

# Pandas

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

```
In [2]: df_infy = pd.read_csv("INFY.csv")
```

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

```
Out[3]:
```

	Date	Open	High	Low	Close	Adj Close	Volume
0	2021-09-17	22.950001	22.990000	22.719999	22.850000	22.401333	8070400
1	2021-09-20	22.750000	22.950001	22.570000	22.760000	22.313101	7148800
2	2021-09-21	23.129999	23.230000	23.100000	23.120001	22.666033	4785000
3	2021-09-22	23.120001	23.379999	23.070000	23.270000	22.813086	6613800
4	2021-09-23	23.680000	23.790001	23.600000	23.730000	23.264053	5323800

```
In [13]: df_infy.sort_values(by = "Volume", ascending = False)
```

```
Out[13]:
```

	Date	Open	High	Low	Close	Adj Close	Volume
251	2022-09-16	17.200001	17.459999	17.090000	17.379999	17.379999	42686600
126	2022-03-18	24.389999	25.040001	24.209999	25.040001	24.763243	29035600
154	2022-04-28	20.410000	20.660000	20.350000	20.500000	20.273422	24229100
20	2021-10-15	22.900000	23.400000	22.760000	23.379999	22.920927	22295600
250	2022-09-15	17.700001	17.809999	17.510000	17.790001	17.790001	21950100
...	...	...	...	...	...	...	...
232	2022-08-19	20.040001	20.040001	19.750000	19.780001	19.780001	2868500
132	2022-03-28	24.730000	24.770000	24.520000	24.719999	24.446779	2834400
68	2021-12-23	24.690001	24.799999	24.500000	24.730000	24.456669	2488600
71	2021-12-29	25.190001	25.379999	25.139999	25.379999	25.099483	2330400
72	2021-12-30	25.540001	25.600000	25.389999	25.410000	25.129152	2293900

252 rows × 7 columns

```
In [14]: df_dict = {'player_id': {0: 401, 1: 401, 2: 401, 3: 401, 4: 401, 5: 401, 6: 401, 7:
```

```
In [19]: df = pd.DataFrame(df_dict)
df
```

	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
10	402	2021-05-14 00:00:00	L
11	402	2021-05-23 00:00:00	L
12	402	2021-05-24 00:00:00	W
13	402	2021-06-01 00:00:00	W
14	402	2021-06-02 00:00:00	W
15	402	2021-07-01 00:00:00	W
16	402	2021-07-11 00:00:00	W
17	402	2021-07-20 00:00:00	L
18	402	2021-07-26 00:00:00	L
19	402	2021-07-30 00:00:00	L
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
30	404	2021-05-10 00:00:00	W
31	404	2021-05-16 00:00:00	W
32	404	2021-05-20 00:00:00	W

	player_id	match_date	match_result
33	404	2021-05-22 00:00:00	W
34	404	2021-05-28 00:00:00	L
35	404	2021-06-06 00:00:00	L
36	404	2021-06-14 00:00:00	W
37	404	2021-07-25 00:00:00	W
38	404	2021-07-26 00:00:00	L
39	405	2021-05-07 00:00:00	L
40	405	2021-05-25 00:00:00	L
41	405	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
44	405	2021-07-01 00:00:00	L
45	405	2021-07-02 00:00:00	L
46	405	2021-07-14 00:00:00	W
47	405	2021-07-16 00:00:00	L
48	405	2021-07-30 00:00:00	L

```
In [20]: # with open("csv/sample.json", "w") as file :  
#     json.dump(df_dict,file) #javascript
```

```
In [22]: df[df[ "player_id" ] == 401]
```

Out[22]:

	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
10	401	2021-05-14 00:00:00	L
11	401	2021-05-23 00:00:00	L
12	401	2021-05-24 00:00:00	W
13	401	2021-06-01 00:00:00	W
14	401	2021-06-02 00:00:00	W
15	401	2021-07-01 00:00:00	W
16	401	2021-07-11 00:00:00	W
17	401	2021-07-20 00:00:00	L
18	401	2021-07-26 00:00:00	L
19	401	2021-07-30 00:00:00	L
20	401	2021-05-03 00:00:00	L
21	401	2021-05-11 00:00:00	W
22	401	2021-05-12 00:00:00	W
23	401	2021-05-13 00:00:00	W
24	401	2021-05-20 00:00:00	W
25	401	2021-05-25 00:00:00	W
26	401	2021-07-06 00:00:00	L
27	401	2021-07-15 00:00:00	L
28	401	2021-07-22 00:00:00	W
29	401	2021-07-23 00:00:00	W
30	401	2021-05-10 00:00:00	W
31	401	2021-05-16 00:00:00	W
32	401	2021-05-20 00:00:00	W

	player_id	match_date	match_result
33	401	2021-05-22 00:00:00	W
34	401	2021-05-28 00:00:00	L
35	401	2021-06-06 00:00:00	L
36	401	2021-06-14 00:00:00	W
37	401	2021-07-25 00:00:00	W
38	401	2021-07-26 00:00:00	L
39	401	2021-05-07 00:00:00	L
40	401	2021-05-25 00:00:00	L
41	401	2021-06-06 00:00:00	L
42	401	2021-06-07 00:00:00	L
43	401	2021-06-14 00:00:00	L
44	401	2021-07-01 00:00:00	L
45	401	2021-07-02 00:00:00	L
46	401	2021-07-14 00:00:00	W
47	401	2021-07-16 00:00:00	L
48	401	2021-07-30 00:00:00	L

In [23]: `df["player_id"] == 401`

```
Out[23]: 0    True
         1    True
         2    True
         3    True
         4    True
         5    True
         6    True
         7    True
         8    True
         9    True
        10   True
        11   True
        12   True
        13   True
        14   True
        15   True
        16   True
        17   True
        18   True
        19   True
        20   True
        21   True
        22   True
        23   True
        24   True
        25   True
        26   True
        27   True
        28   True
        29   True
        30   True
        31   True
        32   True
        33   True
        34   True
        35   True
        36   True
        37   True
        38   True
        39   True
        40   True
        41   True
        42   True
        43   True
        44   True
        45   True
        46   True
        47   True
        48   True
Name: player_id, dtype: bool
```

```
In [24]: df_infy
```

Out[24]:

	Date	Open	High	Low	Close	Adj Close	Volume
0	2021-09-17	22.950001	22.990000	22.719999	22.850000	22.401333	8070400
1	2021-09-20	22.750000	22.950001	22.570000	22.760000	22.313101	7148800
2	2021-09-21	23.129999	23.230000	23.100000	23.120001	22.666033	4785000
3	2021-09-22	23.120001	23.379999	23.070000	23.270000	22.813086	6613800
4	2021-09-23	23.680000	23.790001	23.600000	23.730000	23.264053	5323800
...	...	...	...	...	...	...	...
247	2022-09-12	19.230000	19.410000	19.190001	19.240000	19.240000	3607400
248	2022-09-13	18.760000	18.900000	18.490000	18.559999	18.559999	15227800
249	2022-09-14	18.370001	18.440001	17.969999	18.080000	18.080000	16757300
250	2022-09-15	17.700001	17.809999	17.510000	17.790001	17.790001	21950100
251	2022-09-16	17.200001	17.459999	17.090000	17.379999	17.379999	42686600

252 rows × 7 columns

In [26]: df\_infy["Volume"]

```
Out[26]: 0      8070400
1      7148800
2      4785000
3      6613800
4      5323800
...
247    3607400
248    15227800
249    16757300
250    21950100
251    42686600
Name: Volume, Length: 252, dtype: int64
```

In [28]: df\_infy["Volume"].sum()

Out[28]: 2381188600

how many unique player id are there

In [29]: df["player\_id"].unique()

Out[29]: array([401], dtype=int64)

In [30]: df

	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
10	401	2021-05-14 00:00:00	L
11	401	2021-05-23 00:00:00	L
12	401	2021-05-24 00:00:00	W
13	401	2021-06-01 00:00:00	W
14	401	2021-06-02 00:00:00	W
15	401	2021-07-01 00:00:00	W
16	401	2021-07-11 00:00:00	W
17	401	2021-07-20 00:00:00	L
18	401	2021-07-26 00:00:00	L
19	401	2021-07-30 00:00:00	L
20	401	2021-05-03 00:00:00	L
21	401	2021-05-11 00:00:00	W
22	401	2021-05-12 00:00:00	W
23	401	2021-05-13 00:00:00	W
24	401	2021-05-20 00:00:00	W
25	401	2021-05-25 00:00:00	W
26	401	2021-07-06 00:00:00	L
27	401	2021-07-15 00:00:00	L
28	401	2021-07-22 00:00:00	W
29	401	2021-07-23 00:00:00	W
30	401	2021-05-10 00:00:00	W
31	401	2021-05-16 00:00:00	W
32	401	2021-05-20 00:00:00	W

	player_id	match_date	match_result
33	401	2021-05-22 00:00:00	W
34	401	2021-05-28 00:00:00	L
35	401	2021-06-06 00:00:00	L
36	401	2021-06-14 00:00:00	W
37	401	2021-07-25 00:00:00	W
38	401	2021-07-26 00:00:00	L
39	401	2021-05-07 00:00:00	L
40	401	2021-05-25 00:00:00	L
41	401	2021-06-06 00:00:00	L
42	401	2021-06-07 00:00:00	L
43	401	2021-06-14 00:00:00	L
44	401	2021-07-01 00:00:00	L
45	401	2021-07-02 00:00:00	L
46	401	2021-07-14 00:00:00	W
47	401	2021-07-16 00:00:00	L
48	401	2021-07-30 00:00:00	L

```
In [31]: df = pd.DataFrame(df_dict)
df
```

Out[31]:	player_id	match_date	match_result
<b>0</b>	401	2021-05-04 00:00:00	W
<b>1</b>	401	2021-05-09 00:00:00	L
<b>2</b>	401	2021-05-16 00:00:00	L
<b>3</b>	401	2021-05-18 00:00:00	W
<b>4</b>	401	2021-05-22 00:00:00	L
<b>5</b>	401	2021-06-15 00:00:00	L
<b>6</b>	401	2021-06-16 00:00:00	W
<b>7</b>	401	2021-06-18 00:00:00	W
<b>8</b>	401	2021-07-06 00:00:00	L
<b>9</b>	401	2021-07-13 00:00:00	L
<b>10</b>	402	2021-05-14 00:00:00	L
<b>11</b>	402	2021-05-23 00:00:00	L
<b>12</b>	402	2021-05-24 00:00:00	W
<b>13</b>	402	2021-06-01 00:00:00	W
<b>14</b>	402	2021-06-02 00:00:00	W
<b>15</b>	402	2021-07-01 00:00:00	W
<b>16</b>	402	2021-07-11 00:00:00	W
<b>17</b>	402	2021-07-20 00:00:00	L
<b>18</b>	402	2021-07-26 00:00:00	L
<b>19</b>	402	2021-07-30 00:00:00	L
<b>20</b>	403	2021-05-03 00:00:00	L
<b>21</b>	403	2021-05-11 00:00:00	W
<b>22</b>	403	2021-05-12 00:00:00	W
<b>23</b>	403	2021-05-13 00:00:00	W
<b>24</b>	403	2021-05-20 00:00:00	W
<b>25</b>	403	2021-05-25 00:00:00	W
<b>26</b>	403	2021-07-06 00:00:00	L
<b>27</b>	403	2021-07-15 00:00:00	L
<b>28</b>	403	2021-07-22 00:00:00	W
<b>29</b>	403	2021-07-23 00:00:00	W
<b>30</b>	404	2021-05-10 00:00:00	W
<b>31</b>	404	2021-05-16 00:00:00	W
<b>32</b>	404	2021-05-20 00:00:00	W

	player_id	match_date	match_result
33	404	2021-05-22 00:00:00	W
34	404	2021-05-28 00:00:00	L
35	404	2021-06-06 00:00:00	L
36	404	2021-06-14 00:00:00	W
37	404	2021-07-25 00:00:00	W
38	404	2021-07-26 00:00:00	L
39	405	2021-05-07 00:00:00	L
40	405	2021-05-25 00:00:00	L
41	405	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
44	405	2021-07-01 00:00:00	L
45	405	2021-07-02 00:00:00	L
46	405	2021-07-14 00:00:00	W
47	405	2021-07-16 00:00:00	L
48	405	2021-07-30 00:00:00	L

```
In [32]: df["player_id"].unique() #
```

```
Out[32]: array([401, 402, 403, 404, 405], dtype=int64)
```

```
In [34]: df["player_id"].nunique() # count of unique columns
```

```
Out[34]: 5
```

```
In [35]: spotify_dict = {'id': {0: 303651, 1: 85559, 2: 1046089, 3: 350824, 4: 776822, 5: 46
```

```
In [37]: df_spotify = pd.DataFrame(spotify_dict)
```

```
In [38]: df_spotify
```

Out[38]:

	<b>id</b>	<b>position</b>	<b>trackname</b>	<b>artist</b>	<b>streams</b>	<b>url</b>	<b>date</b>	<b>region</b>
<b>0</b>	303651	52	Heart Won't Forget	Matoma	28047	https://open.spotify.com/track/2of2DM5LqTh7ohm...	2017-02-04 00:00:00	nc
<b>1</b>	85559	160	Someone In The Crowd - From "La La Land" Sound...	Emma Stone	17134	https://open.spotify.com/track/7xE4vKvjqUTtHyJ...	2017-02-26 00:00:00	fi
<b>2</b>	1046089	175	The Greatest	Sia	10060	https://open.spotify.com/track/7xHWNBfM6ObGEQP...	2017-03-06 00:00:00	c
<b>3</b>	350824	25	Unforgettable	French Montana	46603	https://open.spotify.com/track/3B54sVLJ402zGa6...	2017-10-01 00:00:00	nc
<b>4</b>	776822	1	Bad and Boujee (feat. Lil Uzi Vert)	Migos	1823391	https://open.spotify.com/track/4Km5HrUvYTaSUfi...	2017-01-27 00:00:00	us
...	...	...	...	...	...	...	...	..
<b>95</b>	792423	2	DNA.	Kendrick Lamar	3013496	https://open.spotify.com/track/6HZILIRieu8S0iq...	2017-04-15 00:00:00	us
<b>96</b>	792223	2	DNA.	Kendrick Lamar	3643231	https://open.spotify.com/track/6HZILIRieu8S0iq...	2017-04-14 00:00:00	us
<b>97</b>	793422	1	HUMBLE.	Kendrick Lamar	3144482	https://open.spotify.com/track/7KXjTSCq5nL1LoY...	2017-04-20 00:00:00	us
<b>98</b>	793622	1	HUMBLE.	Kendrick Lamar	3172718	https://open.spotify.com/track/7KXjTSCq5nL1LoY...	2017-04-21 00:00:00	us
<b>99</b>	793022	1	HUMBLE.	Kendrick Lamar	3394456	https://open.spotify.com/track/7KXjTSCq5nL1LoY...	2017-04-18 00:00:00	us

100 rows × 8 columns

In [39]: df\_spotify["streams"]

Out[39]:

```
0    28047
1    17134
2    10060
3    46603
4    1823391
      ...
95   3013496
96   3643231
97   3144482
98   3172718
99   3394456
Name: streams, Length: 100, dtype: int64
```

```
In [40]: df_spotify["streams"].sum()
```

```
Out[40]: 39673624
```

```
In [42]: df_spotify["streams"].sum()/100
```

```
Out[42]: 396736.24
```

```
In [50]: host_dict = {'host_id': {0: 0, 1: 1, 2: 2, 3: 3, 4: 4, 5: 5, 6: 6, 7: 7, 8: 8, 9: 9}}
```

```
In [52]: df_host = pd.DataFrame(host_dict)
df_host
```

```
Out[52]:
```

	host_id	nationality	gender	age
0	0	USA	M	28
1	1	USA	F	29
2	2	China	F	31
3	3	China	M	24
4	4	Mali	M	30
...	...	...	...	...
171	7	Luxembourg	F	25
172	6	Luxembourg	M	25
173	7	Luxembourg	F	25
174	6	Luxembourg	M	25
175	7	Luxembourg	F	25

176 rows × 4 columns

```
In [53]: appart_dict = {'host_id': {0: 0, 1: 0, 2: 0, 3: 1, 4: 1, 5: 2, 6: 3, 7: 3, 8: 4, 9: 9}}
```

```
In [54]: df_apprt = pd.DataFrame(appart_dict)
df_apprt
```

Out[54]:

	host_id	apartment_id	apartment_type	n_beds	n_bedrooms	country	city
0	0	A1	Room	1	1	USA	New York
1	0	A2	Room	1	1	USA	New Jersey
2	0	A3	Room	1	1	USA	New Jersey
3	1	A4	Apartment	2	1	USA	Houston
4	1	A5	Apartment	2	1	USA	Las Vegas
5	2	A6	Yurt	3	1	Mongolia	-
6	3	A7	Penthouse	3	3	China	Tianjin
7	3	A8	Penthouse	5	5	China	Beijing
8	4	A9	Apartment	2	1	Mali	Bamako
9	5	A10	Room	3	1	Mali	Segou
10	5	A11	Room	2	1	Mali	Segou
11	6	A12	Penthouse	6	6	Luxembourg	Luxembourg
12	7	A13	Room	4	1	Luxembourg	Luxembourg
13	8	A14	Apartment	2	1	Australia	Perth
14	9	A15	Apartment	2	1	Australia	Perth
15	9	A16	Apartment	2	1	Australia	Perth
16	10	A17	Room	4	1	Brazil	Rio De Janeiro
17	10	A18	Room	4	1	Argentina	Mendoza
18	10	A19	Room	4	2	Uruguay	Mercedes
19	10	A20	Room	4	2	Brazil	Brasilia
20	11	A21	Apartment	2	2	Mexico	Mexico City

In [58]: df\_host\_apprt = pd.merge(df\_host,df\_apprt, on = "host\_id")

In [59]: df\_host\_apprt

Out[59]:

	host_id	nationality	gender	age	apartment_id	apartment_type	n_beds	n_bedrooms	count
0	0	USA	M	28	A1	Room	1	1	U
1	0	USA	M	28	A2	Room	1	1	U
2	0	USA	M	28	A3	Room	1	1	U
3	0	USA	M	28	A1	Room	1	1	U
4	0	USA	M	28	A2	Room	1	1	U
...	...	...	...	...	...	...	...	...	...
347	10	Brazil	M	39	A20	Room	4	2	Bra
348	11	Brazil	F	42	A21	Apartment	2	2	Mexi
349	11	Brazil	F	42	A21	Apartment	2	2	Mexi
350	11	Brazil	F	42	A21	Apartment	2	2	Mexi
351	11	Brazil	F	42	A21	Apartment	2	2	Mexi

352 rows × 10 columns

In [60]: df\_host\_apprt[df\_host\_apprt["host\_id"]==0]

Out[60]:

	host_id	nationality	gender	age	apartment_id	apartment_type	n_beds	n_bedrooms	count
0	0	USA	M	28	A1	Room	1	1	U
1	0	USA	M	28	A2	Room	1	1	U
2	0	USA	M	28	A3	Room	1	1	U
3	0	USA	M	28	A1	Room	1	1	U
4	0	USA	M	28	A2	Room	1	1	U
...	...	...	...	...	...	...	...	...	...
211	0	USA	M	28	A2	Room	1	1	U
212	0	USA	M	28	A3	Room	1	1	U
213	0	USA	M	28	A1	Room	1	1	U
214	0	USA	M	28	A2	Room	1	1	U
215	0	USA	M	28	A3	Room	1	1	U

216 rows × 10 columns

In [61]: df\_host\_apprt[df\_host\_apprt["nationality"] != df\_host\_apprt["country"]]

Out[61]:

	host_id	nationality	gender	age	apartment_id	apartment_type	n_beds	n_bedrooms	cou
232	2	China	F	31	A6	Yurt	3	1	Mong
233	2	China	F	31	A6	Yurt	3	1	Mong
234	2	China	F	31	A6	Yurt	3	1	Mong
235	2	China	F	31	A6	Yurt	3	1	Mong
333	10	Brazil	M	39	A18	Room	4	1	Arger
334	10	Brazil	M	39	A19	Room	4	2	Uruç
337	10	Brazil	M	39	A18	Room	4	1	Arger
338	10	Brazil	M	39	A19	Room	4	2	Uruç
341	10	Brazil	M	39	A18	Room	4	1	Arger
342	10	Brazil	M	39	A19	Room	4	2	Uruç
345	10	Brazil	M	39	A18	Room	4	1	Arger
346	10	Brazil	M	39	A19	Room	4	2	Uruç
348	11	Brazil	F	42	A21	Apartment	2	2	Me
349	11	Brazil	F	42	A21	Apartment	2	2	Me
350	11	Brazil	F	42	A21	Apartment	2	2	Me
351	11	Brazil	F	42	A21	Apartment	2	2	Me

In [63]: df\_host\_apprt[df\_host\_apprt["nationality"] == df\_host\_apprt["country"]]

Out[63]:

	host_id	nationality	gender	age	apartment_id	apartment_type	n_beds	n_bedrooms	count
0	0	USA	M	28	A1	Room	1	1	U
1	0	USA	M	28	A2	Room	1	1	U
2	0	USA	M	28	A3	Room	1	1	U
3	0	USA	M	28	A1	Room	1	1	U
4	0	USA	M	28	A2	Room	1	1	U
...	...	...	...	...	...	...	...	...	...
339	10	Brazil	M	39	A20	Room	4	2	Bra
340	10	Brazil	M	39	A17	Room	4	1	Bra
343	10	Brazil	M	39	A20	Room	4	2	Bra
344	10	Brazil	M	39	A17	Room	4	1	Bra
347	10	Brazil	M	39	A20	Room	4	2	Bra

336 rows × 10 columns

In [64]: `df_host_apprt[df_host_apprt["nationality"] != df_host_apprt["country"]]["host_id"].`

Out[64]: 3

Joins are very important : [https://www.w3schools.com/sql/sql\\_join.asp](https://www.w3schools.com/sql/sql_join.asp)

rename

In [66]: `df_host_apprt.rename(columns = {"host_id" : "renamed_host_id"}, inplace = True )`

In [67]: `df_host_apprt`

Out[67]:

	renamed_host_id	nationality	gender	age	apartment_id	apartment_type	n_beds	n_bedrooms
0	0	USA	M	28	A1	Room	1	
1	0	USA	M	28	A2	Room	1	
2	0	USA	M	28	A3	Room	1	
3	0	USA	M	28	A1	Room	1	
4	0	USA	M	28	A2	Room	1	
...	...	...	...	...	...	...	...	...
347	10	Brazil	M	39	A20	Room	4	
348	11	Brazil	F	42	A21	Apartment	2	
349	11	Brazil	F	42	A21	Apartment	2	
350	11	Brazil	F	42	A21	Apartment	2	
351	11	Brazil	F	42	A21	Apartment	2	

352 rows × 10 columns

In [68]: df\_host\_apprt.columns

Out[68]: Index(['renamed\_host\_id', 'nationality', 'gender', 'age', 'apartment\_id', 'apartment\_type', 'n\_beds', 'n\_bedrooms', 'country', 'city'], dtype='object')

Date wide data is called timeseries data we can use this to predict something

In [ ]: