

```
In [1]: import pandas as pd;
import numpy as np;
```

```
In [2]: data = pd.read_csv('topSubscribed.csv')
```

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

Out[3]:

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Started
0	1	T-Series	234000000	2.129003e+11	18515.0	Music	2006
1	2	YouTube Movies	161000000	NaN	NaN	Film & Animation	2015
2	3	Cocomelon - Nursery Rhymes	152000000	1.490842e+11	846.0	Education	2006
3	4	SET India	150000000	1.378281e+11	103200.0	Shows	2006
4	5	MrBeast	128000000	2.154913e+10	733.0	Entertainment	2012

```
In [4]: data.head(10)
```

Out[4]:

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Sta
0	1	T-Series	234000000	2.129003e+11	18515.0	Music	
1	2	YouTube Movies	161000000	NaN	NaN	Film & Animation	
2	3	Cocomelon - Nursery Rhymes	152000000	1.490842e+11	846.0	Education	
3	4	SET India	150000000	1.378281e+11	103200.0	Shows	
4	5	MrBeast	128000000	2.154913e+10	733.0	Entertainment	
5	6	Music	118000000	NaN	NaN	https://us.youtubers.me/global/all/top-1000-mo...	
6	7	PewDiePie	111000000	2.885188e+10	4694.0	Gaming	
7	8	✿ Kids Diana Show	106000000	8.663857e+10	1056.0	People & Blogs	
8	9	Like Nastya	104000000	8.720294e+10	754.0	People & Blogs	
9	10	Gaming	93300000	NaN	NaN	https://us.youtubers.me/global/all/top-1000-mo...	

In [5]: `data.isnull()`

Out[5]:

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Started
0	False	False	False	False	False	False	False
1	False	False	False	True	True	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False
...
995	False	False	False	False	False	False	False
996	False	False	False	False	False	False	False
997	False	False	False	False	False	False	False
998	False	False	False	False	False	False	False
999	False	False	False	False	False	False	False

1000 rows × 7 columns

In [6]: `data.isnull().count()`

Out[6]:

Rank	1000
Youtube Channel	1000
Subscribers	1000
Video Views	1000
Video Count	1000
Category	1000
Started	1000

dtype: int64

In [7]: `data.isnull().sum()`

Out[7]:

Rank	0
Youtube Channel	0
Subscribers	0
Video Views	9
Video Count	9
Category	0
Started	0

dtype: int64

In [8]: `data.count()`

```
Out[8]: Rank          1000
Youtube Channel      1000
Subscribers          1000
Video Views          991
Video Count          991
Category             1000
Started              1000
dtype: int64
```

In [9]: `data.dtypes`

```
Out[9]: Rank          int64
Youtube Channel      object
Subscribers          int64
Video Views          float64
Video Count          float64
Category             object
Started              int64
dtype: object
```

In [10]: `data.size`

```
Out[10]: 7000
```

In [11]: `data.shape`

```
Out[11]: (1000, 7)
```

In [12]: `data.ndim`

```
Out[12]: 2
```

In [13]: `data.describe()`

```
Out[13]:
```

	Rank	Subscribers	Video Views	Video Count	Started
count	1000.000000	1.000000e+03	9.910000e+02	991.000000	1000.000000
mean	500.500000	2.158140e+07	1.008568e+10	9501.743693	2012.594000
std	288.819436	1.662556e+07	1.302932e+10	32324.321451	4.110238
min	1.000000	1.140000e+07	4.390980e+05	1.000000	1970.000000
25%	250.750000	1.340000e+07	3.959370e+09	380.000000	2010.000000
50%	500.500000	1.660000e+07	6.767680e+09	906.000000	2013.000000
75%	750.250000	2.320000e+07	1.226927e+10	3311.000000	2015.000000
max	1000.000000	2.340000e+08	2.129003e+11	342802.000000	2021.000000

In [14]: data.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 7 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Rank                  1000 non-null  int64
1   Youtube Channel      1000 non-null  object
2   Subscribers          1000 non-null  int64
3   Video Views          991 non-null   float64
4   Video Count          991 non-null   float64
5   Category              1000 non-null  object
6   Started              1000 non-null  int64
dtypes: float64(2), int64(3), object(2)
memory usage: 54.8+ KB
```

In [15]: data.head(15)

Out[15]:

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category
0	1	T-Series	234000000	2.129003e+11	18515.0	Music
1	2	YouTube Movies	161000000	NaN	NaN	Film & Animation
2	3	Cocomelon - Nursery Rhymes	152000000	1.490842e+11	846.0	Education
3	4	SET India	150000000	1.378281e+11	103200.0	Shows
4	5	MrBeast	128000000	2.154913e+10	733.0	Entertainment
5	6	Music	118000000	NaN	NaN	https://us.youtubers.me/global/all/top-1000-mo...
6	7	PewDiePie	111000000	2.885188e+10	4694.0	Gaming
7	8	🌸 Kids Diana Show	106000000	8.663857e+10	1056.0	People & Blogs
8	9	Like Nastya	104000000	8.720294e+10	754.0	People & Blogs
9	10	Gaming	93300000	NaN	NaN	https://us.youtubers.me/global/all/top-1000-mo...
10	11	WWE	92800000	7.357770e+10	66174.0	Sports
11	12	Vlad and Niki	92700000	7.244412e+10	520.0	Entertainment
12	13	Zee Music Company	91600000	5.335260e+10	7530.0	Music
13	14	BLACKPINK	83900000	2.872329e+10	468.0	People & Blogs
14	15	Goldmines	81200000	2.129657e+10	5090.0	Film & Animation

In [16]: `data.isna()`

Out[16]:

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Started
0	False	False	False	False	False	False	False
1	False	False	False	True	True	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False
...
995	False	False	False	False	False	False	False
996	False	False	False	False	False	False	False
997	False	False	False	False	False	False	False
998	False	False	False	False	False	False	False
999	False	False	False	False	False	False	False

1000 rows × 7 columns

In [17]: `data.isna().sum()`

Out[17]:

Rank	0
Youtube Channel	0
Subscribers	0
Video Views	9
Video Count	9
Category	0
Started	0
dtype:	int64

In [18]: `data.isnull().sum()`

Out[18]:

Rank	0
Youtube Channel	0
Subscribers	0
Video Views	9
Video Count	9
Category	0
Started	0
dtype:	int64

```
In [19]: data.isnull()
```

```
Out[19]:
```

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Started
0	False	False	False	False	False	False	False
1	False	False	False	True	True	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False
...
995	False	False	False	False	False	False	False
996	False	False	False	False	False	False	False
997	False	False	False	False	False	False	False
998	False	False	False	False	False	False	False
999	False	False	False	False	False	False	False

1000 rows × 7 columns

```
In [20]: data.isnull().sum()
```

```
Out[20]: Rank          0
Youtube Channel      0
Subscribers          0
Video Views          9
Video Count          9
Category             0
Started              0
dtype: int64
```

```
In [21]: data.replace("music", "dance")
```

```
Out[21]:
```

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Started
0	1	T-Series	234000000	2.129003e+11	18515.0	Music	2006
1	2	YouTube Movies	161000000	NaN	NaN	Film & Animation	2015
2	3	Cocomelon - Nursery Rhymes	152000000	1.490842e+11	846.0	Education	2006
3	4	SET India	150000000	1.378281e+11	103200.0	Shows	2006
4	5	MrBeast	128000000	2.154913e+10	733.0	Entertainment	2012
...
995	996	XO TEAM	11500000	1.147138e+10	2679.0	People & Blogs	2021
996	997	Nirahua Music World	11500000	5.997718e+09	788.0	Music	2014
997	998	YtCrash	11400000	3.825084e+06	1.0	Education	2015
998	999	Painzeiro	11400000	2.143444e+09	946.0	People & Blogs	2015
999	1000	Oum Walid	11400000	2.480065e+09	1598.0	People & Blogs	2015

1000 rows × 7 columns

```
In [22]: data.replace("Music", "Dance")
```

```
Out[22]:
```

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Started
0	1	T-Series	234000000	2.129003e+11	18515.0	Dance	2006
1	2	YouTube Movies	161000000	NaN	NaN	Film & Animation	2015
2	3	Cocomelon - Nursery Rhymes	152000000	1.490842e+11	846.0	Education	2006
3	4	SET India	150000000	1.378281e+11	103200.0	Shows	2006
4	5	MrBeast	128000000	2.154913e+10	733.0	Entertainment	2012
...
995	996	XO TEAM	11500000	1.147138e+10	2679.0	People & Blogs	2021
996	997	Nirahua Music World	11500000	5.997718e+09	788.0	Dance	2014
997	998	YtCrash	11400000	3.825084e+06	1.0	Education	2015
998	999	Painzeiro	11400000	2.143444e+09	946.0	People & Blogs	2015
999	1000	Oum Walid	11400000	2.480065e+09	1598.0	People & Blogs	2015

1000 rows × 7 columns

```
In [23]: data.dropna(axis = 0)
```

```
Out[23]:
```

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Started
0	1	T-Series	234000000	2.129003e+11	18515.0	Music	2006
2	3	Cocomelon - Nursery Rhymes	152000000	1.490842e+11	846.0	Education	2006
3	4	SET India	150000000	1.378281e+11	103200.0	Shows	2006
4	5	MrBeast	128000000	2.154913e+10	733.0	Entertainment	2012
6	7	PewDiePie	111000000	2.885188e+10	4694.0	Gaming	2010
...
995	996	XO TEAM	11500000	1.147138e+10	2679.0	People & Blogs	2021
996	997	Nirahua Music World	11500000	5.997718e+09	788.0	Music	2014
997	998	YtCrash	11400000	3.825084e+06	1.0	Education	2015
998	999	Painzeiro	11400000	2.143444e+09	946.0	People & Blogs	2015
999	1000	Oum Walid	11400000	2.480065e+09	1598.0	People & Blogs	2015

991 rows × 7 columns

In [24]:

data.head(10)

Out[24]:

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Str
0	1	T-Series	234000000	2.129003e+11	18515.0	Music	
1	2	YouTube Movies	161000000	NaN	NaN	Film & Animation	
2	3	Cocomelon - Nursery Rhymes	152000000	1.490842e+11	846.0	Education	
3	4	SET India	150000000	1.378281e+11	103200.0	Shows	
4	5	MrBeast	128000000	2.154913e+10	733.0	Entertainment	
5	6	Music	118000000	NaN	NaN	https://us.youtubers.me/global/all/top-1000-mo...	
6	7	PewDiePie	111000000	2.885188e+10	4694.0		Gaming
7	8	✿ Kids Diana Show	106000000	8.663857e+10	1056.0	People & Blogs	
8	9	Like Nastya	104000000	8.720294e+10	754.0	People & Blogs	
9	10	Gaming	93300000	NaN	NaN	https://us.youtubers.me/global/all/top-1000-mo...	

```
In [25]: data.fillna(data['Rank'].mean())
```

```
Out[25]:
```

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Started
0	1	T-Series	234000000	2.129003e+11	18515.0	Music	2006
1	2	YouTube Movies	161000000	5.005000e+02	500.5	Film & Animation	2015
2	3	Cocomelon - Nursery Rhymes	152000000	1.490842e+11	846.0	Education	2006
3	4	SET India	150000000	1.378281e+11	103200.0	Shows	2006
4	5	MrBeast	128000000	2.154913e+10	733.0	Entertainment	2012
...
995	996	XO TEAM	11500000	1.147138e+10	2679.0	People & Blogs	2021
996	997	Nirahua Music World	11500000	5.997718e+09	788.0	Music	2014
997	998	YtCrash	11400000	3.825084e+06	1.0	Education	2015
998	999	Painzeiro	11400000	2.143444e+09	946.0	People & Blogs	2015
999	1000	Oum Walid	11400000	2.480065e+09	1598.0	People & Blogs	2015

1000 rows × 7 columns

```
In [26]: data.dtypes
```

```
Out[26]: Rank                int64
Youtube Channel            object
Subscribers                int64
Video Views                float64
Video Count                float64
Category                   object
Started                    int64
dtype: object
```

In [27]: `data.convert_dtypes()`

Out[27]:

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Started
0	1	T-Series	234000000	212900271553	18515	Music	2006
1	2	YouTube Movies	161000000	<NA>	<NA>	Film & Animation	2015
2	3	Cocomelon - Nursery Rhymes	152000000	149084178448	846	Education	2006
3	4	SET India	150000000	137828094104	103200	Shows	2006
4	5	MrBeast	128000000	21549128785	733	Entertainment	2012
...
995	996	XO TEAM	11500000	11471384709	2679	People & Blogs	2021
996	997	Nirahua Music World	11500000	5997717755	788	Music	2014
997	998	YtCrash	11400000	3825084	1	Education	2015
998	999	Painzeiro	11400000	2143444281	946	People & Blogs	2015
999	1000	Oum Walid	11400000	2480064722	1598	People & Blogs	2015

1000 rows × 7 columns

In [28]: `data.dtypes`

```
Out[28]: Rank                int64
Youtube Channel            object
Subscribers                int64
Video Views               float64
Video Count               float64
Category                  object
Started                   int64
dtype: object
```

In [29]: `data.isnull().sum()`

```
Out[29]: Rank                0
Youtube Channel            0
Subscribers                0
Video Views                9
Video Count               9
Category                  0
Started                   0
dtype: int64
```

```
In [30]: data.astype({'Rank': 'int32'}).dtypes
```

```
Out[30]: Rank                int32
Youtube Channel            object
Subscribers                int64
Video Views                float64
Video Count                float64
Category                   object
Started                    int64
dtype: object
```

```
In [31]: data.astype({'Started': 'int32'}).dtypes
```

```
Out[31]: Rank                int64
Youtube Channel            object
Subscribers                int64
Video Views                float64
Video Count                float64
Category                   object
Started                    int32
dtype: object
```

```
In [32]: data.isnull()
```

```
Out[32]:
```

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Started
0	False	False	False	False	False	False	False
1	False	False	False	True	True	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False
...
995	False	False	False	False	False	False	False
996	False	False	False	False	False	False	False
997	False	False	False	False	False	False	False
998	False	False	False	False	False	False	False
999	False	False	False	False	False	False	False

1000 rows × 7 columns

```
In [33]: mv = data['Video Views'].mean()
mv
```

```
Out[33]: 10085683561.790112
```

In [34]: `data.head(20)`

Out[34]:

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category
0	1	T-Series	234000000	2.129003e+11	18515.0	Music
1	2	YouTube Movies	161000000	NaN	NaN	Film & Animation
2	3	Cocomelon - Nursery Rhymes	152000000	1.490842e+11	846.0	Education
3	4	SET India	150000000	1.378281e+11	103200.0	Shows
4	5	MrBeast	128000000	2.154913e+10	733.0	Entertainment
5	6	Music	118000000	NaN	NaN	https://us.youtubers.me/global/all/top-1000-mo...
6	7	PewDiePie	111000000	2.885188e+10	4694.0	Gaming
7	8	✿ Kids Diana Show	106000000	8.663857e+10	1056.0	People & Blogs
8	9	Like Nastya	104000000	8.720294e+10	754.0	People & Blogs
9	10	Gaming	93300000	NaN	NaN	https://us.youtubers.me/global/all/top-1000-mo...
10	11	WWE	92800000	7.357770e+10	66174.0	Sports
11	12	Vlad and Niki	92700000	7.244412e+10	520.0	Entertainment
12	13	Zee Music Company	91600000	5.335260e+10	7530.0	Music
13	14	BLACKPINK	83900000	2.872329e+10	468.0	People & Blogs
14	15	Goldmines	81200000	2.129657e+10	5090.0	Film & Animation
15	16	5-Minute Crafts	78500000	2.498343e+10	5937.0	Howto & Style
16	17	Sony SAB	76400000	9.111729e+10	63390.0	Shows
17	18	Sports	75100000	NaN	NaN	https://us.youtubers.me/global/all/top-1000-mo...
18	19	BANGTANTV	73100000	1.924006e+10	2090.0	Music
19	20	Justin Bieber	70800000	2.959865e+10	249.0	Music

In [35]: `d = data.fillna(mv)`

In [36]:

d

Out[36]:

	Rank	Youtube Channel	Subscribers	Video Views	Video Count	Category	Started
0	1	T-Series	234000000	2.129003e+11	1.851500e+04	Music	2006
1	2	YouTube Movies	161000000	1.008568e+10	1.008568e+10	Film & Animation	2015
2	3	Cocomelon - Nursery Rhymes	152000000	1.490842e+11	8.460000e+02	Education	2006
3	4	SET India	150000000	1.378281e+11	1.032000e+05	Shows	2006
4	5	MrBeast	128000000	2.154913e+10	7.330000e+02	Entertainment	2012
...
995	996	XO TEAM	11500000	1.147138e+10	2.679000e+03	People & Blogs	2021
996	997	Nirahua Music World	11500000	5.997718e+09	7.880000e+02	Music	2014
997	998	YtCrash	11400000	3.825084e+06	1.000000e+00	Education	2015
998	999	Painzeiro	11400000	2.143444e+09	9.460000e+02	People & Blogs	2015
999	1000	Oum Walid	11400000	2.480065e+09	1.598000e+03	People & Blogs	2015

1000 rows × 7 columns

In [37]: d.isnull().sum()

```
Out[37]: Rank          0
Youtube Channel      0
Subscribers         0
Video Views         0
Video Count         0
Category            0
Started             0
dtype: int64
```

In [38]: d.dtypes

```
Out[38]: Rank          int64
Youtube Channel      object
Subscribers         int64
Video Views         float64
Video Count         float64
Category            object
Started             int64
dtype: object
```

In []: