EiWD_zadanie_01

October 23, 2021

1 Zadanie 1

[53]:
```
# ładowanie biblioteki Pandas
import pandas as pd
```

[54]:
```
# tworzenie ramki danych ze słownika
dane1 = {
    "liczby": [532, 2532, 523, 2543, 235, 231, 34, 23552, 5324],
    "kategorie": ["a", "b", "c", "d", "e", "f", "g", "h", "i"]
}

df1 = pd.DataFrame(dane1)
df1
```

[54]:

	liczby	kategorie
0	532	a
1	2532	b
2	523	c
3	2543	d
4	235	e
5	231	f
6	34	g
7	23552	h
8	5324	i

[55]:
```
# zapisanie ramki z danymi pobranymi z pliku w pliku .csv
sciezka = r"C:\Users\piotr\Downloads\dane_ze_slownika.csv"

df1.to_csv(sciezka, encoding="utf-8")
```

[56]:
```
# tworzenie ramki danych z listy list
dane2 = [
    [532, 2532, 523, 2543, 235, 231, 34, 23552, 5324],
    ["a", "b", "c", "d", "e", "f", "g", "h", "i"]
]

df2 = pd.DataFrame(dane2)
df2
```

[56]:

	0	1	2	3	4	5	6	7	8
0	532	2532	523	2543	235	231	34	23552	5324
1	a	b	c	d	e	f	g	h	i

[57]:
```
# wczytanie danych z pliku .csv
sciezka = r"C:
 ↪\Users\piotr\Downloads\IHME_DAH_DATABASE_1990_2020_CSV_1\IHME_DAH_DATABASE_1990_2020_Y2021M
 ↪CSV"

df = pd.read_csv(sciezka, low_memory=False, )

# pierwsze 10 wierszy ramki
df.head(10)
```

[57]:

	year	source	channel	recipient_isocode	\
0	1990	Australia	BIL_AUS	AGO	
1	1990	Australia	BIL_AUS	BDI	
2	1990	Australia	BIL_AUS	BEN	
3	1990	Australia	BIL_AUS	BFA	
4	1990	Australia	BIL_AUS	BWA	
5	1990	Australia	BIL_AUS	CAF	
6	1990	Australia	BIL_AUS	CHN	
7	1990	Australia	BIL_AUS	CIV	
8	1990	Australia	BIL_AUS	CMR	
9	1990	Australia	BIL_AUS	COD	

	recipient_country	gbd_location_id	wb_regioncode	\
0	Angola	168	SSA	
1	Burundi	175	SSA	
2	Benin	200	SSA	
3	Burkina Faso	201	SSA	
4	Botswana	193	SSA	
5	Central African Republic	169	SSA	
6	China	6	EAP	
7	Cote d'Ivoire	205	SSA	
8	Cameroon	202	SSA	
9	Democratic Republic of the Congo	171	SSA	

	wb_location_id	gbd_region	gbd_region_id	…	\
0	242	Sub-Saharan Africa, Central	167.0	…	
1	242	Sub-Saharan Africa, Eastern	174.0	…	
2	242	Sub-Saharan Africa, Western	199.0	…	
3	242	Sub-Saharan Africa, Western	199.0	…	
4	242	Sub-Saharan Africa, Southern	192.0	…	
5	242	Sub-Saharan Africa, Central	167.0	…	
6	239	Asia, East	5.0	…	
7	242	Sub-Saharan Africa, Western	199.0	…	

```
8             242   Sub-Saharan Africa, Western          199.0  …
9             242   Sub-Saharan Africa, Central          167.0  …

  other_dah_20  rmh_dah_20  nch_dah_20  ncd_dah_20 hiv_dah_20 mal_dah_20  \
0            0           5           0           0          7          3
1            0           6           0           0          5          1
2            0           6           0           0          5          2
3            0           5           0           0          7          2
4            0           1           0           0         23          -
5            0           1           0           0          2          -
6           38           7           0           0        367          5
7            0           5           0           0         18          1
8            0           2           0           0          8          1
9            0          14           0           0         18          6

  tb_dah_20 swap_hss_total_dah_20 oid_dah_20 unalloc_dah_20
0         0                     0          0             -
1         0                     0          0             0
2         0                     0          0             0
3         0                     0          0             0
4         0                     0          0             -
5         0                     0          0             -
6         0                   130          0             0
7         0                     0          0             0
8         0                     0          0             -
9         0                     0          0             0

[10 rows x 76 columns]
```

[58]:
```
# ostatnie 10 wierszy ramki danych
df.tail(10)
```

[58]:
```
        year         source  channel recipient_isocode  \
384296  2020  United_States  INTLNGO               QZA
384297  2020  United_States      NGO               QZA
384298  2020  United_States     PAHO               QZA
384299  2020  United_States   UNAIDS               QZA
384300  2020  United_States    UNFPA               QZA
384301  2020  United_States   UNICEF               QZA
384302  2020  United_States  UNITAID               QZA
384303  2020  United_States  UNITAID               QZA
384304  2020  United_States   WB_IDA               QZA
384305  2020  United_States      WHO               QZA

              recipient_country  gbd_location_id wb_regioncode  \
384296  Unallocated/Unspecified            44598           NaN
384297  Unallocated/Unspecified            44598           NaN
```

```
384298  Unallocated/Unspecified            44598           NaN
384299  Unallocated/Unspecified            44598           NaN
384300  Unallocated/Unspecified            44598           NaN
384301  Unallocated/Unspecified            44598           NaN
384302  Unallocated/Unspecified            44598           NaN
384303  Unallocated/Unspecified            44598           NaN
384304  Unallocated/Unspecified            44598           NaN
384305  Unallocated/Unspecified            44598           NaN

        wb_location_id              gbd_region  gbd_region_id  …  \
384296           44621  Unallocated/Unspecified        44598.0  …
384297           44621  Unallocated/Unspecified        44598.0  …
384298           44621  Unallocated/Unspecified        44598.0  …
384299           44621  Unallocated/Unspecified        44598.0  …
384300           44621  Unallocated/Unspecified        44598.0  …
384301           44621  Unallocated/Unspecified        44598.0  …
384302           44621  Unallocated/Unspecified        44598.0  …
384303           44621  Unallocated/Unspecified        44598.0  …
384304           44621  Unallocated/Unspecified        44598.0  …
384305           44621  Unallocated/Unspecified        44598.0  …

       other_dah_20  rmh_dah_20  nch_dah_20  ncd_dah_20 hiv_dah_20 mal_dah_20  \
384296       132174        2351        4047         120       5090        623
384297      1293825      230205      148340       11652     264680      20224
384298        13449           0        1121        9526          0          0
384299            0           0           0           0      57674          0
384300            0           0           0           0          0          0
384301            0       17509      265811           0       9398          0
384302            0           0           0           0       1037       1085
384303            0           0           0           0       1250        166
384304        49485       17337       26498        1907       5076        763
384305        17823       13250       32712       22436      11098       8991

       tb_dah_20 swap_hss_total_dah_20 oid_dah_20 unalloc_dah_20
384296       765                    25        640             0
384297     13315                 98494      61832             0
384298         0                 58465      21481             -
384299         0                     0       2885             -
384300         0                     0          0             0
384301         0                     0     116214             -
384302       895                   449        407             -
384303       112                   173        217             0
384304      3424                 46292      47569             -
384305      9457                193266     113937             0

[10 rows x 76 columns]
```

[59]:
```
# transponowanie danych
df_transposed = df.T

df_transposed.head()
```

[59]:

	0	1	2	3	4	\
year	1990	1990	1990	1990	1990	
source	Australia	Australia	Australia	Australia	Australia	
channel	BIL_AUS	BIL_AUS	BIL_AUS	BIL_AUS	BIL_AUS	
recipient_isocode	AGO	BDI	BEN	BFA	BWA	
recipient_country	Angola	Burundi	Benin	Burkina Faso	Botswana	

	5	6	7	\
year	1990	1990	1990	
source	Australia	Australia	Australia	
channel	BIL_AUS	BIL_AUS	BIL_AUS	
recipient_isocode	CAF	CHN	CIV	
recipient_country	Central African Republic	China	Cote d'Ivoire	

	8	9	…	\
year	1990	1990	…	
source	Australia	Australia	…	
channel	BIL_AUS	BIL_AUS	…	
recipient_isocode	CMR	COD	…	
recipient_country	Cameroon	Democratic Republic of the Congo	…	

	384296	384297	\
year	2020	2020	
source	United_States	United_States	
channel	INTLNGO	NGO	
recipient_isocode	QZA	QZA	
recipient_country	Unallocated/Unspecified	Unallocated/Unspecified	

	384298	384299	\
year	2020	2020	
source	United_States	United_States	
channel	PAHO	UNAIDS	
recipient_isocode	QZA	QZA	
recipient_country	Unallocated/Unspecified	Unallocated/Unspecified	

	384300	384301	\
year	2020	2020	
source	United_States	United_States	
channel	UNFPA	UNICEF	
recipient_isocode	QZA	QZA	
recipient_country	Unallocated/Unspecified	Unallocated/Unspecified	

```
                                    384302                    384303  \
year                                  2020                      2020
source                         United_States             United_States
channel                              UNITAID                   UNITAID
recipient_isocode                        QZA                       QZA
recipient_country  Unallocated/Unspecified  Unallocated/Unspecified

                                    384304                    384305
year                                  2020                      2020
source                         United_States             United_States
channel                             WB_IDA                       WHO
recipient_isocode                        QZA                       QZA
recipient_country  Unallocated/Unspecified  Unallocated/Unspecified

[5 rows x 384306 columns]
```

[60]:
```
# wyświetlenie informacji o ramce danych

df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 384306 entries, 0 to 384305
Data columns (total 76 columns):
 #   Column                 Non-Null Count   Dtype
---  ------                 --------------   -----
 0   year                   384306 non-null  int64
 1   source                 384306 non-null  object
 2   channel                384306 non-null  object
 3   recipient_isocode      384306 non-null  object
 4   recipient_country      383773 non-null  object
 5   gbd_location_id        384306 non-null  int64
 6   wb_regioncode          370318 non-null  object
 7   wb_location_id         384306 non-null  int64
 8   gbd_region             383993 non-null  object
 9   gbd_region_id          383993 non-null  float64
 10  gbd_superregion        383993 non-null  object
 11  gbd_superregion_id     383993 non-null  float64
 12  elim_ch                384306 non-null  int64
 13  prelim_est             384306 non-null  int64
 14  dah_20                 384306 non-null  object
 15  rmh_fp_dah_20          384306 non-null  object
 16  rmh_mh_dah_20          384306 non-null  object
 17  rmh_hss_other_dah_20   384306 non-null  object
 18  rmh_hss_hrh_dah_20     384306 non-null  object
 19  rmh_other_dah_20       384306 non-null  object
 20  nch_cnn_dah_20         384306 non-null  object
 21  nch_cnv_dah_20         384306 non-null  object
 22  nch_other_dah_20       384306 non-null  object
```

```
23  nch_hss_other_dah_20   384306 non-null  object
24  nch_hss_hrh_dah_20     384306 non-null  object
25  hiv_treat_dah_20       384306 non-null  object
26  hiv_prev_dah_20        384306 non-null  object
27  hiv_pmtct_dah_20       384306 non-null  object
28  hiv_other_dah_20       384306 non-null  object
29  hiv_ct_dah_20          384306 non-null  object
30  hiv_ovc_dah_20         384306 non-null  object
31  hiv_care_dah_20        384306 non-null  object
32  hiv_hss_other_dah_20   384306 non-null  object
33  hiv_hss_hrh_dah_20     384306 non-null  object
34  hiv_amr_dah_20         384306 non-null  object
35  mal_diag_dah_20        384306 non-null  object
36  mal_hss_other_dah_20   384306 non-null  object
37  mal_hss_hrh_dah_20     384306 non-null  object
38  mal_con_nets_dah_20    384306 non-null  object
39  mal_con_irs_dah_20     384306 non-null  object
40  mal_con_oth_dah_20     384306 non-null  object
41  mal_treat_dah_20       384306 non-null  object
42  mal_comm_con_dah_20    384306 non-null  object
43  mal_other_dah_20       384306 non-null  object
44  mal_amr_dah_20         384306 non-null  object
45  tb_other_dah_20        384306 non-null  object
46  tb_treat_dah_20        384306 non-null  object
47  tb_diag_dah_20         384306 non-null  object
48  tb_hss_other_dah_20    384306 non-null  object
49  tb_hss_hrh_dah_20      384306 non-null  object
50  tb_amr_dah_20          384306 non-null  object
51  oid_hss_other_dah_20   384306 non-null  object
52  oid_hss_hrh_dah_20     384306 non-null  object
53  oid_ebz_dah_20         384306 non-null  object
54  oid_zika_dah_20        384306 non-null  object
55  oid_covid_dah_20       384306 non-null  object
56  oid_other_dah_20       384306 non-null  object
57  oid_amr_dah_20         384306 non-null  object
58  ncd_hss_other_dah_20   384306 non-null  object
59  ncd_hss_hrh_dah_20     384306 non-null  object
60  ncd_tobac_dah_20       384306 non-null  object
61  ncd_mental_dah_20      384306 non-null  object
62  ncd_other_dah_20       384306 non-null  object
63  swap_hss_other_dah_20  384306 non-null  object
64  swap_hss_hrh_dah_20    384306 non-null  object
65  swap_hss_pp_dah_20     384306 non-null  object
66  other_dah_20           384306 non-null  object
67  rmh_dah_20             384306 non-null  object
68  nch_dah_20             384306 non-null  object
69  ncd_dah_20             384306 non-null  object
70  hiv_dah_20             384306 non-null  object
```

```
 71  mal_dah_20             384306 non-null  object
 72  tb_dah_20              384306 non-null  object
 73  swap_hss_total_dah_20  384306 non-null  object
 74  oid_dah_20             384306 non-null  object
 75  unalloc_dah_20         384306 non-null  object
dtypes: float64(2), int64(5), object(69)
memory usage: 222.8+ MB
```

[61]:
```
# wyświetlenie liczby wierszy i kolumn zawartych w ramce danych

rows, cols = df.shape
print(f"Kolumn: {cols}")
print(f"Wierszy: {rows}")
```

```
Kolumn: 76
Wierszy: 384306
```

[62]:
```
# przedstawienie wybranych wierszy i kolumn przy użyciu nazw i indeksów

df["year"]
```

[62]:
```
0         1990
1         1990
2         1990
3         1990
4         1990
          …
384301    2020
384302    2020
384303    2020
384304    2020
384305    2020
Name: year, Length: 384306, dtype: int64
```

[63]:
```
df.year
```

[63]:
```
0         1990
1         1990
2         1990
3         1990
4         1990
          …
384301    2020
384302    2020
384303    2020
384304    2020
384305    2020
Name: year, Length: 384306, dtype: int64
```

[64]:
```
df[["year", "source", "recipient_country"]]
```

[64]:

	year	source	recipient_country
0	1990	Australia	Angola
1	1990	Australia	Burundi
2	1990	Australia	Benin
3	1990	Australia	Burkina Faso
4	1990	Australia	Botswana
…	…	…	…
384301	2020	United_States	Unallocated/Unspecified
384302	2020	United_States	Unallocated/Unspecified
384303	2020	United_States	Unallocated/Unspecified
384304	2020	United_States	Unallocated/Unspecified
384305	2020	United_States	Unallocated/Unspecified

[384306 rows x 3 columns]

[65]:
```
df.loc[:, "year":"channel"] # wszystkie wiersze, kolumny od "year" do "channel" włącznie
```

[65]:

	year	source	channel
0	1990	Australia	BIL_AUS
1	1990	Australia	BIL_AUS
2	1990	Australia	BIL_AUS
3	1990	Australia	BIL_AUS
4	1990	Australia	BIL_AUS
…	…	…	…
384301	2020	United_States	UNICEF
384302	2020	United_States	UNITAID
384303	2020	United_States	UNITAID
384304	2020	United_States	WB_IDA
384305	2020	United_States	WHO

[384306 rows x 3 columns]

[66]:
```
df.iloc[0:10, 2:4] # wiersze od pierwszego do dziesiątego, kolumny od trzeciej do czwartej
```

[66]:

	channel	recipient_isocode
0	BIL_AUS	AGO
1	BIL_AUS	BDI
2	BIL_AUS	BEN
3	BIL_AUS	BFA
4	BIL_AUS	BWA
5	BIL_AUS	CAF
6	BIL_AUS	CHN
7	BIL_AUS	CIV

```
8  BIL_AUS              CMR
9  BIL_AUS              COD
```

[67]:
```
# wyświetlenie podstawowych informacji statystycznych o kolumnach liczbowych␣
↪(liczba wartości niepowtarzalnych, średnia, odchylenie standardowe, minimum,␣
↪maksimum, wartości kwartyli)

liczbowe = df.select_dtypes(include='number')
liczbowe.describe()
```

[67]:
```
                year  gbd_location_id  wb_location_id  gbd_region_id  \
count  384306.000000    384306.000000   384306.000000  383993.000000
mean     2008.127521      1765.935533     2240.752439    1745.812671
std         6.945191      8325.915434     9204.906147    8328.525983
min      1990.000000         1.000000      239.000000       1.000000
25%      2004.000000       110.000000      241.000000      96.000000
50%      2009.000000       169.000000      242.000000     159.000000
75%      2014.000000       200.000000      242.000000     192.000000
max      2020.000000     44598.000000    44621.000000   44598.000000

       gbd_superregion_id        elim_ch     prelim_est
count       383993.000000  384306.000000  384306.000000
mean          1733.144388       0.252052       0.014358
std           8330.949734       0.434191       0.118963
min              1.000000       0.000000       0.000000
25%             64.000000       0.000000       0.000000
50%            158.000000       0.000000       0.000000
75%            166.000000       1.000000       0.000000
max          44598.000000       1.000000       1.000000
```

[68]:
```
# wyświetlenie podstawowych informacji statystycznych o kolumnach␣
↪kategoryzowanych (liczba wartości niepowtarzalnych, wartość najczęstsza,␣
↪liczba wystąpień wartości najczęstszej)

kategoryzowane = df.select_dtypes(exclude='number')
kategoryzowane.describe()
```

[68]:
```
                source channel recipient_isocode       recipient_country  \
count           384306  384306            384306                  383773
unique              32      46               176                     174
top     United_Kingdom   GFATM            INKIND  Administrative expenses
freq             20452   57306              8379                    8379

       wb_regioncode                  gbd_region     gbd_superregion  dah_20  \
count         370318                      383993              383993  384306
unique             7                          22                   9   18258
top              SSA  Sub-Saharan Africa, Western  Sub-Saharan Africa       -
```

```
freq          172395                          72174              168660   54305

       rmh_fp_dah_20 rmh_mh_dah_20  … other_dah_20 rmh_dah_20 nch_dah_20  \
count         384306        384306  …       384306     384306     384306
unique          4066          4370  …         7735       7144       7718
top                0             0  …            0          0          0
freq          350068        314771  …       216332     260201     249416

       ncd_dah_20 hiv_dah_20 mal_dah_20 tb_dah_20 swap_hss_total_dah_20  \
count      384306     384306     384306    384306                384306
unique       2306       7787       4590      3618                  6833
top             0          0          0         0                     0
freq       338336     208921     305384    301461                254927

       oid_dah_20 unalloc_dah_20
count      384306         384306
unique       4217            855
top             0              0
freq       295748         225397

[4 rows x 69 columns]
```

[69]:
```
# usunięcie brakujących wartości z ramki danych

df.dropna(inplace=True) # inplace=True - zmodyfikowane zostaną oryginalne dane
df.head()
```

[69]:
```
   year     source  channel recipient_isocode recipient_country  \
0  1990  Australia  BIL_AUS               AGO            Angola
1  1990  Australia  BIL_AUS               BDI           Burundi
2  1990  Australia  BIL_AUS               BEN             Benin
3  1990  Australia  BIL_AUS               BFA      Burkina Faso
4  1990  Australia  BIL_AUS               BWA          Botswana

   gbd_location_id wb_regioncode  wb_location_id  \
0              168           SSA             242
1              175           SSA             242
2              200           SSA             242
3              201           SSA             242
4              193           SSA             242

                    gbd_region  gbd_region_id  … other_dah_20  rmh_dah_20  \
0   Sub-Saharan Africa, Central          167.0  …            0           5
1   Sub-Saharan Africa, Eastern          174.0  …            0           6
2   Sub-Saharan Africa, Western          199.0  …            0           6
3   Sub-Saharan Africa, Western          199.0  …            0           5
4  Sub-Saharan Africa, Southern          192.0  …            0           1
```

```
   nch_dah_20  ncd_dah_20 hiv_dah_20 mal_dah_20 tb_dah_20  \
0           0           0          7          3         0
1           0           0          5          1         0
2           0           0          5          2         0
3           0           0          7          2         0
4           0           0         23          -         0

  swap_hss_total_dah_20 oid_dah_20 unalloc_dah_20
0                     0          0              -
1                     0          0              0
2                     0          0              0
3                     0          0              0
4                     0          0              -

[5 rows x 76 columns]
```

[70]:
```
# przedstawić  wiersze ramki danych spełniające podany warunek dotyczące␣
↪wybranej kolumny

df[df["recipient_country"] == "Poland"]
```

[70]:

	year	source	channel	recipient_isocode	\
700	1990	Debt_repayments	WB_IBRD	POL	
4369	1991	Debt_repayments	WB_IBRD	POL	
8096	1992	Debt_repayments	WB_IBRD	POL	
10645	1992	Private_other	US_FOUND	POL	
12497	1993	Debt_repayments	WB_IBRD	POL	
15497	1993	Private_other	US_FOUND	POL	
17569	1994	Debt_repayments	WB_IBRD	POL	
21942	1995	Debt_repayments	WB_IBRD	POL	
26712	1996	Debt_repayments	WB_IBRD	POL	
29710	1996	Private_other	US_FOUND	POL	
32012	1997	Debt_repayments	WB_IBRD	POL	
35384	1997	Private_other	US_FOUND	POL	
37780	1998	Debt_repayments	WB_IBRD	POL	
42091	1998	Private_other	US_FOUND	POL	
45270	1999	Debt_repayments	WB_IBRD	POL	
49232	1999	Private_other	US_FOUND	POL	
52973	2000	Debt_repayments	WB_IBRD	POL	
57335	2000	Private_other	US_FOUND	POL	
60891	2001	Debt_repayments	WB_IBRD	POL	
65873	2001	Private_other	US_FOUND	POL	
69771	2002	Debt_repayments	WB_IBRD	POL	
75368	2002	Private_other	US_FOUND	POL	
80075	2003	Debt_repayments	WB_IBRD	POL	
86779	2003	Private_other	US_FOUND	POL	

```
92465   2004           Debt_repayments   WB_IBRD              POL
100248  2004             Private_other  US_FOUND              POL
106477  2005           Debt_repayments   WB_IBRD              POL
114734  2005             Private_other  US_FOUND              POL
122273  2006           Debt_repayments   WB_IBRD              POL
133423  2006             Private_other  US_FOUND              POL
141779  2007           Debt_repayments   WB_IBRD              POL
150228  2007    Non_OECD_DAC_countries       EEA              POL
150762  2007                    Norway       EEA              POL
151596  2007  Other_OECD_DAC_countries       EEA              POL
152756  2007             Private_other       NGO              POL
152904  2007             Private_other  US_FOUND              POL
161587  2008           Debt_repayments   WB_IBRD              POL
170889  2008    Non_OECD_DAC_countries       EEA              POL
171401  2008                    Norway       EEA              POL
172307  2008  Other_OECD_DAC_countries       EEA              POL
173531  2008             Private_other  US_FOUND              POL

       recipient_country  gbd_location_id wb_regioncode  wb_location_id  \
700               Poland               51           ECA             240
4369              Poland               51           ECA             240
8096              Poland               51           ECA             240
10645             Poland               51           ECA             240
12497             Poland               51           ECA             240
15497             Poland               51           ECA             240
17569             Poland               51           ECA             240
21942             Poland               51           ECA             240
26712             Poland               51           ECA             240
29710             Poland               51           ECA             240
32012             Poland               51           ECA             240
35384             Poland               51           ECA             240
37780             Poland               51           ECA             240
42091             Poland               51           ECA             240
45270             Poland               51           ECA             240
49232             Poland               51           ECA             240
52973             Poland               51           ECA             240
57335             Poland               51           ECA             240
60891             Poland               51           ECA             240
65873             Poland               51           ECA             240
69771             Poland               51           ECA             240
75368             Poland               51           ECA             240
80075             Poland               51           ECA             240
86779             Poland               51           ECA             240
92465             Poland               51           ECA             240
100248            Poland               51           ECA             240
106477            Poland               51           ECA             240
114734            Poland               51           ECA             240
```

122273	Poland	51	ECA	240	
133423	Poland	51	ECA	240	
141779	Poland	51	ECA	240	
150228	Poland	51	ECA	240	
150762	Poland	51	ECA	240	
151596	Poland	51	ECA	240	
152756	Poland	51	ECA	240	
152904	Poland	51	ECA	240	
161587	Poland	51	ECA	240	
170889	Poland	51	ECA	240	
171401	Poland	51	ECA	240	
172307	Poland	51	ECA	240	
173531	Poland	51	ECA	240	

	gbd_region	gbd_region_id	…	other_dah_20	rmh_dah_20	\
700	Europe, Central	42.0	…	0	0	
4369	Europe, Central	42.0	…	0	0	
8096	Europe, Central	42.0	…	0	0	
10645	Europe, Central	42.0	…	225	7	
12497	Europe, Central	42.0	…	1536	1491	
15497	Europe, Central	42.0	…	26	0	
17569	Europe, Central	42.0	…	3430	3329	
21942	Europe, Central	42.0	…	3658	3550	
26712	Europe, Central	42.0	…	7902	7670	
29710	Europe, Central	42.0	…	41	146	
32012	Europe, Central	42.0	…	13369	12976	
35384	Europe, Central	42.0	…	41	0	
37780	Europe, Central	42.0	…	320	310	
42091	Europe, Central	42.0	…	56	0	
45270	Europe, Central	42.0	…	1085	1053	
49232	Europe, Central	42.0	…	6	318	
52973	Europe, Central	42.0	…	598	580	
57335	Europe, Central	42.0	…	511	0	
60891	Europe, Central	42.0	…	8029	7793	
65873	Europe, Central	42.0	…	259	0	
69771	Europe, Central	42.0	…	5606	5441	
75368	Europe, Central	42.0	…	0	0	
80075	Europe, Central	42.0	…	0	0	
86779	Europe, Central	42.0	…	18	0	
92465	Europe, Central	42.0	…	0	0	
100248	Europe, Central	42.0	…	333	0	
106477	Europe, Central	42.0	…	0	0	
114734	Europe, Central	42.0	…	294	98	
122273	Europe, Central	42.0	…	0	0	
133423	Europe, Central	42.0	…	426	3	
141779	Europe, Central	42.0	…	0	0	
150228	Europe, Central	42.0	…	0	1	

```
150762  Europe, Central           42.0  …            0        1170
151596  Europe, Central           42.0  …            0           3
152756  Europe, Central           42.0  …            0           0
152904  Europe, Central           42.0  …          646           0
161587  Europe, Central           42.0  …            0           0
170889  Europe, Central           42.0  …            0           1
171401  Europe, Central           42.0  …            0         795
172307  Europe, Central           42.0  …            0           2
173531  Europe, Central           42.0  …          181           0

        nch_dah_20  ncd_dah_20 hiv_dah_20 mal_dah_20 tb_dah_20  \
700              0           0          0          0         0
4369             0           0          0          0         0
8096             0           0          0          0         0
10645           29           0          0          0         0
12497            0           0          0          0         0
15497            0         132          0          0         0
17569            0           0          0          0         0
21942            0           0          0          0         0
26712            0           0          0          0         0
29710            0          43          0          0         0
32012            0           0          0          0         0
35384            0           0          0          0         0
37780            0           0          0          0         0
42091            0         468          0          0         0
45270            0           0          0          0         0
49232            0          24          0          0         0
52973            0           0          0          0         0
57335            0           0          0          0         0
60891            0           0          0          0         0
65873            0         650          0          0         0
69771            0           0          0          0         0
75368           21         335          0          0         0
80075            0           0          0          0         0
86779            0         330          0          0         0
92465            0           0          0          0         0
100248          70          50          0          0         0
106477           0           0          0          0         0
114734         293          33          0          0         0
122273           0           0          0          0         0
133423           3           0          0          0         0
141779           0           0          0          0         0
150228           1           1          0          0         0
150762         907        1266          0          0         0
151596           2           3          0          0         0
152756           0          69          0          0         0
152904           0         246          0          0         0
```

161587	0	0	0	0	0
170889	2	2	0	0	0
171401	1585	2476	0	0	0
172307	4	6	0	0	0
173531	0	0	0	0	0

	swap_hss_total_dah_20	oid_dah_20	unalloc_dah_20
700	0	0	0
4369	0	0	0
8096	0	0	0
10645	0	0	-
12497	1491	0	0
15497	0	0	0
17569	3329	0	0
21942	3550	0	-
26712	7670	0	0
29710	0	0	0
32012	12976	0	-
35384	0	0	0
37780	310	0	-
42091	0	0	0
45270	1053	0	0
49232	0	0	0
52973	580	0	0
57335	0	0	0
60891	7793	0	-
65873	0	0	0
69771	5441	0	-
75368	0	66	0
80075	0	0	0
86779	0	0	0
92465	0	0	0
100248	0	0	0
106477	0	0	0
114734	0	0	0
122273	0	0	0
133423	0	0	0
141779	0	0	0
150228	1	0	0
150762	672	0	-
151596	2	0	0
152756	0	0	0
152904	0	0	0
161587	0	0	0
170889	1	0	-
171401	1268	0	-
172307	3	0	-

```
173531                    0          0              0

[41 rows x 76 columns]
```

[71]:
```
# wybranie wierszy ramki danych spełniających podane warunki jednocześnie

df[(df["recipient_isocode"] == "POL") & (df["year"] >= 2005)]
```

[71]:

	year	source	channel	recipient_isocode	\
106477	2005	Debt_repayments	WB_IBRD	POL	
114734	2005	Private_other	US_FOUND	POL	
122273	2006	Debt_repayments	WB_IBRD	POL	
133423	2006	Private_other	US_FOUND	POL	
141779	2007	Debt_repayments	WB_IBRD	POL	
150228	2007	Non_OECD_DAC_countries	EEA	POL	
150762	2007	Norway	EEA	POL	
151596	2007	Other_OECD_DAC_countries	EEA	POL	
152756	2007	Private_other	NGO	POL	
152904	2007	Private_other	US_FOUND	POL	
161587	2008	Debt_repayments	WB_IBRD	POL	
170889	2008	Non_OECD_DAC_countries	EEA	POL	
171401	2008	Norway	EEA	POL	
172307	2008	Other_OECD_DAC_countries	EEA	POL	
173531	2008	Private_other	US_FOUND	POL	

	recipient_country	gbd_location_id	wb_regioncode	wb_location_id	\
106477	Poland	51	ECA	240	
114734	Poland	51	ECA	240	
122273	Poland	51	ECA	240	
133423	Poland	51	ECA	240	
141779	Poland	51	ECA	240	
150228	Poland	51	ECA	240	
150762	Poland	51	ECA	240	
151596	Poland	51	ECA	240	
152756	Poland	51	ECA	240	
152904	Poland	51	ECA	240	
161587	Poland	51	ECA	240	
170889	Poland	51	ECA	240	
171401	Poland	51	ECA	240	
172307	Poland	51	ECA	240	
173531	Poland	51	ECA	240	

	gbd_region	gbd_region_id	…	other_dah_20	rmh_dah_20	\
106477	Europe, Central	42.0	…	0	0	
114734	Europe, Central	42.0	…	294	98	
122273	Europe, Central	42.0	…	0	0	
133423	Europe, Central	42.0	…	426	3	

```
141779  Europe, Central          42.0  …            0           0
150228  Europe, Central          42.0  …            0           1
150762  Europe, Central          42.0  …            0        1170
151596  Europe, Central          42.0  …            0           3
152756  Europe, Central          42.0  …            0           0
152904  Europe, Central          42.0  …          646           0
161587  Europe, Central          42.0  …            0           0
170889  Europe, Central          42.0  …            0           1
171401  Europe, Central          42.0  …            0         795
172307  Europe, Central          42.0  …            0           2
173531  Europe, Central          42.0  …          181           0

        nch_dah_20  ncd_dah_20 hiv_dah_20 mal_dah_20 tb_dah_20  \
106477           0           0          0          0         0
114734         293          33          0          0         0
122273           0           0          0          0         0
133423           3           0          0          0         0
141779           0           0          0          0         0
150228           1           1          0          0         0
150762         907        1266          0          0         0
151596           2           3          0          0         0
152756           0          69          0          0         0
152904           0         246          0          0         0
161587           0           0          0          0         0
170889           2           2          0          0         0
171401        1585        2476          0          0         0
172307           4           6          0          0         0
173531           0           0          0          0         0

       swap_hss_total_dah_20 oid_dah_20 unalloc_dah_20
106477                     0          0              0
114734                     0          0              0
122273                     0          0              0
133423                     0          0              0
141779                     0          0              0
150228                     1          0              0
150762                   672          0              -
151596                     2          0              0
152756                     0          0              0
152904                     0          0              0
161587                     0          0              0
170889                     1          0              -
171401                  1268          0              -
172307                     3          0              -
173531                     0          0              0

[15 rows x 76 columns]
```

[72]:
```
# wybrać wiersze które zawierają w kolumnie kategoryzowanej określone słowo

df[df["source"]=="Japan"]
```

[72]:

```
        year source  channel recipient_isocode   recipient_country  \
1917    1990  Japan  BIL_JPN               COL            Colombia
1918    1990  Japan  BIL_JPN               DJI            Djibouti
1919    1990  Japan  BIL_JPN               DOM  Dominican Republic
1920    1990  Japan  BIL_JPN               ETH            Ethiopia
1921    1990  Japan  BIL_JPN               GRD             Grenada
…        …     …       …                   …                  …
374701  2018  Japan      WHO               PAK            Pakistan
374702  2018  Japan      WHO               PSE           Palestine
374704  2018  Japan      WHO               SSD         South Sudan
374705  2018  Japan      WHO               SYR               Syria
374706  2018  Japan      WHO               YEM               Yemen

        gbd_location_id wb_regioncode  wb_location_id  \
1917                125           LAC             241
1918                177           MNA             243
1919                111           LAC             241
1920                179           SSA             242
1921                112           LAC             241
…                   …             …               …
374701              165           SAS             244
374702              149           MNA             243
374704              435           SSA             242
374705              153           MNA             243
374706              157           MNA             243

                       gbd_region  gbd_region_id  … other_dah_20  \
1917         Latin America, Central          124.0  …         4926
1918    Sub-Saharan Africa, Eastern          174.0  …         5084
1919                      Caribbean          104.0  …            0
1920    Sub-Saharan Africa, Eastern          174.0  …        24440
1921                      Caribbean          104.0  …        12080
…                               …              …  …            …
374701                 Asia, South           159.0  …            0
374702    North Africa/Middle East           138.0  …            0
374704  Sub-Saharan Africa, Eastern          174.0  …            0
374705    North Africa/Middle East           138.0  …            0
374706    North Africa/Middle East           138.0  …         1548

        rmh_dah_20  nch_dah_20  ncd_dah_20 hiv_dah_20 mal_dah_20 tb_dah_20  \
1917             0           0           0          0          0         0
1918             0           0           0          0          0         0
1919             0           0           0          0          0         0
```

```
1920              0           0           0          0          0          0
1921              0           0           0          0          0          0
…               …           …           …          …          …          …
374701            0           0           0          0          0          0
374702            0           0           0          0          0          0
374704            0           0           0          0          0          0
374705            0        3613        1548          0          0          0
374706            0           0           0          0          0          0

       swap_hss_total_dah_20 oid_dah_20 unalloc_dah_20
1917                       0          0              0
1918                       0          0              0
1919                       0      12709              0
1920                       0          0              0
1921                       0          0              0
…                        …          …              …
374701                     0       3318              0
374702                   516          0              0
374704                  1032          0              0
374705                    71          0              0
374706                     0          0              0

[11008 rows x 76 columns]
```

[73]:
```
# wybrać wiersze które nie zawierają w kolumnie kategoryzowanej określone słowo

selection = df[df["gbd_region"]!="Asia, East"]
selection.head()
```

[73]:
```
   year     source  channel recipient_isocode recipient_country  \
0  1990  Australia  BIL_AUS               AGO            Angola
1  1990  Australia  BIL_AUS               BDI           Burundi
2  1990  Australia  BIL_AUS               BEN             Benin
3  1990  Australia  BIL_AUS               BFA      Burkina Faso
4  1990  Australia  BIL_AUS               BWA          Botswana

   gbd_location_id wb_regioncode  wb_location_id  \
0              168           SSA             242
1              175           SSA             242
2              200           SSA             242
3              201           SSA             242
4              193           SSA             242

                     gbd_region  gbd_region_id  … other_dah_20  rmh_dah_20  \
0   Sub-Saharan Africa, Central          167.0  …             0           5
1   Sub-Saharan Africa, Eastern          174.0  …             0           6
2   Sub-Saharan Africa, Western          199.0  …             0           6
```

```
3   Sub-Saharan Africa, Western          199.0  …            0           5
4  Sub-Saharan Africa, Southern          192.0  …            0           1

   nch_dah_20  ncd_dah_20 hiv_dah_20 mal_dah_20 tb_dah_20  \
0           0           0          7          3         0
1           0           0          5          1         0
2           0           0          5          2         0
3           0           0          7          2         0
4           0           0         23          -         0

  swap_hss_total_dah_20 oid_dah_20 unalloc_dah_20
0                     0          0              -
1                     0          0              0
2                     0          0              0
3                     0          0              0
4                     0          0              -

[5 rows x 76 columns]
```

[74]:
```
# utwórz kolumnę na podstawie istniejącej

recipient_countries = df.recipient_country
recipient_countries
```

[74]:
```
0              Angola
1             Burundi
2               Benin
3        Burkina Faso
4            Botswana
               …
383233         Global
383234          Yemen
383235   South Africa
383236         Zambia
383237       Zimbabwe
Name: recipient_country, Length: 369785, dtype: object
```

[75]:
```
# usuń kolumnę

df_copy = df
df_copy.drop(["source", "channel"], axis=1, inplace=True) # usuwa kolumny␣
↪"source" i "channel"
df_copy.head()
```

[75]:
```
   year recipient_isocode recipient_country  gbd_location_id wb_regioncode  \
0  1990               AGO           Angola              168           SSA
1  1990               BDI          Burundi              175           SSA
```

```
2  1990              BEN             Benin              200          SSA
3  1990              BFA      Burkina Faso              201          SSA
4  1990              BWA          Botswana              193          SSA

   wb_location_id                   gbd_region  gbd_region_id  \
0             242   Sub-Saharan Africa, Central          167.0
1             242   Sub-Saharan Africa, Eastern          174.0
2             242   Sub-Saharan Africa, Western          199.0
3             242   Sub-Saharan Africa, Western          199.0
4             242  Sub-Saharan Africa, Southern          192.0

      gbd_superregion  gbd_superregion_id  …  other_dah_20  rmh_dah_20  \
0  Sub-Saharan Africa               166.0  …             0           5
1  Sub-Saharan Africa               166.0  …             0           6
2  Sub-Saharan Africa               166.0  …             0           6
3  Sub-Saharan Africa               166.0  …             0           5
4  Sub-Saharan Africa               166.0  …             0           1

  nch_dah_20 ncd_dah_20 hiv_dah_20 mal_dah_20 tb_dah_20 swap_hss_total_dah_20  \
0          0          0          7          3         0                      0
1          0          0          5          1         0                      0
2          0          0          5          2         0                      0
3          0          0          7          2         0                      0
4          0          0         23          -         0                      0

  oid_dah_20 unalloc_dah_20
0          0              -
1          0              0
2          0              0
3          0              0
4          0              -

[5 rows x 74 columns]
```

[76]:
```
# zmień nazwę kolumny

df_copy.rename(columns={"year": "rok", "recipient_isocode": "kod_iso_odbiorcy",␣
 ↪"recipient_country": "kraj_odbiorcy"}, inplace=True)
df_copy.head()
```

[76]:
```
    rok kod_iso_odbiorcy kraj_odbiorcy  gbd_location_id wb_regioncode  \
0  1990              AGO        Angola              168          SSA
1  1990              BDI       Burundi              175          SSA
2  1990              BEN         Benin              200          SSA
3  1990              BFA  Burkina Faso              201          SSA
4  1990              BWA      Botswana              193          SSA
```

```
   wb_location_id                    gbd_region  gbd_region_id  \
0             242   Sub-Saharan Africa, Central          167.0
1             242   Sub-Saharan Africa, Eastern          174.0
2             242   Sub-Saharan Africa, Western          199.0
3             242   Sub-Saharan Africa, Western          199.0
4             242  Sub-Saharan Africa, Southern          192.0

      gbd_superregion  gbd_superregion_id  …  other_dah_20  rmh_dah_20  \
0  Sub-Saharan Africa               166.0  …             0           5
1  Sub-Saharan Africa               166.0  …             0           6
2  Sub-Saharan Africa               166.0  …             0           6
3  Sub-Saharan Africa               166.0  …             0           5
4  Sub-Saharan Africa               166.0  …             0           1

  nch_dah_20 ncd_dah_20 hiv_dah_20 mal_dah_20 tb_dah_20 swap_hss_total_dah_20  \
0          0          0          7          3         0                      0
1          0          0          5          1         0                      0
2          0          0          5          2         0                      0
3          0          0          7          2         0                      0
4          0          0         23          -         0                      0

  oid_dah_20 unalloc_dah_20
0          0              -
1          0              0
2          0              0
3          0              0
4          0              -

[5 rows x 74 columns]
```

[77]:
```python
# zachowaj ramkę danych jako plik csv na komputerze

sciezka = r"C:\Users\piotr\Downloads\df_copy.csv"
df.to_csv(sciezka, encoding="utf-8")
```

[78]:
```python
# wyświetlić średnia (maksymalną, minimalną) wartość z jednej kolumny

col = df["elim_ch"]
mean = col.mean()
max_ = col.max()
min_ = col.min()

print(f"średnia: {mean}\nmaksimum: {max_}\nminimum: {min_}")
```

```
średnia: 0.25990237570480146
maksimum: 1
minimum: 0
```

[79]:
```
# wyświetlić liczbę wierszy

df.rok.count()
```

[79]:
```
369785
```

[80]:
```
# wyświetlić wartości unikatowe w kolumnie

df_copy["kraj_odbiorcy"].unique()
```

[80]:
```
array(['Angola', 'Burundi', 'Benin', 'Burkina Faso', 'Botswana',
       'Central African Republic', 'China', "Cote d'Ivoire", 'Cameroon',
       'Democratic Republic of the Congo', 'Congo', 'Cook Islands',
       'Comoros', 'Cape Verde', 'Djibouti', 'Eritrea', 'Ethiopia', 'Fiji',
       'Federated States of Micronesia', 'Gabon', 'Ghana', 'Guinea',
       'The Gambia', 'Guinea-Bissau', 'Equatorial Guinea', 'Indonesia',
       'Kenya', 'Cambodia', 'Kiribati', 'Laos', 'Liberia', 'Lesotho',
       'Madagascar', 'Marshall Islands', 'Mali', 'Mongolia', 'Mozambique',
       'Mauritania', 'Mauritius', 'Malawi', 'Malaysia', 'Namibia',
       'Niger', 'Nigeria', 'Niue', 'Nauru', 'Philippines', 'Palau',
       'Papua New Guinea', 'North Korea', 'Rwanda', 'Sudan', 'Senegal',
       'Solomon Islands', 'Sierra Leone', 'Somalia', 'South Sudan',
       'Sao Tome and Principe', 'Swaziland', 'Seychelles', 'Chad', 'Togo',
       'Thailand', 'Tokelau', 'Timor-Leste', 'Tonga', 'Tuvalu',
       'Tanzania', 'Uganda', 'Vietnam', 'Vanuatu', 'Global',
       'Wallis and Futuna Islands', 'Samoa', 'South Africa', 'Zambia',
       'Zimbabwe', 'Afghanistan', 'Albania', 'Armenia', 'Azerbaijan',
       'Bangladesh', 'Bosnia and Herzegovina', 'Bolivia', 'Bhutan',
       'Dominica', 'Egypt', 'Georgia', 'Grenada', 'Guyana', 'Honduras',
       'Haiti', 'India', 'Kyrgyzstan', 'Saint Lucia', 'Sri Lanka',
       'Moldova', 'Maldives', 'North Macedonia', 'Myanmar', 'Montenegro',
       'Nicaragua', 'Nepal', 'Pakistan', 'Serbia', 'Tajikistan',
       'Tunisia', 'Uzbekistan', 'Saint Vincent and the Grenadines',
       'Yemen', 'Iraq', 'Argentina', 'Brazil', 'Barbados', 'Chile',
       'Colombia', 'Costa Rica', 'Dominican Republic', 'Ecuador',
       'Jamaica', 'Panama', 'Peru', 'Paraguay', 'El Salvador', 'Suriname',
       'Trinidad and Tobago', 'Uruguay', 'Venezuela', 'Belize',
       'Guatemala', 'South Korea', 'Bulgaria', 'Belarus', 'Algeria',
       'Estonia', 'Croatia', 'Hungary', 'Iran', 'Jordan', 'Kazakhstan',
       'Saint Kitts and Nevis', 'Lebanon', 'Lithuania', 'Latvia',
       'Morocco', 'Mexico', 'Oman', 'Poland', 'Romania', 'Russia',
       'Slovakia', 'Slovenia', 'Turkmenistan', 'Turkey', 'Ukraine',
       'Antigua and Barbuda', 'Cuba', 'Libya', 'Montserrat', 'Palestine',
       'Saint Helena', 'Syria', 'Netherlands Antilles',
       'Turks and Caicos Islands', 'Northern Mariana Islands', 'Anguilla',
       'Malta', 'Bahrain', 'Saudi Arabia', 'Mayotte', 'Christmas Island'],
      dtype=object)
```

[81]:
```
# wyświetlić liczby rekordów odpowiadających wartości

df_copy[df_copy["kraj_odbiorcy"]=="Turkey"].rok.count()
```

[81]: 882

[82]:
```
# sortowanie wierszy ramki danych według wartości określonej kolumny(malejąco,␣
↪rosnąco)

df_copy.sort_values(['kraj_odbiorcy'], ascending=True).head()
```

[82]:

	rok	kod_iso_odbiorcy	kraj_odbiorcy	gbd_location_id	wb_regioncode	\
212625	2010	AFG	Afghanistan	160	SAS	
271815	2013	AFG	Afghanistan	160	SAS	
271769	2013	AFG	Afghanistan	160	SAS	
166549	2008	AFG	Afghanistan	160	SAS	
166595	2008	AFG	Afghanistan	160	SAS	

	wb_location_id	gbd_region	gbd_region_id	\
212625	244	North Africa/Middle East	138.0	
271815	244	North Africa/Middle East	138.0	
271769	244	North Africa/Middle East	138.0	
166549	244	North Africa/Middle East	138.0	
166595	244	North Africa/Middle East	138.0	

	gbd_superregion	gbd_superregion_id	…	other_dah_20	\
212625	North Africa and Middle East	137.0	…	0	
271815	North Africa and Middle East	137.0	…	16	
271769	North Africa and Middle East	137.0	…	0	
166549	North Africa and Middle East	137.0	…	0	
166595	North Africa and Middle East	137.0	…	0	

	rmh_dah_20	nch_dah_20	ncd_dah_20	hiv_dah_20	mal_dah_20	tb_dah_20	\
212625	0	139	0	0	0	0	
271815	0	0	59	11	0	0	
271769	0	12	0	0	0	0	
166549	0	0	0	0	0	0	
166595	0	0	0	3	0	0	

	swap_hss_total_dah_20	oid_dah_20	unalloc_dah_20
212625	0	0	0
271815	678	0	0
271769	0	0	0
166549	-	1	0
166595	1	0	0

[5 rows x 74 columns]

[83]:
```
df_copy.sort_values(['kraj_odbiorcy'], ascending=False).head()
```

[83]:

	rok	kod_iso_odbiorcy	kraj_odbiorcy	gbd_location_id	wb_regioncode	\
383237	2018	ZWE	Zimbabwe	198	SSA	
269436	2013	ZWE	Zimbabwe	198	SSA	
89909	2004	ZWE	Zimbabwe	198	SSA	
268851	2013	ZWE	Zimbabwe	198	SSA	
268906	2013	ZWE	Zimbabwe	198	SSA	

	wb_location_id	gbd_region	gbd_region_id	\
383237	242	Sub-Saharan Africa, Southern	192.0	
269436	242	Sub-Saharan Africa, Southern	192.0	
89909	242	Sub-Saharan Africa, Southern	192.0	
268851	242	Sub-Saharan Africa, Southern	192.0	
268906	242	Sub-Saharan Africa, Southern	192.0	

	gbd_superregion	gbd_superregion_id	…	other_dah_20	rmh_dah_20	\
383237	Sub-Saharan Africa	166.0	…	0	120	
269436	Sub-Saharan Africa	166.0	…	0	45	
89909	Sub-Saharan Africa	166.0	…	-	0	
268851	Sub-Saharan Africa	166.0	…	0	0	
268906	Sub-Saharan Africa	166.0	…	1	0	

	nch_dah_20	ncd_dah_20	hiv_dah_20	mal_dah_20	tb_dah_20	\
383237	54	0	0	14	3	
269436	60	0	0	0	0	
89909	0	0	0	0	0	
268851	-	-	0	-	0	
268906	0	0	4	-	1	

	swap_hss_total_dah_20	oid_dah_20	unalloc_dah_20
383237	0	1	0
269436	0	0	0
89909	0	0	0
268851	0	0	-
268906	0	0	0

[5 rows x 74 columns]

[84]:
```
# wyświetlić wierszy dla 10 największych (najmniejszych) wartości określonej␣
↪kolumny

df.nlargest(10, 'elim_ch')[['rok', 'elim_ch']]
```

[84]:

	rok	elim_ch
537	1990	1
2927	1990	1

```
2928  1990         1
2929  1990         1
2930  1990         1
2931  1990         1
2932  1990         1
2933  1990         1
2934  1990         1
2935  1990         1
```

[85]:
```
df.nsmallest(10, 'elim_ch')[['rok', 'elim_ch']]
```

[85]:
```
     rok  elim_ch
0   1990         0
1   1990         0
2   1990         0
3   1990         0
4   1990         0
5   1990         0
6   1990         0
7   1990         0
8   1990         0
9   1990         0
```

[86]:
```
# wyświetlić wierszy dla 10 największych wartości określonej kolumny pod␣
↪warunkiem określonych wartości innej kolumny
df[(df['kraj_odbiorcy'].isin(['Poland', 'Germany'])) & (df['rok'] == 1990)].
↪nlargest(10, 'elim_ch')
```

[86]:
```
      rok kod_iso_odbiorcy kraj_odbiorcy  gbd_location_id wb_regioncode  \
700  1990              POL        Poland               51           ECA

     wb_location_id       gbd_region  gbd_region_id  \
700             240  Europe, Central           42.0

                                    gbd_superregion  gbd_superregion_id  \
700  Central Europe, Eastern Europe, and Central Asia                31.0

     …  other_dah_20  rmh_dah_20 nch_dah_20 ncd_dah_20 hiv_dah_20  \
700  …             0           0          0          0          0

    mal_dah_20 tb_dah_20 swap_hss_total_dah_20 oid_dah_20 unalloc_dah_20
700          0         0                    0          0              0

[1 rows x 74 columns]
```

[87]:
```
# grupowanie wierszy wedlug wartości kolumny kategoryzowanej, potem-␣
↪uśrednienie wartości wszystkich kolumn w grupie - MultiIndex
```

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

[87]:

kraj_odbiorcy	rok	gbd_location_id	wb_location_id	gbd_region_id	gbd_superregion_id	elim_ch	prelim_est
Afghanistan	2008.826382	160.0	244.0	138.0	137.0	0.263207	0.012666
Albania	2007.920669	43.0	240.0	42.0	31.0	0.116028	0.008635
Algeria	2008.934132	139.0	243.0	138.0	137.0	0.308111	0.018508
Angola	2008.284524	168.0	242.0	167.0	166.0	0.293928	0.013971
Anguilla	2000.571429	44598.0	241.0	104.0	103.0	0.071429	0.000000
…	…	…	…	…	…	…	…
Vietnam	2008.299398	20.0	239.0	9.0	4.0	0.183133	0.011446
Wallis and Futuna Islands	2008.653659	423.0	239.0	21.0	4.0	0.351220	0.034146
Yemen	2008.463313	157.0	243.0	138.0	137.0	0.214070	0.012859
Zambia	2008.414866	191.0	242.0	174.0	166.0	0.296521	0.014497
Zimbabwe	2008.462245	198.0	242.0	192.0	166.0	0.316837	0.014541

[171 rows x 7 columns]

[88]:
```
# grupowanie wierszy wedlug wartości kolumny kategoryzowanej, potem uśrednienie wartości dla pewnych kolumn, liczba wartości i medianadla pozostałych kolumn w grupach

kraje = df.groupby('kraj_odbiorcy').agg({'gbd_location_id': ['mean'], 'elim_ch': ['mean'], 'prelim_est': ['mean'], 'rok': ['count', 'median'], 'wb_location_id': ['count', 'median'], 'gbd_superregion_id': ['count', 'median']})
kraje
```

[88]:

	gbd_location_id	elim_ch	prelim_est	rok		\
	mean	mean	mean	count	median	
kraj_odbiorcy						
Afghanistan	160	0.263207	0.012666	3237	2010.0	
Albania	43	0.116028	0.008635	1853	2009.0	
Algeria	139	0.308111	0.018508	1837	2009.0	
Angola	168	0.293928	0.013971	3722	2009.0	
Anguilla	44598	0.071429	0.000000	14	2000.5	
…	…	…	…	…	…	
Vietnam	20	0.183133	0.011446	3320	2009.0	
Wallis and Futuna Islands	423	0.351220	0.034146	205	2010.0	
Yemen	157	0.214070	0.012859	2644	2010.0	
Zambia	191	0.296521	0.014497	3794	2009.0	
Zimbabwe	198	0.316837	0.014541	3920	2009.0	

	wb_location_id		gbd_superregion_id	
	count	median	count	median
kraj_odbiorcy				
Afghanistan	3237	244	3237	137.0
Albania	1853	240	1853	31.0
Algeria	1837	243	1837	137.0
Angola	3722	242	3722	166.0
Anguilla	14	241	14	103.0
…	…	…	…	…
Vietnam	3320	239	3320	4.0
Wallis and Futuna Islands	205	239	205	4.0
Yemen	2644	243	2644	137.0
Zambia	3794	242	3794	166.0
Zimbabwe	3920	242	3920	166.0

[171 rows x 9 columns]

[89]:
```
# wyświetlić nazwy kolumn indeksu złożonego

kraje.index
```

[89]:
```
Index(['Afghanistan', 'Albania', 'Algeria', 'Angola', 'Anguilla',
       'Antigua and Barbuda', 'Argentina', 'Armenia', 'Azerbaijan', 'Bahrain',
       …
       'Ukraine', 'Uruguay', 'Uzbekistan', 'Vanuatu', 'Venezuela', 'Vietnam',
       'Wallis and Futuna Islands', 'Yemen', 'Zambia', 'Zimbabwe'],
      dtype='object', name='kraj_odbiorcy', length=171)
```

[90]:
```
# sortować kolumnę indeksu złożonego

kraje['elim_ch']['mean'].sort_values(ascending=False)
```

[90]:
```
kraj_odbiorcy
Global              0.496842
North Korea         0.436611
Mauritius           0.404205
Namibia             0.391414
Botswana            0.388786
                        …
South Korea         0.000000
Malta               0.000000
Hungary             0.000000
Christmas Island    0.000000
Bahrain             0.000000
Name: mean, Length: 171, dtype: float64
```

[91]:
```
# stworzyć tabelę przystawną (pivot table) na podstawie ramki danych

pivot = df.pivot_table(values='elim_ch', index='kraj_odbiorcy', columns='rok',␣
↪aggfunc='count', margins=False, dropna=True, fill_value=None)
pivot
```

[91]:
```
rok                         1990  1991  1992  1993  1994  1995  1996  1997  \
kraj_odbiorcy
Afghanistan                 22.0  24.0  25.0  25.0  23.0  23.0  28.0  34.0
Albania                     19.0  21.0  24.0  21.0  20.0  22.0  27.0  29.0
Algeria                      5.0   5.0  13.0  11.0  12.0  13.0  16.0  30.0
Angola                      31.0  40.0  48.0  43.0  43.0  33.0  41.0  60.0
Anguilla                     NaN   NaN   1.0   1.0   1.0   1.0   NaN   NaN
…                            …     …     …     …     …     …     …     …
Vietnam                     24.0  26.0  29.0  29.0  33.0  45.0  32.0  53.0
Wallis and Futuna Islands    2.0   1.0   1.0   3.0   2.0   3.0   1.0   2.0
Yemen                       24.0  25.0  28.0  27.0  25.0  23.0  24.0  46.0
Zambia                      41.0  40.0  36.0  45.0  33.0  33.0  39.0  43.0
Zimbabwe                    32.0  33.0  37.0  46.0  47.0  36.0  41.0  61.0

rok                         1998  1999  …   2009   2010   2011   2012   2013  \
kraj_odbiorcy                           …
```

```
Afghanistan                52.0  40.0  …  164.0  200.0  208.0  183.0  209.0
Albania                    38.0  37.0  …  114.0  105.0  101.0  100.0   99.0
Algeria                    17.0  38.0  …  106.0   87.0  109.0  107.0  114.0
Angola                     63.0  65.0  …  237.0  174.0  236.0  220.0  213.0
Anguilla                    1.0   1.0  …    1.0    NaN    NaN    NaN    1.0
…                           …     …    …    …      …      …      …
Vietnam                    57.0  59.0  …  166.0  176.0  207.0  187.0  191.0
Wallis and Futuna Islands   2.0   3.0  …   15.0   10.0   16.0   15.0   15.0
Yemen                      53.0  53.0  …  119.0  135.0  146.0  148.0  155.0
Zambia                     61.0  64.0  …  236.0  177.0  237.0  222.0  227.0
Zimbabwe                   65.0  66.0  …  255.0  194.0  240.0  224.0  229.0

rok                         2014   2015   2016   2017   2018
kraj_odbiorcy
Afghanistan                169.0  202.0  151.0  163.0  171.0
Albania                    102.0   92.0   57.0   96.0   94.0
Algeria                    116.0   99.0   96.0  101.0   86.0
Angola                     227.0  204.0  211.0  171.0  150.0
Anguilla                     NaN    NaN    NaN    NaN    NaN
…                            …      …      …      …      …
Vietnam                    178.0  186.0  172.0  153.0  159.0
Wallis and Futuna Islands   10.0    9.0    7.0   10.0   12.0
Yemen                      150.0  168.0  138.0  139.0  156.0
Zambia                     226.0  205.0  214.0  187.0  168.0
Zimbabwe                   227.0  218.0  227.0  191.0  182.0

[171 rows x 29 columns]
```

[92]:
```
# wyświetlić indeksy i kolumny tabeli przystawnej

pivot.index
```

[92]:
```
Index(['Afghanistan', 'Albania', 'Algeria', 'Angola', 'Anguilla',
       'Antigua and Barbuda', 'Argentina', 'Armenia', 'Azerbaijan', 'Bahrain',
       …
       'Ukraine', 'Uruguay', 'Uzbekistan', 'Vanuatu', 'Venezuela', 'Vietnam',
       'Wallis and Futuna Islands', 'Yemen', 'Zambia', 'Zimbabwe'],
      dtype='object', name='kraj_odbiorcy', length=171)
```

[93]:
```
# utwórz indeks złożony tabeli przystawnej i wyświetl go

pivot2 = df.pivot_table(values='elim_ch', index=['gbd_region',␣
 ↪'kraj_odbiorcy'], columns='rok', aggfunc='count', margins=False,␣
 ↪dropna=True, fill_value=None)
pivot2
```

[93]:
rok		1990	1991	1992	1993
gbd_region	kraj_odbiorcy				
Asia Pacific, high-income	South Korea	5.0	3.0	3.0	4.0
Asia, Central	Armenia	22.0	23.0	24.0	26.0
	Azerbaijan	21.0	22.0	23.0	25.0
	Georgia	22.0	23.0	24.0	26.0
	Kazakhstan	4.0	4.0	5.0	8.0
…		…	…	…	…
Sub-Saharan Africa, Western	Sao Tome and Principe	41.0	28.0	35.0	31.0
	Senegal	31.0	29.0	35.0	43.0
	Sierra Leone	30.0	28.0	36.0	33.0
	The Gambia	30.0	28.0	36.0	32.0
	Togo	30.0	28.0	46.0	30.0

rok		1994	1995	1996	1997
gbd_region	kraj_odbiorcy				
Asia Pacific, high-income	South Korea	4.0	NaN	NaN	NaN
Asia, Central	Armenia	24.0	23.0	28.0	32.0
	Azerbaijan	23.0	20.0	26.0	30.0
	Georgia	25.0	22.0	28.0	32.0
	Kazakhstan	9.0	7.0	9.0	11.0
…		…	…	…	…
Sub-Saharan Africa, Western	Sao Tome and Principe	32.0	45.0	38.0	42.0
	Senegal	31.0	45.0	40.0	59.0
	Sierra Leone	32.0	31.0	52.0	42.0
	The Gambia	32.0	31.0	37.0	42.0
	Togo	32.0	45.0	38.0	42.0

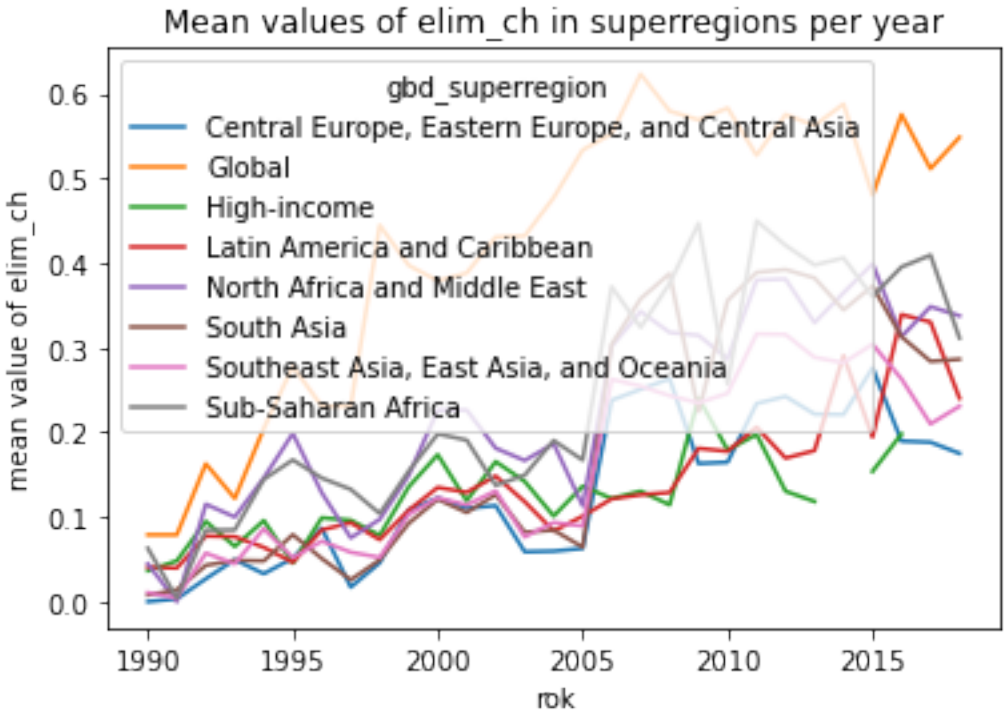
rok		1998	1999	…	2009
gbd_region	kraj_odbiorcy			…	
Asia Pacific, high-income	South Korea	5.0	5.0	…	NaN
Asia, Central	Armenia	53.0	39.0	…	127.0
	Azerbaijan	50.0	38.0	…	122.0
	Georgia	53.0	40.0	…	127.0
	Kazakhstan	30.0	19.0	…	104.0
…		…	…	…	…
Sub-Saharan Africa, Western	Sao Tome and Principe	45.0	61.0	…	232.0
	Senegal	63.0	63.0	…	237.0
	Sierra Leone	60.0	61.0	…	249.0
	The Gambia	44.0	63.0	…	231.0
	Togo	45.0	61.0	…	260.0

rok		2010	2011	2012	2013
gbd_region	kraj_odbiorcy				
Asia Pacific, high-income	South Korea	NaN	NaN	NaN	NaN
Asia, Central	Armenia	133.0	143.0	147.0	154.0
	Azerbaijan	129.0	140.0	140.0	148.0

```
                                Georgia                130.0  155.0  142.0  151.0
                                Kazakhstan             110.0  103.0  121.0  112.0
…                                                      …      …      …      …
Sub-Saharan Africa, Western     Sao Tome and Principe  169.0  234.0  215.0  209.0
                                Senegal                176.0  239.0  246.0  244.0
                                Sierra Leone           172.0  255.0  239.0  249.0
                                The Gambia             197.0  253.0  236.0  234.0
                                Togo                   199.0  254.0  239.0  212.0

rok                                                     2014   2015   2016   2017  \
gbd_region                      kraj_odbiorcy
Asia Pacific, high-income       South Korea              NaN    NaN    NaN    NaN
Asia, Central                   Armenia                139.0  144.0  137.0  132.0
                                Azerbaijan             133.0  138.0  129.0  112.0
                                Georgia                140.0  144.0  135.0  130.0
                                Kazakhstan              99.0  101.0   90.0   69.0
…                                                      …      …      …      …
Sub-Saharan Africa, Western     Sao Tome and Principe  224.0  215.0  222.0  169.0
                                Senegal                251.0  228.0  239.0  202.0
                                Sierra Leone           250.0  239.0  239.0  210.0
                                The Gambia             247.0  224.0  231.0  189.0
                                Togo                   247.0  224.0  232.0  172.0

rok                                                     2018
gbd_region                      kraj_odbiorcy
Asia Pacific, high-income       South Korea              NaN
Asia, Central                   Armenia                123.0
                                Azerbaijan             119.0
                                Georgia                121.0
                                Kazakhstan              75.0
…                                                      …
Sub-Saharan Africa, Western     Sao Tome and Principe  147.0
                                Senegal                179.0
                                Sierra Leone           178.0
                                The Gambia             170.0
                                Togo                   148.0

[171 rows x 29 columns]
```

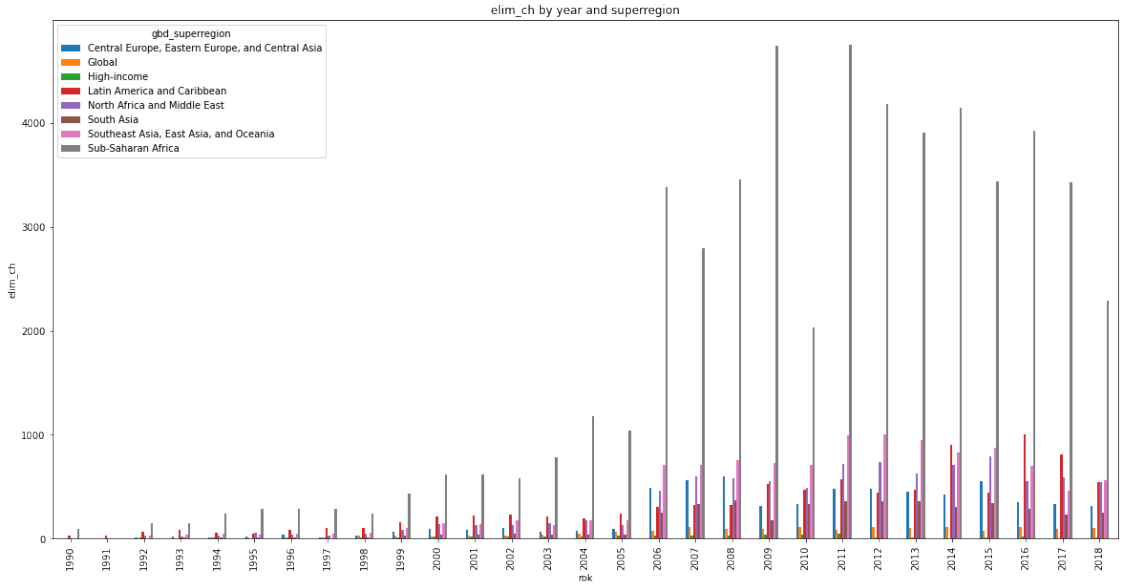
[94]:
```
# zaimportuj moduł pyplot z biblioteki matplotlib
import matplotlib.pyplot as plt

# wskazać, że wykresy należy rysować bezpośrednio w zeszycie, a nie wosobnej␣
 ↪zakładce
%matplotlib inline
```

[95]:
```
# wyświetlić wykres na podstawie tabeli przystawnej

pivot3 = df.pivot_table(values='elim_ch', index='rok',␣
↪columns='gbd_superregion', aggfunc='mean', margins=False, dropna=True,␣
↪fill_value=None)

fig = pivot3.plot(kind='line')
plt.ylabel('mean value of elim_ch')
plt.title('Mean values of elim_ch in superregions per year')
plt.rcParams["figure.figsize"] = (20,10)
#display(fig)

#plt.show()
```



[96]:
```
# narysować histogram na podstawie wartości kolumny

df_bar = df.pivot_table(values='elim_ch', index='rok',␣
↪columns='gbd_superregion', aggfunc='sum', margins=False, fill_value=None,␣
↪dropna=True)
df_bar.plot(kind='bar')
plt.ylabel('elim_ch')
plt.title('elim_ch by year and superregion')
```

[96]: Text(0.5, 1.0, 'elim_ch by year and superregion')



[97]:
```
df.columns
```

[97]:
```
Index(['rok', 'kod_iso_odbiorcy', 'kraj_odbiorcy', 'gbd_location_id',
       'wb_regioncode', 'wb_location_id', 'gbd_region', 'gbd_region_id',
       'gbd_superregion', 'gbd_superregion_id', 'elim_ch', 'prelim_est',
       'dah_20', 'rmh_fp_dah_20', 'rmh_mh_dah_20', 'rmh_hss_other_dah_20',
       'rmh_hss_hrh_dah_20', 'rmh_other_dah_20', 'nch_cnn_dah_20',
       'nch_cnv_dah_20', 'nch_other_dah_20', 'nch_hss_other_dah_20',
       'nch_hss_hrh_dah_20', 'hiv_treat_dah_20', 'hiv_prev_dah_20',
       'hiv_pmtct_dah_20', 'hiv_other_dah_20', 'hiv_ct_dah_20',
       'hiv_ovc_dah_20', 'hiv_care_dah_20', 'hiv_hss_other_dah_20',
       'hiv_hss_hrh_dah_20', 'hiv_amr_dah_20', 'mal_diag_dah_20',
       'mal_hss_other_dah_20', 'mal_hss_hrh_dah_20', 'mal_con_nets_dah_20',
       'mal_con_irs_dah_20', 'mal_con_oth_dah_20', 'mal_treat_dah_20',
       'mal_comm_con_dah_20', 'mal_other_dah_20', 'mal_amr_dah_20',
       'tb_other_dah_20', 'tb_treat_dah_20', 'tb_diag_dah_20',
       'tb_hss_other_dah_20', 'tb_hss_hrh_dah_20', 'tb_amr_dah_20',
       'oid_hss_other_dah_20', 'oid_hss_hrh_dah_20', 'oid_ebz_dah_20',
       'oid_zika_dah_20', 'oid_covid_dah_20', 'oid_other_dah_20',
       'oid_amr_dah_20', 'ncd_hss_other_dah_20', 'ncd_hss_hrh_dah_20',
       'ncd_tobac_dah_20', 'ncd_mental_dah_20', 'ncd_other_dah_20',
       'swap_hss_other_dah_20', 'swap_hss_hrh_dah_20', 'swap_hss_pp_dah_20',
       'other_dah_20', 'rmh_dah_20', 'nch_dah_20', 'ncd_dah_20', 'hiv_dah_20',
       'mal_dah_20', 'tb_dah_20', 'swap_hss_total_dah_20', 'oid_dah_20',
       'unalloc_dah_20'],
```

```
      dtype='object')
```

[98]:
```
# przedstawić sposobyłączenia ramek danych za pomocą metod merge iconcat

part1 = df[['rok', 'kraj_odbiorcy', 'elim_ch', 'gbd_region']]
part2 = df[['rok', 'kraj_odbiorcy', 'prelim_est', 'gbd_superregion']]

pd.merge(part1, part2, on = ['rok', 'kraj_odbiorcy'], how='inner').head()
```

[98]:
```
    rok kraj_odbiorcy  elim_ch                  gbd_region  prelim_est  \
0  1990        Angola        0  Sub-Saharan Africa, Central           0
1  1990        Angola        0  Sub-Saharan Africa, Central           0
2  1990        Angola        0  Sub-Saharan Africa, Central           0
3  1990        Angola        0  Sub-Saharan Africa, Central           0
4  1990        Angola        0  Sub-Saharan Africa, Central           0

      gbd_superregion
0  Sub-Saharan Africa
1  Sub-Saharan Africa
2  Sub-Saharan Africa
3  Sub-Saharan Africa
4  Sub-Saharan Africa
```

[99]:
```
pd.merge(part1, part2, on = [part1.index, part2.index], how='inner').head()
```

[99]:
```
   key_0  key_1  rok_x kraj_odbiorcy_x  elim_ch                  gbd_region  \
0      0      0   1990          Angola        0  Sub-Saharan Africa, Central
1      1      1   1990         Burundi        0  Sub-Saharan Africa, Eastern
2      2      2   1990           Benin        0  Sub-Saharan Africa, Western
3      3      3   1990    Burkina Faso        0  Sub-Saharan Africa, Western
4      4      4   1990        Botswana        0  Sub-Saharan Africa, Southern

   rok_y kraj_odbiorcy_y  prelim_est     gbd_superregion
0   1990          Angola           0  Sub-Saharan Africa
1   1990         Burundi           0  Sub-Saharan Africa
2   1990           Benin           0  Sub-Saharan Africa
3   1990    Burkina Faso           0  Sub-Saharan Africa
4   1990        Botswana           0  Sub-Saharan Africa
```

[103]:
```
pd.concat([part1, part2], axis=0)
```

[103]:
```
         rok kraj_odbiorcy  elim_ch                  gbd_region  prelim_est  \
0       1990        Angola      0.0  Sub-Saharan Africa, Central         NaN
1       1990       Burundi      0.0  Sub-Saharan Africa, Eastern         NaN
2       1990         Benin      0.0  Sub-Saharan Africa, Western         NaN
3       1990  Burkina Faso      0.0  Sub-Saharan Africa, Western         NaN
4       1990      Botswana      0.0  Sub-Saharan Africa, Southern         NaN
```

```
…       …            …       …                                  …          …
383233  2018         Global      NaN                               NaN        0.0
383234  2018          Yemen      NaN                               NaN        0.0
383235  2018   South Africa      NaN                               NaN        0.0
383236  2018         Zambia      NaN                               NaN        0.0
383237  2018       Zimbabwe      NaN                               NaN        0.0

                    gbd_superregion
0                               NaN
1                               NaN
2                               NaN
3                               NaN
4                               NaN
…                               …
383233                       Global
383234  North Africa and Middle East
383235            Sub-Saharan Africa
383236            Sub-Saharan Africa
383237            Sub-Saharan Africa

[739570 rows x 6 columns]
```

[106]:
```
pd.concat([part1, part2], axis=0).shape
```

[106]:
```
(739570, 6)
```

[104]:
```
pd.concat([part1, part2], axis=1)
```

[104]:
```
          rok kraj_odbiorcy  elim_ch                    gbd_region   rok  \
0       1990         Angola        0   Sub-Saharan Africa, Central  1990
1       1990        Burundi        0   Sub-Saharan Africa, Eastern  1990
2       1990          Benin        0   Sub-Saharan Africa, Western  1990
3       1990   Burkina Faso        0   Sub-Saharan Africa, Western  1990
4       1990       Botswana        0  Sub-Saharan Africa, Southern  1990
…       …            …       …                          …    …
383233  2018         Global        1                       Global  2018
383234  2018          Yemen        1      North Africa/Middle East  2018
383235  2018   South Africa        1  Sub-Saharan Africa, Southern  2018
383236  2018         Zambia        1   Sub-Saharan Africa, Eastern  2018
383237  2018       Zimbabwe        1  Sub-Saharan Africa, Southern  2018

       kraj_odbiorcy  prelim_est              gbd_superregion
0             Angola           0            Sub-Saharan Africa
1            Burundi           0            Sub-Saharan Africa
2              Benin           0            Sub-Saharan Africa
3       Burkina Faso           0            Sub-Saharan Africa
4           Botswana           0            Sub-Saharan Africa
```

```
…              …         …                              …
383233        Global           0                         Global
383234         Yemen           0  North Africa and Middle East
383235  South Africa           0            Sub-Saharan Africa
383236        Zambia           0            Sub-Saharan Africa
383237      Zimbabwe           0            Sub-Saharan Africa

[369785 rows x 8 columns]
```

[105]:
```
pd.concat([part1, part2], axis=1).shape
```

[105]:
```
(369785, 8)
```

[100]:
```
# pokazać dodawanie nowych kolumn za pomocą operacji matematycznych
df['sum'] = df['elim_ch'] + df['prelim_est']

df[['elim_ch', 'prelim_est', 'sum']].tail()
```

[100]:
```
        elim_ch  prelim_est  sum
383233        1           0    1
383234        1           0    1
383235        1           0    1
383236        1           0    1
383237        1           0    1
```

[101]:
```
# przedstawić na przykładzie dodawanie nowych kolumn z pomocą funkcjilambda

df['years_ago'] = df['rok'].apply(lambda y: 2021 - int(y))
df[['rok', 'years_ago']]
```

[101]:
```
         rok  years_ago
0       1990         31
1       1990         31
2       1990         31
3       1990         31
4       1990         31
…       …          …
383233  2018          3
383234  2018          3
383235  2018          3
383236  2018          3
383237  2018          3

[369785 rows x 2 columns]
```

[102]:
```
# przedstawić możliwości pracy z dużymi plikami przy użyciu argumentu chunksize
```

```
sciezka = r"C:
 ↪\Users\piotr\Downloads\IHME_DAH_DATABASE_1990_2020_CSV_1\IHME_DAH_DATABASE_1990_2020_Y2021M
 ↪CSV"

chunks = pd.read_csv(sciezka, low_memory=False, chunksize=100_000)

for i, chunk in enumerate(chunks):
    print(f"\n\nChunk number {i+1}:")
    print(chunk.iloc[0:3,0:4])
```

```
Chunk number 1:
   year     source  channel recipient_isocode
0  1990  Australia  BIL_AUS               AGO
1  1990  Australia  BIL_AUS               BDI
2  1990  Australia  BIL_AUS               BEN


Chunk number 2:
        year         source channel recipient_isocode
100000  2004  Private_other     NGO               LBY
100001  2004  Private_other     NGO               LCA
100002  2004  Private_other     NGO               LKA


Chunk number 3:
        year         source  channel recipient_isocode
200000  2009  United_States  INTLNGO               VCT
200001  2009  United_States  INTLNGO               VEN
200002  2009  United_States  INTLNGO               VEN


Chunk number 4:
        year  source channel recipient_isocode
300000  2014  Norway     WHO               SSD
300001  2014  Norway     WHO               STP
300002  2014  Norway     WHO               SWZ
```

[ ]: