Анализ данных уровня самоубийств в период с 1985 по 2016 год

ссылка на датасет

Задачи исследования

- ➔ Узнать количество самоубийств за каждый год во всем мире с 1985 по 2016 год
- ➔ Найти общее количество самоубийств в каждой стране с 1985 по 2016 год
- ➔ Найти 20 стран с минимальным количеством самоубийств за период с 1985 по 2016 год
- ➔ Узнать, какие возрастные группы и пол имеют самый высокий уровень самоубийств за период с 1985 по 2016 год
- ➔ Узнать, в каком году было больше всего самоубийств в каждой стране в период с 1985 по 2016 год
- ➔ Узнать количество самоубийств в России за каждый год в период с 1985 по 2016 год
- ➔ Посмотреть, как с 1989 по 2015 год ВВП влияет на количество самоубийств в России
- ➔ Посмотреть соотношение самоубийств по возрасту и полу в России с 1989 по 2015 год

Инструменты анализа

Python 3.11
Jupiter notebook
SQLlite
Библиотеки:
sqlite3
pandas
matplotlib
seaborn

Описание данных

country	страна
year	год
sex	пол
age	возрастная группа
suicides_no	количество самоубийств
population	численность населения

suicides/100k pop	количество самоубийств на 100 тысяч
country-year	страна-год
HDI for year	индекс человеческого развития за год
gdp_for_year ($)	ввп за год
gdp_per_capita ($)	ввп на душу населения
generation'	поколение

Первичная обработка данных

```
# Подгружаем необходимые библиотеки
import pandas as pd
import sqlite3
```
✓ 0.0s

```
# Загружаем датасет в переменную
df = pd.read_csv('dataset_suicides.csv')
```
✓ 0.3s

```
# Посмотрим данные первых 5-ти строк
df.head()
```
✓ 0.2s

	country	year	sex	age	suicides_no	population	suicides/100k pop	country-year	HDI for year	gdp_for_year ($)	gdp_per_capita ($)	generation
0	Albania	1987	male	15-24 years	21	312900	6.71	Albania1987	NaN	2,156,624,900	796	Generation X
1	Albania	1987	male	35-54 years	16	308000	5.19	Albania1987	NaN	2,156,624,900	796	Silent
2	Albania	1987	female	15-24 years	14	289700	4.83	Albania1987	NaN	2,156,624,900	796	Generation X
3	Albania	1987	male	75+ years	1	21800	4.59	Albania1987	NaN	2,156,624,900	796	G.I. Generation
4	Albania	1987	male	25-34 years	9	274300	3.28	Albania1987	NaN	2,156,624,900	796	Boomers

Первичная обработка данных

```python
# Посмотрим общую информацию о датасете
# Функция info показывает типы данных
# По количеству строк видим, что в  столбце HDI for year отсутствуют некоторые значения
dfName = [x for x in globals() if globals()[x] is df][0]

rows_num,columns_num = df.shape
print(f'Количество записей: {rows_num}')
print(f'Количество столбцов: {columns_num}\n')

print('Общая информация о датасете:\n')
print(df.info())
```

```
Количество записей: 27820
Количество столбцов: 12

Общая информация о датасете:

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 27820 entries, 0 to 27819
Data columns (total 12 columns):
 #   Column              Non-Null Count  Dtype  
---  ------              --------------  -----  
 0   country             27820 non-null  object 
 1   year                27820 non-null  int64  
 2   sex                 27820 non-null  object 
 3   age                 27820 non-null  object 
 4   suicides_no         27820 non-null  int64  
 5   population          27820 non-null  int64  
 6   suicides/100k pop   27820 non-null  float64
 7   country-year        27820 non-null  object 
 8   HDI for year        8364 non-null   float64
 9    gdp_for_year ($)   27820 non-null  object 
 10  gdp_per_capita ($)  27820 non-null  int64  
 11  generation          27820 non-null  object 
dtypes: float64(2), int64(4), object(6)
memory usage: 2.5+ MB
None
```

```python
# Проверим название колонок
list(df.columns)
```

✓ 0.0s

```
['country',
 'year',
 'sex',
 'age',
 'suicides_no',
 'population',
 'suicides/100k pop',
 'country-year',
 'HDI for year',
 ' gdp_for_year ($) ',
 'gdp_per_capita ($)',
 'generation']
```

Первичная обработка данных

```
# Переименуем колонки suicides_no, suicides/100k pop, gdp_for_year ($), gdp_per_capita ($)
df = df.rename(columns={'suicides_no': 'count_of_suicides', 'suicides/100k pop': 'quantity_suidides_per_100k',
                        ' gdp_for_year ($) ': 'gdp_for_year_dollars', 'gdp_per_capita ($)': 'gdp_per_capita_dollars'})
```

✓ 0.0s

```
# Удалим колонку country-year
df = df.drop(columns=['country-year'])
```

✓ 0.0s

+ Code + Markdown

```
# Приведем название колонок к нижнему регистру и заменим пробел на '_'
df.columns = [x.lower().replace(' ', '_') for x in df.columns.values]
df.columns
```

✓ 0.0s

```
Index(['country', 'year', 'sex', 'age', 'count_of_suicides', 'population',
       'quantity_suidides_per_100k', 'hdi_for_year', 'gdp_for_year_dollars',
       'gdp_per_capita_dollars', 'generation'],
      dtype='object')
```

Первичная обработка данных

```
# Посмотрим, сколько пустых значений у нас в колонке hdi_for_year
df.isnull().sum()
```

✓ 0.0s

```
country                          0
year                             0
sex                              0
age                              0
count_of_suicides                0
population                       0
quantity_suidides_per_100k       0
hdi_for_year                 19456
gdp_for_year_dollars             0
gdp_per_capita_dollars           0
generation                       0
dtype: int64
```

```
# Удалим колонку hdi_for_year
df = df.drop(columns=['hdi_for_year'])
```

✓ 0.0s

Создаем БД SQLite и загружаем в нее обработанный датасет

```
# Создаем базу данных
con = sqlite3.connect('D:/db_suicides', timeout=10)
cur = con.cursor()
```

✓ 0.0s

```
# Загружаем таблицу
df.to_sql(con=con, name='db_suicides', if_exists = 'replace', index=False)
```

✓ 1.0s

```
27820
```

```
# подключаем библиотеки seaborn и matplotlib для визуализации данных
import seaborn as sns
import matplotlib.pyplot as plt
```

✓ 0.0s

Анализ датасета в SQL и графических библиотеках Python

1. **Узнать количество самоубийств за каждый год во всем мире с 1985 по 2016 год**

```
SELECT
    year,
    SUM(count_of_suicides) AS total_suicides
FROM db_suicides
GROUP BY year
ORDER BY total_suicides DESC;
```

year	total_suicides
1 999	256 119
2 002	256 095
2 003	256 079
2 000	255 832
2 001	250 652
1 998	249 591
1 996	246 725
1 995	243 544
2 009	243 487
2 004	240 861
1 997	240 745
2 010	238 702
2 011	236 484
2 008	235 447
2 005	234 375
2 007	233 408
2 006	233 361
1 994	232 063
2 012	230 160
2 013	223 199
2 014	222 984
1 993	221 565
1 992	211 473
2 015	203 640

Как мы видим, больше всего самоубийств в мире было в:

1. 1999 году: 256119 человек,
2. 2002 году: 256095 человек,
3. 2003 году: 256079 человек,
4. 2000 году: 255832 человек.

Запускаем SQL запрос из Python и строим визуализацию

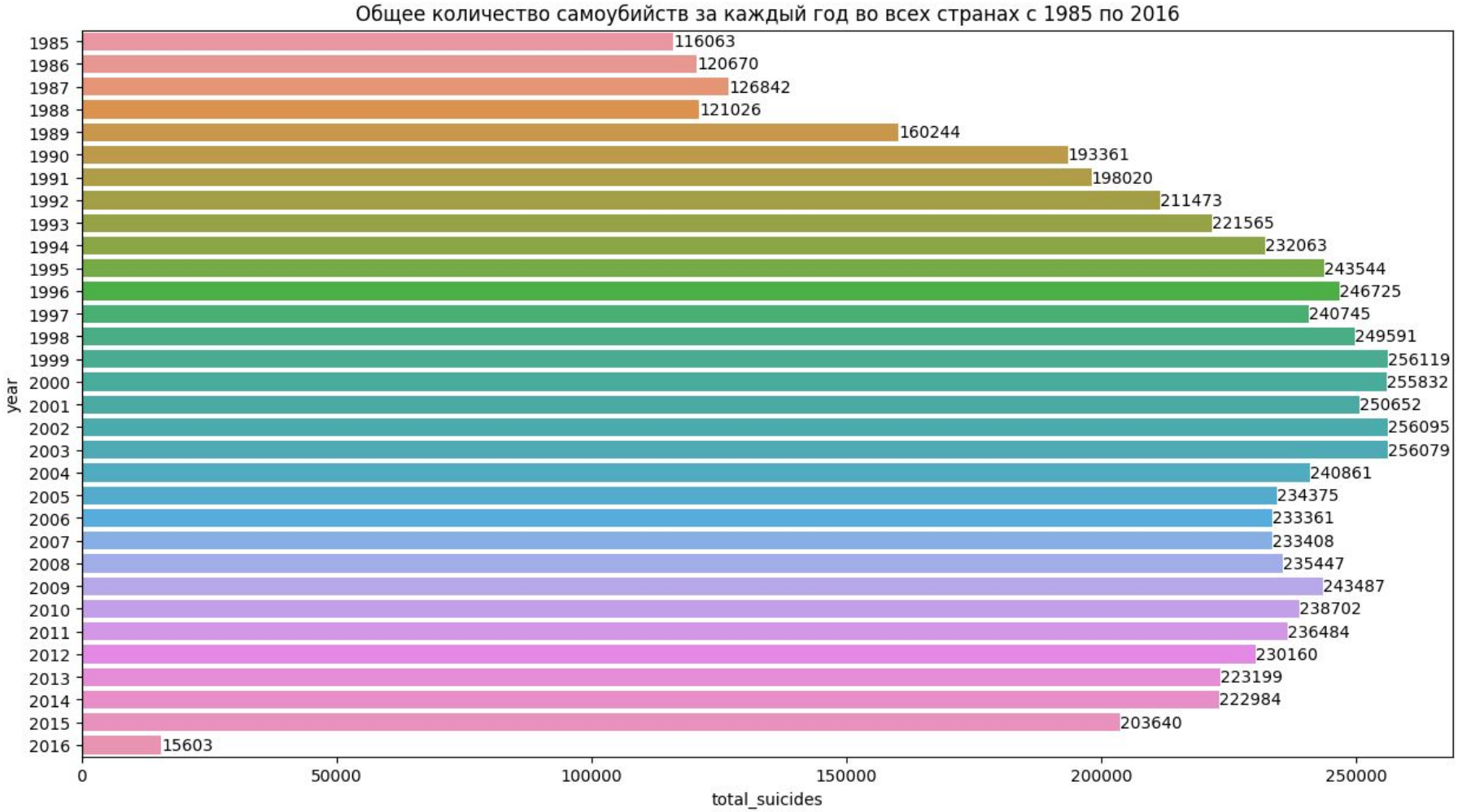
```
plt.figure(figsize=(15, 8))
plt.title('Общее количество самоубийств за каждый год во всех странах с 1985 по 2016')
sql_query = """
SELECT
    year,
    SUM(count_of_suicides) AS total_suicides
FROM db_suicides
GROUP BY year
ORDER BY total_suicides DESC;
"""
df_from_sql = pd.read_sql_query(sql_query, con=con)

my_plot = sns.barplot(data=df_from_sql, y="year", x="total_suicides", orient="h")

my_plot.set_xticklabels(my_plot.get_xticklabels())

for i in my_plot.containers:
    my_plot.bar_label(i)
```

✓ 0.9s


Общее количество самоубийств за каждый год во всех странах с 1985 по 2016
year
1985 116063
1986 120670
1987 126842
1988 121026
1989 160244
1990 193361
1991 198020
1992 211473
1993 221565
1994 232063
1995 243544
1996 246725
1997 240745
1998 249591
1999 256119
2000 255832
2001 250652
2002 256095
2003 256079
2004 240861
2005 234375
2006 233361
2007 233408
2008 235447
2009 243487
2010 238702
2011 236484
2012 230160
2013 223199
2014 222984
2015 203640
2016 15603
0
50000
100000
150000
200000
250000
total_suicides

Анализ датасета в SQL и графических библиотеках Python

2. Найти общее количество самоубийств в каждой стране с 1985 по 2016 год и отфильтровать первые 20 по убыванию

```
SELECT
    country,
    SUM(count_of_suicides) AS total_suicides
FROM db_suicides
GROUP BY country
ORDER BY total_suicides DESC
LIMIT 20;
```

country	total_suicides
Russian Federation	1 209 742
United States	1 034 013
Japan	806 902
France	329 127
Ukraine	319 950
Germany	291 262
Republic of Korea	261 730
Brazil	226 613
Poland	139 098
United Kingdom	136 805
Italy	132 060
Mexico	111 139
Thailand	110 643
Canada	107 561
Kazakhstan	101 546
Spain	100 202
Argentina	82 219
Hungary	73 891
Romania	72 777
Australia	70 111

По числу суицидов с 1985 года по 2016 год лидирует Россия: 1209742 человека

На втором месте США: 1034013 человек

Тройку замыкает Япония: 806902 человека

Запускаем SQL запрос из Python и строим визуализацию

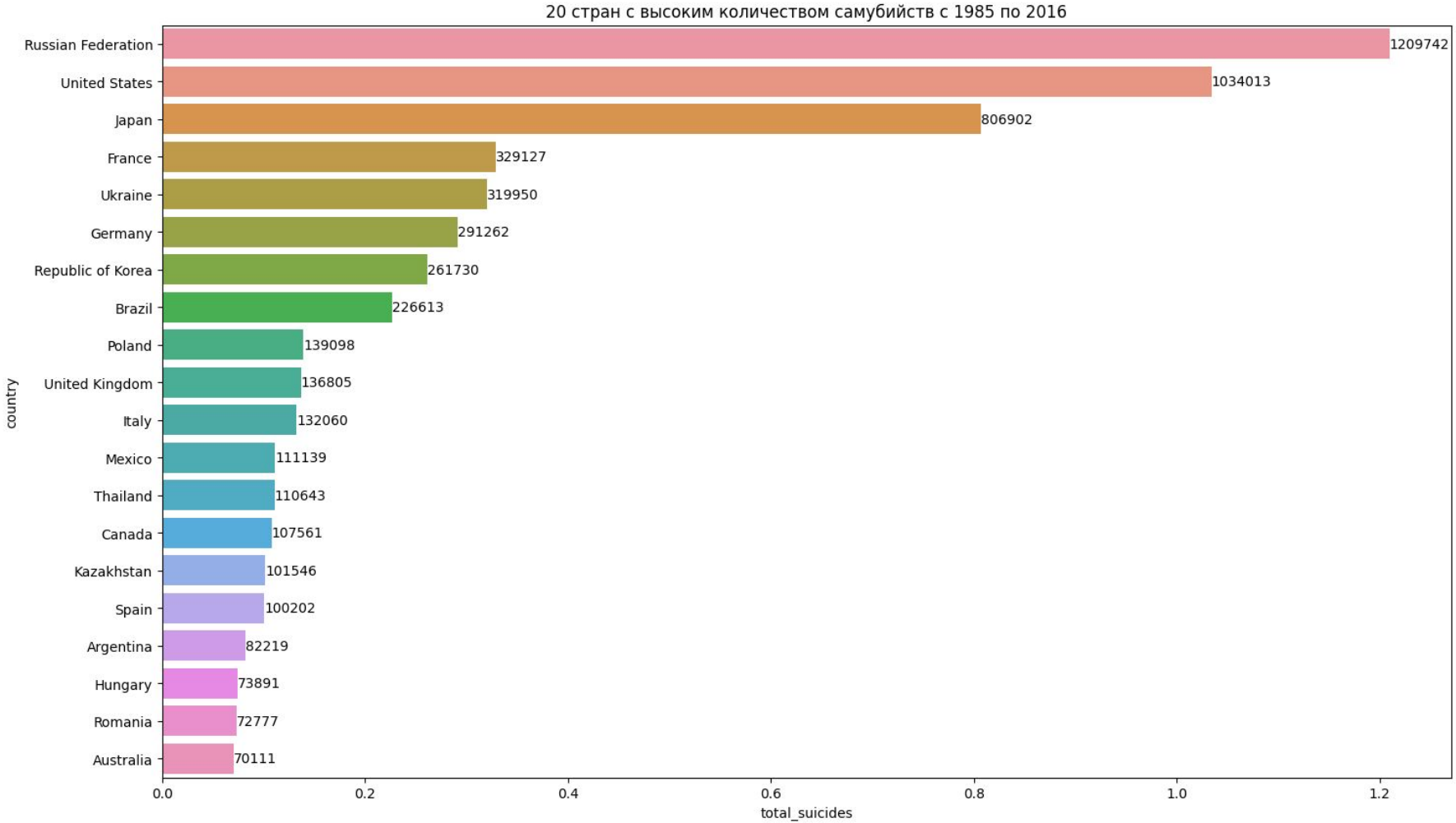
```
plt.figure(figsize=(17, 10))
plt.title('20 стран с высоким количеством самубийств с 1985 по 2016')
sql_query = """
SELECT
    country,
    SUM(count_of_suicides) AS total_suicides
FROM db_suicides
GROUP BY country
ORDER BY total_suicides DESC
LIMIT 20;
"""
df_from_sql = pd.read_sql_query(sql_query, con=con)

my_plot = sns.barplot(data=df_from_sql, y="country", x="total_suicides", orient="h")

my_plot.set_xticklabels(my_plot.get_xticklabels())

for i in my_plot.containers:
    my_plot.bar_label(i, fmt="%d")
```

✓ 0.7s


20 стран с высоким количеством самубийств с 1985 по 2016
Russian Federation
United States
Japan
France
Ukraine
Germany
Republic of Korea
Brazil
Poland
United Kingdom
Italy
Mexico
Thailand
Canada
Kazakhstan
Spain
Argentina
Hungary
Romania
Australia
1209742
1034013
806902
329127
319950
291262
261730
226613
139098
136805
132060
111139
110643
107561
101546
100202
82219
73891
72777
70111
country
0.0
0.2
0.4
0.6
0.8
1.0
1.2
total_suicides

Анализ датасета в SQL и графических библиотеках Python

3. Найти 20 стран с минимальным количеством самоубийств за период с 1985 по 2016 год

```
SELECT
    country,
    SUM(count_of_suicides) AS total_suicides
FROM db_suicides
GROUP BY country
ORDER BY total_suicides ASC
LIMIT 20;
```

country	total_suicides
Dominica	0
Saint Kitts and Nevis	0
San Marino	4
Antigua and Barbud	11
Maldives	20
Macau	27
Oman	33
Grenada	38
Cabo Verde	42
Kiribati	53
Bahamas	93
Seychelles	98
Aruba	101
Saint Vincent and Gı	124
Barbados	177
Jamaica	184
Saint Lucia	230
Fiji	304
Bosnia and Herzego	318
Belize	348

Как видим, с 1985 по 2016 год в таких странах, как Доминика и Сент-Китс и Невис, вообще не зарегистрировано ни одного случая суицида.

Запускаем SQL запрос из Python и строим визуализацию

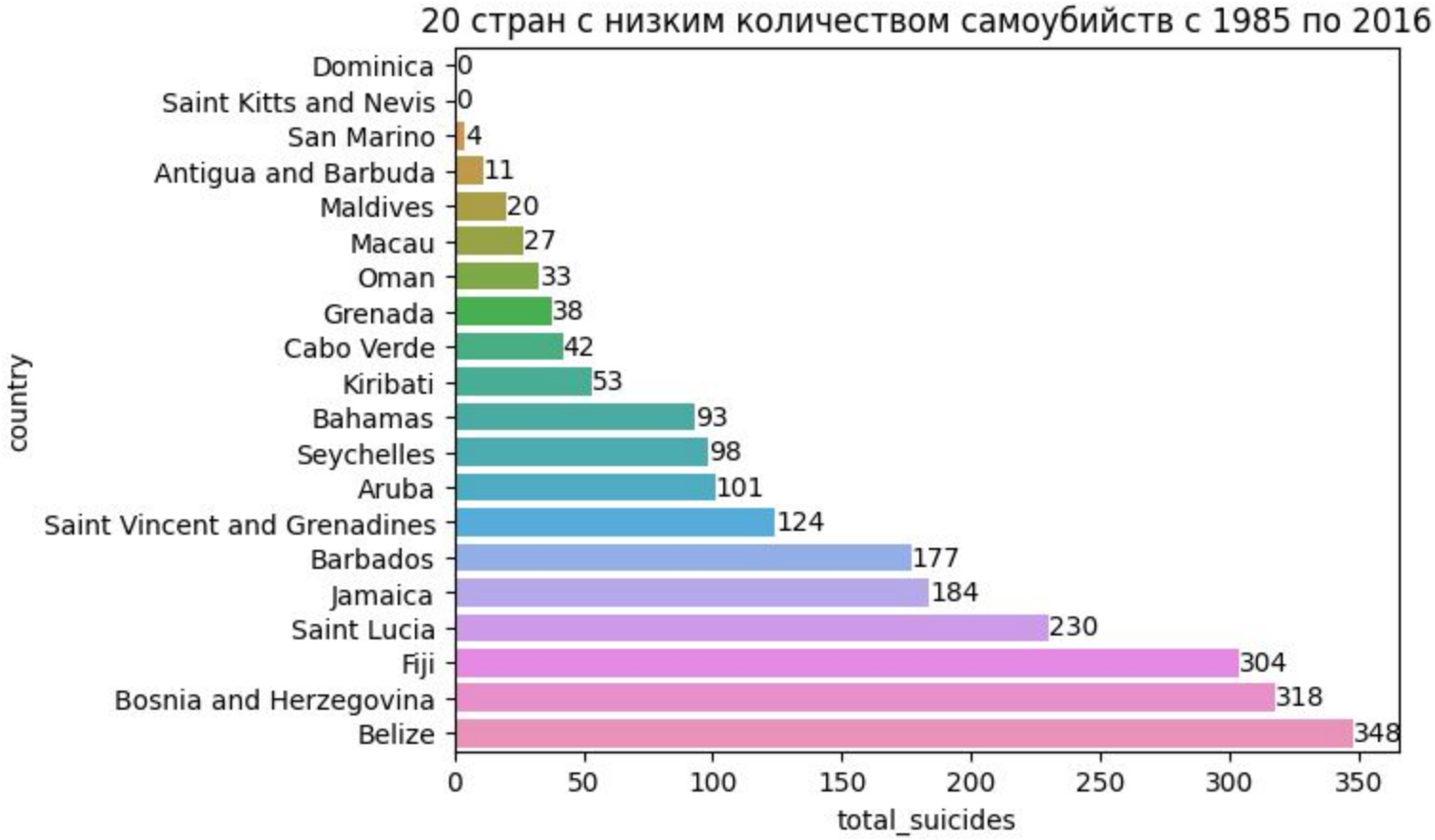
```
sql_query = """
SELECT
    country,
    SUM(count_of_suicides) AS total_suicides
FROM db_suicides
GROUP BY country
ORDER BY total_suicides ASC
LIMIT 20;
"""
df_from_sql = pd.read_sql_query(sql_query, con=con)

my_plot = sns.barplot(data=df_from_sql, y="country", x="total_suicides", orient="h")

my_plot.set_xticklabels(my_plot.get_xticklabels())
plt.title('20 стран с низким количеством самоубийств с 1985 по 2016')

for i in my_plot.containers:
    my_plot.bar_label(i)
```

✓ 0.6s


20 стран с низким количеством самоубийств с 1985 по 2016
Dominica
Saint Kitts and Nevis
San Marino
Antigua and Barbuda
Maldives
Macau
Oman
Grenada
Cabo Verde
Kiribati
Bahamas
Seychelles
Aruba
Saint Vincent and Grenadines
Barbados
Jamaica
Saint Lucia
Fiji
Bosnia and Herzegovina
Belize
0
0
4
11
20
27
33
38
42
53
93
98
101
124
177
184
230
304
318
348
country
0
50
100
150
200
250
300
350
total_suicides

Анализ датасета в SQL и графических библиотеках Python

4. Узнать, какие возрастные группы и пол имеют самый высокий уровень самоубийств за период с 1985 по 2016 год

```sql
SELECT
    age,
    sex,
    SUM(count_of_suicides) AS total_suicides
FROM db_suicides
GROUP BY age, sex
ORDER BY total_suicides DESC;
```

ABC age	ABC sex	123 total_suicides
35-54 years	male	1 945 908
55-74 years	male	1 228 407
25-34 years	male	915 089
15-24 years	male	633 105
35-54 years	female	506 233
75+ years	male	431 134
55-74 years	female	430 036
75+ years	female	221 984
25-34 years	female	208 823
15-24 years	female	175 437
5-14 years	male	35 267
5-14 years	female	16 997

На графике видно, что по миру превалируют самоубийства мужского пола в возрасте от 35 до 54 лет: 1945908

Относительно женщин аналогичной возрастной группы, мужчины совершают суицид в 3,8 раза чаще.

Запускаем SQL запрос из Python и строим визуализацию

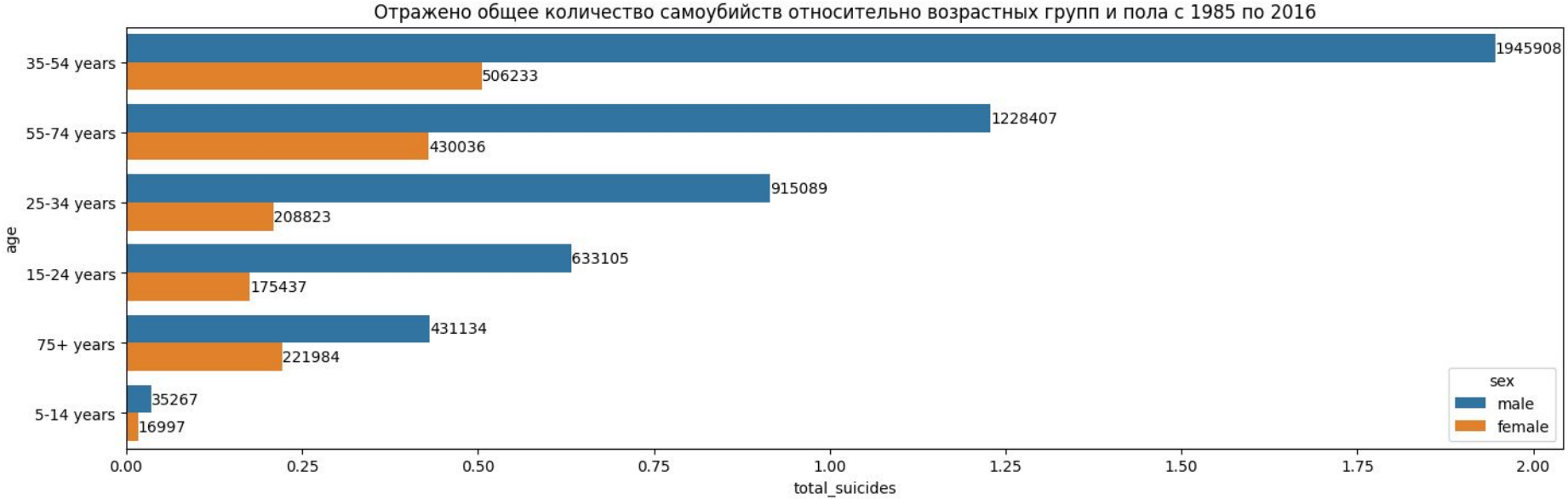
```
plt.figure(figsize=(17, 5))
plt.title('Отражено общее количество самоубийств относительно возрастных групп и пола с 1985 по 2016')
sql_query = """
SELECT
    age,
    sex,
    SUM(count_of_suicides) AS total_suicides
FROM db_suicides
GROUP BY age, sex
ORDER BY total_suicides DESC;
"""
df_from_sql = pd.read_sql_query(sql_query, con=con)

my_plot = sns.barplot(data=df_from_sql, y="age", x="total_suicides", hue="sex", orient="h")

my_plot.set_xticklabels(my_plot.get_xticklabels())

for i in my_plot.containers:
    my_plot.bar_label(i, fmt="%d")
```

✓ 0.6s



Было доказано, что половые различия в уровне самоубийств значительны. Показатели совершенных самоубийств и суицидального поведения у мужчин и женщин различны. В то время как у женщин чаще возникают суицидальные мысли, мужчины чаще кончают жизнь самоубийством. Это также известно как гендерный парадокс самоубийств.Многие исследователи пытались найти объяснение тому, почему пол является столь значимым показателем для самоубийства. Распространенное объяснение опирается на социальные конструкции гегемонной маскулинности и феминности. Согласно литературе по гендеру и самоубийствам, уровень самоубийств среди мужчин объясняется с точки зрения традиционных гендерных ролей. Мужские гендерные роли, как правило, подчеркивают более высокий уровень силы, независимости, рискованного поведения, экономического статуса и индивидуализма. Укрепление этой гендерной роли часто мешает мужчинам обращаться за помощью в связи с суицидальными настроениями и депрессией.

Источник: https://ru.wikipedia.org/wiki/Гендерные_различия_в_самоубийствах

Анализ датасета в SQL и графических библиотеках Python

5. Узнать, в каком году было больше всего самоубийств в каждой стране в период с 1985 по 2016 год

```sql
WITH suicides_per_year AS (
    SELECT
        country,
        year,
        SUM(count_of_suicides) AS suicides_count,
        RANK() OVER (PARTITION BY country ORDER BY SUM(count_of_suicides) DESC) AS max_suicides
    FROM db_suicides
    GROUP BY country, year)
SELECT
    country,
    year AS year_with_most_suicides,
    suicides_count AS max_suicides_count
FROM suicides_per_year
WHERE max_suicides = 1
ORDER BY max_suicides_count DESC;
```

country	year_with_most_suicides	max_suicides_count
Russian Federation	1 994	61 420
United States	2 015	44 189
Japan	2 003	31 881
Republic of Korea	2 011	15 906
Ukraine	1 996	15 160
Germany	1 991	14 010
France	1 986	12 529
Brazil	2 015	11 163
Poland	2 009	6 477
Mexico	2 015	6 234
Sri Lanka	1 997	5 887
Thailand	1 999	5 276
United Kingdom	1 985	5 105
Kazakhstan	1 996	4 773
Italy	1 985	4 759
Canada	1 999	4 074
Hungary	1 992	4 000
Spain	2 014	3 911
Belarus	1 996	3 627
Argentina	2 003	3 289
Romania	2 002	3 067
Australia	2 015	3 027
Philippines	2 011	2 449
Colombia	2 001	2 412

Мы видим, как значимые события в странах не могут не отражаться на количестве самоубийств.
Например:
В России в 1994 году началась чеченская война. 11 октября был черный вторник.
В США в 2015 году был экономический кризис и множество стихийных бедствий.
С 1997 по 2003 год в Японии был экономический кризис.

Анализ датасета в SQL и графических библиотеках Python

6. Узнать количество самоубийств в России за каждый год в период с 1985 по 2016 год

```
SELECT
    country,
    year,
    SUM(count_of_suicides) AS suicides_count
FROM db_suicides
WHERE country LIKE 'Russian Federation'
GROUP BY year
ORDER BY year ASC;
```

country	year	suicides_count
Russian Federation	1 989	37 921
Russian Federation	1 990	39 028
Russian Federation	1 991	39 281
Russian Federation	1 992	45 923
Russian Federation	1 993	55 846
Russian Federation	1 994	61 420
Russian Federation	1 995	60 548
Russian Federation	1 996	57 511
Russian Federation	1 997	54 746
Russian Federation	1 998	51 518
Russian Federation	1 999	56 974
Russian Federation	2 000	56 619
Russian Federation	2 001	56 958
Russian Federation	2 002	55 024
Russian Federation	2 003	51 445
Russian Federation	2 004	49 096
Russian Federation	2 005	45 802
Russian Federation	2 006	42 614
Russian Federation	2 007	41 149
Russian Federation	2 008	38 211
Russian Federation	2 009	37 408
Russian Federation	2 010	33 356
Russian Federation	2 011	31 038
Russian Federation	2 012	29 643

На графике видно, что в датасете данные о самоубийствах в России только с 1989 по 2015 год.

Мы видим, как отражается тяжелый период 90-х и начало 2000-ых. Черный вторник 1994. Чеченская война,. Дефолт 1998.

Запускаем SQL запрос из Python и строим визуализацию

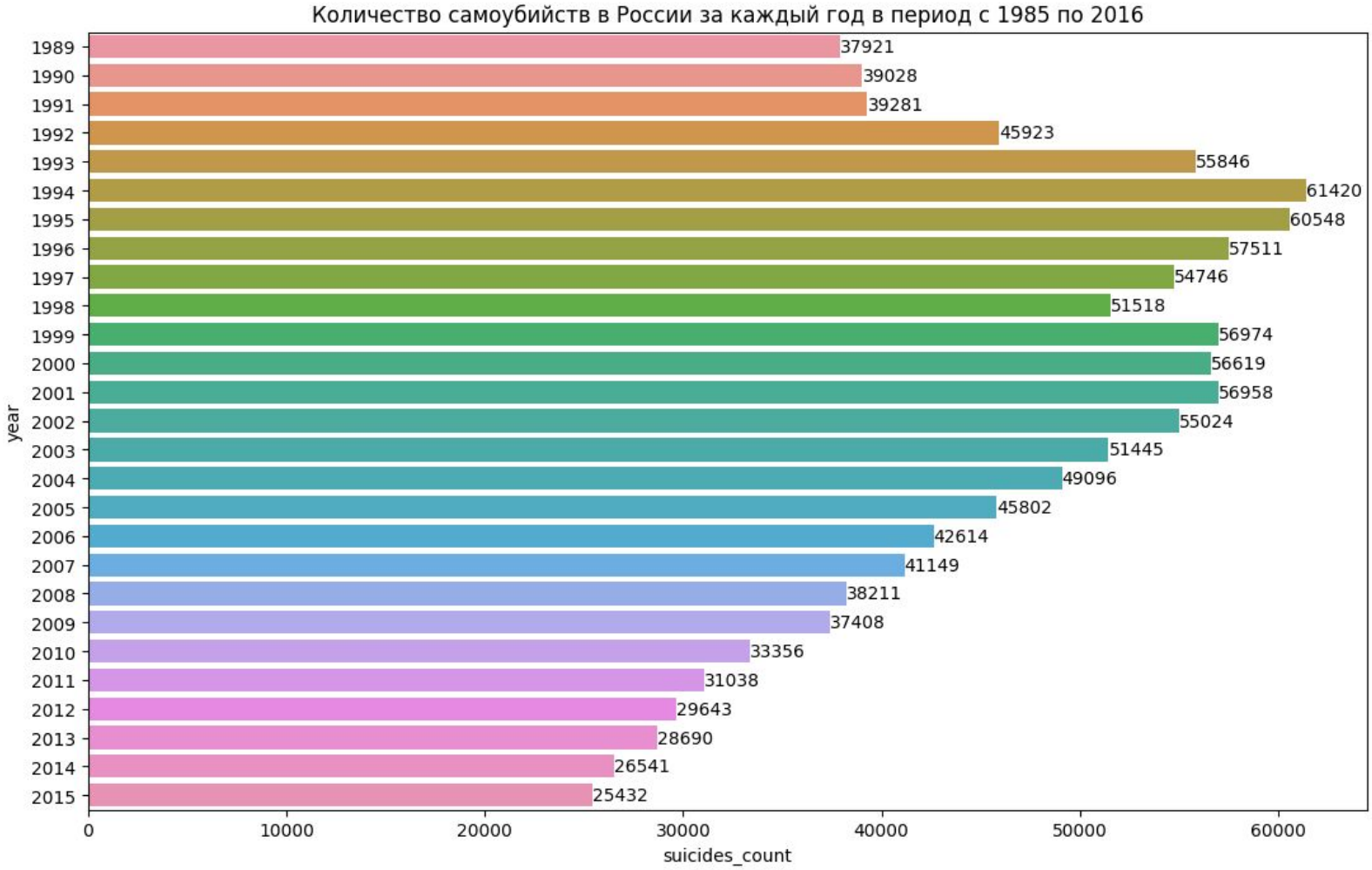
```
plt.figure(figsize=(13, 8))
plt.title('Количество самоубийств в России за каждый год в период с 1985 по 2016')
sql_query = """
SELECT
    country,
    year,
    SUM(count_of_suicides) AS suicides_count
FROM db_suicides
WHERE country LIKE 'Russian Federation'
GROUP BY year
ORDER BY year ASC;
"""
df_from_sql = pd.read_sql_query(sql_query, con=con)

my_plot = sns.barplot(data=df_from_sql, y="year", x="suicides_count", orient="h")

my_plot.set_xticklabels(my_plot.get_xticklabels())

for i in my_plot.containers:
    my_plot.bar_label(i, fmt="%d")
```

✓ 0.8s


Количество самоубийств в России за каждый год в период с 1985 по 2016
year
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
37921
39028
39281
45923
55846
61420
60548
57511
54746
51518
56974
56619
56958
55024
51445
49096
45802
42614
41149
38211
37408
33356
31038
29643
28690
26541
25432
0
10000
20000
30000
40000
50000
60000
suicides_count

Анализ датасета в SQL и графических библиотеках Python

7. Посмотреть, как с 1989 по 2015 год ВВП влияет на количество самоубийств в России

```
SELECT
    country,
    year,
    SUM(count_of_suicides) AS suicides_count,
    gdp_per_capita_dollars AS gdp_capita
FROM db_suicides
WHERE country LIKE 'Russian Federation'
GROUP BY year
ORDER BY year ASC;
```

country	year	suicides_count	gdp_capita
Russian Federation	1 989	37 921	3 740
Russian Federation	1 990	39 028	3 789
Russian Federation	1 991	39 281	3 773
Russian Federation	1 992	45 923	3 333
Russian Federation	1 993	55 846	3 160
Russian Federation	1 994	61 420	2 853
Russian Federation	1 995	60 548	2 844
Russian Federation	1 996	57 511	2 813
Russian Federation	1 997	54 746	2 907
Russian Federation	1 998	51 518	1 948
Russian Federation	1 999	56 974	1 412
Russian Federation	2 000	56 619	1 879
Russian Federation	2 001	56 958	2 229
Russian Federation	2 002	55 024	2 527
Russian Federation	2 003	51 445	3 141
Russian Federation	2 004	49 096	4 312
Russian Federation	2 005	45 802	5 611
Russian Federation	2 006	42 614	7 313
Russian Federation	2 007	41 149	9 643
Russian Federation	2 008	38 211	12 359
Russian Federation	2 009	37 408	9 118
Russian Federation	2 010	33 356	11 307
Russian Federation	2 011	31 038	15 226
Russian Federation	2 012	29 643	16 413

На предыдущем графике мы видели, как после 2003 года количество самоубийств плавно идет на снижение, а на этом графике мы видим, как идет плавное увеличение ВВП на душу населения. Нельзя не заметить отражение событий на графике 2008-2009 годов(кризис на финансовых рынках России) и 2014-2015 годов(валютный кризис).

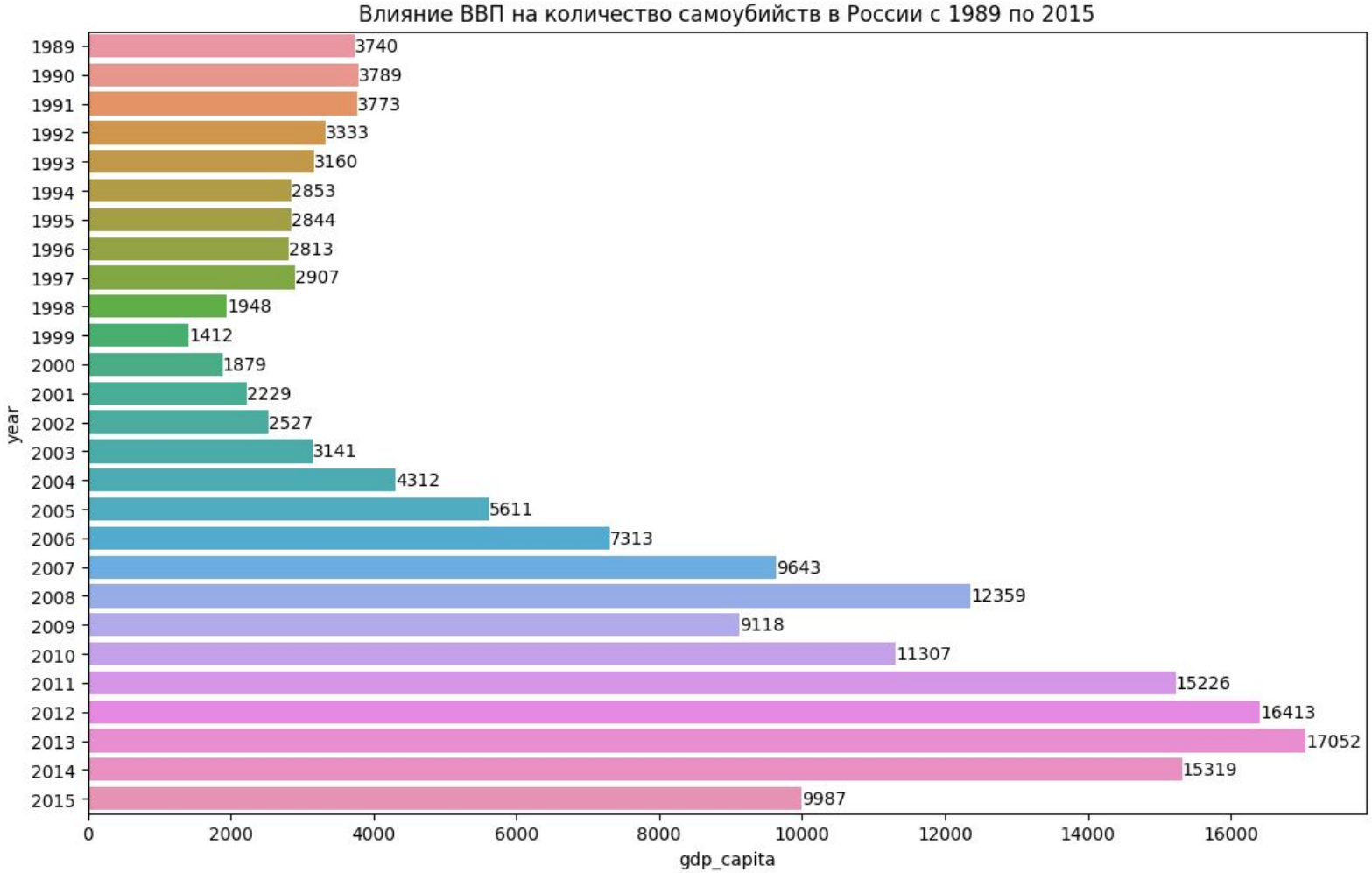
Запускаем SQL запрос из Python и строим визуализацию

```
plt.figure(figsize=(13, 8))
plt.title('Влияние ВВП на количество самоубийств в России с 1989 по 2015')
sql_query = """
SELECT
    country,
    year,
    SUM(count_of_suicides) AS suicides_count,
    gdp_per_capita_dollars AS gdp_capita
FROM db_suicides
WHERE country LIKE 'Russian Federation'
GROUP BY year
ORDER BY year ASC
"""
df_from_sql = pd.read_sql_query(sql_query, con=con)

my_plot = sns.barplot(data=df_from_sql, y="year", x="gdp_capita", orient="h")
my_plot.set_xticklabels(my_plot.get_xticklabels())

for i in my_plot.containers:
    my_plot.bar_label(i)
```

✓ 0.8s


Влияние ВВП на количество самоубийств в России с 1989 по 2015
year
gdp_capita
1989 3740
1990 3789
1991 3773
1992 3333
1993 3160
1994 2853
1995 2844
1996 2813
1997 2907
1998 1948
1999 1412
2000 1879
2001 2229
2002 2527
2003 3141
2004 4312
2005 5611
2006 7313
2007 9643
2008 12359
2009 9118
2010 11307
2011 15226
2012 16413
2013 17052
2014 15319
2015 9987
0
2000
4000
6000
8000
10000
12000
14000
16000

Анализ датасета в SQL и графических библиотеках Python

8. Посмотреть соотношение самоубийств по возрасту и полу в России с 1989 по 2015 год

```
SELECT
    country,
    sex,
    age,
    SUM(count_of_suicides) AS suicides_count
FROM db_suicides
WHERE country LIKE 'Russian Federation'
GROUP BY sex, age
ORDER BY suicides_count DESC;
```

country	sex	age	suicides_count
Russian Federation	male	35-54 years	414 090
Russian Federation	male	55-74 years	205 284
Russian Federation	male	25-34 years	204 787
Russian Federation	male	15-24 years	126 379
Russian Federation	female	35-54 years	65 050
Russian Federation	female	55-74 years	62 469
Russian Federation	male	75+ years	38 034
Russian Federation	female	75+ years	36 177
Russian Federation	female	25-34 years	26 400
Russian Federation	female	15-24 years	22 232
Russian Federation	male	5-14 years	6 838
Russian Federation	female	5-14 years	2 002

Разрыв по количеству самоубийств 75+ между полами в России сравнялись относительно результатов по миру.

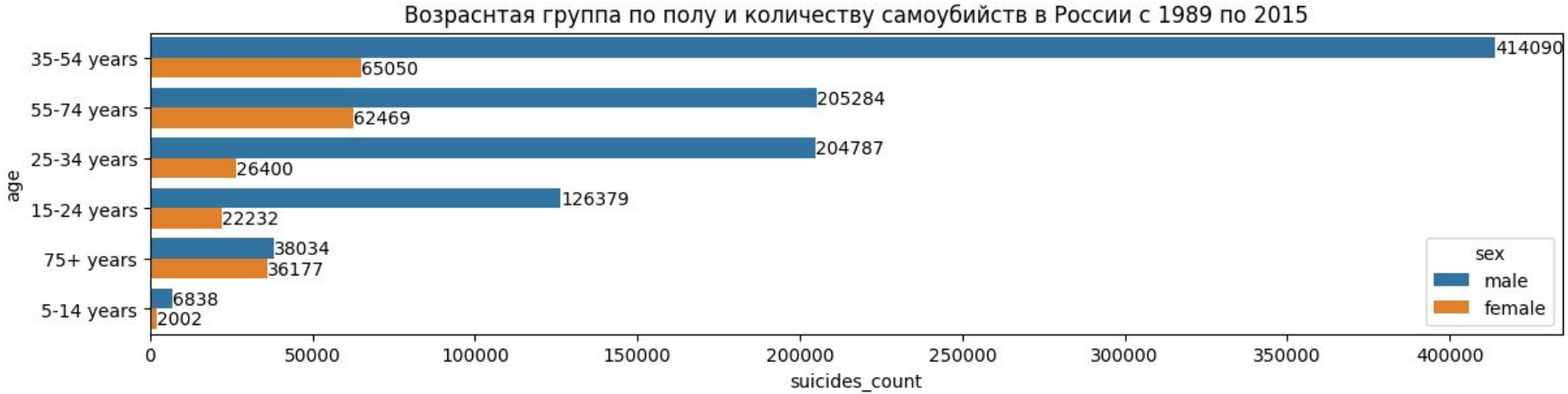
Запускаем SQL запрос из Python и строим визуализацию

```
plt.figure(figsize=(14, 3))
plt.title('Возрастная группа по полу и количеству самоубийств в России с 1989 по 2015')
sql_query = """
SELECT
    country,
    sex,
    age,
    SUM(count_of_suicides) AS suicides_count
FROM db_suicides
WHERE country LIKE 'Russian Federation'
GROUP BY sex, age
ORDER BY suicides_count DESC;
"""
df_from_sql = pd.read_sql_query(sql_query, con=con)

my_plot = sns.barplot(data=df_from_sql, y="age", x="suicides_count", hue='sex', orient="h")
my_plot.set_xticklabels(my_plot.get_xticklabels())

for i in my_plot.containers:
    my_plot.bar_label(i, fmt="%d")
```

✓ 0.6s



Выводы исследования

- Каждое значимое событие в стране или мире влияет на уровень самоубийств
- Социальные конструкции и гендерные роли мешают мужскому полу обращаться за помощью в связи с суицидальными настроениями и депрессией
- В исследуемом интервале Россия лидирует по количеству самоубийств. Россия еще в пути к нормализации обращения за психологической помощью в обществе
- Что-то с этим миром не так, если такое количество человек не хочет здесь жить