In [1]:
```python
import pandas as pd
import numpy as np
from tqdm import tqdm
```

In [2]:
```python
df_dict = {'player_id': {0: 401, 1: 401, 2:  401, 3: 401, 4: 401, 5: 401, 6: 401,
df = pd.DataFrame(df_dict)
```

In [5]:
```python
#player id records for player_id=403
df[df['player_id']==403]
```

Out[5]:

	player_id	match_date	match_result
20	403	2021-05-03 00:00:00	L
21	403	2021-05-11 00:00:00	W
22	403	2021-05-12 00:00:00	W
23	403	2021-05-13 00:00:00	W
24	403	2021-05-20 00:00:00	W
25	403	2021-05-25 00:00:00	W
26	403	2021-07-06 00:00:00	L
27	403	2021-07-15 00:00:00	L
28	403	2021-07-22 00:00:00	W
29	403	2021-07-23 00:00:00	W

In [8]:
```python
df[df.player_id==403]
```

Out[8]:

	player_id	match_date	match_result
20	403	2021-05-03 00:00:00	L
21	403	2021-05-11 00:00:00	W
22	403	2021-05-12 00:00:00	W
23	403	2021-05-13 00:00:00	W
24	403	2021-05-20 00:00:00	W
25	403	2021-05-25 00:00:00	W
26	403	2021-07-06 00:00:00	L
27	403	2021-07-15 00:00:00	L
28	403	2021-07-22 00:00:00	W
29	403	2021-07-23 00:00:00	W

In [9]:
```python
# Players id =403 and give me wins/losses
```

In [12]:
```python
df_403_wins = df[ (df['player_id']==403) & (df['match_result']=='W')]
```

In [13]:
```python
df_403_wins
```

Out[13]:

	player_id	match_date	match_result
21	403	2021-05-11 00:00:00	W
22	403	2021-05-12 00:00:00	W
23	403	2021-05-13 00:00:00	W
24	403	2021-05-20 00:00:00	W
25	403	2021-05-25 00:00:00	W
28	403	2021-07-22 00:00:00	W
29	403	2021-07-23 00:00:00	W

In [14]:
```python
df_403_losses = df[ (df['player_id']==403) & (df['match_result']=='L')]
```

In [15]:
```python
df_403_losses
```

Out[15]:

	player_id	match_date	match_result
20	403	2021-05-03 00:00:00	L
26	403	2021-07-06 00:00:00	L
27	403	2021-07-15 00:00:00	L

In [18]:
```python
df_403_wins.match_result.count()/df_403_losses.match_result.count()
```

Out[18]: 2.3333333333333335

In [20]:
```python
df_403_wins.shape[0]/df_403_losses.shape[0]
```

Out[20]: 2.3333333333333335

In [22]:
```python
df[df['player_id']==403].groupby('match_result').count()
```

Out[22]:

	player_id	match_date
match_result		
L	3	3
W	7	7

In [23]:
```python
df.groupby(['player_id','match_result']).count()
```

Out[23]:

player_id	match_result	match_date
401	L	6
	W	4
402	L	5
	W	5
403	L	3
	W	7
404	L	3
	W	6
405	L	9
	W	1

In [24]:
```python
df.groupby(['match_result','player_id']).count()
```

Out[24]:

match_result	player_id	match_date
L	401	6
	402	5
	403	3
	404	3
	405	9
W	401	4
	402	5
	403	7
	404	6
	405	1

In [25]:
```python
# Accessing elements in dataframe
```

In [26]:
```python
df.head(10)
```

Out[26]:

	player_id	match_date	match_result
0	401	2021-05-04 00:00:00	W
1	401	2021-05-09 00:00:00	L
2	401	2021-05-16 00:00:00	L
3	401	2021-05-18 00:00:00	W
4	401	2021-05-22 00:00:00	L
5	401	2021-06-15 00:00:00	L
6	401	2021-06-16 00:00:00	W
7	401	2021-06-18 00:00:00	W
8	401	2021-07-06 00:00:00	L
9	401	2021-07-13 00:00:00	L

In [30]:
```python
df.iloc[:,[0,2]]
```

Out[30]:

	player_id	match_result
0	401	W
1	401	L
2	401	L
3	401	W
4	401	L
5	401	L
6	401	W
7	401	W
8	401	L
9	401	L
10	402	L
11	402	L
12	402	W
13	402	W
14	402	W
15	402	W
16	402	W
17	402	L
18	402	L
19	402	L
20	403	L
21	403	W
22	403	W
23	403	W
24	403	W
25	403	W
26	403	L
27	403	L
28	403	W
29	403	W
30	404	W
31	404	W
32	404	W
33	404	W
34	404	L
35	404	L

	player_id	match_result
36	404	W
37	404	W
38	404	L
39	405	L
40	405	L
41	405	L
42	405	L
43	405	L
44	405	L
45	405	L
46	405	W
47	405	L
48	405	L

In [35]:
```python
df.iloc[40:47,1:]
```

Out[35]:

	match_date	match_result
40	2021-05-25 00:00:00	L
41	2021-06-06 00:00:00	L
42	2021-06-07 00:00:00	L
43	2021-06-14 00:00:00	L
44	2021-07-01 00:00:00	L
45	2021-07-02 00:00:00	L
46	2021-07-14 00:00:00	W

In [37]:
```python
df.sort_values(by='match_date')
```

Out[37]:

	player_id	match_date	match_result
20	403	2021-05-03 00:00:00	L
0	401	2021-05-04 00:00:00	W
39	405	2021-05-07 00:00:00	L
1	401	2021-05-09 00:00:00	L
30	404	2021-05-10 00:00:00	W
21	403	2021-05-11 00:00:00	W
22	403	2021-05-12 00:00:00	W
23	403	2021-05-13 00:00:00	W
10	402	2021-05-14 00:00:00	L
2	401	2021-05-16 00:00:00	L
31	404	2021-05-16 00:00:00	W
3	401	2021-05-18 00:00:00	W
32	404	2021-05-20 00:00:00	W
24	403	2021-05-20 00:00:00	W
4	401	2021-05-22 00:00:00	L
33	404	2021-05-22 00:00:00	W
11	402	2021-05-23 00:00:00	L
12	402	2021-05-24 00:00:00	W
40	405	2021-05-25 00:00:00	L
25	403	2021-05-25 00:00:00	W
34	404	2021-05-28 00:00:00	L
13	402	2021-06-01 00:00:00	W
14	402	2021-06-02 00:00:00	W
41	405	2021-06-06 00:00:00	L
35	404	2021-06-06 00:00:00	L
42	405	2021-06-07 00:00:00	L
43	405	2021-06-14 00:00:00	L
36	404	2021-06-14 00:00:00	W
5	401	2021-06-15 00:00:00	L
6	401	2021-06-16 00:00:00	W
7	401	2021-06-18 00:00:00	W
44	405	2021-07-01 00:00:00	L
15	402	2021-07-01 00:00:00	W
45	405	2021-07-02 00:00:00	L
26	403	2021-07-06 00:00:00	L
8	401	2021-07-06 00:00:00	L

	player_id	match_date	match_result
16	402	2021-07-11 00:00:00	W
9	401	2021-07-13 00:00:00	L
46	405	2021-07-14 00:00:00	W
27	403	2021-07-15 00:00:00	L
47	405	2021-07-16 00:00:00	L
17	402	2021-07-20 00:00:00	L
28	403	2021-07-22 00:00:00	W
29	403	2021-07-23 00:00:00	W
37	404	2021-07-25 00:00:00	W
18	402	2021-07-26 00:00:00	L
38	404	2021-07-26 00:00:00	L
19	402	2021-07-30 00:00:00	L
48	405	2021-07-30 00:00:00	L

```python
In [38]: df['numeric_value'] = np.where(df['match_result']=='W',1,-1)
```

```python
In [40]: df[['numeric_value','match_result']]
```

Out[40]:

	numeric_value	match_result
0	1	W
1	-1	L
2	-1	L
3	1	W
4	-1	L
5	-1	L
6	1	W
7	1	W
8	-1	L
9	-1	L
10	-1	L
11	-1	L
12	1	W
13	1	W
14	1	W
15	1	W
16	1	W
17	-1	L
18	-1	L
19	-1	L
20	-1	L
21	1	W
22	1	W
23	1	W
24	1	W
25	1	W
26	-1	L
27	-1	L
28	1	W
29	1	W
30	1	W
31	1	W
32	1	W
33	1	W
34	-1	L
35	-1	L

	numeric_value	match_result
36	1	W
37	1	W
38	-1	L
39	-1	L
40	-1	L
41	-1	L
42	-1	L
43	-1	L
44	-1	L
45	-1	L
46	1	W
47	-1	L
48	-1	L

```
In [42]: df_401  = df[df['player_id']==401]
```

```
In [45]: df_401.sort_values(by='match_date',inplace=True)
```

C:\Users\rahul\AppData\Local\Temp\ipykernel_18236\2248478404.py:1: SettingWithCopy
Warning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stabl
e/user_guide/indexing.html#returning-a-view-versus-a-copy
  df_401.sort_values(by='match_date',inplace=True)

```
In [48]: df_401
```

Out[48]:

	player_id	match_date	match_result	numeric_value
0	401	2021-05-04 00:00:00	W	1
1	401	2021-05-09 00:00:00	L	-1
2	401	2021-05-16 00:00:00	L	-1
3	401	2021-05-18 00:00:00	W	1
4	401	2021-05-22 00:00:00	L	-1
5	401	2021-06-15 00:00:00	L	-1
6	401	2021-06-16 00:00:00	W	1
7	401	2021-06-18 00:00:00	W	1
8	401	2021-07-06 00:00:00	L	-1
9	401	2021-07-13 00:00:00	L	-1

```
In [59]: streak = 0
max_streak = -1
for index,item in df_403.iterrows():
```

```
#      print(index," : ",item['player_id'])
    if item['match_result']=='W':
        streak+=1
    elif item['match_result'] =='L':
        streak=0

    if streak>=max_streak:
        max_streak = streak
```

In [60]: `max_streak`

Out[60]: 5

In [77]: `df_403 = df[df['player_id']==403].sort_values(by='match_date')`

In [78]: `df_403['shifted_result'] = df_403['match_result'].shift()`

In [79]: `df_403['start_streak'] = df_403['match_result']!=df_403['shifted_result']`

In [80]: `df_403`

Out[80]:

	player_id	match_date	match_result	numeric_value	shifted_result	start_streak
20	403	2021-05-03 00:00:00	L	-1	NaN	True
21	403	2021-05-11 00:00:00	W	1	L	True
22	403	2021-05-12 00:00:00	W	1	W	False
23	403	2021-05-13 00:00:00	W	1	W	False
24	403	2021-05-20 00:00:00	W	1	W	False
25	403	2021-05-25 00:00:00	W	1	W	False
26	403	2021-07-06 00:00:00	L	-1	W	True
27	403	2021-07-15 00:00:00	L	-1	L	False
28	403	2021-07-22 00:00:00	W	1	L	True
29	403	2021-07-23 00:00:00	W	1	W	False

In [81]: `df_403['streak_id'] = df_403['start_streak'].cumsum()`

In [82]: `df_403`

Out[82]:

	player_id	match_date	match_result	numeric_value	shifted_result	start_streak	streak_id
20	403	2021-05-03 00:00:00	L	-1	NaN	True	1
21	403	2021-05-11 00:00:00	W	1	L	True	2
22	403	2021-05-12 00:00:00	W	1	W	False	2
23	403	2021-05-13 00:00:00	W	1	W	False	2
24	403	2021-05-20 00:00:00	W	1	W	False	2
25	403	2021-05-25 00:00:00	W	1	W	False	2
26	403	2021-07-06 00:00:00	L	-1	W	True	3
27	403	2021-07-15 00:00:00	L	-1	L	False	3
28	403	2021-07-22 00:00:00	W	1	L	True	4
29	403	2021-07-23 00:00:00	W	1	W	False	4

In [84]:
```python
df_403_grp = df_403.groupby(['streak_id','match_result'])['player_id'].count().res
```

In [86]:
```python
df_403_grp[df_403_grp['match_result']=='W'].sort_values('player_id',ascending=Fals
```

Out[86]: 5

In [ ]: