

```
In [3]: import pandas as pd
```

```
In [4]: df=pd.read_csv('world_population.csv')
```

```
In [5]: df
```

Out[5]:

	Rank	CCA3	Country/Territory	Capital	Continent	2022 Population	2020 Population	2015 Population	Pop
0	36	AFG	Afghanistan	Kabul	Asia	41128771	38972230	33753499	28
1	138	ALB	Albania	Tirana	Europe	2842321	2866849	2882481	2
2	34	DZA	Algeria	Algiers	Africa	44903225	43451666	39543154	35
3	213	ASM	American Samoa	Pago Pago	Oceania	44273	46189	51368	
4	203	AND	Andorra	Andorra la Vella	Europe	79824	77700	71746	
...	
229	226	WLF	Wallis and Futuna	Mata- Utu	Oceania	11572	11655	12182	
230	172	ESH	Western Sahara	El Aaiún	Africa	575986	556048	491824	
231	46	YEM	Yemen	Sanaa	Asia	33696614	32284046	28516545	24
232	63	ZMB	Zambia	Lusaka	Africa	20017675	18927715	16248230	13
233	74	ZWE	Zimbabwe	Harare	Africa	16320537	15669666	14154937	12

234 rows × 17 columns



```
In [6]: df.head()
```

Out[6]:

	Rank	CCA3	Country/Territory	Capital	Continent	2022 Population	2020 Population	2015 Population	Popul
0	36	AFG	Afghanistan	Kabul	Asia	41128771	38972230	33753499	2818
1	138	ALB	Albania	Tirana	Europe	2842321	2866849	2882481	291
2	34	DZA	Algeria	Algiers	Africa	44903225	43451666	39543154	3585
3	213	ASM	American Samoa	Pago Pago	Oceania	44273	46189	51368	5
4	203	AND	Andorra	Andorra la Vella	Europe	79824	77700	71746	7



In [7]: `df.tail()`

Out[7]:

	Rank	CCA3	Country/Territory	Capital	Continent	2022 Population	2020 Population	2015 Population	Pop
229	226	WLF	Wallis and Futuna	Mata-Utu	Oceania	11572	11655	12182	
230	172	ESH	Western Sahara	El Aaiún	Africa	575986	556048	491824	
231	46	YEM	Yemen	Sanaa	Asia	33696614	32284046	28516545	24
232	63	ZMB	Zambia	Lusaka	Africa	20017675	18927715	16248230	13
233	74	ZWE	Zimbabwe	Harare	Africa	16320537	15669666	14154937	12

In [8]: `df.sample(5)`

Out[8]:

	Rank	CCA3	Country/Territory	Capital	Continent	2022 Population	2020 Population	2015 Population	Pe
11	55	AUS	Australia	Canberra	Oceania	26177413	25670051	23820236	2
117	141	LTU	Lithuania	Vilnius	Europe	2750055	2820267	2963765	
133	135	MDA	Moldova	Chisinau	Europe	3272996	3084847	3277388	
155	127	OMN	Oman	Muscat	Asia	4576298	4543399	4191776	
105	66	KAZ	Kazakhstan	Nursultan	Asia	19397998	18979243	17835909	

In [9]: df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 234 entries, 0 to 233
Data columns (total 17 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Rank                                  234 non-null    int64
1   CCA3                                  234 non-null    object
2   Country/Territory                    234 non-null    object
3   Capital                              234 non-null    object
4   Continent                            234 non-null    object
5   2022 Population                       234 non-null    int64
6   2020 Population                       234 non-null    int64
7   2015 Population                       234 non-null    int64
8   2010 Population                       234 non-null    int64
9   2000 Population                       234 non-null    int64
10  1990 Population                       234 non-null    int64
11  1980 Population                       234 non-null    int64
12  1970 Population                       234 non-null    int64
13  Area (km²)                           234 non-null    int64
14  Density (per km²)                     234 non-null    float64
15  Growth Rate                           234 non-null    float64
16  World Population Percentage            234 non-null    float64
dtypes: float64(3), int64(10), object(4)
memory usage: 31.2+ KB
```

In [10]: df.dtypes

```
Out[10]: Rank                int64
CCA3                object
Country/Territory    object
Capital              object
Continent            object
2022 Population      int64
2020 Population      int64
2015 Population      int64
2010 Population      int64
2000 Population      int64
1990 Population      int64
1980 Population      int64
1970 Population      int64
Area (km²)           int64
Density (per km²)    float64
Growth Rate          float64
World Population Percentage float64
dtype: object
```

In [11]:

df.describe()

Out[11]:

	Rank	2022 Population	2020 Population	2015 Population	2010 Population	2000 Population	Pc
count	234.000000	2.340000e+02	2.340000e+02	2.340000e+02	2.340000e+02	2.340000e+02	2.340
mean	117.500000	3.407441e+07	3.350107e+07	3.172996e+07	2.984524e+07	2.626947e+07	2.271
std	67.694165	1.367664e+08	1.355899e+08	1.304050e+08	1.242185e+08	1.116982e+08	9.783
min	1.000000	5.100000e+02	5.200000e+02	5.640000e+02	5.960000e+02	6.510000e+02	7.000
25%	59.250000	4.197385e+05	4.152845e+05	4.046760e+05	3.931490e+05	3.272420e+05	2.647
50%	117.500000	5.559944e+06	5.493074e+06	5.307400e+06	4.942770e+06	4.292907e+06	3.825
75%	175.750000	2.247650e+07	2.144798e+07	1.973085e+07	1.915957e+07	1.576230e+07	1.186
max	234.000000	1.425887e+09	1.424930e+09	1.393715e+09	1.348191e+09	1.264099e+09	1.153

In [12]:

df.isnull()

Out[12]:

	Rank	CCA3	Country/Territory	Capital	Continent	2022 Population	2020 Population	2015 Population	Pop
0	False	False	False	False	False	False	False	False	
1	False	False	False	False	False	False	False	False	
2	False	False	False	False	False	False	False	False	
3	False	False	False	False	False	False	False	False	
4	False	False	False	False	False	False	False	False	
...	
229	False	False	False	False	False	False	False	False	
230	False	False	False	False	False	False	False	False	
231	False	False	False	False	False	False	False	False	
232	False	False	False	False	False	False	False	False	
233	False	False	False	False	False	False	False	False	

234 rows × 17 columns

In [13]: `df.isnull().sum()`

```
Out[13]: Rank                0
CCA3                0
Country/Territory      0
Capital              0
Continent            0
2022 Population        0
2020 Population        0
2015 Population        0
2010 Population        0
2000 Population        0
1990 Population        0
1980 Population        0
1970 Population        0
Area (km²)            0
Density (per km²)      0
Growth Rate           0
World Population Percentage  0
dtype: int64
```

In [14]: `df.shape`

Out[14]: (234, 17)

In [15]: `df.isna()`

Out[15]:

	Rank	CCA3	Country/Territory	Capital	Continent	2022 Population	2020 Population	2015 Population	Pop
0	False	False	False	False	False	False	False	False	
1	False	False	False	False	False	False	False	False	
2	False	False	False	False	False	False	False	False	
3	False	False	False	False	False	False	False	False	
4	False	False	False	False	False	False	False	False	
...	
229	False	False	False	False	False	False	False	False	
230	False	False	False	False	False	False	False	False	
231	False	False	False	False	False	False	False	False	
232	False	False	False	False	False	False	False	False	
233	False	False	False	False	False	False	False	False	

234 rows × 17 columns



```
In [16]: df_continent=pd.get_dummies(df['Continent'])
```

```
In [17]: df_continent
```

```
Out[17]:
```

	Africa	Asia	Europe	North America	Oceania	South America
0	0	1	0	0	0	0
1	0	0	1	0	0	0
2	1	0	0	0	0	0
3	0	0	0	0	1	0
4	0	0	1	0	0	0
...
229	0	0	0	0	1	0
230	1	0	0	0	0	0
231	0	1	0	0	0	0
232	1	0	0	0	0	0
233	1	0	0	0	0	0

234 rows × 6 columns

```
In [18]: df["Country/Territory"].astype("string")
```

```
Out[18]: 0      Afghanistan
1      Albania
2      Algeria
3      American Samoa
4      Andorra
...
229    Wallis and Futuna
230    Western Sahara
231      Yemen
232      Zambia
233      Zimbabwe
Name: Country/Territory, Length: 234, dtype: string
```

```
In [ ]:
```

```
In [ ]:
```