pandas-lecture-1-dec-batch

June 3, 2023

[1]:
```
import numpy as np
```

[2]:
```
a = np.arange(6)
a.shape
```

[2]:
```
(6,)
```

[3]:
```
a = a.reshape(1, -1)
a.shape
```

[3]:
```
(1, 6)
```

[4]:
```
a
```

[4]:
```
array([[0, 1, 2, 3, 4, 5]])
```

[7]:
```
a = np.arange(6)
np.expand_dims(a, axis=0).shape
```

[7]:
```
(1, 6)
```

[9]:
```
np.expand_dims(a, axis=1)
```

[9]:
```
array([[0],
       [1],
       [2],
       [3],
       [4],
       [5]])
```

[10]:
```
a = np.arange(6)
a[np.newaxis, :] #Equavalent of expand_dims( axis=0)
```

[10]:
```
array([[0, 1, 2, 3, 4, 5]])
```

[11]:
```
a[:, np.newaxis] #Equavalent of expand_dims( axis=1)
```

[11]:
```
array([[0],
       [1],
       [2],
       [3],
       [4],
       [5]])
```

[12]:
```
b = np.arange(6).reshape(2, 3)
b.shape
```

[12]:
```
(2, 3)
```

[13]:
```
np.expand_dims(b, axis=0).shape
```

[13]:
```
(1, 2, 3)
```

[15]:
```
np.expand_dims(b, axis=1)
```

[15]:
```
array([[[0, 1, 2]],

       [[3, 4, 5]]])
```

[16]:
```
a = np.arange(9).reshape(1, 1, 9)
a
```

[16]:
```
array([[[0, 1, 2, 3, 4, 5, 6, 7, 8]]])
```

[18]:
```
np.squeeze(a)
```

[18]:
```
array([0, 1, 2, 3, 4, 5, 6, 7, 8])
```

[20]:
```
np.squeeze(a, axis=1).shape
```

[20]:
```
(1, 9)
```

[21]:
```
np.squeeze(a, axis=2)
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
Input In [21], in <cell line: 1>()
----> 1 np.squeeze(a, axis=2)

File <__array_function__ internals>:180, in squeeze(*args, **kwargs)

File /usr/local/lib/python3.9/site-packages/numpy/core/fromnumeric.py:1545, in␣
 ↪squeeze(a, axis)
   1543     return squeeze()
```

```
   1544 else:
-> 1545     return squeeze(axis=axis)

ValueError: cannot select an axis to squeeze out which has size not equal to on
```

0.0.1 Pandas

[22]:
```
!pip install pandas
```

```
DEPRECATION: Configuring installation scheme with distutils config files is
deprecated and will no longer work in the near future. If you are using a
Homebrew or Linuxbrew Python, please see discussion at
https://github.com/Homebrew/homebrew-core/issues/76621
Requirement already satisfied: pandas in /usr/local/lib/python3.9/site-
packages (1.4.1)
Requirement already satisfied: python-dateutil>=2.8.1 in
/usr/local/lib/python3.9/site-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.9/site-
packages (from pandas) (2021.3)
Requirement already satisfied: numpy>=1.18.5 in /usr/local/lib/python3.9/site-
packages (from pandas) (1.22.3)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.9/site-
packages (from python-dateutil>=2.8.1->pandas) (1.16.0)
DEPRECATION: Configuring installation scheme with distutils config files is
deprecated and will no longer work in the near future. If you are using a
Homebrew or Linuxbrew Python, please see discussion at
https://github.com/Homebrew/homebrew-core/issues/76621
```

[23]:
```
import pandas as pd
```

[24]:
```
!gdown 1E3bwvYGf1ig32RmcYiWc0IXPN-mD_bI_
```

```
Downloading…
From: https://drive.google.com/uc?id=1E3bwvYGf1ig32RmcYiWc0IXPN-mD_bI_
To: /Users/satish/Desktop/scaler/Dec Tue Batch - DAV-1/mckinsey.csv
100%|                    | 83.8k/83.8k [00:00<00:00, 10.2MB/s]
```

[25]:
```
df = pd.read_csv('mckinsey.csv')
type(df)
```

[25]:
```
pandas.core.frame.DataFrame
```

[26]:
```
df
```

[26]:
```
          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 x 6 columns]
```

[28]:
```
type(df['country'])
```

[28]:
```
pandas.core.series.Series
```

[29]:
```
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
```

[30]:
```
df.head()
```

[30]:
```
       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
```

[31]:
```
df.head(10)
```

[31]:

	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

[32]:
```
df
```

[32]:

	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 x 6 columns]

[33]:
```
df.tail()
```

[33]:

	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

[34]:
```
df.tail(10)
```

[34]:

	country	year	population	continent	life_exp	gdp_cap
1694	Zimbabwe	1962	4277736	Africa	52.358	527.272182
1695	Zimbabwe	1967	4995432	Africa	53.995	569.795071
1696	Zimbabwe	1972	5861135	Africa	55.635	799.362176
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  Zimbabwe  2002    11926563    Africa    39.989  672.038623
1703  Zimbabwe  2007    12311143    Africa    43.487  469.709298
```

[35]:
```
df.head(-10)
```

[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
…             …     …           …         …         …           …
1689       Zambia  1997     9417789    Africa    40.238  1071.353818
1690       Zambia  2002    10595811    Africa    39.193  1071.613938
1691       Zambia  2007    11746035    Africa    42.384  1271.211593
1692     Zimbabwe  1952     3080907    Africa    48.451   406.884115
1693     Zimbabwe  1957     3646340    Africa    50.469   518.764268

[1694 rows x 6 columns]
```

[36]:
```
df.shape
```

[36]:
```
(1704, 6)
```

[37]:
```
df
```

[37]:
```
          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 x 6 columns]
```

0.0.2 Create a Data frame from scratch

[40]:
```
a = pd.DataFrame([['Afghanistan',1952,8425333,"Asia",28.801,779.445314],
                  ['Afghanistan',1953,8425333,"Asia",28.801,779.445314]],
                  columns=['Country', 'Year', 'Population', 'Continent', 'life_exp',␣
 ↪'gdp_cap'])
a
```

[40]:

	Country	Year	Population	Continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1953	8425333	Asia	28.801	779.445314

[41]:
```
a = pd.DataFrame(['Afghanistan',1952,8425333,"Asia",28.801,779.445314],
                    columns=['Country', 'Year', 'Population', 'Continent',␣
 ↪'life_exp', 'gdp_cap'])
a
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
Input In [41], in <cell line: 1>()
----> 1 a = pd.DataFrame(['Afghanistan',1952,8425333,"Asia",28.801,779.445314],
      2␣
 ↪                  columns=['Country', 'Year', 'Population', 'Continent', 'life_exp', 'gdp_ca
      3 a

File /usr/local/lib/python3.9/site-packages/pandas/core/frame.py:737, in␣
 ↪DataFrame.__init__(self, data, index, columns, dtype, copy)
    729         mgr = arrays_to_mgr(
    730             arrays,
    731             columns,
   (…)
    734             typ=manager,
    735         )
    736     else:
--> 737         mgr = ndarray_to_mgr(
    738             data,
    739             index,
    740             columns,
    741             dtype=dtype,
    742             copy=copy,
    743             typ=manager,
    744         )
    745 else:
    746     mgr = dict_to_mgr(
    747         {},
    748         index,
   (…)
```

```
    751         typ=manager,
    752     )

File /usr/local/lib/python3.9/site-packages/pandas/core/internals/construction.
 ↪py:351, in ndarray_to_mgr(values, index, columns, dtype, copy, typ)
    346 # _prep_ndarray ensures that values.ndim == 2 at this point
    347 index, columns = _get_axes(
    348     values.shape[0], values.shape[1], index=index, columns=columns
    349 )
--> 351 _check_values_indices_shape_match(values, index, columns)
    353 if typ == "array":
    355     if issubclass(values.dtype.type, str):

File /usr/local/lib/python3.9/site-packages/pandas/core/internals/construction.
 ↪py:422, in _check_values_indices_shape_match(values, index, columns)
    420 passed = values.shape
    421 implied = (len(index), len(columns))
--> 422 raise ValueError(f"Shape of passed values is {passed}, indices imply␣
 ↪{implied}")

ValueError: Shape of passed values is (6, 1), indices imply (6, 6)
```

[42]:
```
a = pd.DataFrame({'Country': ['India', 'India'], 'Population': [345678766,␣
 ↪87654567]})
a
```

[42]:

	Country	Population
0	India	345678766
1	India	87654567

0.0.3 Basic Operations on Columns

[43]:
```
df.columns
```

[43]:
```
Index(['country', 'year', 'population', 'continent', 'life_exp', 'gdp_cap'],
dtype='object')
```

[44]:
```
df.keys()
```

[44]:
```
Index(['country', 'year', 'population', 'continent', 'life_exp', 'gdp_cap'],
dtype='object')
```

[47]:
```
type(df[['country', 'population']])
```

[47]:
```
pandas.core.frame.DataFrame
```

[48]:
```
type(df[['country']])
```

[48]:
```
pandas.core.frame.DataFrame
```

[50]:
```
type(df['country'])
```

[50]:
```
pandas.core.series.Series
```

[51]:
```
#Find unique countries present in a column called country
```

[53]:
```
df['country'].unique()
```

[53]:
```
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)
```

[54]:
```
df['country'].value_counts()
```

[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
```

Change the name of the column

[55]:
```
df.rename({'country': 'Country', 'population': 'Population'}, axis=1)
```

[55]:

	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 x 6 columns]

[56]:
```
df
```

[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 x 6 columns]

[57]:
```
df.rename({'country': 'Country', 'population': 'Population'}, axis=1,␣
↪inplace=True)
```

[58]:
```
df
```

[58]:

	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 x 6 columns]

[61]:
```
df.rename(columns={'Country': 'country'}, inplace=True)
```

[62]:
```
df
```

[62]:

	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 x 6 columns]

[63]:
```
df.country
```

[63]:
```
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
```

0.0.4 Delete a column in Pandas

[66]:
```
df.drop('continent', axis=1, inplace=True)
```

[67]:
```
df
```

[67]:
```
          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 x 5 columns]
```

[69]:
```
df.drop(columns=['country', 'Population'])
```

[69]:
```
      year  life_exp     gdp_cap
0     1952    28.801  779.445314
1     1957    30.332  820.853030
2     1962    31.997  853.100710
3     1967    34.020  836.197138
4     1972    36.088  739.981106
…     …          …           …
1699  1987    62.351  706.157306
1700  1992    60.377  693.420786
1701  1997    46.809  792.449960
1702  2002    39.989  672.038623
1703  2007    43.487  469.709298

[1704 rows x 3 columns]
```

1 Add New column to a Data frame

[70]:
```
df
```

[70]:
```
          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 x 5 columns]
```

[71]:
```
df['year+7'] = df['year'] + 7
df
```

[71]:
```
          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 x 6 columns]
```

[72]:
```
df['gdp'] = df['Population']*df['gdp_cap']
df
```

[72]:
```
          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

                gdp
0     6.567086e+09
1     7.585449e+09
2     8.758856e+09
3     9.648014e+09
4     9.678553e+09
…              …
1699  6.508241e+09
1700  7.422612e+09
1701  9.037851e+09
1702  8.015111e+09
1703  5.782658e+09

[1704 rows x 7 columns]
```

[73]:
```
df['own'] = list(range(1, df.shape[0]+1))
df
```

[73]:
```
           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

                gdp   own
0     6.567086e+09     1
1     7.585449e+09     2
2     8.758856e+09     3
3     9.648014e+09     4
4     9.678553e+09     5
…              …   …
1699  6.508241e+09  1700
1700  7.422612e+09  1701
1701  9.037851e+09  1702
```

```
1702  8.015111e+09  1703
1703  5.782658e+09  1704

[1704 rows x 8 columns]
```