Pandas

- **Installation of pandas** ✅
  - Importing pandas ✅
  - Importing the dataset ✅
  - Dataframe/Series ✅
- **Basic ops on a DataFrame** ✅
  - df.info() ✅
  - df.head() ✅
  - df.tail() ✅
  - df.shape() ✅
- **Creating Dataframe from Scratch** ✅
- **Basic ops on columns** ✅
  - Different ways of accessing cols ✅
  - Check for Unique values ✅
  - Rename column
  - Deleting col
  - Creating new cols

In [1]:
```
!pip install pandas
```

```
Requirement already satisfied: pandas in /Users/ritnil/anaconda3/lib/python
3.10/site-packages (1.5.3)
Requirement already satisfied: python-dateutil>=2.8.1 in /Users/ritnil/anaco
nda3/lib/python3.10/site-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /Users/ritnil/anaconda3/lib/p
ython3.10/site-packages (from pandas) (2022.7)
Requirement already satisfied: numpy>=1.21.0 in /Users/ritnil/anaconda3/lib/
python3.10/site-packages (from pandas) (1.23.5)
Requirement already satisfied: six>=1.5 in /Users/ritnil/anaconda3/lib/pytho
n3.10/site-packages (from python-dateutil>=2.8.1->pandas) (1.16.0)
```

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

In [3]:
```
!gdown 1E3bwvYGf1ig32RmcYiWc0IXPN-mD_bI_
```

```
Downloading...
From: https://drive.google.com/uc?id=1E3bwvYGf1ig32RmcYiWc0IXPN-mD_bI_ (http
s://drive.google.com/uc?id=1E3bwvYGf1ig32RmcYiWc0IXPN-mD_bI_)
To: /Users/ritnil/Scaler-cohorts/mckinsey.csv
100%|████████████████████████████████████████| 83.8k/83.8k [00:00<00:00, 510k
B/s]
```

In [6]:
```
df = pd.read_csv("/Users/ritnil/Scaler-cohorts/mckinsey.csv")
df
```

Out[6]:

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106
...	...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	Africa	62.351	706.157306
1700	Zimbabwe	1992	10704340	Africa	60.377	693.420786
1701	Zimbabwe	1997	11404948	Africa	46.809	792.449960
1702	Zimbabwe	2002	11926563	Africa	39.989	672.038623
1703	Zimbabwe	2007	12311143	Africa	43.487	469.709298

1704 rows × 6 columns

In [7]:
```
type(df)
```

Out[7]:
```
pandas.core.frame.DataFrame
```

In [8]:
```
df["country"] #recommended method 1 to access the column
```

Out[8]:
```
0       Afghanistan
1       Afghanistan
2       Afghanistan
3       Afghanistan
4       Afghanistan
           ...
1699       Zimbabwe
1700       Zimbabwe
1701       Zimbabwe
1702       Zimbabwe
1703       Zimbabwe
Name: country, Length: 1704, dtype: object
```

In [9]:
```
type(df["country"])
```

Out[9]:
```
pandas.core.series.Series
```

In [10]:
```
1 df
```

Out[10]:

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106
...	...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	Africa	62.351	706.157306
1700	Zimbabwe	1992	10704340	Africa	60.377	693.420786
1701	Zimbabwe	1997	11404948	Africa	46.809	792.449960
1702	Zimbabwe	2002	11926563	Africa	39.989	672.038623
1703	Zimbabwe	2007	12311143	Africa	43.487	469.709298

1704 rows × 6 columns

In [11]:
```
1 df.shape
```

Out[11]:
```
(1704, 6)
```

In [12]:
```
1 df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1704 entries, 0 to 1703
Data columns (total 6 columns):
 #   Column      Non-Null Count  Dtype  
---  ------      --------------  -----  
 0   country     1704 non-null   object 
 1   year        1704 non-null   int64  
 2   population  1704 non-null   int64  
 3   continent   1704 non-null   object 
 4   life_exp    1704 non-null   float64
 5   gdp_cap     1704 non-null   float64
dtypes: float64(2), int64(2), object(2)
memory usage: 80.0+ KB
```

In [13]:
```
1 df.describe()
```

Out[13]:

	year	population	life_exp	gdp_cap
count	1704.00000	1.704000e+03	1704.000000	1704.000000
mean	1979.50000	2.960121e+07	59.474439	7215.327081
std	17.26533	1.061579e+08	12.917107	9857.454543
min	1952.00000	6.001100e+04	23.599000	241.165876
25%	1965.75000	2.793664e+06	48.198000	1202.060309
50%	1979.50000	7.023596e+06	60.712500	3531.846988
75%	1993.25000	1.958522e+07	70.845500	9325.462346
max	2007.00000	1.318683e+09	82.603000	113523.132900

In [14]:
```
1 df.describe(include="all")
```

Out[14]:

	country	year	population	continent	life_exp	gdp_cap
count	1704	1704.00000	1.704000e+03	1704	1704.000000	1704.000000
unique	142	NaN	NaN	5	NaN	NaN
top	Afghanistan	NaN	NaN	Africa	NaN	NaN
freq	12	NaN	NaN	624	NaN	NaN
mean	NaN	1979.50000	2.960121e+07	NaN	59.474439	7215.327081
std	NaN	17.26533	1.061579e+08	NaN	12.917107	9857.454543
min	NaN	1952.00000	6.001100e+04	NaN	23.599000	241.165876
25%	NaN	1965.75000	2.793664e+06	NaN	48.198000	1202.060309
50%	NaN	1979.50000	7.023596e+06	NaN	60.712500	3531.846988
75%	NaN	1993.25000	1.958522e+07	NaN	70.845500	9325.462346
max	NaN	2007.00000	1.318683e+09	NaN	82.603000	113523.132900

In [15]:
```
1 df.head()
```

Out[15]:

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106

In [16]:
```
1 df.head(10)
```

Out[16]:

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106
5	Afghanistan	1977	14880372	Asia	38.438	786.113360
6	Afghanistan	1982	12881816	Asia	39.854	978.011439
7	Afghanistan	1987	13867957	Asia	40.822	852.395945
8	Afghanistan	1992	16317921	Asia	41.674	649.341395
9	Afghanistan	1997	22227415	Asia	41.763	635.341351

In [17]:
```
1 df.tail()
```

Out[17]:

	country	year	population	continent	life_exp	gdp_cap
1699	Zimbabwe	1987	9216418	Africa	62.351	706.157306
1700	Zimbabwe	1992	10704340	Africa	60.377	693.420786
1701	Zimbabwe	1997	11404948	Africa	46.809	792.449960
1702	Zimbabwe	2002	11926563	Africa	39.989	672.038623
1703	Zimbabwe	2007	12311143	Africa	43.487	469.709298

In [19]:
```
df.head(-2)
```

Out[19]:

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106
...	...	...	...	...	...	...
1697	Zimbabwe	1977	6642107	Africa	57.674	685.587682
1698	Zimbabwe	1982	7636524	Africa	60.363	788.855041
1699	Zimbabwe	1987	9216418	Africa	62.351	706.157306
1700	Zimbabwe	1992	10704340	Africa	60.377	693.420786
1701	Zimbabwe	1997	11404948	Africa	46.809	792.449960

1702 rows × 6 columns

In [20]:
```
#Create our own dataframes
```

In [ ]:
```
#1st method --> Row oriented
```

In [21]:
```
df.head(3)
```

Out[21]:

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710

In [22]:
```
df1=pd.DataFrame([["Afghanistan",1952,8425333,"Asia",28.801,779.445314],
                  ["Afghanistan",1957,9240934,"Asia",30.332,820.853030],
                  ["Afghanistan",1962,10267083,"Asia",31.997,853.100710]],
                 columns=["country","year","population","continent","life_exp'
```

In [23]:
```
df1
```

Out[23]:

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710

In [31]:
```
df2=pd.DataFrame([["Afghanistan",1952,8425333,"Asia",28.801,779.445314]],
                 columns=["country","year","population","continent","life_exp"
df2
```

Out[31]:

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314

In [32]:
```
#2nd method --> Column oriented
```

In [33]:
```
df3=pd.DataFrame({"country":["Afghanistan","Afghanistan","Afghanistan"],
                  "year":[1952,1957,1962],
                  "population":[8425333,9240934,10267083],
                  "continent":["Asia","Asia","Asia"],
                  "life_exp":[28.801,30.332,31.997],
                  "gdp_cap":[779.445314,820.853030,853.100710]})
```

In [34]:
```
df3
```

Out[34]:

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710

In [35]:
```
df
```

Out[35]:

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106
...	...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	Africa	62.351	706.157306
1700	Zimbabwe	1992	10704340	Africa	60.377	693.420786
1701	Zimbabwe	1997	11404948	Africa	46.809	792.449960
1702	Zimbabwe	2002	11926563	Africa	39.989	672.038623
1703	Zimbabwe	2007	12311143	Africa	43.487	469.709298

1704 rows × 6 columns

In [ ]:
```
1
```

In [36]:

```
df4=pd.DataFrame({"country":["Afghanistan","Afghanistan","Afghanistan","A
                  "year":[1952,1957,1962],
                  "population":[8425333,9240934,10267083],
                  "continent":["Asia","Asia","Asia"],
                  "life_exp":[28.801,30.332,31.997],
                  "gdp_cap":[779.445314,820.853030,853.100710]})
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
Cell In[36], line 1
----> 1 df4=pd.DataFrame({"country":["Afghanistan","Afghanistan","Afghanista
n","Afghanistan"],
      2              "year":[1952,1957,1962],
      3              "population":[8425333,9240934,10267083],
      4              "continent":["Asia","Asia","Asia"],
      5              "life_exp":[28.801,30.332,31.997],
      6              "gdp_cap":[779.445314,820.853030,853.100710]})

File ~/anaconda3/lib/python3.10/site-packages/pandas/core/frame.py:664, in D
ataFrame.__init__(self, data, index, columns, dtype, copy)
    658     mgr = self._init_mgr(
    659         data, axes={"index": index, "columns": columns}, dtype=dtyp
e, copy=copy
    660     )
    662 elif isinstance(data, dict):
    663     # GH#38939 de facto copy defaults to False only in non-dict case
s
--> 664     mgr = dict_to_mgr(data, index, columns, dtype=dtype, copy=copy,
typ=manager)
    665 elif isinstance(data, ma.MaskedArray):
    666     import numpy.ma.mrecords as mrecords

File ~/anaconda3/lib/python3.10/site-packages/pandas/core/internals/construc
tion.py:493, in dict_to_mgr(data, index, columns, dtype, typ, copy)
    489     else:
    490         # dtype check to exclude e.g. range objects, scalars
    491         arrays = [x.copy() if hasattr(x, "dtype") else x for x in ar
rays]
--> 493 return arrays_to_mgr(arrays, columns, index, dtype=dtype, typ=typ, c
onsolidate=copy)

File ~/anaconda3/lib/python3.10/site-packages/pandas/core/internals/construc
tion.py:118, in arrays_to_mgr(arrays, columns, index, dtype, verify_integrit
y, typ, consolidate)
    115 if verify_integrity:
    116     # figure out the index, if necessary
    117     if index is None:
--> 118         index = _extract_index(arrays)
    119     else:
    120         index = ensure_index(index)

File ~/anaconda3/lib/python3.10/site-packages/pandas/core/internals/construc
tion.py:666, in _extract_index(data)
    664 lengths = list(set(raw_lengths))
    665 if len(lengths) > 1:
--> 666     raise ValueError("All arrays must be of the same length")
    668 if have_dicts:
    669     raise ValueError(
    670         "Mixing dicts with non-Series may lead to ambiguous orderin
g."
    671     )
```

```
ValueError: All arrays must be of the same length
```

Basic ops on columns**

- Different ways of accessing cols
- Check for Unique values

- Rename column
- Deleting col

- Creating new cols

In [37]:
```
1 df["country"] #recommended
```

Out[37]:
```
0       Afghanistan
1       Afghanistan
2       Afghanistan
3       Afghanistan
4       Afghanistan
           ...
1699       Zimbabwe
1700       Zimbabwe
1701       Zimbabwe
1702       Zimbabwe
1703       Zimbabwe
Name: country, Length: 1704, dtype: object
```

In [38]:
```
1 df.country #not recommended
```

Out[38]:
```
0       Afghanistan
1       Afghanistan
2       Afghanistan
3       Afghanistan
4       Afghanistan
           ...
1699       Zimbabwe
1700       Zimbabwe
1701       Zimbabwe
1702       Zimbabwe
1703       Zimbabwe
Name: country, Length: 1704, dtype: object
```

In [39]:
```
1 df.columns
```

Out[39]:
```
Index(['country', 'year', 'population', 'continent', 'life_exp', 'gdp_cap'],
dtype='object')
```

In [40]:
```
df.keys() #Alternative method
```

Out[40]:
```
Index(['country', 'year', 'population', 'continent', 'life_exp', 'gdp_cap'],
dtype='object')
```

In [41]:
```
df["country"].unique()
```

Out[41]:
```
array(['Afghanistan', 'Albania', 'Algeria', 'Angola', 'Argentina',
       'Australia', 'Austria', 'Bahrain', 'Bangladesh', 'Belgium',
       'Benin', 'Bolivia', 'Bosnia and Herzegovina', 'Botswana', 'Brazil',
       'Bulgaria', 'Burkina Faso', 'Burundi', 'Cambodia', 'Cameroon',
       'Canada', 'Central African Republic', 'Chad', 'Chile', 'China',
       'Colombia', 'Comoros', 'Congo, Dem. Rep.', 'Congo, Rep.',
       'Costa Rica', "Cote d'Ivoire", 'Croatia', 'Cuba', 'Czech Republic',
       'Denmark', 'Djibouti', 'Dominican Republic', 'Ecuador', 'Egypt',
       'El Salvador', 'Equatorial Guinea', 'Eritrea', 'Ethiopia',
       'Finland', 'France', 'Gabon', 'Gambia', 'Germany', 'Ghana',
       'Greece', 'Guatemala', 'Guinea', 'Guinea-Bissau', 'Haiti',
       'Honduras', 'Hong Kong, China', 'Hungary', 'Iceland', 'India',
       'Indonesia', 'Iran', 'Iraq', 'Ireland', 'Israel', 'Italy',
       'Jamaica', 'Japan', 'Jordan', 'Kenya', 'Korea, Dem. Rep.',
       'Korea, Rep.', 'Kuwait', 'Lebanon', 'Lesotho', 'Liberia', 'Libya',
       'Madagascar', 'Malawi', 'Malaysia', 'Mali', 'Mauritania',
       'Mauritius', 'Mexico', 'Mongolia', 'Montenegro', 'Morocco',
       'Mozambique', 'Myanmar', 'Namibia', 'Nepal', 'Netherlands',
       'New Zealand', 'Nicaragua', 'Niger', 'Nigeria', 'Norway', 'Oman',
       'Pakistan', 'Panama', 'Paraguay', 'Peru', 'Philippines', 'Poland',
       'Portugal', 'Puerto Rico', 'Reunion', 'Romania', 'Rwanda',
       'Sao Tome and Principe', 'Saudi Arabia', 'Senegal', 'Serbia',
       'Sierra Leone', 'Singapore', 'Slovak Republic', 'Slovenia',
       'Somalia', 'South Africa', 'Spain', 'Sri Lanka', 'Sudan',
       'Swaziland', 'Sweden', 'Switzerland', 'Syria', 'Taiwan',
       'Tanzania', 'Thailand', 'Togo', 'Trinidad and Tobago', 'Tunisia',
       'Turkey', 'Uganda', 'United Kingdom', 'United States', 'Uruguay',
       'Venezuela', 'Vietnam', 'West Bank and Gaza', 'Yemen, Rep.',
       'Zambia', 'Zimbabwe'], dtype=object)
```

In [42]:
```
df["country"].nunique()
```

Out[42]:
```
142
```

In [45]:
```
df[["country","continent"]].head()
```

Out[45]:

	country	continent
0	Afghanistan	Asia
1	Afghanistan	Asia
2	Afghanistan	Asia
3	Afghanistan	Asia
4	Afghanistan	Asia

In [49]:
```
1 df[["country","continent"]]
```

Out[49]:

	country	continent
0	Afghanistan	Asia
1	Afghanistan	Asia
2	Afghanistan	Asia
3	Afghanistan	Asia
4	Afghanistan	Asia
...	...	...
1699	Zimbabwe	Africa
1700	Zimbabwe	Africa
1701	Zimbabwe	Africa
1702	Zimbabwe	Africa
1703	Zimbabwe	Africa

1704 rows × 2 columns

In [ ]:
```
1
```

In [52]:
```
1 #df["country"]
2 type(df["country"])
```

Out[52]:
```
pandas.core.series.Series
```

In [53]:
```
1 #df[["country"]]
2 type(df[["country"]])
```

Out[53]:
```
pandas.core.frame.DataFrame
```

In [54]:
```
1 df["country"].value_counts() #frequency of occurences
```

Out[54]:
```
Afghanistan          12
Pakistan             12
New Zealand          12
Nicaragua            12
Niger                12
                     ..
Eritrea              12
Equatorial Guinea    12
El Salvador          12
Egypt                12
Zimbabwe             12
Name: country, Length: 142, dtype: int64
```

In [55]:
```
df["continent"].value_counts()
```

Out[55]:
```
Africa      624
Asia        396
Europe      360
Americas    300
Oceania      24
Name: continent, dtype: int64
```

In [ ]:
```
```

In [59]:
```
df.rename({"country":"Country"},axis=1)
```

Out[59]:

	Country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106
...	...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	Africa	62.351	706.157306
1700	Zimbabwe	1992	10704340	Africa	60.377	693.420786
1701	Zimbabwe	1997	11404948	Africa	46.809	792.449960
1702	Zimbabwe	2002	11926563	Africa	39.989	672.038623
1703	Zimbabwe	2007	12311143	Africa	43.487	469.709298

1704 rows × 6 columns

In [56]:
```
df
```

Out[56]:

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106
...	...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	Africa	62.351	706.157306
1700	Zimbabwe	1992	10704340	Africa	60.377	693.420786
1701	Zimbabwe	1997	11404948	Africa	46.809	792.449960
1702	Zimbabwe	2002	11926563	Africa	39.989	672.038623
1703	Zimbabwe	2007	12311143	Africa	43.487	469.709298

1704 rows × 6 columns

In [ ]:
```
df.rename({"country":"Country"},axis=1) #method 1 to modify the column nc
```

In [63]:
```
df.rename({"country":"Country"},axis=1,inplace=True) #method 2
```

In [61]:
```
df
```

Out[61]:

	Country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106
...	...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	Africa	62.351	706.157306
1700	Zimbabwe	1992	10704340	Africa	60.377	693.420786
1701	Zimbabwe	1997	11404948	Africa	46.809	792.449960
1702	Zimbabwe	2002	11926563	Africa	39.989	672.038623
1703	Zimbabwe	2007	12311143	Africa	43.487	469.709298

1704 rows × 6 columns

In [64]:
```
df.rename(columns={"life_exp":"Life_exp",
                   "population":"Population",
                   "continent":"Continent"},inplace=True) #another way to
```

In [65]:
```
df
```

Out[65]:

	Country	year	Population	Continent	Life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106
...	...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	Africa	62.351	706.157306
1700	Zimbabwe	1992	10704340	Africa	60.377	693.420786
1701	Zimbabwe	1997	11404948	Africa	46.809	792.449960
1702	Zimbabwe	2002	11926563	Africa	39.989	672.038623
1703	Zimbabwe	2007	12311143	Africa	43.487	469.709298

1704 rows × 6 columns

In [66]:
```
df.columns
```

Out[66]:
```
Index(['Country', 'year', 'Population', 'Continent', 'Life_exp', 'gdp_cap'],
dtype='object')
```

In [67]:
```
df.columns[-1]
```

Out[67]:
```
'gdp_cap'
```

In [68]:
```
1 df.drop("Continent",axis=1)
```

Out[68]:

	Country	year	Population	Life_exp	gdp_cap
0	Afghanistan	1952	8425333	28.801	779.445314
1	Afghanistan	1957	9240934	30.332	820.853030
2	Afghanistan	1962	10267083	31.997	853.100710
3	Afghanistan	1967	11537966	34.020	836.197138
4	Afghanistan	1972	13079460	36.088	739.981106
...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	62.351	706.157306
1700	Zimbabwe	1992	10704340	60.377	693.420786
1701	Zimbabwe	1997	11404948	46.809	792.449960
1702	Zimbabwe	2002	11926563	39.989	672.038623
1703	Zimbabwe	2007	12311143	43.487	469.709298

1704 rows × 5 columns

In [70]:
```
1 df
```

Out[70]:

	Country	year	Population	Continent	Life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106
...	...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	Africa	62.351	706.157306
1700	Zimbabwe	1992	10704340	Africa	60.377	693.420786
1701	Zimbabwe	1997	11404948	Africa	46.809	792.449960
1702	Zimbabwe	2002	11926563	Africa	39.989	672.038623
1703	Zimbabwe	2007	12311143	Africa	43.487	469.709298

1704 rows × 6 columns

In [71]:
```
df.drop(["Continent","year"],axis=1)
```

Out[71]:

	Country	Population	Life_exp	gdp_cap
0	Afghanistan	8425333	28.801	779.445314
1	Afghanistan	9240934	30.332	820.853030
2	Afghanistan	10267083	31.997	853.100710
3	Afghanistan	11537966	34.020	836.197138
4	Afghanistan	13079460	36.088	739.981106
...	...	...	...	...
1699	Zimbabwe	9216418	62.351	706.157306
1700	Zimbabwe	10704340	60.377	693.420786
1701	Zimbabwe	11404948	46.809	792.449960
1702	Zimbabwe	11926563	39.989	672.038623
1703	Zimbabwe	12311143	43.487	469.709298

1704 rows × 4 columns

In [72]:
```
df.drop(columns=["Continent","year"])
```

Out[72]:

	Country	Population	Life_exp	gdp_cap
0	Afghanistan	8425333	28.801	779.445314
1	Afghanistan	9240934	30.332	820.853030
2	Afghanistan	10267083	31.997	853.100710
3	Afghanistan	11537966	34.020	836.197138
4	Afghanistan	13079460	36.088	739.981106
...	...	...	...	...
1699	Zimbabwe	9216418	62.351	706.157306
1700	Zimbabwe	10704340	60.377	693.420786
1701	Zimbabwe	11404948	46.809	792.449960
1702	Zimbabwe	11926563	39.989	672.038623
1703	Zimbabwe	12311143	43.487	469.709298

1704 rows × 4 columns

In [73]:
```
df.drop(columns=["Continent"],inplace=True)
```

In [74]:
```
1 df
```

Out[74]:

	Country	year	Population	Life_exp	gdp_cap
0	Afghanistan	1952	8425333	28.801	779.445314
1	Afghanistan	1957	9240934	30.332	820.853030
2	Afghanistan	1962	10267083	31.997	853.100710
3	Afghanistan	1967	11537966	34.020	836.197138
4	Afghanistan	1972	13079460	36.088	739.981106
...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	62.351	706.157306
1700	Zimbabwe	1992	10704340	60.377	693.420786
1701	Zimbabwe	1997	11404948	46.809	792.449960
1702	Zimbabwe	2002	11926563	39.989	672.038623
1703	Zimbabwe	2007	12311143	43.487	469.709298

1704 rows × 5 columns

In [ ]:
```
1
```

In [75]:
```
1 df["year+7"] = df["year"]+7
```

In [76]:
```
1 df
```

Out[76]:

	Country	year	Population	Life_exp	gdp_cap	year+7
0	Afghanistan	1952	8425333	28.801	779.445314	1959
1	Afghanistan	1957	9240934	30.332	820.853030	1964
2	Afghanistan	1962	10267083	31.997	853.100710	1969
3	Afghanistan	1967	11537966	34.020	836.197138	1974
4	Afghanistan	1972	13079460	36.088	739.981106	1979
...	...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	62.351	706.157306	1994
1700	Zimbabwe	1992	10704340	60.377	693.420786	1999
1701	Zimbabwe	1997	11404948	46.809	792.449960	2004
1702	Zimbabwe	2002	11926563	39.989	672.038623	2009
1703	Zimbabwe	2007	12311143	43.487	469.709298	2014

1704 rows × 6 columns

In [78]:
```
df["year + Gdp_cap"] =  ( df["year"] + df["Life_exp"] )*2
```

In [79]:
```
df
```

Out[79]:

	Country	year	Population	Life_exp	gdp_cap	year+7	year + Gdp_cap
0	Afghanistan	1952	8425333	28.801	779.445314	1959	3961.602
1	Afghanistan	1957	9240934	30.332	820.853030	1964	3974.664
2	Afghanistan	1962	10267083	31.997	853.100710	1969	3987.994
3	Afghanistan	1967	11537966	34.020	836.197138	1974	4002.040
4	Afghanistan	1972	13079460	36.088	739.981106	1979	4016.176
...	...	...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	62.351	706.157306	1994	4098.702
1700	Zimbabwe	1992	10704340	60.377	693.420786	1999	4104.754
1701	Zimbabwe	1997	11404948	46.809	792.449960	2004	4087.618
1702	Zimbabwe	2002	11926563	39.989	672.038623	2009	4083.978
1703	Zimbabwe	2007	12311143	43.487	469.709298	2014	4100.974

1704 rows × 7 columns

In [83]:
```
df["own"]=[i for i in range(1,1705)] #Length of value should match length
```

In [81]:
```
df
```

Out[81]:

	Country	year	Population	Life_exp	gdp_cap	year+7	year + Gdp_cap	own
0	Afghanistan	1952	8425333	28.801	779.445314	1959	3961.602	1
1	Afghanistan	1957	9240934	30.332	820.853030	1964	3974.664	2
2	Afghanistan	1962	10267083	31.997	853.100710	1969	3987.994	3
3	Afghanistan	1967	11537966	34.020	836.197138	1974	4002.040	4
4	Afghanistan	1972	13079460	36.088	739.981106	1979	4016.176	5
...	...	...	...	...	...	...	...	...
1699	Zimbabwe	1987	9216418	62.351	706.157306	1994	4098.702	1700
1700	Zimbabwe	1992	10704340	60.377	693.420786	1999	4104.754	1701
1701	Zimbabwe	1997	11404948	46.809	792.449960	2004	4087.618	1702
1702	Zimbabwe	2002	11926563	39.989	672.038623	2009	4083.978	1703
1703	Zimbabwe	2007	12311143	43.487	469.709298	2014	4100.974	1704

1704 rows × 8 columns

In [ ]:
```
1
```