Zadanie 1 ładowanie biblioteki Pandas

```python
import pandas as pd
```

Zadanie 2 tworzenie ramki danych ze slownika

```python
dict_user = {"User" : ["A", "B", "C"],
             "Score" : [5,   10,   15]}

pd.DataFrame(dict_user)
```

```
   User  Score
0     A      5
1     B     10
2     C     15
```

Zadanie 3 zachowanie ramki danych pobranych z pliku w formacie csv (xlsx)

```python
df = pd.read_csv("IHME-GBD_2019_DATA-15798851-2.csv", encoding='utf-8', engine='python')
```

Zadanie 4 tworzenie ramki danych z listy list

```python
lists_user = [["A", "B", "C"],
[5,   10,   15]]

pd.DataFrame(lists_user)
```

```
   0   1   2
0  A   B   C
1  5  10  15
```

Zadanie 5 transponowanie (wymieniamy kolumny na wiersze)

```python
pd.DataFrame(lists_user).T
```

```
   0   1
0  A   5
1  B  10
2  C  15
```

Zadanie 6 wyswietlic pierwsze 10 wierszy ramki danych

```python
df.head(10)
```

```
                                measure location     sex
age  \
0  DALYs (Disability-Adjusted Life Years)   Gambia  Female  All Ages

1  DALYs (Disability-Adjusted Life Years)   Gambia    Both  All Ages

2  DALYs (Disability-Adjusted Life Years)   Gambia    Male  All Ages
```

```
3   DALYs (Disability-Adjusted Life Years)   Gambia   Female   All Ages

4   DALYs (Disability-Adjusted Life Years)   Gambia     Both   All Ages

5   DALYs (Disability-Adjusted Life Years)   Gambia     Male   All Ages

6   DALYs (Disability-Adjusted Life Years)   Gambia   Female   All Ages

7   DALYs (Disability-Adjusted Life Years)   Gambia     Both   All Ages

8   DALYs (Disability-Adjusted Life Years)   Gambia     Male   All Ages

9   DALYs (Disability-Adjusted Life Years)   Gambia   Female   All Ages


                                 cause    metric   year          val
upper  \
0  Maternal and neonatal disorders     Rate   2012   7475.212700
9104.773541
1  Maternal and neonatal disorders     Rate   2012   7814.344518
9667.960848
2         Substance use disorders   Number   2012   1659.038707
2126.829521
3         Substance use disorders   Number   2012    874.432466
1186.560596
4         Substance use disorders   Number   2012   2533.471173
3231.220866
5         Substance use disorders   Percent  2012      0.003799
0.004849
6         Substance use disorders   Percent  2012      0.002202
0.002889
7         Substance use disorders   Percent  2012      0.003038
0.003885
8         Substance use disorders     Rate   2012    179.493660
230.104586
9         Substance use disorders     Rate   2012     91.400543
124.025911

          lower
0   6157.428603
1   6289.146375
2   1239.172699
3    618.271780
4   1868.204609
5      0.002863
6      0.001585
7      0.002298
```

```
8    134.067784
9     64.625204
```

Zadanie 7 wyswietlic ostatnie 10 wierszy ramki danych

```
df.tail(10)
```

```
                                 measure location     sex
age  \
307830  DALYs (Disability-Adjusted Life Years)   Malawi    Both  All
Ages
307831  DALYs (Disability-Adjusted Life Years)   Malawi    Male  All
Ages
307832  DALYs (Disability-Adjusted Life Years)   Malawi  Female  All
Ages
307833  DALYs (Disability-Adjusted Life Years)   Malawi    Both  All
Ages
307834  DALYs (Disability-Adjusted Life Years)   Malawi    Male  All
Ages
307835  DALYs (Disability-Adjusted Life Years)   Malawi  Female  All
Ages
307836  DALYs (Disability-Adjusted Life Years)   Malawi    Both  All
Ages
307837  DALYs (Disability-Adjusted Life Years)   Malawi    Male  All
Ages
307838  DALYs (Disability-Adjusted Life Years)   Malawi  Female  All
Ages
307839  DALYs (Disability-Adjusted Life Years)   Malawi    Both  All
Ages

                                 cause   metric  year              val  \
307830  Chronic respiratory diseases     Rate  2019       635.274862
307831         Unintentional injuries   Number  2019    118281.113928
307832         Unintentional injuries   Number  2019     77026.152292
307833         Unintentional injuries   Number  2019    195307.266221
307834         Unintentional injuries  Percent  2019         0.029346
307835         Unintentional injuries  Percent  2019         0.022152
307836         Unintentional injuries  Percent  2019         0.026012
307837         Unintentional injuries     Rate  2019      1318.844977
307838         Unintentional injuries     Rate  2019       813.052440
307839         Unintentional injuries     Rate  2019      1059.021476

                 upper           lower
307830      755.292643      530.146531
307831   154072.043418    93047.621026
307832    98717.013916    60499.526888
307833   246103.104036   156121.639397
307834        0.036745        0.024719
307835        0.026646        0.018384
307836        0.031307        0.022140
307837     1717.917036     1037.489279
```

```
307838      1042.011143      638.605025
307839      1334.453539      846.543870
```

Zadanie 8 wyswietlic informację o ramce danych

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 307840 entries, 0 to 307839
Data columns (total 10 columns):
 #   Column    Non-Null Count    Dtype
---  ------    --------------    -----
 0   measure   307840 non-null   object
 1   location  307840 non-null   object
 2   sex       307840 non-null   object
 3   age       307840 non-null   object
 4   cause     307840 non-null   object
 5   metric    307840 non-null   object
 6   year      307840 non-null   int64
 7   val       307840 non-null   float64
 8   upper     307840 non-null   float64
 9   lower     307840 non-null   float64
dtypes: float64(3), int64(1), object(6)
memory usage: 23.5+ MB
```

Zadanie 9 wyswietlic, ile wierszy i kolumn znajduje się w ramce danych

```
df.shape
```

```
(307840, 10)
```

Zadanie 10 wyswietlic informację statystyczną o kolumnach liczbowych (wartosci niepowtarzalne, ́srednia, odchylenie standardowe, minimum, kwartyle, maksimum)

```
df.describe()
```

```
              year            val          upper          lower
count  307840.000000  3.078400e+05  3.078400e+05  3.078400e+05
mean     2015.666980  1.256793e+05  1.582079e+05  9.936609e+04
std         2.218973  1.159277e+06  1.379991e+06  9.791267e+05
min      2011.000000  5.539435e-05  9.085092e-05  2.626711e-05
25%      2014.000000  5.728504e-02  7.223016e-02  4.511736e-02
50%      2016.000000  7.406817e+02  9.757076e+02  5.325397e+02
75%      2018.000000  6.713425e+03  8.757922e+03  5.114418e+03
max      2019.000000  9.193312e+07  1.041249e+08  8.013535e+07
```

Zadanie 11 wyswietlic informację statystyczną o kolumnach kategoryzowanych (ile unikalnych wartosci, top - jaka jest najpopularniejsza wartosc, freq - jak czesto najpopularniejsza)

```
df.describe(include = 'all')
```

	measure	location	sex	age
count	307840	307840	307840	307840
unique	1	204	3	1
top	DALYs (Disability-Adjusted Life Years)	Monaco	Female	All Ages
freq	307840	1620	102614	307840
mean	NaN	NaN	NaN	NaN
std	NaN	NaN	NaN	NaN
min	NaN	NaN	NaN	NaN
25%	NaN	NaN	NaN	NaN
50%	NaN	NaN	NaN	NaN
75%	NaN	NaN	NaN	NaN
max	NaN	NaN	NaN	NaN

	cause	metric	year	val
count	307840	307840	307840.000000	3.078400e+05
unique	22	3	NaN	NaN
top	Chronic respiratory diseases	Rate	NaN	NaN
freq	14076	102617	NaN	NaN
mean	NaN	NaN	2015.666980	1.256793e+05
std	NaN	NaN	2.218973	1.159277e+06
min	NaN	NaN	2011.000000	5.539435e-05
25%	NaN	NaN	2014.000000	5.728504e-02
50%	NaN	NaN	2016.000000	7.406817e+02
75%	NaN	NaN	2018.000000	6.713425e+03
max	NaN	NaN	2019.000000	9.193312e+07

```
                upper            lower
count      3.078400e+05    3.078400e+05
unique              NaN             NaN
top                 NaN             NaN
freq                NaN             NaN
mean       1.582079e+05    9.936609e+04
std        1.379991e+06    9.791267e+05
min        9.085092e-05    2.626711e-05
25%        7.223016e-02    4.511736e-02
50%        9.757076e+02    5.325397e+02
75%        8.757922e+03    5.114418e+03
max        1.041249e+08    8.013535e+07
```

Zadanie 12 usunac brakujace wartosci w ramce danych

```
df.dropna(inplace=True)
```

Zadanie 13 przedstawic wybor wierszy i kolumny uzywajac nazw oraz indeksow na rozne sposoby

```
df["location"]
```

```
0          Gambia
1          Gambia
2          Gambia
3          Gambia
4          Gambia
            ...
307835     Malawi
307836     Malawi
307837     Malawi
307838     Malawi
307839     Malawi
Name: location, Length: 307840, dtype: object
```

```
df.location
```

```
0          Gambia
1          Gambia
2          Gambia
3          Gambia
4          Gambia
            ...
307835     Malawi
307836     Malawi
307837     Malawi
307838     Malawi
307839     Malawi
Name: location, Length: 307840, dtype: object
```

```
df[["location","age"]]
```

```
        location       age
0         Gambia  All Ages
1         Gambia  All Ages
2         Gambia  All Ages
3         Gambia  All Ages
4         Gambia  All Ages
...          ...       ...
307835    Malawi  All Ages
307836    Malawi  All Ages
307837    Malawi  All Ages
307838    Malawi  All Ages
307839    Malawi  All Ages

[307840 rows x 2 columns]
```

df.loc[:, "measure":"age"]

```
                                        measure location     sex
age
0       DALYs (Disability-Adjusted Life Years)   Gambia  Female  All
Ages
1       DALYs (Disability-Adjusted Life Years)   Gambia    Both  All
Ages
2       DALYs (Disability-Adjusted Life Years)   Gambia    Male  All
Ages
3       DALYs (Disability-Adjusted Life Years)   Gambia  Female  All
Ages
4       DALYs (Disability-Adjusted Life Years)   Gambia    Both  All
Ages
...                                         ...      ...     ...
...
307835  DALYs (Disability-Adjusted Life Years)   Malawi  Female  All
Ages
307836  DALYs (Disability-Adjusted Life Years)   Malawi    Both  All
Ages
307837  DALYs (Disability-Adjusted Life Years)   Malawi    Male  All
Ages
307838  DALYs (Disability-Adjusted Life Years)   Malawi  Female  All
Ages
307839  DALYs (Disability-Adjusted Life Years)   Malawi    Both  All
Ages

[307840 rows x 4 columns]
```

df.loc[2000:2010, "measure":"age"]

```
                                      measure location     sex
age
2000  DALYs (Disability-Adjusted Life Years)    Benin  Female  All
Ages
2001  DALYs (Disability-Adjusted Life Years)    Benin    Both  All
```

```
Ages
2002  DALYs (Disability-Adjusted Life Years)    Benin    Male  All
Ages
2003  DALYs (Disability-Adjusted Life Years)    Benin  Female  All
Ages
2004  DALYs (Disability-Adjusted Life Years)    Benin    Both  All
Ages
2005  DALYs (Disability-Adjusted Life Years)    Benin    Male  All
Ages
2006  DALYs (Disability-Adjusted Life Years)    Benin  Female  All
Ages
2007  DALYs (Disability-Adjusted Life Years)    Benin    Both  All
Ages
2008  DALYs (Disability-Adjusted Life Years)    Benin    Male  All
Ages
2009  DALYs (Disability-Adjusted Life Years)    Benin  Female  All
Ages
2010  DALYs (Disability-Adjusted Life Years)    Benin    Both  All
Ages
```

```python
df.iloc[1000:1010, 1:4]
```

```
      location     sex       age
1000    Gambia    Male  All Ages
1001    Gambia  Female  All Ages
1002    Gambia    Both  All Ages
1003    Gambia    Male  All Ages
1004    Gambia  Female  All Ages
1005    Gambia    Both  All Ages
1006    Gambia    Male  All Ages
1007    Gambia  Female  All Ages
1008    Gambia    Both  All Ages
1009     China    Male  All Ages
```

Zadanie 14 przedstawic wybor wierszy z ramki danych pod warunkiem odnosnie okreslonej wartosci kolumny

```python
df[df["location"] == "China"]
```

```
                                    measure location     sex
age  \
1009    DALYs (Disability-Adjusted Life Years)    China    Male  All
Ages
1010    DALYs (Disability-Adjusted Life Years)    China  Female  All
Ages
1011    DALYs (Disability-Adjusted Life Years)    China    Both  All
Ages
1012    DALYs (Disability-Adjusted Life Years)    China    Male  All
Ages
1013    DALYs (Disability-Adjusted Life Years)    China  Female  All
Ages
```

```
...                                                  ...       ...     ...
...
287789  DALYs (Disability-Adjusted Life Years)    China   Female   All
Ages
287790  DALYs (Disability-Adjusted Life Years)    China     Both   All
Ages
287791  DALYs (Disability-Adjusted Life Years)    China     Male   All
Ages
287792  DALYs (Disability-Adjusted Life Years)    China   Female   All
Ages
287793  DALYs (Disability-Adjusted Life Years)    China     Both   All
Ages

                                   cause     metric  year
val \
1009              Neurological disorders     Number  2012  6.647544e+06

1010              Neurological disorders     Number  2012  8.690736e+06

1011              Neurological disorders     Number  2012  1.533828e+07

1012              Neurological disorders    Percent  2012  3.231099e-02

1013              Neurological disorders    Percent  2012  5.528421e-02

...                                  ...        ...   ...           ...

287789  Maternal and neonatal disorders   Percent  2018  2.106277e-02

287790  Maternal and neonatal disorders   Percent  2018  2.016140e-02

287791  Maternal and neonatal disorders      Rate  2018  5.671537e+02

287792  Maternal and neonatal disorders      Rate  2018  5.018387e+02

287793  Maternal and neonatal disorders      Rate  2018  5.351365e+02


               upper          lower
1009     1.054803e+07   4.028985e+06
1010     1.524511e+07   4.364572e+06
1011     2.560954e+07   8.487202e+06
1012     4.975241e-02   2.008808e-02
1013     9.241568e-02   2.910098e-02
...               ...            ...
287789   2.453323e-02   1.777037e-02
287790   2.317598e-02   1.728540e-02
287791   6.433731e+02   5.021475e+02
287792   5.687035e+02   4.445986e+02
```

```
287793  6.039113e+02  4.756582e+02
```

[1449 rows x 10 columns]

Zadanie 15 przedstawic wybor wierszy z ramki danych pod warunkiem spelnieniakilku warunkow jednoczesnie

```python
df[(df["location"] == "China") & (df["sex"] == "Male") & (df["year"] == 2012)]
```

	measure	location	sex	age
1009	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
1012	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
1015	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
1018	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
1021	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
1024	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
3996	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
3999	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
4002	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
11001	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
11004	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
11007	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
12811	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
12814	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
12817	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
13000	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
13003	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
13006	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages
13072	DALYs (Disability-Adjusted Life Years)	China	Male	All Ages

```
13075   DALYs (Disability-Adjusted Life Years)   China  Male  All Ages

13078   DALYs (Disability-Adjusted Life Years)   China  Male  All Ages


                                    cause   metric  year          val  \
1009            Neurological disorders   Number  2012  6.647544e+06
1012            Neurological disorders  Percent  2012  3.231099e-02
1015            Neurological disorders     Rate  2012  9.452140e+02
1018                   Mental disorders   Number  2012  8.964378e+06
1021                   Mental disorders  Percent  2012  4.356153e-02
1024                   Mental disorders     Rate  2012  1.274645e+03
3996        Chronic respiratory diseases   Number  2012  1.299454e+07
3999        Chronic respiratory diseases  Percent  2012  6.332585e-02
4002        Chronic respiratory diseases     Rate  2012  1.847693e+03
11001           Substance use disorders   Number  2012  3.772109e+06
11004           Substance use disorders  Percent  2012  1.834607e-02
11007           Substance use disorders     Rate  2012  5.363560e+02
12811  Skin and subcutaneous diseases   Number  2012  3.812229e+06
12814  Skin and subcutaneous diseases  Percent  2012  1.849193e-02
12817  Skin and subcutaneous diseases     Rate  2012  5.420607e+02
13000               Digestive diseases   Number  2012  6.567530e+06
13003               Digestive diseases  Percent  2012  3.198003e-02
13006               Digestive diseases     Rate  2012  9.338369e+02
13072             Sense organ diseases   Number  2012  6.760562e+06
13075             Sense organ diseases  Percent  2012  3.279114e-02
13078             Sense organ diseases     Rate  2012  9.612840e+02

               upper          lower
1009   1.054803e+07  4.028985e+06
1012   4.975241e-02  2.008808e-02
1015   1.499824e+03  5.728813e+02
1018   1.171418e+07  6.700420e+06
1021   5.443070e-02  3.374576e-02
1024   1.665639e+03  9.527325e+02
3996   1.460151e+07  1.172592e+07
3999   6.856550e-02  5.851620e-02
4002   2.076188e+03  1.667309e+03
11001  4.727222e+06  2.939533e+06
11004  2.191061e-02  1.495828e-02
11007  6.721635e+02  4.179721e+02
12811  5.680035e+06  2.495356e+06
12814  2.588049e-02  1.281703e-02
12817  8.076440e+02  3.548146e+02
13000  7.404799e+06  5.751166e+06
13003  3.398019e-02  3.036259e-02
13006  1.052888e+03  8.177580e+02
13072  9.755552e+06  4.527608e+06
```

```
13075   4.509270e-02   2.364315e-02
13078   1.387142e+03   6.437804e+02
```

Zadanie 16 wybrac wiersze ktore zawieraja w kolumnie kategoryzowanej okreslone slowo

```
df[df["cause"].str.contains("disorder")]
```

```
                                            measure      location     sex
age  \
0        DALYs (Disability-Adjusted Life Years)      Gambia  Female
All Ages
1        DALYs (Disability-Adjusted Life Years)      Gambia    Both
All Ages
2        DALYs (Disability-Adjusted Life Years)      Gambia    Male
All Ages
3        DALYs (Disability-Adjusted Life Years)      Gambia  Female
All Ages
4        DALYs (Disability-Adjusted Life Years)      Gambia    Both
All Ages
...                                              ...         ...     ...
...
307808  DALYs (Disability-Adjusted Life Years)  Madagascar  Female
All Ages
307809  DALYs (Disability-Adjusted Life Years)  Madagascar    Both
All Ages
307810  DALYs (Disability-Adjusted Life Years)  Madagascar    Male
All Ages
307811  DALYs (Disability-Adjusted Life Years)  Madagascar  Female
All Ages
307812  DALYs (Disability-Adjusted Life Years)  Madagascar    Both
All Ages

                                    cause    metric  year          val  \
0        Maternal and neonatal disorders      Rate  2012   7475.212700
1        Maternal and neonatal disorders      Rate  2012   7814.344518
2                Substance use disorders    Number  2012   1659.038707
3                Substance use disorders    Number  2012    874.432466
4                Substance use disorders    Number  2012   2533.471173
...                                  ...       ...   ...          ...
307808           Substance use disorders   Percent  2019      0.002730
307809           Substance use disorders   Percent  2019      0.004518
307810           Substance use disorders      Rate  2019    257.460586
307811           Substance use disorders      Rate  2019    105.102128
307812           Substance use disorders      Rate  2019    181.300662

              upper         lower
0        9104.773541   6157.428603
1        9667.960848   6289.146375
2        2126.829521   1239.172699
3        1186.560596    618.271780
4        3231.220866   1868.204609
```

```
...              ...         ...
307808       0.003577    0.001976
307809       0.005898    0.003337
307810     342.981851  185.670120
307811     140.677940   74.284308
307812     241.019620  130.522645

[70049 rows x 10 columns]
```

Zadanie 17 wybrac wiersze ktore nie zawieraja w kolumnie kategoryzowanej okreslone slowo

```
df[df["cause"].str.contains("disorder") == False]
```

```
                                        measure location     sex
age  \
11      DALYs (Disability-Adjusted Life Years)   Gambia    Male  All
Ages
12      DALYs (Disability-Adjusted Life Years)   Gambia  Female  All
Ages
13      DALYs (Disability-Adjusted Life Years)   Gambia    Both  All
Ages
14      DALYs (Disability-Adjusted Life Years)   Gambia    Male  All
Ages
15      DALYs (Disability-Adjusted Life Years)   Gambia  Female  All
Ages
...                                        ...     ...     ...
...
307835  DALYs (Disability-Adjusted Life Years)   Malawi  Female  All
Ages
307836  DALYs (Disability-Adjusted Life Years)   Malawi    Both  All
Ages
307837  DALYs (Disability-Adjusted Life Years)   Malawi    Male  All
Ages
307838  DALYs (Disability-Adjusted Life Years)   Malawi  Female  All
Ages
307839  DALYs (Disability-Adjusted Life Years)   Malawi    Both  All
Ages

                               cause   metric  year           val  \
11        Diabetes and kidney diseases   Number  2012   8646.107047
12        Diabetes and kidney diseases   Number  2012   7619.998216
13        Diabetes and kidney diseases   Number  2012  16266.105263
14        Diabetes and kidney diseases  Percent  2012      0.019778
15        Diabetes and kidney diseases  Percent  2012      0.019198
...                                ...      ...   ...           ...
307835           Unintentional injuries  Percent  2019      0.022152
307836           Unintentional injuries  Percent  2019      0.026012
307837           Unintentional injuries     Rate  2019   1318.844977
307838           Unintentional injuries     Rate  2019    813.052440
```

```
307839         Unintentional injuries     Rate  2019   1059.021476

                upper           lower
11       10775.702769     6850.455885
12        9230.731994     6143.150103
13       19428.265694    13230.611712
14           0.024331        0.015917
15           0.022440        0.016115
...               ...             ...
307835       0.026646        0.018384
307836       0.031307        0.022140
307837    1717.917036     1037.489279
307838    1042.011143      638.605025
307839    1334.453539      846.543870

[237791 rows x 10 columns]
```

Zadanie 18 utworz kolumne na podstawie istniejacych

```
df["val_int"] = df["val"].astype(int)
df.head()
```

```
                                    measure location      sex
age  \
0  DALYs (Disability-Adjusted Life Years)   Gambia  Female  All Ages

1  DALYs (Disability-Adjusted Life Years)   Gambia    Both  All Ages

2  DALYs (Disability-Adjusted Life Years)   Gambia    Male  All Ages

3  DALYs (Disability-Adjusted Life Years)   Gambia  Female  All Ages

4  DALYs (Disability-Adjusted Life Years)   Gambia    Both  All Ages


                               cause  metric  year          val
upper  \
0  Maternal and neonatal disorders    Rate   2012   7475.212700
9104.773541
1  Maternal and neonatal disorders    Rate   2012   7814.344518
9667.960848
2         Substance use disorders   Number   2012   1659.038707
2126.829521
3         Substance use disorders   Number   2012    874.432466
1186.560596
4         Substance use disorders   Number   2012   2533.471173
3231.220866


        lower  val_int
0  6157.428603     7475
```

```
1  6289.146375      7814
2  1239.172699      1659
3   618.271780       874
4  1868.204609      2533
```

Zadanie 19 usun kolumne

```python
df.drop("val_int", axis=1, inplace = True)
```

Zadanie 20 zmien nazwe kolumny

```python
df.rename(columns = {"location": "country"}, inplace = True)
```

Zadanie 21 zachowaj ramke danych jako plik csv na komputerze

```python
df.to_csv("lab1.csv")
```

```python
df["val"].mean()
```

```
125679.26443763783
```

```python
df['val'].max()
```

```
91933122.11887452
```

```python
df['val'].min()
```

```
5.5394346233794366e-05
```

Zadanie 23 wyswietlic liczbe wierszy

```python
df['year'].count()
```

```
307840
```

Zadanie 24 wyswietlic wartosci unikatowe w kolumnie

```python
df['year'].unique()
```

```
array([2012, 2011, 2013, 2014, 2015, 2016, 2017, 2018, 2019],
dtype=int64)
```

Zadanie 25 wyswietlic liczby rekordow odpowiadajacych do wartosci

```python
df['year'].value_counts()
```

```
2014    40392
2015    40392
2016    40392
2017    40392
2018    40392
2019    40392
2013    39483
2012    24781
```

```
2011      1224
Name: year, dtype: int64
```

Zadanie 26 sortowanie wierszy ramki danych wedlug wartosci okreslonej kolumny (malejaco, rosnaco)

```
df.sort_values(['year'], ascending = True)
```

```
                                        measure      country     sex
age  \
7411    DALYs (Disability-Adjusted Life Years)     Tuvalu  Female
All Ages
3539    DALYs (Disability-Adjusted Life Years)     Comoros    Both
All Ages
3540    DALYs (Disability-Adjusted Life Years)     Comoros    Male
All Ages
3541    DALYs (Disability-Adjusted Life Years)     Comoros  Female
All Ages
3542    DALYs (Disability-Adjusted Life Years)     Comoros    Both
All Ages
...                                         ...          ...     ...
...
273492  DALYs (Disability-Adjusted Life Years)  Azerbaijan    Male
All Ages
273493  DALYs (Disability-Adjusted Life Years)  Azerbaijan  Female
All Ages
273494  DALYs (Disability-Adjusted Life Years)  Azerbaijan    Both
All Ages
273486  DALYs (Disability-Adjusted Life Years)  Azerbaijan    Male
All Ages
307839  DALYs (Disability-Adjusted Life Years)     Malawi    Both
All Ages

                                        cause    metric  year
val  \
7411    Respiratory infections and tuberculosis   Number  2011
111.096155
3539                    Neurological disorders   Number  2011
5577.948609
3540                    Neurological disorders  Percent  2011
0.016442
3541                    Neurological disorders  Percent  2011
0.022104
3542                    Neurological disorders  Percent  2011
0.019182
...                                        ...      ...   ...
...
273492                       Digestive diseases     Rate  2019
1609.174252
273493                       Digestive diseases     Rate  2019
1050.915679
```

```
273494                Digestive diseases      Rate  2019
1330.182663
273486                Digestive diseases    Number  2019
82741.684819
307839             Unintentional injuries      Rate  2019
1059.021476

                  upper
7411         138.731425
3539        8824.083838
3540           0.024979
3541           0.035332
3542           0.029746
...                 ...
273492      2045.271473
273493      1303.737374
273494      1585.542898
273486    105165.246949
307839      1334.453539

[307840 rows x 9 columns]

df.sort_values(['year'], ascending = False)

                                          measure      country     sex
age  \
307839  DALYs (Disability-Adjusted Life Years)      Malawi     Both
All Ages
273487  DALYs (Disability-Adjusted Life Years)  Azerbaijan   Female
All Ages
273495  DALYs (Disability-Adjusted Life Years)  Azerbaijan     Male
All Ages
273494  DALYs (Disability-Adjusted Life Years)  Azerbaijan     Both
All Ages
273493  DALYs (Disability-Adjusted Life Years)  Azerbaijan   Female
All Ages
...                                        ...         ...     ...
...
1056    DALYs (Disability-Adjusted Life Years)    Pakistan     Both
All Ages
1057    DALYs (Disability-Adjusted Life Years)    Pakistan     Male
All Ages
1058    DALYs (Disability-Adjusted Life Years)    Pakistan   Female
All Ages
1059    DALYs (Disability-Adjusted Life Years)    Pakistan     Both
All Ages
7412    DALYs (Disability-Adjusted Life Years)      Tuvalu     Both
All Ages

                                          cause   metric  year
```

```
                                    val  \
307839                  Unintentional injuries     Rate  2019
1.059021e+03
273487                      Digestive diseases   Number  2019
5.398345e+04
273495                  Neurological disorders   Number  2019
5.175890e+04
273494                      Digestive diseases     Rate  2019
1.330183e+03
273493                      Digestive diseases     Rate  2019
1.050916e+03
...                                         ...      ...   ...
...
1056    Respiratory infections and tuberculosis   Number  2011
9.647647e+06
1057    Respiratory infections and tuberculosis  Percent  2011
9.870503e-02
1058    Respiratory infections and tuberculosis  Percent  2011
1.081019e-01
1059    Respiratory infections and tuberculosis  Percent  2011
1.031574e-01
7412    Respiratory infections and tuberculosis   Number  2011
2.359526e+02


                upper
307839   1.334454e+03
273487   6.697040e+04
273495   8.345070e+04
273494   1.585543e+03
273493   1.303737e+03
...               ...
1056     1.101207e+07
1057     1.151398e-01
1058     1.213359e-01
1059     1.145292e-01
7412     2.921261e+02

[307840 rows x 9 columns]
```

Zadanie 27 wyswietlic wierszy dla 10 najwiekszych (najmniejszych) wartosci okreslonej kolumny

```python
df.nlargest(10,'year')
```

```
                                  measure              country
sex  \
221626  DALYs (Disability-Adjusted Life Years)  Trinidad and Tobago
Male
221627  DALYs (Disability-Adjusted Life Years)  Trinidad and Tobago
Female
```

```
221628  DALYs (Disability-Adjusted Life Years)  Trinidad and Tobago
Both
221629  DALYs (Disability-Adjusted Life Years)  Trinidad and Tobago
Male
221630  DALYs (Disability-Adjusted Life Years)  Trinidad and Tobago
Female
221631  DALYs (Disability-Adjusted Life Years)  Trinidad and Tobago
Both
221632  DALYs (Disability-Adjusted Life Years)  Trinidad and Tobago
Male
221633  DALYs (Disability-Adjusted Life Years)  Trinidad and Tobago
Female
221634  DALYs (Disability-Adjusted Life Years)  Trinidad and Tobago
Both
222418  DALYs (Disability-Adjusted Life Years)              Slovenia
Male


              age                                   cause    metric
year  \
221626  All Ages  Neglected tropical diseases and malaria   Number
2019
221627  All Ages  Neglected tropical diseases and malaria   Number
2019
221628  All Ages  Neglected tropical diseases and malaria   Number
2019
221629  All Ages  Neglected tropical diseases and malaria  Percent
2019
221630  All Ages  Neglected tropical diseases and malaria  Percent
2019
221631  All Ages  Neglected tropical diseases and malaria  Percent
2019
221632  All Ages  Neglected tropical diseases and malaria     Rate
2019
221633  All Ages  Neglected tropical diseases and malaria     Rate
2019
221634  All Ages  Neglected tropical diseases and malaria     Rate
2019
222418  All Ages                  Cardiovascular diseases   Number
2019


                 val         upper
221626    733.297685   1085.100546
221627    794.179351   1146.207365
221628   1527.477036   2207.352916
221629      0.002903      0.004233
221630      0.003834      0.005421
221631      0.003322      0.004783
221632    105.331751    155.865132
221633    114.885679    165.809915
```

```
221634     110.091840     159.093419
222418   59811.786386   74598.500667

df.nsmallest(10,'year')
```

	measure	country	sex	age
20	DALYs (Disability-Adjusted Life Years)	Eswatini	Male	All Ages
21	DALYs (Disability-Adjusted Life Years)	Eswatini	Female	All Ages
22	DALYs (Disability-Adjusted Life Years)	Eswatini	Both	All Ages
23	DALYs (Disability-Adjusted Life Years)	Eswatini	Male	All Ages
24	DALYs (Disability-Adjusted Life Years)	Eswatini	Female	All Ages
25	DALYs (Disability-Adjusted Life Years)	Eswatini	Both	All Ages
26	DALYs (Disability-Adjusted Life Years)	Eswatini	Male	All Ages
27	DALYs (Disability-Adjusted Life Years)	Eswatini	Female	All Ages
28	DALYs (Disability-Adjusted Life Years)	Eswatini	Both	All Ages
38	DALYs (Disability-Adjusted Life Years)	Ghana	Male	All Ages

	cause	metric	year	val	upper
20	Cardiovascular diseases	Number	2011	22301.330403	28259.013463
21	Cardiovascular diseases	Number	2011	17657.116393	25364.135521
22	Cardiovascular diseases	Number	2011	39958.446796	52308.684711
23	Cardiovascular diseases	Percent	2011	0.047998	0.057634
24	Cardiovascular diseases	Percent	2011	0.042699	0.057047
25	Cardiovascular diseases	Percent	2011	0.045459	0.055033
26	Cardiovascular diseases	Rate	2011	4306.077971	5456.424042
27	Cardiovascular diseases	Rate	2011	3135.307033	4503.813120
28	Cardiovascular diseases	Rate	2011	3696.181251	4838.585961

```
38  Substance use disorders   Number  2011  35417.948406  45513.639625
```

Zadanie 28 wyswietlic wierszy dla 10 najwiekszych wartosci okreslonej kolumny pod warunkiem okreslonych wartosci innej kolumny

```
df[df['year'] == 2017].nlargest(10,'val')
```

```
                                         measure country     sex
age  \
163438  DALYs (Disability-Adjusted Life Years)   China    Both  All
Ages
182391  DALYs (Disability-Adjusted Life Years)   China    Both  All
Ages
240306  DALYs (Disability-Adjusted Life Years)   India    Both  All
Ages
163436  DALYs (Disability-Adjusted Life Years)   China    Male  All
Ages
198054  DALYs (Disability-Adjusted Life Years)   India    Both  All
Ages
182389  DALYs (Disability-Adjusted Life Years)   China    Male  All
Ages
217639  DALYs (Disability-Adjusted Life Years)   India    Both  All
Ages
163437  DALYs (Disability-Adjusted Life Years)   China  Female  All
Ages
240304  DALYs (Disability-Adjusted Life Years)   India    Male  All
Ages
179874  DALYs (Disability-Adjusted Life Years)   China    Both  All
Ages

                                         cause  metric  year
val  \
163438                     Cardiovascular diseases  Number  2017
8.907722e+07
182391                                   Neoplasms  Number  2017
6.412602e+07
240306                     Cardiovascular diseases  Number  2017
6.199285e+07
163436                     Cardiovascular diseases  Number  2017
5.280536e+07
198054         Maternal and neonatal disorders  Number  2017
4.944075e+07
182389                                   Neoplasms  Number  2017
4.163113e+07
217639  Respiratory infections and tuberculosis  Number  2017
3.756084e+07
163437                     Cardiovascular diseases  Number  2017
3.627186e+07
240304                     Cardiovascular diseases  Number  2017
```

```
3.579288e+07
179874              Musculoskeletal disorders  Number  2017
2.884718e+07
```

```
               upper
163438   9.963379e+07
182391   7.246892e+07
240306   6.884950e+07
163436   6.167300e+07
198054   5.780675e+07
182389   4.917715e+07
217639   4.159206e+07
163437   4.142893e+07
240304   4.070805e+07
179874   3.846334e+07
```

Zadanie 29 grupowanie wierszy wedlug wartosci kolumny kategoryzowanej, potem - usrednienie wartosci wszystkich kolumn w grupie - MultiIndex

```
df.groupby('year').agg('mean')
```

```
C:\Users\Mikołaj\AppData\Local\Temp\ipykernel_8808\1612254600.py:1:
FutureWarning: The default value of numeric_only in
DataFrameGroupBy.mean is deprecated. In a future version, numeric_only
will default to False. Either specify numeric_only or select only
columns which should be valid for the function.
  df.groupby('year').agg('mean')
```

```
              val          upper
year
2011  128072.010514  167694.736793
2012  121301.008264  151381.245415
2013  127595.006984  158634.371413
2014  125758.367947  156849.433086
2015  125897.239359  157614.635786
2016  125761.032250  158048.741593
2017  125589.932835  158689.396038
2018  125696.438186  159978.127268
2019  126113.554017  161551.212355
```

Zadanie 30 grupowanie wierszy wedlug wartosci kolumny kategoryzowanej, potem - usrednienie wartosci dla pewnych kolumn, liczba wartosci i mediana dla pozostalych kolumn w grupach

```
df.groupby('country').agg({'val': ['count'], 'year': ['mean',
'median']})
```

	val	year	
	count	mean	median
country			
Afghanistan	1530	2015.623529	2016.0

```
Albania                           1521  2015.639053  2016.0
Algeria                           1548  2015.581395  2016.0
American Samoa                    1575  2015.508571  2016.0
Andorra                           1449  2015.782609  2016.0
...                                ...          ...     ...
Venezuela (Bolivarian Republic of) 1521  2015.644970  2016.0
Viet Nam                          1467  2015.760736  2016.0
Yemen                             1539  2015.590643  2016.0
Zambia                            1548  2015.569767  2016.0
Zimbabwe                          1449  2015.819876  2016.0

[204 rows x 3 columns]
```

Zadanie 31 wyswietlic nazwy kolumn indeksu zlozonego

```
df2 = df.groupby('country').agg({'val': ['count'], 'year': ['mean',
'median']})
df2.columns

MultiIndex([( 'val',  'count'),
            ('year',   'mean'),
            ('year', 'median')],
           )
```

Zadanie 32 sortowac kolumne indeksu zlozonego

```
df2['val']['count'].sort_values(ascending = False)

country
Monaco                    1620
Benin                     1602
Puerto Rico               1602
Tokelau                   1593
South Sudan               1593
                          ...
Dominican Republic        1431
Kazakhstan                1422
Marshall Islands          1422
Georgia                   1422
Taiwan (Province of China) 1377
Name: count, Length: 204, dtype: int64
```

Zadanie 33 stworzyc tabele przystawna (pivot table) na podstawie ramki danych

```
p_table = df.pivot_table(values='val', index='country',
columns='year', aggfunc='mean',
                margins=False, dropna=True, fill_value=None)
p_table

year                                 2011           2012  \
country
Afghanistan                           NaN   200514.658753
```

```
Albania                              NaN    6422.308807
Algeria                              NaN   81050.750711
American Samoa                981.747175     488.919759
Andorra                              NaN     417.137023
...                                  ...            ...
Venezuela (Bolivarian Republic of)   NaN   85697.191704
Viet Nam                             NaN  239101.129358
Yemen                        50641.336830  140491.688202
Zambia                       33674.818321   91480.389694
Zimbabwe                     39861.178832   35873.510289

year                                2013           2014  \
country
Afghanistan                 166648.277422  170652.779370
Albania                       6979.330969    7821.721755
Algeria                      97761.413017   99187.033780
American Samoa                 584.788261     586.694740
Andorra                        748.270090     578.489642
...                                  ...            ...
Venezuela (Bolivarian Republic of)  81237.996311   82803.589521
Viet Nam                    244222.280082  248373.168669
Yemen                       106344.136263  106959.432953
Zambia                       85125.387107   84321.100478
Zimbabwe                     81210.306050   79508.990677

year                                2015           2016  \
country
Afghanistan                 174423.111532  177983.969843
Albania                       7936.527336    8000.395533
Algeria                     100167.334265  100470.020248
American Samoa                 593.198440     600.746742
Andorra                        585.546372     593.402917
...                                  ...            ...
Venezuela (Bolivarian Republic of)  83801.790943   83061.677288
Viet Nam                    250516.566160  253189.204425
Yemen                       110163.771749  116349.337898
Zambia                       83963.659685   83430.871402
Zimbabwe                     78166.092953   76688.696834

year                                2017           2018  \
country
Afghanistan                 173229.279062  175144.303825
Albania                       8051.663681    8091.695092
Algeria                     100821.983873  101919.739030
American Samoa                 606.808598     612.239655
Andorra                        603.262468     612.539106
...                                  ...            ...
Venezuela (Bolivarian Republic of)  81719.352883   82046.894505
Viet Nam                    255789.044772  258356.470676
Yemen                       119133.677920  122905.988392
```

```
Zambia                                 83234.039033   82178.940517
Zimbabwe                               75477.670094   73612.881938


year                                            2019
country
Afghanistan                            173580.297349
Albania                                  8160.947768
Algeria                                102610.345089
American Samoa                            617.790296
Andorra                                   621.084123
...                                             ...
Venezuela (Bolivarian Republic of)      84380.157668
Viet Nam                               261128.451582
Yemen                                  117978.610849
Zambia                                  81450.482362
Zimbabwe                                73414.169442

[204 rows x 9 columns]
```

Zadanie 34 wyswietlic indeksy i kolumny tabeli przystawnej

```
p_table.index
```

```
Index(['Afghanistan', 'Albania', 'Algeria', 'American Samoa',
'Andorra',
       'Angola', 'Antigua and Barbuda', 'Argentina', 'Armenia',
'Australia',
       ...
       'United States Virgin Islands', 'United States of America',
'Uruguay',
       'Uzbekistan', 'Vanuatu', 'Venezuela (Bolivarian Republic of)',
       'Viet Nam', 'Yemen', 'Zambia', 'Zimbabwe'],
      dtype='object', name='country', length=204)
```

```
p_table.columns
```

```
Int64Index([2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019],
dtype='int64', name='year')
```

Zadanie 35 utworz indeks zlozony tabeli przystawnej i wyswietl go

```
p_table = df.pivot_table(values='val', index=['country', 'measure'],
columns='year', aggfunc='mean',
                         margins=False, dropna=True, fill_value=None)
p_table
```

```
year
2011  \
country                                measure

Afghanistan                            DALYs (Disability-Adjusted Life
Years)              NaN
```

```
Albania                            DALYs (Disability-Adjusted Life
Years)            NaN
Algeria                            DALYs (Disability-Adjusted Life
Years)            NaN
American Samoa                     DALYs (Disability-Adjusted Life
Years)     981.747175
Andorra                            DALYs (Disability-Adjusted Life
Years)            NaN
...
...
Venezuela (Bolivarian Republic of) DALYs (Disability-Adjusted Life
Years)            NaN
Viet Nam                           DALYs (Disability-Adjusted Life
Years)            NaN
Yemen                              DALYs (Disability-Adjusted Life
Years)  50641.336830
Zambia                             DALYs (Disability-Adjusted Life
Years)  33674.818321
Zimbabwe                           DALYs (Disability-Adjusted Life
Years)  39861.178832

year
2012  \
country                            measure

Afghanistan                        DALYs (Disability-Adjusted Life
Years)  200514.658753
Albania                            DALYs (Disability-Adjusted Life
Years)    6422.308807
Algeria                            DALYs (Disability-Adjusted Life
Years)   81050.750711
American Samoa                     DALYs (Disability-Adjusted Life
Years)     488.919759
Andorra                            DALYs (Disability-Adjusted Life
Years)     417.137023
...
...
Venezuela (Bolivarian Republic of) DALYs (Disability-Adjusted Life
Years)   85697.191704
Viet Nam                           DALYs (Disability-Adjusted Life
Years)  239101.129358
Yemen                              DALYs (Disability-Adjusted Life
Years)  140491.688202
Zambia                             DALYs (Disability-Adjusted Life
Years)   91480.389694
Zimbabwe                           DALYs (Disability-Adjusted Life
Years)   35873.510289

year
2013  \
```

```
country                                 measure

Afghanistan                             DALYs (Disability-Adjusted Life
Years)  166648.277422
Albania                                 DALYs (Disability-Adjusted Life
Years)    6979.330969
Algeria                                 DALYs (Disability-Adjusted Life
Years)   97761.413017
American Samoa                          DALYs (Disability-Adjusted Life
Years)     584.788261
Andorra                                 DALYs (Disability-Adjusted Life
Years)     748.270090
...
...
Venezuela (Bolivarian Republic of) DALYs (Disability-Adjusted Life
Years)   81237.996311
Viet Nam                                DALYs (Disability-Adjusted Life
Years)  244222.280082
Yemen                                   DALYs (Disability-Adjusted Life
Years)  106344.136263
Zambia                                  DALYs (Disability-Adjusted Life
Years)   85125.387107
Zimbabwe                                DALYs (Disability-Adjusted Life
Years)   81210.306050

year
2014  \
country                                 measure

Afghanistan                             DALYs (Disability-Adjusted Life
Years)  170652.779370
Albania                                 DALYs (Disability-Adjusted Life
Years)    7821.721755
Algeria                                 DALYs (Disability-Adjusted Life
Years)   99187.033780
American Samoa                          DALYs (Disability-Adjusted Life
Years)     586.694740
Andorra                                 DALYs (Disability-Adjusted Life
Years)     578.489642
...
...
Venezuela (Bolivarian Republic of) DALYs (Disability-Adjusted Life
Years)   82803.589521
Viet Nam                                DALYs (Disability-Adjusted Life
Years)  248373.168669
Yemen                                   DALYs (Disability-Adjusted Life
Years)  106959.432953
Zambia                                  DALYs (Disability-Adjusted Life
Years)   84321.100478
Zimbabwe                                DALYs (Disability-Adjusted Life
```

```
Years)     79508.990677

year
2015    \
country                                   measure

Afghanistan                               DALYs (Disability-Adjusted Life
Years)   174423.111532
Albania                                   DALYs (Disability-Adjusted Life
Years)     7936.527336
Algeria                                   DALYs (Disability-Adjusted Life
Years)   100167.334265
American Samoa                            DALYs (Disability-Adjusted Life
Years)      593.198440
Andorra                                   DALYs (Disability-Adjusted Life
Years)      585.546372
                                          ...
                                          ...
Venezuela (Bolivarian Republic of) DALYs (Disability-Adjusted Life
Years)    83801.790943
Viet Nam                                  DALYs (Disability-Adjusted Life
Years)   250516.566160
Yemen                                     DALYs (Disability-Adjusted Life
Years)   110163.771749
Zambia                                    DALYs (Disability-Adjusted Life
Years)    83963.659685
Zimbabwe                                  DALYs (Disability-Adjusted Life
Years)    78166.092953

year
2016    \
country                                   measure

Afghanistan                               DALYs (Disability-Adjusted Life
Years)   177983.969843
Albania                                   DALYs (Disability-Adjusted Life
Years)     8000.395533
Algeria                                   DALYs (Disability-Adjusted Life
Years)   100470.020248
American Samoa                            DALYs (Disability-Adjusted Life
Years)      600.746742
Andorra                                   DALYs (Disability-Adjusted Life
Years)      593.402917
                                          ...
                                          ...
Venezuela (Bolivarian Republic of) DALYs (Disability-Adjusted Life
Years)    83061.677288
Viet Nam                                  DALYs (Disability-Adjusted Life
Years)   253189.204425
Yemen                                     DALYs (Disability-Adjusted Life
```

```
Years)  116349.337898
Zambia                                DALYs (Disability-Adjusted Life
Years)   83430.871402
Zimbabwe                              DALYs (Disability-Adjusted Life
Years)   76688.696834

year
2017  \
country                                measure

Afghanistan                           DALYs (Disability-Adjusted Life
Years)  173229.279062
Albania                               DALYs (Disability-Adjusted Life
Years)    8051.663681
Algeria                               DALYs (Disability-Adjusted Life
Years)  100821.983873
American Samoa                        DALYs (Disability-Adjusted Life
Years)     606.808598
Andorra                               DALYs (Disability-Adjusted Life
Years)     603.262468
...
...
Venezuela (Bolivarian Republic of) DALYs (Disability-Adjusted Life
Years)   81719.352883
Viet Nam                              DALYs (Disability-Adjusted Life
Years)  255789.044772
Yemen                                 DALYs (Disability-Adjusted Life
Years)  119133.677920
Zambia                                DALYs (Disability-Adjusted Life
Years)   83234.039033
Zimbabwe                              DALYs (Disability-Adjusted Life
Years)   75477.670094

year
2018  \
country                                measure

Afghanistan                           DALYs (Disability-Adjusted Life
Years)  175144.303825
Albania                               DALYs (Disability-Adjusted Life
Years)    8091.695092
Algeria                               DALYs (Disability-Adjusted Life
Years)  101919.739030
American Samoa                        DALYs (Disability-Adjusted Life
Years)     612.239655
Andorra                               DALYs (Disability-Adjusted Life
Years)     612.539106
...
...
Venezuela (Bolivarian Republic of) DALYs (Disability-Adjusted Life
```

```
Years)   82046.894505
Viet Nam                                DALYs (Disability-Adjusted Life
Years)  258356.470676
Yemen                                   DALYs (Disability-Adjusted Life
Years)  122905.988392
Zambia                                  DALYs (Disability-Adjusted Life
Years)   82178.940517
Zimbabwe                                DALYs (Disability-Adjusted Life
Years)   73612.881938

year
2019
country                                 measure

Afghanistan                             DALYs (Disability-Adjusted Life
Years)  173580.297349
Albania                                 DALYs (Disability-Adjusted Life
Years)    8160.947768
Algeria                                 DALYs (Disability-Adjusted Life
Years)  102610.345089
American Samoa                          DALYs (Disability-Adjusted Life
Years)     617.790296
Andorra                                 DALYs (Disability-Adjusted Life
Years)     621.084123
...
...
Venezuela (Bolivarian Republic of) DALYs (Disability-Adjusted Life
Years)   84380.157668
Viet Nam                                DALYs (Disability-Adjusted Life
Years)  261128.451582
Yemen                                   DALYs (Disability-Adjusted Life
Years)  117978.610849
Zambia                                  DALYs (Disability-Adjusted Life
Years)   81450.482362
Zimbabwe                                DALYs (Disability-Adjusted Life
Years)   73414.169442
```

[204 rows x 9 columns]

Zadanie 36 zaimportuj modul pyplot z biblioteki matplotlib

```
import matplotlib.pyplot as plt
```

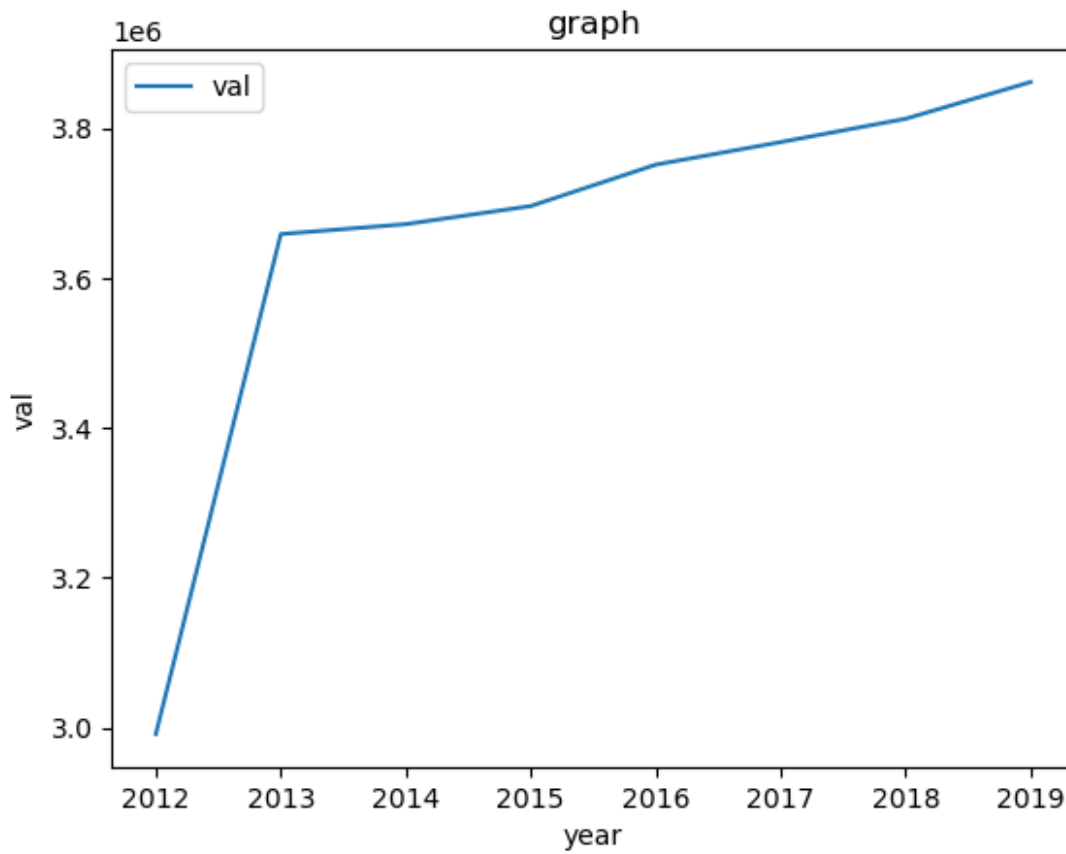Zadaie 37 wskazac, ze wykresy nalezy rysowac bezposrednio w zeszycie, a nie w osobnej zakladce

```
%matplotlib inline
```

Zadanie 38 wyswietlic wykres na podstawie tabeli przystawnej

```
df[(df['country'] == 'China')].pivot_table(values='val', index='year',
aggfunc='mean',
                                           fill_value=None,
margins=False, dropna=True).plot(kind = 'line')
plt.ylabel('val')
plt.title('graph')
plt.show()
```
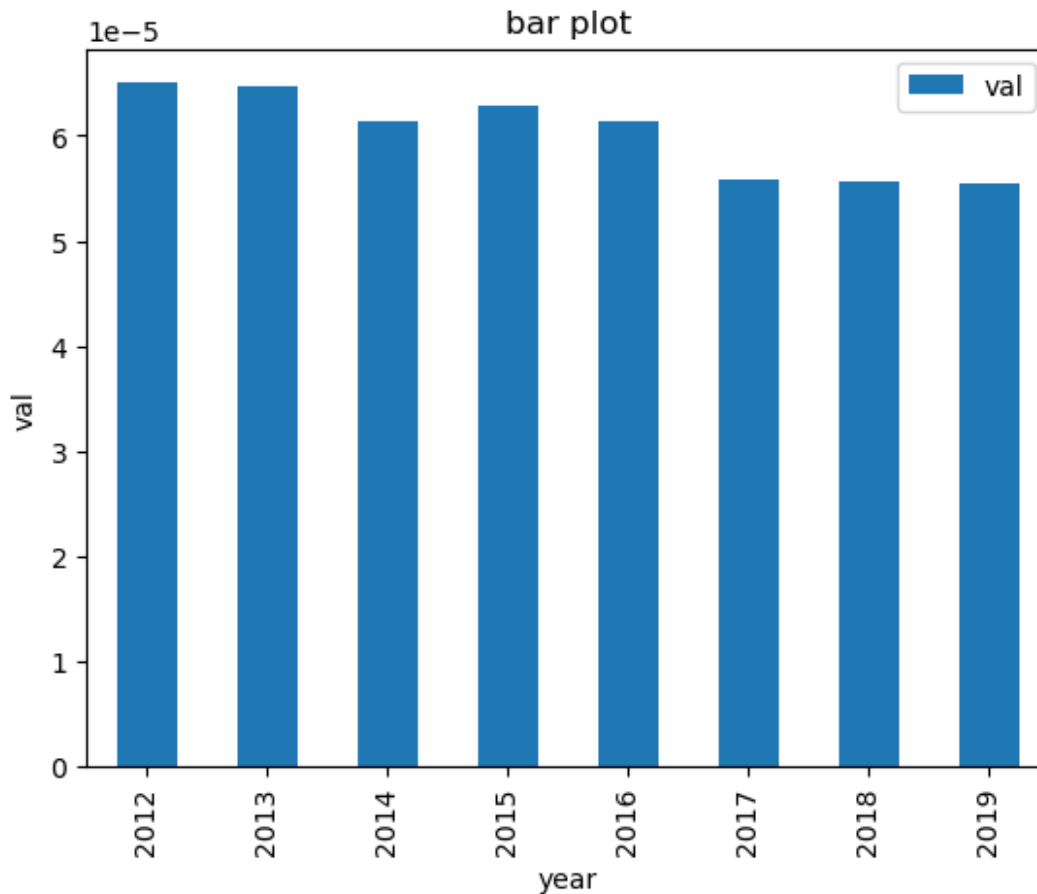

Zadanie 39 narysowac histogram na podstawie wartosci kolumny

```
df_bar = df[(df['year'] >
2011)].pivot_table(values='val',index='year',
aggfunc='min',fill_value=None, margins=False, dropna=True)
df_bar.plot(kind = 'bar')
plt.ylabel('val')
plt.title('bar plot')
```

```
Text(0.5, 1.0, 'bar plot')
```

Zadanie 40 przedstawic sposoby laczenia ramek danych za pomocą metod merge i concat

```python
df1 = pd.read_csv("IHME-GBD_2019_DATA-15798851-2.csv", encoding='utf-
8', engine='python')
df2 = pd.read_csv("IHME-GBD_2019_DATA-ff08d9bc-1.csv", encoding='utf-
8', engine='python')

df1.head()
```

```
                                     measure location     sex
age  \
0  DALYs (Disability-Adjusted Life Years)   Gambia  Female  All Ages

1  DALYs (Disability-Adjusted Life Years)   Gambia    Both  All Ages

2  DALYs (Disability-Adjusted Life Years)   Gambia    Male  All Ages

3  DALYs (Disability-Adjusted Life Years)   Gambia  Female  All Ages

4  DALYs (Disability-Adjusted Life Years)   Gambia    Both  All Ages


                    cause   metric  year           val
```

```
                                 upper  \
0  Maternal and neonatal disorders     Rate   2012   7475.212700
9104.773541
1  Maternal and neonatal disorders     Rate   2012   7814.344518
9667.960848
2          Substance use disorders   Number   2012   1659.038707
2126.829521
3          Substance use disorders   Number   2012    874.432466
1186.560596
4          Substance use disorders   Number   2012   2533.471173
3231.220866

           lower
0   6157.428603
1   6289.146375
2   1239.172699
3    618.271780
4   1868.204609
```

`df2.head()`

```
   measure location    sex       age                            cause
metric  \
0  Deaths    Samoa    Male  All Ages    Chronic respiratory diseases
Rate
1  Deaths    Samoa  Female  All Ages    Chronic respiratory diseases
Rate
2  Deaths    Samoa    Both  All Ages    Chronic respiratory diseases
Rate
3  Deaths    Samoa    Male  All Ages  Skin and subcutaneous diseases
Rate
4  Deaths    Samoa  Female  All Ages  Skin and subcutaneous diseases
Rate

   year        val      upper       lower
0  2000  64.470214  81.808307   53.476793
1  2000  55.234399  77.883497   39.978647
2  2000  60.039961  76.013089   49.941986
3  2000   2.246741   3.467454    1.438979
4  2000   1.368385   1.945448    0.866099
```

```python
df1.rename(columns = {'val': 'val1', 'upper':'upper1',
'lower':'lower1'}, inplace = True)
df2.rename(columns = {'val': 'val2', 'upper': 'upper2',
'lower':'lower2'}, inplace = True)
df_both = pd.merge(df1, df2, on = ['location', 'sex', 'age', 'cause',
'metric', 'year'], how = 'inner')
df_both.head()
```

```
                               measure_x location      sex
age  \
```

```
0   DALYs (Disability-Adjusted Life Years)    Gambia  Female  All Ages

1   DALYs (Disability-Adjusted Life Years)    Gambia    Both  All Ages

2   DALYs (Disability-Adjusted Life Years)    Gambia    Male  All Ages

3   DALYs (Disability-Adjusted Life Years)    Gambia  Female  All Ages

4   DALYs (Disability-Adjusted Life Years)    Gambia    Both  All Ages


                              cause metric  year         val1
upper1  \
0  Maternal and neonatal disorders    Rate  2012  7475.212700
9104.773541
1  Maternal and neonatal disorders    Rate  2012  7814.344518
9667.960848
2          Substance use disorders    Rate  2012   179.493660
230.104586
3          Substance use disorders    Rate  2012    91.400543
124.025911
4          Substance use disorders    Rate  2012   134.688036
171.782808

        lower1 measure_y        val2      upper2      lower2
0  6157.428603    Deaths   89.867238  110.176286   74.001324
1  6289.146375    Deaths   89.365457  111.497307   71.539316
2   134.067784    Deaths    1.101136    1.545407    0.786638
3    64.625204    Deaths    0.227136    0.296917    0.173921
4    99.320179    Deaths    0.656605    0.896138    0.490594
```

```python
#concat
df_part1 = df_both.iloc[:50000,:]
df_part2 = df_both.iloc[50000:,:]
df_both_2 = pd.concat([df_part1, df_part2], axis = 0)
df_both_2.head()
```

```
                              measure_x location     sex
age  \
0  DALYs (Disability-Adjusted Life Years)    Gambia  Female  All Ages

1  DALYs (Disability-Adjusted Life Years)    Gambia    Both  All Ages

2  DALYs (Disability-Adjusted Life Years)    Gambia    Male  All Ages

3  DALYs (Disability-Adjusted Life Years)    Gambia  Female  All Ages

4  DALYs (Disability-Adjusted Life Years)    Gambia    Both  All Ages
```

```
                             cause metric   year         val1
upper1  \
0  Maternal and neonatal disorders   Rate   2012   7475.212700
9104.773541
1  Maternal and neonatal disorders   Rate   2012   7814.344518
9667.960848
2         Substance use disorders   Rate   2012    179.493660
230.104586
3         Substance use disorders   Rate   2012     91.400543
124.025911
4         Substance use disorders   Rate   2012    134.688036
171.782808

        lower1 measure_y       val2     upper2      lower2
0   6157.428603    Deaths  89.867238  110.176286  74.001324
1   6289.146375    Deaths  89.365457  111.497307  71.539316
2    134.067784    Deaths   1.101136    1.545407   0.786638
3     64.625204    Deaths   0.227136    0.296917   0.173921
4     99.320179    Deaths   0.656605    0.896138   0.490594
```

Zadanie 41 pokazac dodawanie nowych kolumn za pomoca operacji matematycznych

```python
df_both["val2_round"] = df_both["val2"].round(decimals = 1)
df_both.head()
```

```
                            measure_x location     sex
age  \
0  DALYs (Disability-Adjusted Life Years)   Gambia  Female  All Ages

1  DALYs (Disability-Adjusted Life Years)   Gambia    Both  All Ages

2  DALYs (Disability-Adjusted Life Years)   Gambia    Male  All Ages

3  DALYs (Disability-Adjusted Life Years)   Gambia  Female  All Ages

4  DALYs (Disability-Adjusted Life Years)   Gambia    Both  All Ages


                             cause metric   year         val1
upper1  \
0  Maternal and neonatal disorders   Rate   2012   7475.212700
9104.773541
1  Maternal and neonatal disorders   Rate   2012   7814.344518
9667.960848
2         Substance use disorders   Rate   2012    179.493660
230.104586
3         Substance use disorders   Rate   2012     91.400543
124.025911
4         Substance use disorders   Rate   2012    134.688036
```

171.782808

	lower1	measure_y	val2	upper2	lower2	val2_round
0	6157.428603	Deaths	89.867238	110.176286	74.001324	89.9
1	6289.146375	Deaths	89.365457	111.497307	71.539316	89.4
2	134.067784	Deaths	1.101136	1.545407	0.786638	1.1
3	64.625204	Deaths	0.227136	0.296917	0.173921	0.2
4	99.320179	Deaths	0.656605	0.896138	0.490594	0.7

```python
df_both["sum"] = df_both["val2"] + df_both["upper2"] +
df_both["upper1"]
df_both.head()
```

	measure_x	location	sex	age
0	DALYs (Disability-Adjusted Life Years)	Gambia	Female	All Ages
1	DALYs (Disability-Adjusted Life Years)	Gambia	Both	All Ages
2	DALYs (Disability-Adjusted Life Years)	Gambia	Male	All Ages
3	DALYs (Disability-Adjusted Life Years)	Gambia	Female	All Ages
4	DALYs (Disability-Adjusted Life Years)	Gambia	Both	All Ages

	cause	metric	year	val1	upper1
0	Maternal and neonatal disorders	Rate	2012	7475.212700	9104.773541
1	Maternal and neonatal disorders	Rate	2012	7814.344518	9667.960848
2	Substance use disorders	Rate	2012	179.493660	230.104586
3	Substance use disorders	Rate	2012	91.400543	124.025911
4	Substance use disorders	Rate	2012	134.688036	171.782808

	lower1	measure_y	val2	upper2	lower2	val2_round
0	6157.428603	Deaths	89.867238	110.176286	74.001324	89.9

```
1   6289.146375    Deaths   89.365457   111.497307   71.539316          89.4

2    134.067784    Deaths    1.101136     1.545407    0.786638           1.1

3     64.625204    Deaths    0.227136     0.296917    0.173921           0.2

4     99.320179    Deaths    0.656605     0.896138    0.490594           0.7


            sum
0   9304.817065
1   9868.823613
2    232.751130
3    124.549965
4    173.335552
```

```python
df_both["sum2"] = df_both.loc[:, "val2":"lower2"].sum(axis = 1)
df_both.head()
```

```
                                  measure_x location     sex
age  \
0  DALYs (Disability-Adjusted Life Years)   Gambia  Female  All Ages

1  DALYs (Disability-Adjusted Life Years)   Gambia    Both  All Ages

2  DALYs (Disability-Adjusted Life Years)   Gambia    Male  All Ages

3  DALYs (Disability-Adjusted Life Years)   Gambia  Female  All Ages

4  DALYs (Disability-Adjusted Life Years)   Gambia    Both  All Ages


                            cause metric  year          val1
upper1  \
0  Maternal and neonatal disorders   Rate  2012  7475.212700
9104.773541
1  Maternal and neonatal disorders   Rate  2012  7814.344518
9667.960848
2          Substance use disorders   Rate  2012   179.493660
230.104586
3          Substance use disorders   Rate  2012    91.400543
124.025911
4          Substance use disorders   Rate  2012   134.688036
171.782808

        lower1 measure_y        val2       upper2      lower2  val2_round
\
0  6157.428603    Deaths  89.867238   110.176286   74.001324        89.9
```

```
1   6289.146375     Deaths   89.365457   111.497307   71.539316           89.4

2    134.067784     Deaths    1.101136     1.545407    0.786638            1.1

3     64.625204     Deaths    0.227136     0.296917    0.173921            0.2

4     99.320179     Deaths    0.656605     0.896138    0.490594            0.7


         sum          sum2
0  9304.817065   274.044849
1  9868.823613   272.402080
2   232.751130     3.433182
3   124.549965     0.697975
4   173.335552     2.043338
```

Zadanie 42 przedstawic na przykladzie dodawanie nowych kolumn z pomoca funkcji lambda

```python
c = ['Poland', 'Ukraine', 'Italy', 'Germany', 'Hungary', 'Spain',
'Romania']

df_both['c'] = df_both['location'].apply(lambda x: True if x in c else
False )
df_both[df_both['c'] == True]
```

```
                                    measure_x location      sex
age  \
176    DALYs (Disability-Adjusted Life Years)  Romania     Male  All
Ages
177    DALYs (Disability-Adjusted Life Years)  Romania   Female  All
Ages
178    DALYs (Disability-Adjusted Life Years)  Romania     Both  All
Ages
526    DALYs (Disability-Adjusted Life Years)  Romania     Male  All
Ages
527    DALYs (Disability-Adjusted Life Years)  Romania   Female  All
Ages
...                                       ...      ...      ...   .
..
95835  DALYs (Disability-Adjusted Life Years)  Ukraine   Female  All
Ages
95836  DALYs (Disability-Adjusted Life Years)  Ukraine     Both  All
Ages
95837  DALYs (Disability-Adjusted Life Years)  Ukraine     Male  All
Ages
95838  DALYs (Disability-Adjusted Life Years)  Ukraine   Female  All
Ages
95839  DALYs (Disability-Adjusted Life Years)  Ukraine     Both  All
Ages
```

	cause	metric	year	val1
176	Digestive diseases	Rate	2013	2697.039450
177	Digestive diseases	Rate	2013	1688.287627
178	Digestive diseases	Rate	2013	2179.809770
526	HIV/AIDS and sexually transmitted infections	Rate	2012	62.211141
527	HIV/AIDS and sexually transmitted infections	Rate	2012	48.638280
...	...	...	...	...
95835	Substance use disorders	Rate	2019	524.280165
95836	Substance use disorders	Rate	2019	1379.972560
95837	Diabetes and kidney diseases	Rate	2019	859.632605
95838	Diabetes and kidney diseases	Rate	2019	837.557575
95839	Diabetes and kidney diseases	Rate	2019	847.727158

	upper1	lower1	measure_y	val2	upper2	lower2
176	2834.479777	2582.161628	Deaths	85.838243	88.899467	82.905470
177	1864.552118	1540.753908	Deaths	54.320030	56.658301	51.582158
178	2326.710260	2053.355455	Deaths	69.677523	72.085742	67.023604
526	67.330638	57.963919	Deaths	0.980503	1.028893	0.938133
527	63.003805	39.937237	Deaths	0.490602	0.521284	0.461069
...	...	...	...	...	...	...
95835	647.200517	417.812871	Deaths	6.579509	8.642640	4.910730
95836	1639.511644	1161.683891	Deaths	20.873053	25.610960	16.900271
95837	1054.725757	690.504711	Deaths	12.709024	15.294875	10.358222
95838	1045.022774	649.471184	Deaths	12.318098	14.897062	9.951266
95839	1044.016288	678.838388	Deaths	12.498191	14.465368	10.848537

```
           c
176    True
177    True
178    True
526    True
527    True
...     ...
95835  True
95836  True
95837  True
95838  True
95839  True

[3348 rows x 15 columns]
```

Zadanie 43 przedstawic mozliwosci pracy z duzymi plikami przy uzyciu argumentu chunksize

```python
df_both.to_csv('df_both.csv')

for chunk_df in pd.read_csv('df_both.csv',
                  chunksize = 50000):
    print("CHUNK DF")
    print(chunk_df.tail())
```

```
CHUNK DF
       Unnamed: 0                                measure_x location
sex  \
49995       49995  DALYs (Disability-Adjusted Life Years)    India
Male
49996       49996  DALYs (Disability-Adjusted Life Years)    India
Female
49997       49997  DALYs (Disability-Adjusted Life Years)    India
Both
49998       49998  DALYs (Disability-Adjusted Life Years)    India
Male
49999       49999  DALYs (Disability-Adjusted Life Years)    India
Female


           age                     cause metric  year
val1  \
49995  All Ages        Enteric infections   Rate  2016  1710.169269

49996  All Ages        Enteric infections   Rate  2016  2375.331799

49997  All Ages        Enteric infections   Rate  2016  2033.915925

49998  All Ages  Other infectious diseases   Rate  2016   907.819033
```

49999 All Ages Other infectious diseases   Rate 2016 1012.889263

             upper1   lower1 measure_y   val2   upper2
lower2 \
49995 2676.033040 1276.800636   Deaths 40.978011   77.739106
28.491105
49996 3653.659090 1547.850615   Deaths 67.649038 121.943819
33.118259
49997 2792.605894 1509.210547   Deaths 53.959283   85.027199
34.567153
49998 1183.267342   708.664221   Deaths 13.481944   17.237128
10.830031
49999 1359.953116   794.781479   Deaths 14.679474   18.969057
11.730752

           c
49995 False
49996 False
49997 False
49998 False
49999 False
CHUNK DF
     Unnamed: 0   measure_x location
sex \
97971   97971 DALYs (Disability-Adjusted Life Years)   Malawi
Female
97972   97972 DALYs (Disability-Adjusted Life Years)   Malawi
Both
97973   97973 DALYs (Disability-Adjusted Life Years)   Malawi
Male
97974   97974 DALYs (Disability-Adjusted Life Years)   Malawi
Female
97975   97975 DALYs (Disability-Adjusted Life Years)   Malawi
Both

           age   cause metric year
val1 \
97971 All Ages Chronic respiratory diseases   Rate 2019
602.342558
97972 All Ages Chronic respiratory diseases   Rate 2019
635.274862
97973 All Ages   Unintentional injuries   Rate 2019
1318.844977
97974 All Ages   Unintentional injuries   Rate 2019
813.052440
97975 All Ages   Unintentional injuries   Rate 2019
1059.021476

           upper1   lower1 measure_y   val2   upper2

```
       lower2  \
97971   732.856075    491.019574    Deaths   10.783192   13.484120
8.267941
97972   755.292643    530.146531    Deaths   13.025678   15.412885
10.664151
97973  1717.917036   1037.489279    Deaths   20.824550   27.026330
16.310791
97974  1042.011143    638.605025    Deaths   11.865525   15.430914
9.199714
97975  1334.453539    846.543870    Deaths   16.222336   20.617369
12.836190

           c
97971  False
97972  False
97973  False
97974  False
97975  False
```