Problem Statement:

This purpose of this project is to carry out a detailed analysis and visualization of death causes across the globe using the 'Death Cause Reason by Country' datasets. The dataset can be found here

Introduction

The key points to consider are as follows:

- Total number of deaths per cause across the globe
- Analysis of the leading cause of death
- Analysis of the top 25 countries with the most recorded deaths from the leading cause

In [1]:
```
import numpy as np
import pandas as pd
from wordcloud import WordCloud

import matplotlib.pyplot as plt
import seaborn as sns
from PIL import Image


import warnings
warnings.filterwarnings("ignore")
```

In [3]:
```
df = pd.read_csv("Death Cause Reason by Country.csv")
pd.set_option('display.max_columns',50)
```

Death causes by country EDA and Stats

In [4]:
```
df.describe()
```

Out[4]:

	Covid-19 Deaths	Cardiovascular diseases	Respiratory diseases	Kidney diseases	Neonatal disorders	Meningitis	
count	188.000000	1.910000e+02	1.910000e+02	191.000000	191.000000	191.000000	19
mean	10461.404255	9.808936e+04	2.104242e+04	7490.712042	9759.513089	1217.586387	327
std	35107.423980	3.956782e+05	1.162353e+05	24085.215522	39729.671493	4486.652415	1531
min	1.000000	8.000000e+00	1.000000e+00	1.000000	0.000000	0.000000	
25%	84.750000	4.648500e+03	5.295000e+02	342.000000	104.500000	13.500000	
50%	612.500000	1.815700e+04	2.529000e+03	1813.000000	549.000000	83.000000	
75%	3986.750000	5.737250e+04	7.118000e+03	4773.500000	5296.500000	474.000000	92

	Covid-19 Deaths	Cardiovascular diseases	Respiratory diseases	Kidney diseases	Neonatal disorders	Meningitis	
max	353948.000000	4.584273e+06	1.168381e+06	222922.000000	438004.000000	44914.000000	191106

In [5]:
```
df.shape
```

Out[5]:
```
(191, 32)
```

In [6]:
```
df.head()
```

Out[6]:

	Country Name	Covid-19 Deaths	Cardiovascular diseases	Respiratory diseases	Kidney diseases	Neonatal disorders	Meningitis	Malaria	Interperso viole
0	Afghanistan	2201.0	61995	7082	5637	23701	1563	530	5
1	Albania	1181.0	12904	815	329	161	13	0	
2	Algeria	2762.0	97931	7528	8201	8756	292	0	
3	Andorra	84.0	169	39	16	0	0	0	
4	Angola	33.0	25724	3934	2464	18189	2520	10784	

In [7]:
```
df.tail(15)
```

Out[7]:

	Country Name	Covid-19 Deaths	Cardiovascular diseases	Respiratory diseases	Kidney diseases	Neonatal disorders	Meningitis	Malaria	Inter
176	Turkmenistan	NaN	18323	442	712	1113	52	0	
177	Tuvalu	3.0	42	7	5	1	0	0	
178	Uganda	265.0	28149	5480	3386	30955	3941	22587	
179	Ukraine	19437.0	449376	12181	2660	1305	399	0	
180	United Arab Emirates	671.0	8621	1778	1433	83	54	0	
181	United Kingdom	74185.0	188113	51879	7776	1484	265	0	
182	United States	353948.0	957455	224988	106954	10994	1146	0	
183	Uruguay	193.0	10003	2155	989	148	14	0	

	Country Name	Covid-19 Deaths	Cardiovascular diseases	Respiratory diseases	Kidney diseases	Neonatal disorders	Meningitis	Malaria	Inter
184	Uzbekistan	614.0	107394	3837	3971	5257	175	0	
185	Venezuela	1030.0	61510	7062	10405	3338	254	667	
186	Vietnam	35.0	240161	36229	21485	6468	747	59	
187	Wales	512.0	10519	2892	426	57	14	0	
188	Yemen	610.0	55752	5738	2676	18982	479	5089	
189	Zambia	390.0	18157	2388	1931	9851	2065	4777	
190	Zimbabwe	369.0	17810	2891	2292	8609	1450	2068	

In [8]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 191 entries, 0 to 190
Data columns (total 32 columns):
 #   Column                                  Non-Null Count  Dtype
---  ------                                  --------------  -----
 0   Country Name                            191 non-null    object
 1   Covid-19 Deaths                         188 non-null    float64
 2   Cardiovascular diseases                 191 non-null    int64
 3   Respiratory diseases                    191 non-null    int64
 4   Kidney diseases                         191 non-null    int64
 5   Neonatal disorders                      191 non-null    int64
 6   Meningitis                              191 non-null    int64
 7   Malaria                                 191 non-null    int64
 8   Interpersonal violence                  191 non-null    int64
 9   HIV/AIDS                                191 non-null    int64
 10  Tuberculosis                            191 non-null    int64
 11  Maternal disorders                      191 non-null    int64
 12  Lower respiratory infections            191 non-null    int64
 13  Alcohol use disorders                   191 non-null    int64
 14  Diarrheal diseases                      191 non-null    int64
 15  Poisoning                               191 non-null    int64
 16  Nutritional deficiencies                191 non-null    int64
 17   Alzheimer's disease                    191 non-null    int64
 18  Parkinson's disease                     191 non-null    int64
 19   Acute hepatitis                        191 non-null    int64
 20  Digestive diseases                      191 non-null    int64
 21   Cirrhosis and other chronic liver diseases  191 non-null    int64
 22  Protein-energy malnutrition             191 non-null    int64
 23  Neoplasms                               191 non-null    int64
 24  Fire, heat                              191 non-null    int64
 25  Drowning                                191 non-null    int64
 26  Drug use disorders                      191 non-null    int64
 27  Road injuries                           191 non-null    int64
 28  Environmental heat and cold exposure    191 non-null    int64
 29  Self-harm                               191 non-null    int64
 30   Conflict and terrorism                 191 non-null    int64
```

```
 31  Diabetes                                        191 non-null    int64
dtypes: float64(1), int64(30), object(1)
memory usage: 47.9+ KB
```

Let's check the top 20 countries with the highest number of deaths from a given disease

In [10]:
```
causes = list(df.columns)
causes = causes[1:-2]
causes
```

Out[10]:
```
['Covid-19 Deaths',
 'Cardiovascular diseases',
 'Respiratory diseases ',
 'Kidney diseases',
 'Neonatal disorders ',
 'Meningitis ',
 'Malaria ',
 'Interpersonal violence',
 'HIV/AIDS',
 'Tuberculosis',
 'Maternal disorders',
 'Lower respiratory infections',
 'Alcohol use disorders',
 'Diarrheal diseases',
 'Poisoning',
 'Nutritional deficiencies',
 " Alzheimer's disease",
 "Parkinson's disease",
 ' Acute hepatitis',
 'Digestive diseases',
 ' Cirrhosis and other chronic liver diseases',
 'Protein-energy malnutrition',
 'Neoplasms',
 'Fire, heat',
 'Drowning',
 'Drug use disorders',
 'Road injuries',
 'Environmental heat and cold exposure',
 'Self-harm']
```

In [16]:
```
#Top 10 countries with the highest number of deaths from a given disease
causes_10 = {}
for cause in causes:
    top_ten = df[["Country Name", cause]].sort_values(ascending=False,by=cause)
    causes_10[cause] = top_ten[:10]
    print(top_ten[:10])
```

```
       Country Name  Covid-19 Deaths
182   United States         353948.0
24           Brazil         195541.0
75            India         149218.0
104          Mexico         126507.0
133            Peru          93231.0
81            Italy          74621.0
54          England          74185.0
181  United Kingdom          74185.0
59           France          64777.0
```

```
140          Russia          56798.0
      Country Name  Cardiovascular diseases
35           China                 4584273
75           India                 2574410
140         Russia                 1004931
182  United States                  957455
76       Indonesia                  651481
179        Ukraine                  449376
24          Brazil                  397993
83           Japan                  372483
63         Germany                  364285
127       Pakistan                  341108
       Country Name  Respiratory diseases
75            India               1168381
35            China               1085273
182   United States                224988
76        Indonesia                100930
127        Pakistan                 82758
24           Brazil                 76551
13        Bangladesh                 70893
83            Japan                 55794
181  United Kingdom                 51879
63          Germany                 45165
      Country Name  Kidney diseases
75            India           222922
35            China           196726
182  United States           106954
104         Mexico            72539
83            Japan            45324
127       Pakistan            44390
24          Brazil            42336
76        Indonesia            42131
63          Germany            35220
134    Philippines             33358
                    Country Name  Neonatal disorders
75                         India              438004
127                     Pakistan              249785
120                      Nigeria              195397
56                      Ethiopia               77436
35                         China               44950
46   Democratic Republic of Congo               42976
13                     Bangladesh               40175
168                     Tanzania               39594
76                     Indonesia               38125
100                         Mali               32006
                    Country Name  Meningitis
120                      Nigeria       44914
75                         India       34736
127                     Pakistan       17987
56                      Ethiopia       11283
119                        Niger        7772
35                         China        6465
100                         Mali        6260
46   Democratic Republic of Congo        6147
27                   Burkina Faso        5258
76                      Indonesia        4715
                    Country Name  Malaria
120                      Nigeria   191106
46   Democratic Republic of Congo    57160
75                         India    33372
```

```
41                 Cote d'Ivoire     29398
27                  Burkina Faso     26305
119                        Niger     23822
30                      Cameroon     22652
178                       Uganda     22587
64                         Ghana     21597
111                     Mozambique     20528
      Country Name  Interpersonal violence
24          Brazil                   65907
75           India                   40954
104         Mexico                   30673
140         Russia                   22285
156   South Africa                   19122
182  United States                   17709
36        Colombia                   17061
134    Philippines                   15570
127       Pakistan                   13446
120        Nigeria                   12958
     Country Name  HIV/AIDS
156  South Africa    143851
120       Nigeria     82270
111    Mozambique     66304
86          Kenya     51135
75          India     46298
35          China     31746
168      Tanzania     27125
56       Ethiopia     26591
30       Cameroon     23172
189        Zambia     22540
                   Country Name  Tuberculosis
75                        India        422634
76                    Indonesia         76549
127                    Pakistan         62774
46   Democratic Republic of Congo         46145
120                     Nigeria         45278
35                        China         36566
56                      Ethiopia         29874
134                  Philippines         29181
13                    Bangladesh         29086
111                   Mozambique         20700
                   Country Name  Maternal disorders
75                        India               41843
120                      Nigeria               17650
127                     Pakistan               16162
46   Democratic Republic of Congo               10119
56                      Ethiopia                7634
13                    Bangladesh                6598
168                     Tanzania                5506
76                    Indonesia                5264
0                    Afghanistan                4038
86                        Kenya                3767
                   Country Name  Lower respiratory infections
75                        India                        433661
35                        China                        185264
120                      Nigeria                        172978
83                         Japan                        118801
24                        Brazil                         88641
182                 United States                         81921
127                      Pakistan                         68193
134                   Philippines                         58413
```

```
56                      Ethiopia                         46301
46   Democratic Republic of Congo                        44646
      Country Name  Alcohol use disorders
140         Russia                  28460
75           India                  23261
35           China                  19459
182  United States                  14044
24          Brazil                   9705
179        Ukraine                   7716
63         Germany                   6379
104         Mexico                   4518
135         Poland                   4380
59          France                   3748
                    Country Name  Diarrheal diseases
75                         India              632345
120                      Nigeria              181138
127                     Pakistan               77024
76                     Indonesia               59589
56                      Ethiopia               50773
46   Democratic Republic of Congo               34003
13                    Bangladesh               32927
33                          Chad               30372
119                        Niger               28694
86                         Kenya               18564
                   Country Name  Poisoning
35                         China      27084
140                       Russia       5805
75                         India       3945
120                      Nigeria       3640
56                      Ethiopia       2623
127                     Pakistan       2106
179                      Ukraine       1539
182                United States       1528
46   Democratic Republic of Congo       1268
168                     Tanzania       1020
    Country Name  Nutritional deficiencies
75         India                     26868
76     Indonesia                     20348
35         China                     16863
100         Mali                     14865
127     Pakistan                     14631
56      Ethiopia                      8989
24        Brazil                      8221
155      Somalia                      7626
104       Mexico                      7558
168     Tanzania                      6887
       Country Name   Alzheimer's disease
35            China                320715
83            Japan                164874
182   United States                143919
75            India                129011
24           Brazil                 54594
63          Germany                 49557
81            Italy                 48763
140          Russia                 48437
59           France                 46008
181  United Kingdom                 32425
       Country Name  Parkinson's disease
35             China                76990
75             India                45305
```

```
182   United States               32211
83            Japan               16750
63          Germany               12875
24           Brazil                9564
140          Russia                9390
81            Italy                8036
59           France                7997
181  United Kingdom                7906
    Country Name   Acute hepatitis
75         India             30987
127     Pakistan              8516
76     Indonesia              5560
13    Bangladesh              4437
35         China              3726
120      Nigeria              2540
0    Afghanistan              1940
56      Ethiopia              1626
86         Kenya              1476
52         Egypt              1356
     Country Name  Digestive diseases
75           India              464914
35           China              277142
182  United States              129343
76       Indonesia              126347
140         Russia               90887
24           Brazil               78484
104         Mexico               71124
120        Nigeria               70077
52            Egypt               69216
127       Pakistan               62351
     Country Name   Cirrhosis and other chronic liver diseases
75           India                                       270037
35           China                                       152262
76       Indonesia                                        88670
182  United States                                        67286
52            Egypt                                       62635
104         Mexico                                       46328
140         Russia                                       46102
127       Pakistan                                       44070
120        Nigeria                                       40381
24           Brazil                                       38029
    Country Name  Protein-energy malnutrition
76     Indonesia                       18641
100         Mali                       14535
75         India                       13384
35         China                       13099
56      Ethiopia                        8603
127     Pakistan                        8452
24        Brazil                        7648
155      Somalia                        7476
104       Mexico                        6745
168     Tanzania                        6613
      Country Name  Neoplasms
35             China    2716551
75             India     932587
182   United States     773895
83             Japan     442474
140           Russia     296822
63           Germany     274088
24            Brazil     266014
```

```
76        Indonesia     229524
59           France     197313
181  United Kingdom     195999
      Country Name  Fire, heat
75           India       25876
35           China       11096
140         Russia        6599
120        Nigeria        4017
182  United States        3616
127       Pakistan        2778
179        Ukraine        2198
156   South Africa        2179
56         Ethiopia        2101
77             Iran        1763
    Country Name  Drowning
35         China     56524
75         India     54046
83         Japan      7874
13    Bangladesh      7581
140       Russia      7086
24        Brazil      6110
186      Vietnam      6098
76     Indonesia      5011
134  Philippines      4792
169     Thailand      4519
       Country Name  Drug use disorders
182   United States               65717
35            China               11445
75            India                8465
140          Russia                5877
77             Iran                2918
181  United Kingdom                2654
31           Canada                2526
54          England                1914
59           France                1839
63          Germany                1678
                    Country Name  Road injuries
35                         China         250025
75                         India         211975
24                        Brazil          44529
182                 United States          41362
76                      Indonesia          37004
46   Democratic Republic of Congo          29543
52                         Egypt          29490
186                       Vietnam          24153
147                  Saudi Arabia          21316
77                           Iran          21122
      Country Name  Environmental heat and cold exposure
75           India                                11348
140         Russia                                 9104
35           China                                 4966
179        Ukraine                                 2733
83           Japan                                 2134
59          France                                 1875
182  United States                                 1378
120        Nigeria                                  927
85       Kazakhstan                                 741
56         Ethiopia                                 694
      Country Name  Self-harm
75           India     195336
```

```
35           China     121217
182  United States      45349
140         Russia      39040
83           Japan      24773
127       Pakistan      17688
157    South Korea      14575
179        Ukraine      13690
24          Brazil      13503
63         Germany      11664
```

In [17]:
```
print("Total null values of all columns\n",df.isnull().sum())
```

```
Total null values of all columns
 Country Name                                  0
Covid-19 Deaths                                3
Cardiovascular diseases                        0
Respiratory diseases                           0
Kidney diseases                                0
Neonatal disorders                             0
Meningitis                                     0
Malaria                                        0
Interpersonal violence                         0
HIV/AIDS                                       0
Tuberculosis                                   0
Maternal disorders                             0
Lower respiratory infections                   0
Alcohol use disorders                          0
Diarrheal diseases                             0
Poisoning                                      0
Nutritional deficiencies                       0
 Alzheimer's disease                           0
Parkinson's disease                            0
 Acute hepatitis                               0
Digestive diseases                             0
 Cirrhosis and other chronic liver diseases    0
Protein-energy malnutrition                    0
Neoplasms                                      0
Fire, heat                                     0
Drowning                                       0
Drug use disorders                             0
Road injuries                                  0
Environmental heat and cold exposure           0
Self-harm                                      0
 Conflict and terrorism                        0
Diabetes                                       0
dtype: int64
```

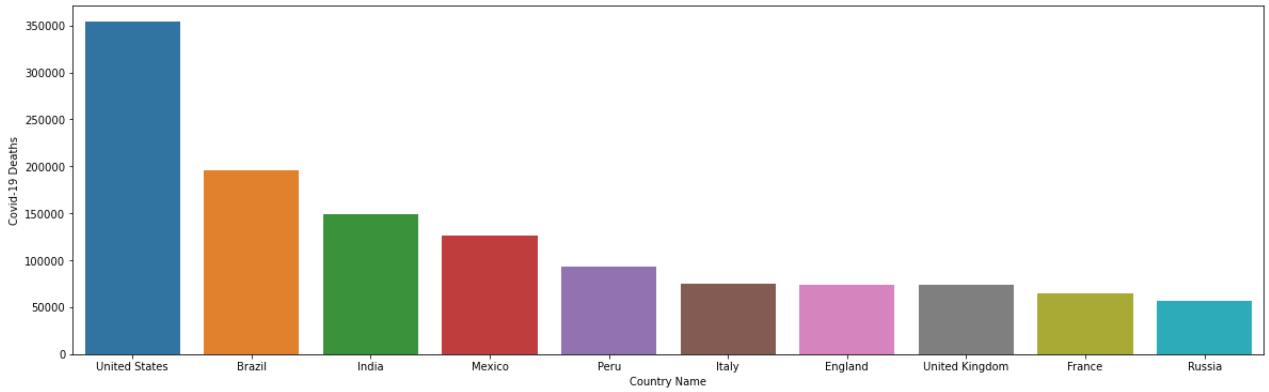
In [ ]:
```
```

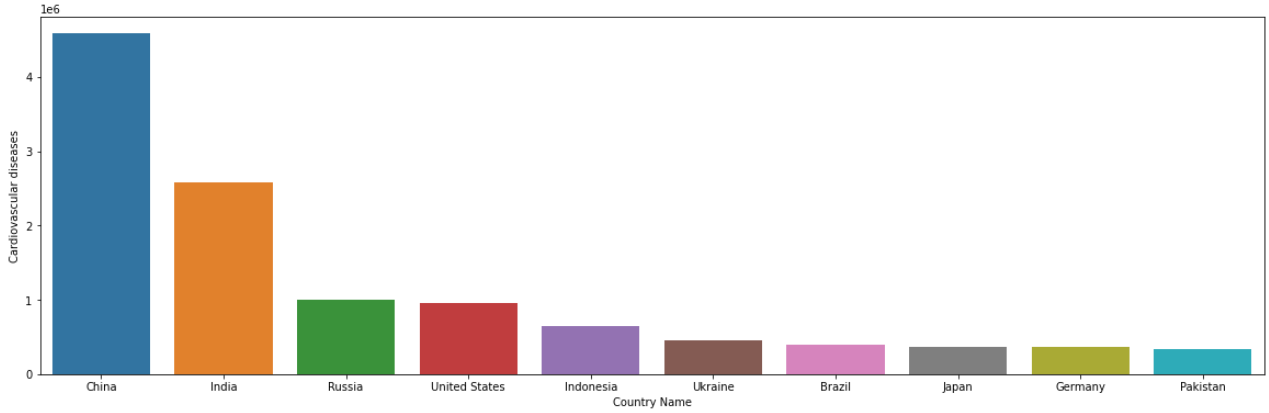
Let's make some visualizations

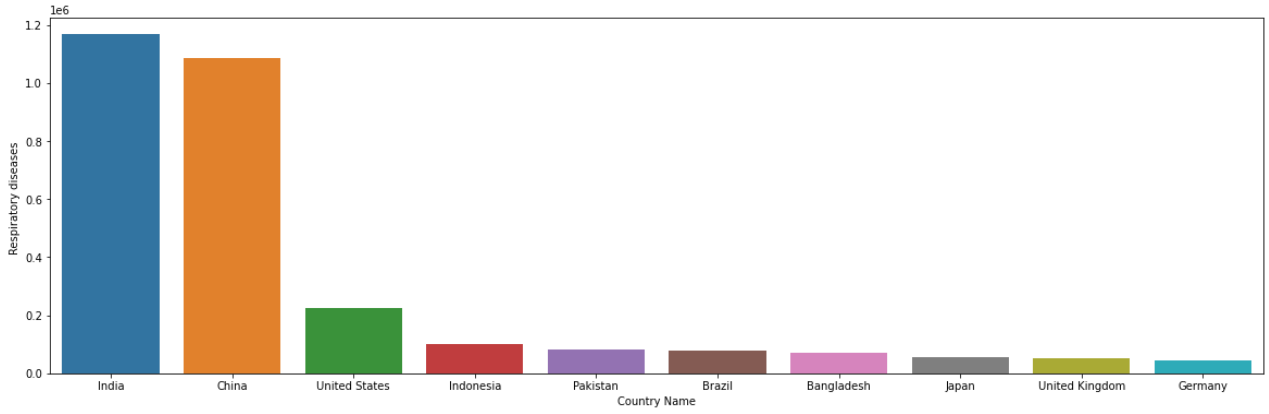
TOP 20 countries for affected mostly by each disease

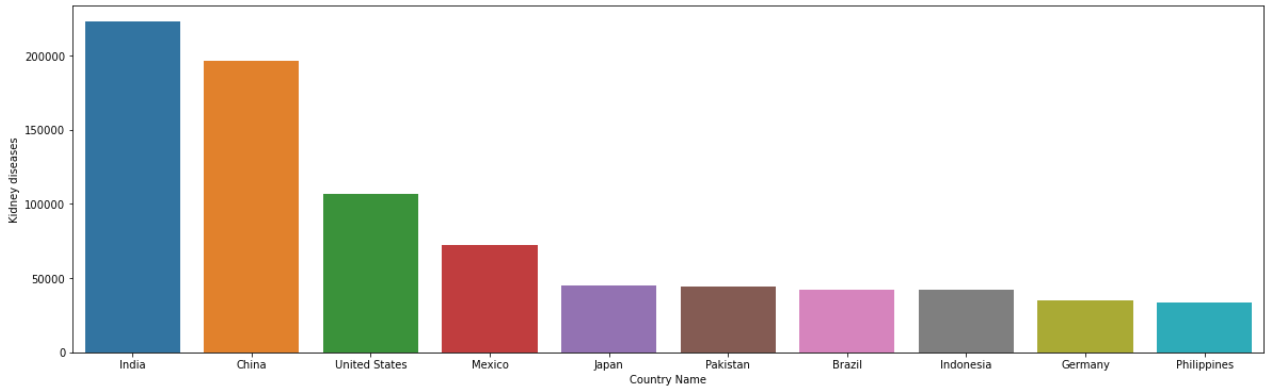
In [18]:
```
#Barplot
for cause in causes[:11]:
    f = plt.figure(figsize=(20, 6))
```

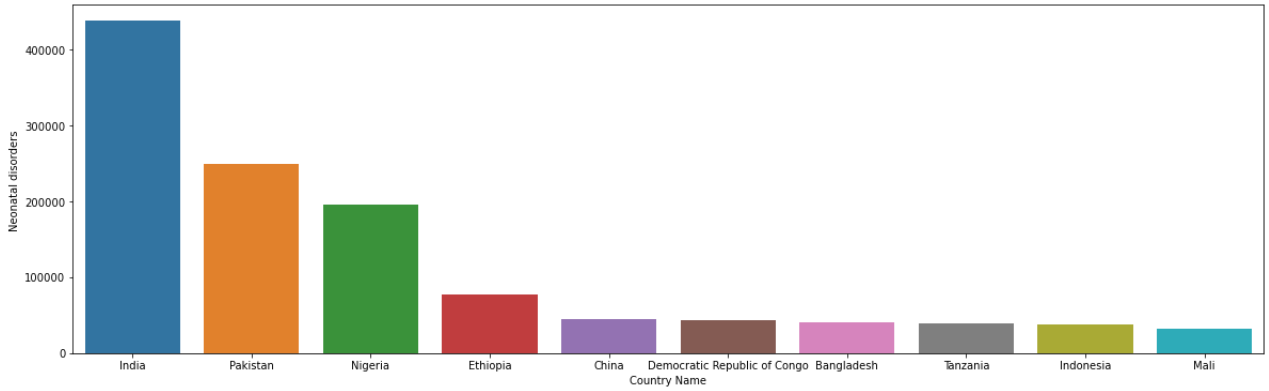
```
    sns.barplot(x = "Country Name", y = cause, data = causes_10[cause])

gs = f.add_gridspec(6, 6)
f.tight_layout()
```

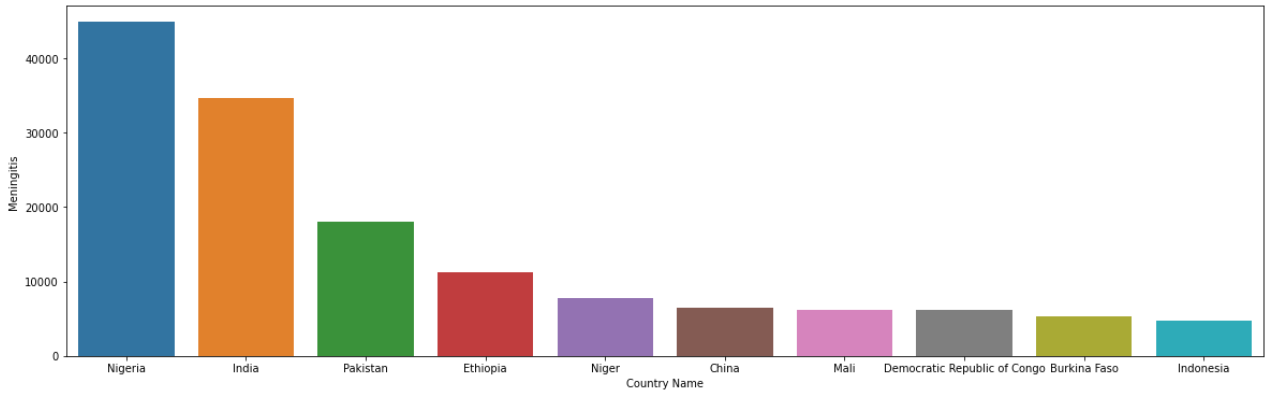


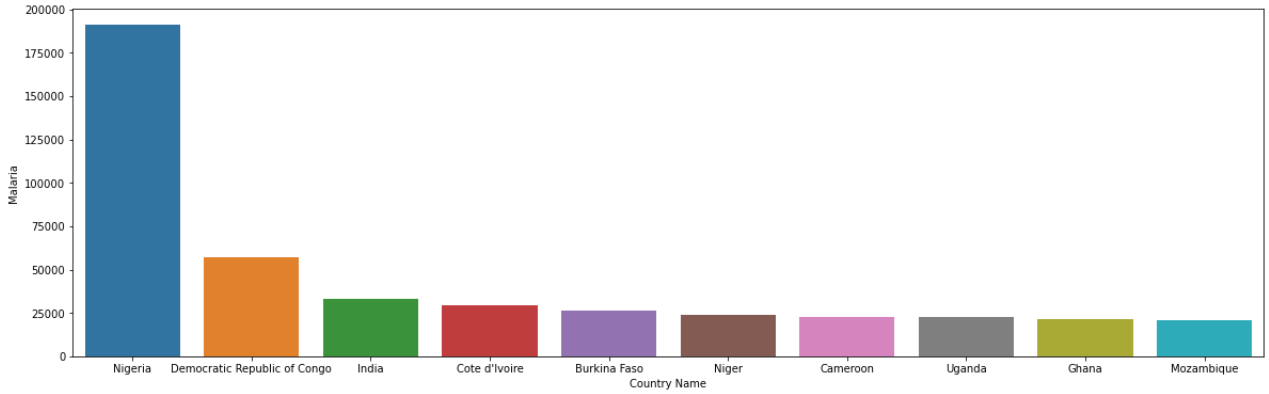


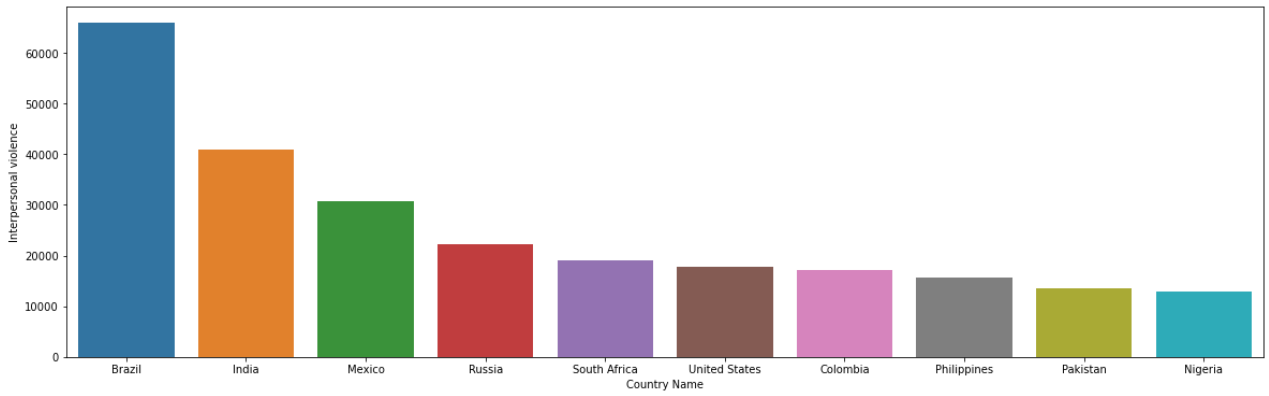


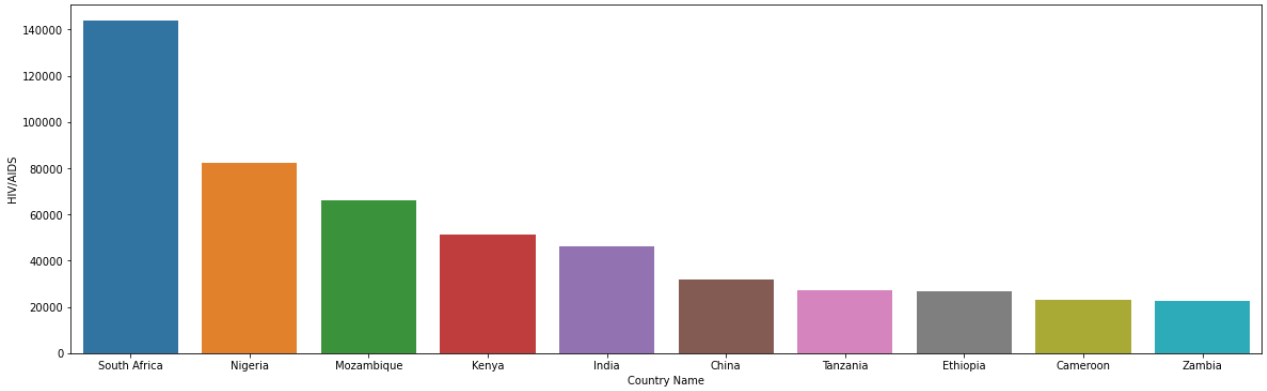


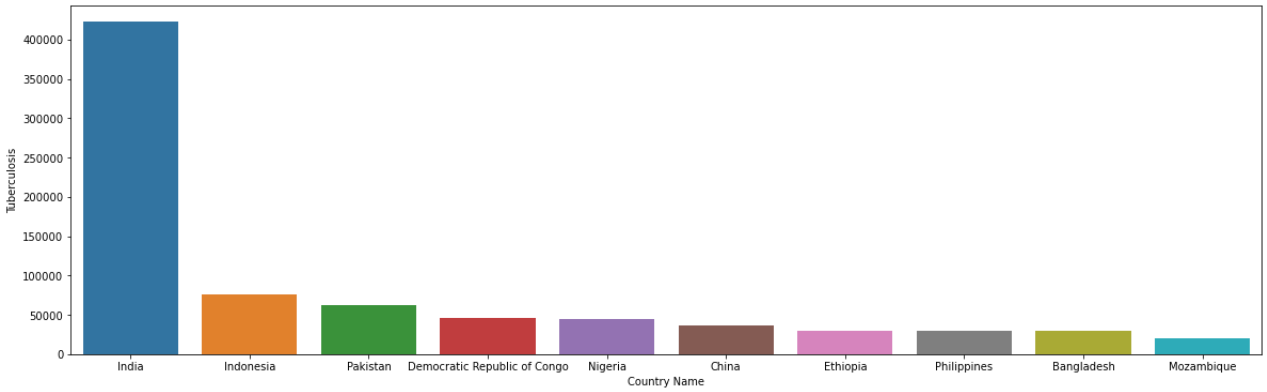

Neonatal disorders
400000
300000
200000
100000
0
India
Pakistan
Nigeria
Ethiopia
China
Democratic Republic of Congo
Bangladesh
Tanzania
Indonesia
Mali
Country Name

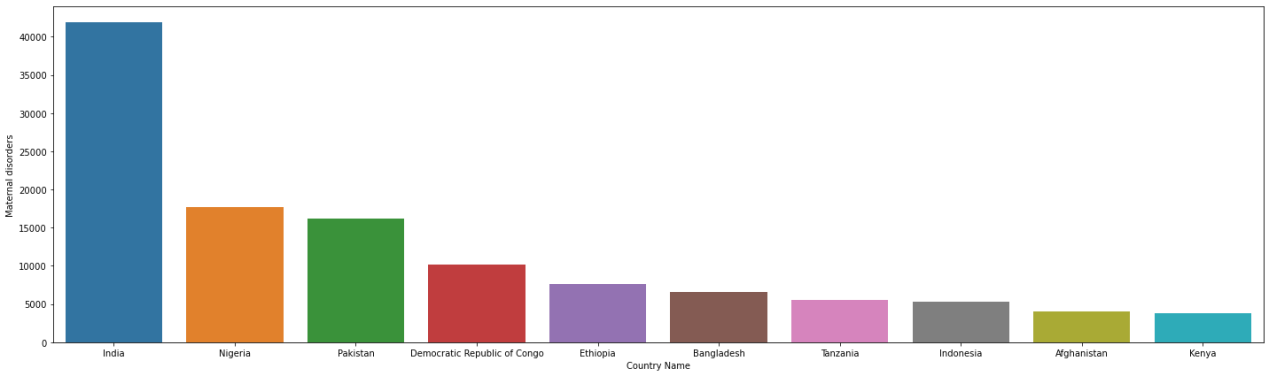

Meningitis
40000
30000
20000
10000
0
Nigeria
India
Pakistan
Ethiopia
Niger
China
Mali
Democratic Republic of Congo
Burkina Faso
Indonesia
Country Name


Malaria
200000
175000
150000
125000
100000
75000
50000
25000
0
Nigeria
Democratic Republic of Congo
India
Cote d'Ivoire
Burkina Faso
Niger
Cameroon
Uganda
Ghana
Mozambique
Country Name


Interpersonal violence
60000
50000
40000
30000
20000
10000
0
Brazil
India
Mexico
Russia
South Africa
United States
Colombia
Philippines
Pakistan
Nigeria
Country Name







In [21]:

```
words = ''
for i in list(df["Country Name"]):
        i = i.replace(" ", "_") + " "
        words+= i

wordcloud = WordCloud(background_color='white',
                      max_font_size = 25,
                      random_state = 55,
                      mask = np.array(Image.open('loc.png'))
                       )

wordcloud.generate(str(words))
plt.figure(figsize=(15,12))
plt.title("Countries in the world",fontsize= 30, x=0.5, y=1.1)
plt.axis('off');
plt.tight_layout()
plt.imshow(wordcloud);
```

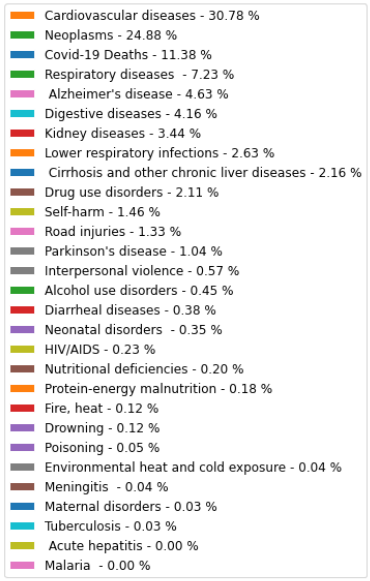
Countries in the world

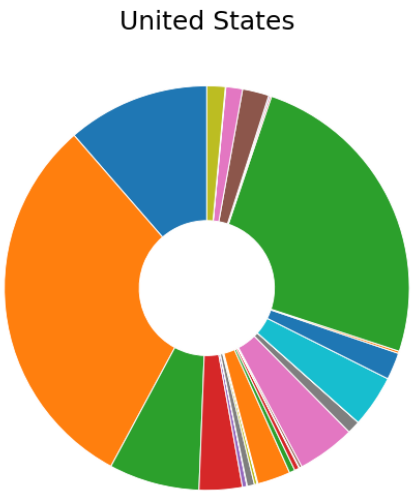
Let's show the top 10 countries that have the greatest number of deaths due to given disease.

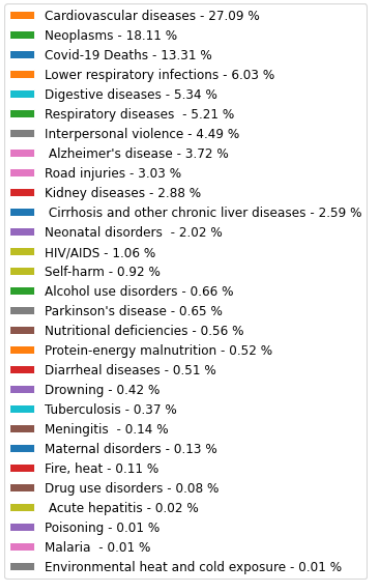
In [22]:

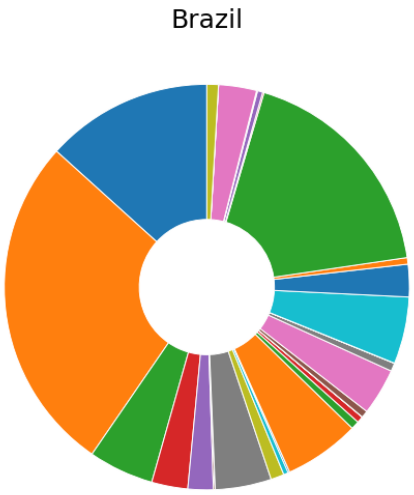
```
def my_autopct(pct):
    return ('%.2f' % pct) if pct > 5 else ''


covid_10 = list(causes_10["Covid-19 Deaths"]["Country Name"])

for country in covid_10:

    causes_death = []
    value = []
    causes_death = list(df[df["Country Name"] == country].columns)
    causes_death = causes_death[1: -2]
```

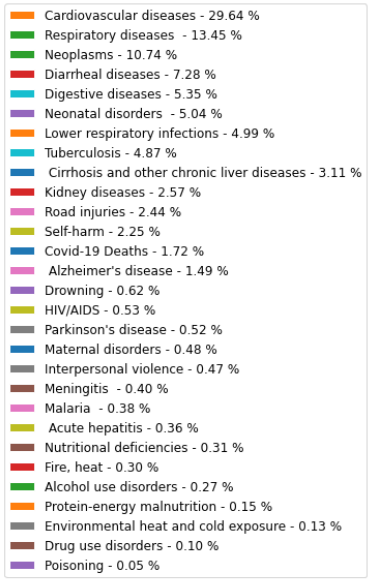
```
    value = df[df["Country Name"] == country].values.tolist()
    value = value[0][1: -2]
    number = value
    per_val = np.array(number)
    percent = 100.*per_val/per_val.sum()

    size = 0.4
    plt.figure(figsize=(15, 15))
    patches, texts = plt.pie(number,startangle=90, radius=1 - size, wedgeprops=dict(wid
    labels = ['{0} - {1:1.2f} %'.format(i,j) for i,j in zip(causes_death, percent)]

    sort_legend = True
    if sort_legend:
        patches, labels, dummy =  zip(*sorted(zip(patches, labels, number),
                                              key=lambda causes_death: causes_death[2],
                                              reverse=True))

    plt.legend(patches, labels, loc='best', bbox_to_anchor=(-0.05, 1), fontsize=12)
    plt.rcParams['axes.titley'] = 0.8
    plt.title(country, fontsize=25)


    plt.show()
```

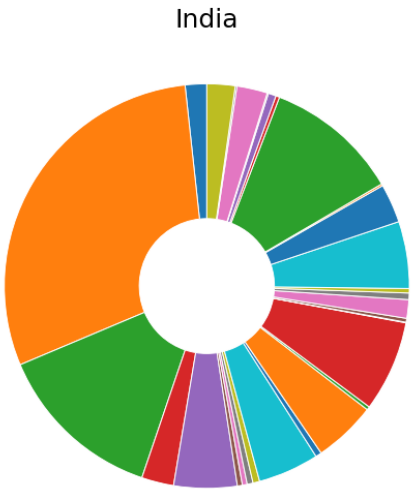


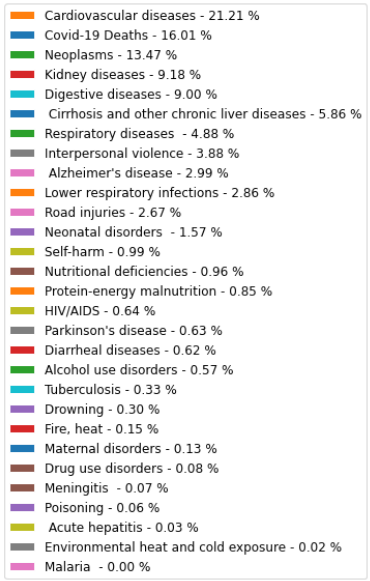


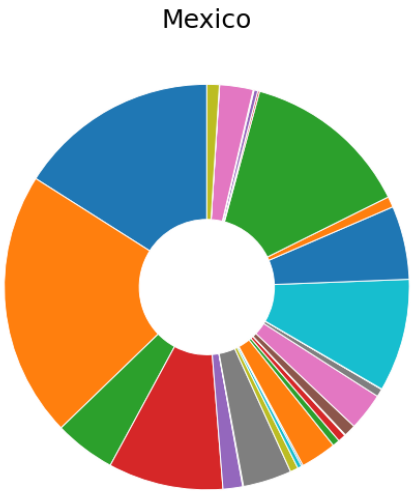

Cardiovascular diseases - 27.09 %
Neoplasms - 18.11 %
Covid-19 Deaths - 13.31 %
Lower respiratory infections - 6.03 %
Digestive diseases - 5.34 %
Respiratory diseases - 5.21 %
Interpersonal violence - 4.49 %
Alzheimer's disease - 3.72 %
Road injuries - 3.03 %
Kidney diseases - 2.88 %
Cirrhosis and other chronic liver diseases - 2.59 %
Neonatal disorders - 2.02 %
HIV/AIDS - 1.06 %
Self-harm - 0.92 %
Alcohol use disorders - 0.66 %
Parkinson's disease - 0.65 %
Nutritional deficiencies - 0.56 %
Protein-energy malnutrition - 0.52 %
Diarrheal diseases - 0.51 %
Drowning - 0.42 %
Tuberculosis - 0.37 %
Meningitis - 0.14 %
Maternal disorders - 0.13 %
Fire, heat - 0.11 %
Drug use disorders - 0.08 %
Acute hepatitis - 0.02 %
Poisoning - 0.01 %
Malaria - 0.01 %
Environmental heat and cold exposure - 0.01 %

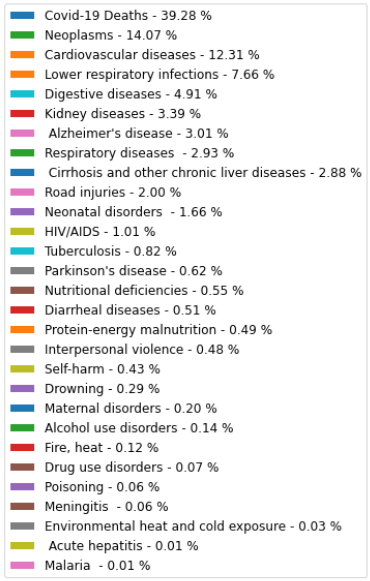

Brazil

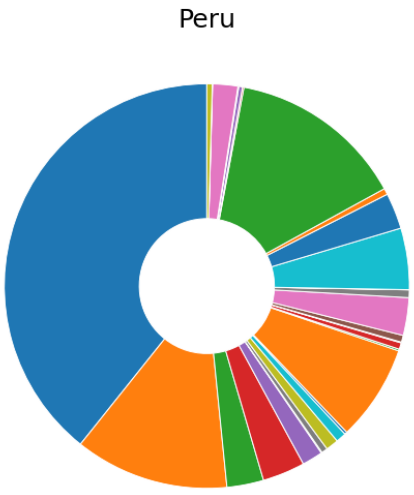

Cardiovascular diseases - 29.64 %
Respiratory diseases - 13.45 %
Neoplasms - 10.74 %
Diarrheal diseases - 7.28 %
Digestive diseases - 5.35 %
Neonatal disorders - 5.04 %
Lower respiratory infections - 4.99 %
Tuberculosis - 4.87 %
Cirrhosis and other chronic liver diseases - 3.11 %
Kidney diseases - 2.57 %
Road injuries - 2.44 %
Self-harm - 2.25 %
Covid-19 Deaths - 1.72 %
Alzheimer's disease - 1.49 %
Drowning - 0.62 %
HIV/AIDS - 0.53 %
Parkinson's disease - 0.52 %
Maternal disorders - 0.48 %
Interpersonal violence - 0.47 %
Meningitis - 0.40 %
Malaria - 0.38 %
Acute hepatitis - 0.36 %
Nutritional deficiencies - 0.31 %
Fire, heat - 0.30 %
Alcohol use disorders - 0.27 %
Protein-energy malnutrition - 0.15 %
Environmental heat and cold exposure - 0.13 %
Drug use disorders - 0.10 %
Poisoning - 0.05 %

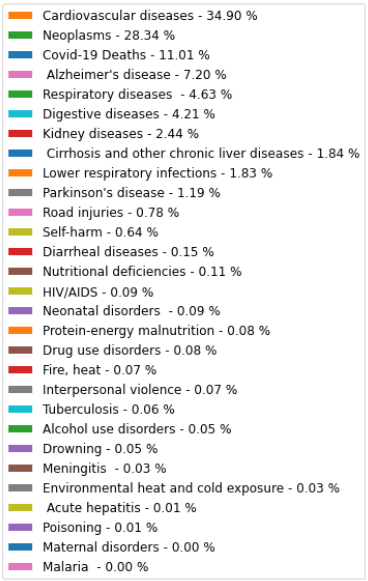

India

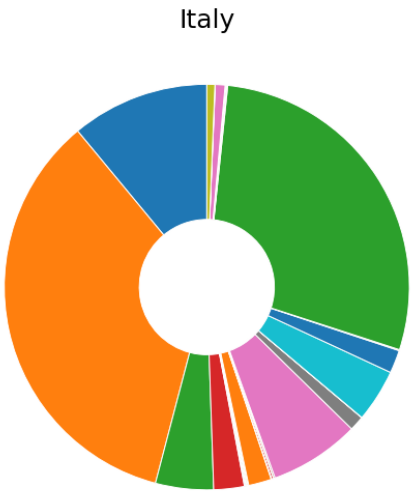

Cardiovascular diseases - 21.21 %
Covid-19 Deaths - 16.01 %
Neoplasms - 13.47 %
Kidney diseases - 9.18 %
Digestive diseases - 9.00 %
Cirrhosis and other chronic liver diseases - 5.86 %
Respiratory diseases - 4.88 %
Interpersonal violence - 3.88 %
Alzheimer's disease - 2.99 %
Lower respiratory infections - 2.86 %
Road injuries - 2.67 %
Neonatal disorders - 1.57 %
Self-harm - 0.99 %
Nutritional deficiencies - 0.96 %
Protein-energy malnutrition - 0.85 %
HIV/AIDS - 0.64 %
Parkinson's disease - 0.63 %
Diarrheal diseases - 0.62 %
Alcohol use disorders - 0.57 %
Tuberculosis - 0.33 %
Drowning - 0.30 %
Fire, heat - 0.15 %
Maternal disorders - 0.13 %
Drug use disorders - 0.08 %
Meningitis - 0.07 %
Poisoning - 0.06 %
Acute hepatitis - 0.03 %
Environmental heat and cold exposure - 0.02 %
Malaria - 0.00 %

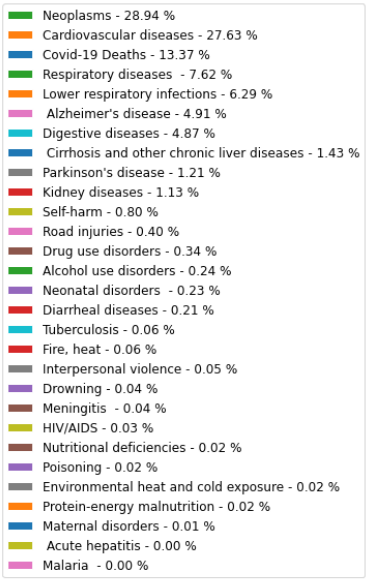

Mexico

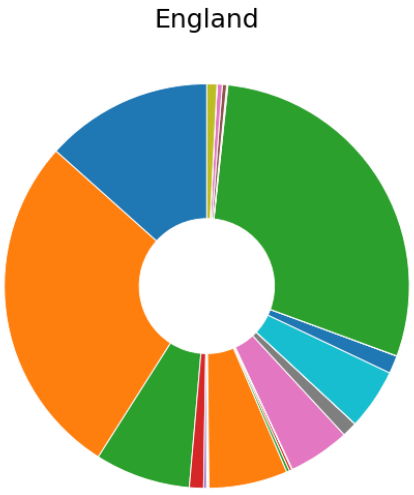

Covid-19 Deaths - 39.28 %
Neoplasms - 14.07 %
Cardiovascular diseases - 12.31 %
Lower respiratory infections - 7.66 %
Digestive diseases - 4.91 %
Kidney diseases - 3.39 %
Alzheimer's disease - 3.01 %
Respiratory diseases - 2.93 %
Cirrhosis and other chronic liver diseases - 2.88 %
Road injuries - 2.00 %
Neonatal disorders - 1.66 %
HIV/AIDS - 1.01 %
Tuberculosis - 0.82 %
Parkinson's disease - 0.62 %
Nutritional deficiencies - 0.55 %
Diarrheal diseases - 0.51 %
Protein-energy malnutrition - 0.49 %
Interpersonal violence - 0.48 %
Self-harm - 0.43 %
Drowning - 0.29 %
Maternal disorders - 0.20 %
Alcohol use disorders - 0.14 %
Fire, heat - 0.12 %
Drug use disorders - 0.07 %
Poisoning - 0.06 %
Meningitis - 0.06 %
Environmental heat and cold exposure - 0.03 %
Acute hepatitis - 0.01 %
Malaria - 0.01 %

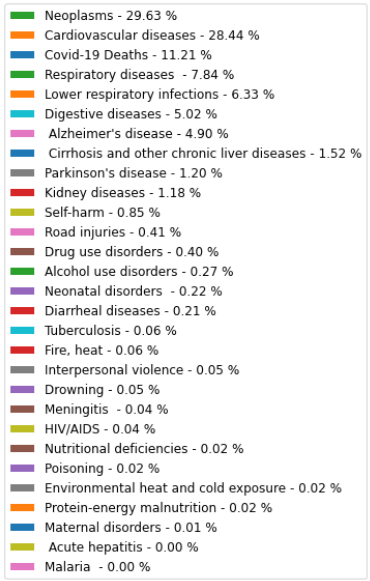

Peru

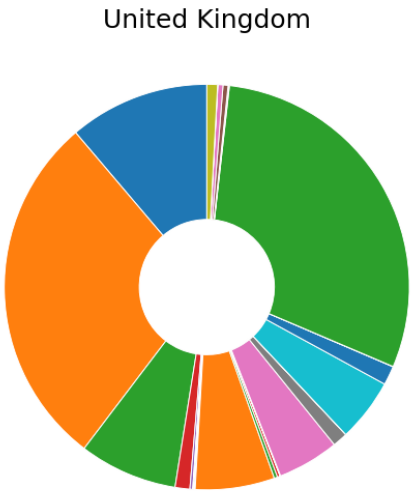

Cardiovascular diseases - 34.90 %
Neoplasms - 28.34 %
Covid-19 Deaths - 11.01 %
Alzheimer's disease - 7.20 %
Respiratory diseases - 4.63 %
Digestive diseases - 4.21 %
Kidney diseases - 2.44 %
Cirrhosis and other chronic liver diseases - 1.84 %
Lower respiratory infections - 1.83 %
Parkinson's disease - 1.19 %
Road injuries - 0.78 %
Self-harm - 0.64 %
Diarrheal diseases - 0.15 %
Nutritional deficiencies - 0.11 %
HIV/AIDS - 0.09 %
Neonatal disorders - 0.09 %
Protein-energy malnutrition - 0.08 %
Drug use disorders - 0.08 %
Fire, heat - 0.07 %
Interpersonal violence - 0.07 %
Tuberculosis - 0.06 %
Alcohol use disorders - 0.05 %
Drowning - 0.05 %
Meningitis - 0.03 %
Environmental heat and cold exposure - 0.03 %
Acute hepatitis - 0.01 %
Poisoning - 0.01 %
Maternal disorders - 0.00 %
Malaria - 0.00 %

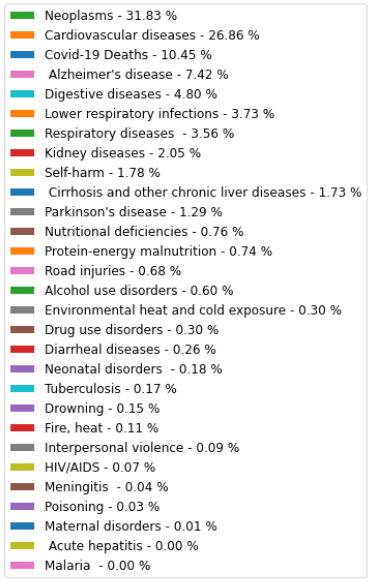

Italy

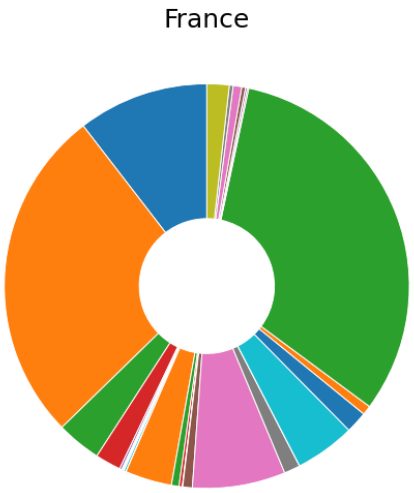

Neoplasms - 28.94 %
Cardiovascular diseases - 27.63 %
Covid-19 Deaths - 13.37 %
Respiratory diseases - 7.62 %
Lower respiratory infections - 6.29 %
Alzheimer's disease - 4.91 %
Digestive diseases - 4.87 %
Cirrhosis and other chronic liver diseases - 1.43 %
Parkinson's disease - 1.21 %
Kidney diseases - 1.13 %
Self-harm - 0.80 %
Road injuries - 0.40 %
Drug use disorders - 0.34 %
Alcohol use disorders - 0.24 %
Neonatal disorders - 0.23 %
Diarrheal diseases - 0.21 %
Tuberculosis - 0.06 %
Fire, heat - 0.06 %
Interpersonal violence - 0.05 %
Drowning - 0.04 %
Meningitis - 0.04 %
HIV/AIDS - 0.03 %
Nutritional deficiencies - 0.02 %
Poisoning - 0.02 %
Environmental heat and cold exposure - 0.02 %
Protein-energy malnutrition - 0.02 %
Maternal disorders - 0.01 %
Acute hepatitis - 0.00 %
Malaria - 0.00 %

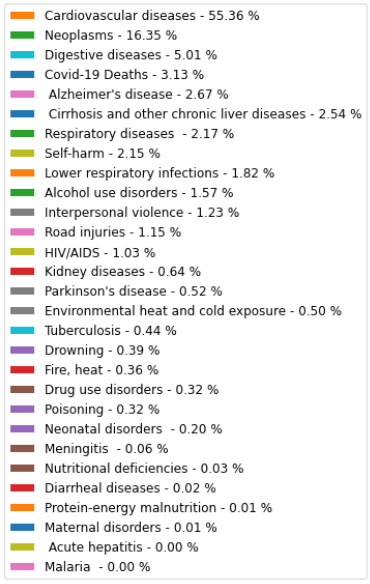

England

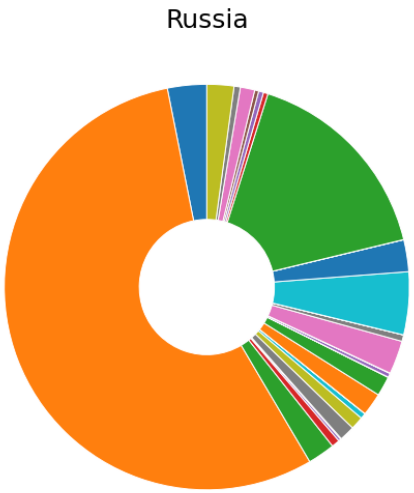

Neoplasms - 29.63 %
Cardiovascular diseases - 28.44 %
Covid-19 Deaths - 11.21 %
Respiratory diseases - 7.84 %
Lower respiratory infections - 6.33 %
Digestive diseases - 5.02 %
Alzheimer's disease - 4.90 %
Cirrhosis and other chronic liver diseases - 1.52 %
Parkinson's disease - 1.20 %
Kidney diseases - 1.18 %
Self-harm - 0.85 %
Road injuries - 0.41 %
Drug use disorders - 0.40 %
Alcohol use disorders - 0.27 %
Neonatal disorders - 0.22 %
Diarrheal diseases - 0.21 %
Tuberculosis - 0.06 %
Fire, heat - 0.06 %
Interpersonal violence - 0.05 %
Drowning - 0.05 %
Meningitis - 0.04 %
HIV/AIDS - 0.04 %
Nutritional deficiencies - 0.02 %
Poisoning - 0.02 %
Environmental heat and cold exposure - 0.02 %
Protein-energy malnutrition - 0.02 %
Maternal disorders - 0.01 %
Acute hepatitis - 0.00 %
Malaria - 0.00 %


United Kingdom


Neoplasms - 31.83 %
Cardiovascular diseases - 26.86 %
Covid-19 Deaths - 10.45 %
Alzheimer's disease - 7.42 %
Digestive diseases - 4.80 %
Lower respiratory infections - 3.73 %
Respiratory diseases - 3.56 %
Kidney diseases - 2.05 %
Self-harm - 1.78 %
Cirrhosis and other chronic liver diseases - 1.73 %
Parkinson's disease - 1.29 %
Nutritional deficiencies - 0.76 %
Protein-energy malnutrition - 0.74 %
Road injuries - 0.68 %
Alcohol use disorders - 0.60 %
Environmental heat and cold exposure - 0.30 %
Drug use disorders - 0.30 %
Diarrheal diseases - 0.26 %
Neonatal disorders - 0.18 %
Tuberculosis - 0.17 %
Drowning - 0.15 %
Fire, heat - 0.11 %
Interpersonal violence - 0.09 %
HIV/AIDS - 0.07 %
Meningitis - 0.04 %
Poisoning - 0.03 %
Maternal disorders - 0.01 %
Acute hepatitis - 0.00 %
Malaria - 0.00 %

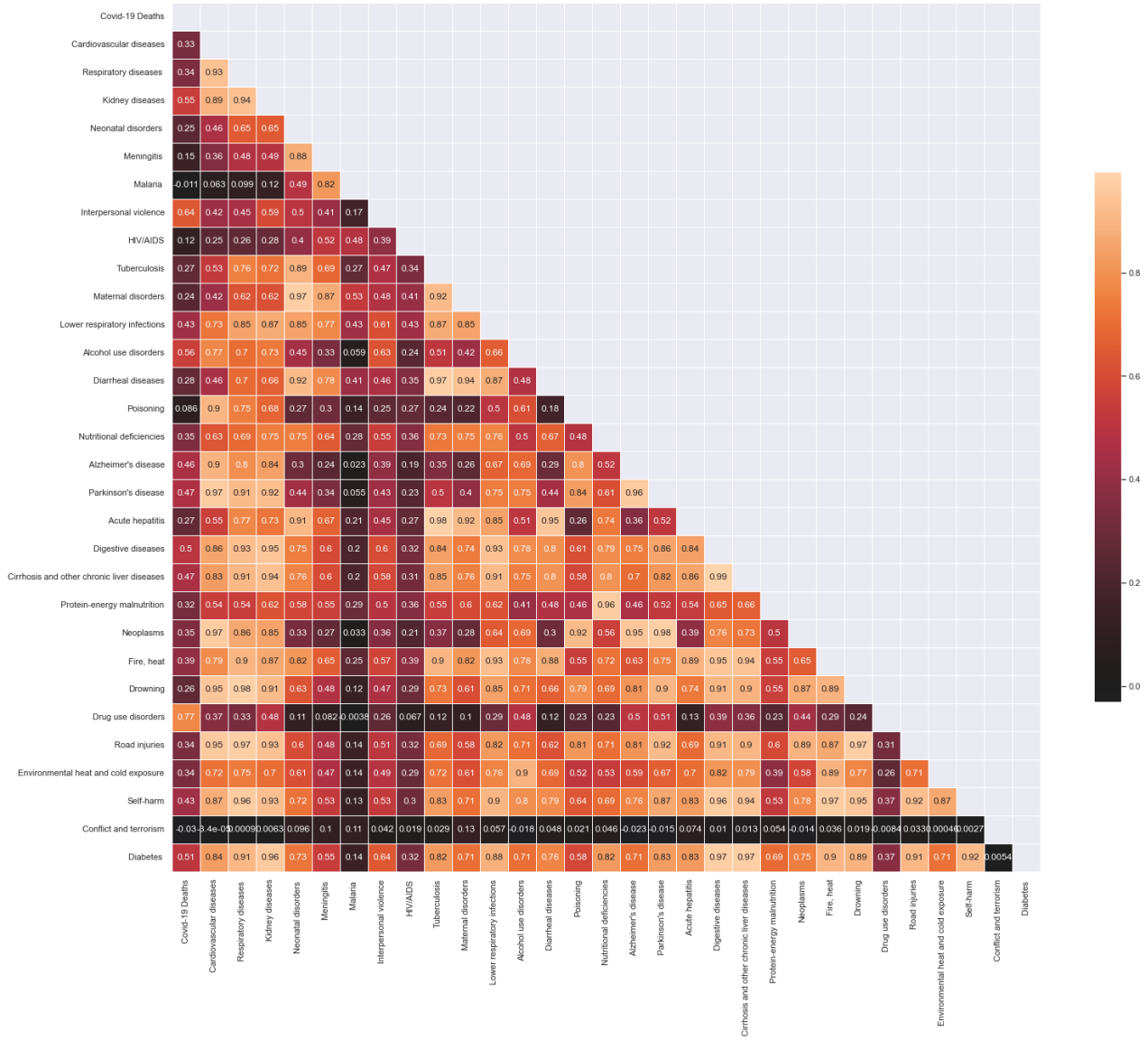

France





In [23]:
```
sns.set_theme(style="darkgrid")

corr = df.corr()
mask = np.triu(np.ones_like(corr, dtype=bool))

f, ax = plt.subplots(figsize=(25, 25))
cmap = sns.diverging_palette(230, 20)

sns.heatmap(corr, mask=mask, center=0, annot=True,
            square=True, linewidths=.5,cbar_kws={"shrink": 0.5})
```

Out[23]: <AxesSubplot:>



In [24]:

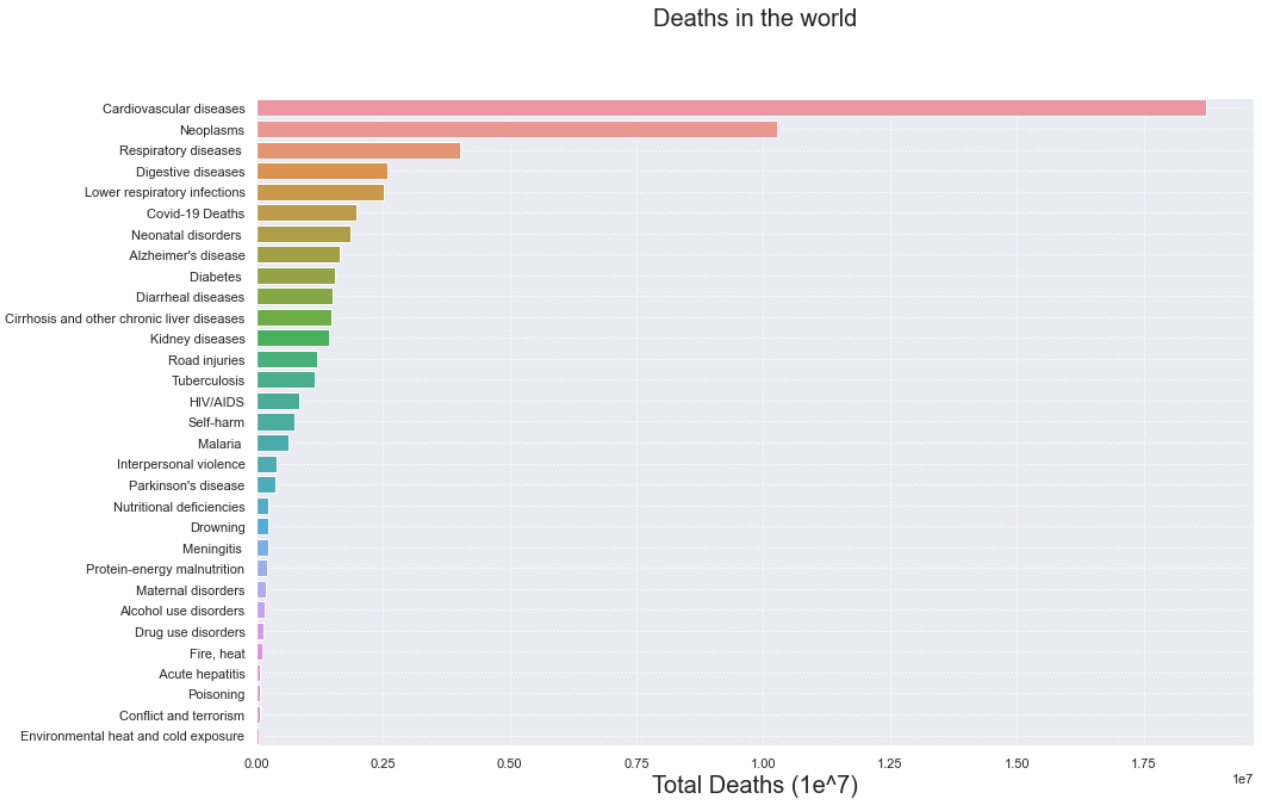
```
countries= df.sum(axis=0)[1:].sort_values(ascending=True).index
total_death = df.sum(axis=0)[1:].sort_values(ascending=True).values

print("\033[1m"+"Causes of death (sorted ascending) "+"\033[0m")
for i in range(0,len(countries)):
    print(countries[i],"=",f'{total_death[i]:,}')

countries= df.sum(axis=0)[1:].sort_values(ascending=False).index
total_death = df.sum(axis=0)[1:].sort_values(ascending=False).values

print("\033[1m"+"\nTotal death ="+"\n",
     f'{ np.sum(df.sum(axis=0)[1:].sort_values(ascending=False).values) :,}')
```

Causes of death (sorted ascending)
Environmental heat and cold exposure = 47,381
Conflict and terrorism = 62,594
Poisoning = 76,684
Acute hepatitis = 78,711
Fire, heat = 110,886
Drug use disorders = 130,660
Alcohol use disorders = 169,466
Maternal disorders = 192,852

```
Protein-energy malnutrition = 208,395
Meningitis  = 232,559
Drowning = 236,366
Nutritional deficiencies = 247,638
Parkinson's disease = 370,490
Interpersonal violence = 412,055
Malaria  = 625,532
Self-harm = 761,569
HIV/AIDS = 836,303
Tuberculosis = 1,159,685
Road injuries = 1,193,057
Kidney diseases = 1,430,726
 Cirrhosis and other chronic liver diseases = 1,477,440
Diarrheal diseases = 1,515,372
Diabetes  = 1,551,054
 Alzheimer's disease = 1,655,193
Neonatal disorders  = 1,864,067
Covid-19 Deaths = 1,966,744.0
Lower respiratory infections = 2,511,920
Digestive diseases = 2,581,914
Respiratory diseases  = 4,019,102
Neoplasms = 10,262,088
Cardiovascular diseases = 18,735,068

Total death =
 56,723,571.0
```

In [25]:

```
fig, ax = plt.subplots(figsize =(15, 10))
sns.barplot(total_death,countries)
plt.grid(linestyle='--', linewidth=.6)
plt.title("Deaths in the world",fontsize=20, x=0.5, y=1.1)
plt.xlabel("Total Deaths (1e^7)",fontsize=20)
plt.show()
```



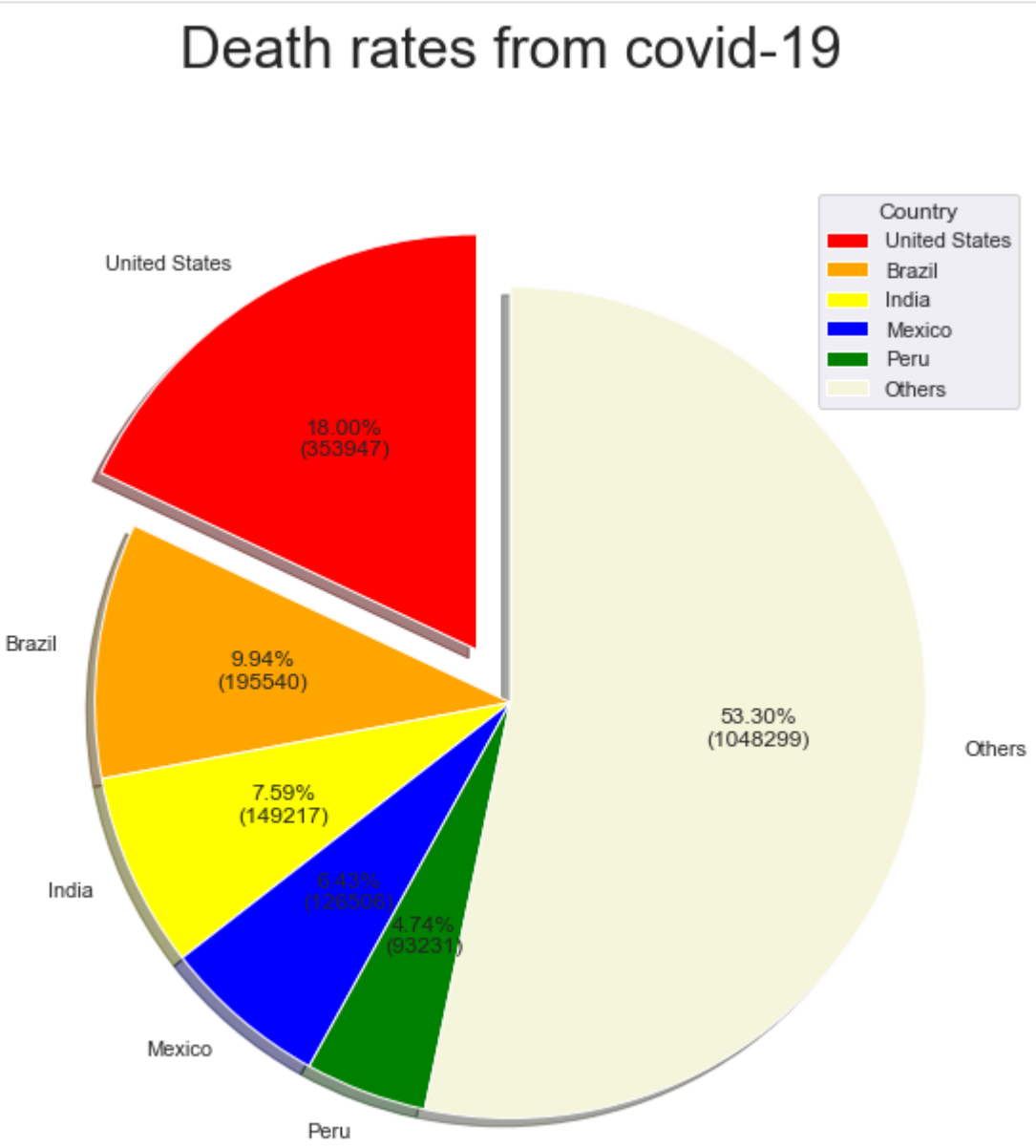
In [26]:
```
countries = df.loc[ df["Covid-19 Deaths"].sort_values(ascending=True)[:5].index]["Count
total_death = df["Covid-19 Deaths"].sort_values(ascending=True)[:5].values

print("\033[1m"+"5 countries with the lowest death rates from covid-19"+"\033[0m")
for i in range(0,len(countries)):
    print(countries[i],"=",f'{total_death[i]:,}')

countries = df.loc[ df["Covid-19 Deaths"].sort_values(ascending=False)[:5].index]["Coun
total_death = df["Covid-19 Deaths"].sort_values(ascending=False)[:5].values

print("\033[1m"+"\n5 countries with the highest death rates from covid-19"+"\033[0m")
for i in range(0,len(countries)):
    print(countries[i],"=",f'{total_death[i]:,}')

print("\033[1m"+"\nTotal death toll from covid-19 ="+"\033[0m",
     f'{ df["Covid-19 Deaths"].sort_values(ascending=False).sum() :,}')
```

5 countries with the lowest death rates from covid-19
Mongolia = 1.0
Fiji = 2.0
Dominica = 2.0
Solomon Islands = 2.0
Micronesia = 2.0

5 countries with the highest death rates from covid-19
United States = 353,948.0
Brazil = 195,541.0
India = 149,218.0
Mexico = 126,507.0
Peru = 93,231.0

Total death toll from covid-19 = 1,966,744.0

In [27]:
```
all_others_covid = df["Covid-19 Deaths"].sort_values(ascending=False)[5:].sum()
total_death = np.append(total_death,all_others_covid)
countries = np.append(countries,"Others")

colors = ( "red", "orange", "yellow",
          "blue", "green", "beige")

wp = { 'linewidth' : 1, 'edgecolor' : "green" }

def func(pct, allvalues):
    absolute = int(pct / 100.*np.sum(allvalues))
    return "{:.2f}%\n({:d})".format(pct, absolute)

fig, ax = plt.subplots(figsize =(10, 10))

wedges, texts, autotexts = ax.pie(total_death,
                                  autopct = lambda pct: func(pct, total_death),
                                  labels = countries ,
                                   explode = [0.15, 0, 0, 0, 0, 0],
                                  shadow = True,
                                  colors = colors,
                                  startangle = 90,)

plt.title("Death rates from covid-19",fontsize= 30, x=0.5, y=1.1)
plt.legend(title = "Country",loc ="upper right")
plt.show()
```



In [28]:

```
countries = df.loc[ df["Cardiovascular diseases"].sort_values(ascending=True)[:5].index
total_death = df["Cardiovascular diseases"].sort_values(ascending=True)[:5].values

print("\033[1m"+"5 countries with the lowest death rates from Cardiovascular diseases"+
for i in range(0,len(countries)):
    print(countries[i],"=",total_death[i])

countries = df.loc[ df["Cardiovascular diseases"].sort_values(ascending=False)[:5].inde
total_death = df["Cardiovascular diseases"].sort_values(ascending=False)[:5].values

print("\033[1m"+"\n5 countries with the highest death rates from Cardiovascular disease
for i in range(0,len(countries)):
    print(countries[i],"=",f'{total_death[i]:,}')

print("\033[1m"+"\nTotal death toll from Cardiovascular diseases ="+"\033[0m",
      f'{ df["Cardiovascular diseases"].sort_values(ascending=False).sum() :,}')
```

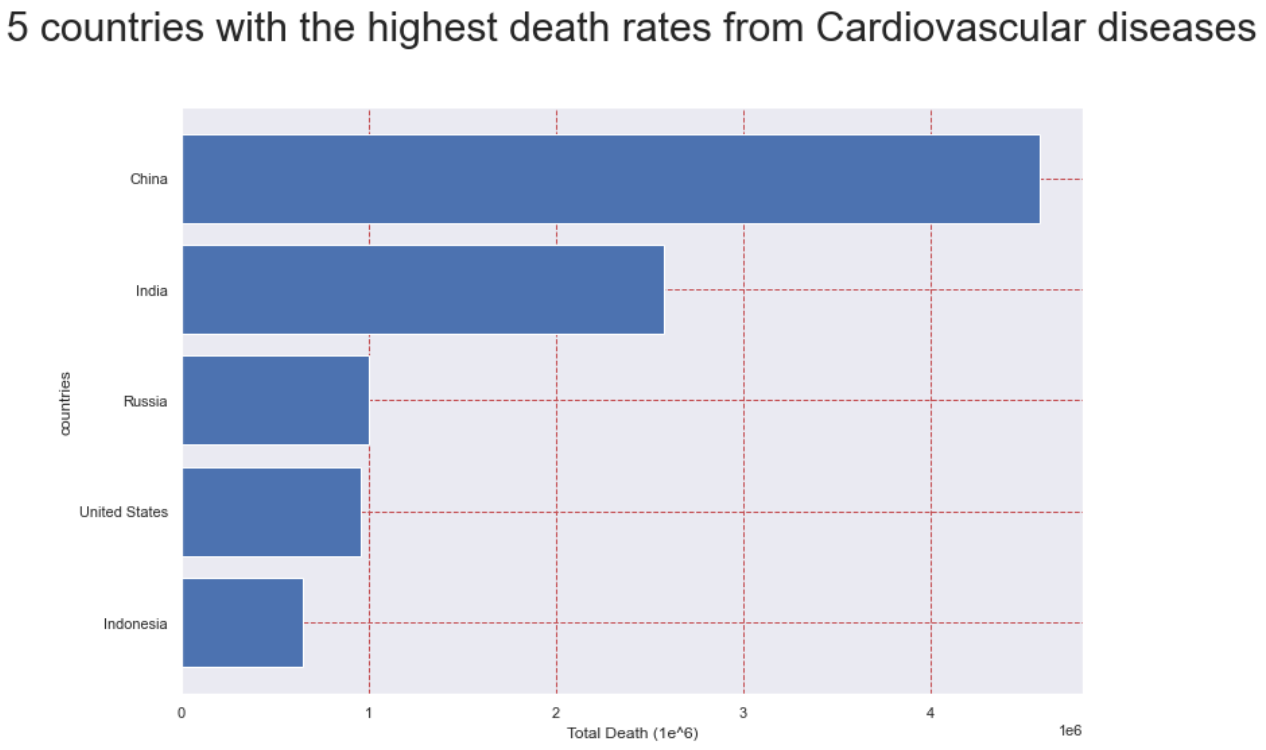
5 countries with the lowest death rates from Cardiovascular diseases

Niue = 8
Nauru = 23
Tuvalu = 42
Cook Islands = 58
Palau = 65

5 countries with the highest death rates from Cardiovascular diseases
China = 4,584,273
India = 2,574,410
Russia = 1,004,931
United States = 957,455
Indonesia = 651,481

Total death toll from Cardiovascular diseases = 18,735,068

In [29]:
```
fig, ax = plt.subplots(figsize =(12,8))

plt.grid(color='r', linestyle='--', linewidth=1)
plt.barh(countries, total_death )

plt.title("5 countries with the highest death rates from Cardiovascular diseases",fonts
plt.ylabel('countries')
plt.xlabel('Total Death (1e^6)')
ax.invert_yaxis()

plt.show()
```



In [ ]: