

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

In [5]: populations = pd.read_csv("population.csv")

In [6]: populations

Out[6]:
```

	Country (or dependency)	Population (2020)	Yearly Change	Net Change	Density (P/Km²)	Land Area (Km²)	Migrants (net)	Fert. Rate	Med. Age	Urban Pop %	World Share
0	China	1440297825	0.39%	5540090	153	9388211	-348399.0	1.7	38	61%	18.47%
1	India	1382345085	0.99%	13586631	464	2973190	-532687.0	2.2	28	35%	17.70%
2	United States	331341050	0.59%	1937734	36	9147420	954806.0	1.8	38	83%	4.25%
3	Indonesia	274021604	1.07%	2898047	151	1811570	-98955.0	2.3	30	56%	3.51%
4	Pakistan	221612785	2.00%	4327022	287	770880	-233379.0	3.6	23	35%	2.83%
...	...	...	...	...	...	...	...	...	...	...	...
230	Montserrat	4993	0.06%	3	50	100	NaN	N.A.	N.A.	10%	0.00%
231	Falkland Islands	3497	3.05%	103	0	12170	NaN	N.A.	N.A.	66%	0.00%
232	Niue	1628	0.68%	11	6	260	NaN	N.A.	N.A.	46%	0.00%
233	Tokelau	1360	1.27%	17	136	10	NaN	N.A.	N.A.	0%	0.00%
234	Holy See	801	0.25%	2	2003	0	NaN	N.A.	N.A.	N.A.	0.00%

235 rows × 11 columns

Describing data

```
In [7]: populations.dtypes

Out[7]:
```

Country (or dependency)	object
Population (2020)	int64
Yearly Change	object
Net Change	int64
Density (P/Km²)	int64
Land Area (Km²)	int64
Migrants (net)	float64
Fert. Rate	object
Med. Age	object
Urban Pop %	object
World Share	object
dtype:	object

```
In [10]: populations.index
Out[10]: RangeIndex(start=0, stop=235, step=1)

In [11]: populations.describe()
```

	Population (2020)	Net Change	Density (P/Km²)	Land Area (Km²)	Migrants (net)
count	2.350000e+02	2.350000e+02	235.000000	2.350000e+02	201.000000
mean	3.322744e+07	3.460878e+05	475.770213	5.535918e+05	6.283582
std	1.353034e+08	1.128260e+06	2331.285935	1.687796e+06	123291.887548
min	8.010000e+02	-3.838400e+05	0.000000	0.000000e+00	-653249.000000
25%	3.994905e+05	4.240000e+02	37.000000	2.545000e+03	-10047.000000
50%	5.460109e+06	3.917000e+04	95.000000	7.724000e+04	-852.000000
75%	2.067170e+07	2.496600e+05	239.500000	4.038200e+05	9741.000000
max	1.440298e+09	1.358663e+07	26337.000000	1.637687e+07	954806.000000

```
In [12]: populations.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 235 entries, 0 to 234
Data columns (total 11 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Country (or dependency)               235 non-null   object
1   Population (2020)                     235 non-null   int64
2   Yearly Change                         235 non-null   object
3   Net Change                           235 non-null   int64
4   Density (P/Km²)                       235 non-null   int64
5   Land Area (Km²)                       235 non-null   int64
6   Migrants (net)                        201 non-null   float64
7   Fert. Rate                           235 non-null   object
8   Med. Age                             235 non-null   object
9   Urban Pop %                           235 non-null   object
10  World Share                           235 non-null   object
dtypes: float64(1), int64(4), object(6)
memory usage: 20.3+ KB

In [11]: populations.mean()
```

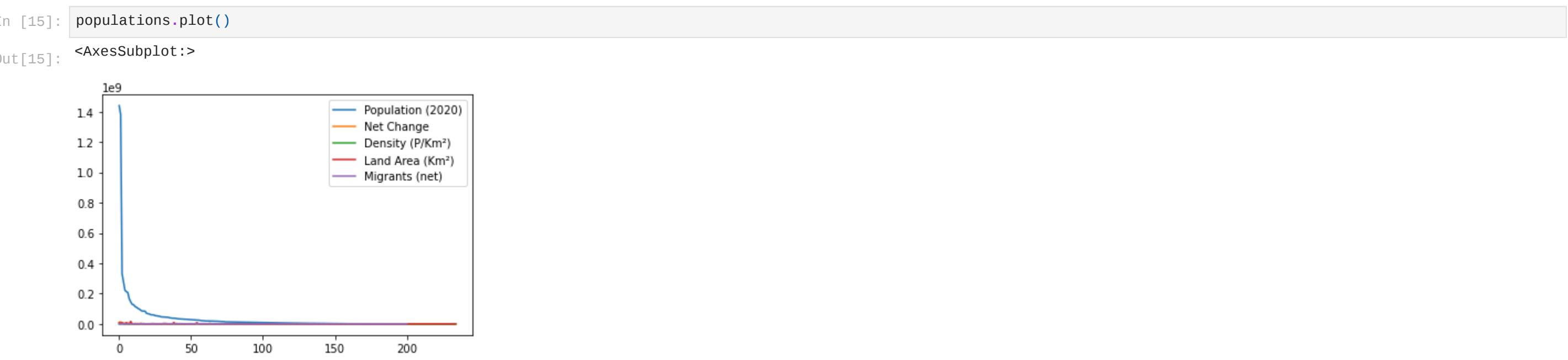
```
C:\Users\Blessing N\AppData\Local\Temp\ipykernel_15588\94686229.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric_only=None') is deprecated, in a future version this will raise TypeError.  Select only valid columns before calling the reduction.
  populations.mean()
```

Population (2020)	3.322744e+07
Net Change	3.460878e+05
Density (P/Km²)	4.757702e+02
Land Area (Km²)	5.535918e+05
Migrants (net)	6.283582e+00
dtype:	float64

```
In [13]: populations.sum()
```

Country (or dependency)	ChinaIndiaUnited StatesIndonesiaPakistanBrazil...
Population (2020)	7808449406
Yearly Change	0.39%0.99%0.59%1.07%2.00%0.72%2.58%1.01%0.04%1...
Net Change	81330639
Density (P/Km²)	111806
Land Area (Km²)	130094083
Migrants (net)	1263.0
Fert. Rate	1.72.21.82.33.61.75.42.11.82.11.44.32.63.32.16...
Med. Age	3828383023331828402948192625321732324640404247...
Urban Pop %	61%35%83%56%35%88%52%39%74%84%92%21%47%43%38%4...
World Share	18.47%17.70%4.25%3.51%2.83%2.73%2.64%2.11%1.87...
dtype:	object

```
In [14]: len(populations)
Out[14]: 235
```



```
In [16]: populations

Out[16]:
```

	Country (or dependency)	Population (2020)	Yearly Change	Net Change	Density (P/Km²)	Land Area (Km²)	Migrants (net)	Fert. Rate	Med. Age	Urban Pop %	World Share
0	China	1440297825	0.39%	5540090	153	9388211	-348399.0	1.7	38	61%	18.47%
1	India	1382345085	0.99%	13586631	464	2973190	-532687.0	2.2	28	35%	17.70%
2	United States	331341050	0.59%	1937734	36	9147420	954806.0	1.8	38	83%	4.25%
3	Indonesia	274021604	1.07%	2898047	151	1811570	-98955.0	2.3	30	56%	3.51%
4	Pakistan	221612785	2.00%	4327022	287	770880	-233379.0	3.6	23	35%	2.83%
...	...	...	...	...	...	...	...	...	...	...	...
230	Montserrat	4993	0.06%	3	50	100	NaN	N.A.	N.A.	10%	0.00%
231	Falkland Islands	3497	3.05%	103	0	12170	NaN	N.A.	N.A.	66%	0.00%
232	Niue	1628	0.68%	11	6	260	NaN	N.A.	N.A.	46%	0.00%
233	Tokelau	1360	1.27%	17	136	10	NaN	N.A.	N.A.	0%	0.00%
234	Holy See	801	0.25%	2	2003	0	NaN	N.A.	N.A.	N.A.	0.00%

235 rows × 11 columns

```
In [17]: #populations['Urban Pop %'] = populations['Urban Pop %'].object.replace('[\%,\.]', '').astype(int)
```

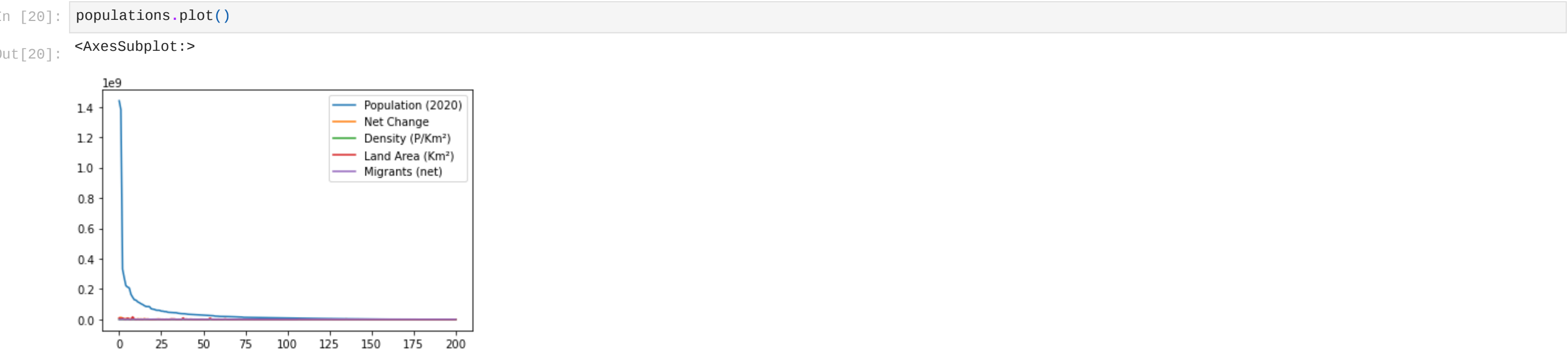
```
In [18]: populations.dropna(inplace=True)
```

```
In [19]: populations

Out[19]:
```

	Country (or dependency)	Population (2020)	Yearly Change	Net Change	Density (P/Km²)	Land Area (Km²)	Migrants (net)	Fert. Rate	Med. Age	Urban Pop %	World Share
0	China	1440297825	0.39%	5540090	153	9388211	-348399.0	1.7	38	61%	18.47%
1	India	1382345085	0.99%	13586631	464	2973190	-532687.0	2.2	28	35%	17.70%
2	United States	331341050	0.59%	1937734	36	9147420	954806.0	1.8	38	83%	4.25%
3	Indonesia	274021604	1.07%	2898047	151	1811570	-98955.0	2.3	30	56%	3.51%
4	Pakistan	221612785	2.00%	4327022	287	770880	-233379.0	3.6	23	35%	2.83%
...	...	...	...	...	...	...	...	...	...	...	...
196	Aruba	106845	0.43%	452	593	180	201.0	1.9	41	44%	0.00%
197	Tonga	105901	1.15%	1201	147	720	-800.0	3.6	22	24%	0.00%
198	U.S. Virgin Islands	104398	-0.15%	-153	298	350	-451.0	2	43	96%	0.00%
199	Seychelles	98453	0.62%	608	214	460	-200.0	2.5	34	56%	0.00%
200	Antigua and Barbuda	98069	0.84%	811	223	440	0.0	2	34	26%	0.00%

201 rows × 11 columns



```
In [21]: populations.sample(frac=0.5)

Out[21]:
```

	Country (or dependency)	Population (2020)	Yearly Change	Net Change	Density (P/Km²)	Land Area (Km²)	Migrants (net)	Fert. Rate	Med. Age	Urban Pop %	World Share
185	Mayotte	273905	2.50%	6665	728	375	0.0	3.7	20	46%	0.00%
198	U.S. Virgin Islands	104398	-0.15%	-153	298	350	-451.0	2	43	96%	0.00%
62	Chile	19144605	0.87%	164163	26	743532	111708.0	1.7	35	85%	0.25%
33	Sudan	44019263	2.42%	1036022	25	1765048	-50000.0	4.4	20	35%	0.56%
117	Slovakia	5460109	0.05%	2629	114	48088	1485.0	1.5	41	54%	0.07%
...	...	...	...	...	...	...	...	...	...	...	...
69	Senegal	16816539	2.75%	447563	87	192530	-20000.0	4.7	19	49%	0.21%
195	St. Vincent & Grenadines	111002	0.32%	351	284	390	-200.0	1.9	33	53%	0.00%
36	Afghanistan	39074280	2.33%	886592	60	652860	-62920.0	4.6	18	25%	0.50%
38	Canada	37799407	0.89%	331107	4	9093510	242032.0	1.5	41	81%	0.48%
63	Kazakhstan	18815231	1.21%	225280	7	2699700	-18000.0	2.8	31	58%	0.24%

100 rows × 11 columns

```
In [22]: populations.sample(frac=1)

Out[22]:
```

	Country (or dependency)	Population (2020)	Yearly Change	Net Change	Density (P/Km²)	Land Area (Km²)	Migrants (net)	Fert. Rate	Med. Age	Urban Pop %	World Share
136	Armenia	2964219	0.19%	5512	104	28470	-4998.0	1.8	35	63%	0.04%
28	Colombia	50976248	1.08%	543448	46	1109500	204796.0	1.8	31	80%	0.65%
133	Uruguay	3475842	0.35%	11996	20	175020	-3000.0	2	36	96%	0.04%
93	Hungary	9655983	-0.25%	-24328	107	90530	6000.0	1.5	43	72%	0.12%
84	Dominican Republic	10866667	1.01%	108952	225	48320	-30000.0	2.4	28	85%	0.14%
...	...	...	...	...	...	...	...	...	...	...	...
1	India	1382345085	0.99%	13586631	464	2973190	-532687.0	2.2	28	35%	17.70%
102	Sierra Leone	8004158	2.10%	163768	111	72180	-4200.0	4.3	19	43%	0.10%
183	New Caledonia	285972	0.97%	2748	16	18280	502.0	2	34	72%	0.00%
150	Latvia	1882408	-1.08%	-20545	30	62200	-14837.0	1.7	44	69%	0.02%
173	Malta	441750	0.27%	1171	1380	320	900.0	1.5	43	93%	0.01%

201 rows × 11 columns