\npyarrow              : 20.0.0\npyreadstat
: None\npytest              : None\npython-calamine      : None\npyxlsb
: None\ns3fs                : None\nscipy                : None\nsqlalchemy
: None\ntables            : None\ntabulate            : None\nxarray
: None\nxlrd                : None\nxlswriter            : None\nzstandard
: None\ntzdata            : 2025.2\nqtpy                : None\npyqt5
: None\n</details>\n', 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61570/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61570/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61569', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61569/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61569/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61569/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61569', 'id': 3122042354,
'node_id': 'PR_kwD0AA0YD86ZPoQa', 'number': 61569, 'title': 'BLD: Build wheels
for 3.9', 'user': {'login': 'mroeschke', 'id': 10647082, 'node_id':
'MDQ6VXNlcjEwNjQ3MDgy', 'avatar_url':
'https://avatars.githubusercontent.com/u/10647082?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/mroeschke', 'html_url':
'https://github.com/mroeschke', 'followers_url':
'https://api.github.com/users/mroeschke/followers', 'following_url':
'https://api.github.com/users/mroeschke/following{/other_user}', 'gists_url':
'https://api.github.com/users/mroeschke/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/mroeschke/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/mroeschke/subscriptions',
'organizations_url': 'https://api.github.com/users/mroeschke/orgs', 'repos_url':
'https://api.github.com/users/mroeschke/repos', 'events_url':
'https://api.github.com/users/mroeschke/events{/privacy}',
'received_events_url': 'https://api.github.com/users/mroeschke/received_events',
'type': 'User', 'user_view_type': 'public', 'site_admin': False}, 'labels':
[{'id': 129350, 'node_id': 'MDU6TGFiZWwzMjkzNTA=', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Build', 'name': 'Build',
'color': '75507B', 'default': False, 'description': 'Library building on various
platforms'}], 'state': 'open', 'locked': False, 'assignee': None, 'assignees':
[], 'milestone': None, 'comments': 1, 'created_at': '2025-06-05T17:29:26Z',

```

```

'updated_at': '2025-06-05T18:54:07Z', 'closed_at': None, 'author_association':
'MEMBER', 'type': None, 'active_lock_reason': None, 'draft': False,
'pull_request': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/pulls/61569', 'html_url': 'https://github.com/pandas-
dev/pandas/pull/61569', 'diff_url': 'https://github.com/pandas-
dev/pandas/pull/61569.diff', 'patch_url': 'https://github.com/pandas-
dev/pandas/pull/61569.patch', 'merged_at': None}, 'body': '@lithomas1 would I
need to re-tag the 2.3.x branch if/when we merge this?\r\n\r\nxref
https://github.com/pandas-dev/pandas/issues/61563', 'closed_by': None,
'reactions': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61569/reactions', 'total_count': 0, '+1': 0, '-1': 0, 'laugh':
0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 0},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61569/timeline', 'performed_via_github_app': None,
'state_reason': None}, {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61568', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61568/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61568/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61568/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61568', 'id':
3121700295, 'node_id': 'I_kwDOAAOYD866EV3H', 'number': 61568, 'title': 'BUG:
Inconsistent behavior surrounding pd.fillna', 'user': {'login': 'anna-
intellegens', 'id': 98812604, 'node_id': 'U_kgD0BePCvA', 'avatar_url':
'https://avatars.githubusercontent.com/u/98812604?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/anna-intellegens', 'html_url':
'https://github.com/anna-intellegens', 'followers_url':
'https://api.github.com/users/anna-intellegens/followers', 'following_url':
'https://api.github.com/users/anna-intellegens/following{/other_user}',
'gists_url': 'https://api.github.com/users/anna-intellegens/gists{/gist_id}',
'starred_url': 'https://api.github.com/users/anna-
intellegens/starred{/owner}/{repo}', 'subscriptions_url':
'https://api.github.com/users/anna-intellegens/subscriptions',
'organizations_url': 'https://api.github.com/users/anna-intellegens/orgs',
'repos_url': 'https://api.github.com/users/anna-intellegens/repos',
'events_url': 'https://api.github.com/users/anna-intellegens/events{/privacy}',
'received_events_url': 'https://api.github.com/users/anna-
intellegens/received_events', 'type': 'User', 'user_view_type': 'public',
'site_admin': False}, 'labels': [{'id': 76811, 'node_id':
'MDU6TGFiZWw3NjgxMQ==', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Bug', 'name': 'Bug', 'color': 'e10c02', 'default': False,
'description': None}, {'id': 1954720290, 'node_id': 'MDU6TGFiZWwxOTUONzIwMjkw',
'url': 'https://api.github.com/repos/pandas-dev/pandas/labels/Needs%20Triage',
'name': 'Needs Triage', 'color': '0052cc', 'default': False, 'description':
'Issue that has not been reviewed by a pandas team member'}], 'state': 'open',
'locked': False, 'assignee': None, 'assignees': [], 'milestone': None,
'comments': 1, 'created_at': '2025-06-05T15:33:32Z', 'updated_at':

```

```

'2025-06-05T18:13:49Z', 'closed_at': None, 'author_association': 'NONE', 'type':
None, 'active_lock_reason': None, 'sub_issues_summary': {'total': 0,
'completed': 0, 'percent_completed': 0}, 'body': '### Pandas version checks\n\n-
[x] I have checked that this issue has not already been reported.\n\n- [x] I
have confirmed this bug exists on the [latest
version](https://pandas.pydata.org/docs/whatsnew/index.html) of pandas.\n\n- [ ]
I have confirmed this bug exists on the [main branch](https://pandas.pydata.org/
docs/dev/getting_started/install.html#installing-the-development-version-of-
pandas) of pandas.\n\n\n### Reproducible Example\n\n```\npython\nimport pandas as
pd\nimport numpy as np\n\nempty = pd.DataFrame([[None, None, None, None], [None,
None, None, None], [None, None, None, None], [None, None, None, None]],
columns=list("ABCD"), dtype=np.float64)\nprint(empty.dtypes)\n# A    float64\n# B    float64\n# C    float64\n# D    float64\n# dtype: object\n\nfull_a =
pd.DataFrame([[1.0, 2.0, "3.0", 4.0], [5.0, 6.0, "7.0", 8.0],
[9.0, 10.0, "11.0", 12.0], [13.0, 14.0, "15.0", 16.0]],
columns=list("ABCD"))\nprint(full_a.dtypes)\n# A    float64\n# B    float64\n# C
object\n# D    float64\n# dtype: object\n\n\nfull_b = pd.DataFrame([[1.5, 2.0,
"3.0", 4.0], [5.0, 6.5, "7.0", 8.0], [9.0, 10.0, "11.0", 12.0],
[13.0, 14.0, "15.0", 16.5]], columns=list("ABCD"))\nprint(full_b.dtypes)\n# A
float64\n# B    float64\n# C    object\n# D    float64\n# dtype:
object\n\ncombined_1 = empty.fillna(full_a)\nprint(combined_1.dtypes)\n# A
int64\n# B    int64\n# C    object\n# D    int64\n# dtype:
object\n\ncombined_2 = empty.fillna(full_b)\nprint(combined_2.dtypes)\n# A
object\n# B    object\n# C    object\n# D    object\n# dtype: object\n\n\n###
Issue Description\n\nThe returned types of pandas dataframe fillna method gives
inconsistent resulting types between a column that contains integral float
values, and ones that don't. This leads to very confusing behavior, where the
exact values of the input data (even if it was correctly starting as float64s in
both dataframes) can affect the output types. In particular, if both the
starting column and the merging column have the float64 dtype, as a user I would
expect the output column to have a float64 dtype, but instead I get an int64 if
all the values happen to be integral, otherwise I get an object dtype?! This
behavior is further only observed if one of the other columns happen to be
(correctly) an object dtype, when again, I expected the types of unrelated
columns not to affect each other.\n\nI know there are currently changes
undergoing surrounding casting of types, but here as all types are being
inputted correctly I didn't expect any casting to be being performed as part of
this operation?\n\n\n### Expected Behavior\n\nIn the above example, I expected
both `combined_1` and `combined_2` to have the same dtypes as each other.\nI
also expected both of them to actually have dtypes of `float64` for cols `A`,
`B` and `D`, given the input types are `float64`. The `object` type for those
columns of `combined_2` is particularly confusing in this case\n\n\n### Installed
Versions\n\n<details>\n\nINSTALLED VERSIONS\n-----\ncommit
: 0691c5cf90477d3503834d983f69350f250a6ff7\npython
:
3.12.8\npython-bits
: 64\nOS
: Linux\nOS-release
: 6.11.0-26-generic\nVersion
: #26~24.04.1-Ubuntu SMP
PREEMPT_DYNAMIC Thu Apr 17 19:20:47 UTC 2\nmachine
:
x86_64\nprocessor
: x86_64\nbyteorder
: little\nLC_ALL

```

```

: None\nLANG          : en_GB.UTF-8\nLOCALE          :
en_GB.UTF-8\npandas    : 2.2.3\nnumpy              : 2.2.6\npytz
: 2025.2\ndateutil    : 2.9.0.post0\npip              :
None\nCython          : None\nsphinx              : None\nIPython
: None\nadbc-driver-postgresql: None\nadbc-driver-sqlite  : None\nbs4
: None\nblosc          : None\nbottleneck          : None\ndataframe-
api-compat : None\nfastparquet      : None\nfsspec          :
None\nhtml5lib          : None\nhypothesis          : None\nngcsfs
: None\njinja2          : 3.1.6\nlxml.etree          : None\nmatplotlib
: None\nnumba          : None\nnumexpr          : None\nodfpy
: None\nopenpyxl        : None\npandas_gbq          : None\npsycopg2
: 2.9.10\npymysql       : None\npyarrow          : None\npyreadstat
: None\npytest          : 8.3.5\npython-calamine      : None\npyxlsb
: None\ns3fs            : None\nscipy            : 1.15.3\nsqlalchemy
: 2.0.41\ntables        : None\ntabulate          : None\nxarray
: None\nxlrd            : None\nxlswriter        : None\nzstandard
: None\ntzdata          : 2025.2\nqtpy            : None\npyqt5
: None\n\n</details>\n\nAlso tested on 2.3.0 (sorry, website still says 2.2.3 is
latest):\n\n<details>\n\nINSTALLED VERSIONS\n-----\ncommit
: 2cc37625532045f4ac55b27176454bbbc9baf213\npython          :
3.12.8\npython-bits      : 64\nOS                  : Linux\nOS-release
: 6.11.0-26-generic\nVersion          : #26~24.04.1-Ubuntu SMP
PREEMPT_DYNAMIC Thu Apr 17 19:20:47 UTC 2\nmachine          :
x86_64\nprocessor        : x86_64\nbyteorder          : little\nLC_ALL
: None\nLANG            : en_GB.UTF-8\nLOCALE          :
en_GB.UTF-8\npandas    : 2.3.0\nnumpy              : 2.2.6\npytz
: 2025.2\ndateutil    : 2.9.0.post0\npip              :
None\nCython          : None\nsphinx              : None\nIPython
: None\nadbc-driver-postgresql: None\nadbc-driver-sqlite  : None\nbs4
: None\nblosc          : None\nbottleneck          : None\ndataframe-
api-compat : None\nfastparquet      : None\nfsspec          :
None\nhtml5lib          : None\nhypothesis          : None\nngcsfs
: None\njinja2          : 3.1.6\nlxml.etree          : None\nmatplotlib
: None\nnumba          : None\nnumexpr          : None\nodfpy
: None\nopenpyxl        : None\npandas_gbq          : None\npsycopg2
: 2.9.10\npymysql       : None\npyarrow          : None\npyreadstat
: None\npytest          : 8.3.5\npython-calamine      : None\npyxlsb
: None\ns3fs            : None\nscipy            : 1.15.3\nsqlalchemy
: 2.0.41\ntables        : None\ntabulate          : None\nxarray
: None\nxlrd            : None\nxlswriter        : None\nzstandard
: None\ntzdata          : 2025.2\nqtpy            : None\npyqt5
: None\n\n</details>', 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61568/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61568/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61567', 'repository_url':

```

```

'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61567/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61567/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61567/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61567', 'id': 3121077669,
'node_id': 'PR_kwDOAAOYD86ZMg54', 'number': 61567, 'title': 'BUILD: Add wheels
for musllinux_aarch64', 'user': {'login': 'cdce8p', 'id': 30130371, 'node_id':
'MDQ6VXNlcjMwMTMwMzcx', 'avatar_url':
'https://avatars.githubusercontent.com/u/30130371?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/cdce8p', 'html_url':
'https://github.com/cdce8p', 'followers_url':
'https://api.github.com/users/cdce8p/followers', 'following_url':
'https://api.github.com/users/cdce8p/following{/other_user}', 'gists_url':
'https://api.github.com/users/cdce8p/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/cdce8p/starred{/owner}/{repo}',
'subscriptions_url': 'https://api.github.com/users/cdce8p/subscriptions',
'organizations_url': 'https://api.github.com/users/cdce8p/orgs', 'repos_url':
'https://api.github.com/users/cdce8p/repos', 'events_url':
'https://api.github.com/users/cdce8p/events{/privacy}', 'received_events_url':
'https://api.github.com/users/cdce8p/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [], 'state': 'open',
'locked': False, 'assignee': None, 'assignees': [], 'milestone': None,
'comments': 0, 'created_at': '2025-06-05T12:41:27Z', 'updated_at':
'2025-06-05T12:41:27Z', 'closed_at': None, 'author_association': 'CONTRIBUTOR',
'type': None, 'active_lock_reason': None, 'draft': False, 'pull_request':
{'url': 'https://api.github.com/repos/pandas-dev/pandas/pulls/61567',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61567', 'diff_url':
'https://github.com/pandas-dev/pandas/pull/61567.diff', 'patch_url':
'https://github.com/pandas-dev/pandas/pull/61567.patch', 'merged_at': None},
'body': 'It seems the `musllinux_aarch64` wheels got accidentally removed in the
transition from circleci to GHA.\r\nWould be great if this could be backported
to `2.3.x` as well.\r\n\r\n- [x] closes https://github.com/pandas-
dev/pandas/issues/55645#issuecomment-2943818815\r\n- [ ] [Tests added and
passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#writing-tests) if fixing a bug
or adding a new feature\r\n- [ ] All [code checks
passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#pre-commit).\r\n- [ ] Added
[type annotations](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#type-hints) to new
arguments/methods/functions.\r\n- [ ] Added an entry in the latest
`doc/source/whatsnew/vX.X.X.rst` file if fixing a bug or adding a new
feature.\r\n', 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61567/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61567/timeline',

```

```

'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61566', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61566/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61566/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61566/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61566', 'id':
3120672132, 'node_id': 'I_kwDOAAOYD866Aa2E', 'number': 61566, 'title': 'BUILD:
Installation issue on Mac with M1 Pro arm64 processor.
pandas_parser.cpython-311-darwin.so is using x86_64 arch', 'user': {'login':
'Leviann', 'id': 2115421, 'node_id': 'MDQ6VXNlcjIxMTUOMjE=', 'avatar_url':
'https://avatars.githubusercontent.com/u/2115421?v=4', 'gravatar_id': '', 'url':
'https://api.github.com/users/Leviann', 'html_url':
'https://github.com/Leviann', 'followers_url':
'https://api.github.com/users/Leviann/followers', 'following_url':
'https://api.github.com/users/Leviann/following{/other_user}', 'gists_url':
'https://api.github.com/users/Leviann/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/Leviann/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/Leviann/subscriptions',
'organizations_url': 'https://api.github.com/users/Leviann/orgs', 'repos_url':
'https://api.github.com/users/Leviann/repos', 'events_url':
'https://api.github.com/users/Leviann/events{/privacy}', 'received_events_url':
'https://api.github.com/users/Leviann/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{ 'id': 129350,
'node_id': 'MDU6TGFiZWwxMjkzNTA=', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Build', 'name': 'Build', 'color': '75507B', 'default': False,
'description': 'Library building on various platforms'}, { 'id': 1954720290,
'node_id': 'MDU6TGFiZWwxOTUONzIwMjkw', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Needs%20Triage', 'name':
'Needs Triage', 'color': '0052cc', 'default': False, 'description': 'Issue that
has not been reviewed by a pandas team member'}], 'state': 'open', 'locked':
False, 'assignee': None, 'assignees': [], 'milestone': None, 'comments': 0,
'created_at': '2025-06-05T10:26:33Z', 'updated_at': '2025-06-05T10:26:33Z',
'closed_at': None, 'author_association': 'NONE', 'type': None,
'active_lock_reason': None, 'sub_issues_summary': {'total': 0, 'completed': 0,
'percent_completed': 0}, 'body': "### Installation check\n\n [x] I have read
the [installation guide](https://pandas.pydata.org/pandas-
docs/stable/getting_started/install.html#installing-pandas).\n\n\n###
Platform\n\nmacOS-15.5-arm64-arm-64bit\n\n\n### Installation Method\n\nBuilt from
source\n\n\n### pandas Version\n\npandas-3.0.0.dev0+2147.g1f6f42ac55\n\n\n### Python
Version\n\n3.11.13\n\n\n### Installation Logs\n\nI can't use pandas installed
either using pip or from source, on my macbook with M1 Pro arm64 processor.\nI
am not using rosetta in terminal and my python install is also amd64.\n\nPlease
help. It looks like one library file is built with x86-x64
arch.\n\n<details>\n\n    import pandas as
pd\n\n../Library/Python/3.11/lib/python/site-packages/pandas/__init__.py:45: in
<module>\n    from pandas.core.api import

```

```
(\n../Library/Python/3.11/lib/python/site-packages/pandas/core/api.py:1: in
<module>\n    from pandas._libs import
(\n../Library/Python/3.11/lib/python/site-
packages/pandas/_libs/_init__.py:16: in <module>\n    import
pandas._libs.pandas_parser # isort: skip # type: ignore[reportUnusedImport]\n
~~~~~\nE ImportError:
dlopen(/Users/leviann/Library/Python/3.11/lib/python/site-
packages/pandas/_libs/pandas_parser.cpython-311-darwin.so, 0x0002): tried:
'/Users/leviann/Library/Python/3.11/lib/python/site-
packages/pandas/_libs/pandas_parser.cpython-311-darwin.so' (mach-o file, but is
an incompatible architecture (have 'x86_64', need 'arm64e' or 'arm64')), '/System/Volumes/Preboot/Cryptexes/OS/Users/leviann/Library/Python/3.11/lib/python/site-
packages/pandas/_libs/pandas_parser.cpython-311-darwin.so' (no such file),
'/Users/leviann/Library/Python/3.11/lib/python/site-
packages/pandas/_libs/pandas_parser.cpython-311-darwin.so' (mach-o file, but is
an incompatible architecture (have 'x86_64', need 'arm64e' or
'arm64'))\n\n</details>\n", 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61566/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61566/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61565', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61565/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61565/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61565/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61565', 'id':
3120501218, 'node_id': 'I_kwDOAAOYD865_xHi', 'number': 61565, 'title': 'BUG:
RecursionError when apply generic alias as a func', 'user': {'login':
'MacroBull', 'id': 2204917, 'node_id': 'MDQ6VXNlcjIyMDQ5MTc=', 'avatar_url':
'https://avatars.githubusercontent.com/u/2204917?v=4', 'gravatar_id': '', 'url':
'https://api.github.com/users/MacroBull', 'html_url':
'https://github.com/MacroBull', 'followers_url':
'https://api.github.com/users/MacroBull/followers', 'following_url':
'https://api.github.com/users/MacroBull/following{/other_user}', 'gists_url':
'https://api.github.com/users/MacroBull/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/MacroBull/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/MacroBull/subscriptions',
'organizations_url': 'https://api.github.com/users/MacroBull/orgs', 'repos_url':
'https://api.github.com/users/MacroBull/repos', 'events_url':
'https://api.github.com/users/MacroBull/events{/privacy}',
'received_events_url': 'https://api.github.com/users/MacroBull/received_events',
'type': 'User', 'user_view_type': 'public', 'site_admin': False}, 'labels':
[{'id': 76811, 'node_id': 'MDU6TGFiZWw3NjgxMQ==', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Bug', 'name': 'Bug',
'color': 'e10c02', 'default': False, 'description': None}, {'id': 1954720290,
```

```

'node_id': 'MDU6TGFiZWwzOTUONzIwMjkw', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Needs%20Triage', 'name':
'Needs Triage', 'color': '0052cc', 'default': False, 'description': 'Issue that
has not been reviewed by a pandas team member']], 'state': 'open', 'locked':
False, 'assignee': None, 'assignees': [], 'milestone': None, 'comments': 0,
'created_at': '2025-06-05T09:30:30Z', 'updated_at': '2025-06-05T09:30:30Z',
'closed_at': None, 'author_association': 'NONE', 'type': None,
'active_lock_reason': None, 'sub_issues_summary': {'total': 0, 'completed': 0,
'percent_completed': 0}, 'body': '### Pandas version checks\n\n- [x] I have
checked that this issue has not already been reported.\n\n- [x] I have confirmed
this bug exists on the [latest
version](https://pandas.pydata.org/docs/whatsnew/index.html) of pandas.\n\n- [ ]
I have confirmed this bug exists on the [main branch](https://pandas.pydata.org/
docs/dev/getting_started/install.html#installing-the-development-version-of-
pandas) of pandas.\n\n\n### Reproducible Example\n\n```python\nimport pandas as
pd\npd.DataFrame({'x': [1], 'y': [2]}).apply(list,
axis="columns")\npd.DataFrame({'x': [1], 'y': [2]}).apply(list[int],
axis="columns")\n``` \n\n### Issue Description\n\nTraceback:\n\n```python\n[...
skipping similar frames: Series.apply at line 4935 (593 times),
NDFrameApply.agg_or_apply_list_like at line 744 (592 times), SeriesApply.apply
at line 1412 (592 times), Apply.apply_list_or_dict_like at line 630 (592 times),
Apply.compute_list_like at line 369 (592 times)]\n\nFile
/opt/homebrew/envs/pandera/lib/python3.12/site-
packages/pandas/core/apply.py:1412, in SeriesApply.apply(self)\n    ... \n\nFile
/opt/homebrew/envs/pandera/lib/python3.12/site-
packages/pandas/core/apply.py:630, in Apply.apply_list_or_dict_like(self)\n
... \n\nFile /opt/homebrew/envs/pandera/lib/python3.12/site-
packages/pandas/core/apply.py:744, in NDFrameApply.agg_or_apply_list_like(self,
op_name)\n    ... \n\nFile /opt/homebrew/envs/pandera/lib/python3.12/site-
packages/pandas/core/apply.py:369, in Apply.compute_list_like(self, op_name,
selected_obj, kwargs)\n    ... \n\nFile
/opt/homebrew/envs/pandera/lib/python3.12/site-
packages/pandas/core/series.py:4935, in Series.apply(self, func, convert_dtype,
args, by_row, **kwargs)\n    ... \n\nFile
/opt/homebrew/envs/pandera/lib/python3.12/site-
packages/pandas/core/apply.py:1407, in SeriesApply.apply(self)\n    1404 def
apply(self) -> DataFrame | Series:\n    1405     obj = self.obj\n-> 1407     if
len(obj) == 0:\n    1408         return self.apply_empty_result()\n    1410     #
dispatch to handle list-like or dict-like\n\nFile
/opt/homebrew/envs/pandera/lib/python3.12/site-
packages/pandas/core/series.py:918, in Series.__len__(self)\n    914 def
__len__(self) -> int:\n    915     """\n    916     Return the length of the
Series.\n    917     """\n-> 918     return len(self._mgr)\n\nFile
/opt/homebrew/envs/pandera/lib/python3.12/site-
packages/pandas/core/internals/base.py:76, in DataManager.__len__(self)\n    74
@final\n    75 def __len__(self) -> int:\n--> 76     return
len(self.items)\n\nRecursionError: maximum recursion depth exceeded\n>
/opt/homebrew/envs/pandera/lib/python3.12/site-

```



```

packages/pandas/core/internals/base.py(76).__len__()\n      74      @final\n      75
def __len__(self) -> int:\n----> 76          return len(self.items)\n      77 \n
78      @property\n\nnipdb>\n```\n\nwhere `self.func` is regarded as list of
funcs, which lead to the bug \n\n```\npython\n>
/opt/homebrew/envs/pandera/lib/python3.12/site-
packages/pandas/core/apply.py(1412)apply()\n      1410          # dispatch to handle
list-like or dict-like\n      1411          if is_list_like(self.func):\n--> 1412
return self.apply_list_or_dict_like()\n      1413 \n      1414          if
isinstance(self.func, str):\n\nnipdb> p self.func\n*list[int]\n\nnipdb> p
is_list_like(self.func)\nTrue\n\nnipdb>\n```\n\n### Expected Behavior\n\nExpected
output:\n> 0      [1, 2]\n> dtype: object\n\n### Installed
Versions\n\n<details>\n\nINSTALLED VERSIONS\n-----\ncommit
: 2cc37625532045f4ac55b27176454bbbc9baf213\npython
:
3.12.9\npython-bits      : 64\nOS      : Darwin\nOS-release
: 24.5.0\nVersion      : Darwin Kernel Version 24.5.0: Tue Apr 22
19:54:33 PDT 2025; root:xnu-11417.121.6~2/RELEASE_ARM64_T8122\nmachine
: arm64\nprocessor      : arm\nbyteorder : little\nLC_ALL
: None\nLANG      : zh_CN.UTF-8\nLOCALE      :
zh_CN.UTF-8\n\npandas      : 2.3.0\nnumpy      :
2.2.6\npytz      : 2025.2\ndateutil      : 2.9.0.post0\npip
: 25.0.1\nCython      : None\nsphinx      : None\nIPython
: 8.31.0\nadbc-driver-postgresql: None\nadbc-driver-sqlite : None\nbs4
: None\nblosc      : None\nbottleneck : None\ndataframe-
api-compat : None\nfastparquet      : None\nfsspec      :
None\nhtml5lib      : None\nhypothesis : None\nngcsfs
: None\njinja2      : None\nlxml.etree : None\nmatplotlib
: None\nnumba      : None\nnumexpr    : None\nnodfpy
: None\nopenpyxl    : None\npandas_gbq : None\npsycopg2
: None\npymysql     : None\npyarrow   : None\npyreadstat
: None\npytest      : 8.3.4\npython-calamine : None\npyxlsb
: None\ns3fs        : None\nscipy     : None\nsqlalchemy
: None\ntables      : None\nxtabulate : None\nxarray
: None\nxlrd        : None\nxlswriter : None\nzstandard
: 0.23.0\ntzdata    : 2025.2\nqtpy    : None\npyqt5
: None\n\n</details>\n', 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61565/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61565/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61563', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61563/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61563/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61563/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61563', 'id':
3119772338, 'node_id': 'I_kwDOAAOYD8658_Ky', 'number': 61563, 'title': 'Failed

```

```

to install pandas==2.3.0 with Python 3.9', 'user': {'login': 'ryanchao2012',
'id': 7628311, 'node_id': 'MDQ6VXNlcjc2MjgzMTE=', 'avatar_url':
'https://avatars.githubusercontent.com/u/7628311?v=4', 'gravatar_id': '', 'url':
'https://api.github.com/users/ryanchao2012', 'html_url':
'https://github.com/ryanchao2012', 'followers_url':
'https://api.github.com/users/ryanchao2012/followers', 'following_url':
'https://api.github.com/users/ryanchao2012/following{/other_user}', 'gists_url':
'https://api.github.com/users/ryanchao2012/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/ryanchao2012/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/ryanchao2012/subscriptions',
'organizations_url': 'https://api.github.com/users/ryanchao2012/orgs',
'repos_url': 'https://api.github.com/users/ryanchao2012/repos', 'events_url':
'https://api.github.com/users/ryanchao2012/events{/privacy}',
'received_events_url':
'https://api.github.com/users/ryanchao2012/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{'id': 129350,
'node_id': 'MDU6TGFiZWwzMjkzNTA=', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Build', 'name': 'Build', 'color': '75507B', 'default': False,
'description': 'Library building on various platforms'}, {'id': 1954720290,
'node_id': 'MDU6TGFiZWwzOTUONzIwMjkw', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Needs%20Triage', 'name':
'Needs Triage', 'color': '0052cc', 'default': False, 'description': 'Issue that
has not been reviewed by a pandas team member'}], 'state': 'open', 'locked':
False, 'assignee': None, 'assignees': [], 'milestone': None, 'comments': 16,
'created_at': '2025-06-05T04:11:02Z', 'updated_at': '2025-06-05T20:38:43Z',
'closed_at': None, 'author_association': 'NONE', 'type': None,
'active_lock_reason': None, 'sub_issues_summary': {'total': 0, 'completed': 0,
'percent_completed': 0}, 'body': '### Installation check\n\n- [x] I have read
the [installation guide](https://pandas.pydata.org/pandas-
docs/stable/getting_started/install.html#installing-pandas).\n\n\n###
Platform\n\nLinux-5.10.195-1.20230921.el7.x86_64-x86_64-with-glibc2.17\n\n###
Installation Method\n\npip install\n\n### pandas Version\n\n2.3.0\n\n### Python
Version\n\n3.9.15\n\n### Installation Logs\n\n<details>\n(base)
[root@64bf929a621d7dafeb18b348 ~]# python -c \'import platform;
print(platform.platform())\'
Linux-5.10.195-1.20230921.el7.x86_64-x86_64-with-
glibc2.17\n\n(base) [root@64bf929a621d7dafeb18b348 ~]# pip install pandas
-U\nRequirement already satisfied: pandas in /opt/conda/lib/python3.9/site-
packages (1.5.0)\nCollecting pandas\n Downloading pandas-2.3.0.tar.gz (4.5
MB)\n
4.5/4.5 MB 90.9 MB/s eta
0:00:00\n Installing build dependencies ... done\n Getting requirements to
build wheel ... done\n Installing backend dependencies ... done\n Preparing
metadata (pyproject.toml) ... error\n error: subprocess-exited-with-error\n\n
× Preparing metadata (pyproject.toml) did not run successfully.\n exit code:
1\n > [152 lines of output]\n + meson setup /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86 /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/.mesonpy-lsu89q1a
-Dbuildtype=release -Db_ndebug=if-release -Db_vscrt=md --vsenv --native-
file=/tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/.mesonpy-

```

```

lsu89q1a/meson-python-native-file.ini\n      The Meson build system\n
Version: 1.8.1\n      Source dir: /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86\n      Build dir: /tmp/pip-
install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/.mesonpy-lsu89q1a\n
Build type: native build\n      Project name: pandas\n      Project version:
2.3.0\n      C compiler for the host machine: cc (gcc 4.8.5 "cc (GCC) 4.8.5
20150623 (Red Hat 4.8.5-44)")\n      C linker for the host machine: cc ld.bfd
2.27-44\n      C++ compiler for the host machine: c++ (gcc 4.8.5 "c++ (GCC)
4.8.5 20150623 (Red Hat 4.8.5-44)")\n      C++ linker for the host machine: c++
ld.bfd 2.27-44\n      Cython compiler for the host machine: cython (cython
3.1.1)\n      Host machine cpu family: x86_64\n      Host machine cpu: x86_64\n
Program python found: YES (/opt/conda/bin/python)\n      Found pkg-config: YES
(/usr/bin/pkg-config) 0.27.1\n      Run-time dependency python found: YES 3.9\n
Build targets in project: 53\n\n      pandas 2.3.0\n\n      User defined
options\n      Native files: /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/.mesonpy-lsu89q1a/meson-python-
native-file.ini\n      b_ndebug      : if-release\n      b_vsctrl      :
md\n      buildtype      : release\n      vsenv      : true\n\n      Found
ninja-1.11.1.git.kitware.jobserver-1 at /tmp/pip-build-env-
ltjugjgm/normal/bin/ninja\n\n      Visual Studio environment is needed to run
Ninja. It is recommended to use Meson wrapper:\n      /tmp/pip-build-env-
ltjugjgm/overlay/bin/meson compile -C .\n      + /tmp/pip-build-env-
ltjugjgm/normal/bin/ninja\n      [1/151] Generating
pandas/_libs/intervaltree_helper_pxi with a custom command\n      [2/151]
Generating pandas/_libs/hashtable_func_helper_pxi with a custom command\n
[3/151] Generating pandas/_libs/algos_common_helper_pxi with a custom command\n
[4/151] Generating pandas/_libs/khash_primitive_helper_pxi with a custom
command\n      [5/151] Generating pandas/_libs/index_class_helper_pxi with a
custom command\n      [6/151] Generating pandas/_libs/algos_take_helper_pxi with
a custom command\n      [7/151] Generating
pandas/_libs/hashtable_class_helper_pxi with a custom command\n      [8/151]
Copying file pandas/_init_.py\n      [9/151] Generating
pandas/_libs/sparse_op_helper_pxi with a custom command\n      [10/151]
Compiling C object pandas/_libs/json.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_ujson_lib_ultrajsonenc.c.o\n      FAILED:
pandas/_libs/json.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_ujson_lib_ultrajsonenc.c.o\n      cc
-Ipandas/_libs/json.cpython-39-x86_64-linux-gnu.so.p -Ipandas/_libs
-I../pandas/_libs -I../../../../pip-build-env-ltjugjgm/overlay/lib/python3.9/site-
packages/numpy/_core/include -I../pandas/_libs/include
-I/opt/conda/include/python3.9 -fvisibility=hidden -DNDEBUG
-D_FILE_OFFSET_BITS=64 -Wall -Winvalid-pch -Wextra -std=c11 -O3
-DNPY_NO_DEPRECATED_API=0 -DNPY_TARGET_VERSION=NPY_1_21_API_VERSION -fPIC -MD
-MQ pandas/_libs/json.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_ujson_lib_ultrajsonenc.c.o -MF
pandas/_libs/json.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_ujson_lib_ultrajsonenc.c.o.d -o
pandas/_libs/json.cpython-39-x86_64-linux-

```

```

gnu.so.p/src_vendored_ujson_lib_ultrajsonenc.c.o -c
../pandas/_libs/src/vendored/ujson/lib/ultrajsonenc.c\n      In file included
from ../pandas/_libs/src/vendored/ujson/lib/ultrajsonenc.c:43:0:\n
../pandas/_libs/include/pandas/portable.h:31:22: error: missing binary operator
before token "("\n      #elif __has_attribute(__fallthrough__)\n
^\n      [11/151] Compiling C object
pandas/_libs/pandas_parser.cpython-39-x86_64-linux-gnu.so.p/src_parser_io.c.o\n
[12/151] Compiling C object
pandas/_libs/pandas_datetime.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_numpy_datetime_np_datetime.c.o\n      FAILED:
pandas/_libs/pandas_datetime.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_numpy_datetime_np_datetime.c.o\n      cc
-Ipandas/_libs/pandas_datetime.cpython-39-x86_64-linux-gnu.so.p -Ipandas/_libs
-I../pandas/_libs -I../.../pip-build-env-ltjugjgm/overlay/lib/python3.9/site-
packages/numpy/_core/include -I../pandas/_libs/include
-I/opt/conda/include/python3.9 -fvisibility=hidden -DNDEBUG
-D_FILE_OFFSET_BITS=64 -Wall -Winvalid-pch -Wextra -std=c11 -O3
-DNPY_NO_DEPRECATED_API=0 -DNPY_TARGET_VERSION=NPY_1_21_API_VERSION -fPIC -MD
-MQ pandas/_libs/pandas_datetime.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_numpy_datetime_np_datetime.c.o -MF
pandas/_libs/pandas_datetime.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_numpy_datetime_np_datetime.c.o.d -o
pandas/_libs/pandas_datetime.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_numpy_datetime_np_datetime.c.o -c
../pandas/_libs/src/vendored/numpy/datetime/np_datetime.c\n
../pandas/_libs/src/vendored/numpy/datetime/np_datetime.c:57:1: error: static
assertion failed: "__has_builtin not detected; please try a newer compiler"\n
_Static_assert(0, "__has_builtin not detected; please try a newer compiler");\n
^\n      ../pandas/_libs/src/vendored/numpy/datetime/np_datetime.c: In function
'scaleYearToEpoch':\n
../pandas/_libs/src/vendored/numpy/datetime/np_datetime.c:343:3: warning:
implicit declaration of function 'checked_int64_sub' [-Wimplicit-function-
declaration]\n      return checked_int64_sub(year, 1970, result);\n
^\n      ../pandas/_libs/src/vendored/numpy/datetime/np_datetime.c: In function
'scaleYearsToMonths':\n
../pandas/_libs/src/vendored/numpy/datetime/np_datetime.c:347:3: warning:
implicit declaration of function 'checked_int64_mul' [-Wimplicit-function-
declaration]\n      return checked_int64_mul(years, 12, result);\n
^\n      ../pandas/_libs/src/vendored/numpy/datetime/np_datetime.c: In function
'numpy_datetimestruct_to_datetime':\n
../pandas/_libs/src/vendored/numpy/datetime/np_datetime.c:425:5: warning:
implicit declaration of function 'checked_int64_add' [-Wimplicit-function-
declaration]\n      PD_CHECK_OVERFLOW(checked_int64_add(months,
months_adder, &months));\n      ^\n      [13/151] Compiling C object
pandas/_libs/pandas_parser.cpython-39-x86_64-linux-
gnu.so.p/src_parser_pd_parser.c.o\n      [14/151] Compiling C object
pandas/_libs/pandas_datetime.cpython-39-x86_64-linux-
gnu.so.p/src_datetime_date_conversions.c.o\n      [15/151] Compiling C object

```

```

pandas/_libs/json.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_ujson_python_JSONtoObj.c.o\n      [16/151] Compiling C
object pandas/_libs/pandas_datetime.cpython-39-x86_64-linux-
gnu.so.p/src_datetime_pd_datetime.c.o\n      [17/151] Compiling C object
pandas/_libs/json.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_ujson_python_ujson.c.o\n      [18/151] Compiling C object
pandas/_libs/pandas_datetime.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_numpy_datetime_np_datetime_strings.c.o\n      [19/151]
Compiling C object pandas/_libs/json.cpython-39-x86_64-linux-
gnu.so.p/src_vendored_ujson_python_objToJSON.c.o\n      [20/151] Compiling C
object pandas/_libs/tslibs/parsing.cpython-39-x86_64-linux-
gnu.so.p/...src_parser_tokenizer.c.o\n      [21/151] Compiling Cython source
/tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/indexing.pyx\n
[22/151] Compiling C object pandas/_libs/pandas_parser.cpython-39-x86_64-linux-
gnu.so.p/src_parser_tokenizer.c.o\n      [23/151] Compiling C object
pandas/_libs/lib.cpython-39-x86_64-linux-gnu.so.p/src_parser_tokenizer.c.o\n
[24/151] Compiling Cython source /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d
0e8725d2463640de86/pandas/_libs/tslibs/ccalendar.pyx\n      [25/151] Compiling
Cython source /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/base.pyx\n
[26/151] Compiling Cython source /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d
0e8725d2463640de86/pandas/_libs/tslibs/np_datetime.pyx\n      warning: /tmp/pip-
install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/uti
l.pxd:188:38: noexcept clause is ignored for function returning Python object\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:193:40: noexcept clause is ignored for function
returning Python object\n      [27/151] Compiling Cython source /tmp/pip-
install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/missing.pyx\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:188:38: noexcept clause is ignored for function
returning Python object\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342
ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept clause
is ignored for function returning Python object\n      [28/151] Compiling Cython
source /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/
_libs/tslibs/dtypes.pyx\n      [29/151] Compiling Cython source /tmp/pip-
install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/arrays.pyx\n
[30/151] Compiling Cython source /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/hashing.pyx\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:188:38: noexcept clause is ignored for function
returning Python object\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342
ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept clause
is ignored for function returning Python object\n      [31/151] Compiling Cython
source /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/
_libs/tslibs/natype.pyx\n      warning: /tmp/pip-install-gp_gpioe/pandas_860834

```

```

2ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:188:38: noexcept clause
is ignored for function returning Python object\n      warning: /tmp/pip-
install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/uti
l.pxd:193:40: noexcept clause is ignored for function returning Python object\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/natype.pyx:79:0: Global name __nat_unpickle matched from within
class scope in contradiction to to Python \'class private name\' rules. This may
change in a future release.\n      warning: /tmp/pip-install-gp_gpioe/pandas_860
8342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/natype.pyx:79:0: Global name
__nat_unpickle matched from within class scope in contradiction to to Python
\'class private name\' rules. This may change in a future release.\n
[32/151] Compiling Cython source /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d
0e8725d2463640de86/pandas/_libs/tslibs/vectorized.pyx\n      [33/151] Compiling
Cython source /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/
pandas/_libs/tslibs/fields.pyx\n      [34/151] Compiling Cython source /tmp/pip-
install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/internals.pyx\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:188:38: noexcept clause is ignored for function
returning Python object\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342
ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept clause
is ignored for function returning Python object\n      [35/151] Compiling Cython
source /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/
_libs/tslibs/conversion.pyx\n      warning: /tmp/pip-install-gp_gpioe/pandas_860
8342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:188:38: noexcept
clause is ignored for function returning Python object\n      warning: /tmp/pip-
install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/uti
l.pxd:193:40: noexcept clause is ignored for function returning Python object\n
[36/151] Compiling Cython source /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d
0e8725d2463640de86/pandas/_libs/tslibs/parsing.pyx\n      warning: /tmp/pip-
install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/uti
l.pxd:188:38: noexcept clause is ignored for function returning Python object\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:193:40: noexcept clause is ignored for function
returning Python object\n      [37/151] Compiling Cython source /tmp/pip-
install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/tim
ezones.pyx\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725
d2463640de86/pandas/_libs/tslibs/util.pxd:188:38: noexcept clause is ignored for
function returning Python object\n      warning: /tmp/pip-install-gp_gpioe/panda
s_8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept
clause is ignored for function returning Python object\n      [38/151] Compiling
Cython source /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/
pandas/_libs/tslibs/tzconversion.pyx\n      [39/151] Compiling Cython source
/tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/t
slibs/strptime.pyx\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb16
4d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:188:38: noexcept clause is
ignored for function returning Python object\n      warning: /tmp/pip-install-gp
_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:

```

```

40: noexcept clause is ignored for function returning Python object\n
[40/151] Compiling Cython source /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/parsers.pyx\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:188:38: noexcept clause is ignored for function
returning Python object\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342
ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept clause
is ignored for function returning Python object\n      warning: /tmp/pip-
install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/parsers.py
x:1605:18: noexcept clause is ignored for function returning Python object\n
[41/151] Compiling Cython source /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d
0e8725d2463640de86/pandas/_libs/tslibs/timestamps.pyx\n      warning: /tmp/pip-
install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/uti
l.pxd:188:38: noexcept clause is ignored for function returning Python object\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:193:40: noexcept clause is ignored for function
returning Python object\n      [42/151] Compiling Cython source /tmp/pip-
install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/per
iod.pyx\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d24
63640de86/pandas/_libs/tslibs/util.pxd:188:38: noexcept clause is ignored for
function returning Python object\n      warning: /tmp/pip-install-gp_gpioe/panda
s_8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept
clause is ignored for function returning Python object\n      [43/151] Compiling
Cython source /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/
pandas/_libs/tslibs/timedeltas.pyx\n      warning: /tmp/pip-install-gp_gpioe/pan
das_8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:188:38:
noexcept clause is ignored for function returning Python object\n      warning:
/tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/t
slibs/util.pxd:193:40: noexcept clause is ignored for function returning Python
object\n      [44/151] Compiling Cython source /tmp/pip-install-gp_gpioe/pandas_
8608342ddb164d0e8725d2463640de86/pandas/_libs/tslibs/offsets.pyx\n      warning:
/tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/t
slibs/util.pxd:188:38: noexcept clause is ignored for function returning Python
object\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d246
3640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept clause is ignored for
function returning Python object\n      [45/151] Compiling Cython source
/tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/lib.pyx\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:188:38: noexcept clause is ignored for function
returning Python object\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342
ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept clause
is ignored for function returning Python object\n      [46/151] Compiling Cython
source /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/index.pyx\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:188:38: noexcept clause is ignored for function
returning Python object\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342

```

```

ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept clause
is ignored for function returning Python object\n      [47/151] Compiling Cython
source /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/interval.pyx\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:188:38: noexcept clause is ignored for function
returning Python object\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342
ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept clause
is ignored for function returning Python object\n      [48/151] Compiling Cython
source /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/join.pyx\n
[49/151] Compiling Cython source /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/hashtable.pyx\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:188:38: noexcept clause is ignored for function
returning Python object\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342
ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept clause
is ignored for function returning Python object\n      [50/151] Compiling Cython
source /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/algos.pyx\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:188:38: noexcept clause is ignored for function
returning Python object\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342
ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept clause
is ignored for function returning Python object\n      [51/151] Compiling Cython
source /tmp/pip-install-
gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/pandas/_libs/groupby.pyx\n
warning: /tmp/pip-install-gp_gpioe/pandas_8608342ddb164d0e8725d2463640de86/panda
s/_libs/tslibs/util.pxd:188:38: noexcept clause is ignored for function
returning Python object\n      warning: /tmp/pip-install-gp_gpioe/pandas_8608342
ddb164d0e8725d2463640de86/pandas/_libs/tslibs/util.pxd:193:40: noexcept clause
is ignored for function returning Python object\n      ninja: build stopped:
subcommand failed.\n      [end of output]\n\n note: This error originates from
a subprocess, and is likely not a problem with pip.\nerror: metadata-generation-
failed\n\nx Encountered error while generating package metadata.\n > See above
for output.\n\nnote: This is an issue with the package mentioned above, not
pip.\nhint: See above for details.\n\n\n</details>\n', 'closed_by': None,
'reactions': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61563/reactions', 'total_count': 13, '+1': 11, '-1': 0,
'laugh': 0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 2},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61563/timeline', 'performed_via_github_app': None,
'state_reason': None}, {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61557', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61557/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61557/comments', 'events_url':

```



```

'https://api.github.com/repos/pandas-dev/pandas/issues/61557/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61557', 'id':
3118522709, 'node_id': 'I_kwDOAAOYD86540FV', 'number': 61557, 'title': 'BUG:
regex match in compliance tests no longer match pytest expected inputs', 'user':
{'login': 'chalmerlowe', 'id': 7291104, 'node_id': 'MDQ6VXNlcjcyOTExMDQ=',
'avatar_url': 'https://avatars.githubusercontent.com/u/7291104?v=4',
'gravatar_id': '', 'url': 'https://api.github.com/users/chalmerlowe',
'html_url': 'https://github.com/chalmerlowe', 'followers_url':
'https://api.github.com/users/chalmerlowe/followers', 'following_url':
'https://api.github.com/users/chalmerlowe/following{/other_user}', 'gists_url':
'https://api.github.com/users/chalmerlowe/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/chalmerlowe/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/chalmerlowe/subscriptions',
'organizations_url': 'https://api.github.com/users/chalmerlowe/orgs',
'repos_url': 'https://api.github.com/users/chalmerlowe/repos', 'events_url':
'https://api.github.com/users/chalmerlowe/events{/privacy}',
'received_events_url':
'https://api.github.com/users/chalmerlowe/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{'id': 127685,
'node_id': 'MDU6TGFiZWwzMjc2ODU=', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Testing', 'name': 'Testing', 'color': 'C4A000', 'default':
False, 'description': 'pandas testing functions or related to the test suite'},
{'id': 42670965, 'node_id': 'MDU6TGFiZWw0MjY3MDk2NQ==', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Error%20Reporting',
'name': 'Error Reporting', 'color': 'ffa0ff', 'default': False, 'description':
'Incorrect or improved errors from pandas'}, {'id': 717120670, 'node_id':
'MDU6TGFiZWw3MTcxMjA2NzA=', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/good%20first%20issue', 'name': 'good first issue', 'color':
'0e8a16', 'default': True, 'description': None}], 'state': 'open', 'locked':
False, 'assignee': None, 'assignees': [], 'milestone': None, 'comments': 2,
'created_at': '2025-06-04T17:11:35Z', 'updated_at': '2025-06-05T17:46:38Z',
'closed_at': None, 'author_association': 'CONTRIBUTOR', 'type': None,
'active_lock_reason': None, 'sub_issues_summary': {'total': 0, 'completed': 0,
'percent_completed': 0}, 'body': '### Pandas version checks\n\n [x] I have
checked that this issue has not already been reported.\n\n [x] I have confirmed
this bug exists on the [latest
version](https://pandas.pydata.org/docs/whatsnew/index.html) of pandas.\n\n [ ]
I have confirmed this bug exists on the [main branch](https://pandas.pydata.org/
docs/dev/getting_started/install.html#installing-the-development-version-of-
pandas) of pandas.\n\n\n### Reproducible Example\n\n`python\nN/A\n```\n\n###
Issue Description\n\nWhen I run compliance tests in python-db-dtype-pandas (a
support file used by python-bigquery) I am getting multiple warnings (which
cause test failures) due to a recent update in how pytest handles regex
matches.\n\nIn pandas release 2.2.3 there is a snippet of code:\n```\ndef
test_take_pandas_style_negative_raises(self, data, na_value):\n    with
pytest.raises(ValueError, match=""):\n```\n\nPytest returns this
Warning:\n\n```\npytest.PytestWarning: matching against an empty string will
*always* pass. If you want to check for an empty message you need to pass

```

```

\`$`. If you don't want to match you should pass `None` or leave out the
parameter.\n``\n\nThis warning shows up in association with each of these
pandas tests (it may occur with other tests, but these are the only ones that my
tests revealed.):`n``\nFAILED
...:TestGetitem::test_take_pandas_style_negative_raises\nFAILED
...:TestMethods::test_argmax_argmin_no_skipna_notimplemented\nFAILED
...:TestSetitem::test_setitem_invalid\nFAILED
...:TestJSONArrayGetitem::test_take_pandas_style_negative_raises\nFAILED
...:TestJSONArrayMethods::test_argmax_argmin_no_skipna_notimplemented\nFAILED
...:TestJSONArraySetitem::test_setitem_invalid\n``\n\n\n### Expected
Behavior\n\nN/A\n\n### Installed Versions\n\nN/A', 'closed_by': None,
'reactions': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61557/reactions', 'total_count': 0, '+1': 0, '-1': 0, 'laugh':
0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 0},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61557/timeline', 'performed_via_github_app': None,
'state_reason': None}, {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61555', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61555/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61555/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61555/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61555', 'id': 3117071656,
'node_id': 'PR_kwD0AA0YD86Y-985', 'number': 61555, 'title': 'Upgrade Python
version for development environment', 'user': {'login': 'datapythonista', 'id':
10058240, 'node_id': 'MDQ6VXNlcjEwMDU4MjQw', 'avatar_url':
'https://avatars.githubusercontent.com/u/10058240?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/datapythonista', 'html_url':
'https://github.com/datapythonista', 'followers_url':
'https://api.github.com/users/datapythonista/followers', 'following_url':
'https://api.github.com/users/datapythonista/following{/other_user}',
'gists_url': 'https://api.github.com/users/datapythonista/gists{/gist_id}',
'starred_url':
'https://api.github.com/users/datapythonista/starred{/owner}/{/repo}',
'subscriptions_url':
'https://api.github.com/users/datapythonista/subscriptions',
'organizations_url': 'https://api.github.com/users/datapythonista/orgs',
'repos_url': 'https://api.github.com/users/datapythonista/repos', 'events_url':
'https://api.github.com/users/datapythonista/events{/privacy}',
'received_events_url':
'https://api.github.com/users/datapythonista/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{'id': 48070600,
'node_id': 'MDU6TGFiZWw0ODAzMDYwMA==', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/CI', 'name': 'CI',
'color': 'a2bca7', 'default': False, 'description': 'Continuous Integration'},
{'id': 527603109, 'node_id': 'MDU6TGFiZWw1Mjc2MDMxMDk=', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Dependencies', 'name':

```

```

'Dependencies', 'color': 'd93f0b', 'default': False, 'description': 'Required
and optional dependencies'}], 'state': 'open', 'locked': False, 'assignee':
None, 'assignees': [], 'milestone': None, 'comments': 2, 'created_at':
'2025-06-04T09:04:21Z', 'updated_at': '2025-06-04T23:38:12Z', 'closed_at': None,
'author_association': 'MEMBER', 'type': None, 'active_lock_reason': None,
'draft': False, 'pull_request': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/pulls/61555', 'html_url': 'https://github.com/pandas-
dev/pandas/pull/61555', 'diff_url': 'https://github.com/pandas-
dev/pandas/pull/61555.diff', 'patch_url': 'https://github.com/pandas-
dev/pandas/pull/61555.patch', 'merged_at': None}, 'body': "I guess the only
reason we require Python 3.10 for the development environments and run the docs
is because it wasn't updated for a long time. Not sure if it'd be better to
unpin, which I assume would get us the latest version of Python. Happy to do
that here too.", 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61555/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61555/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61554', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61554/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61554/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61554/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61554', 'id':
3117048901, 'node_id': 'I_kwDOAAOYD865ymRF', 'number': 61554, 'title': 'BUG:
duplicated() raises error with singleton set as subset', 'user': {'login':
'camold', 'id': 2109454, 'node_id': 'MDQ6VXNlcjIxMDkONTQ=', 'avatar_url':
'https://avatars.githubusercontent.com/u/2109454?v=4', 'gravatar_id': '', 'url':
'https://api.github.com/users/camold', 'html_url': 'https://github.com/camold',
'followers_url': 'https://api.github.com/users/camold/followers',
'following_url': 'https://api.github.com/users/camold/following{/other_user}',
'gists_url': 'https://api.github.com/users/camold/gists{/gist_id}',
'starred_url': 'https://api.github.com/users/camold/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/camold/subscriptions',
'organizations_url': 'https://api.github.com/users/camold/orgs', 'repos_url':
'https://api.github.com/users/camold/repos', 'events_url':
'https://api.github.com/users/camold/events{/privacy}', 'received_events_url':
'https://api.github.com/users/camold/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{'id': 76811,
'node_id': 'MDU6TGFiZWw3NjgxMQ==', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Bug', 'name': 'Bug', 'color': 'e10c02', 'default': False,
'description': None}, {'id': 1954720290, 'node_id': 'MDU6TGFiZWw3OTU0NzIwMjkw',
'url': 'https://api.github.com/repos/pandas-dev/pandas/labels/Needs%20Triage',
'name': 'Needs Triage', 'color': '0052cc', 'default': False, 'description':
'Issue that has not been reviewed by a pandas team member'}], 'state': 'open',
'locked': False, 'assignee': None, 'assignees': [], 'milestone': None,

```

```

'comments': 0, 'created_at': '2025-06-04T08:57:11Z', 'updated_at':
'2025-06-04T08:57:11Z', 'closed_at': None, 'author_association': 'NONE', 'type':
None, 'active_lock_reason': None, 'sub_issues_summary': {'total': 0,
'completed': 0, 'percent_completed': 0}, 'body': '### Pandas version checks\n\n-
[x] I have checked that this issue has not already been reported.\n\n- [x] I
have confirmed this bug exists on the [latest
version](https://pandas.pydata.org/docs/whatsnew/index.html) of pandas.\n\n- [ ]
I have confirmed this bug exists on the [main branch](https://pandas.pydata.org/
docs/dev/getting_started/install.html#installing-the-development-version-of-
pandas) of pandas.\n\n\n### Reproducible Example\n\n```\npython\nimport pandas as
pd\nndf = pd.DataFrame([{"a": "foo", "b": "bar"}])\nndf.duplicated(subset={"a"}) #
raises error\nndf.duplicated(subset=["a"]) # works\nndf.duplicated(subset=("a",))
# works\nndf.duplicates(subset=["a","b"]) # works\n```\n\n\n### Issue
Description\n\nProviding a singleton set to the subset parameter raises an
error.\n\n\n### Expected Behavior\n\nShould work normally without having to
convert the input to list or tuple.\n\n\n### Installed
Versions\n\n<details>\n\n\nINSTALLED VERSIONS\n\n-----\n\ncommit
: 0691c5cf90477d3503834d983f69350f250a6ff7\n\npython
: 3.12.3\n\npython-bits
: 64\n\nOS
: Linux\n\nOS-release
: 6.11.0-26-generic\n\nVersion
: #26~24.04.1-Ubuntu SMP
PREEMPT_DYNAMIC Thu Apr 17 19:20:47 UTC 2\n\nmachine
: x86_64\n\nprocessor
: x86_64\n\nbyteorder
: little\n\nLC_ALL
: None\n\nLANG
: de_DE.UTF-8\n\nLOCALE
: de_DE.UTF-8\n\n\npandas
: 2.2.3\n\nnumpy
: 2.2.0\n\npytz
: 2024.2\n\ndateutil
: 2.9.0.post0\n\npip
: 24.0\n\nCython
: None\n\nsphinx
: None\n\nIPython
: 8.29.0\n\nadbc-driver-postgresql: None\n\nadbc-driver-sqlite
: None\n\nbs4
: 4.12.3\n\nblosc
: None\n\nbottleneck
: None\n\ndataframe-
api-compat : None\n\nfastparquet
: None\n\nfsspec
: None\n\nhtml5lib
: None\n\nhypothesis
: None\n\nngcsfs
: None\n\njinja2
: 3.1.4\n\nlxml.etree
: 5.3.0\n\nmatplotlib
: 3.9.2\n\nnumba
: None\n\nnumexpr
: None\n\nodfpy
: None\n\nopenpyxl
: 3.1.5\n\npandas_gbq
: None\n\npsycopy2
: None\n\npymysql
: None\n\npyarrow
: None\n\npyreadstat
: None\n\npytest
: 8.3.3\n\npython-calamine
: None\n\npyxlsb
: None\n\ns3fs
: None\n\nscipy
: None\n\nsqlalchemy
: 2.0.36\n\ntables
: None\n\nxtabulate
: None\n\nxarray
: None\n\nxlrd
: None\n\nxlsxwriter
: None\n\nzstandard
: None\n\nntzdata
: 2024.2\n\nnqtpy
: None\n\npyqt5
: None\n\n\n\n</details>\n\n', 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61554/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61554/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61552', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61552/labels{/name}',

```

```

'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61552/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61552/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61552', 'id': 3116474516,
'node_id': 'PR_kwD0AA0YD86Y9COH', 'number': 61552, 'title': "DOC: Add note on
inference behavior of apply with result_type='expand'", 'user': {'login':
'pinkgradient', 'id': 101425679, 'node_id': 'U_kgDOBguiDw', 'avatar_url':
'https://avatars.githubusercontent.com/u/101425679?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/pinkgradient', 'html_url':
'https://github.com/pinkgradient', 'followers_url':
'https://api.github.com/users/pinkgradient/followers', 'following_url':
'https://api.github.com/users/pinkgradient/following{/other_user}', 'gists_url':
'https://api.github.com/users/pinkgradient/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/pinkgradient/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/pinkgradient/subscriptions',
'organizations_url': 'https://api.github.com/users/pinkgradient/orgs',
'repos_url': 'https://api.github.com/users/pinkgradient/repos', 'events_url':
'https://api.github.com/users/pinkgradient/events{/privacy}',
'received_events_url':
'https://api.github.com/users/pinkgradient/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{'id': 134699,
'node_id': 'MDU6TGFiZWwzMzQ2OTk=', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Docs', 'name': 'Docs', 'color': '3465A4', 'default': False,
'description': None}, {'id': 697792067, 'node_id': 'MDU6TGFiZWw20Tc30TIwNjc=',
'url': 'https://api.github.com/repos/pandas-dev/pandas/labels/Apply', 'name':
'Apply', 'color': 'fbca04', 'default': False, 'description': 'Apply, Aggregate,
Transform, Map'}], 'state': 'open', 'locked': False, 'assignee': None,
'assignees': [], 'milestone': None, 'comments': 3, 'created_at':
'2025-06-04T05:20:11Z', 'updated_at': '2025-06-04T11:01:04Z', 'closed_at': None,
'author_association': 'NONE', 'type': None, 'active_lock_reason': None, 'draft':
False, 'pull_request': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/pulls/61552', 'html_url': 'https://github.com/pandas-
dev/pandas/pull/61552', 'diff_url': 'https://github.com/pandas-
dev/pandas/pull/61552.diff', 'patch_url': 'https://github.com/pandas-
dev/pandas/pull/61552.patch', 'merged_at': None}, 'body': "Closes
#61057\r\n\r\n- This PR adds a docstring note about the inference behavior of
apply with result_type='expand' when function returns NaN-like values, as
discussed in issue #61057.\r\n- Added docstring to distinguish behavior for
'apply' method.", 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61552/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61552/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61551', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61551/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-

```

```

dev/pandas/issues/61551/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61551/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61551', 'id': 3116401342,
'node_id': 'PR_kwD0AA0YD86Y80iP', 'number': 61551, 'title': 'Parallelize
test_sql.py - Issue #60378', 'user': {'login': 'dShcherbakov1', 'id': 119084461,
'node_id': 'U_kgD0BxkVrQ', 'avatar_url':
'https://avatars.githubusercontent.com/u/119084461?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/dShcherbakov1', 'html_url':
'https://github.com/dShcherbakov1', 'followers_url':
'https://api.github.com/users/dShcherbakov1/followers', 'following_url':
'https://api.github.com/users/dShcherbakov1/following{/other_user}',
'gists_url': 'https://api.github.com/users/dShcherbakov1/gists{/gist_id}',
'starred_url':
'https://api.github.com/users/dShcherbakov1/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/dShcherbakov1/subscriptions',
'organizations_url': 'https://api.github.com/users/dShcherbakov1/orgs',
'repos_url': 'https://api.github.com/users/dShcherbakov1/repos', 'events_url':
'https://api.github.com/users/dShcherbakov1/events{/privacy}',
'received_events_url':
'https://api.github.com/users/dShcherbakov1/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [], 'state': 'open',
'locked': False, 'assignee': None, 'assignees': [], 'milestone': None,
'comments': 1, 'created_at': '2025-06-04T04:47:40Z', 'updated_at':
'2025-06-04T21:39:56Z', 'closed_at': None, 'author_association': 'NONE', 'type':
None, 'active_lock_reason': None, 'draft': False, 'pull_request': {'url':
'https://api.github.com/repos/pandas-dev/pandas/pulls/61551', 'html_url':
'https://github.com/pandas-dev/pandas/pull/61551', 'diff_url':
'https://github.com/pandas-dev/pandas/pull/61551.diff', 'patch_url':
'https://github.com/pandas-dev/pandas/pull/61551.patch', 'merged_at': None},
'body': "- [ ] Closes #60378\r\n- [ ] Builds off of https://github.com/pandas-
dev/pandas/pull/60595\r\n- [ ] Extends @UmbertoFasci's per-table UUID work\r\n-
[ ] Solves https://github.com/pandas-dev/pandas/pull/60595 concurrency
issue\r\n\r\n~~~ [x] Parallelizes test_sql.py via per-worker_DBs (approach
superseded due to complexity)~~~\r\n", 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61551/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61551/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61550', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61550/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61550/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61550/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61550', 'id': 3116382404,
'node_id': 'PR_kwD0AA0YD86Y8ynF', 'number': 61550, 'title': 'DOC: Remove and
Update out of date Docker Image issue with #61511', 'user': {'login':

```

```

'jacksnnn', 'id': 175151559, 'node_id': 'U_kgD0Cn CZxw', 'avatar_url':
'https://avatars.githubusercontent.com/u/175151559?v=4', 'gravavatar_id': '',
'url': 'https://api.github.com/users/jacksnnn', 'html_url':
'https://github.com/jacksnnn', 'followers_url':
'https://api.github.com/users/jacksnnn/followers', 'following_url':
'https://api.github.com/users/jacksnnn/following{/other_user}', 'gists_url':
'https://api.github.com/users/jacksnnn/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/jacksnnn/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/jacksnnn/subscriptions',
'organizations_url': 'https://api.github.com/users/jacksnnn/orgs', 'repos_url':
'https://api.github.com/users/jacksnnn/repos', 'events_url':
'https://api.github.com/users/jacksnnn/events{/privacy}', 'received_events_url':
'https://api.github.com/users/jacksnnn/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{ 'id': 134699,
'node_id': 'MDU6TGFiZWwzMzQ2OTk=', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Docs', 'name': 'Docs', 'color': '3465A4', 'default': False,
'description': None}], 'state': 'open', 'locked': False, 'assignee': None,
'assignees': [], 'milestone': None, 'comments': 1, 'created_at':
'2025-06-04T04:40:39Z', 'updated_at': '2025-06-04T12:49:55Z', 'closed_at': None,
'author_association': 'NONE', 'type': None, 'active_lock_reason': None, 'draft':
False, 'pull_request': { 'url': 'https://api.github.com/repos/pandas-
dev/pandas/pulls/61550', 'html_url': 'https://github.com/pandas-
dev/pandas/pull/61550', 'diff_url': 'https://github.com/pandas-
dev/pandas/pull/61550.diff', 'patch_url': 'https://github.com/pandas-
dev/pandas/pull/61550.patch', 'merged_at': None}, 'body': '- [ ] Addresses &
closes [DOC: Docker image provided on "Debugging C extensions" is out of date
#61511](https://github.com/pandas-dev/pandas/issues/61511) \r\n- [ ] All [code
checks passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#pre-commit).\r\n\r\n',
'closed_by': None, 'reactions': { 'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61550/reactions', 'total_count': 0, '+1': 0, '-1': 0, 'laugh':
0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 0},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61550/timeline', 'performed_via_github_app': None,
'state_reason': None}, { 'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61546', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61546/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61546/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61546/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61546', 'id': 3115813278,
'node_id': 'PR_kwD0AA0YD86Y69PY', 'number': 61546, 'title': 'ENH: Adding
DataFrame plotting benchmarks for large datasets', 'user': { 'login': 'shadnikn',
'id': 147014887, 'node_id': 'U_kgD0CMNE5w', 'avatar_url':
'https://avatars.githubusercontent.com/u/147014887?v=4', 'gravavatar_id': '',
'url': 'https://api.github.com/users/shadnikn', 'html_url':
'https://github.com/shadnikn', 'followers_url':

```

```

'https://api.github.com/users/shadnikn/followers', 'following_url':
'https://api.github.com/users/shadnikn/following{/other_user}', 'gists_url':
'https://api.github.com/users/shadnikn/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/shadnikn/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/shadnikn/subscriptions',
'organizations_url': 'https://api.github.com/users/shadnikn/orgs', 'repos_url':
'https://api.github.com/users/shadnikn/repos', 'events_url':
'https://api.github.com/users/shadnikn/events{/privacy}', 'received_events_url':
'https://api.github.com/users/shadnikn/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{ 'id': 732775912,
'node_id': 'MDU6TGFiZWw3MzI3NzU5MTI=', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Benchmark', 'name':
'Benchmark', 'color': 'ae68cc', 'default': False, 'description': 'Performance
(ASV) benchmarks'}], 'state': 'open', 'locked': False, 'assignee': None,
'assignees': [], 'milestone': None, 'comments': 1, 'created_at':
'2025-06-04T00:21:17Z', 'updated_at': '2025-06-04T16:40:20Z', 'closed_at': None,
'author_association': 'NONE', 'type': None, 'active_lock_reason': None, 'draft':
False, 'pull_request': { 'url': 'https://api.github.com/repos/pandas-
dev/pandas/pulls/61546', 'html_url': 'https://github.com/pandas-
dev/pandas/pull/61546', 'diff_url': 'https://github.com/pandas-
dev/pandas/pull/61546.diff', 'patch_url': 'https://github.com/pandas-
dev/pandas/pull/61546.patch', 'merged_at': None}, 'body': "- [ ] related to
#61532 - Adding in performance benchmarks for DataFrame plotting with large
datasets.\r\n- [ ] Description: Added 'DataFramePlottingLarge' benchmark class
to track performance issues related to bottlenecks in #61398 and #61532. Tests
multiple DataFrame sizes with/w/o DatetimeIndex & provides a baseline single-
column comparison.\r\n- [ ] Intended to cover:\r\n          - DataFrame sizes:
(1000,10) to (10000,10)\r\n          - DatetimeIndex vs. regular index
comparison\r\n          - Multi-column vs. single-column plotting.\r\n",
'closed_by': None, 'reactions': { 'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61546/reactions', 'total_count': 0, '+1': 0, '-1': 0, 'laugh':
0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 0},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61546/timeline', 'performed_via_github_app': None,
'state_reason': None}, { 'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61543', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61543/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61543/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61543/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61543', 'id': 3115248818,
'node_id': 'PR_kwDOAAOYD86Y5Ety', 'number': 61543, 'title': 'DOC: Fix docs for
BusinessDay constructor #52431', 'user': { 'login': 'camacluc', 'id': 114314487,
'node_id': 'U_kgDOBtBM9w', 'avatar_url':
'https://avatars.githubusercontent.com/u/114314487?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/camacluc', 'html_url':
'https://github.com/camacluc', 'followers_url':

```



```

'https://api.github.com/users/camacluc/followers', 'following_url':
'https://api.github.com/users/camacluc/following{/other_user}', 'gists_url':
'https://api.github.com/users/camacluc/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/camacluc/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/camacluc/subscriptions',
'organizations_url': 'https://api.github.com/users/camacluc/orgs', 'repos_url':
'https://api.github.com/users/camacluc/repos', 'events_url':
'https://api.github.com/users/camacluc/events{/privacy}', 'received_events_url':
'https://api.github.com/users/camacluc/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False, 'labels': [], 'state': 'open',
'locked': False, 'assignee': None, 'assignees': [], 'milestone': None,
'comments': 0, 'created_at': '2025-06-03T19:50:59Z', 'updated_at':
'2025-06-03T19:51:05Z', 'closed_at': None, 'author_association': 'NONE', 'type':
None, 'active_lock_reason': None, 'draft': False, 'pull_request': {'url':
'https://api.github.com/repos/pandas-dev/pandas/pulls/61543', 'html_url':
'https://github.com/pandas-dev/pandas/pull/61543', 'diff_url':
'https://github.com/pandas-dev/pandas/pull/61543.diff', 'patch_url':
'https://github.com/pandas-dev/pandas/pull/61543.patch', 'merged_at': None},
'body': 'fix constructor for offset BusinessDay\r\n\r\n- [ ] addresses #52431
\r\n- [ ] All [code checks passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#pre-commit).\r\n', 'closed_by':
None, 'reactions': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61543/reactions', 'total_count': 0, '+1': 0, '-1': 0, 'laugh':
0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 0},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61543/timeline', 'performed_via_github_app': None,
'state_reason': None}, {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61542', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61542/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61542/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61542/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61542', 'id': 3114631618,
'node_id': 'PR_kwD0AA0YD86Y2-S4', 'number': 61542, 'title': 'CI: Debug slow
environment solve times', 'user': {'login': 'mroeschke', 'id': 10647082,
'node_id': 'MDQ6VXNlclcjEwNjQ3MDgy', 'avatar_url':
'https://avatars.githubusercontent.com/u/10647082?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/mroeschke', 'html_url':
'https://github.com/mroeschke', 'followers_url':
'https://api.github.com/users/mroeschke/followers', 'following_url':
'https://api.github.com/users/mroeschke/following{/other_user}', 'gists_url':
'https://api.github.com/users/mroeschke/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/mroeschke/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/mroeschke/subscriptions',
'organizations_url': 'https://api.github.com/users/mroeschke/orgs', 'repos_url':
'https://api.github.com/users/mroeschke/repos', 'events_url':
'https://api.github.com/users/mroeschke/events{/privacy}'},

```

```

'received_events_url': 'https://api.github.com/users/mroeschke/received_events',
'type': 'User', 'user_view_type': 'public', 'site_admin': False}, 'labels': [],
'state': 'open', 'locked': False, 'assignee': None, 'assignees': [],
'milestone': None, 'comments': 1, 'created_at': '2025-06-03T16:19:34Z',
'updated_at': '2025-06-04T09:05:48Z', 'closed_at': None, 'author_association':
'MEMBER', 'type': None, 'active_lock_reason': None, 'draft': True,
'pull_request': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/pulls/61542', 'html_url': 'https://github.com/pandas-
dev/pandas/pull/61542', 'diff_url': 'https://github.com/pandas-
dev/pandas/pull/61542.diff', 'patch_url': 'https://github.com/pandas-
dev/pandas/pull/61542.patch', 'merged_at': None}, 'body': '- [ ] closes #xxxx
(Replace xxxx with the GitHub issue number)\r\n- [ ] [Tests added and
passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#writing-tests) if fixing a bug
or adding a new feature\r\n- [ ] All [code checks
passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#pre-commit).\r\n- [ ] Added
[type annotations](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#type-hints) to new
arguments/methods/functions.\r\n- [ ] Added an entry in the latest
`doc/source/whatsnew/vX.X.X.rst` file if fixing a bug or adding a new
feature.\r\n', 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61542/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61542/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61541', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61541/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61541/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61541/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61541', 'id': 3114329620,
'node_id': 'PR_kwDOAAOYD86Y1-gs', 'number': 61541, 'title': 'Fix Index.equals
between object and string', 'user': {'login': 'sanggon6107', 'id': 68040183,
'node_id': 'MDQ6VXNlcjY4MDQwMTgz', 'avatar_url':
'https://avatars.githubusercontent.com/u/68040183?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/sanggon6107', 'html_url':
'https://github.com/sanggon6107', 'followers_url':
'https://api.github.com/users/sanggon6107/followers', 'following_url':
'https://api.github.com/users/sanggon6107/following{/other_user}', 'gists_url':
'https://api.github.com/users/sanggon6107/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/sanggon6107/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/sanggon6107/subscriptions',
'organizations_url': 'https://api.github.com/users/sanggon6107/orgs',
'repos_url': 'https://api.github.com/users/sanggon6107/repos', 'events_url':
'https://api.github.com/users/sanggon6107/events{/privacy}',

```

```

'received_events_url':
'https://api.github.com/users/sanggon6107/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [], 'state': 'open',
'locked': False, 'assignee': None, 'assignees': [], 'milestone': None,
'comments': 0, 'created_at': '2025-06-03T14:54:11Z', 'updated_at':
'2025-06-04T13:38:41Z', 'closed_at': None, 'author_association': 'NONE', 'type':
None, 'active_lock_reason': None, 'draft': False, 'pull_request': {'url':
'https://api.github.com/repos/pandas-dev/pandas/pulls/61541', 'html_url':
'https://github.com/pandas-dev/pandas/pull/61541', 'diff_url':
'https://github.com/pandas-dev/pandas/pull/61541.diff', 'patch_url':
'https://github.com/pandas-dev/pandas/pull/61541.patch', 'merged_at': None},
'body': '- [X] closes #61099 \r\n- [X] [Tests added and
passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#writing-tests) if fixing a bug
or adding a new feature\r\n- [X] All [code checks
passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#pre-commit).\r\n- [ ] Added
[type annotations](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#type-hints) to new
arguments/methods/functions.\r\n- [ ] Added an entry in the latest
`doc/source/whatsnew/vX.X.X.rst` file if fixing a bug or adding a new
feature.\r\n\r\n## Description of the code change on `Index.equals`\r\n\r\nOn
the main branch, `Index.equals` casts `self` to `object` only when
`self.dtype.na_value` is `np.nan`. The comparison actually succeeds when
`self.dtype.na_value` is `np.nan` as below.\r\n```\r\npython\r\n>>> import pandas as
pd\r\n>>> import numpy as np\r\n\r\n>>> s1 = pd.Series([1, 2, 3], index=['a',
'b', 'c'])\r\n>>> s2 = pd.Series([4, 5, 6], index=['a', 'b',
'c'])\r\n>>> s2.index = s2.index.astype(pd.StringDtype(storage="pyarrow",
na_value=np.nan))\r\n\r\n>>> print(s1 < s2)\r\na    True\r\nb    True\r\nc
True\r\ndtype: bool\r\n```\r\nHowever, since doc stated that `dtype` is not
compared, `self` should be casted regardless of `self.dtype.na_value` so that
`self` could be compared with other dtypes as desired.\r\n\r\n\r\n## Description
of the code change on `test_mixed_col_index_dtype`\r\n\r\n`using_infer_string`
has been removed since I think that `result` should be `string` regardless of
`using_infer_string`. This is because of the code change made on `Index.equals` -
since `Index.equals` consider `df1.columns` is equal to `df2.columns`,
`Index.intersection` returns `self` (which is `string`). You could see the result
becomes `object` (which is the dtype of `df2`) in case of `result = df2 + df1`.
On the main branch, on the other hand, `Index.intersection` returns `object`
because `Index.equals` returns `False`, and then both `self` and `other` are
cast to `object` by `_find_common_type_compat`. (see `L3287` at
pandas/core/indexes/base.py)\r\n\r\nhttps://github.com/pandas-dev/pandas/blob/25
e64629ec317fba2fc1c2834b20362fa6c1fd89/pandas/core/indexes/base.py#L3286-L3290\r
\r\n\r\n* I created this pull request since @MayurKishorKumar doesn't seem to
work on this issue anymore, but please let me know if there is going to be
further actions on the previous PR and I am supposed to close this one.',
'closed_by': None, 'reactions': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61541/reactions', 'total_count': 0, '+1': 0, '-1': 0, 'laugh':

```

```

0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 0},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61541/timeline', 'performed_via_github_app': None,
'state_reason': None}, {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61540', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61540/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61540/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61540/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61540', 'id': 3113693938,
'node_id': 'PR_kwD0AA0YD86Yzzcs', 'number': 61540, 'title': 'ENH pandas-
dev#60693: shift operations for Series and DataFrames', 'user': {'login':
'David-msggc', 'id': 95498196, 'node_id': 'U_kgDOBbEv1A', 'avatar_url':
'https://avatars.githubusercontent.com/u/95498196?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/David-msggc', 'html_url':
'https://github.com/David-msggc', 'followers_url':
'https://api.github.com/users/David-msggc/followers', 'following_url':
'https://api.github.com/users/David-msggc/following{/other_user}', 'gists_url':
'https://api.github.com/users/David-msggc/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/David-msggc/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/David-msggc/subscriptions',
'organizations_url': 'https://api.github.com/users/David-msggc/orgs',
'repos_url': 'https://api.github.com/users/David-msggc/repos', 'events_url':
'https://api.github.com/users/David-msggc/events{/privacy}',
'received_events_url': 'https://api.github.com/users/David-
msggc/received_events', 'type': 'User', 'user_view_type': 'public',
'site_admin': False}, 'labels': [], 'state': 'open', 'locked': False,
'assignee': None, 'assignees': [], 'milestone': None, 'comments': 0,
'created_at': '2025-06-03T12:06:30Z', 'updated_at': '2025-06-03T12:06:30Z',
'closed_at': None, 'author_association': 'CONTRIBUTOR', 'type': None,
'active_lock_reason': None, 'draft': False, 'pull_request': {'url':
'https://api.github.com/repos/pandas-dev/pandas/pulls/61540', 'html_url':
'https://github.com/pandas-dev/pandas/pull/61540', 'diff_url':
'https://github.com/pandas-dev/pandas/pull/61540.diff', 'patch_url':
'https://github.com/pandas-dev/pandas/pull/61540.patch', 'merged_at': None},
'body': "This commit introduces the rshift and lshift method for both Series and
DataFrames in pandas. It also adds the corresponding in-place methods.\r\nThese
methdos don't work between a Series and a Dataframe, or if the two Series or
DataFrames differ in size.\r\n\r\n- [x] closes #60693\r\n- [x] [Tests added and
passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#writing-tests) if fixing a bug
or adding a new feature\r\n- [x] All [code checks
passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#pre-commit).\r\n- [x] Added
[type annotations](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#type-hints) to new
arguments/methods/functions.\r\n- [x] Added an entry in the latest

```

```

`doc/source/whatsnew/vX.X.X.rst` file if fixing a bug or adding a new
feature.\r\n", 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61540/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61540/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61539', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61539/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61539/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61539/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61539', 'id':
3113587665, 'node_id': 'I_kwDOAAOYD865lZPR', 'number': 61539, 'title': 'ENH:
Columns formatted as "Text" in Excel are read as numbers', 'user': {'login':
'pranay-sa', 'id': 116664160, 'node_id': 'U_kgDOBvQnYA', 'avatar_url':
'https://avatars.githubusercontent.com/u/116664160?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/pranay-sa', 'html_url':
'https://github.com/pranay-sa', 'followers_url':
'https://api.github.com/users/pranay-sa/followers', 'following_url':
'https://api.github.com/users/pranay-sa/following{/other_user}', 'gists_url':
'https://api.github.com/users/pranay-sa/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/pranay-sa/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/pranay-sa/subscriptions',
'organizations_url': 'https://api.github.com/users/pranay-sa/orgs', 'repos_url':
'https://api.github.com/users/pranay-sa/repos', 'events_url':
'https://api.github.com/users/pranay-sa/events{/privacy}',
'received_events_url': 'https://api.github.com/users/pranay-sa/received_events',
'type': 'User', 'user_view_type': 'public', 'site_admin': False}, 'labels':
[{'id': 76812, 'node_id': 'MDU6TGFiZWw3NjgxMg==', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Enhancement', 'name':
'Enhancement', 'color': '4E9A06', 'default': False, 'description': None}, {'id':
1954720290, 'node_id': 'MDU6TGFiZWwzOTUONzlwMjkw', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Needs%20Triage', 'name':
'Needs Triage', 'color': '0052cc', 'default': False, 'description': 'Issue that
has not been reviewed by a pandas team member'}], 'state': 'open', 'locked':
False, 'assignee': None, 'assignees': [], 'milestone': None, 'comments': 0,
'created_at': '2025-06-03T11:31:43Z', 'updated_at': '2025-06-03T11:38:17Z',
'closed_at': None, 'author_association': 'NONE', 'type': None,
'active_lock_reason': None, 'sub_issues_summary': {'total': 0, 'completed': 0,
'percent_completed': 0}, 'body': '### Feature Type\n\n [x] Adding new
functionality to pandas\n\n [ ] Changing existing functionality in pandas\n\n
[ ] Removing existing functionality in pandas\n\n\n### Problem
Description\n\nWhen reading Excel files, pandas ignores Excel\'s "Text" cell
formatting and converts text-formatted numbers (e.g., IDs, codes) to numeric
types (int/float). \nThis requires manual conversion back to strings, which can
be inefficient , for a huge dataset and prone to

```

```

errors.\n\n! [Image] (https://github.com/user-
attachments/assets/cbf6cedb-29d4-4b75-ae29-76fbc04cfcb) \n\n### Feature
Description\n\nAdd an option in pd.read_excel() to respect Excel's cell
formatting\n(e.g., dtype_from_format=True), or set it to true by default ,
preserving text-formatted columns as strings.\n\n### Alternative
Solutions\n\nOpenPyXL/Xlrd Engine + Format Detection\nRead cell formats directly
(requires manual parsing):\n\nfrom openpyxl import load_workbook\n\nwb =
load_workbook("data.xlsx", data_only=False)\nsheet = wb.active\ntext_columns =
[col for col in sheet.columns if sheet.cell(row=1,
column=col[0].column).number_format == "@"]\n\n### Additional Context\n\nNo
response_', 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61539/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61539/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61537', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61537/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61537/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61537/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61537', 'id':
3111287144, 'node_id': 'I_kwDOAAOYD865cnlo', 'number': 61537, 'title': 'BUG: use
of iloc with heterogeneous DataFrame sometimes performs undocumented
conversions', 'user': {'login': 'illbebach', 'id': 1385858, 'node_id':
'MDQ6VXNlcjEzODU4NTg=', 'avatar_url':
'https://avatars.githubusercontent.com/u/1385858?v=4', 'gravatar_id': '', 'url':
'https://api.github.com/users/illbebach', 'html_url':
'https://github.com/illbebach', 'followers_url':
'https://api.github.com/users/illbebach/followers', 'following_url':
'https://api.github.com/users/illbebach/following{/other_user}', 'gists_url':
'https://api.github.com/users/illbebach/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/illbebach/starred{/owner}/{repo}',
'subscriptions_url': 'https://api.github.com/users/illbebach/subscriptions',
'organizations_url': 'https://api.github.com/users/illbebach/orgs', 'repos_url':
'https://api.github.com/users/illbebach/repos', 'events_url':
'https://api.github.com/users/illbebach/events{/privacy}',
'received_events_url': 'https://api.github.com/users/illbebach/received_events',
'type': 'User', 'user_view_type': 'public', 'site_admin': False}, 'labels':
[{'id': 76811, 'node_id': 'MDU6TGFiZWw3NjgxMQ==', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Bug', 'name': 'Bug',
'color': 'e10c02', 'default': False, 'description': None}, {'id': 1954720290,
'node_id': 'MDU6TGFiZWwzOTUONzIwMjkw', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Needs%20Triage', 'name':
'Needs Triage', 'color': '0052cc', 'default': False, 'description': 'Issue that
has not been reviewed by a pandas team member'}], 'state': 'open', 'locked':
False, 'assignee': None, 'assignees': [], 'milestone': None, 'comments': 0,

```

```

'created_at': '2025-06-02T18:59:10Z', 'updated_at': '2025-06-02T18:59:10Z',
'closed_at': None, 'author_association': 'NONE', 'type': None,
'active_lock_reason': None, 'sub_issues_summary': {'total': 0, 'completed': 0,
'percent_completed': 0}, 'body': "### Pandas version checks\n\n- [x] I have
checked that this issue has not already been reported.\n\n- [x] I have confirmed
this bug exists on the [latest
version](https://pandas.pydata.org/docs/whatsnew/index.html) of pandas.\n\n- [ ]
I have confirmed this bug exists on the [main branch](https://pandas.pydata.org/
docs/dev/getting_started/install.html#installing-the-development-version-of-
pandas) of pandas.\n\n### Reproducible Example\n\n```python\nimport pandas as
pd\nnd = {\n    'Name': ['Bob', 'John', 'Alice'],\n    'Age': [25, 41, 30],\n
'Result': [1.2, 0.5, 0.3],\n    'Ok': [True, False, True],\n}\nndf =
pd.DataFrame(data=d)\n\nprint()\nprint('----- test 1
-----')\nprint(df)\nprint()\nprint('first row')\nprint(df.iloc[0])\n\nnd = {\n
'Age': [25, 41, 30],\n    'Result': [1.2, 0.5, 0.3],\n    'Ok': [True, False,
True],\n}\nndf = pd.DataFrame(data=d)\n\nprint()\nprint('----- test 2
-----')\nprint(df)\nprint()\nprint('first row')\nprint(df.iloc[0])\n\nnd = {\n
'Age': [25, 41, 30],\n    'Result': [1.2, 0.5, 0.3],\n}\nndf =
pd.DataFrame(data=d)\n\nprint()\nprint('----- test 3
-----')\nprint(df)\nprint()\nprint('first row')\nprint(df.iloc[0])\n\n```\n\n###
Issue Description\n\nPandas sometimes performs type conversions on returned data
from the `.iloc` function.\n\nIn the first two cases, a Series of object is
returned. In the last case pandas decided to promote all values to float, and
returns a Series of float, which is not what I want.\n\nProgram
output\n\n```\n\n----- test 1 -----\n    Name Age Result    Ok\n0    Bob    25
1.2    True\n1    John    41    0.5    False\n2    Alice    30    0.3    True\n\nfirst
row\nName    Bob\nAge    25\nResult    1.2\nOk    True\nName: 0,
dtype: object\n\n----- test 2 -----\n    Age Result    Ok\n0    25    1.2
True\n1    41    0.5    False\n2    30    0.3    True\n\nfirst row\nAge
25\nResult    1.2\nOk    True\nName: 0, dtype: object\n\n----- test 3
-----\n    Age Result\n0    25    1.2\n1    41    0.5\n2    30    0.3\n\nfirst
row\nAge    25.0\nResult    1.2\nName: 0, dtype: float64\n\n```\n\n1. To me,
it is undesirable, in any circumstance, for pandas to apply type conversions on
returned data. I realize this conversion is probably there for historical or
performance reasons, and may not be changed.\n3. At a minimum, the documentation
should mention exactly what circumstances a type conversion will occur. The
[.iloc documentation](https://pandas.pydata.org/docs/reference/api/pandas.DataF
rame.iloc.html#pandas.DataFrame.iloc) makes no mention of what will occur when a
DataFrame containing heterogeneous data is indexed via `.iloc`\n\nFor
completeness, there is a similar bug logged long ago. #5256\n\n### Expected
Behavior\n\n1. My preference is to never perform type conversions. I realize
changing this behavior could break some existing code that depends on such a
conversion.\n2. My second recommendation is to update the documentation to
describe exactly when and what data conversions will be performed by pandas. At
a minimum there should be a 'warning' or 'note' about the type
conversions.\n\n### Installed Versions\n\n<details>\n\nINSTALLED
VERSIONS\n\n-----\n\ncommit
:
0691c5cf90477d3503834d983f69350f250a6ff7\n\npython
:

```

```

3.12.10\npython-bits          : 64\nOS                      : Windows\nOS-release
: 10\nVersion                  : 10.0.19045\nmachine                    :
AMD64\nprocessor                : Intel64 Family 6 Model 154 Stepping 3,
GenuineIntel\nbyteorder       : little\nLC_ALL                  : None\nLANG
: None\nLOCALE                : English_United States.1252\n\nnpandas
: 2.2.3\nnumpy                : 2.2.6\npytz                      : 2025.2\ndateutil
: 2.9.0.post0\npip            : 25.1.1\nCython                    :
None\nsphinx                 : None\nIPython                    : None\nadbc-driver-
postgresql: None\nadbc-driver-sqlite : None\nbs4
None\nblosc                   : None\nbottleneck                 : None\ndataframe-api-
compat : None\nfastparquet      : None\nfsspec                    :
None\nhtml5lib                 : None\nhypothesis                 : None\nngcsfs
: None\njinja2                 : None\nlxml.etree                 : None\nmatplotlib
: 3.10.3\nnumba                : None\nnumexpr                    : None\nodfpy
: None\nopenpyxl               : 3.1.5\npandas_gbq                 : None\npsycopg2
: None\npymysql                : None\npyarrow                    : None\npyreadstat
: None\npytest                 : None\npython-calamine           : None\npyxlsb
: None\ns3fs                    : None\nscipy                      : 1.15.3\nsqlalchemy
: None\ntables                  : None\nTABULATE                 : None\nxarray
: None\nxlrd                    : None\nxlswriter                 : None\nzstandard
: None\ntzdata                  : 2025.2\nqtptpy                    : None\npyqt5
: None\n</details>\n", 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61537/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61537/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61535', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61535/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61535/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61535/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61535', 'id':
3108703108, 'node_id': 'I_kwDOAAOYD865SwuE', 'number': 61535, 'title': 'ENH:
read_csv tz option', 'user': {'login': 'hasandiwan', 'id': 44325453, 'node_id':
'MDQ6VXNlcjQOMzI1NDUz', 'avatar_url':
'https://avatars.githubusercontent.com/u/44325453?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/hasandiwan', 'html_url':
'https://github.com/hasandiwan', 'followers_url':
'https://api.github.com/users/hasandiwan/followers', 'following_url':
'https://api.github.com/users/hasandiwan/following{/other_user}', 'gists_url':
'https://api.github.com/users/hasandiwan/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/hasandiwan/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/hasandiwan/subscriptions',
'organizations_url': 'https://api.github.com/users/hasandiwan/orgs',
'repos_url': 'https://api.github.com/users/hasandiwan/repos', 'events_url':
'https://api.github.com/users/hasandiwan/events{/privacy}',

```



```

'received_events_url':
'https://api.github.com/users/hasandiwan/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{ 'id': 76812,
'node_id': 'MDU6TGFiZWw3NjgxMg==', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Enhancement', 'name': 'Enhancement', 'color': '4E9A06',
'default': False, 'description': None}, { 'id': 1954720290, 'node_id':
'MDU6TGFiZWw3OTU0NzIwMjkw', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Needs%20Triage', 'name': 'Needs Triage', 'color': '0052cc',
'default': False, 'description': 'Issue that has not been reviewed by a pandas
team member'}], 'state': 'open', 'locked': False, 'assignee': None, 'assignees':
[], 'milestone': None, 'comments': 0, 'created_at': '2025-06-02T05:57:05Z',
'updated_at': '2025-06-02T05:59:51Z', 'closed_at': None, 'author_association':
'NONE', 'type': None, 'active_lock_reason': None, 'sub_issues_summary':
{'total': 0, 'completed': 0, 'percent_completed': 0}, 'body': "### Feature
Type\n\n [x] Adding new functionality to pandas\n\n [ ] Changing existing
functionality in pandas\n\n [ ] Removing existing functionality in
pandas\n\n\n### Problem Description\n\nI use pd.read_csv to grab a series of
timestamp'd links interactively from a remote website. \n\n### Feature
Description\n\nI would like it to convert the columns specified by parse_dates
to the timezone specified by wherever the /etc/localtime link points to by
default in a non-deprecated manner:\n\n>
[frame.loc[:,c].dt.tz_convert('/'.join([os.getenv('TZ',
os.path.realpath('/etc/localtime').split('/')[ -2:]))[0])) for c in
frame.select_dtypes('datetime64[ns, UTC]')]` \n\nI'd like to propose this
functionality as the tz parameter to read_csv. I suspect the implementation is
not python, and can't find it in my git checkout of pandas. \n\n### Alternative
Solutions\n\nCovered above\n\n### Additional Context", 'closed_by': None,
'reactions': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61535/reactions', 'total_count': 0, '+1': 0, '-1': 0, 'laugh':
0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 0},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61535/timeline', 'performed_via_github_app': None,
'state_reason': None}, { 'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61533', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61533/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61533/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61533/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61533', 'id': 3107929504,
'node_id': 'PR_kwDOAAOYD86YgpaA', 'number': 61533, 'title': 'BUG: Fix
Series.str.zfill for ArrowDtype string arrays #61485', 'user': {'login':
'iabhi4', 'id': 61010675, 'node_id': 'MDQ6VXNlcjYxMDEwNjc1', 'avatar_url':
'https://avatars.githubusercontent.com/u/61010675?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/iabhi4', 'html_url':
'https://github.com/iabhi4', 'followers_url':
'https://api.github.com/users/iabhi4/followers', 'following_url':
'https://api.github.com/users/iabhi4/following{/other_user}', 'gists_url':

```

```

'https://api.github.com/users/iabhi4/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/iabhi4/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/iabhi4/subscriptions',
'organizations_url': 'https://api.github.com/users/iabhi4/orgs', 'repos_url':
'https://api.github.com/users/iabhi4/repos', 'events_url':
'https://api.github.com/users/iabhi4/events{/privacy}', 'received_events_url':
'https://api.github.com/users/iabhi4/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False, 'labels': [], 'state': 'open',
'locked': False, 'assignee': None, 'assignees': [], 'milestone': None,
'comments': 0, 'created_at': '2025-06-01T22:09:25Z', 'updated_at':
'2025-06-05T17:18:11Z', 'closed_at': None, 'author_association': 'CONTRIBUTOR',
'type': None, 'active_lock_reason': None, 'draft': False, 'pull_request':
{'url': 'https://api.github.com/repos/pandas-dev/pandas/pulls/61533',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61533', 'diff_url':
'https://github.com/pandas-dev/pandas/pull/61533.diff', 'patch_url':
'https://github.com/pandas-dev/pandas/pull/61533.patch', 'merged_at': None},
'body': "Implemented `_str_zfill` for `ArrowExtensionArray` to support
`Series.str.zfill` on Arrow-backed string arrays (`ArrowDtype(pa.string())`).
This fixes an AttributeError due to the method relying on `_str_map`, which
wasn't implemented. Used `_apply_elementwise` to match the approach of other
string methods. Added tests under `test_string_array.py` and confirmed they
pass. Also confirmed no other relevant test files are broken and the change
aligns with how other string accessors are handled.\r\n\r\n- [x] closes
#61485\r\n- [x] Tests added and passed\r\n- [x] All code checks passed with pre-
commit\r\n- [x] Added a changelog entry in v3.0.0.rst", 'closed_by': None,
'reactions': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61533/reactions', 'total_count': 0, '+1': 0, '-1': 0, 'laugh':
0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 0},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61533/timeline', 'performed_via_github_app': None,
'state_reason': None}, {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61532', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61532/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61532/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61532/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61532', 'id':
3107872632, 'node_id': 'I_kwDOAAOYD865Pl94', 'number': 61532, 'title': 'ENH:
speed up `DataFrame.plot` using `LineCollection`', 'user': {'login':
'Abdelgha-4', 'id': 72174073, 'node_id': 'MDQ6VXNlcjcyMTc0MDcz', 'avatar_url':
'https://avatars.githubusercontent.com/u/72174073?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/Abdelgha-4', 'html_url':
'https://github.com/Abdelgha-4', 'followers_url':
'https://api.github.com/users/Abdelgha-4/followers', 'following_url':
'https://api.github.com/users/Abdelgha-4/following{/other_user}', 'gists_url':
'https://api.github.com/users/Abdelgha-4/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/Abdelgha-4/starred{/owner}/{/repo}',

```

```

'subscriptions_url': 'https://api.github.com/users/Abdelgha-4/subscriptions',
'organizations_url': 'https://api.github.com/users/Abdelgha-4/orgs',
'repos_url': 'https://api.github.com/users/Abdelgha-4/repos', 'events_url':
'https://api.github.com/users/Abdelgha-4/events{/privacy}',
'received_events_url':
'https://api.github.com/users/Abdelgha-4/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{ 'id': 2413328,
'node_id': 'MDU6TGFiZWwvNDEzMzI4', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Visualization', 'name': 'Visualization', 'color': '8AE234',
'default': False, 'description': 'plotting'}, { 'id': 8935311, 'node_id':
'MDU6TGFiZWwvNDQ0MTMzEx', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Performance', 'name': 'Performance', 'color': 'a10c02',
'default': False, 'description': 'Memory or execution speed performance'}],
'state': 'open', 'locked': False, 'assignee': { 'login': 'shadnikn', 'id':
147014887, 'node_id': 'U_kgDOCMNE5w', 'avatar_url':
'https://avatars.githubusercontent.com/u/147014887?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/shadnikn', 'html_url':
'https://github.com/shadnikn', 'followers_url':
'https://api.github.com/users/shadnikn/followers', 'following_url':
'https://api.github.com/users/shadnikn/following{/other_user}', 'gists_url':
'https://api.github.com/users/shadnikn/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/shadnikn/starred{/owner}/{repo}',
'subscriptions_url': 'https://api.github.com/users/shadnikn/subscriptions',
'organizations_url': 'https://api.github.com/users/shadnikn/orgs', 'repos_url':
'https://api.github.com/users/shadnikn/repos', 'events_url':
'https://api.github.com/users/shadnikn/events{/privacy}', 'received_events_url':
'https://api.github.com/users/shadnikn/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'assignees': [{ 'login':
'shadnikn', 'id': 147014887, 'node_id': 'U_kgDOCMNE5w', 'avatar_url':
'https://avatars.githubusercontent.com/u/147014887?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/shadnikn', 'html_url':
'https://github.com/shadnikn', 'followers_url':
'https://api.github.com/users/shadnikn/followers', 'following_url':
'https://api.github.com/users/shadnikn/following{/other_user}', 'gists_url':
'https://api.github.com/users/shadnikn/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/shadnikn/starred{/owner}/{repo}',
'subscriptions_url': 'https://api.github.com/users/shadnikn/subscriptions',
'organizations_url': 'https://api.github.com/users/shadnikn/orgs', 'repos_url':
'https://api.github.com/users/shadnikn/repos', 'events_url':
'https://api.github.com/users/shadnikn/events{/privacy}', 'received_events_url':
'https://api.github.com/users/shadnikn/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}], 'milestone': None,
'comments': 5, 'created_at': '2025-06-01T21:25:38Z', 'updated_at':
'2025-06-04T00:23:14Z', 'closed_at': None, 'author_association': 'NONE', 'type':
None, 'active_lock_reason': None, 'sub_issues_summary': { 'total': 0,
'completed': 0, 'percent_completed': 0}, 'body': '**Description:**\n\nWhen
plotting line charts with many columns or rows, DataFrame.plot() currently adds
one Line2D object per column. This incurs significant overhead in large

```

datasets.\n\nReplacing this with a single LineCollection (from matplotlib.collections) can yield substantial speedups. In my benchmarks, plotting via LineCollection was ~2.5x faster on large DataFrames with many columns.\n\n\n\*\*Minimal example:\*\*\n``python\n# Imports and data generation\nimport itertools\nimport matplotlib.pyplot as plt\nimport numpy as np\nimport pandas as pd\nfrom matplotlib.collections import LineCollection\n\nnum\_rows = 500\nnum\_cols = 2000\nntest\_df = pd.DataFrame(np.random.randn(num\_rows, num\_cols).cumsum(axis=0))\n\n# Simply using DataFrame.plot, (5.6 secs)\ntest\_df.plot(legend=False, figsize=(12, 8))\nplt.show()\n\n# Optimized version using LineCollection (2.2 secs)\nidx = np.arange(len(test\_df.index))\nlines = [np.column\_stack([x, test\_df[col].values]) for col in test\_df.columns]\ndefault\_colors = plt.rcParams["axes.prop\_cycle"].by\_key()["color"]\ncolor\_cycle = list(itertools.islice(itertools.cycle(default\_colors), len(lines)))\n\nline\_collection = LineCollection(lines, colors=color\_cycle)\nfig, ax = plt.subplots(figsize=(12, 8))\nax.add\_collection(line\_collection)\nax.margins(0.05)\nplt.show()\n``\n\n\n\*\*Note:\*\* the ~2.5x speed improvement is specific to dataframes with integer index. For dataframes with `DatetimeIndex` the actual speed improvement is ~27x when combined with the workaround here: #61398\n\nThank you for considering this suggestion!',\n'closed\_by': None, 'reactions': {'url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61532/reactions', 'total\_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline\_url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61532/timeline', 'performed\_via\_github\_app': None, 'state\_reason': None}, {'url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61531', 'repository\_url': 'https://api.github.com/repos/pandas-dev/pandas', 'labels\_url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61531/labels{/name}', 'comments\_url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61531/comments', 'events\_url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61531/events', 'html\_url': 'https://github.com/pandas-dev/pandas/issues/61531', 'id': 3107305231, 'node\_id': 'I\_kwDOAAOYD865NbcP', 'number': 61531, 'title': 'CI: Micromamba taking too long to resolve the environments in the CI', 'user': {'login': 'datapythonista', 'id': 10058240, 'node\_id': 'MDQ6VXNlcjEwMDU4MjQw', 'avatar\_url': 'https://avatars.githubusercontent.com/u/10058240?v=4', 'gravatar\_id': '', 'url': 'https://api.github.com/users/datapythonista', 'html\_url': 'https://github.com/datapythonista', 'followers\_url': 'https://api.github.com/users/datapythonista/followers', 'following\_url': 'https://api.github.com/users/datapythonista/following{/other\_user}', 'gists\_url': 'https://api.github.com/users/datapythonista/gists{/gist\_id}', 'starred\_url': 'https://api.github.com/users/datapythonista/starred{/owner}/{/repo}', 'subscriptions\_url': 'https://api.github.com/users/datapythonista/subscriptions', 'organizations\_url': 'https://api.github.com/users/datapythonista/orgs', 'repos\_url': 'https://api.github.com/users/datapythonista/repos', 'events\_url':

```

'https://api.github.com/users/datapythonista/events{/privacy}',
'received_events_url':
'https://api.github.com/users/datapythonista/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{ 'id': 48070600,
'node_id': 'MDU6TGFiZWw0ODAzMDYwMA==', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/CI', 'name': 'CI',
'color': 'a2bca7', 'default': False, 'description': 'Continuous Integration'},
{ 'id': 527603109, 'node_id': 'MDU6TGFiZWw1Mjc2MDMxMDk=', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/Dependencies', 'name':
'Dependencies', 'color': 'd93f0b', 'default': False, 'description': 'Required
and optional dependencies'}, { 'id': 717120670, 'node_id':
'MDU6TGFiZWw3MTcxMjA2NzA=', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/good%20first%20issue', 'name': 'good first issue', 'color':
'Oe8a16', 'default': True, 'description': None}], 'state': 'open', 'locked':
False, 'assignee': { 'login': 'microslaw', 'id': 91637238, 'node_id':
'U_kgDOBXZF9g', 'avatar_url':
'https://avatars.githubusercontent.com/u/91637238?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/microslaw', 'html_url':
'https://github.com/microslaw', 'followers_url':
'https://api.github.com/users/microslaw/followers', 'following_url':
'https://api.github.com/users/microslaw/following{/other_user}', 'gists_url':
'https://api.github.com/users/microslaw/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/microslaw/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/microslaw/subscriptions',
'organizations_url': 'https://api.github.com/users/microslaw/orgs', 'repos_url':
'https://api.github.com/users/microslaw/repos', 'events_url':
'https://api.github.com/users/microslaw/events{/privacy}',
'received_events_url': 'https://api.github.com/users/microslaw/received_events',
'type': 'User', 'user_view_type': 'public', 'site_admin': False}, 'assignees':
[{ 'login': 'microslaw', 'id': 91637238, 'node_id': 'U_kgDOBXZF9g', 'avatar_url':
'https://avatars.githubusercontent.com/u/91637238?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/microslaw', 'html_url':
'https://github.com/microslaw', 'followers_url':
'https://api.github.com/users/microslaw/followers', 'following_url':
'https://api.github.com/users/microslaw/following{/other_user}', 'gists_url':
'https://api.github.com/users/microslaw/gists{/gist_id}', 'starred_url':
'https://api.github.com/users/microslaw/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/microslaw/subscriptions',
'organizations_url': 'https://api.github.com/users/microslaw/orgs', 'repos_url':
'https://api.github.com/users/microslaw/repos', 'events_url':
'https://api.github.com/users/microslaw/events{/privacy}',
'received_events_url': 'https://api.github.com/users/microslaw/received_events',
'type': 'User', 'user_view_type': 'public', 'site_admin': False}], 'milestone':
None, 'comments': 6, 'created_at': '2025-06-01T14:08:04Z', 'updated_at':
'2025-06-04T13:19:10Z', 'closed_at': None, 'author_association': 'MEMBER',
'type': None, 'active_lock_reason': None, 'sub_issues_summary': { 'total': 0,
'completed': 0, 'percent_completed': 0}, 'body': "Our CI jobs are frequently
failing now as they timeout after 90 minutes of execution. Of those 90 minutes,

```

25 are spent on micromamba resolving the environment.\n\nIn the past we have fixed this by limiting the number of packages to be considered. For example, if the environment just says `numpy`, maybe there are 200 versions that will be considered. While if we say `numpy >= 2` the number can be limited to few.\n\nI'm not sure which packages have lots of options, and we don't want to filter out the versions that make sense to install. But we should have a look and see if by adding few constraints we can get a reasonable time to solve the environment."

```
{
  'closed_by': None,
  'reactions': {'url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61531/reactions',
    'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
    'heart': 0, 'rocket': 0, 'eyes': 0},
  'timeline_url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61531/timeline',
  'performed_via_github_app': None,
  'state_reason': None},
  {'url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61530',
    'repository_url': 'https://api.github.com/repos/pandas-dev/pandas',
    'labels_url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61530/labels{/name}',
    'comments_url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61530/comments',
    'events_url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61530/events',
    'html_url': 'https://github.com/pandas-dev/pandas/pull/61530',
    'id': 3107129953,
    'node_id': 'PR_kwD0AA0YD86YePF9',
    'number': 61530,
    'title': 'API: Replace na_action parameter in Series/DataFrame/Index.map by the standard skipna',
    'user': {'login': 'datapythonista', 'id': 10058240, 'node_id': 'MDQ6VXNlcjEwMDU4MjQw',
      'avatar_url': 'https://avatars.githubusercontent.com/u/10058240?v=4',
      'gravatar_id': '',
      'url': 'https://api.github.com/users/datapythonista',
      'html_url': 'https://github.com/datapythonista',
      'followers_url': 'https://api.github.com/users/datapythonista/followers',
      'following_url': 'https://api.github.com/users/datapythonista/following{/other_user}',
      'gists_url': 'https://api.github.com/users/datapythonista/gists{/gist_id}',
      'starred_url': 'https://api.github.com/users/datapythonista/starred{/owner}/{/repo}',
      'subscriptions_url': 'https://api.github.com/users/datapythonista/subscriptions',
      'organizations_url': 'https://api.github.com/users/datapythonista/orgs',
      'repos_url': 'https://api.github.com/users/datapythonista/repos',
      'events_url': 'https://api.github.com/users/datapythonista/events{/privacy}',
      'received_events_url': 'https://api.github.com/users/datapythonista/received_events',
      'type': 'User',
      'user_view_type': 'public',
      'site_admin': False},
    'labels': [{'id': 35818298,
      'node_id': 'MDU6TGFiZWwzNTgxODI5OA==',
      'url': 'https://api.github.com/repos/pandas-dev/pandas/labels/API%20Design',
      'name': 'API Design',
      'color': 'AD7FA8',
      'default': False,
      'description': None},
      {'id': 697792067,
      'node_id': 'MDU6TGFiZWwzNTgxODI5OA==',
      'url': 'https://api.github.com/repos/pandas-dev/pandas/labels/Apply',
      'name': 'Apply',
      'color': 'fbca04',
      'default': False,
      'description': 'Apply, Aggregate, Transform, Map'}],
    'state': 'open',
    'locked': False,
    'assignee': None,
    'assignees': [],
    'milestone': None,
    'comments': 0,
    'created_at':
```

```

'2025-06-01T11:47:23Z', 'updated_at': '2025-06-02T21:04:10Z', 'closed_at': None,
'author_association': 'MEMBER', 'type': None, 'active_lock_reason': None,
'draft': False, 'pull_request': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/pulls/61530', 'html_url': 'https://github.com/pandas-
dev/pandas/pull/61530', 'diff_url': 'https://github.com/pandas-
dev/pandas/pull/61530.diff', 'patch_url': 'https://github.com/pandas-
dev/pandas/pull/61530.patch', 'merged_at': None}, 'body': '- [X] xref
#61128\r\n- [X] [Tests added and passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#writing-tests) if fixing a bug
or adding a new feature\r\n- [X] All [code checks
passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#pre-commit).\r\n- [X] Added
[type annotations](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#type-hints) to new
arguments/methods/functions.\r\n- [X] Added an entry in the latest
`doc/source/whatsnew/vX.X.X.rst` file if fixing a bug or adding a new
feature.\r\n\r\nWe\'ve been consistently using a boolean `skipna` parameter to
let users decide whether missing values should be ignored or not in different
methods. The `.map` method has an awkward `na_action=None | "ignore"` for the
same purpose. I add the standard `skipna` parameter to the methods, and start
the deprecation of the `na_action` one.', 'closed_by': None, 'reactions':
{'url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61530/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61530/timeline',
'performed_via_github_app': None, 'state_reason': None}, {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61529', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61529/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61529/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61529/events',
'html_url': 'https://github.com/pandas-dev/pandas/issues/61529', 'id':
3106434267, 'node_id': 'I_kwDOAAOYD865KGzb', 'number': 61529, 'title': 'raise
TypeError when function returns object-like results', 'user': {'login':
'michael2015tse', 'id': 44341134, 'node_id': 'MDQ6VXNlcjQOMzQxMTM0',
'avatar_url': 'https://avatars.githubusercontent.com/u/44341134?v=4',
'gravatar_id': '', 'url': 'https://api.github.com/users/michael2015tse',
'html_url': 'https://github.com/michael2015tse', 'followers_url':
'https://api.github.com/users/michael2015tse/followers', 'following_url':
'https://api.github.com/users/michael2015tse/following{/other_user}',
'gists_url': 'https://api.github.com/users/michael2015tse/gists{/gist_id}',
'starred_url':
'https://api.github.com/users/michael2015tse/starred{/owner}/{/repo}',
'subscriptions_url':
'https://api.github.com/users/michael2015tse/subscriptions',
'organizations_url': 'https://api.github.com/users/michael2015tse/orgs',
'repos_url': 'https://api.github.com/users/michael2015tse/repos', 'events_url':

```

```

'https://api.github.com/users/michael2015tse/events{/privacy}',
'received_events_url':
'https://api.github.com/users/michael2015tse/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [], 'state': 'open',
'locked': False, 'assignee': None, 'assignees': [], 'milestone': None,
'comments': 0, 'created_at': '2025-06-01T02:41:40Z', 'updated_at':
'2025-06-01T02:43:23Z', 'closed_at': None, 'author_association': 'NONE', 'type':
None, 'active_lock_reason': None, 'sub_issues_summary': {'total': 0,
'completed': 0, 'percent_completed': 0}, 'body': 'This roll_apply raises
`TypeError: must be real number, not str` when my customized function returns
str / list or other object-like results. Because here it is: `ndarray[float64_t]
output`\n\nhttps://github.com/pandas-dev/pandas/blob/0691c5cf90477d3503834d983f6
9350f250a6ff7/pandas/_libs/window/aggregations.pyx#L1396', 'closed_by': None,
'reactions': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61529/reactions', 'total_count': 0, '+1': 0, '-1': 0, 'laugh':
0, 'hooray': 0, 'confused': 0, 'heart': 0, 'rocket': 0, 'eyes': 0},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61529/timeline', 'performed_via_github_app': None,
'state_reason': None}, {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61527', 'repository_url':
'https://api.github.com/repos/pandas-dev/pandas', 'labels_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61527/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61527/comments', 'events_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61527/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61527', 'id': 3105461098,
'node_id': 'PR_kwDOAAOYD86YZGj2', 'number': 61527, 'title': 'ENH: Implement
DataFrame.select', 'user': {'login': 'datapythonista', 'id': 10058240,
'node_id': 'MDQ6VXNlcjEwMDU4MjQw', 'avatar_url':
'https://avatars.githubusercontent.com/u/10058240?v=4', 'gravatar_id': '',
'url': 'https://api.github.com/users/datapythonista', 'html_url':
'https://github.com/datapythonista', 'followers_url':
'https://api.github.com/users/datapythonista/followers', 'following_url':
'https://api.github.com/users/datapythonista/following{/other_user}',
'gists_url': 'https://api.github.com/users/datapythonista/gists{/gist_id}',
'starred_url':
'https://api.github.com/users/datapythonista/starred{/owner}/{/repo}',
'subscriptions_url':
'https://api.github.com/users/datapythonista/subscriptions',
'organizations_url': 'https://api.github.com/users/datapythonista/orgs',
'repos_url': 'https://api.github.com/users/datapythonista/repos', 'events_url':
'https://api.github.com/users/datapythonista/events{/privacy}',
'received_events_url':
'https://api.github.com/users/datapythonista/received_events', 'type': 'User',
'user_view_type': 'public', 'site_admin': False}, 'labels': [{'id': 76812,
'node_id': 'MDU6TGFiZWw3NjgXMg==', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Enhancement', 'name': 'Enhancement', 'color': '4E9A06',
'default': False, 'description': None}, {'id': 2822098, 'node_id':

```



```
'MDU6TGFiZWwyODIyMDk4', 'url': 'https://api.github.com/repos/pandas-
dev/pandas/labels/Indexing', 'name': 'Indexing', 'color': '0b02e1', 'default':
False, 'description': 'Related to indexing on series/frames, not to indexes
themselves'}, {'id': 35818298, 'node_id': 'MDU6TGFiZWwzNTgxODI5OA==', 'url':
'https://api.github.com/repos/pandas-dev/pandas/labels/API%20Design', 'name':
'API Design', 'color': 'AD7FA8', 'default': False, 'description': None}],
'state': 'open', 'locked': False, 'assignee': None, 'assignees': [],
'milestone': None, 'comments': 15, 'created_at': '2025-05-31T13:10:03Z',
'updated_at': '2025-06-04T22:32:07Z', 'closed_at': None, 'author_association':
'MEMBER', 'type': None, 'active_lock_reason': None, 'draft': False,
'pull_request': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/pulls/61527', 'html_url': 'https://github.com/pandas-
dev/pandas/pull/61527', 'diff_url': 'https://github.com/pandas-
dev/pandas/pull/61527.diff', 'patch_url': 'https://github.com/pandas-
dev/pandas/pull/61527.patch', 'merged_at': None}, 'body': '- [X] closes
#61522\r\n- [X] [Tests added and passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#writing-tests) if fixing a bug
or adding a new feature\r\n- [X] All [code checks
passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#pre-commit).\r\n- [X] Added
[type annotations](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#type-hints) to new
arguments/methods/functions.\r\n- [X] Added an entry in the latest
`doc/source/whatsnew/vX.X.X.rst` file if fixing a bug or adding a new
feature.\r\n\r\nBased on the feedback in #61522 and on the last devs call, I
implemented `DataFrame.select` in the most simple way. It does work with
`MultiIndex`, but it does not support equivalents to `filter(regex=)` or
`filter(like=)` directly. I added examples in the docs, so users can do that
easily in Python (I can add one for regex if people think it\'s worth
it).\r\n\r\nThe examples in the docs and the tests should make quite clear
what\'s the behavior, feedback welcome.\r\n\r\nFor context, this is added so we
can make `DataFrame.filter` focus on filtering rows, for
example:\r\n\r\n``python\r\nndf = df.select("name", "age")\r\nndf =
df.filter(df.age >= 18)\r\n``\r\n\r\nor\r\n\r\n``python\r\n(df.select("name",
"age")\r\n    .filter(lambda df: df.age >= 18))\r\n``\r\n\r\nCC: @pandas-
dev/pandas-core ', 'closed_by': None, 'reactions': {'url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61527/reactions',
'total_count': 0, '+1': 0, '-1': 0, 'laugh': 0, 'hooray': 0, 'confused': 0,
'heart': 0, 'rocket': 0, 'eyes': 0}, 'timeline_url':
'https://api.github.com/repos/pandas-dev/pandas/issues/61527/timeline',
'performed_via_github_app': None, 'state_reason': None}]
```

Getting a First record

```
[172]: js[0]
```

```
[172]: {'url': 'https://api.github.com/repos/pandas-dev/pandas/issues/61576',
'repository_url': 'https://api.github.com/repos/pandas-dev/pandas',
```

```

'labels_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61576/labels{/name}',
'comments_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61576/comments',
'events_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61576/events',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61576',
'id': 3122405296,
'node_id': 'PR_kwDOAAOYD86ZQ5dU',
'number': 61576,
'title': 'DataFrames Class update',
'user': {'login': 'SaraTammame',
'id': 116681400,
'node_id': 'U_kgDOBvRquA',
'avatar_url': 'https://avatars.githubusercontent.com/u/116681400?v=4',
'gravatar_id': '',
'url': 'https://api.github.com/users/SaraTammame',
'html_url': 'https://github.com/SaraTammame',
'followers_url': 'https://api.github.com/users/SaraTammame/followers',
'following_url':
'https://api.github.com/users/SaraTammame/following{/other_user}',
'gists_url': 'https://api.github.com/users/SaraTammame/gists{/gist_id}',
'starred_url':
'https://api.github.com/users/SaraTammame/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/SaraTammame/subscriptions',
'organizations_url': 'https://api.github.com/users/SaraTammame/orgs',
'repos_url': 'https://api.github.com/users/SaraTammame/repos',
'events_url': 'https://api.github.com/users/SaraTammame/events{/privacy}',
'received_events_url':
'https://api.github.com/users/SaraTammame/received_events',
'type': 'User',
'user_view_type': 'public',
'site_admin': False},
'labels': [],
'state': 'open',
'locked': False,
'assignee': None,
'assignees': [],
'milestone': None,
'comments': 0,
'created_at': '2025-06-05T19:32:17Z',
'updated_at': '2025-06-05T19:32:17Z',
'closed_at': None,
'author_association': 'NONE',
'type': None,
'active_lock_reason': None,
'draft': False,

```

```

'pull_request': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/pulls/61576',
'html_url': 'https://github.com/pandas-dev/pandas/pull/61576',
'diff_url': 'https://github.com/pandas-dev/pandas/pull/61576.diff',
'patch_url': 'https://github.com/pandas-dev/pandas/pull/61576.patch',
'merged_at': None},
'body': '- [ ] closes #xxxx (Replace xxxx with the GitHub issue number)\r\n- [
] [Tests added and passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#writing-tests) if fixing a bug
or adding a new feature\r\n- [ ] All [code checks
passed](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#pre-commit).\r\n- [ ] Added
[type annotations](https://pandas.pydata.org/pandas-
docs/dev/development/contributing_codebase.html#type-hints) to new
arguments/methods/functions.\r\n- [ ] Added an entry in the latest
`doc/source/whatsnew/vX.X.X.rst` file if fixing a bug or adding a new
feature.\r\n',
'closed_by': None,
'reactions': {'url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61576/reactions',
'total_count': 0,
'+1': 0,
'-1': 0,
'laugh': 0,
'hooray': 0,
'confused': 0,
'heart': 0,
'rocket': 0,
'eyes': 0},
'timeline_url': 'https://api.github.com/repos/pandas-
dev/pandas/issues/61576/timeline',
'performed_via_github_app': None,
'state_reason': None}

```

```
[173]: js[0]['url']
```

```
[173]: 'https://api.github.com/repos/pandas-dev/pandas/issues/61576'
```

```
[174]: js[0]['title']
```

```
[174]: 'DataFrames Class update'
```

```
[175]: #Printing All Titles
for i in range(len(js)):
    print(js[i]['title'])
```

DataFrames Class update

BUG: 2.3.0 didn't publish wheels for musl-aarch64 (arm)

WEB: Test that our versions JSON is valid  
 BUG: comparing strings of different dtypes errors  
 BLD: Build wheels for 3.9  
 BUG: Inconsistent behavior surrounding pd.fillna  
 BUILD: Add wheels for musllinux\_aarch64  
 BUILD: Installation issue on Mac with M1 Pro arm64 processor.  
 pandas\_parser.cpython-311-darwin.so is using x86\_64 arch  
 BUG: RecursionError when apply generic alias as a func  
 Failed to install pandas==2.3.0 with Python 3.9  
 BUG: regex match in compliance tests no longer match pytest expected inputs  
 Upgrade Python version for development environment  
 BUG: duplicated() raises error with singleton set as subset  
 DOC: Add note on inference behavior of apply with result\_type='expand'  
 Parallelize test\_sql.py - Issue #60378  
 DOC: Remove and Update out of date Docker Image issue with #61511  
 ENH: Adding DataFrame plotting benchmarks for large datasets  
 DOC: Fix docs for BusinessDay constructor #52431  
 CI: Debug slow environment solve times  
 Fix Index.equals between object and string  
 ENH pandas-dev#60693: shift operations for Series and DataFrames  
 ENH: Columns formatted as "Text" in Excel are read as numbers  
 BUG: use of iloc with heterogeneous DataFrame sometimes performs undocumented conversions  
 ENH: read\_csv tz option  
 BUG: Fix Series.str.zfill for ArrowDtype string arrays #61485  
 ENH: speed up `DataFrame.plot` using `LineCollection`  
 CI: Micromamba taking too long to resolve the environments in the CI  
 API: Replace na\_action parameter in Series/DataFrame/Index.map by the standard skipna  
 raise TypeError when function returns object-like results  
 ENH: Implement DataFrame.select

```
[176]: # Get and display the number of issues (i.e., the number of dictionaries) in
      ↪ the list 'js'
      len(js)
```

[176]: 30

```
[177]: # Create a DataFrame from the list of GitHub issues (js) and extract only the
      ↪ 'user', 'url', and 'title' columns
      pd.DataFrame(js, columns=['user', 'url', 'title'])
```

```
[177]:
```

	user	url	title
0	{'login': 'SaraTammame', 'id': 116681400, 'nod...		
1	{'login': 'virajmehta', 'id': 9244420, 'node_i...		
2	{'login': 'datapythonista', 'id': 10058240, 'n...		
3	{'login': 'a-reich', 'id': 73507369, 'node_id'...		

```

4  {'login': 'mroeschke', 'id': 10647082, 'node_i...
5  {'login': 'anna-intellegens', 'id': 98812604, ...
6  {'login': 'cdce8p', 'id': 30130371, 'node_id':...
7  {'login': 'Leviann', 'id': 2115421, 'node_id':...
8  {'login': 'MacroBull', 'id': 2204917, 'node_id...
9  {'login': 'ryanchao2012', 'id': 7628311, 'node...
10 {'login': 'chalmerlowe', 'id': 7291104, 'node_...
11 {'login': 'datapythonista', 'id': 10058240, 'n...
12 {'login': 'camold', 'id': 2109454, 'node_id': ...
13 {'login': 'pinkgradient', 'id': 101425679, 'no...
14 {'login': 'dShcherbakov1', 'id': 119084461, 'n...
15 {'login': 'jacksnnn', 'id': 175151559, 'node_i...
16 {'login': 'shadnikn', 'id': 147014887, 'node_i...
17 {'login': 'camacluc', 'id': 114314487, 'node_i...
18 {'login': 'mroeschke', 'id': 10647082, 'node_i...
19 {'login': 'sanggon6107', 'id': 68040183, 'node...
20 {'login': 'David-msggc', 'id': 95498196, 'node...
21 {'login': 'pranay-sa', 'id': 116664160, 'node_...
22 {'login': 'illbebach', 'id': 1385858, 'node_id...
23 {'login': 'hasandiwan', 'id': 44325453, 'node_...
24 {'login': 'iabhi4', 'id': 61010675, 'node_id':...
25 {'login': 'Abdelgha-4', 'id': 72174073, 'node_...
26 {'login': 'datapythonista', 'id': 10058240, 'n...
27 {'login': 'datapythonista', 'id': 10058240, 'n...
28 {'login': 'michael2015tse', 'id': 44341134, 'n...
29 {'login': 'datapythonista', 'id': 10058240, 'n...

```

url \

```

0  https://api.github.com/repos/pandas-dev/pandas...
1  https://api.github.com/repos/pandas-dev/pandas...
2  https://api.github.com/repos/pandas-dev/pandas...
3  https://api.github.com/repos/pandas-dev/pandas...
4  https://api.github.com/repos/pandas-dev/pandas...
5  https://api.github.com/repos/pandas-dev/pandas...
6  https://api.github.com/repos/pandas-dev/pandas...
7  https://api.github.com/repos/pandas-dev/pandas...
8  https://api.github.com/repos/pandas-dev/pandas...
9  https://api.github.com/repos/pandas-dev/pandas...
10 https://api.github.com/repos/pandas-dev/pandas...
11 https://api.github.com/repos/pandas-dev/pandas...
12 https://api.github.com/repos/pandas-dev/pandas...
13 https://api.github.com/repos/pandas-dev/pandas...
14 https://api.github.com/repos/pandas-dev/pandas...
15 https://api.github.com/repos/pandas-dev/pandas...
16 https://api.github.com/repos/pandas-dev/pandas...
17 https://api.github.com/repos/pandas-dev/pandas...
18 https://api.github.com/repos/pandas-dev/pandas...

```

```

19 https://api.github.com/repos/pandas-dev/pandas...
20 https://api.github.com/repos/pandas-dev/pandas...
21 https://api.github.com/repos/pandas-dev/pandas...
22 https://api.github.com/repos/pandas-dev/pandas...
23 https://api.github.com/repos/pandas-dev/pandas...
24 https://api.github.com/repos/pandas-dev/pandas...
25 https://api.github.com/repos/pandas-dev/pandas...
26 https://api.github.com/repos/pandas-dev/pandas...
27 https://api.github.com/repos/pandas-dev/pandas...
28 https://api.github.com/repos/pandas-dev/pandas...
29 https://api.github.com/repos/pandas-dev/pandas...

                                title
0                                DataFrames Class update
1  BUG: 2.3.0 didn't publish wheels for musl-aarc...
2                                WEB: Test that our versions JSON is valid
3  BUG: comparing strings of different dtypes errors
4                                BLD: Build wheels for 3.9
5  BUG: Inconsistent behavior surrounding pd.fillna
6                                BUILD: Add wheels for musllinux_aarch64
7  BUILD: Installation issue on Mac with M1 Pro a...
8  BUG: RecursionError when apply generic alias a...
9    Failed to install pandas==2.3.0 with Python 3.9
10 BUG: regex match in compliance tests no longer...
11 Upgrade Python version for development environ...
12 BUG: duplicated() raises error with singleton s...
13 DOC: Add note on inference behavior of apply w...
14                                Parallelize test_sql.py - Issue #60378
15 DOC: Remove and Update out of date Docker Imag...
16 ENH: Adding DataFrame plotting benchmarks for ...
17 DOC: Fix docs for BusinessDay constructor #52431
18                                CI: Debug slow environment solve times
19                                Fix Index.equals between object and string
20 ENH pandas-dev#60693: shift operations for Ser...
21 ENH: Columns formatted as "Text" in Excel are ...
22 BUG: use of iloc with heterogeneous DataFrame ...
23                                ENH: read_csv tz option
24 BUG: Fix Series.str.zfill for ArrowDtype strin...
25 ENH: speed up `DataFrame.plot` using `LineColl...
26 CI: Micromamba taking too long to resolve the ...
27 API: Replace na_action parameter in Series/Dat...
28 raise TypeError when function returns object-l...
29                                ENH: Implement DataFrame.select

```

```

[178]: # View the first few rows of the dataset
pd.DataFrame(js).head()

```

```

[178]:
                                url \
0 https://api.github.com/repos/pandas-dev/pandas...
1 https://api.github.com/repos/pandas-dev/pandas...
2 https://api.github.com/repos/pandas-dev/pandas...
3 https://api.github.com/repos/pandas-dev/pandas...
4 https://api.github.com/repos/pandas-dev/pandas...

                                repository_url \
0 https://api.github.com/repos/pandas-dev/pandas
1 https://api.github.com/repos/pandas-dev/pandas
2 https://api.github.com/repos/pandas-dev/pandas
3 https://api.github.com/repos/pandas-dev/pandas
4 https://api.github.com/repos/pandas-dev/pandas

                                labels_url \
0 https://api.github.com/repos/pandas-dev/pandas...
1 https://api.github.com/repos/pandas-dev/pandas...
2 https://api.github.com/repos/pandas-dev/pandas...
3 https://api.github.com/repos/pandas-dev/pandas...
4 https://api.github.com/repos/pandas-dev/pandas...

                                comments_url \
0 https://api.github.com/repos/pandas-dev/pandas...
1 https://api.github.com/repos/pandas-dev/pandas...
2 https://api.github.com/repos/pandas-dev/pandas...
3 https://api.github.com/repos/pandas-dev/pandas...
4 https://api.github.com/repos/pandas-dev/pandas...

                                events_url \
0 https://api.github.com/repos/pandas-dev/pandas...
1 https://api.github.com/repos/pandas-dev/pandas...
2 https://api.github.com/repos/pandas-dev/pandas...
3 https://api.github.com/repos/pandas-dev/pandas...
4 https://api.github.com/repos/pandas-dev/pandas...

                                html_url      id \
0 https://github.com/pandas-dev/pandas/pull/61576 3122405296
1 https://github.com/pandas-dev/pandas/issues/61574 3122363438
2 https://github.com/pandas-dev/pandas/issues/61573 3122306508
3 https://github.com/pandas-dev/pandas/issues/61570 3122126379
4 https://github.com/pandas-dev/pandas/pull/61569 3122042354

                                node_id  number \
0 PR_kwDOAAOYD86ZQ5dU 61576
1 I_kwDOAAOYD866G3wu 61574
2 I_kwDOAAOYD866Gp3M 61573
3 I_kwDOAAOYD866F94r 61570

```

4 PR\_kwD0AA0YD86ZPoQa 61569

	title	...	active_lock_reason	\
0	DataFrames Class update	...	None	
1	BUG: 2.3.0 didn't publish wheels for musl-aarc...	...	None	
2	WEB: Test that our versions JSON is valid	...	None	
3	BUG: comparing strings of different dtypes errors	...	None	
4	BLD: Build wheels for 3.9	...	None	

	draft	pull_request	\
0	False	{'url': 'https://api.github.com/repos/pandas-d...	
1	NaN	NaN	
2	NaN	NaN	
3	NaN	NaN	
4	False	{'url': 'https://api.github.com/repos/pandas-d...	

	body	closed_by	\
0	- [ ] closes #xxxx (Replace xxxx with the GitH...	None	
1	### Pandas version checks\n\n- [x] I have chec...	None	
2	See #61572\n\nI think this can be as simple as...	None	
3	### Pandas version checks\n\n- [x] I have chec...	None	
4	@lithomas1 would I need to re-tag the 2.3.x br...	None	

	reactions	\
0	{'url': 'https://api.github.com/repos/pandas-d...	
1	{'url': 'https://api.github.com/repos/pandas-d...	
2	{'url': 'https://api.github.com/repos/pandas-d...	
3	{'url': 'https://api.github.com/repos/pandas-d...	
4	{'url': 'https://api.github.com/repos/pandas-d...	

	timeline_url	\
0	https://api.github.com/repos/pandas-dev/pandas...	
1	https://api.github.com/repos/pandas-dev/pandas...	
2	https://api.github.com/repos/pandas-dev/pandas...	
3	https://api.github.com/repos/pandas-dev/pandas...	
4	https://api.github.com/repos/pandas-dev/pandas...	

	performed_via_github_app	state_reason	\
0	None	None	
1	None	None	
2	None	None	
3	None	None	
4	None	None	

	sub_issues_summary	
0	NaN	
1	{'total': 0, 'completed': 0, 'percent_complete...	



```

2 {'total': 0, 'completed': 0, 'percent_complete...
3 {'total': 0, 'completed': 0, 'percent_complete...
4                                     NaN

```

[5 rows x 33 columns]

```

[179]: # Storing the parsed JSON data (GitHub issues) into a CSV file

# Convert the list of dictionaries `js` into a DataFrame - each key becomes a
↪column
df4 = pd.DataFrame(js)

# Save the DataFrame to a CSV file named 'json_dump.csv'
df4.to_csv('json_dump.csv')

```

```

[180]: # Access the 'user' field from the first issue in the list
# The value of 'user' is itself a dictionary containing user details
js[0]['user']

```

```

[180]: {'login': 'SaraTammame',
'id': 116681400,
'node_id': 'U_kgDOBvRquA',
'avatar_url': 'https://avatars.githubusercontent.com/u/116681400?v=4',
'gravatar_id': '',
'url': 'https://api.github.com/users/SaraTammame',
'html_url': 'https://github.com/SaraTammame',
'followers_url': 'https://api.github.com/users/SaraTammame/followers',
'following_url':
'https://api.github.com/users/SaraTammame/following{/other_user}',
'gists_url': 'https://api.github.com/users/SaraTammame/gists{/gist_id}',
'starred_url':
'https://api.github.com/users/SaraTammame/starred{/owner}/{/repo}',
'subscriptions_url': 'https://api.github.com/users/SaraTammame/subscriptions',
'organizations_url': 'https://api.github.com/users/SaraTammame/orgs',
'repos_url': 'https://api.github.com/users/SaraTammame/repos',
'events_url': 'https://api.github.com/users/SaraTammame/events{/privacy}',
'received_events_url':
'https://api.github.com/users/SaraTammame/received_events',
'type': 'User',
'user_view_type': 'public',
'site_admin': False}

```

```

[181]: # Load JSON data directly into a DataFrame (basic structure, may not flatten
↪nested fields like 'user')
pd.read_json("https://api.github.com/repos/pandas-dev/pandas/issues")

```

[181]:

url \

```
0 https://api.github.com/repos/pandas-dev/pandas...
1 https://api.github.com/repos/pandas-dev/pandas...
2 https://api.github.com/repos/pandas-dev/pandas...
3 https://api.github.com/repos/pandas-dev/pandas...
4 https://api.github.com/repos/pandas-dev/pandas...
5 https://api.github.com/repos/pandas-dev/pandas...
6 https://api.github.com/repos/pandas-dev/pandas...
7 https://api.github.com/repos/pandas-dev/pandas...
8 https://api.github.com/repos/pandas-dev/pandas...
9 https://api.github.com/repos/pandas-dev/pandas...
10 https://api.github.com/repos/pandas-dev/pandas...
11 https://api.github.com/repos/pandas-dev/pandas...
12 https://api.github.com/repos/pandas-dev/pandas...
13 https://api.github.com/repos/pandas-dev/pandas...
14 https://api.github.com/repos/pandas-dev/pandas...
15 https://api.github.com/repos/pandas-dev/pandas...
16 https://api.github.com/repos/pandas-dev/pandas...
17 https://api.github.com/repos/pandas-dev/pandas...
18 https://api.github.com/repos/pandas-dev/pandas...
19 https://api.github.com/repos/pandas-dev/pandas...
20 https://api.github.com/repos/pandas-dev/pandas...
21 https://api.github.com/repos/pandas-dev/pandas...
22 https://api.github.com/repos/pandas-dev/pandas...
23 https://api.github.com/repos/pandas-dev/pandas...
24 https://api.github.com/repos/pandas-dev/pandas...
25 https://api.github.com/repos/pandas-dev/pandas...
26 https://api.github.com/repos/pandas-dev/pandas...
27 https://api.github.com/repos/pandas-dev/pandas...
28 https://api.github.com/repos/pandas-dev/pandas...
29 https://api.github.com/repos/pandas-dev/pandas...
```

repository\_url \

```
0 https://api.github.com/repos/pandas-dev/pandas
1 https://api.github.com/repos/pandas-dev/pandas
2 https://api.github.com/repos/pandas-dev/pandas
3 https://api.github.com/repos/pandas-dev/pandas
4 https://api.github.com/repos/pandas-dev/pandas
5 https://api.github.com/repos/pandas-dev/pandas
6 https://api.github.com/repos/pandas-dev/pandas
7 https://api.github.com/repos/pandas-dev/pandas
8 https://api.github.com/repos/pandas-dev/pandas
9 https://api.github.com/repos/pandas-dev/pandas
10 https://api.github.com/repos/pandas-dev/pandas
11 https://api.github.com/repos/pandas-dev/pandas
12 https://api.github.com/repos/pandas-dev/pandas
13 https://api.github.com/repos/pandas-dev/pandas
```





12 <https://api.github.com/repos/pandas-dev/pandas...>  
 13 <https://api.github.com/repos/pandas-dev/pandas...>  
 14 <https://api.github.com/repos/pandas-dev/pandas...>  
 15 <https://api.github.com/repos/pandas-dev/pandas...>  
 16 <https://api.github.com/repos/pandas-dev/pandas...>  
 17 <https://api.github.com/repos/pandas-dev/pandas...>  
 18 <https://api.github.com/repos/pandas-dev/pandas...>  
 19 <https://api.github.com/repos/pandas-dev/pandas...>  
 20 <https://api.github.com/repos/pandas-dev/pandas...>  
 21 <https://api.github.com/repos/pandas-dev/pandas...>  
 22 <https://api.github.com/repos/pandas-dev/pandas...>  
 23 <https://api.github.com/repos/pandas-dev/pandas...>  
 24 <https://api.github.com/repos/pandas-dev/pandas...>  
 25 <https://api.github.com/repos/pandas-dev/pandas...>  
 26 <https://api.github.com/repos/pandas-dev/pandas...>  
 27 <https://api.github.com/repos/pandas-dev/pandas...>  
 28 <https://api.github.com/repos/pandas-dev/pandas...>  
 29 <https://api.github.com/repos/pandas-dev/pandas...>

	html_url	id \
0	<a href="https://github.com/pandas-dev/pandas/pull/61576">https://github.com/pandas-dev/pandas/pull/61576</a>	3122405296
1	<a href="https://github.com/pandas-dev/pandas/issues/61574">https://github.com/pandas-dev/pandas/issues/61574</a>	3122363438
2	<a href="https://github.com/pandas-dev/pandas/issues/61573">https://github.com/pandas-dev/pandas/issues/61573</a>	3122306508
3	<a href="https://github.com/pandas-dev/pandas/issues/61570">https://github.com/pandas-dev/pandas/issues/61570</a>	3122126379
4	<a href="https://github.com/pandas-dev/pandas/pull/61569">https://github.com/pandas-dev/pandas/pull/61569</a>	3122042354
5	<a href="https://github.com/pandas-dev/pandas/issues/61568">https://github.com/pandas-dev/pandas/issues/61568</a>	3121700295
6	<a href="https://github.com/pandas-dev/pandas/pull/61567">https://github.com/pandas-dev/pandas/pull/61567</a>	3121077669
7	<a href="https://github.com/pandas-dev/pandas/issues/61566">https://github.com/pandas-dev/pandas/issues/61566</a>	3120672132
8	<a href="https://github.com/pandas-dev/pandas/issues/61565">https://github.com/pandas-dev/pandas/issues/61565</a>	3120501218
9	<a href="https://github.com/pandas-dev/pandas/issues/61563">https://github.com/pandas-dev/pandas/issues/61563</a>	3119772338
10	<a href="https://github.com/pandas-dev/pandas/issues/61557">https://github.com/pandas-dev/pandas/issues/61557</a>	3118522709
11	<a href="https://github.com/pandas-dev/pandas/pull/61555">https://github.com/pandas-dev/pandas/pull/61555</a>	3117071656
12	<a href="https://github.com/pandas-dev/pandas/issues/61554">https://github.com/pandas-dev/pandas/issues/61554</a>	3117048901
13	<a href="https://github.com/pandas-dev/pandas/pull/61552">https://github.com/pandas-dev/pandas/pull/61552</a>	3116474516
14	<a href="https://github.com/pandas-dev/pandas/pull/61551">https://github.com/pandas-dev/pandas/pull/61551</a>	3116401342
15	<a href="https://github.com/pandas-dev/pandas/pull/61550">https://github.com/pandas-dev/pandas/pull/61550</a>	3116382404
16	<a href="https://github.com/pandas-dev/pandas/pull/61546">https://github.com/pandas-dev/pandas/pull/61546</a>	3115813278
17	<a href="https://github.com/pandas-dev/pandas/pull/61543">https://github.com/pandas-dev/pandas/pull/61543</a>	3115248818
18	<a href="https://github.com/pandas-dev/pandas/pull/61542">https://github.com/pandas-dev/pandas/pull/61542</a>	3114631618
19	<a href="https://github.com/pandas-dev/pandas/pull/61541">https://github.com/pandas-dev/pandas/pull/61541</a>	3114329620
20	<a href="https://github.com/pandas-dev/pandas/pull/61540">https://github.com/pandas-dev/pandas/pull/61540</a>	3113693938
21	<a href="https://github.com/pandas-dev/pandas/issues/61539">https://github.com/pandas-dev/pandas/issues/61539</a>	3113587665
22	<a href="https://github.com/pandas-dev/pandas/issues/61537">https://github.com/pandas-dev/pandas/issues/61537</a>	3111287144
23	<a href="https://github.com/pandas-dev/pandas/issues/61535">https://github.com/pandas-dev/pandas/issues/61535</a>	3108703108
24	<a href="https://github.com/pandas-dev/pandas/pull/61533">https://github.com/pandas-dev/pandas/pull/61533</a>	3107929504
25	<a href="https://github.com/pandas-dev/pandas/issues/61532">https://github.com/pandas-dev/pandas/issues/61532</a>	3107872632
26	<a href="https://github.com/pandas-dev/pandas/issues/61531">https://github.com/pandas-dev/pandas/issues/61531</a>	3107305231

27	<a href="https://github.com/pandas-dev/pandas/pull/61530">https://github.com/pandas-dev/pandas/pull/61530</a>	3107129953
28	<a href="https://github.com/pandas-dev/pandas/issues/61529">https://github.com/pandas-dev/pandas/issues/61529</a>	3106434267
29	<a href="https://github.com/pandas-dev/pandas/pull/61527">https://github.com/pandas-dev/pandas/pull/61527</a>	3105461098

	node_id	number	\
0	PR_kwDOAAOYD86ZQ5dU	61576	
1	I_kwDOAAOYD866G3wu	61574	
2	I_kwDOAAOYD866Gp3M	61573	
3	I_kwDOAAOYD866F94r	61570	
4	PR_kwDOAAOYD86ZPoQa	61569	
5	I_kwDOAAOYD866EV3H	61568	
6	PR_kwDOAAOYD86ZMg54	61567	
7	I_kwDOAAOYD866Aa2E	61566	
8	I_kwDOAAOYD865_xHi	61565	
9	I_kwDOAAOYD8658_Ky	61563	
10	I_kwDOAAOYD86540FV	61557	
11	PR_kwDOAAOYD86Y-985	61555	
12	I_kwDOAAOYD865ymRF	61554	
13	PR_kwDOAAOYD86Y9COH	61552	
14	PR_kwDOAAOYD86Y80iP	61551	
15	PR_kwDOAAOYD86Y8ynF	61550	
16	PR_kwDOAAOYD86Y69PY	61546	
17	PR_kwDOAAOYD86Y5Ety	61543	
18	PR_kwDOAAOYD86Y2-S4	61542	
19	PR_kwDOAAOYD86Y1-gs	61541	
20	PR_kwDOAAOYD86Yzzcs	61540	
21	I_kwDOAAOYD8651ZPR	61539	
22	I_kwDOAAOYD865cnlo	61537	
23	I_kwDOAAOYD865SwuE	61535	
24	PR_kwDOAAOYD86YgpaA	61533	
25	I_kwDOAAOYD865P194	61532	
26	I_kwDOAAOYD865NbcP	61531	
27	PR_kwDOAAOYD86YePF9	61530	
28	I_kwDOAAOYD865KGzb	61529	
29	PR_kwDOAAOYD86YZGj2	61527	

	title	...	active_lock_reason	\
0	DataFrames Class update	...	NaN	
1	BUG: 2.3.0 didn't publish wheels for musl-aarc...	...	NaN	
2	WEB: Test that our versions JSON is valid	...	NaN	
3	BUG: comparing strings of different dtypes errors	...	NaN	
4	BLD: Build wheels for 3.9	...	NaN	
5	BUG: Inconsistent behavior surrounding pd.fillna	...	NaN	
6	BUILD: Add wheels for musllinux_aarch64	...	NaN	
7	BUILD: Installation issue on Mac with M1 Pro a...	...	NaN	
8	BUG: RecursionError when apply generic alias a...	...	NaN	
9	Failed to install pandas==2.3.0 with Python 3.9	...	NaN	

10	BUG: regex match in compliance tests no longer...	NaN
11	Upgrade Python version for development environ...	NaN
12	BUG: duplicated() raises error with singleton s...	NaN
13	DOC: Add note on inference behavior of apply w...	NaN
14	Parallelize test_sql.py - Issue #60378	NaN
15	DOC: Remove and Update out of date Docker Imag...	NaN
16	ENH: Adding DataFrame plotting benchmarks for ...	NaN
17	DOC: Fix docs for BusinessDay constructor #52431	NaN
18	CI: Debug slow environment solve times	NaN
19	Fix Index.equals between object and string	NaN
20	ENH pandas-dev#60693: shift operations for Ser...	NaN
21	ENH: Columns formatted as "Text" in Excel are ...	NaN
22	BUG: use of iloc with heterogeneous DataFrame ...	NaN
23	ENH: read_csv tz option	NaN
24	BUG: Fix Series.str.zfill for ArrowDtype strin...	NaN
25	ENH: speed up `DataFrame.plot` using `LineColl...	NaN
26	CI: Micromamba taking too long to resolve the ...	NaN
27	API: Replace na_action parameter in Series/Dat...	NaN
28	raise TypeError when function returns object-l...	NaN
29	ENH: Implement DataFrame.select	NaN

	draft	pull_request \
0	0.0	{'url': 'https://api.github.com/repos/pandas-d...
1	NaN	NaN
2	NaN	NaN
3	NaN	NaN
4	0.0	{'url': 'https://api.github.com/repos/pandas-d...
5	NaN	NaN
6	0.0	{'url': 'https://api.github.com/repos/pandas-d...
7	NaN	NaN
8	NaN	NaN
9	NaN	NaN
10	NaN	NaN
11	0.0	{'url': 'https://api.github.com/repos/pandas-d...
12	NaN	NaN
13	0.0	{'url': 'https://api.github.com/repos/pandas-d...
14	0.0	{'url': 'https://api.github.com/repos/pandas-d...
15	0.0	{'url': 'https://api.github.com/repos/pandas-d...
16	0.0	{'url': 'https://api.github.com/repos/pandas-d...
17	0.0	{'url': 'https://api.github.com/repos/pandas-d...
18	1.0	{'url': 'https://api.github.com/repos/pandas-d...
19	0.0	{'url': 'https://api.github.com/repos/pandas-d...
20	0.0	{'url': 'https://api.github.com/repos/pandas-d...
21	NaN	NaN
22	NaN	NaN
23	NaN	NaN
24	0.0	{'url': 'https://api.github.com/repos/pandas-d...

25	NaN	NaN
26	NaN	NaN
27	0.0	{'url': 'https://api.github.com/repos/pandas-d...
28	NaN	NaN
29	0.0	{'url': 'https://api.github.com/repos/pandas-d...

		body closed_by \
0	- [ ] closes #xxxx (Replace xxxx with the GitH...	NaN
1	### Pandas version checks\n\n- [x] I have chec...	NaN
2	See #61572\n\nI think this can be as simple as...	NaN
3	### Pandas version checks\n\n- [x] I have chec...	NaN
4	@lithomas1 would I need to re-tag the 2.3.x br...	NaN
5	### Pandas version checks\n\n- [x] I have chec...	NaN
6	It seems the `musllinux_aarch64` wheels got ac...	NaN
7	### Installation check\n\n- [x] I have read th...	NaN
8	### Pandas version checks\n\n- [x] I have chec...	NaN
9	### Installation check\n\n- [x] I have read th...	NaN
10	### Pandas version checks\n\n- [x] I have chec...	NaN
11	I guess the only reason we require Python 3.10...	NaN
12	### Pandas version checks\n\n- [x] I have chec...	NaN
13	Closes #61057\r\n\r\n- This PR adds a docstrin...	NaN
14	- [ ] Closes #60378\r\n- [ ] Builds off of htt...	NaN
15	- [ ] Addresses & closes [DOC: Docker image pr...	NaN
16	- [ ] related to #61532 - Adding in performan...	NaN
17	fix constructor for offset BusinessDay\r\n\r\n...	NaN
18	- [ ] closes #xxxx (Replace xxxx with the GitH...	NaN
19	- [X] closes #61099 \r\n- [X] [Tests added and...	NaN
20	This commit introduces the rshift and lshift m...	NaN
21	### Feature Type\n\n- [x] Adding new functiona...	NaN
22	### Pandas version checks\n\n- [x] I have chec...	NaN
23	### Feature Type\n\n- [x] Adding new functiona...	NaN
24	Implemented `_str_zfill` for `ArrowExtensionAr...	NaN
25	**Description:**\n\nWhen plotting line charts ...	NaN
26	Our CI jobs are frequently failing now as they...	NaN
27	- [X] xref #61128\r\n- [X] [Tests added and pa...	NaN
28	This roll_apply raises `TypeError: must be rea...	NaN
29	- [X] closes #61522\r\n- [X] [Tests added and ...	NaN

		reactions \
0	{'url': 'https://api.github.com/repos/pandas-d...	
1	{'url': 'https://api.github.com/repos/pandas-d...	
2	{'url': 'https://api.github.com/repos/pandas-d...	
3	{'url': 'https://api.github.com/repos/pandas-d...	
4	{'url': 'https://api.github.com/repos/pandas-d...	
5	{'url': 'https://api.github.com/repos/pandas-d...	
6	{'url': 'https://api.github.com/repos/pandas-d...	
7	{'url': 'https://api.github.com/repos/pandas-d...	





```

23 https://api.github.com/repos/pandas-dev/pandas...
24 https://api.github.com/repos/pandas-dev/pandas...
25 https://api.github.com/repos/pandas-dev/pandas...
26 https://api.github.com/repos/pandas-dev/pandas...
27 https://api.github.com/repos/pandas-dev/pandas...
28 https://api.github.com/repos/pandas-dev/pandas...
29 https://api.github.com/repos/pandas-dev/pandas...

```

	performed_via_github_app	state_reason	\
0	NaN	NaN	
1	NaN	NaN	
2	NaN	NaN	
3	NaN	NaN	
4	NaN	NaN	
5	NaN	NaN	
6	NaN	NaN	
7	NaN	NaN	
8	NaN	NaN	
9	NaN	NaN	
10	NaN	NaN	
11	NaN	NaN	
12	NaN	NaN	
13	NaN	NaN	
14	NaN	NaN	
15	NaN	NaN	
16	NaN	NaN	
17	NaN	NaN	
18	NaN	NaN	
19	NaN	NaN	
20	NaN	NaN	
21	NaN	NaN	
22	NaN	NaN	
23	NaN	NaN	
24	NaN	NaN	
25	NaN	NaN	
26	NaN	NaN	
27	NaN	NaN	
28	NaN	NaN	
29	NaN	NaN	

	sub_issues_summary
0	NaN
1	{'total': 0, 'completed': 0, 'percent_complete...
2	{'total': 0, 'completed': 0, 'percent_complete...
3	{'total': 0, 'completed': 0, 'percent_complete...
4	NaN
5	{'total': 0, 'completed': 0, 'percent_complete...

```

6                                     NaN
7  {'total': 0, 'completed': 0, 'percent_complete...
8  {'total': 0, 'completed': 0, 'percent_complete...
9  {'total': 0, 'completed': 0, 'percent_complete...
10 {'total': 0, 'completed': 0, 'percent_complete...
11                                     NaN
12 {'total': 0, 'completed': 0, 'percent_complete...
13                                     NaN
14                                     NaN
15                                     NaN
16                                     NaN
17                                     NaN
18                                     NaN
19                                     NaN
20                                     NaN
21 {'total': 0, 'completed': 0, 'percent_complete...
22 {'total': 0, 'completed': 0, 'percent_complete...
23 {'total': 0, 'completed': 0, 'percent_complete...
24                                     NaN
25 {'total': 0, 'completed': 0, 'percent_complete...
26 {'total': 0, 'completed': 0, 'percent_complete...
27                                     NaN
28 {'total': 0, 'completed': 0, 'percent_complete...
29                                     NaN

```

[30 rows x 33 columns]

[182]: *# Extracting the 'user' field from each issue and converting it into a DataFrame*

```

import requests

# Send GET request to the GitHub API
res = requests.get('https://api.github.com/repos/pandas-dev/pandas/issues')

# Parse JSON response into Python list of dictionaries
js = res.json()

# Create a list to store user info from each issue
user_list = []

# Loop through each issue and extract the nested 'user' dictionary
for i in range(len(js)):
    user_list.append(js[i]['user'])

# Convert the list of user dictionaries into a DataFrame
pd.DataFrame(user_list)

```

[182]:

	login	id	node_id \		avatar_url	gravatar_id \
0	SaraTamame	116681400	U_kgDOBvRquA		<a href="https://avatars.githubusercontent.com/u/116681...">https://avatars.githubusercontent.com/u/116681...</a>	
1	virajmehta	9244420	MDQ6VXNlcjkyNDQOMjA=		<a href="https://avatars.githubusercontent.com/u/924442...">https://avatars.githubusercontent.com/u/924442...</a>	
2	datapythonista	10058240	MDQ6VXNlcjEwMDU4MjQw		<a href="https://avatars.githubusercontent.com/u/100582...">https://avatars.githubusercontent.com/u/100582...</a>	
3	a-reich	73507369	MDQ6VXNlcjczNTA3MzY5		<a href="https://avatars.githubusercontent.com/u/735073...">https://avatars.githubusercontent.com/u/735073...</a>	
4	mroeschke	10647082	MDQ6VXNlcjEwNjQ3MDgy		<a href="https://avatars.githubusercontent.com/u/106470...">https://avatars.githubusercontent.com/u/106470...</a>	
5	anna-intellegens	98812604	U_kgDOBePCvA		<a href="https://avatars.githubusercontent.com/u/988126...">https://avatars.githubusercontent.com/u/988126...</a>	
6	cdce8p	30130371	MDQ6VXNlcjMwMTMwMzcw		<a href="https://avatars.githubusercontent.com/u/301303...">https://avatars.githubusercontent.com/u/301303...</a>	
7	Leviann	2115421	MDQ6VXNlcjIxMTU0MjE=		<a href="https://avatars.githubusercontent.com/u/211542...">https://avatars.githubusercontent.com/u/211542...</a>	
8	MacroBull	2204917	MDQ6VXNlcjIyMDQ5MTc=		<a href="https://avatars.githubusercontent.com/u/220491...">https://avatars.githubusercontent.com/u/220491...</a>	
9	ryanchao2012	7628311	MDQ6VXNlcjcyMjgzMTE=		<a href="https://avatars.githubusercontent.com/u/762831...">https://avatars.githubusercontent.com/u/762831...</a>	
10	chalmerlowe	7291104	MDQ6VXNlcjcyOTExMDQ=		<a href="https://avatars.githubusercontent.com/u/729110...">https://avatars.githubusercontent.com/u/729110...</a>	
11	datapythonista	10058240	MDQ6VXNlcjEwMDU4MjQw		<a href="https://avatars.githubusercontent.com/u/100582...">https://avatars.githubusercontent.com/u/100582...</a>	
12	camold	2109454	MDQ6VXNlcjIxMDk0NTQ=		<a href="https://avatars.githubusercontent.com/u/210945...">https://avatars.githubusercontent.com/u/210945...</a>	
13	pinkgradient	101425679	U_kgDOBguiDw		<a href="https://avatars.githubusercontent.com/u/101425...">https://avatars.githubusercontent.com/u/101425...</a>	
14	dShcherbakov1	119084461	U_kgDOBxkVrQ		<a href="https://avatars.githubusercontent.com/u/119084...">https://avatars.githubusercontent.com/u/119084...</a>	
15	jacksnnn	175151559	U_kgDOCnCZxw		<a href="https://avatars.githubusercontent.com/u/175151...">https://avatars.githubusercontent.com/u/175151...</a>	
16	shadnikn	147014887	U_kgDOCMNE5w		<a href="https://avatars.githubusercontent.com/u/147014...">https://avatars.githubusercontent.com/u/147014...</a>	
17	camacluc	114314487	U_kgDOBtBM9w		<a href="https://avatars.githubusercontent.com/u/114314...">https://avatars.githubusercontent.com/u/114314...</a>	
18	mroeschke	10647082	MDQ6VXNlcjEwNjQ3MDgy		<a href="https://avatars.githubusercontent.com/u/106470...">https://avatars.githubusercontent.com/u/106470...</a>	
19	sanggon6107	68040183	MDQ6VXNlcjY4MDQwMTgz		<a href="https://avatars.githubusercontent.com/u/680401...">https://avatars.githubusercontent.com/u/680401...</a>	
20	David-msggc	95498196	U_kgDOBbEv1A		<a href="https://avatars.githubusercontent.com/u/954981...">https://avatars.githubusercontent.com/u/954981...</a>	
21	pranay-sa	116664160	U_kgDOBvQnYA		<a href="https://avatars.githubusercontent.com/u/116664...">https://avatars.githubusercontent.com/u/116664...</a>	
22	illbebach	1385858	MDQ6VXNlcjEzODU4NTg=		<a href="https://avatars.githubusercontent.com/u/138585...">https://avatars.githubusercontent.com/u/138585...</a>	
23	hasandiwan	44325453	MDQ6VXNlcjQOMzI1NDUz		<a href="https://avatars.githubusercontent.com/u/443254...">https://avatars.githubusercontent.com/u/443254...</a>	
24	iabhi4	61010675	MDQ6VXNlcjYxMDEwNjc1		<a href="https://avatars.githubusercontent.com/u/610106...">https://avatars.githubusercontent.com/u/610106...</a>	
25	Abdelgha-4	72174073	MDQ6VXNlcjcyMTc0MDcz		<a href="https://avatars.githubusercontent.com/u/721740...">https://avatars.githubusercontent.com/u/721740...</a>	
26	datapythonista	10058240	MDQ6VXNlcjEwMDU4MjQw		<a href="https://avatars.githubusercontent.com/u/100582...">https://avatars.githubusercontent.com/u/100582...</a>	
27	datapythonista	10058240	MDQ6VXNlcjEwMDU4MjQw		<a href="https://avatars.githubusercontent.com/u/100582...">https://avatars.githubusercontent.com/u/100582...</a>	
28	michael2015tse	44341134	MDQ6VXNlcjQOMzQxMTM0		<a href="https://avatars.githubusercontent.com/u/443411...">https://avatars.githubusercontent.com/u/443411...</a>	
29	datapythonista	10058240	MDQ6VXNlcjEwMDU4MjQw		<a href="https://avatars.githubusercontent.com/u/100582...">https://avatars.githubusercontent.com/u/100582...</a>	

14 <https://avatars.githubusercontent.com/u/119084...>  
 15 <https://avatars.githubusercontent.com/u/175151...>  
 16 <https://avatars.githubusercontent.com/u/147014...>  
 17 <https://avatars.githubusercontent.com/u/114314...>  
 18 <https://avatars.githubusercontent.com/u/106470...>  
 19 <https://avatars.githubusercontent.com/u/680401...>  
 20 <https://avatars.githubusercontent.com/u/954981...>  
 21 <https://avatars.githubusercontent.com/u/116664...>  
 22 <https://avatars.githubusercontent.com/u/138585...>  
 23 <https://avatars.githubusercontent.com/u/443254...>  
 24 <https://avatars.githubusercontent.com/u/610106...>  
 25 <https://avatars.githubusercontent.com/u/721740...>  
 26 <https://avatars.githubusercontent.com/u/100582...>  
 27 <https://avatars.githubusercontent.com/u/100582...>  
 28 <https://avatars.githubusercontent.com/u/443411...>  
 29 <https://avatars.githubusercontent.com/u/100582...>

url \

0 <https://api.github.com/users/SaraTammame>  
 1 <https://api.github.com/users/virajmehta>  
 2 <https://api.github.com/users/datapythonista>  
 3 <https://api.github.com/users/a-reich>  
 4 <https://api.github.com/users/mroeschke>  
 5 <https://api.github.com/users/anna-intellegens>  
 6 <https://api.github.com/users/cdce8p>  
 7 <https://api.github.com/users/Leviann>  
 8 <https://api.github.com/users/MacroBull>  
 9 <https://api.github.com/users/ryanchao2012>  
 10 <https://api.github.com/users/chalmerlowe>  
 11 <https://api.github.com/users/datapythonista>  
 12 <https://api.github.com/users/camold>  
 13 <https://api.github.com/users/pinkgradient>  
 14 <https://api.github.com/users/dShcherbakov1>  
 15 <https://api.github.com/users/jacksnnn>  
 16 <https://api.github.com/users/shadnikn>  
 17 <https://api.github.com/users/camacluc>  
 18 <https://api.github.com/users/mroeschke>  
 19 <https://api.github.com/users/sanggon6107>  
 20 <https://api.github.com/users/David-msggc>  
 21 <https://api.github.com/users/pranay-sa>  
 22 <https://api.github.com/users/illbebach>  
 23 <https://api.github.com/users/hasandiwan>  
 24 <https://api.github.com/users/iabhi4>  
 25 <https://api.github.com/users/Abdelgha-4>  
 26 <https://api.github.com/users/datapythonista>  
 27 <https://api.github.com/users/datapythonista>  
 28 <https://api.github.com/users/michael2015tse>

29     <https://api.github.com/users/datapythonista>

```
                                html_url  \
0      https://github.com/SaraTammame
1      https://github.com/virajmehta
2      https://github.com/datapythonista
3      https://github.com/a-reich
4      https://github.com/mroeschke
5      https://github.com/anna-intellegens
6      https://github.com/cdce8p
7      https://github.com/Leviann
8      https://github.com/MacroBull
9      https://github.com/ryanchao2012
10     https://github.com/chalmerlowe
11     https://github.com/datapythonista
12     https://github.com/camold
13     https://github.com/pinkgradient
14     https://github.com/dShcherbakov1
15     https://github.com/jacksnnn
16     https://github.com/shadnikn
17     https://github.com/camacluc
18     https://github.com/mroeschke
19     https://github.com/sanggon6107
20     https://github.com/David-msggc
21     https://github.com/pranay-sa
22     https://github.com/illbebach
23     https://github.com/hasandiwan
24     https://github.com/iabhi4
25     https://github.com/Abdelgha-4
26     https://github.com/datapythonista
27     https://github.com/datapythonista
28     https://github.com/michael2015tse
29     https://github.com/datapythonista
```

```
                                followers_url  \
0      https://api.github.com/users/SaraTammame/followers
1      https://api.github.com/users/virajmehta/followers
2      https://api.github.com/users/datapythonista/followers
3      https://api.github.com/users/a-reich/followers
4      https://api.github.com/users/mroeschke/followers
5      https://api.github.com/users/anna-intellegens/followers
6      https://api.github.com/users/cdce8p/followers
7      https://api.github.com/users/Leviann/followers
8      https://api.github.com/users/MacroBull/followers
9      https://api.github.com/users/ryanchao2012/followers
10     https://api.github.com/users/chalmerlowe/followers
11     https://api.github.com/users/datapythonista/followers
```

```

12     https://api.github.com/users/camold/followers
13 https://api.github.com/users/pinkgradient/foll...
14 https://api.github.com/users/dShcherbakov1/fol...
15     https://api.github.com/users/jacksnnn/followers
16     https://api.github.com/users/shadnikn/followers
17     https://api.github.com/users/camacluc/followers
18     https://api.github.com/users/mroeschke/followers
19 https://api.github.com/users/sanggon6107/follo...
20 https://api.github.com/users/David-msggc/follo...
21     https://api.github.com/users/pranay-sa/followers
22     https://api.github.com/users/illbebach/followers
23 https://api.github.com/users/hasandiwan/followers
24     https://api.github.com/users/iabhi4/followers
25 https://api.github.com/users/Abdelgha-4/followers
26 https://api.github.com/users/datapythonista/fo...
27 https://api.github.com/users/datapythonista/fo...
28 https://api.github.com/users/michael2015tse/fo...
29 https://api.github.com/users/datapythonista/fo...

```

following\_url \

```

0 https://api.github.com/users/SaraTammame/follo...
1 https://api.github.com/users/virajmehta/follow...
2 https://api.github.com/users/datapythonista/fo...
3 https://api.github.com/users/a-reich/following...
4 https://api.github.com/users/mroeschke/followi...
5 https://api.github.com/users/anna-intellegens/...
6 https://api.github.com/users/cdce8p/following{...
7 https://api.github.com/users/Leviann/following...
8 https://api.github.com/users/MacroBull/followi...
9 https://api.github.com/users/ryanchao2012/foll...
10 https://api.github.com/users/chalmerlowe/follo...
11 https://api.github.com/users/datapythonista/fo...
12 https://api.github.com/users/camold/following{...
13 https://api.github.com/users/pinkgradient/foll...
14 https://api.github.com/users/dShcherbakov1/fol...
15 https://api.github.com/users/jacksnnn/followin...
16 https://api.github.com/users/shadnikn/followin...
17 https://api.github.com/users/camacluc/followin...
18 https://api.github.com/users/mroeschke/followi...
19 https://api.github.com/users/sanggon6107/follo...
20 https://api.github.com/users/David-msggc/follo...
21 https://api.github.com/users/pranay-sa/followi...
22 https://api.github.com/users/illbebach/followi...
23 https://api.github.com/users/hasandiwan/follow...
24 https://api.github.com/users/iabhi4/following{...
25 https://api.github.com/users/Abdelgha-4/follow...
26 https://api.github.com/users/datapythonista/fo...

```

```

27 https://api.github.com/users/datapythonista/fo...
28 https://api.github.com/users/michael2015tse/fo...
29 https://api.github.com/users/datapythonista/fo...

```

gists\_url \

```

0 https://api.github.com/users/SaraTammame/gists...
1 https://api.github.com/users/virajmehta/gists{...
2 https://api.github.com/users/datapythonista/gi...
3 https://api.github.com/users/a-reich/gists{/gi...
4 https://api.github.com/users/mroeschke/gists{/...
5 https://api.github.com/users/anna-intellegens/...
6 https://api.github.com/users/cdce8p/gists{/gis...
7 https://api.github.com/users/Leviann/gists{/gi...
8 https://api.github.com/users/MacroBull/gists{/...
9 https://api.github.com/users/ryanchao2012/gist...
10 https://api.github.com/users/chalmerlowe/gists...
11 https://api.github.com/users/datapythonista/gi...
12 https://api.github.com/users/camold/gists{/gis...
13 https://api.github.com/users/pinkgradient/gist...
14 https://api.github.com/users/dShcherbakov1/gis...
15 https://api.github.com/users/jacksnnn/gists{/g...
16 https://api.github.com/users/shadnikn/gists{/g...
17 https://api.github.com/users/camacluc/gists{/g...
18 https://api.github.com/users/mroeschke/gists{/...
19 https://api.github.com/users/sanggon6107/gists...
20 https://api.github.com/users/David-msggc/gists...
21 https://api.github.com/users/pranay-sa/gists{/...
22 https://api.github.com/users/illbebach/gists{/...
23 https://api.github.com/users/hasandiwan/gists{...
24 https://api.github.com/users/iabhi4/gists{/gis...
25 https://api.github.com/users/Abdelgha-4/gists{...
26 https://api.github.com/users/datapythonista/gi...
27 https://api.github.com/users/datapythonista/gi...
28 https://api.github.com/users/michael2015tse/gi...
29 https://api.github.com/users/datapythonista/gi...

```

starred\_url \

```

0 https://api.github.com/users/SaraTammame/starr...
1 https://api.github.com/users/virajmehta/starre...
2 https://api.github.com/users/datapythonista/st...
3 https://api.github.com/users/a-reich/starred{/...
4 https://api.github.com/users/mroeschke/starred...
5 https://api.github.com/users/anna-intellegens/...
6 https://api.github.com/users/cdce8p/starred{/o...
7 https://api.github.com/users/Leviann/starred{/...
8 https://api.github.com/users/MacroBull/starred...
9 https://api.github.com/users/ryanchao2012/star...

```



```

10 https://api.github.com/users/chalmerlowe/starr...
11 https://api.github.com/users/datapythonista/st...
12 https://api.github.com/users/camold/starred{/o...
13 https://api.github.com/users/pinkgradient/star...
14 https://api.github.com/users/dShcherbakov1/sta...
15 https://api.github.com/users/jacksnnn/starred{...
16 https://api.github.com/users/shadnikn/starred{...
17 https://api.github.com/users/camacluc/starred{...
18 https://api.github.com/users/mroeschke/starred...
19 https://api.github.com/users/sanggon6107/starr...
20 https://api.github.com/users/David-msggc/starr...
21 https://api.github.com/users/pranay-sa/starred...
22 https://api.github.com/users/illbebach/starred...
23 https://api.github.com/users/hasandiwan/starre...
24 https://api.github.com/users/iabhi4/starred{/o...
25 https://api.github.com/users/Abdelgha-4/starre...
26 https://api.github.com/users/datapythonista/st...
27 https://api.github.com/users/datapythonista/st...
28 https://api.github.com/users/michael2015tse/st...
29 https://api.github.com/users/datapythonista/st...

```

subscriptions\_url \

```

0 https://api.github.com/users/SaraTammame/subsc...
1 https://api.github.com/users/virajmehta/subscr...
2 https://api.github.com/users/datapythonista/su...
3 https://api.github.com/users/a-reich/subscript...
4 https://api.github.com/users/mroeschke/subscri...
5 https://api.github.com/users/anna-intellegens/...
6 https://api.github.com/users/cdce8p/subscriptions
7 https://api.github.com/users/Leviann/subscript...
8 https://api.github.com/users/MacroBull/subscri...
9 https://api.github.com/users/ryanchao2012/subs...
10 https://api.github.com/users/chalmerlowe/subsc...
11 https://api.github.com/users/datapythonista/su...
12 https://api.github.com/users/camold/subscriptions
13 https://api.github.com/users/pinkgradient/subs...
14 https://api.github.com/users/dShcherbakov1/sub...
15 https://api.github.com/users/jacksnnn/subscrip...
16 https://api.github.com/users/shadnikn/subscrip...
17 https://api.github.com/users/camacluc/subscrip...
18 https://api.github.com/users/mroeschke/subscri...
19 https://api.github.com/users/sanggon6107/subsc...
20 https://api.github.com/users/David-msggc/subsc...
21 https://api.github.com/users/pranay-sa/subscri...
22 https://api.github.com/users/illbebach/subscri...
23 https://api.github.com/users/hasandiwan/subscr...
24 https://api.github.com/users/iabhi4/subscriptions

```

```

25 https://api.github.com/users/Abdelgha-4/subscr...
26 https://api.github.com/users/datapythonista/su...
27 https://api.github.com/users/datapythonista/su...
28 https://api.github.com/users/michael2015tse/su...
29 https://api.github.com/users/datapythonista/su...

```

```

                                organizations_url \
0      https://api.github.com/users/SaraTammame/orgs
1      https://api.github.com/users/virajmehta/orgs
2      https://api.github.com/users/datapythonista/orgs
3      https://api.github.com/users/a-reich/orgs
4      https://api.github.com/users/mroeschke/orgs
5      https://api.github.com/users/anna-intellegens/...
6      https://api.github.com/users/cdce8p/orgs
7      https://api.github.com/users/Leviann/orgs
8      https://api.github.com/users/MacroBull/orgs
9      https://api.github.com/users/ryanchao2012/orgs
10     https://api.github.com/users/chalmerlowe/orgs
11     https://api.github.com/users/datapythonista/orgs
12     https://api.github.com/users/camold/orgs
13     https://api.github.com/users/pinkgradient/orgs
14     https://api.github.com/users/dShcherbakov1/orgs
15     https://api.github.com/users/jacksnnn/orgs
16     https://api.github.com/users/shadnikn/orgs
17     https://api.github.com/users/camacluc/orgs
18     https://api.github.com/users/mroeschke/orgs
19     https://api.github.com/users/sanggon6107/orgs
20     https://api.github.com/users/David-msggc/orgs
21     https://api.github.com/users/pranay-sa/orgs
22     https://api.github.com/users/illbebach/orgs
23     https://api.github.com/users/hasandiwana/orgs
24     https://api.github.com/users/iabhi4/orgs
25     https://api.github.com/users/Abdelgha-4/orgs
26     https://api.github.com/users/datapythonista/orgs
27     https://api.github.com/users/datapythonista/orgs
28     https://api.github.com/users/michael2015tse/orgs
29     https://api.github.com/users/datapythonista/orgs

```

```

                                repos_url \
0      https://api.github.com/users/SaraTammame/repos
1      https://api.github.com/users/virajmehta/repos
2      https://api.github.com/users/datapythonista/repos
3      https://api.github.com/users/a-reich/repos
4      https://api.github.com/users/mroeschke/repos
5      https://api.github.com/users/anna-intellegens/...
6      https://api.github.com/users/cdce8p/repos
7      https://api.github.com/users/Leviann/repos

```

```

8      https://api.github.com/users/MacroBull/repos
9      https://api.github.com/users/ryanchao2012/repos
10     https://api.github.com/users/chalmerlowe/repos
11     https://api.github.com/users/datapythonista/repos
12     https://api.github.com/users/camold/repos
13     https://api.github.com/users/pinkgradient/repos
14     https://api.github.com/users/dShcherbakov1/repos
15     https://api.github.com/users/jacksnnn/repos
16     https://api.github.com/users/shadnikn/repos
17     https://api.github.com/users/camacluc/repos
18     https://api.github.com/users/mroeschke/repos
19     https://api.github.com/users/sanggon6107/repos
20     https://api.github.com/users/David-msggc/repos
21     https://api.github.com/users/pranay-sa/repos
22     https://api.github.com/users/illbebach/repos
23     https://api.github.com/users/hasandiwan/repos
24     https://api.github.com/users/iabhi4/repos
25     https://api.github.com/users/Abdelgha-4/repos
26     https://api.github.com/users/datapythonista/repos
27     https://api.github.com/users/datapythonista/repos
28     https://api.github.com/users/michael2015tse/repos
29     https://api.github.com/users/datapythonista/repos

```

```

                                events_url \
0      https://api.github.com/users/SaraTammame/event...
1      https://api.github.com/users/virajmehta/events...
2      https://api.github.com/users/datapythonista/ev...
3      https://api.github.com/users/a-reich/events{/p...
4      https://api.github.com/users/mroeschke/events{...
5      https://api.github.com/users/anna-intellegens/...
6      https://api.github.com/users/cdce8p/events{/pr...
7      https://api.github.com/users/Leviann/events{/p...
8      https://api.github.com/users/MacroBull/events{...
9      https://api.github.com/users/ryanchao2012/even...
10     https://api.github.com/users/chalmerlowe/event...
11     https://api.github.com/users/datapythonista/ev...
12     https://api.github.com/users/camold/events{/pr...
13     https://api.github.com/users/pinkgradient/even...
14     https://api.github.com/users/dShcherbakov1/eve...
15     https://api.github.com/users/jacksnnn/events{/...
16     https://api.github.com/users/shadnikn/events{/...
17     https://api.github.com/users/camacluc/events{/...
18     https://api.github.com/users/mroeschke/events{...
19     https://api.github.com/users/sanggon6107/event...
20     https://api.github.com/users/David-msggc/event...
21     https://api.github.com/users/pranay-sa/events{...
22     https://api.github.com/users/illbebach/events{...

```

```

23 https://api.github.com/users/hasandiwan/events...
24 https://api.github.com/users/iabhi4/events{/pr...
25 https://api.github.com/users/Abdelgha-4/events...
26 https://api.github.com/users/datapythonista/ev...
27 https://api.github.com/users/datapythonista/ev...
28 https://api.github.com/users/michael2015tse/ev...
29 https://api.github.com/users/datapythonista/ev...

```

	received_events_url	type	user_view_type	\
0	https://api.github.com/users/SaraTammame/recei...	User	public	
1	https://api.github.com/users/virajmehta/receiv...	User	public	
2	https://api.github.com/users/datapythonista/re...	User	public	
3	https://api.github.com/users/a-reich/received_...	User	public	
4	https://api.github.com/users/mroeschke/receive...	User	public	
5	https://api.github.com/users/anna-intellegens/...	User	public	
6	https://api.github.com/users/cdce8p/received_e...	User	public	
7	https://api.github.com/users/Leviann/received_...	User	public	
8	https://api.github.com/users/MacroBull/receive...	User	public	
9	https://api.github.com/users/ryanchao2012/rece...	User	public	
10	https://api.github.com/users/chalmerlowe/recei...	User	public	
11	https://api.github.com/users/datapythonista/re...	User	public	
12	https://api.github.com/users/camold/received_e...	User	public	
13	https://api.github.com/users/pinkgradient/rece...	User	public	
14	https://api.github.com/users/dShcherbakov1/rec...	User	public	
15	https://api.github.com/users/jacksnnn/received...	User	public	
16	https://api.github.com/users/shadnikn/received...	User	public	
17	https://api.github.com/users/camacluc/received...	User	public	
18	https://api.github.com/users/mroeschke/receive...	User	public	
19	https://api.github.com/users/sanggon6107/recei...	User	public	
20	https://api.github.com/users/David-msggc/recei...	User	public	
21	https://api.github.com/users/pranay-sa/receive...	User	public	
22	https://api.github.com/users/illbebach/receive...	User	public	
23	https://api.github.com/users/hasandiwan/receiv...	User	public	
24	https://api.github.com/users/iabhi4/received_e...	User	public	
25	https://api.github.com/users/Abdelgha-4/receiv...	User	public	
26	https://api.github.com/users/datapythonista/re...	User	public	
27	https://api.github.com/users/datapythonista/re...	User	public	
28	https://api.github.com/users/michael2015tse/re...	User	public	
29	https://api.github.com/users/datapythonista/re...	User	public	

	site_admin
0	False
1	False
2	False
3	False
4	False
5	False

```

6      False
7      False
8      False
9      False
10     False
11     False
12     False
13     False
14     False
15     False
16     False
17     False
18     False
19     False
20     False
21     False
22     False
23     False
24     False
25     False
26     False
27     False
28     False
29     False

```

```

[183]: user = []
res=requests.get('https://api.github.com/repos/pandas-dev/pandas/issues')
js=res.json()
user.append(js[i]['user'])
pd.DataFrame(user)

```

```

[183]:          login          id          node_id  \
0  datapythonista  10058240  MDQ6VXNlcjEwMDU4MjQw

          avatar_url  gravatar_id  \
0  https://avatars.githubusercontent.com/u/100582...

          url  \
0  https://api.github.com/users/datapythonista

          html_url  \
0  https://github.com/datapythonista

          followers_url  \
0  https://api.github.com/users/datapythonista/fo...

          following_url  \

```

```

0 https://api.github.com/users/datapythonista/fo...
                                gists_url \
0 https://api.github.com/users/datapythonista/gi...
                                starred_url \
0 https://api.github.com/users/datapythonista/st...
                                subscriptions_url \
0 https://api.github.com/users/datapythonista/su...
                                organizations_url \
0 https://api.github.com/users/datapythonista/orgs
                                repos_url \
0 https://api.github.com/users/datapythonista/repos
                                events_url \
0 https://api.github.com/users/datapythonista/ev...
                                received_events_url type user_view_type \
0 https://api.github.com/users/datapythonista/re... User public
                                site_admin
0 False

```

```

[184]: l = []
        l.append(js[i]['user'])
        pd.DataFrame(l)

```

```

[184]: login id node_id \
0 datapythonista 10058240 MDQ6VXNlcjEwMDU4MjQw

                                avatar_url gravatar_id \
0 https://avatars.githubusercontent.com/u/100582...

                                url \
0 https://api.github.com/users/datapythonista

                                html_url \
0 https://github.com/datapythonista

                                followers_url \
0 https://api.github.com/users/datapythonista/fo...

                                following_url \
0 https://api.github.com/users/datapythonista/fo...

```

```

                                gists_url \
0  https://api.github.com/users/datapythonista/gi...

                                starred_url \
0  https://api.github.com/users/datapythonista/st...

                                subscriptions_url \
0  https://api.github.com/users/datapythonista/su...

                                organizations_url \
0  https://api.github.com/users/datapythonista/orgs

                                repos_url \
0  https://api.github.com/users/datapythonista/repos

                                events_url \
0  https://api.github.com/users/datapythonista/ev...

                                received_events_url type user_view_type \
0  https://api.github.com/users/datapythonista/re... User      public

    site_admin
0      False

```

```

[185]: json_data = pd.read_json("https://api.github.com/repos/pandas-dev/pandas/
↪issues")
data = pd.DataFrame.from_records(json_data['user'])
#View the first few rows of the dataset
data.head()

```

```

[185]:
      login      id      node_id \
0  SaraTammame 116681400      U_kgDOBvRquA
1   virajmehta  9244420  MDQ6VXNlcjkyNDQOMjA=
2  datapythonista 10058240  MDQ6VXNlcjEwMDU4MjQw
3    a-reich  73507369  MDQ6VXNlcjczNTA3MzY5
4   mroeschke 10647082  MDQ6VXNlcjEwNjQ3MDgy

                                avatar_url gravatar_id \
0  https://avatars.githubusercontent.com/u/116681...
1  https://avatars.githubusercontent.com/u/924442...
2  https://avatars.githubusercontent.com/u/100582...
3  https://avatars.githubusercontent.com/u/735073...
4  https://avatars.githubusercontent.com/u/106470...

                                url \
0  https://api.github.com/users/SaraTammame

```

```

1      https://api.github.com/users/virajmehta
2 https://api.github.com/users/datapythonista
3      https://api.github.com/users/a-reich
4      https://api.github.com/users/mroeschke

      html_url \
0      https://github.com/SaraTammame
1      https://github.com/virajmehta
2 https://github.com/datapythonista
3      https://github.com/a-reich
4      https://github.com/mroeschke

      followers_url \
0 https://api.github.com/users/SaraTammame/follo...
1 https://api.github.com/users/virajmehta/followers
2 https://api.github.com/users/datapythonista/fo...
3      https://api.github.com/users/a-reich/followers
4 https://api.github.com/users/mroeschke/followers

      following_url \
0 https://api.github.com/users/SaraTammame/follo...
1 https://api.github.com/users/virajmehta/follow...
2 https://api.github.com/users/datapythonista/fo...
3 https://api.github.com/users/a-reich/following...
4 https://api.github.com/users/mroeschke/followi...

      gists_url \
0 https://api.github.com/users/SaraTammame/gists...
1 https://api.github.com/users/virajmehta/gists{...
2 https://api.github.com/users/datapythonista/gi...
3 https://api.github.com/users/a-reich/gists{/gi...
4 https://api.github.com/users/mroeschke/gists{/...

      starred_url \
0 https://api.github.com/users/SaraTammame/starr...
1 https://api.github.com/users/virajmehta/starre...
2 https://api.github.com/users/datapythonista/st...
3 https://api.github.com/users/a-reich/starred{/...
4 https://api.github.com/users/mroeschke/starred...

      subscriptions_url \
0 https://api.github.com/users/SaraTammame/subsc...
1 https://api.github.com/users/virajmehta/subscr...
2 https://api.github.com/users/datapythonista/su...
3 https://api.github.com/users/a-reich/subscript...
4 https://api.github.com/users/mroeschke/subscri...

```



```

                                organizations_url \
0      https://api.github.com/users/SaraTammame/orgs
1      https://api.github.com/users/virajmehta/orgs
2      https://api.github.com/users/datapythonista/orgs
3      https://api.github.com/users/a-reich/orgs
4      https://api.github.com/users/mroeschke/orgs

                                repos_url \
0      https://api.github.com/users/SaraTammame/repos
1      https://api.github.com/users/virajmehta/repos
2      https://api.github.com/users/datapythonista/repos
3      https://api.github.com/users/a-reich/repos
4      https://api.github.com/users/mroeschke/repos

                                events_url \
0      https://api.github.com/users/SaraTammame/event...
1      https://api.github.com/users/virajmehta/events...
2      https://api.github.com/users/datapythonista/ev...
3      https://api.github.com/users/a-reich/events{/p...
4      https://api.github.com/users/mroeschke/events{...

                                received_events_url  type  user_view_type \
0      https://api.github.com/users/SaraTammame/recei...  User      public
1      https://api.github.com/users/virajmehta/receiv...  User      public
2      https://api.github.com/users/datapythonista/re...  User      public
3      https://api.github.com/users/a-reich/received_...  User      public
4      https://api.github.com/users/mroeschke/receive...  User      public

                                site_admin
0      False
1      False
2      False
3      False
4      False

```

```

[186]: #Read the dataset from a CSV file into a DataFrame
form_file = pd.read_csv("html_data1.csv")

#Serialize (save) the DataFrame into a pickle file - fast, efficient binary
↪format
form_file.to_pickle("my_pickle")

#Check the current working directory to know where files are being saved
import os
print(os.getcwd())

#Display the DataFrame to verify contents

```

form\_file

C:\Users\ISHA\Pandas

```
[186]:
```

	Season	Age	Team	Lg	Pos	G	GS	\
0	2009-10	20	OKC	NBA	PF	6.0	0.0	
1	2010-11	21	OKC	NBA	PF	17.0	17.0	
2	2011-12	22	OKC	NBA	PF	20.0	20.0	
3	2012-13	23	OKC	NBA	PF	11.0	11.0	
4	2013-14	24	OKC	NBA	PF	17.0	17.0	
5	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
6	2015-16	26	OKC	NBA	PF	18.0	18.0	
7	2016-17	27	TOR	NBA	PF	10.0	10.0	
8	2017-18	28	TOR	NBA	PF	10.0	9.0	
9	2018-19	29	TOR	NBA	C	24.0	0.0	
10	2019-20	30	TOR	NBA	C	11.0	0.0	
11	2020-21	31	LAC	NBA	C	2.0	0.0	
12	2021-22	32	MIL	NBA	C	6.0	0.0	
13	12 Yrs	12 Yrs	12 Yrs	12 Yrs	NaN	152.0	102.0	
14	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
15	OKC (6 Yrs)	OKC (6 Yrs)	OKC (6 Yrs)	OKC (6 Yrs)	NaN	89.0	83.0	
16	TOR (4 Yrs)	TOR (4 Yrs)	TOR (4 Yrs)	TOR (4 Yrs)	NaN	55.0	19.0	
17	MIL (1 Yr)	MIL (1 Yr)	MIL (1 Yr)	MIL (1 Yr)	NaN	6.0	0.0	
18	LAC (1 Yr)	LAC (1 Yr)	LAC (1 Yr)	LAC (1 Yr)	NaN	2.0	0.0	

	MP	FG	FGA	...	ORB	DRB	TRB	AST	STL	BLK	TOV	PF	PTS	Awards
0	25.5	3.3	5.8	...	2.2	4.3	6.5	0.3	0.3	2.0	1.0	3.2	7.8	NaN
1	28.8	3.9	8.5	...	2.8	4.5	7.3	0.2	0.2	3.1	1.2	3.8	9.8	NaN
2	28.5	4.2	8.0	...	2.4	3.4	5.8	0.6	0.6	3.0	0.6	2.9	9.8	NaN
3	33.3	5.4	12.3	...	3.5	4.8	8.4	0.7	0.0	3.0	0.6	3.5	12.8	NaN
4	34.2	5.1	8.6	...	1.8	5.0	6.8	0.7	0.6	2.5	1.2	3.2	12.1	NaN
5	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6	33.4	4.7	9.1	...	1.3	5.0	6.3	0.6	0.8	1.3	0.5	2.8	12.0	NaN
7	30.7	6.0	13.0	...	0.9	5.6	6.5	1.4	0.4	1.7	1.1	3.7	14.3	NaN
8	26.0	3.0	7.2	...	1.3	4.6	5.9	1.1	0.1	1.3	1.6	2.1	8.7	NaN
9	20.8	3.8	8.0	...	1.5	4.5	6.0	0.9	0.5	1.0	1.0	3.3	9.4	NaN
10	22.8	5.7	10.0	...	1.5	6.2	7.7	1.2	0.2	1.3	1.1	2.5	14.8	NaN
11	9.0	1.5	3.0	...	0.5	1.5	2.0	1.0	0.5	1.5	0.5	0.5	5.0	NaN
12	3.7	0.7	2.7	...	0.3	1.3	1.7	0.0	0.0	0.0	0.3	0.2	1.5	NaN
13	27.1	4.3	8.6	...	1.8	4.5	6.3	0.7	0.4	1.9	0.9	3.0	10.6	NaN
14	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
15	31.0	4.5	8.8	...	2.2	4.5	6.7	0.5	0.5	2.5	0.8	3.2	10.9	NaN
16	24.0	4.5	9.2	...	1.3	5.1	6.4	1.1	0.3	1.2	1.2	3.0	11.2	NaN
17	3.7	0.7	2.7	...	0.3	1.3	1.7	0.0	0.0	0.0	0.3	0.2	1.5	NaN
18	9.0	1.5	3.0	...	0.5	1.5	2.0	1.0	0.5	1.5	0.5	0.5	5.0	NaN

[19 rows x 31 columns]

## Reading the Titanic Dataset

```
[187]: # Read the dataset into a DataFrame
df = pd.read_csv("https://raw.githubusercontent.com/datasciencedojo/datasets/
↳master/titanic.csv")
```

```
[188]: # Statistical summary of numeric columns
df.describe()
```

```
[188]:
```

	PassengerId	Survived	Pclass	Age	SibSp \
count	891.000000	891.000000	891.000000	714.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008
std	257.353842	0.486592	0.836071	14.526497	1.102743
min	1.000000	0.000000	1.000000	0.420000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000
50%	446.000000	0.000000	3.000000	28.000000	0.000000
75%	668.500000	1.000000	3.000000	38.000000	1.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000

  

	Parch	Fare
count	891.000000	891.000000
mean	0.381594	32.204208
std	0.806057	49.693429
min	0.000000	0.000000
25%	0.000000	7.910400
50%	0.000000	14.454200
75%	0.000000	31.000000
max	6.000000	512.329200

```
[189]: # Select only the columns from DataFrame 'df' that have data type 'object'
↳(usually strings)
# and store them in a new DataFrame 'df1'
df1 = df[df.dtypes[df.dtypes == 'object'].index]
```

```
[190]: # Statistical summary of numeric columns
df1.describe()
```

```
[190]:
```

	Name	Sex	Ticket	Cabin	Embarked
count	891	891	891	204	889
unique	891	2	681	147	3
top	Braund, Mr. Owen Harris	male	347082	B96 B98	S
freq	1	577	7	4	644

```
[191]: #Printing object data types which we stored in data frame1
df1.dtypes
```

```
[191]: Name      object
Sex        object
```

```
Ticket      object
Cabin       object
Embarked    object
dtype: object
```

```
[192]: #Printing object data types which we stored in data frame1
df2 = df[df.dtypes[df.dtypes == 'int64'].index]
```

```
[193]: df2.dtypes
```

```
[193]: PassengerId    int64
Survived            int64
Pclass              int64
SibSp               int64
Parch               int64
dtype: object
```

```
[194]: df
```

```
[194]:
```

	PassengerId	Survived	Pclass	\
0	1	0	3	
1	2	1	1	
2	3	1	3	
3	4	1	1	
4	5	0	3	
..	...	...	...	
886	887	0	2	
887	888	1	1	
888	889	0	3	
889	890	1	1	
890	891	0	3	

	Name	Sex	Age	SibSp	\
0	Braund, Mr. Owen Harris	male	22.0	1	
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	
2	Heikkinen, Miss. Laina	female	26.0	0	
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	
4	Allen, Mr. William Henry	male	35.0	0	
..	...	...	...	...	
886	Montvila, Rev. Juozas	male	27.0	0	
887	Graham, Miss. Margaret Edith	female	19.0	0	
888	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	
889	Behr, Mr. Karl Howell	male	26.0	0	
890	Dooley, Mr. Patrick	male	32.0	0	

	Parch	Ticket	Fare	Cabin	Embarked
0	0	A/5 21171	7.2500	NaN	S

1	0	PC	17599	71.2833	C85	C
2	0	STON/O2.	3101282	7.9250	NaN	S
3	0		113803	53.1000	C123	S
4	0		373450	8.0500	NaN	S
..	...		...	...	...	
886	0		211536	13.0000	NaN	S
887	0		112053	30.0000	B42	S
888	2	W./C.	6607	23.4500	NaN	S
889	0		111369	30.0000	C148	C
890	0		370376	7.7500	NaN	Q

[891 rows x 12 columns]

```
[195]: # Check which columns in DataFrame 'df' have the data type 'float64'
df.dtypes=='float64'
```

```
[195]: PassengerId    False
Survived          False
Pclass            False
Name              False
Sex               False
Age               True
SibSp             False
Parch            False
Ticket            False
Fare              True
Cabin             False
Embarked          False
dtype: bool
```

```
[196]: # Select and return only the columns from DataFrame 'df' that have data type
↳ 'float64'
df[df.dtypes[df.dtypes=='float64'].index]
```

```
[196]:      Age      Fare
0    22.0    7.2500
1    38.0   71.2833
2    26.0    7.9250
3    35.0   53.1000
4    35.0    8.0500
..    ...    ...
886  27.0   13.0000
887  19.0   30.0000
888   NaN   23.4500
889  26.0   30.0000
890  32.0    7.7500
```

[891 rows x 2 columns]

```
[197]: #printing age and Fare column
df[["Age", "Fare"]]
```

```
[197]:
```

	Age	Fare
0	22.0	7.2500
1	38.0	71.2833
2	26.0	7.9250
3	35.0	53.1000
4	35.0	8.0500
..	...	...
886	27.0	13.0000
887	19.0	30.0000
888	NaN	23.4500
889	26.0	30.0000
890	32.0	7.7500

[891 rows x 2 columns]

```
[198]: #Slice the 'Name' column from index 5 to 9 (step size of 2 - i.e., rows 5, 7,
      ↪and 9)
df['Name'][5:10:2]
```

```
[198]: 5          Moran, Mr. James
      7      Palsson, Master. Gosta Leonard
      9  Nasser, Mrs. Nicholas (Adele Achem)
      Name: Name, dtype: object
```

```
[199]: # Slice the 'Name' column from index 5 to 9 (rows 5, 6, 7, 8, 9)
df['Name'][5:10]
```

```
[199]: 5          Moran, Mr. James
      6      McCarthy, Mr. Timothy J
      7      Palsson, Master. Gosta Leonard
      8  Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)
      9      Nasser, Mrs. Nicholas (Adele Achem)
      Name: Name, dtype: object
```

```
[200]: #Slice rows 5 to 15 (inclusive of 5, exclusive of 16) from the 'Ticket' column
      ↪(as a DataFrame)
df[['Ticket']][5:16]
```

```
[200]:
```

	Ticket
5	330877
6	17463
7	349909

```

8      347742
9      237736
10     PP 9549
11     113783
12    A/5. 2151
13     347082
14     350406
15     248706

```

```

[201]: # Loop through all column names in the DataFrame and print each column as a
        ↪ DataFrame
for i in df.columns :
    print(df[[i]])

```

```

    PassengerId
0              1
1              2
2              3
3              4
4              5
..           ...
886           887
887           888
888           889
889           890
890           891

```

[891 rows x 1 columns]

```

    Survived
0           0
1           1
2           1
3           1
4           0
..         ...
886         0
887         1
888         0
889         1
890         0

```

[891 rows x 1 columns]

```

    Pclass
0         3
1         1
2         3
3         1
4         3

```

```

..      ...
886      2
887      1
888      3
889      1
890      3

```

[891 rows x 1 columns]

```

                                     Name
0          Braund, Mr. Owen Harris
1  Cumings, Mrs. John Bradley (Florence Briggs Th...
2          Heikkinen, Miss. Laina
3  Futrelle, Mrs. Jacques Heath (Lily May Peel)
4          Allen, Mr. William Henry
..      ...
886          Montvila, Rev. Juozas
887          Graham, Miss. Margaret Edith
888  Johnston, Miss. Catherine Helen "Carrie"
889          Behr, Mr. Karl Howell
890          Dooley, Mr. Patrick

```

[891 rows x 1 columns]

```

      Sex
0    male
1  female
2  female
3  female
4    male
..      ...
886    male
887  female
888  female
889    male
890    male

```

[891 rows x 1 columns]

```

      Age
0    22.0
1    38.0
2    26.0
3    35.0
4    35.0
..      ...
886    27.0
887    19.0
888    NaN
889    26.0
890    32.0

```



[891 rows x 1 columns]

	SibSp
0	1
1	1
2	0
3	1
4	0
..	...
886	0
887	0
888	1
889	0
890	0

[891 rows x 1 columns]

	Parch
0	0
1	0
2	0
3	0
4	0
..	...
886	0
887	0
888	2
889	0
890	0

[891 rows x 1 columns]

	Ticket
0	A/5 21171
1	PC 17599
2	STON/O2. 3101282
3	113803
4	373450
..	...
886	211536
887	112053
888	W./C. 6607
889	111369
890	370376

[891 rows x 1 columns]

	Fare
0	7.2500
1	71.2833
2	7.9250

```

3      53.1000
4       8.0500
..      ...
886    13.0000
887    30.0000
888    23.4500
889    30.0000
890     7.7500

```

[891 rows x 1 columns]

```

Cabin
0      NaN
1      C85
2      NaN
3     C123
4      NaN
..      ...
886     NaN
887     B42
888     NaN
889    C148
890     NaN

```

[891 rows x 1 columns]

```

Embarked
0         S
1         C
2         S
3         S
4         S
..      ...
886        S
887        S
888        S
889        C
890        Q

```

[891 rows x 1 columns]

```

[202]: #sorting the tickets
sorted(df['Ticket'])

```

```

[202]: ['110152',
        '110152',
        '110152',
        '110413',
        '110413',

```

'110413',  
'110465',  
'110465',  
'110564',  
'110813',  
'111240',  
'111320',  
'111361',  
'111361',  
'111369',  
'111426',  
'111427',  
'111428',  
'112050',  
'112052',  
'112053',  
'112058',  
'112059',  
'112277',  
'112379',  
'113028',  
'113043',  
'113050',  
'113051',  
'113055',  
'113056',  
'113059',  
'113501',  
'113503',  
'113505',  
'113505',  
'113509',  
'113510',  
'113514',  
'113572',  
'113572',  
'113760',  
'113760',  
'113760',  
'113760',  
'113767',  
'113773',  
'113776',  
'113776',  
'113781',  
'113781',  
'113781',

'113781',  
'113783',  
'113784',  
'113786',  
'113787',  
'113788',  
'113789',  
'113789',  
'113792',  
'113794',  
'113796',  
'113798',  
'113798',  
'113800',  
'113803',  
'113803',  
'113804',  
'113806',  
'113806',  
'113807',  
'11668',  
'11668',  
'11751',  
'11751',  
'11752',  
'11753',  
'11755',  
'11765',  
'11767',  
'11767',  
'11769',  
'11771',  
'11774',  
'11813',  
'11967',  
'11967',  
'12233',  
'12460',  
'12749',  
'12749',  
'13049',  
'13213',  
'13214',  
'13502',  
'13502',  
'13502',  
'13507',

'13507',  
'13509',  
'13567',  
'13568',  
'14311',  
'14312',  
'14313',  
'14973',  
'1601',  
'1601',  
'1601',  
'1601',  
'1601',  
'1601',  
'1601',  
'1601',  
'16966',  
'16966',  
'16988',  
'17421',  
'17421',  
'17421',  
'17421',  
'17453',  
'17453',  
'17463',  
'17464',  
'17465',  
'17466',  
'17474',  
'17474',  
'17764',  
'19877',  
'19877',  
'19928',  
'19928',  
'19943',  
'19943',  
'19947',  
'19950',  
'19950',  
'19950',  
'19950',  
'19952',  
'19972',  
'19988',  
'19996',  
'19996',

'2003',  
'211536',  
'21440',  
'218629',  
'219533',  
'220367',  
'220845',  
'220845',  
'2223',  
'223596',  
'226593',  
'226875',  
'228414',  
'229236',  
'230080',  
'230080',  
'230080',  
'230136',  
'230136',  
'230433',  
'230433',  
'230434',  
'231919',  
'231919',  
'231945',  
'233639',  
'233866',  
'234360',  
'234604',  
'234686',  
'234818',  
'236171',  
'236852',  
'236853',  
'237442',  
'237565',  
'237668',  
'237671',  
'237736',  
'237736',  
'237789',  
'237798',  
'239853',  
'239853',  
'239853',  
'239854',  
'239855',

'239856',  
'239865',  
'239865',  
'240929',  
'24160',  
'24160',  
'24160',  
'243847',  
'243847',  
'243880',  
'244252',  
'244252',  
'244270',  
'244278',  
'244310',  
'244358',  
'244361',  
'244367',  
'244367',  
'244373',  
'248698',  
'248706',  
'248723',  
'248727',  
'248727',  
'248727',  
'248731',  
'248733',  
'248738',  
'248738',  
'248740',  
'248747',  
'250643',  
'250644',  
'250644',  
'250646',  
'250647',  
'250647',  
'250648',  
'250649',  
'250649',  
'250651',  
'250652',  
'250653',  
'250655',  
'250655',  
'2620',

'2623',  
'2624',  
'2625',  
'2626',  
'2627',  
'2627',  
'2628',  
'2629',  
'2631',  
'26360',  
'26360',  
'2641',  
'2647',  
'2648',  
'2649',  
'2650',  
'2651',  
'2651',  
'2653',  
'2653',  
'2659',  
'2659',  
'2661',  
'2661',  
'2662',  
'2663',  
'2664',  
'2665',  
'2665',  
'2666',  
'2666',  
'2666',  
'2666',  
'2667',  
'2668',  
'2668',  
'2669',  
'26707',  
'2671',  
'2672',  
'2674',  
'2677',  
'2678',  
'2678',  
'2680',  
'2683',  
'2685',



'2686',  
'2687',  
'2689',  
'2690',  
'2691',  
'2691',  
'2693',  
'2694',  
'2695',  
'2697',  
'2699',  
'2699',  
'2700',  
'27042',  
'27267',  
'27849',  
'28134',  
'28206',  
'28213',  
'28220',  
'28228',  
'28403',  
'28403',  
'28424',  
'28425',  
'28551',  
'28664',  
'28665',  
'29011',  
'2908',  
'2908',  
'29103',  
'29104',  
'29105',  
'29106',  
'29106',  
'29106',  
'29108',  
'2926',  
'29750',  
'29750',  
'29751',  
'3101264',  
'3101265',  
'3101267',  
'3101276',  
'3101277',

'3101278',  
'3101278',  
'3101281',  
'3101295',  
'3101295',  
'3101295',  
'3101295',  
'3101295',  
'3101295',  
'3101296',  
'3101298',  
'31027',  
'31027',  
'31028',  
'312991',  
'312992',  
'312993',  
'31418',  
'315037',  
'315082',  
'315084',  
'315086',  
'315088',  
'315089',  
'315090',  
'315093',  
'315094',  
'315096',  
'315097',  
'315098',  
'315151',  
'315153',  
'323592',  
'323951',  
'324669',  
'330877',  
'330909',  
'330919',  
'330923',  
'330931',  
'330932',  
'330935',  
'330958',  
'330959',  
'330979',  
'330980',  
'334912',

'335097',  
'335677',  
'33638',  
'336439',  
'3411',  
'341826',  
'34218',  
'342826',  
'343095',  
'343120',  
'343275',  
'343276',  
'345364',  
'345572',  
'345763',  
'345764',  
'345764',  
'345765',  
'345767',  
'345769',  
'345770',  
'345773',  
'345773',  
'345773',  
'345774',  
'345777',  
'345778',  
'345779',  
'345780',  
'345781',  
'345783',  
'3460',  
'347054',  
'347054',  
'347060',  
'347061',  
'347062',  
'347063',  
'347064',  
'347067',  
'347068',  
'347069',  
'347071',  
'347073',  
'347074',  
'347076',  
'347077',

'347077',  
'347077',  
'347077',  
'347078',  
'347080',  
'347080',  
'347081',  
'347082',  
'347082',  
'347082',  
'347082',  
'347082',  
'347082',  
'347082',  
'347083',  
'347085',  
'347087',  
'347088',  
'347088',  
'347088',  
'347088',  
'347088',  
'347088',  
'347088',  
'347089',  
'3474',  
'347464',  
'347466',  
'347468',  
'347470',  
'347742',  
'347742',  
'347742',  
'347743',  
'348121',  
'348123',  
'348124',  
'349201',  
'349203',  
'349204',  
'349205',  
'349206',  
'349207',  
'349208',  
'349209',  
'349210',  
'349212',  
'349213',

'349214',  
'349215',  
'349216',  
'349217',  
'349218',  
'349219',  
'349221',  
'349222',  
'349223',  
'349224',  
'349225',  
'349227',  
'349228',  
'349231',  
'349233',  
'349234',  
'349236',  
'349237',  
'349237',  
'349239',  
'349240',  
'349241',  
'349242',  
'349243',  
'349244',  
'349245',  
'349246',  
'349247',  
'349248',  
'349249',  
'349251',  
'349252',  
'349253',  
'349254',  
'349256',  
'349257',  
'349909',  
'349909',  
'349909',  
'349909',  
'349910',  
'349912',  
'350025',  
'350026',  
'350029',  
'350034',  
'350035',

'350036',  
'350042',  
'350043',  
'350046',  
'350047',  
'350048',  
'350050',  
'350052',  
'350060',  
'350404',  
'350406',  
'350407',  
'350417',  
'35273',  
'35273',  
'35273',  
'35281',  
'35281',  
'35851',  
'35852',  
'358585',  
'358585',  
'36209',  
'362316',  
'363291',  
'363291',  
'363291',  
'363294',  
'363592',  
'364498',  
'364499',  
'364500',  
'364506',  
'364511',  
'364512',  
'364516',  
'364516',  
'364846',  
'364848',  
'364849',  
'364849',  
'364850',  
'364851',  
'365222',  
'365226',  
'36568',  
'367226',

'367226',  
'367228',  
'367229',  
'367230',  
'367230',  
'367231',  
'367232',  
'367655',  
'368323',  
'36864',  
'36865',  
'36866',  
'368703',  
'36928',  
'36928',  
'36947',  
'36947',  
'36963',  
'36967',  
'36973',  
'36973',  
'370129',  
'370129',  
'370365',  
'370365',  
'370369',  
'370370',  
'370371',  
'370372',  
'370373',  
'370375',  
'370376',  
'370377',  
'371060',  
'371110',  
'371110',  
'371110',  
'371362',  
'372622',  
'373450',  
'374746',  
'374887',  
'374910',  
'376564',  
'376564',  
'376566',  
'382649',

'382651',  
'382652',  
'382652',  
'382652',  
'382652',  
'382652',  
'383121',  
'384461',  
'386525',  
'392091',  
'392092',  
'392096',  
'392096',  
'394140',  
'4133',  
'4133',  
'4133',  
'4133',  
'4134',  
'4135',  
'4136',  
'4137',  
'4138',  
'4579',  
'54636',  
'54636',  
'5727',  
'65303',  
'65304',  
'65306',  
'6563',  
'693',  
'695',  
'7267',  
'7534',  
'7534',  
'7540',  
'7545',  
'7546',  
'7552',  
'7553',  
'7598',  
'8471',  
'8475',  
'9234',  
'A./5. 2152',  
'A./5. 3235',



'A.5. 11206',  
'A.5. 18509',  
'A/4 45380',  
'A/4 48871',  
'A/4 48871',  
'A/4. 20589',  
'A/4. 34244',  
'A/4. 39886',  
'A/5 21171',  
'A/5 21172',  
'A/5 21173',  
'A/5 21174',  
'A/5 2466',  
'A/5 2817',  
'A/5 3536',  
'A/5 3540',  
'A/5 3594',  
'A/5 3902',  
'A/5. 10482',  
'A/5. 13032',  
'A/5. 2151',  
'A/5. 3336',  
'A/5. 3336',  
'A/5. 3337',  
'A/5. 851',  
'A/S 2816',  
'A4. 54510',  
'C 17369',  
'C 4001',  
'C 7075',  
'C 7076',  
'C 7077',  
'C.A. 17248',  
'C.A. 18723',  
'C.A. 2315',  
'C.A. 2315',  
'C.A. 24579',  
'C.A. 24580',  
'C.A. 2673',  
'C.A. 2673',  
'C.A. 29178',  
'C.A. 29395',  
'C.A. 29566',  
'C.A. 31026',  
'C.A. 31921',  
'C.A. 31921',  
'C.A. 31921',

'C.A. 33111',  
'C.A. 33112',  
'C.A. 33112',  
'C.A. 33595',  
'C.A. 34260',  
'C.A. 34651',  
'C.A. 34651',  
'C.A. 34651',  
'C.A. 37671',  
'C.A. 37671',  
'C.A. 5547',  
'C.A. 6212',  
'C.A./SOTON 34068',  
'CA 2144',  
'CA 2144',  
'CA 2144',  
'CA 2144',  
'CA 2144',  
'CA 2144',  
'CA. 2314',  
'CA. 2343',  
'CA. 2343',  
'CA. 2343',  
'CA. 2343',  
'CA. 2343',  
'CA. 2343',  
'CA. 2343',  
'CA. 2343',  
'F.C. 12750',  
'F.C.C. 13528',  
'F.C.C. 13529',  
'F.C.C. 13529',  
'F.C.C. 13529',  
'F.C.C. 13531',  
'Fa 265302',  
'LINE',  
'LINE',  
'LINE',  
'LINE',  
'P/PP 3381',  
'P/PP 3381',  
'PC 17318',  
'PC 17473',  
'PC 17474',  
'PC 17475',  
'PC 17476',  
'PC 17477',  
'PC 17477',

'PC 17482',  
'PC 17483',  
'PC 17485',  
'PC 17485',  
'PC 17558',  
'PC 17558',  
'PC 17569',  
'PC 17569',  
'PC 17572',  
'PC 17572',  
'PC 17572',  
'PC 17582',  
'PC 17582',  
'PC 17582',  
'PC 17585',  
'PC 17590',  
'PC 17592',  
'PC 17593',  
'PC 17593',  
'PC 17595',  
'PC 17596',  
'PC 17597',  
'PC 17599',  
'PC 17600',  
'PC 17601',  
'PC 17603',  
'PC 17604',  
'PC 17604',  
'PC 17605',  
'PC 17608',  
'PC 17608',  
'PC 17609',  
'PC 17610',  
'PC 17611',  
'PC 17611',  
'PC 17612',  
'PC 17754',  
'PC 17755',  
'PC 17755',  
'PC 17755',  
'PC 17756',  
'PC 17757',  
'PC 17757',  
'PC 17757',  
'PC 17757',  
'PC 17758',  
'PC 17758',

'PC 17759',  
'PC 17760',  
'PC 17760',  
'PC 17760',  
'PC 17761',  
'PC 17761',  
'PP 4348',  
'PP 9549',  
'PP 9549',  
'S.C./A.4. 23567',  
'S.C./PARIS 2079',  
'S.C./PARIS 2079',  
'S.O./P.P. 3',  
'S.O./P.P. 3',  
'S.O./P.P. 751',  
'S.O.C. 14879',  
'S.O.C. 14879',  
'S.O.C. 14879',  
'S.O.C. 14879',  
'S.O.C. 14879',  
'S.O.P. 1166',  
'S.P. 3464',  
'S.W./PP 752',  
'SC 1748',  
'SC/AH 29037',  
'SC/AH 3085',  
'SC/AH Basle 541',  
'SC/PARIS 2131',  
'SC/PARIS 2133',  
'SC/PARIS 2146',  
'SC/PARIS 2149',  
'SC/PARIS 2167',  
'SC/Paris 2123',  
'SC/Paris 2123',  
'SC/Paris 2123',  
'SC/Paris 2163',  
'SCO/W 1585',  
'SO/C 14885',  
'SOTON/O.Q. 3101305',  
'SOTON/O.Q. 3101306',  
'SOTON/O.Q. 3101307',  
'SOTON/O.Q. 3101310',  
'SOTON/O.Q. 3101311',  
'SOTON/O.Q. 3101312',  
'SOTON/O.Q. 392078',  
'SOTON/O.Q. 392087',  
'SOTON/O2 3101272',

```

'SOTON/O2 3101287',
'SOTON/OQ 3101316',
'SOTON/OQ 3101317',
'SOTON/OQ 392076',
'SOTON/OQ 392082',
'SOTON/OQ 392086',
'SOTON/OQ 392089',
'SOTON/OQ 392090',
'STON/O 2. 3101269',
'STON/O 2. 3101273',
'STON/O 2. 3101274',
'STON/O 2. 3101275',
'STON/O 2. 3101280',
'STON/O 2. 3101285',
'STON/O 2. 3101286',
'STON/O 2. 3101288',
'STON/O 2. 3101289',
'STON/O 2. 3101292',
'STON/O 2. 3101293',
'STON/O 2. 3101294',
'STON/O2. 3101271',
'STON/O2. 3101279',
'STON/O2. 3101279',
'STON/O2. 3101282',
'STON/O2. 3101283',
'STON/O2. 3101290',
'SW/PP 751',
'W./C. 14258',
'W./C. 14263',
'W./C. 6607',
'W./C. 6607',
'W./C. 6608',
'W./C. 6608',
'W./C. 6608',
'W./C. 6608',
'W./C. 6609',
'W.E.P. 5734',
'W/C 14208',
'WE/P 5735',
'WE/P 5735']

```

```

[218]: # Convert the 'Sex' column into a categorical data type (useful for memory
      ↪ efficiency and encoding)
pd.Categorical(df['Sex'])

```

```

[218]: ['male', 'female', 'female', 'female', 'male', ..., 'male', 'female', 'female',
      'male', 'male']

```

```
Length: 891
Categories (2, object): ['female', 'male']
```

```
[219]: #Add a new column 'isha' to the DataFrame and assign a default value 'myrecord'
      ↪to all rows
      df['isha']="myrecord"
```

```
[220]: # Overwrite the 'isha' column with the values from the 'PassengerId' column
      df['isha']=df['PassengerId']
```

```
[206]: # Convert the 'Pclass' column into a categorical data type
      pd.Categorical(df['Pclass'])
```

```
[206]: [3, 1, 3, 1, 3, ..., 2, 1, 3, 1, 3]
      Length: 891
      Categories (3, int64): [1, 2, 3]
```

```
[221]: #Get unique values
      df['Pclass'].unique()
```

```
[221]: array([3, 1, 2], dtype=int64)
```

```
[222]: # Check for missing values
      df['Pclass'].isnull()
```

```
[222]: 0      False
      1      False
      2      False
      3      False
      4      False
      ...
      886    False
      887    False
      888    False
      889    False
      890    False
      Name: Pclass, Length: 891, dtype: bool
```

```
[223]: # Check for missing values
      df['Cabin'].isnull()
```

```
[223]: 0      True
      1     False
      2      True
      3     False
      4      True
      ...
      886    True
```

```

887     False
888     True
889     False
890     True
Name: Cabin, Length: 891, dtype: bool

```

```

[228]: #Create a new column 'new_col' by adding values from 'PassengerId' and 'Age'
      ↪ columns
df['new_col'] = df['PassengerId'] + df['Age']

```

```

[229]: # Check which rows in the 'Fare' column have the maximum fare value
df['Fare']==max(df['Fare'])

```

```

[229]: 0     False
      1     False
      2     False
      3     False
      4     False
      ...
      886    False
      887    False
      888    False
      889    False
      890    False
Name: Fare, Length: 891, dtype: bool

```

```

[230]: # Filter and display the row(s) where the 'Fare' is equal to the maximum fare
      ↪ value in the DataFrame
df[df['Fare']==max(df['Fare'])]

```

```

[230]:      PassengerId  Survived  Pclass                                Name \
258             259         1        1                      Ward, Miss. Anna
679             680         1        1  Cardeza, Mr. Thomas Drake Martinez
737             738         1        1      Lesurer, Mr. Gustave J

      Sex  Age  SibSp  Parch    Ticket     Fare      Cabin Embarked \
258  female  35.0    0     0  PC 17755  512.3292      NaN        C
679   male  36.0    0     1  PC 17755  512.3292  B51 B53 B55        C
737   male  35.0    0     0  PC 17755  512.3292    B101        C

      isha  new_col
258    259    294.0
679    680    716.0
737    738    773.0

```

```

[231]: # Import NumPy for numerical operations (arrays, math functions, etc.)
import numpy as np

```

```
# Import the os module to interact with the operating system (e.g.,
↳ directories, files)
import os

# Import matplotlib's pyplot module for data visualization (plots, charts,
↳ graphs)
import matplotlib.pyplot as plt

# Enable inline plotting in Jupyter Notebook (so plots appear directly below
↳ the code cells)
%matplotlib inline
```

```
[233]: # Read the dataset into a DataFrame
titanic_train = pd.read_csv("https://raw.githubusercontent.com/datasciencedojo/
↳ datasets/master/titanic.csv")
```

```
[234]: #print dataset
titanic_train
```

```
[234]:
```

	PassengerId	Survived	Pclass	\
0	1	0	3	
1	2	1	1	
2	3	1	3	
3	4	1	1	
4	5	0	3	
..	...	...	...	
886	887	0	2	
887	888	1	1	
888	889	0	3	
889	890	1	1	
890	891	0	3	

	Name	Sex	Age	SibSp	\
0	Braund, Mr. Owen Harris	male	22.0	1	
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	
2	Heikkinen, Miss. Laina	female	26.0	0	
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	
4	Allen, Mr. William Henry	male	35.0	0	
..	...	...	...	...	
886	Montvila, Rev. Juozas	male	27.0	0	
887	Graham, Miss. Margaret Edith	female	19.0	0	
888	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	
889	Behr, Mr. Karl Howell	male	26.0	0	
890	Dooley, Mr. Patrick	male	32.0	0	

Parch	Ticket	Fare	Cabin	Embarked
-------	--------	------	-------	----------



0	0	A/5	21171	7.2500	NaN	S
1	0	PC	17599	71.2833	C85	C
2	0	STON/O2.	3101282	7.9250	NaN	S
3	0		113803	53.1000	C123	S
4	0		373450	8.0500	NaN	S
..	...		...	...	...	
886	0		211536	13.0000	NaN	S
887	0		112053	30.0000	B42	S
888	2	W./C.	6607	23.4500	NaN	S
889	0		111369	30.0000	C148	C
890	0		370376	7.7500	NaN	Q

[891 rows x 12 columns]

```
[235]: # Printing last 10 records
titanic_train.tail(10)
```

```
[235]:      PassengerId  Survived  Pclass                                Name \
881           882         0         3                        Markun, Mr. Johann
882           883         0         3          Dahlberg, Miss. Gerda Ulrika
883           884         0         2      Banfield, Mr. Frederick James
884           885         0         3          Sutehall, Mr. Henry Jr
885           886         0         3      Rice, Mrs. William (Margaret Norton)
886           887         0         2          Montvila, Rev. Juozas
887           888         1         1          Graham, Miss. Margaret Edith
888           889         0         3  Johnston, Miss. Catherine Helen "Carrie"
889           890         1         1          Behr, Mr. Karl Howell
890           891         0         3          Dooley, Mr. Patrick
```

	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
881	male	33.0	0	0	349257	7.8958	NaN	S
882	female	22.0	0	0	7552	10.5167	NaN	S
883	male	28.0	0	0	C.A./SOTON 34068	10.5000	NaN	S
884	male	25.0	0	0	SOTON/OQ 392076	7.0500	NaN	S
885	female	39.0	0	5	382652	29.1250	NaN	Q
886	male	27.0	0	0	211536	13.0000	NaN	S
887	female	19.0	0	0	112053	30.0000	B42	S
888	female	NaN	1	2	W./C. 6607	23.4500	NaN	S
889	male	26.0	0	0	111369	30.0000	C148	C
890	male	32.0	0	0	370376	7.7500	NaN	Q

```
[236]: #Printing data types
titanic_train.dtypes
```

```
[236]: PassengerId      int64
Survived            int64
Pclass              int64
```

```

Name          object
Sex           object
Age          float64
SibSp         int64
Parch         int64
Ticket        object
Fare          float64
Cabin         object
Embarked      object
dtype: object

```

```

[89]: #where the data type is object that will be true
      titanic_train.dtypes=="object"

```

```

[89]: PassengerId    False
      Survived      False
      Pclass        False
      Name          True
      Sex           True
      Age          False
      SibSp         False
      Parch         False
      Ticket        True
      Fare          False
      Cabin         True
      Embarked      True
      dtype: bool

```

```

[237]: #selecting the columns with object data type
      titanic_train.dtypes[titanic_train.dtypes=="object"]

```

```

[237]: Name          object
      Sex           object
      Ticket        object
      Cabin         object
      Embarked      object
      dtype: object

```

```

[240]: # Get the names of all columns in the DataFrame 'titanic_train' that have data
      ↪type 'object' (usually strings)
      titanic_train.dtypes[titanic_train.dtypes=="object"].index

```

```

[240]: Index(['Name', 'Sex', 'Ticket', 'Cabin', 'Embarked'], dtype='object')

```

```

[239]: # Get the names of all columns in 'titanic_train' that have data type 'object'
      ↪(usually strings)
      a = titanic_train.dtypes[titanic_train.dtypes == "object"].index

```

```
#Display the list of object-type column names
a
# Display statistical summary (count, unique, top, freq) for all object-type
↳(categorical/text) columns
titanic_train[a].describe()
```

```
[239]:
```

	Name	Sex	Ticket	Cabin	Embarked
count	891	891	891	204	889
unique	891	2	681	147	3
top	Braund, Mr. Owen Harris	male	347082	B96 B98	S
freq	1	577	7	4	644

```
[241]: # Count of unique values
titanic_train.Sex.value_counts()
```

```
[241]: Sex
male      577
female    314
Name: count, dtype: int64
```

```
[242]: # Statistical summary of numeric columns
titanic_train.describe()
```

```
[242]:
```

	PassengerId	Survived	Pclass	Age	SibSp \
count	891.000000	891.000000	891.000000	714.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008
std	257.353842	0.486592	0.836071	14.526497	1.102743
min	1.000000	0.000000	1.000000	0.420000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000
50%	446.000000	0.000000	3.000000	28.000000	0.000000
75%	668.500000	1.000000	3.000000	38.000000	1.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000

  

	Parch	Fare
count	891.000000	891.000000
mean	0.381594	32.204208
std	0.806057	49.693429
min	0.000000	0.000000
25%	0.000000	7.910400
50%	0.000000	14.454200
75%	0.000000	31.000000
max	6.000000	512.329200

```
[243]: # Get the data types of columns where the dtype is 'object' (i.e., text or
↳categorical columns)
titanic_train.dtypes[list(titanic_train.dtypes=="object")].values
```

```
[243]: array([dtype('O'), dtype('O'), dtype('O'), dtype('O'), dtype('O')],
      dtype=object)
```

```
[244]: #List all column names in the DataFrame
      titanic_train.columns
```

```
[244]: Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',
      'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],
      dtype='object')
```

```
[245]: #Select and display the 'Name', 'Sex', and 'Age' columns from the
      ↪ 'titanic_train' DataFrame
      titanic_train[['Name', 'Sex', 'Age']]
```

```
[245]:
```

	Name	Sex	Age
0	Braund, Mr. Owen Harris	male	22.0
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0
2	Heikkinen, Miss. Laina	female	26.0
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0
4	Allen, Mr. William Henry	male	35.0
..	...	...	...
886	Montvila, Rev. Juozas	male	27.0
887	Graham, Miss. Margaret Edith	female	19.0
888	Johnston, Miss. Catherine Helen "Carrie"	female	NaN
889	Behr, Mr. Karl Howell	male	26.0
890	Dooley, Mr. Patrick	male	32.0

[891 rows x 3 columns]

```
[246]: # Define a list of required column names (as strings) from the Titanic dataset
      required_column_list = ['PassengerId', 'Survived', 'Pclass']

      # Display the list
      required_column_list
```

```
[246]: ['PassengerId', 'Survived', 'Pclass']
```

```
[247]: titanic_train['Name']
```

```
[247]:
```

0	Braund, Mr. Owen Harris
1	Cumings, Mrs. John Bradley (Florence Briggs Th...
2	Heikkinen, Miss. Laina
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)
4	Allen, Mr. William Henry
...	...
886	Montvila, Rev. Juozas
887	Graham, Miss. Margaret Edith
888	Johnston, Miss. Catherine Helen "Carrie"

```

889                                Behr, Mr. Karl Howell
890                                Dooley, Mr. Patrick
Name: Name, Length: 891, dtype: object

```

```

[249]: #printing the column with name
titanic_train[['Name']]

```

```

[249]:
                                Name
0                Braund, Mr. Owen Harris
1    Cumings, Mrs. John Bradley (Florence Briggs Th...
2                Heikkinen, Miss. Laina
3        Futrelle, Mrs. Jacques Heath (Lily May Peel)
4                Allen, Mr. William Henry
..
886                Montvila, Rev. Juozas
887                Graham, Miss. Margaret Edith
888        Johnston, Miss. Catherine Helen "Carrie"
889                Behr, Mr. Karl Howell
890                Dooley, Mr. Patrick

[891 rows x 1 columns]

```

```

[250]: #Sort the 'Name' column alphabetically and select every 2nd name from index 5
↳to 9 (i.e., 5 and 7)
sorted(titanic_train["Name"])[5:10:2]

```

```

[250]: ['Adahl, Mr. Mauritz Nils Martin',
'Ahlin, Mrs. Johan (Johanna Persdotter Larsson)',
'Albimona, Mr. Nassef Cassem']

```

```

[251]: # Sort the entire DataFrame by the 'Name' column (in-place is optional)
titanic_train.sort_values(by='Name')

```

```

[251]:
   PassengerId  Survived  Pclass                                Name \
845          846         0      3                Abbing, Mr. Anthony
746          747         0      3            Abbott, Mr. Rossmore Edward
279          280         1      3        Abbott, Mrs. Stanton (Rosa Hunt)
308          309         0      2                Abelson, Mr. Samuel
874          875         1      2    Abelson, Mrs. Samuel (Hannah Wizosky)
..          ...         ...      ...
286          287         1      3            de Mulder, Mr. Theodore
282          283         0      3            de Pelsmaecker, Mr. Alfons
361          362         0      2            del Carlo, Mr. Sebastiano
153          154         0      3    van Billiard, Mr. Austin Blyler
868          869         0      3            van Melkebeke, Mr. Philemon

   Sex  Age  SibSp  Parch              Ticket     Fare Cabin Embarked

```

845	male	42.0	0	0	C.A. 5547	7.5500	NaN	S
746	male	16.0	1	1	C.A. 2673	20.2500	NaN	S
279	female	35.0	1	1	C.A. 2673	20.2500	NaN	S
308	male	30.0	1	0	P/PP 3381	24.0000	NaN	C
874	female	28.0	1	0	P/PP 3381	24.0000	NaN	C
..	...	...	...	...	...	...	...	...
286	male	30.0	0	0	345774	9.5000	NaN	S
282	male	16.0	0	0	345778	9.5000	NaN	S
361	male	29.0	1	0	SC/PARIS 2167	27.7208	NaN	C
153	male	40.5	0	2	A/5. 851	14.5000	NaN	S
868	male	NaN	0	0	345777	9.5000	NaN	S

[891 rows x 12 columns]

```
[252]: #Get the first 10 names from the 'Name' column after sorting them
↳alphabetically (A-Z)
sorted(titanic_train["Name"])[10]
```

```
[252]: ['Abbing, Mr. Anthony',
'Abbott, Mr. Rossmore Edward',
'Abbott, Mrs. Stanton (Rosa Hunt)',
'Abelson, Mr. Samuel',
'Abelson, Mrs. Samuel (Hannah Wizesky)',
'Adahl, Mr. Mauritz Nils Martin',
'Adams, Mr. John',
'Ahlin, Mrs. Johan (Johanna Persdotter Larsson)',
'Aks, Mrs. Sam (Leah Rosen)',
'Albimona, Mr. Nassef Cassem']
```

```
[253]: # Get the first 50 names from the 'Name' column after sorting them in
↳alphabetical order (A-Z)
sorted(titanic_train["Name"])[50]
```

```
[253]: ['Abbing, Mr. Anthony',
'Abbott, Mr. Rossmore Edward',
'Abbott, Mrs. Stanton (Rosa Hunt)',
'Abelson, Mr. Samuel',
'Abelson, Mrs. Samuel (Hannah Wizesky)',
'Adahl, Mr. Mauritz Nils Martin',
'Adams, Mr. John',
'Ahlin, Mrs. Johan (Johanna Persdotter Larsson)',
'Aks, Mrs. Sam (Leah Rosen)',
'Albimona, Mr. Nassef Cassem',
'Alexander, Mr. William',
'Alhomaki, Mr. Ilmari Rudolf',
'Ali, Mr. Ahmed',
'Ali, Mr. William',
```

'Allen, Miss. Elisabeth Walton',  
 'Allen, Mr. William Henry',  
 'Allison, Master. Hudson Trevor',  
 'Allison, Miss. Helen Loraine',  
 'Allison, Mrs. Hudson J C (Bessie Waldo Daniels)',  
 'Allum, Mr. Owen George',  
 'Andersen-Jensen, Miss. Carla Christine Nielsine',  
 'Anderson, Mr. Harry',  
 'Andersson, Master. Sigvard Harald Elias',  
 'Andersson, Miss. Ebba Iris Alfrida',  
 'Andersson, Miss. Ellis Anna Maria',  
 'Andersson, Miss. Erna Alexandra',  
 'Andersson, Miss. Ingeborg Constanzia',  
 'Andersson, Miss. Sigrid Elisabeth',  
 'Andersson, Mr. Anders Johan',  
 'Andersson, Mr. August Edvard ("Wennerstrom")',  
 'Andersson, Mrs. Anders Johan (Alfrida Konstantia Brogren)',  
 'Andreasson, Mr. Paul Edvin',  
 'Andrew, Mr. Edgardo Samuel',  
 'Andrews, Miss. Kornelia Theodosia',  
 'Andrews, Mr. Thomas Jr',  
 'Angle, Mrs. William A (Florence "Mary" Agnes Hughes)',  
 'Appleton, Mrs. Edward Dale (Charlotte Lamson)',  
 'Arnold-Franchi, Mr. Josef',  
 'Arnold-Franchi, Mrs. Josef (Josefine Franchi)',  
 'Artagaveytia, Mr. Ramon',  
 'Asim, Mr. Adola',  
 'Asplund, Master. Clarence Gustaf Hugo',  
 'Asplund, Master. Edvin Rojj Felix',  
 'Asplund, Miss. Lillian Gertrud',  
 'Asplund, Mrs. Carl Oscar (Selma Augusta Emilia Johansson)',  
 'Astor, Mrs. John Jacob (Madeleine Talmadge Force)',  
 'Attalah, Miss. Malake',  
 'Attalah, Mr. Sleiman',  
 'Aubart, Mme. Leontine Pauline',  
 'Augustsson, Mr. Albert']

```
[255]: #Get the top 10 names from the 'Name' column sorted in reverse (Z-A)
      ↪ alphabetical order
sorted(titanic_train["Name"],reverse=True)[:10]
```

```
[255]: ['van Melkebeke, Mr. Philemon',
      'van Billiard, Mr. Austin Blyler',
      'del Carlo, Mr. Sebastiano',
      'de Pelsmaecker, Mr. Alfons',
      'de Mulder, Mr. Theodore',
      'de Messemaeker, Mrs. Guillaume Joseph (Emma)',
```

```
'Zimmerman, Mr. Leo',  
'Zabour, Miss. Thamine',  
'Zabour, Miss. Hileni',  
'Yrois, Miss. Henriette ("Mrs Harbeck")']
```

```
[256]: sorted(titanic_train["Name"],reverse=True)[:50]
```

```
[256]: ['van Melkebeke, Mr. Philemon',  
'van Billiard, Mr. Austin Blyler',  
'del Carlo, Mr. Sebastiano',  
'de Pelsmaecker, Mr. Alfons',  
'de Mulder, Mr. Theodore',  
'de Messemaeker, Mrs. Guillaume Joseph (Emma)',  
'Zimmerman, Mr. Leo',  
'Zabour, Miss. Thamine',  
'Zabour, Miss. Hileni',  
'Yrois, Miss. Henriette ("Mrs Harbeck")',  
'Yousseff, Mr. Gerious',  
'Yousif, Mr. Wazli',  
'Youseff, Mr. Gerious',  
'Young, Miss. Marie Grice',  
'Yasbeck, Mrs. Antoni (Selini Alexander)',  
'Yasbeck, Mr. Antoni',  
'Wright, Mr. George',  
'Woolner, Mr. Hugh',  
'Wiseman, Mr. Phillippe',  
'Windelov, Mr. Einar',  
'Williams-Lambert, Mr. Fletcher Fellows',  
'Williams, Mr. Leslie',  
'Williams, Mr. Howard Hugh "Harry"',  
'Williams, Mr. Charles Eugene',  
'Williams, Mr. Charles Duane',  
'Willey, Mr. Edward',  
'Wilhelms, Mr. Charles',  
'Wiklund, Mr. Jakob Alfred',  
'Widener, Mr. Harry Elkins',  
'Widegren, Mr. Carl/Charles Peter',  
'Wick, Mrs. George Dennick (Mary Hitchcock)',  
'Wick, Miss. Mary Natalie',  
'White, Mr. Richard Frasar',  
'White, Mr. Percival Wayland',  
'Wheadon, Mr. Edward H',  
'West, Mrs. Edwy Arthur (Ada Mary Worth)',  
'West, Mr. Edwy Arthur',  
'West, Miss. Constance Mirium',  
'Wells, Miss. Joan',  
'Weisz, Mrs. Leopold (Mathilde Francoise Pede)']
```



```
'Weir, Col. John',
'Webber, Mr. James',
'Webber, Miss. Susan',
'Watt, Mrs. James (Elizabeth "Bessie" Inglis Milne)',
'Watson, Mr. Ennis Hastings',
'Warren, Mrs. Frank Manley (Anna Sophia Atkinson)',
'Ward, Miss. Anna',
'Walker, Mr. William Anderson',
'Waelens, Mr. Achille',
'Vovk, Mr. Janko']
```

```
[260]: #Statistical summary of numeric columns
titanic_train["Name"].describe()
```

```
[260]: count          891
unique          891
top      Braund, Mr. Owen Harris
freq              1
Name: Name, dtype: object
```

```
[261]: #Select the 'Ticket' column as a DataFrame and display rows at index positions
↳0 and 2 (every 2nd row from the first 5)
titanic_train[["Ticket"]][:5:2]
```

```
[261]:      Ticket
0      A/5 21171
2  STON/O2. 3101282
4      373450
```

```
[262]: # Display the statistical summary of all columns, including both numeric and
↳non-numeric (categorical/object) types
titanic_train.describe(include="all")
```

```
[262]:
```

	PassengerId	Survived	Pclass	Name	Sex	\
count	891.000000	891.000000	891.000000	891	891	
unique	NaN	NaN	NaN	891	2	
top	NaN	NaN	NaN	Braund, Mr. Owen Harris	male	
freq	NaN	NaN	NaN	1	577	
mean	446.000000	0.383838	2.308642	NaN	NaN	
std	257.353842	0.486592	0.836071	NaN	NaN	
min	1.000000	0.000000	1.000000	NaN	NaN	
25%	223.500000	0.000000	2.000000	NaN	NaN	
50%	446.000000	0.000000	3.000000	NaN	NaN	
75%	668.500000	1.000000	3.000000	NaN	NaN	
max	891.000000	1.000000	3.000000	NaN	NaN	

  

	Age	SibSp	Parch	Ticket	Fare	Cabin	\
--	-----	-------	-------	--------	------	-------	---

count	714.000000	891.000000	891.000000	891	891.000000	204
unique	NaN	NaN	NaN	681	NaN	147
top	NaN	NaN	NaN	347082	NaN	B96 B98
freq	NaN	NaN	NaN	7	NaN	4
mean	29.699118	0.523008	0.381594	NaN	32.204208	NaN
std	14.526497	1.102743	0.806057	NaN	49.693429	NaN
min	0.420000	0.000000	0.000000	NaN	0.000000	NaN
25%	20.125000	0.000000	0.000000	NaN	7.910400	NaN
50%	28.000000	0.000000	0.000000	NaN	14.454200	NaN
75%	38.000000	1.000000	0.000000	NaN	31.000000	NaN
max	80.000000	8.000000	6.000000	NaN	512.329200	NaN

	Embarked
count	889
unique	3
top	S
freq	644
mean	NaN
std	NaN
min	NaN
25%	NaN
50%	NaN
75%	NaN
max	NaN

```
[263]: #adding new column
titanic_train["isha"]="isha"
```

```
[264]: #Convert the 'Pclass' column in the Titanic dataset to a pandas Categorical
        ↳object (unordered by default)
new_Pclass = pd.Categorical(titanic_train["Pclass"])
#Check the data type of the converted object - should return <class 'pandas.
        ↳core.arrays.categorical.Categorical'>
type(new_Pclass)
```

```
[264]: pandas.core.arrays.categorical.Categorical
```

```
[265]: #Get unique values
titanic_train["Pclass"].unique()
```

```
[265]: array([3, 1, 2], dtype=int64)
```

```
[266]: # Create an ordered categorical variable with defined category order: "one" <
        ↳"two" < "three"
pd.
        ↳Categorical(["one","two","three","one","two","three"],categories=["one","two","three"],order=
```

```
[266]: ['one', 'two', 'three', 'one', 'two', 'three']
Categories (3, object): ['one' < 'two' < 'three']
```

```
[267]: #Get unique values
titanic_train["Cabin"].unique()
```

```
[267]: array([nan, 'C85', 'C123', 'E46', 'G6', 'C103', 'D56', 'A6',
        'C23 C25 C27', 'B78', 'D33', 'B30', 'C52', 'B28', 'C83', 'F33',
        'F G73', 'E31', 'A5', 'D10 D12', 'D26', 'C110', 'B58 B60', 'E101',
        'F E69', 'D47', 'B86', 'F2', 'C2', 'E33', 'B19', 'A7', 'C49', 'F4',
        'A32', 'B4', 'B80', 'A31', 'D36', 'D15', 'C93', 'C78', 'D35',
        'C87', 'B77', 'E67', 'B94', 'C125', 'C99', 'C118', 'D7', 'A19',
        'B49', 'D', 'C22 C26', 'C106', 'C65', 'E36', 'C54',
        'B57 B59 B63 B66', 'C7', 'E34', 'C32', 'B18', 'C124', 'C91', 'E40',
        'T', 'C128', 'D37', 'B35', 'E50', 'C82', 'B96 B98', 'E10', 'E44',
        'A34', 'C104', 'C111', 'C92', 'E38', 'D21', 'E12', 'E63', 'A14',
        'B37', 'C30', 'D20', 'B79', 'E25', 'D46', 'B73', 'C95', 'B38',
        'B39', 'B22', 'C86', 'C70', 'A16', 'C101', 'C68', 'A10', 'E68',
        'B41', 'A20', 'D19', 'D50', 'D9', 'A23', 'B50', 'A26', 'D48',
        'E58', 'C126', 'B71', 'B51 B53 B55', 'D49', 'B5', 'B20', 'F G63',
        'C62 C64', 'E24', 'C90', 'C45', 'E8', 'B101', 'D45', 'C46', 'D30',
        'E121', 'D11', 'E77', 'F38', 'B3', 'D6', 'B82 B84', 'D17', 'A36',
        'B102', 'B69', 'E49', 'C47', 'D28', 'E17', 'A24', 'C50', 'B42',
        'C148'], dtype=object)
```

```
[268]: #Printing values from column Cabin
titanic_train["Cabin"]
```

```
[268]: 0      NaN
      1      C85
      2      NaN
      3     C123
      4      NaN
      ...
     886     NaN
     887     B42
     888     NaN
     889     C148
     890     NaN
      Name: Cabin, Length: 891, dtype: object
```

```
[269]: #Convert all values in the 'Cabin' column to string type (including NaN, which
      ↪ becomes the string 'nan')
titanic_train["Cabin"].astype(str)
```

```
[269]: 0      nan
      1     C85
```

```

2      nan
3     C123
4      nan
...
886    nan
887    B42
888    nan
889    C148
890    nan
Name: Cabin, Length: 891, dtype: object

```

```

[271]: # Convert all entries in the 'Cabin' column to string type (NaN becomes 'nan')
char_cabin = titanic_train["Cabin"].astype(str)

# Extract the first character from each cabin entry (e.g., 'C85' → 'C') using
↳ list comprehension
new_Cabin = [cabin[0] for cabin in char_cabin]

# Convert the list of first letters into a categorical variable
new_Cabin = pd.Categorical(new_Cabin)

# View the resulting Categorical object - typically deck letters like 'A',
↳ 'B', 'C', etc., and possibly 'n' for 'nan'
new_Cabin

```

```

[271]: ['n', 'C', 'n', 'C', 'n', ..., 'n', 'B', 'n', 'C', 'n']
Length: 891
Categories (9, object): ['A', 'B', 'C', 'D', ..., 'F', 'G', 'T', 'n']

```

```

[272]: # Replace the original 'Cabin' column in the Titanic dataset with the new
↳ simplified version (first character only, as categorical)
titanic_train["Cabin"] = new_Cabin

```

```

[273]: # View the first few rows of the dataset
titanic_train.head()

```

```

[273]:
  PassengerId  Survived  Pclass \
0             1         0       3
1             2         1       1
2             3         1       3
3             4         1       1
4             5         0       3

      Name      Sex  Age  SibSp \
0  Braund, Mr. Owen Harris    male  22.0      1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
2    Heikkinen, Miss. Laina   female  26.0      0

```

3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1
4	Allen, Mr. William Henry	male	35.0	0

	Parch	Ticket	Fare	Cabin	Embarked	isha
0	0	A/5 21171	7.2500	n	S	isha
1	0	PC 17599	71.2833	C	C	isha
2	0	STON/O2. 3101282	7.9250	n	S	isha
3	0	113803	53.1000	C	S	isha
4	0	373450	8.0500	n	S	isha

```
[274]: #Check for non-missing (non-null) values in the entire DataFrame - returns True
      ↪for filled cells, False for NaNs
titanic_train.notnull()
```

```
[274]:
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	\
0	True	True	True	True	True	True	True	True	True	
1	True	True	True	True	True	True	True	True	True	
2	True	True	True	True	True	True	True	True	True	
3	True	True	True	True	True	True	True	True	True	
4	True	True	True	True	True	True	True	True	True	
..	...	...	...	...	...	...	...	...	...	
886	True	True	True	True	True	True	True	True	True	
887	True	True	True	True	True	True	True	True	True	
888	True	True	True	True	True	False	True	True	True	
889	True	True	True	True	True	True	True	True	True	
890	True	True	True	True	True	True	True	True	True	

  

	Fare	Cabin	Embarked	isha
0	True	True	True	True
1	True	True	True	True
2	True	True	True	True
3	True	True	True	True
4	True	True	True	True
..	...	...	...	...
886	True	True	True	True
887	True	True	True	True
888	True	True	True	True
889	True	True	True	True
890	True	True	True	True

[891 rows x 13 columns]

```
[275]: # Check for missing values
titanic_train.isnull().sum()
```

```
[275]: PassengerId      0
      Survived         0
```

```

Pclass      0
Name        0
Sex         0
Age        177
SibSp       0
Parch       0
Ticket      0
Fare        0
Cabin       0
Embarked    2
isha        0
dtype: int64

```

```

[276]: # Check for missing values
titanic_train.isnull()

```

```

[276]:
   PassengerId  Survived  Pclass  Name  Sex  Age  SibSp  Parch  Ticket  \
0          False     False   False  False  False  False  False  False  False
1          False     False   False  False  False  False  False  False  False
2          False     False   False  False  False  False  False  False  False
3          False     False   False  False  False  False  False  False  False
4          False     False   False  False  False  False  False  False  False
..          ...      ...      ...      ...      ...      ...      ...      ...
886         False     False   False  False  False  False  False  False  False
887         False     False   False  False  False  False  False  False  False
888         False     False   False  False  False  True   False  False  False
889         False     False   False  False  False  False  False  False  False
890         False     False   False  False  False  False  False  False  False

   Fare  Cabin  Embarked  isha
0   False  False     False  False
1   False  False     False  False
2   False  False     False  False
3   False  False     False  False
4   False  False     False  False
..      ...      ...      ...      ...
886  False  False     False  False
887  False  False     False  False
888  False  False     False  False
889  False  False     False  False
890  False  False     False  False

```

[891 rows x 13 columns]

Remove the specific column

```

[124]: #Drop specified column(s)
titanic_train.drop(['PassengerId'],axis=1)

```

```
[124]:
```

	Survived	Pclass	Name \
0	0	3	Braund, Mr. Owen Harris
1	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...
2	1	3	Heikkinen, Miss. Laina
3	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)
4	0	3	Allen, Mr. William Henry
..	...	...	...
886	0	2	Montvila, Rev. Juozas
887	1	1	Graham, Miss. Margaret Edith
888	0	3	Johnston, Miss. Catherine Helen "Carrie"
889	1	1	Behr, Mr. Karl Howell
890	0	3	Dooley, Mr. Patrick

	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked \
0	male	22.0	1	0	A/5 21171	7.2500	n	S
1	female	38.0	1	0	PC 17599	71.2833	C	C
2	female	26.0	0	0	STON/O2. 3101282	7.9250	n	S
3	female	35.0	1	0	113803	53.1000	C	S
4	male	35.0	0	0	373450	8.0500	n	S
..	...	...	...	...	...	...	...	...
886	male	27.0	0	0	211536	13.0000	n	S
887	female	19.0	0	0	112053	30.0000	B	S
888	female	NaN	1	2	W./C. 6607	23.4500	n	S
889	male	26.0	0	0	111369	30.0000	C	C
890	male	32.0	0	0	370376	7.7500	n	Q

```

    isha
0    isha
1    isha
2    isha
3    isha
4    isha
..    ...
886  isha
887  isha
888  isha
889  isha
890  isha

```

```
[891 rows x 12 columns]
```

```
titanic_tarin.drop(['PassengerId'],axis=1,inplace =True)
```

```
[278]: # Check for missing values
missing=np.where(titanic_train["Age"].isnull()==True)
missing
```

```
[278]: (array([ 5, 17, 19, 26, 28, 29, 31, 32, 36, 42, 45, 46, 47,
              48, 55, 64, 65, 76, 77, 82, 87, 95, 101, 107, 109, 121,
              126, 128, 140, 154, 158, 159, 166, 168, 176, 180, 181, 185, 186,
              196, 198, 201, 214, 223, 229, 235, 240, 241, 250, 256, 260, 264,
              270, 274, 277, 284, 295, 298, 300, 301, 303, 304, 306, 324, 330,
              334, 335, 347, 351, 354, 358, 359, 364, 367, 368, 375, 384, 388,
              409, 410, 411, 413, 415, 420, 425, 428, 431, 444, 451, 454, 457,
              459, 464, 466, 468, 470, 475, 481, 485, 490, 495, 497, 502, 507,
              511, 517, 522, 524, 527, 531, 533, 538, 547, 552, 557, 560, 563,
              564, 568, 573, 578, 584, 589, 593, 596, 598, 601, 602, 611, 612,
              613, 629, 633, 639, 643, 648, 650, 653, 656, 667, 669, 674, 680,
              692, 697, 709, 711, 718, 727, 732, 738, 739, 740, 760, 766, 768,
              773, 776, 778, 783, 790, 792, 793, 815, 825, 826, 828, 832, 837,
              839, 846, 849, 859, 863, 868, 878, 888], dtype=int64),)
```

```
[126]: # Check for missing values
titanic_train['Age'][titanic_train.Age.isnull()]
```

```
[126]: 5      NaN
      17      NaN
      19      NaN
      26      NaN
      28      NaN
      ..
      859     NaN
      863     NaN
      868     NaN
      878     NaN
      888     NaN
      Name: Age, Length: 177, dtype: float64
```