

# India Census 2011 Dataset

```
In [1]: # The data is about total population, Demography, Literacy, Districts,
# States, Workers, Religion, Education, Age
```

```
In [2]: import pandas as pd
```

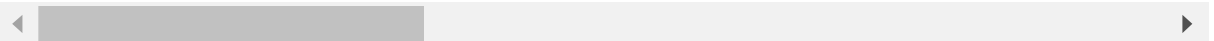
```
In [3]: data = pd.read_csv('Censes.csv')
```

```
In [4]: data
```

```
Out[4]:
```

	District_code	State_name	District_name	Population	Male	Female	Literate	Workers
0	1	JAMMU AND KASHMIR	Kupwara	870354	474190	396164	439654	229064
1	2	JAMMU AND KASHMIR	Badgam	753745	398041	355704	335649	214866
2	3	JAMMU AND KASHMIR	Leh(Ladakh)	133487	78971	54516	93770	75079
3	4	JAMMU AND KASHMIR	Kargil	140802	77785	63017	86236	51873
4	5	JAMMU AND KASHMIR	Punch	476835	251899	224936	261724	161393
...	...	...	...	...	...	...	...	...
635	636	PONDICHERRY	Mahe	41816	19143	22673	36470	11802
636	637	PONDICHERRY	Karaikal	200222	97809	102413	154916	68301
637	638	ANDAMAN AND NICOBAR ISLANDS	Nicobars	36842	20727	16115	25332	17125
638	639	ANDAMAN AND NICOBAR ISLANDS	North AND Middle Andaman	105597	54861	50736	78683	38579
639	640	ANDAMAN AND NICOBAR ISLANDS	South Andaman	238142	127283	110859	190266	96831

640 rows × 25 columns



## 1. How will you hide the indexes of the dataframe.

In [5]: `#df.style.hide_index()`

In [6]: `data.style.hide_index()`

C:\Users\91958\AppData\Local\Temp\ipykernel\_9608\1655106480.py:1: FutureWarning: this method is deprecated in favour of `Styler.hide(axis="index")`  
`data.style.hide_index()`

Out[6]:

	District_code	State_name	District_name	Population	Male	Female	Literate	V
	1	JAMMU AND KASHMIR	Kupwara	870354	474190	396164	439654	
	2	JAMMU AND KASHMIR	Badgam	753745	398041	355704	335649	
	3	JAMMU AND KASHMIR	Leh(Ladakh)	133487	78971	54516	93770	
	4	JAMMU AND KASHMIR	Kargil	140802	77785	63017	86236	
	5	JAMMU AND KASHMIR	Punch	476835	251899	224936	261724	
	6	JAMMU AND KASHMIR	Rajouri	642415	345351	297064	364100	

## 2.How can we set the caption/heading on the dataframe.

In [7]: `#df.style.set_caption('Caption')`

In [8]: `data.style.set_caption('India Census 2011 Dataset')`

Out[8]:

India Census 2011 Dataset

	District_code	State_name	District_name	Population	Male	Female	Literate
0	1	JAMMU AND KASHMIR	Kupwara	870354	474190	396164	439654
1	2	JAMMU AND KASHMIR	Badgam	753745	398041	355704	335649
2	3	JAMMU AND KASHMIR	Leh(Ladakh)	133487	78971	54516	93770
3	4	JAMMU AND KASHMIR	Kargil	140802	77785	63017	86236
4	5	JAMMU AND KASHMIR	Punch	476835	251899	224936	261724
5	6	JAMMU AND KASHMIR	Rajouri	642415	345351	297064	364100

### 3. Show the records related with the district - New Delhi, Lucknow, Jaipur

```
In [9]: #df[df['Col_name'].isin(['Item_1', 'Item_2', 'Item_3'])]
```

```
In [10]: data[data['District_name'].isin(['New Delhi', 'Lucknow', 'Jaipur'])]
```

```
Out[10]:
```

	District_code	State_name	District_name	Population	Male	Female	Literate	Workers
93	94	NCT OF DELHI	New Delhi	142004	77942	64062	114179	59541
109	110	RAJASTHAN	Jaipur	6626178	3468507	3157671	4300965	2464893
156	157	UTTAR PRADESH	Lucknow	4589838	2394476	2195362	3127260	1542806

3 rows × 25 columns

### 4. Calculate state-wise:

#### A. Total number of population.

#### B. Total no. of population with different religion.

```
In [11]: data.head(2)
```

```
Out[11]:
```

	District_code	State_name	District_name	Population	Male	Female	Literate	Workers	Mal
0	1	JAMMU AND KASHMIR	Kupwara	870354	474190	396164	439654	229064	
1	2	JAMMU AND KASHMIR	Badgam	753745	398041	355704	335649	214866	

2 rows × 25 columns

```
In [12]: data.groupby('State_name').Population.sum().sort_values(ascending = True)
```

```
Out[12]: State_name
LAKSHADWEEP                64473
DAMAN AND DIU              243247
DADRA AND NAGAR HAVELI     343709
ANDAMAN AND NICOBAR ISLANDS 380581
SIKKIM                     610577
CHANDIGARH                 1055450
MIZORAM                    1097206
PONDICHERRY                1247953
ARUNACHAL PRADESH          1383727
GOA                        1458545
NAGALAND                   1978502
MANIPUR                    2855794
MEGHALAYA                  2966889
TRIPURA                   3673917
HIMACHAL PRADESH           6864602
UTTARAKHAND                10086292
JAMMU AND KASHMIR          12541302
NCT OF DELHI               16787941
HARYANA                    25351462
CHHATTISGARH              25545198
PUNJAB                     27743338
ASSAM                      31205576
JHARKHAND                  32988134
KERALA                     33406061
ORISSA                     41974218
GUJARAT                    60439692
KARNATAKA                  61095297
RAJASTHAN                  68548437
TAMIL NADU                 72147030
MADHYA PRADESH             72626809
ANDHRA PRADESH             84580777
WEST BENGAL                91276115
BIHAR                     104099452
MAHARASHTRA                112374333
UTTAR PRADESH              199812341
Name: Population, dtype: int64
```

```
In [13]: data.columns
```

```
Out[13]: Index(['District_code', 'State_name', 'District_name', 'Population', 'Male',
                'Female', 'Literate', 'Workers', 'Male_Workers', 'Female_Workers',
                'Cultivator_Workers', 'Agricultural_Workers', 'Household_Workers',
                'Hindus', 'Muslims', 'Christians', 'Sikhs', 'Buddhists', 'Jains',
                'Secondary_Education', 'Higher_Education', 'Graduate_Education',
                'Age_Group_0_29', 'Age_Group_30_49', 'Age_Group_50'],
                dtype='object')
```

```
In [14]: #B.  
data.groupby('State_name')['Hindus', 'Muslims', 'Christians', 'Sikhs', 'Buddhis  
C:\Users\91958\AppData\Local\Temp\ipykernel_9608\87213495.py:2: FutureWarnin  
g: Indexing with multiple keys (implicitly converted to a tuple of keys) will  
be deprecated, use a list instead.  
    data.groupby('State_name')['Hindus', 'Muslims', 'Christians', 'Sikhs', 'Budd  
hists', 'Jains'].sum()
```

Out[14]:

	Hindus	Muslims	Christians	Sikhs	Buddhists	Jains
State_name						
ANDAMAN AND NICOBAR ISLANDS	264296	32413	80984	1286	338	31
ANDHRA PRADESH	74824149	8082412	1129784	40244	36692	53849
ARUNACHAL PRADESH	401876	27045	418732	3287	162815	771
ASSAM	19180759	10679345	1165867	20672	54993	25949
BIHAR	86078686	17557809	129247	23779	25453	18914
CHANDIGARH	852574	51447	8720	138329	1160	1960
CHHATTISGARH	23819789	514998	490542	70036	70467	61510
DADRA AND NAGAR HAVELI	322857	12922	5113	217	634	1186
DAMAN AND DIU	220150	19277	2820	172	217	287
GOA	963877	121564	366130	1473	1095	1109
GUJARAT	53533988	5846761	316178	58246	30483	579654
HARYANA	22171128	1781342	50353	1243752	7514	52613
HIMACHAL PRADESH	6532765	149881	12646	79896	78659	1805
JAMMU AND KASHMIR	3566674	8567485	35631	234848	112584	2490
JHARKHAND	22376051	4793994	1418608	71422	8956	14974
KARNATAKA	51317472	7893065	1142647	28773	95710	440280
KERALA	18282492	8873472	6141269	3814	4752	4489
LAKSHADWEEP	1788	62268	317	8	10	11
MADHYA PRADESH	66007121	4774695	213282	151412	216052	567028
MAHARASHTRA	89703057	12971152	1080073	223247	6531200	1400349
MANIPUR	1181876	239836	1179043	1527	7084	1692
MEGHALAYA	342078	130399	2213027	3045	9864	627
MIZORAM	30136	14832	956331	286	93411	376
NAGALAND	173054	48963	1739651	1890	6759	2655
NCT OF DELHI	13712100	2158684	146093	570581	18449	166231
ORISSA	39300341	911670	1161708	21991	13852	9420
PONDICHERY	1089409	75556	78550	297	451	1400
PUNJAB	10678138	535489	348230	16004754	33237	45040
RAJASTHAN	60657103	6215377	96430	872930	12185	622023
SIKKIM	352662	9867	60522	1868	167216	314
TAMIL NADU	63188168	4229479	4418331	14601	11186	89265
TRIPURA	3063903	316042	159882	1070	125385	860
UTTAR PRADESH	159312654	38483967	356448	643500	206285	213267
UTTARAKHAND	8368636	1406825	37781	236340	14926	9183

	Hindus	Muslims	Christians	Sikhs	Buddhists	Jains
State_name						
WEST BENGAL	64385546	24654825	658618	63523	282898	60141

```
In [15]: data.groupby('State_name')['Hindus', 'Muslims', 'Christians', 'Sikhs', 'Buddhis
```

C:\Users\91958\AppData\Local\Temp\ipykernel\_9608\1475905639.py:1: FutureWarning: Indexing with multiple keys (implicitly converted to a tuple of keys) will be deprecated, use a list instead.

```
data.groupby('State_name')['Hindus', 'Muslims', 'Christians', 'Sikhs', 'Buddhists', 'Jains'].sum().sort_values('Hindus')
```



Out[15]:

	Hindus	Muslims	Christians	Sikhs	Buddhists	Jains
State_name						
LAKSHADWEEP	1788	62268	317	8	10	11
MIZORAM	30136	14832	956331	286	93411	376
NAGALAND	173054	48963	1739651	1890	6759	2655
DAMAN AND DIU	220150	19277	2820	172	217	287
ANDAMAN AND NICOBAR ISLANDS	264296	32413	80984	1286	338	31
DADRA AND NAGAR HAVELI	322857	12922	5113	217	634	1186
MEGHALAYA	342078	130399	2213027	3045	9864	627
SIKKIM	352662	9867	60522	1868	167216	314
ARUNACHAL PRADESH	401876	27045	418732	3287	162815	771
CHANDIGARH	852574	51447	8720	138329	1160	1960
GOA	963877	121564	366130	1473	1095	1109
PONDICHERRY	1089409	75556	78550	297	451	1400
MANIPUR	1181876	239836	1179043	1527	7084	1692
TRIPURA	3063903	316042	159882	1070	125385	860
JAMMU AND KASHMIR	3566674	8567485	35631	234848	112584	2490
HIMACHAL PRADESH	6532765	149881	12646	79896	78659	1805
UTTARAKHAND	8368636	1406825	37781	236340	14926	9183
PUNJAB	10678138	535489	348230	16004754	33237	45040
NCT OF DELHI	13712100	2158684	146093	570581	18449	166231
KERALA	18282492	8873472	6141269	3814	4752	4489
ASSAM	19180759	10679345	1165867	20672	54993	25949
HARYANA	22171128	1781342	50353	1243752	7514	52613
JHARKHAND	22376051	4793994	1418608	71422	8956	14974
CHHATTISGARH	23819789	514998	490542	70036	70467	61510
ORISSA	39300341	911670	1161708	21991	13852	9420
KARNATAKA	51317472	7893065	1142647	28773	95710	440280
GUJARAT	53533988	5846761	316178	58246	30483	579654
RAJASTHAN	60657103	6215377	96430	872930	12185	622023
TAMIL NADU	63188168	4229479	4418331	14601	11186	89265
WEST BENGAL	64385546	24654825	658618	63523	282898	60141
MADHYA PRADESH	66007121	4774695	213282	151412	216052	567028
ANDHRA PRADESH	74824149	8082412	1129784	40244	36692	53849
BIHAR	86078686	17557809	129247	23779	25453	18914
MAHARASHTRA	89703057	12971152	1080073	223247	6531200	1400349

Hindus   Muslims   Christians   Sikhs   Buddhists   Jains

State_name									
UTTAR PRADESH	159312654	38483967	356448	643500	206285	213267			

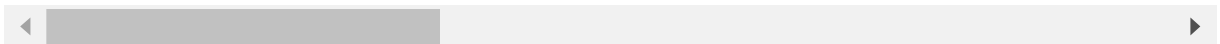
## 5. How many Male workers were there in Maharashtra state?

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

Out[16]:

	District_code	State_name	District_name	Population	Male	Female	Literate	Workers	Mal
0	1	JAMMU AND KASHMIR	Kupwara	870354	474190	396164	439654	229064	
1	2	JAMMU AND KASHMIR	Badgam	753745	398041	355704	335649	214866	
2	3	JAMMU AND KASHMIR	Leh(Ladakh)	133487	78971	54516	93770	75079	
3	4	JAMMU AND KASHMIR	Kargil	140802	77785	63017	86236	51873	
4	5	JAMMU AND KASHMIR	Punch	476835	251899	224936	261724	161393	

5 rows × 25 columns



In [17]: `# df[df.Col_name_1 == 'Item']['Col_name_2'].sum()  
# data[data.State_name == 'Maharashtra']['Male_Workers'].sum()`

In [21]: `data[data.State_name == 'MAHARASHTRA']['Male_Workers'].sum()`

Out[21]: 32616875

In [26]: `data[data.State_name == 'UTTAR PRADESH'].Female_Workers.sum()`

Out[26]: 15967953

## 6. How to set a column as index of the dataframe?

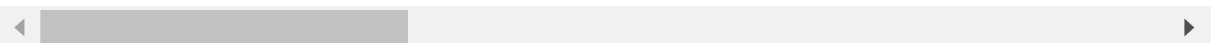
```
In [27]: # df = df.set_index('col_name')
```

```
In [30]: data.set_index('District_code')
```

```
Out[30]:
```

	State_name	District_name	Population	Male	Female	Literate	Workers	Mal
District_code								
1	JAMMU AND KASHMIR	Kupwara	870354	474190	396164	439654	229064	
2	JAMMU AND KASHMIR	Badgam	753745	398041	355704	335649	214866	
3	JAMMU AND KASHMIR	Leh(Ladakh)	133487	78971	54516	93770	75079	
4	JAMMU AND KASHMIR	Kargil	140802	77785	63017	86236	51873	
5	JAMMU AND KASHMIR	Punch	476835	251899	224936	261724	161393	
...	...	...	...	...	...	...	...	...
636	PONDICHERRY	Mahe	41816	19143	22673	36470	11802	
637	PONDICHERRY	Karaikal	200222	97809	102413	154916	68301	
638	ANDAMAN AND NICOBAR ISLANDS	Nicobars	36842	20727	16115	25332	17125	
639	ANDAMAN AND NICOBAR ISLANDS	North AND Middle Andaman	105597	54861	50736	78683	38579	
640	ANDAMAN AND NICOBAR ISLANDS	South Andaman	238142	127283	110859	190266	96831	

640 rows × 24 columns



## 7a. Add a Suffix to the column names.

## 7b. Add a Prefix to the column names.

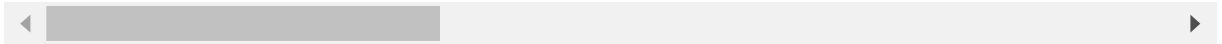
```
In [31]: # df = df.add_suffix('_value')
# df = df.add_prefix('value_')
```

In [32]: data.head(2)

Out[32]:

	District_code	State_name	District_name	Population	Male	Female	Literate	Workers	Mal
0	1	JAMMU AND KASHMIR	Kupwara	870354	474190	396164	439654	229064	
1	2	JAMMU AND KASHMIR	Badgam	753745	398041	355704	335649	214866	

2 rows × 25 columns

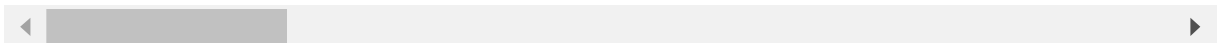


In [33]: data.add\_suffix('\_rightone')

Out[33]:

	District_code_rightone	State_name_rightone	District_name_rightone	Population_rightone	M
0	1	JAMMU AND KASHMIR	Kupwara	870354	
1	2	JAMMU AND KASHMIR	Badgam	753745	
2	3	JAMMU AND KASHMIR	Leh(Ladakh)	133487	
3	4	JAMMU AND KASHMIR	Kargil	140802	
4	5	JAMMU AND KASHMIR	Punch	476835	
...	...	...	...	...	...
635	636	PONDICHERRY	Mahe	41816	
636	637	PONDICHERRY	Karaikal	200222	
637	638	ANDAMAN AND NICOBAR ISLANDS	Nicobars	36842	
638	639	ANDAMAN AND NICOBAR ISLANDS	North AND Middle Andaman	105597	
639	640	ANDAMAN AND NICOBAR ISLANDS	South Andaman	238142	

640 rows × 25 columns

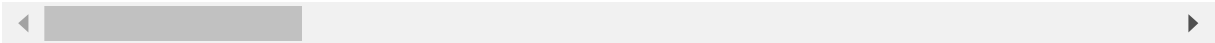


```
In [34]: data.add_prefix('leftone_')
```

Out[34]:

	leftone_District_code	leftone_State_name	leftone_District_name	leftone_Population	leftone_
0	1	JAMMU AND KASHMIR	Kupwara	870354	47
1	2	JAMMU AND KASHMIR	Badgam	753745	39
2	3	JAMMU AND KASHMIR	Leh(Ladakh)	133487	7
3	4	JAMMU AND KASHMIR	Kargil	140802	7
4	5	JAMMU AND KASHMIR	Punch	476835	25
...	...	...	...	...	
635	636	PONDICHERRY	Mahe	41816	1
636	637	PONDICHERRY	Karaikal	200222	9
637	638	ANDAMAN AND NICOBAR ISLANDS	Nicobars	36842	2
638	639	ANDAMAN AND NICOBAR ISLANDS	North AND Middle Andaman	105597	5
639	640	ANDAMAN AND NICOBAR ISLANDS	South Andaman	238142	12

640 rows × 25 columns



```
In [ ]:
```