

PROJECT TITLE

*SUICIDE TRENDS
ACROSS THE WORLD*

PROJECT DESCRIPTION

A) Problem Statement

- To study and visualize the Suicide records of 101 countries through the GDP per Capita, Population, , GDP per year , Sex, Age records from the dataset ranging between the years 1985 to 2016

B) Problem Analysis :

The fundamental Computational issue for this problem is the development of an algorithm and see What countries rank the first in overall suicides over the years and relation of the suicides to the GDP of the country.

B) Problem Design :

I. Pre-Requisites

- Knowledge of python
- The data set
- Anaconda with Jupyter Notebook

2.DATA FEATURE LIST

- country-name of the country whose data is being displayed
- Year-The particular year for which data is given
- Sex-Whether the person is male or female
- Age-Specifies the various age groups
- suicides_no-Number of suicides taking place
- Population-It depicts the effective populations of the country
- suicides/ 100k pop- Number of suicides taking place/ 100k
- gdp for year - GDP of that country for that particular year
- gdp per capita- measure of a country's economic output that accounts for its number of people
- Generation-specifies the generation on the basis of age group.
Namely- *Generation X ; Boomer ; Silent ; G . I . Generation*
(15-24 yrs.) (25-34 yrs.) (35-54 yrs.) (75+ yrs.)

PROGRAMMING STEPS

- **STEP 1:**

Import the Dataset(csv file) to the jupyter notebook.

- **STEP 2:**

Analyze the Dataset.

- **STEP 3:**

Gather the relevant information from the Dataset.

- **STEP 4:**

Make the Queries from the information gathered.

- **STEP 5:**

Find the solutions of the queries and visualize the result



IMPLEMENTING QUERRIES AND OUTPUT

CHECKING HEAD OF THE DATAFRAME:

```
df.head()
```

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

DISPLAYING ROWS AND COLUMNS OF THE DATAFRAME

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 27820 entries, 0 to 27819  
Data columns (total 12 columns):  
country                27820 non-null object  
year                   27820 non-null int64  
sex                    27820 non-null object  
age                    27820 non-null object  
suicides_no            27820 non-null int64  
population              27820 non-null int64  
suicides/100k pop      27820 non-null float64  
country-year           27820 non-null object  
HDI for year           8364 non-null float64  
  gdp_for_year ($)     27820 non-null object  
gdp_per_capita ($)     27820 non-null int64  
generation             27820 non-null object  
dtypes: float64(2), int64(4), object(6)  
memory usage: 2.5+ MB
```

CHECK AND DROP NULL VALUES

```
df.isnull().sum()
```

```
country      0
year         0
sex          0
age          0
suicides_no  0
population   0
suicides/100k pop  0
country-year  0
HDI for year 19456
gdp_for_year ($)  0
gdp_per_capita ($)  0
generation   0
dtype: int64
```

Checking for the null values

```
df.drop(['HDI for year', 'country-year'],axis=1,inplace=True)
```

```
df.columns
```

```
Index(['country', 'year', 'sex', 'age', 'suicides_no', 'population',
      'suicides/100k pop', 'gdp_for_year ($)', 'gdp_per_capita ($)',
      'generation'],
      dtype='object')
```

Dropping for the null values

COUNTRY WITH MOST SUICIDES IN A PARTICULAR YEAR

Q.5 Which country has maximum number of suicides and in which years ?

```
df[df['suicides_no']==df['suicides_no'].max()][['country','year']]
```

	country	year
20996	Russian Federation	1994

```

In [15]: df[df['suicides_no'] == df['suicides_no'].min()][['country']]
Out[15]:
9          Albania
10         Albania
11         Albania
22         Albania
23         Albania
33         Albania
34         Albania
35         Albania
44         Albania
45         Albania
46         Albania
47         Albania
59         Albania
71         Albania
107        Albania
119        Albania
131        Albania
142        Albania
143        Albania
179        Albania
192        Albania
193        Albania
194        Albania
195        Albania
196        Albania
197        Albania
198        Albania
199        Albania
200        Albania
201        Albania

...
26427      United Arab Emirates
26436      United Arab Emirates
26437      United Arab Emirates
26438      United Arab Emirates
26439      United Arab Emirates
26447      United Arab Emirates
26448      United Arab Emirates
26449      United Arab Emirates
26450      United Arab Emirates
26451      United Arab Emirates
26460      United Arab Emirates
26461      United Arab Emirates
26462      United Arab Emirates
26463      United Arab Emirates
26473      United Arab Emirates
26474      United Arab Emirates
26475      United Arab Emirates
26547      United Kingdom
26571      United Kingdom
26619      United Kingdom
27243      Uruguay
27267      Uruguay
27279      Uruguay
27315      Uruguay
27327      Uruguay
27363      Uruguay
27459      Uruguay
27471      Uruguay
27495      Uruguay
27543      Uruguay
Name: country, Length: 4281, dtype: object

```

COUNTRY WITH LEAST SUICIDES

1. Albania
2. United Arab Emirates
3. United Kingdom
4. Uruguay

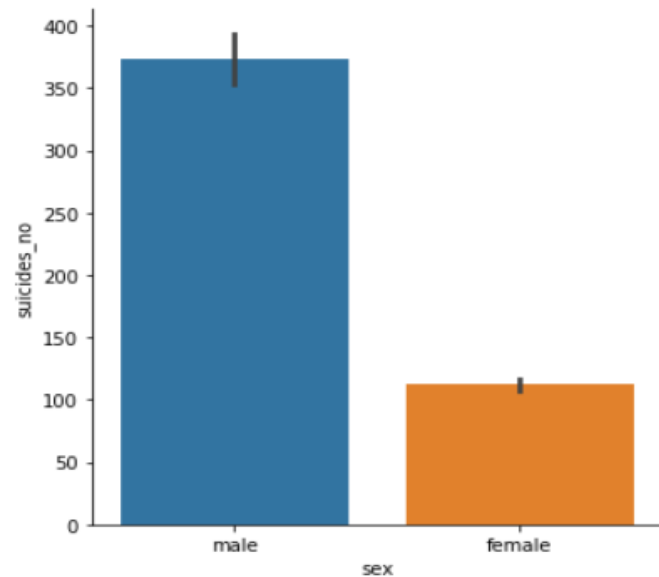
MALE AND FEMALE SUICIDES COMPARISION

```
sns.factorplot(x='sex',y='suicides_no',data=df,kind='bar')
```

C:\Users\ritik\Anaconda3\lib\site-packages\seaborn\categorical.py:3666: UserWarning: The `factorplot` function has been renamed to `catplot`. The original name will be removed in a future release. Please update your code. Note that the default `kind` in `factorplot` (`'point'`) has changed to `strip` in `catplot`.

```
warnings.warn(msg)
```

<seaborn.axisgrid.FacetGrid at 0x24cfc157860>



NUMBER OF SUICIDES IN VARIOUS AGE GROUPS

```
df.groupby('age')['suicides_no'].count()
```

age

15-24 years 4642

25-34 years 4642

35-54 years 4642

5-14 years 4610

55-74 years 4642

75+ years 4642

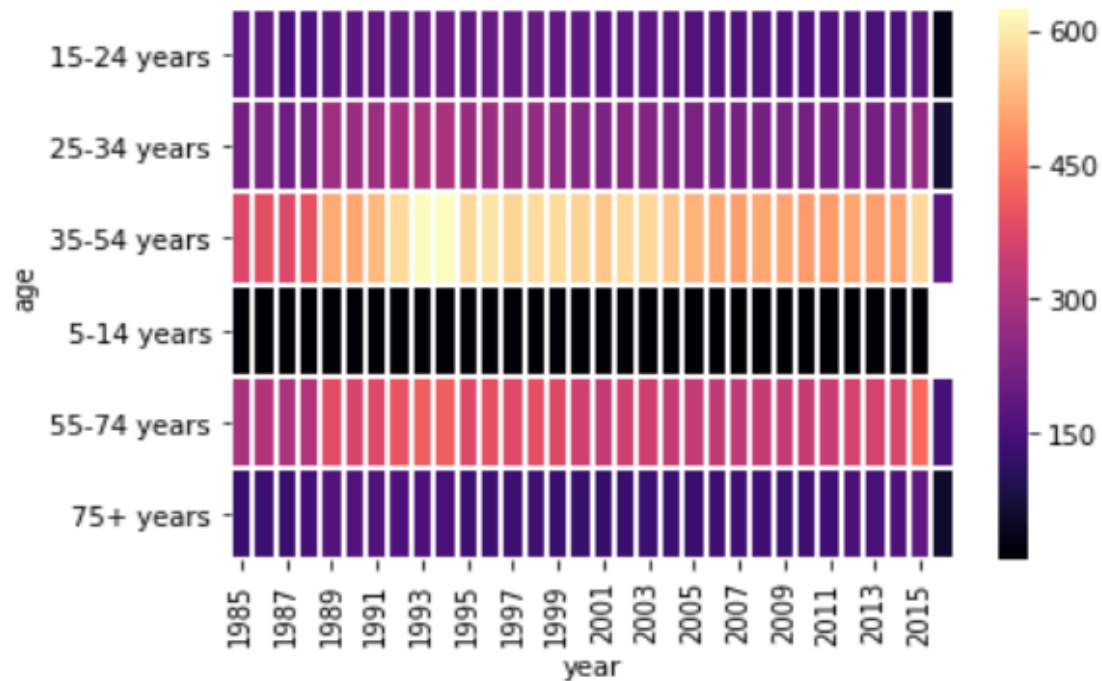
Name: suicides_no, dtype: int64

AGE GROUP HAVING MAX SUICIDAL TENDENCY

```
pv = df.pivot_table(values='suicides_no',index='age',columns='year')
```

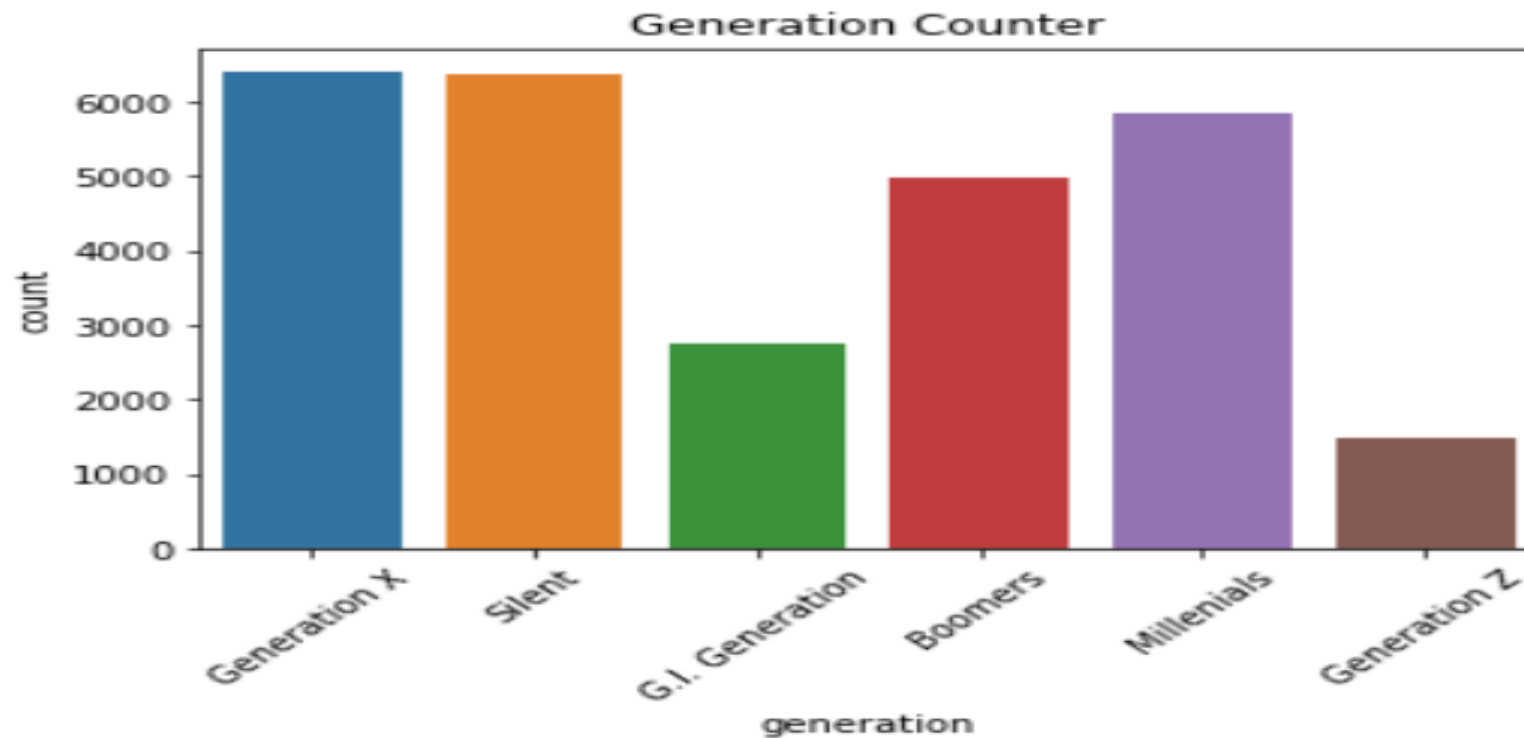
```
sns.heatmap(pv,cmap='magma',linecolor='white',linewidths=1)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x21186f15b70>
```



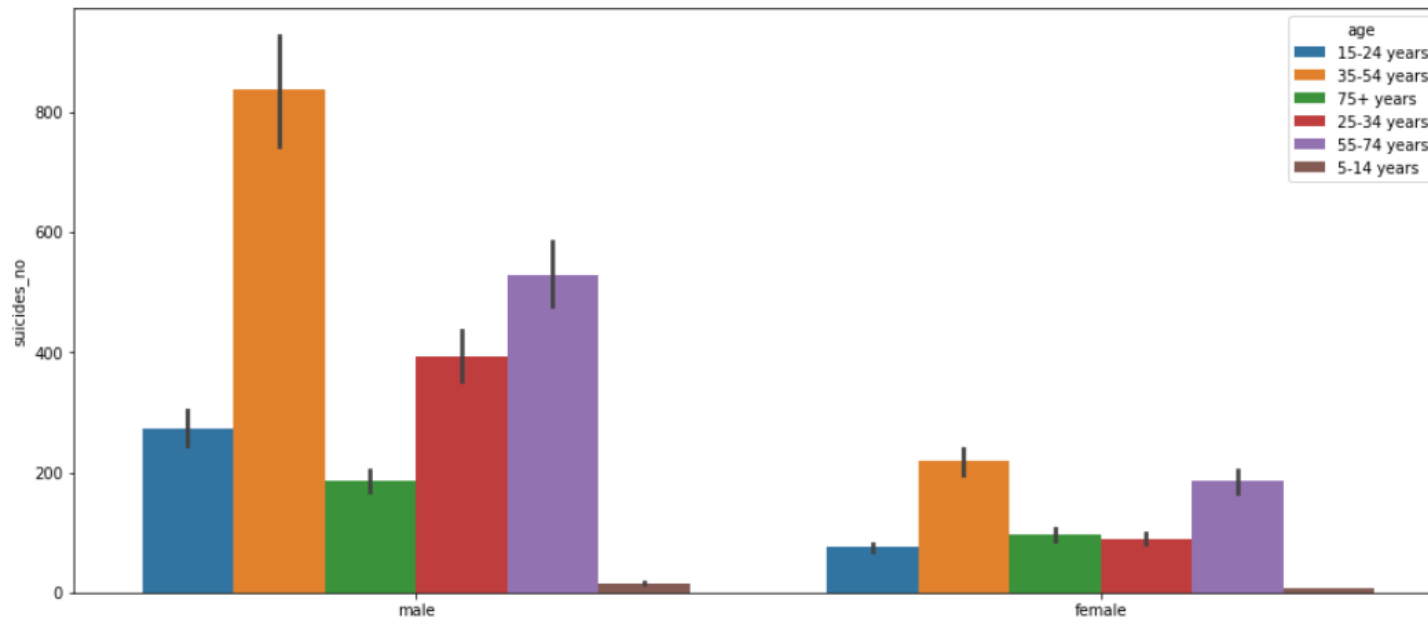
GENERATION HAVING MAX SUICIDAL TENDENCY

```
sns.countplot(df.generation)
plt.title('Generation Counter')
plt.xticks(rotation=45)
plt.show()
```



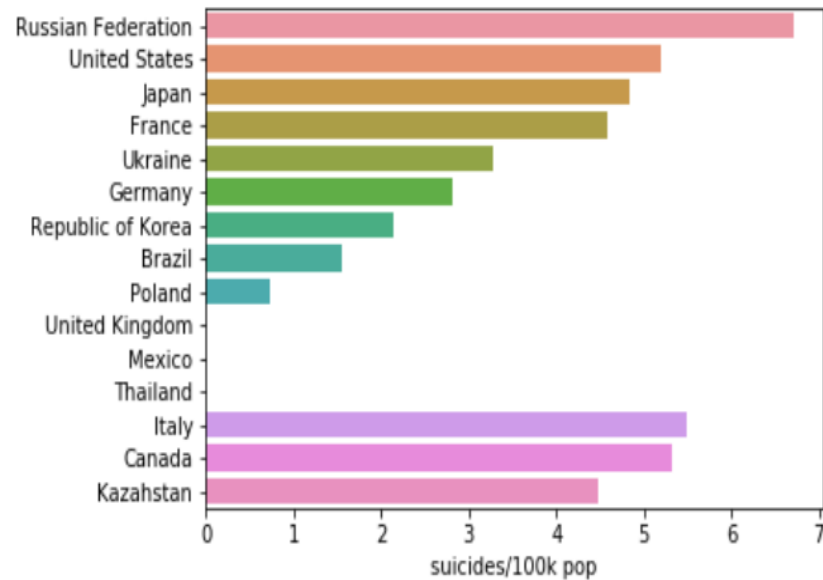
COMPARISON OF SUICIDES BASED ON AGE GROUP AND SEX

```
plt.figure(figsize=(16,7))  
bar_age = sns.barplot(x = 'sex', y = 'suicides_no', hue = 'age', data = df)
```



COUNTRY HAVING MOST SUICIDES PER 100K POPULATION

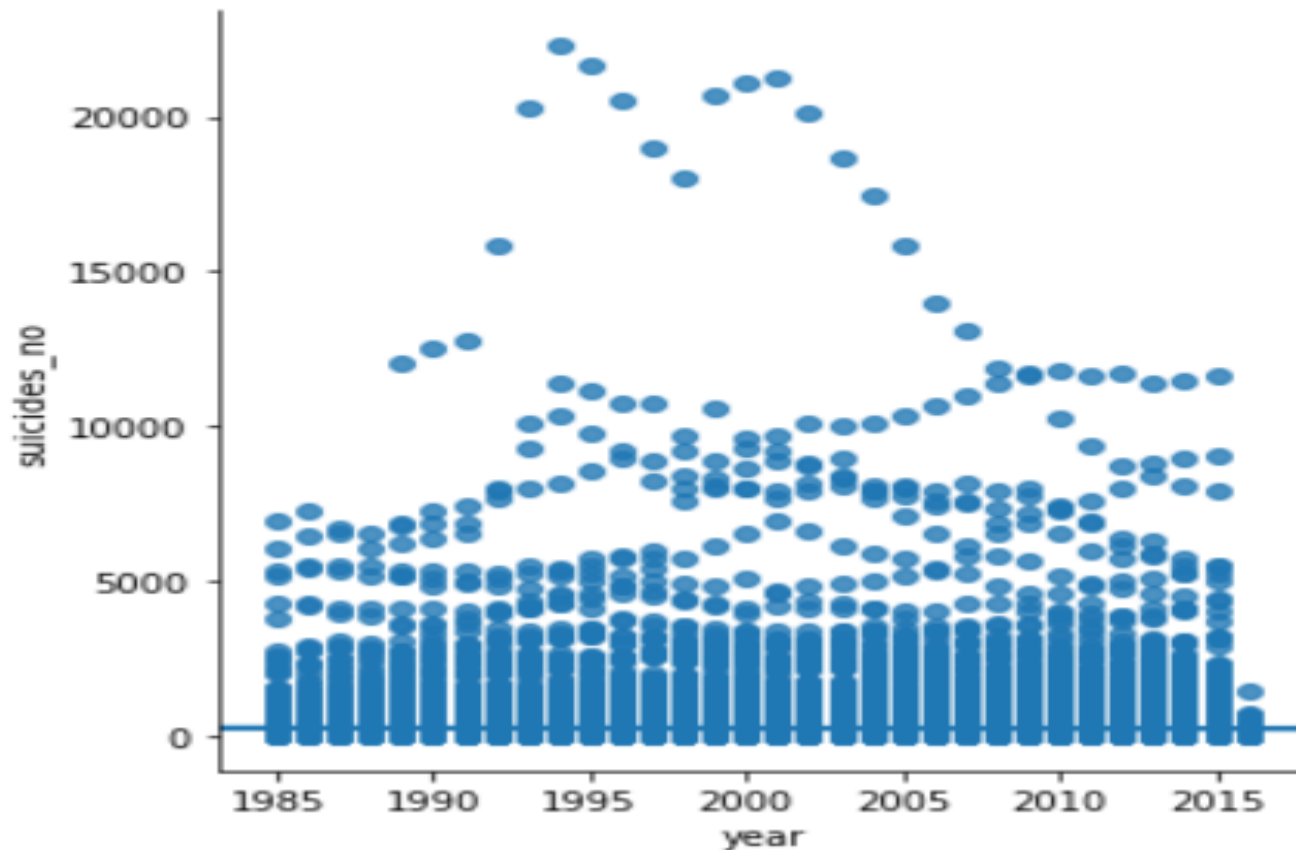
```
country=['Russian Federation','United States','Japan','France','Ukraine','Germany','Republic of Korea','Brazil',  
suicide=df['suicides/100k pop']  
sns.barplot(y=country[:15],x=suicide[:15])  
plt.show()
```



NUMBER OF SUICIDES PER YEAR

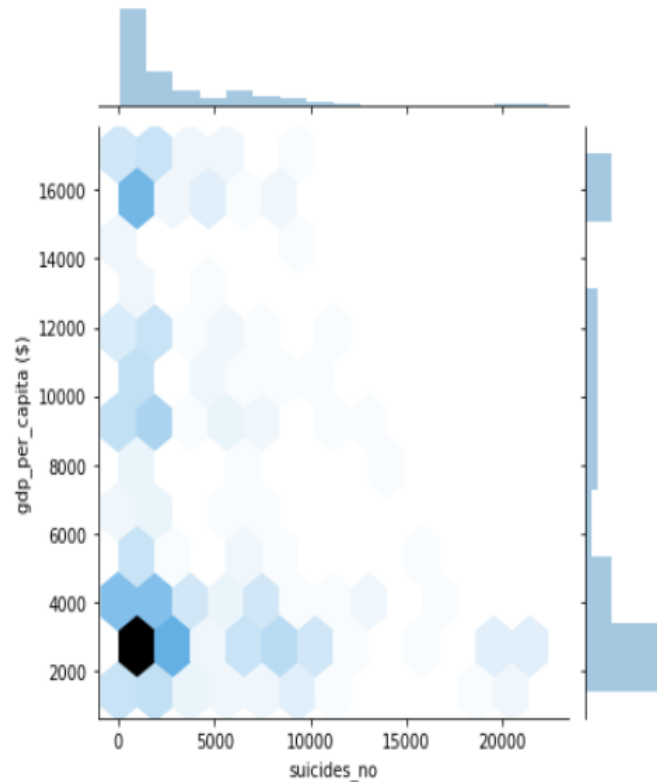
```
sns.lmplot(x='year',y='suicides_no',data=df)
```

```
<seaborn.axisgrid.FacetGrid at 0x21187d22ac8>
```



AFFECT ON THE GDP OF COUNTRY WITH MAX NUMBER OF SUICIDES

```
fig=sns.jointplot(y='gdp_per_capita ($)',x='suicides_no',kind='hex',data=df[df['country']=='Russian Federation'])  
plt.show()
```



GDP PER CAPITA OF MOST AFFECTED COUNTRIES

```
df[df['suicides_no']==df['suicides_no'].max()][['country','year','gdp_per_capita ($)']]
```

	country	year	gdp_per_capita (\$)
20996	Russian Federation	1994	2853

GDP PER CAPITA OF LEAST AFFECTED COUNTRIES

```
gk=df[df['suicides_no']==df['suicides_no'].min()][['country','year','gdp_per_capita ($)']]
gk=gk.groupby(['country'])
gk=gk.first()
gk.head()
```

	year	gdp_per_capita (\$)
country		
Albania	1987	796
Antigua and Barbuda	1985	3850
Armenia	1990	756
Aruba	1995	17949
Australia	1990	19665

THANK YOU

EFFORTS BY-

- RITIKA RAO(18CSUI77)