

```
In [1]: ▶ import pandas as pd
```

```
In [2]: ▶ data = pd.read_excel('wikipedia_dataset.xlsx')
data
```

Out[2]:

		Page	2016-01-01 00:00:00	2016-01-02 00:00:00	2016-01-03 00:00:00	2016-01-04 00:00:00
0	.xxx_en.wikipedia.org_all-access_all-agents		7089	7592	7436	703
1	.xxx_en.wikipedia.org_mobile-web_all-agents		6182	6447	6437	591
2	1._Juli_de.wikipedia.org_desktop_all-agents		20	49	33	4
3	2014_FIFA_World_Cup_en.wikipedia.org_all-acces...		5363	6280	6297	628
4	2014_FIFA_World_Cup_en.wikipedia.org_desktop_a...		3186	3784	3832	405
...	...		...	...	...	.
1495	Zac_Efron_en.wikipedia.org_all-access_all-agents		14396	15501	13176	1100
1496	Zayn_Malik_en.wikipedia.org_all-access_all-agents		31663	18883	19065	1797
1497	Zendaya_en.wikipedia.org_all-access_all-agents		7091	6062	4719	430
1498	Zendaya_en.wikipedia.org_mobile-web_all-agents		5221	4344	3272	275
1499	Zodiac_Killer_en.wikipedia.org_all-access_all-...		4251	5850	5833	573

1500 rows × 367 columns



```
In [4]: ▶ wiki_1 = pd.melt(data, id_vars='Page',
                        var_name= 'Dates',
                        value_name= 'Daily_Page_Visit_Count'
                        )
wiki_1
```

Out[4]:

	Page	Dates	Daily_Page_Visit_Count
0	.xxx_en.wikipedia.org_all-access_all-agents	2016-01-01	7089.0
1	.xxx_en.wikipedia.org_mobile-web_all-agents	2016-01-01	6182.0
2	1._Juli_de.wikipedia.org_desktop_all-agents	2016-01-01	20.0
3	2014_FIFA_World_Cup_en.wikipedia.org_all-acces...	2016-01-01	5363.0
4	2014_FIFA_World_Cup_en.wikipedia.org_desktop_a...	2016-01-01	3186.0
...	...	...	...
548995	Zac_Efron_en.wikipedia.org_all-access_all-agents	2016-12-31	13146.0
548996	Zayn_Malik_en.wikipedia.org_all-access_all-agents	2016-12-31	12597.0
548997	Zendaya_en.wikipedia.org_all-access_all-agents	2016-12-31	6217.0
548998	Zendaya_en.wikipedia.org_mobile-web_all-agents	2016-12-31	4650.0
548999	Zodiac_Killer_en.wikipedia.org_all-access_all-...	2016-12-31	6659.0

549000 rows × 3 columns

```
In [9]: ▶ wiki_2 = wiki_1.rename(columns = {'Page' : 'Search_Topics'})
wiki_2.head()
```

Out[9]:

	Search_Topics	Dates	Daily_Page_Visit_Count
0	.xxx_en.wikipedia.org_all-access_all-agents	2016-01-01	7089.0
1	.xxx_en.wikipedia.org_mobile-web_all-agents	2016-01-01	6182.0
2	1._Juli_de.wikipedia.org_desktop_all-agents	2016-01-01	20.0
3	2014_FIFA_World_Cup_en.wikipedia.org_all-acces...	2016-01-01	5363.0
4	2014_FIFA_World_Cup_en.wikipedia.org_desktop_a...	2016-01-01	3186.0

In [10]: `wiki_2.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 549000 entries, 0 to 548999
Data columns (total 3 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Search_Topics                        549000 non-null object
1   Dates                               549000 non-null datetime64[ns]
2   Daily_Page_Visit_Count              548913 non-null float64
dtypes: datetime64[ns](1), float64(1), object(1)
memory usage: 12.6+ MB
```

In [12]: `wiki_2.dropna(inplace = True)`  
`wiki_2.info()`

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 548913 entries, 0 to 548999
Data columns (total 3 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Search_Topics                        548913 non-null object
1   Dates                               548913 non-null datetime64[ns]
2   Daily_Page_Visit_Count              548913 non-null float64
dtypes: datetime64[ns](1), float64(1), object(1)
memory usage: 16.8+ MB
```

In [13]: `wiki_tidy = (wiki_2`  
 `.assign(Month = lambda x: x['Dates']`  
 `.dt.month_name())`  
 `.assign(Day = lambda x: x['Dates']`  
 `.dt.day)`  
`)`  
`wiki_tidy.head()`

Out[13]:

	Search_Topics	Dates	Daily_Page_Visit_Count	Month
0	.xxx_en.wikipedia.org_all-access_all-agents	2016-01-01	7089.0	January
1	.xxx_en.wikipedia.org_mobile-web_all-agents	2016-01-01	6182.0	January
2	1._Juli_de.wikipedia.org_desktop_all-agents	2016-01-01	20.0	January
3	2014_FIFA_World_Cup_en.wikipedia.org_all-acces...	2016-01-01	5363.0	January
4	2014_FIFA_World_Cup_en.wikipedia.org_desktop_a...	2016-01-01	3186.0	January

```
In [14]: > wiki_tidy.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 548913 entries, 0 to 548999
Data columns (total 5 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Search_Topics                        548913 non-null object
1   Dates                               548913 non-null datetime64[ns]
2   Daily_Page_Visit_Count              548913 non-null float64
3   Month                               548913 non-null object
4   Day                                 548913 non-null int64
dtypes: datetime64[ns](1), float64(1), int64(1), object(2)
memory usage: 25.1+ MB
```

## SUBSETTING

```
In [16]: > Jan_wiki = wiki_tidy[wiki_tidy['Month'] == 'January']
Jan_wiki
```

Out[16]:

	Search_Topics	Dates	Daily_Page_Visit_Count	Month
0	.xxx_en.wikipedia.org_all-access_all-agents	2016-01-01	7089.0	January
1	.xxx_en.wikipedia.org_mobile-web_all-agents	2016-01-01	6182.0	January
2	1._Juli_de.wikipedia.org_desktop_all-agents	2016-01-01	20.0	January
3	2014_FIFA_World_Cup_en.wikipedia.org_all-access_all-agents	2016-01-01	5363.0	January
4	2014_FIFA_World_Cup_en.wikipedia.org_desktop_all-agents	2016-01-01	3186.0	January
...	...	...	...	...
46495	Zac_Efron_en.wikipedia.org_all-access_all-agents	2016-01-31	17548.0	January
46496	Zayn_Malik_en.wikipedia.org_all-access_all-agents	2016-01-31	56947.0	January
46497	Zendaya_en.wikipedia.org_all-access_all-agents	2016-01-31	6298.0	January
46498	Zendaya_en.wikipedia.org_mobile-web_all-agents	2016-01-31	4021.0	January
46499	Zodiac_Killer_en.wikipedia.org_all-access_all-agents	2016-01-31	8044.0	January

46500 rows × 5 columns



```
In [17]: ▶ Jan_1 = (Jan_wiki[Jan_wiki['Day']==1])
Jan_1
```

Out[17]:

	Search_Topics	Dates	Daily_Page_Visit_Count	Mor
0	.xxx_en.wikipedia.org_all-access_all-agents	2016-01-01	7089.0	Janu:
1	.xxx_en.wikipedia.org_mobile-web_all-agents	2016-01-01	6182.0	Janu:
2	1._Juli_de.wikipedia.org_desktop_all-agents	2016-01-01	20.0	Janu:
3	2014_FIFA_World_Cup_en.wikipedia.org_all-acces...	2016-01-01	5363.0	Janu:
4	2014_FIFA_World_Cup_en.wikipedia.org_desktop_a...	2016-01-01	3186.0	Janu:
...	...	...	...	
1495	Zac_Efron_en.wikipedia.org_all-access_all-agents	2016-01-01	14396.0	Janu:
1496	Zayn_Malik_en.wikipedia.org_all-access_all-agents	2016-01-01	31663.0	Janu:
1497	Zendaya_en.wikipedia.org_all-access_all-agents	2016-01-01	7091.0	Janu:
1498	Zendaya_en.wikipedia.org_mobile-web_all-agents	2016-01-01	5221.0	Janu:
1499	Zodiac_Killer_en.wikipedia.org_all-access_all-...	2016-01-01	4251.0	Janu:

1500 rows × 5 columns



```
In [28]: Jan = pd.DataFrame(Jan_1.groupby(["Daily_Page_Visit_Count"])[ "Search_Topic"
Jan
```

```
Out[28]:
```

	0	1
0	5.0	38 ā,¬āf¼ā,¼ā,āf¼āfž_ja.wikipedia.org_all-...
1	7.0	467 è..."ā@mq—ţâ€™ç³¼mq_zh.wikipedia.org_deskto...
2	10.0	621 Guns_N'_Roses_es.wikipedia.org_all-acce...
3	11.0	46 āmqšè¥¿èè±ç"·_ja.wikipedia.org_desktop_al...
4	12.0	466 è..."ā@mq—ţâ€™ç³¼mq_zh.wikipedia.org_all-ac...
...	...	...
1445	2306896.0	1463 Wikipedia:Hauptseite_de.wikipedia.org_...
1446	3062989.0	1461 Wikipedia:Hauptseite_de.wikipedia.org_...
1447	6920930.0	915 Main_Page_en.wikipedia.org_mobile-web_a...
1448	9170612.0	914 Main_Page_en.wikipedia.org_desktop_all-...
1449	16357307.0	912 Main_Page_en.wikipedia.org_all-access_a...

1450 rows × 2 columns

```
In [30]: df_Jan1 = (Jan.rename(columns = {0: "Daily_Page_Visit_Count", 1: 'Search_T
df_Jan1
```

```
Out[30]:
```

	Daily_Page_Visit_Count	Search_Topics
0	5.0	38 ā,¬āf¼ā,¼ā,āf¼āfž_ja.wikipedia.org_all-...
1	7.0	467 è..."ā@mq—ţâ€™ç³¼mq_zh.wikipedia.org_deskto...
2	10.0	621 Guns_N'_Roses_es.wikipedia.org_all-acce...
3	11.0	46 āmqšè¥¿èè±ç"·_ja.wikipedia.org_desktop_al...
4	12.0	466 è..."ā@mq—ţâ€™ç³¼mq_zh.wikipedia.org_all-ac...
...	...	...
1445	2306896.0	1463 Wikipedia:Hauptseite_de.wikipedia.org_...
1446	3062989.0	1461 Wikipedia:Hauptseite_de.wikipedia.org_...
1447	6920930.0	915 Main_Page_en.wikipedia.org_mobile-web_a...
1448	9170612.0	914 Main_Page_en.wikipedia.org_desktop_all-...
1449	16357307.0	912 Main_Page_en.wikipedia.org_all-access_a...

1450 rows × 2 columns

MOST TRENDING IN JANUARY

```
In [32]: df_Jan1.tail(10)
```

Out[32]:

	Daily_Page_Visit_Count	Search_Topics
1440	1214388.0	1464 Wikipedia:Portada_es.wikipedia.org_all...
1441	1253720.0	1449 WikipÃ©dia:Accueil_principal_fr.wikiped...
1442	1401667.0	1250 Special:Search_en.wikipedia.org_deskto...
1443	1532934.0	1447 WikipÃ©dia:Accueil_principal_fr.wikiped...
1444	1794204.0	1249 Special:Search_en.wikipedia.org_all-ac...
1445	2306896.0	1463 Wikipedia:Hauptseite_de.wikipedia.org_...
1446	3062989.0	1461 Wikipedia:Hauptseite_de.wikipedia.org_...
1447	6920930.0	915 Main_Page_en.wikipedia.org_mobile-web_a...
1448	9170612.0	914 Main_Page_en.wikipedia.org_desktop_all-...
1449	16357307.0	912 Main_Page_en.wikipedia.org_all-access_a...

#### MOST TRENDING FOR NOVEMBER 08

```
In [33]: Nov_wiki = wiki_tidy[wiki_tidy['Month'] == 'November']
Nov_wiki
```

Out[33]:

	Search_Topics	Dates	Daily_Page_Visit_Count	
457500	.xxx_en.wikipedia.org_all-access_all-agents	2016-11-01	4788.0	No
457501	.xxx_en.wikipedia.org_mobile-web_all-agents	2016-11-01	4210.0	No
457502	1._Juli_de.wikipedia.org_desktop_all-agents	2016-11-01	45.0	No
457503	2014_FIFA_World_Cup_en.wikipedia.org_all-acces...	2016-11-01	3614.0	No
457504	2014_FIFA_World_Cup_en.wikipedia.org_desktop_a...	2016-11-01	1944.0	No
...	...	...	...	
502495	Zac_Efron_en.wikipedia.org_all-access_all-agents	2016-11-30	8551.0	No
502496	Zayn_Malik_en.wikipedia.org_all-access_all-agents	2016-11-30	13840.0	No
502497	Zendaya_en.wikipedia.org_all-access_all-agents	2016-11-30	21273.0	No
502498	Zendaya_en.wikipedia.org_mobile-web_all-agents	2016-11-30	18220.0	No
502499	Zodiac_Killer_en.wikipedia.org_all-access_all-...	2016-11-30	8491.0	No

44970 rows × 5 columns



```
In [35]: ▶ Nov_8 = (Nov_wiki[Nov_wiki['Day']==8])
Nov_8
```

Out[35]:

	Search_Topics	Dates	Daily_Page_Visit_Count	
<b>468000</b>	.xxx_en.wikipedia.org_all-access_all-agents	2016-11-08	3896.0	No
<b>468001</b>	.xxx_en.wikipedia.org_mobile-web_all-agents	2016-11-08	3415.0	No
<b>468002</b>	1._Juli_de.wikipedia.org_desktop_all-agents	2016-11-08	27.0	No
<b>468003</b>	2014_FIFA_World_Cup_en.wikipedia.org_all-acces...	2016-11-08	3703.0	No
<b>468004</b>	2014_FIFA_World_Cup_en.wikipedia.org_desktop_a...	2016-11-08	2075.0	No
...	...	...	...	
<b>469495</b>	Zac_Efron_en.wikipedia.org_all-access_all-agents	2016-11-08	8537.0	No
<b>469496</b>	Zayn_Malik_en.wikipedia.org_all-access_all-agents	2016-11-08	8403.0	No
<b>469497</b>	Zendaya_en.wikipedia.org_all-access_all-agents	2016-11-08	4602.0	No
<b>469498</b>	Zendaya_en.wikipedia.org_mobile-web_all-agents	2016-11-08	2892.0	No
<b>469499</b>	Zodiac_Killer_en.wikipedia.org_all-access_all-...	2016-11-08	10503.0	No

1499 rows × 5 columns





```
In [37]: Nov = pd.DataFrame(Nov_8.groupby(["Daily_Page_Visit_Count"])[ "Search_Topic"
Nov
```

```
Out[37]:
```

	0	1
0	10.0	468621 Guns_N'_Roses_es.wikipedia.org_all-a...
1	20.0	468046 ああ§è¥¿è±ç"._ja.wikipedia.org_deskto...
2	26.0	468338 Copa_AmÃ©rica_2015_fr.wikipedia.org_...
3	27.0	468002 1._Juli_de.wikipedia.org_desktop_all...
4	33.0	468467 è..."â®¤ç—†â€™ç³¼¤_zh.wikipedia.org_des...
...	...	...
1438	2246978.0	469249 Special:Search_en.wikipedia.org_all-...
1439	2829085.0	469461 Wikipedia:Hauptseite_de.wikipedia.or...
1440	4161059.0	468915 Main_Page_en.wikipedia.org_mobile-we...
1441	20367321.0	468914 Main_Page_en.wikipedia.org_desktop_a...
1442	24600978.0	468912 Main_Page_en.wikipedia.org_all-acces...

1443 rows × 2 columns

```
In [38]: df_Nov8 = (Nov.rename(columns = {0: "Daily_Page_Visit_Count", 1: 'Search_T
df_Nov8
```

```
Out[38]:
```

	Daily_Page_Visit_Count	Search_Topics
0	10.0	468621 Guns_N'_Roses_es.wikipedia.org_all-a...
1	20.0	468046 ああ§è¥¿è±ç"._ja.wikipedia.org_deskto...
2	26.0	468338 Copa_AmÃ©rica_2015_fr.wikipedia.org_...
3	27.0	468002 1._Juli_de.wikipedia.org_desktop_all...
4	33.0	468467 è..."â®¤ç—†â€™ç³¼¤_zh.wikipedia.org_des...
...	...	...
1438	2246978.0	469249 Special:Search_en.wikipedia.org_all-...
1439	2829085.0	469461 Wikipedia:Hauptseite_de.wikipedia.or...
1440	4161059.0	468915 Main_Page_en.wikipedia.org_mobile-we...
1441	20367321.0	468914 Main_Page_en.wikipedia.org_desktop_a...
1442	24600978.0	468912 Main_Page_en.wikipedia.org_all-acces...

1443 rows × 2 columns

In [39]: `df_Nov8.tail(10)`

Out[39]:

	Daily_Page_Visit_Count	Search_Topics
1433	1182277.0	469464 Wikipedia:Portada_es.wikipedia.org_a...
1434	1379041.0	468408 Ð—Ð°Ð³Ð»»Ð°Ð²Ð½Ñ_Ñ_Ñ,Ñ€Ð½Ð,Ñ†Ð°_...
1435	1464684.0	469447 WikipÃ©dia:Accueil_principal_fr.wiki...
1436	1796203.0	469250 Special:Search_en.wikipedia.org_desk...
1437	1796959.0	469463 Wikipedia:Hauptseite_de.wikipedia.or...
1438	2246978.0	469249 Special:Search_en.wikipedia.org_all-...
1439	2829085.0	469461 Wikipedia:Hauptseite_de.wikipedia.or...
1440	4161059.0	468915 Main_Page_en.wikipedia.org_mobile-we...
1441	20367321.0	468914 Main_Page_en.wikipedia.org_desktop_a...
1442	24600978.0	468912 Main_Page_en.wikipedia.org_all-acces...