Suicide Data Analysis

By Charles Zheng

Introduction

Suicide is a sensitive topic to discuss, but it is important to acknowledge suicide as a serious public health problem. To better understand suicide and its possible factors, looking at historical data can be incredibly insightful. This project aims to examine key trends among a global dataset spanning 30 years.

Data Source

The data used from this project comes from the dataset Suicide Rates Overview 1985 to 2016 on Kaggle, which largely consists of data collected by WHO.

Importing and Cleaning Data

The data is stored in a csv file, which is uploaded to the Github repository. Packages and the file are imported below.

In [1]:
```
import numpy as np
import pandas as pd
import matplotlib as mpl
import matplotlib.pyplot as plt
import matplotlib.ticker as ticker
from mpl_toolkits.axes_grid1 import host_subplot
from mpl_toolkits import axisartist #
import seaborn as sns

%matplotlib inline

url = "C:\\Users\\charl\\Data_Bootcamp\\master.csv"
```

In [2]:
```
df = pd.read_csv(url)
df = df.replace(to_replace = "//N", value = np.nan) #replace all null values with NaN
df.columns = df.columns.str.capitalize() #capitalize column names
df
```

Out[2]:

	Country	Year	Sex	Age	Suicides_no	Population	Suicides/100k pop	Country-year	Hdi for year
0	Albania	1987	male	15-24 years	21	312900	6.71	Albania1987	NaN

	Country	Year	Sex	Age	Suicides_no	Population	Suicides/100k pop	Country-year	Hdi for year	
1	Albania	1987	male	35-54 years	16	308000	5.19	Albania1987	NaN	
2	Albania	1987	female	15-24 years	14	289700	4.83	Albania1987	NaN	
3	Albania	1987	male	75+ years	1	21800	4.59	Albania1987	NaN	
4	Albania	1987	male	25-34 years	9	274300	3.28	Albania1987	NaN	
...	...	...	...	...	...	...	...	...	...	
27815	Uzbekistan	2014	female	35-54 years	107	3620833	2.96	Uzbekistan2014	0.675	6
27816	Uzbekistan	2014	female	75+ years	9	348465	2.58	Uzbekistan2014	0.675	6
27817	Uzbekistan	2014	male	5-14 years	60	2762158	2.17	Uzbekistan2014	0.675	6
27818	Uzbekistan	2014	female	5-14 years	44	2631600	1.67	Uzbekistan2014	0.675	6
27819	Uzbekistan	2014	female	55-74 years	21	1438935	1.46	Uzbekistan2014	0.675	6

27820 rows × 12 columns

Global Trends

I first want to take a look at the data on a global scale. I created a dataframe indexed by year and aggregated total suicides and population as well as finding the average suicide ratio.

In [3]:
```
year_list = df["Year"].unique() #adding to list all unique years
year_list.sort() #sorting years chronologically
globaltotal_df = pd.DataFrame(year_list) #initializing dataframe for global totals
globaltotal_df.columns = ["Year"]
globaltotal_df = globaltotal_df.set_index(["Year"]) #setting year as index
globaltotal_df
```

Out[3]:

Year
1985
1986

Year
1987
1988
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
2016

In [4]:
```
suicide_list = [] #empty list for aggregate suicide numbers
for year in year_list:
    temp_value = df[df["Year"] == year]["Suicides_no"].values.sum() #summing all suicid
    suicide_list.append(temp_value)
```

```
globaltotal_df["Total # Suicides"] = suicide_list

population_list = [] #empty list for aggregate population numbers
for year in year_list:
    temp_value = df[df["Year"] == year]["Population"].values.sum() #summing population
    population_list.append(temp_value)
globaltotal_df["Total Population"] = population_list
globaltotal_df["Suicide Ratio per 100k"] = (globaltotal_df["Total # Suicides"]*100000)/
globaltotal_df
```

Out[4]:

Year	Total # Suicides	Total Population	Suicide Ratio per 100k
1985	116063	1008600086	11.507336
1986	120670	1029909613	11.716562
1987	126842	1095029726	11.583430
1988	121026	1054094424	11.481514
1989	160244	1225514347	13.075653
1990	193361	1466620100	13.184123
1991	198020	1489988384	13.290036
1992	211473	1569539447	13.473570
1993	221565	1530416654	14.477430
1994	232063	1548749372	14.983896
1995	243544	1591559103	15.302228
1996	246725	1662267662	14.842676
1997	240745	1702991519	14.136594
1998	249591	1725181351	14.467522
1999	256119	1776363155	14.418167
2000	255832	1799227908	14.218988
2001	250652	1755565489	14.277565
2002	256095	1822152815	14.054529
2003	256079	1838458020	13.929010
2004	240861	1745246613	13.800972
2005	234375	1734909645	13.509349
2006	233361	1840908837	12.676402
2007	233408	1859564353	12.551757
2008	235447	1860620851	12.654217
2009	243487	1976228366	12.320793
2010	238702	1997297329	11.951250
2011	236484	1993362332	11.863573

Year	Total # Suicides	Total Population	Suicide Ratio per 100k
2012	230160	1912812088	12.032546
2013	223199	1890161710	11.808461
2014	222984	1912057309	11.661994
2015	203640	1774657932	11.474887
2016	15603	132101896	11.811337

The total population can be seen trending upwards, but the number fluctuates unexpectedly at various points - for example, the drop from 1,912,057,309 in 2014 to 1,774,657,932 in 2015. The statistics for 2016 are vastly lower than previous years, so I decided to add another column tracking the number of unique countries with available data each year.

In [5]:
```
uniquecountries_list = []
for year in year_list:
    temp_value = df[df["Year"] == year]["Country"].nunique()
    uniquecountries_list.append(temp_value)
globaltotal_df["Unique Countries"] = uniquecountries_list
globaltotal_df
```

Out[5]:

Year	Total # Suicides	Total Population	Suicide Ratio per 100k	Unique Countries
1985	116063	1008600086	11.507336	48
1986	120670	1029909613	11.716562	48
1987	126842	1095029726	11.583430	54
1988	121026	1054094424	11.481514	49
1989	160244	1225514347	13.075653	52
1990	193361	1466620100	13.184123	64
1991	198020	1489988384	13.290036	64
1992	211473	1569539447	13.473570	65
1993	221565	1530416654	14.477430	65
1994	232063	1548749372	14.983896	68
1995	243544	1591559103	15.302228	78
1996	246725	1662267662	14.842676	77
1997	240745	1702991519	14.136594	77
1998	249591	1725181351	14.467522	79
1999	256119	1776363155	14.418167	83
2000	255832	1799227908	14.218988	86
2001	250652	1755565489	14.277565	88

Year	Total # Suicides	Total Population	Suicide Ratio per 100k	Unique Countries
2002	256095	1822152815	14.054529	86
2003	256079	1838458020	13.929010	86
2004	240861	1745246613	13.800972	84
2005	234375	1734909645	13.509349	84
2006	233361	1840908837	12.676402	85
2007	233408	1859564353	12.551757	86
2008	235447	1860620851	12.654217	85
2009	243487	1976228366	12.320793	89
2010	238702	1997297329	11.951250	88
2011	236484	1993362332	11.863573	86
2012	230160	1912812088	12.032546	81
2013	223199	1890161710	11.808461	80
2014	222984	1912057309	11.661994	78
2015	203640	1774657932	11.474887	62
2016	15603	132101896	11.811337	16

In [6]:
```
globaltotal_df = globaltotal_df.drop(globaltotal_df.index[31]) #dropping 2016 data
globaltotal_df
```

Out[6]:

Year	Total # Suicides	Total Population	Suicide Ratio per 100k	Unique Countries
1985	116063	1008600086	11.507336	48
1986	120670	1029909613	11.716562	48
1987	126842	1095029726	11.583430	54
1988	121026	1054094424	11.481514	49
1989	160244	1225514347	13.075653	52
1990	193361	1466620100	13.184123	64
1991	198020	1489988384	13.290036	64
1992	211473	1569539447	13.473570	65
1993	221565	1530416654	14.477430	65
1994	232063	1548749372	14.983896	68
1995	243544	1591559103	15.302228	78
1996	246725	1662267662	14.842676	77
1997	240745	1702991519	14.136594	77

Year	Total # Suicides	Total Population	Suicide Ratio per 100k	Unique Countries
1998	249591	1725181351	14.467522	79
1999	256119	1776363155	14.418167	83
2000	255832	1799227908	14.218988	86
2001	250652	1755565489	14.277565	88
2002	256095	1822152815	14.054529	86
2003	256079	1838458020	13.929010	86
2004	240861	1745246613	13.800972	84
2005	234375	1734909645	13.509349	84
2006	233361	1840908837	12.676402	85
2007	233408	1859564353	12.551757	86
2008	235447	1860620851	12.654217	85
2009	243487	1976228366	12.320793	89
2010	238702	1997297329	11.951250	88
2011	236484	1993362332	11.863573	86
2012	230160	1912812088	12.032546	81
2013	223199	1890161710	11.808461	80
2014	222984	1912057309	11.661994	78
2015	203640	1774657932	11.474887	62

It seems like the variance between the numbers each year are largely impacted by the number of countries with data available. 2016 only has data for 16 countries, so let's remove that year.

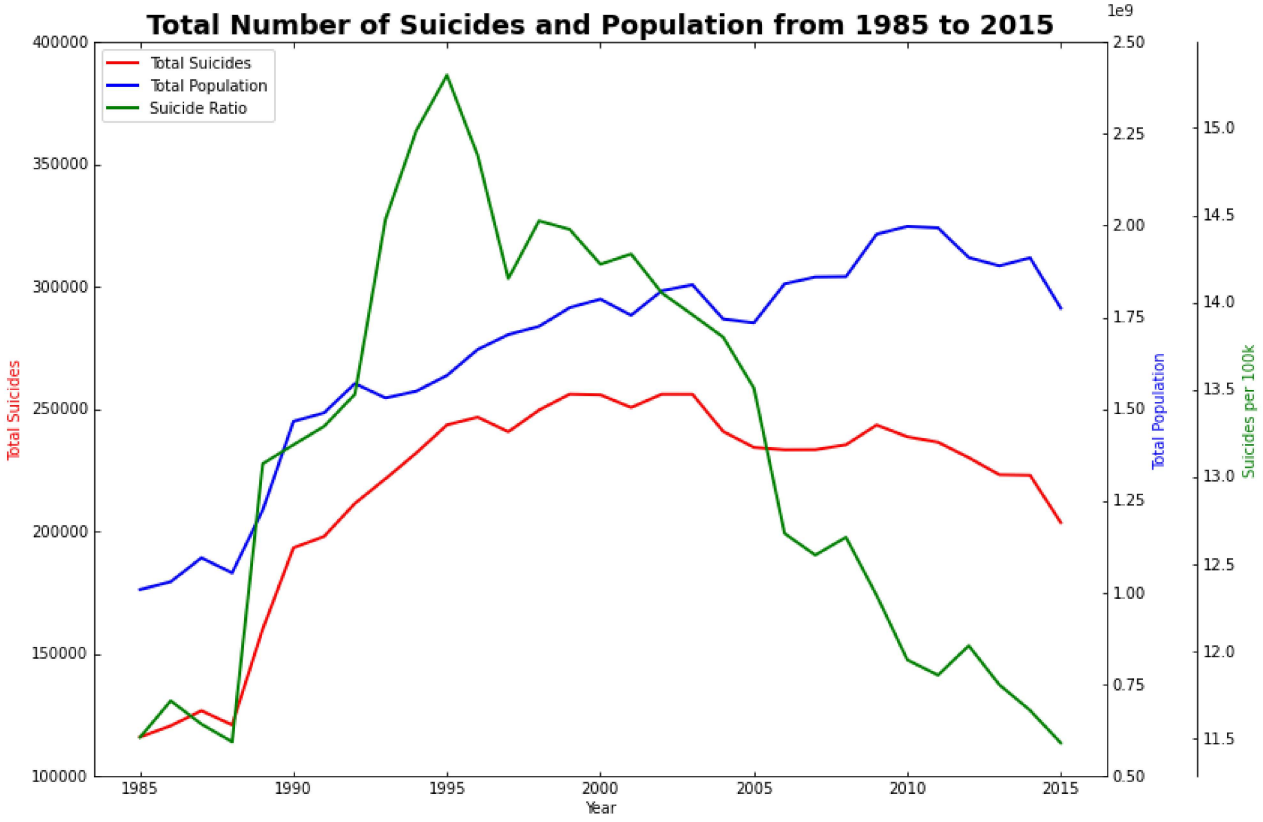
It is important to note that the total population never reaches over 2 billion. This is largely because among the countries with missing data are China and India, which make up nearly 3 billion of the world population alone.

Now with the dataframe filled with the data we want to look at, we can create a line plot tracking each statistic from 1985 to 2005.

In [7]:
```
plt.figure(figsize=[15,9])
host = host_subplot(111, axes_class=axisartist.Axes)
plt.subplots_adjust(right = 0.75)

par1 = host.twinx()
par2 = host.twinx()

par2.axis["right"] = par2.new_fixed_axis(loc = "right", offset = (60,0)) #shifting spin

par1.axis["right"].toggle(all = True)
par2.axis["left"].toggle(all = True)
```

```
p1, = host.plot(globaltotal_df.index, globaltotal_df["Total # Suicides"], color = "red"
p2, = par1.plot(globaltotal_df.index, globaltotal_df["Total Population"], color = "blue
p3, = par2.plot(globaltotal_df.index, globaltotal_df["Suicide Ratio per 100k"], color =

host.set_title("Total Number of Suicides and Population from 1985 to 2015", size = 18,
host.set_xlabel("Year")
host.set_ylabel("Total Suicides", color = "red")
host.set_ylim(100000, 400000)

par1.set_ylabel("Total Population", color = "blue")
par1.set_ylim(500000000, 2500000000)

par2.set_ylabel("Suicides per 100k", color = "green")

host.legend(loc = "upper left")
```

Out[7]: <matplotlib.legend.Legend at 0x1eda7b12670>



Plotting the data over a line graph, we can see that the recorded suicides shot up between 1985 and 1995, but has trended slightly downwards since. However, since this graph tracks total suicides globally, the data is sensitive to variances of the number of countries recorded in a given year - both the population and suicide numbers may be inaccurate for assuming trends because of this. Suicides per 100k is more insightful to look at since it's a ratio. Interestingly, the average suicide rate peaked near 1995 at over 15 per 100k and is down to 11.5 in 2015.

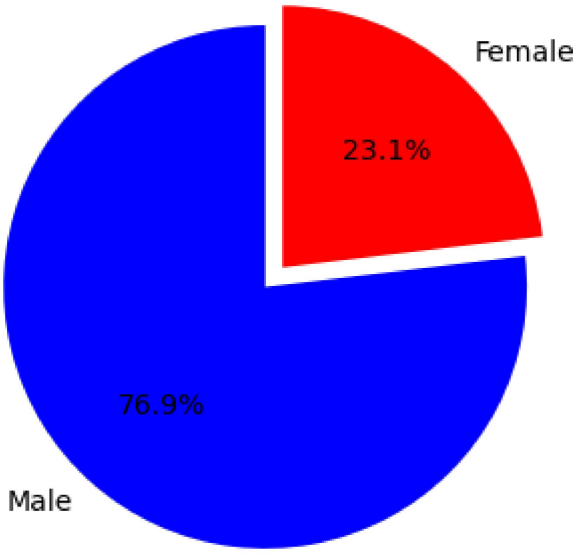
In [8]:
```
gender = {"Gender": ["Male", "Female"]} #creating dataframe for suicide statistics amon
gender_df = pd.DataFrame(gender)
gender_df = gender_df.set_index("Gender")

malesum = df[df["Sex"] == "male"]["Suicides_no"].values.sum() #summing global male suic
femalesum = df[df["Sex"] == "female"]["Suicides_no"].values.sum() #summing global femal
```

```
gendercount = [malesum, femalesum]
gender_df["Total Suicides"] = gendercount
gender_df
```

Out[8]:

Gender	Total Suicides
Male	5188910
Female	1559510

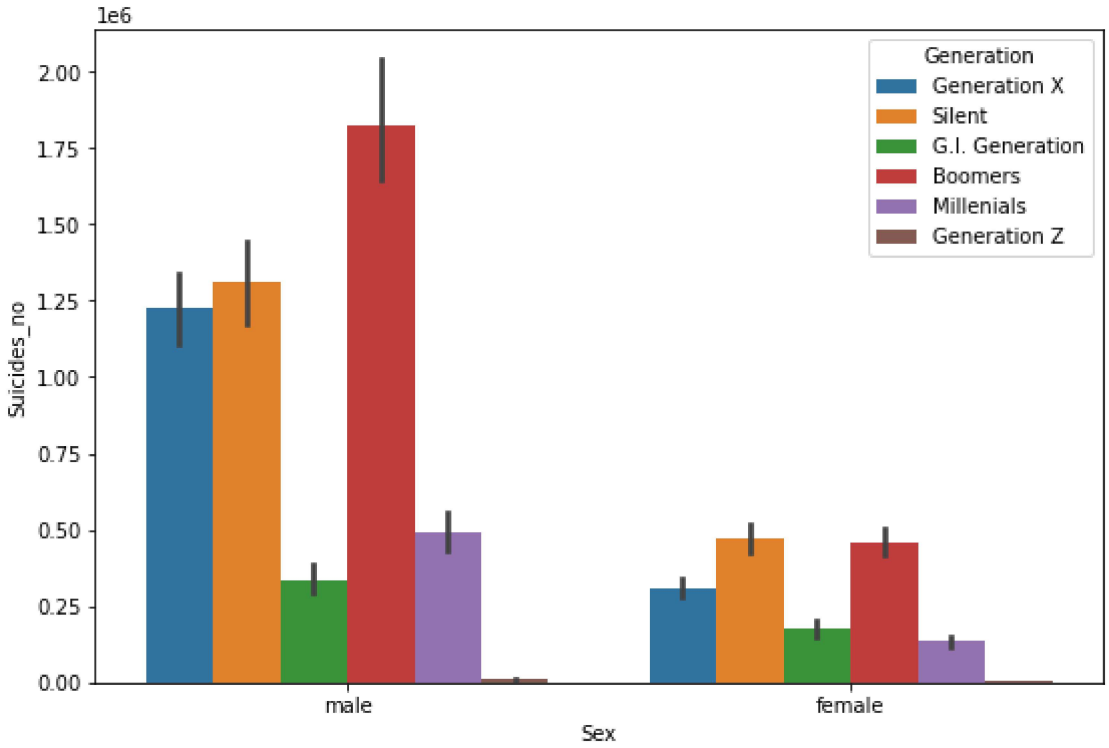
In [9]:
```
fig, ax = plt.subplots(figsize = (9,6))
genderlabels = ["Male", "Female"]
gendercolors = ["blue", "red"]
genderexplode = [0.1, 0]
plt.pie(gender_df["Total Suicides"], labels = genderlabels, colors = gendercolors, expl
plt.show()
```



To investigate how suicide statistics compare between gender, I created a dataframe indexed by gender and plotted the data over a pie chart. Males are more likely to commit suicide than females, as they account for over 3/4ths of suicides over the time range.

In [10]:
```
fig, ax = plt.subplots(figsize = (9,6))
sns.barplot(ax = ax, x = "Sex", y = "Suicides_no", hue = "Generation", data = df, estim
```

Out[10]:
```
<AxesSubplot:xlabel='Sex', ylabel='Suicides_no'>
```



I also plotted the total number of suicides by both gender and generation. As confirmed by the pie chart, males have shown higher numbers of suicide across all the generations. Boomers, Generation X, and the Silent Generation account for the top three generations for suicide for both males and females. Intriguingly, while there are more female G.I. Generation suicides than female Millienials, suicide among male Millenials is higher than the male G.I. Generation.

Country Analysis

Shifting from looking at aggregate statistics, let's begin breaking down suicide statistics by country. Note that many have incomplete data from the date range - nonetheless, suicide rates will be important to examine. We now make a dataframe indexed by country.

In [11]:
```
country_list = df["Country"].unique() #adding to list all unique countries
country_list.sort() #sorting list alphabetically
country_df = pd.DataFrame(country_list) #initializing dataframe by country
country_df.columns = ["Country"]
country_df = country_df.set_index(["Country"])

countrysuicide_list = []
for country in country_list:
    temp_value = df[df["Country"] == country]["Suicides_no"].values.sum()
    temp_value2 = df[df["Country"] == country]["Population"].values.sum()
    countrysuicide_list.append(temp_value*100000/temp_value2)
country_df["Suicides per 100k"] = countrysuicide_list
country_df
```

Out[11]:

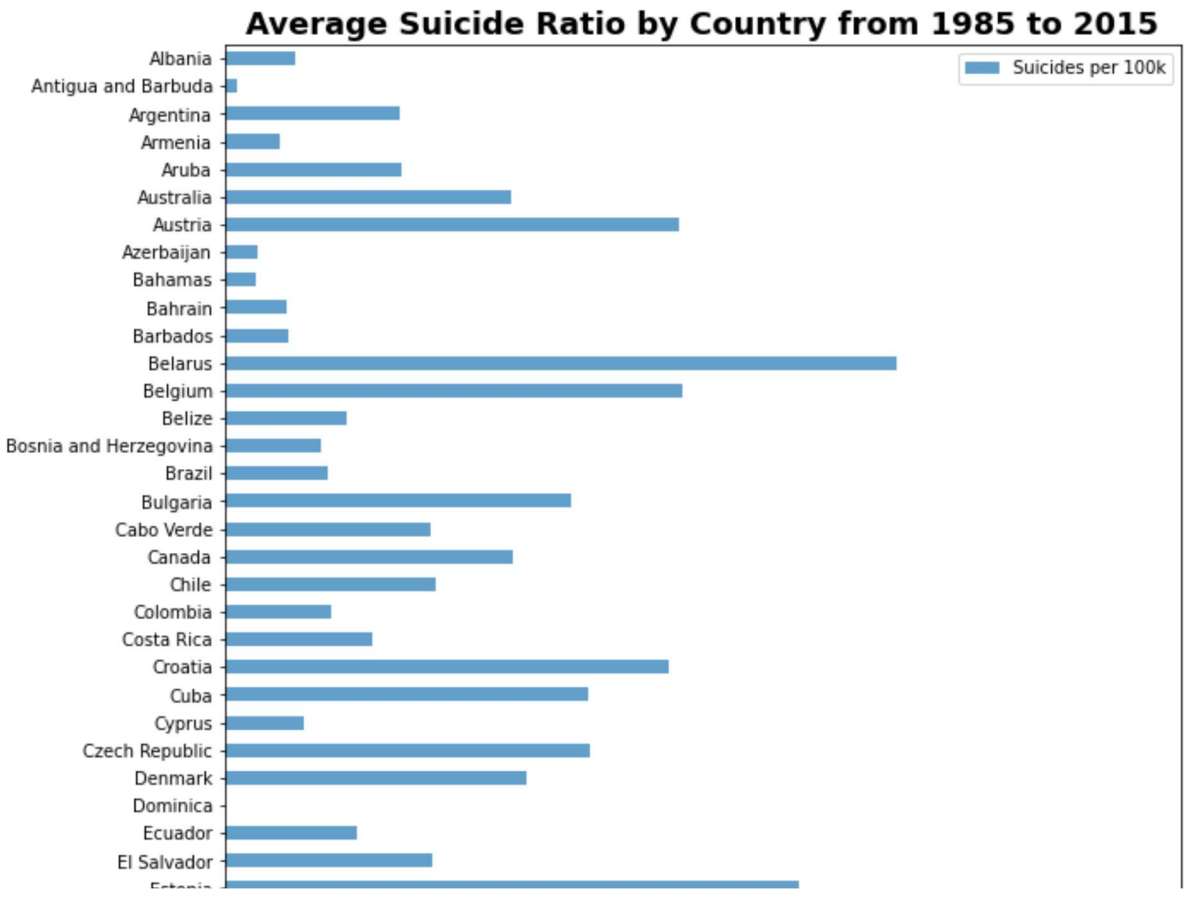
Country	Suicides per 100k
Albania	3.160827

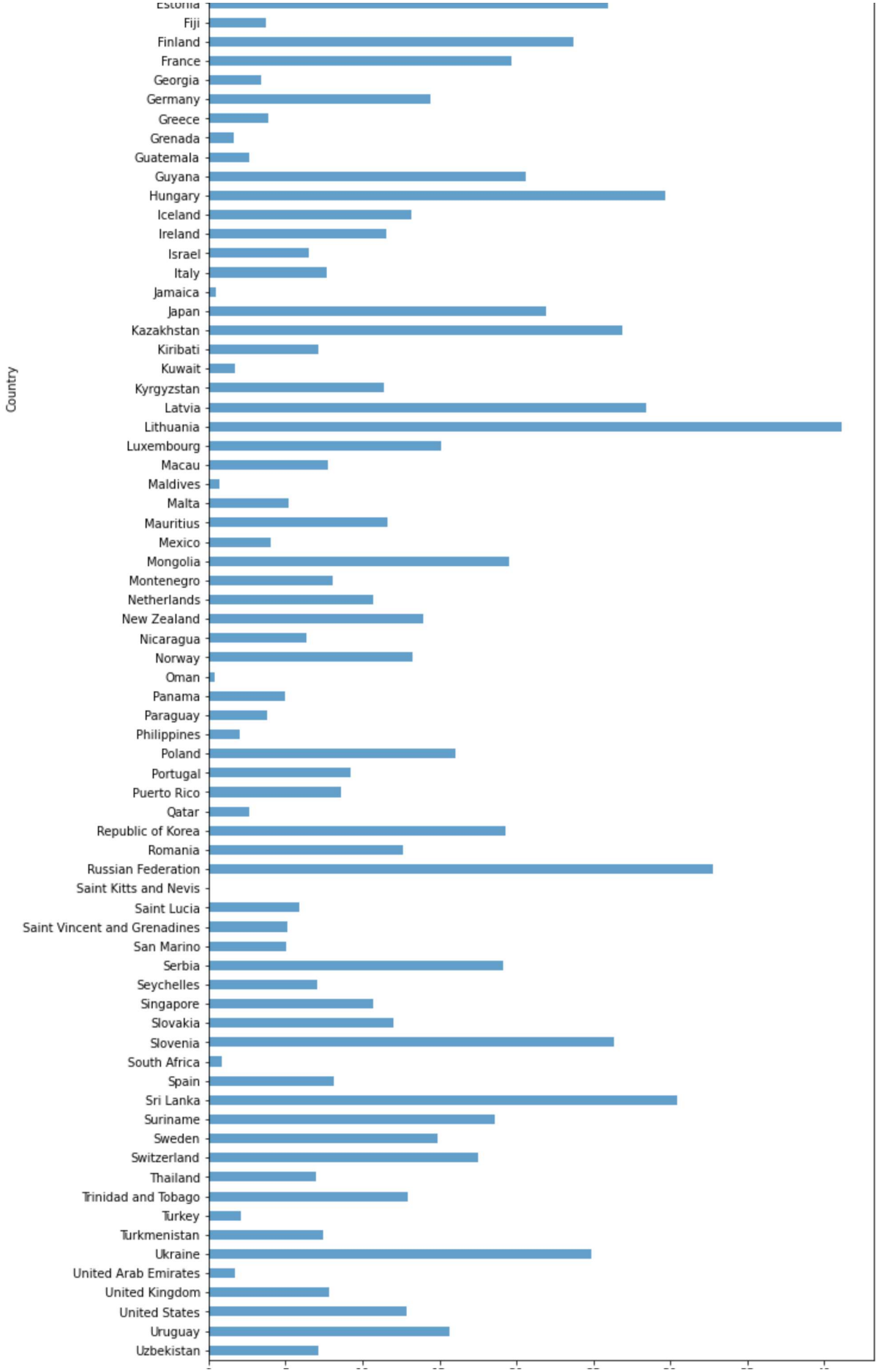
Country	Suicides per 100k
Antigua and Barbuda	0.552700
Argentina	7.936309
Armenia	2.462889
Aruba	8.017928
...	...
United Arab Emirates	1.704003
United Kingdom	7.867928
United States	12.838459
Uruguay	15.627650
Uzbekistan	7.154891

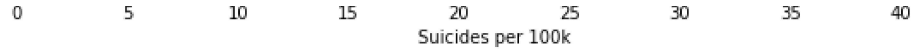
101 rows × 1 columns

In [12]:
```
fig, ax = plt.subplots(figsize = (10,30))
ax.set_title("Average Suicide Ratio by Country from 1985 to 2015", size = 18, fontweigh
country_df = country_df.sort_index(ascending = False)
country_df.plot(ax = ax, kind = "barh", alpha = 0.7) #plotting suicide ratios by countr
ax.set_xlabel("Suicides per 100k")
ax.set_ylabel("Country")
```

Out[12]:
```
Text(0, 0.5, 'Country')
```




Estonia
Fiji
Finland
France
Georgia
Germany
Greece
Grenada
Guatemala
Guyana
Hungary
Iceland
Ireland
Israel
Italy
Jamaica
Japan
Kazakhstan
Kiribati
Kuwait
Kyrgyzstan
Latvia
Lithuania
Luxembourg
Macau
Maldives
Malta
Mauritius
Mexico
Mongolia
Montenegro
Netherlands
New Zealand
Nicaragua
Norway
Oman
Panama
Paraguay
Philippines
Poland
Portugal
Puerto Rico
Qatar
Republic of Korea
Romania
Russian Federation
Saint Kitts and Nevis
Saint Lucia
Saint Vincent and Grenadines
San Marino
Serbia
Seychelles
Singapore
Slovakia
Slovenia
South Africa
Spain
Sri Lanka
Suriname
Sweden
Switzerland
Thailand
Trinidad and Tobago
Turkey
Turkmenistan
Ukraine
United Arab Emirates
United Kingdom
United States
Uruguay
Uzbekistan
Country



We can see that suicide rates vary tremendously between countries. This graph is too cluttered, so let's sort the countries by ratio and look at the extremes.

In [13]:
```
country_df = country_df.sort_values(by = ["Suicides per 100k"], ascending = False) #sor
country_df
```

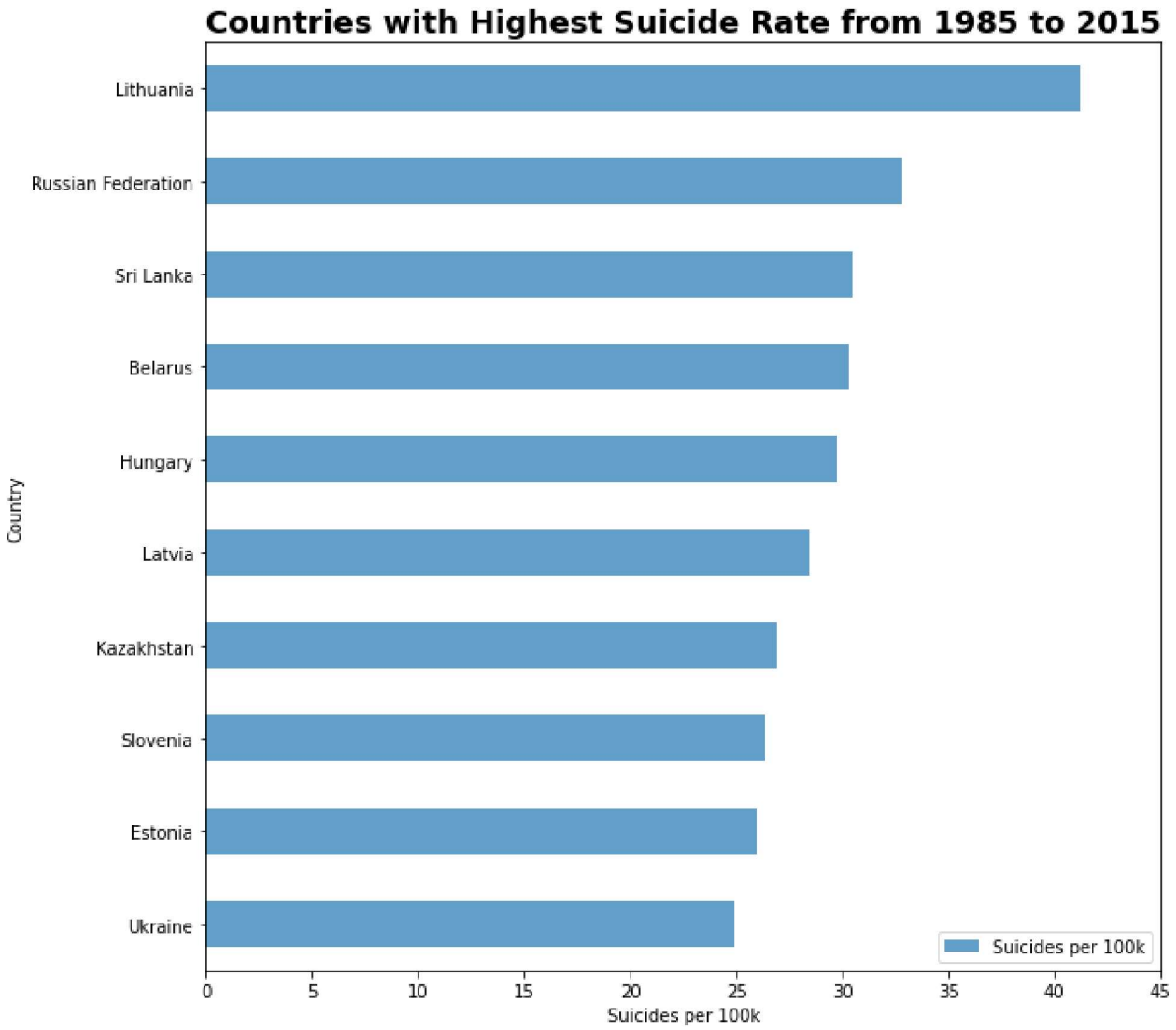
Out[13]:

Country	Suicides per 100k
Lithuania	41.182219
Russian Federation	32.777207
Sri Lanka	30.483939
Belarus	30.344685
Hungary	29.717558
...	...
Antigua and Barbuda	0.552700
Jamaica	0.466037
Oman	0.367194
Saint Kitts and Nevis	0.000000
Dominica	0.000000

101 rows × 1 columns

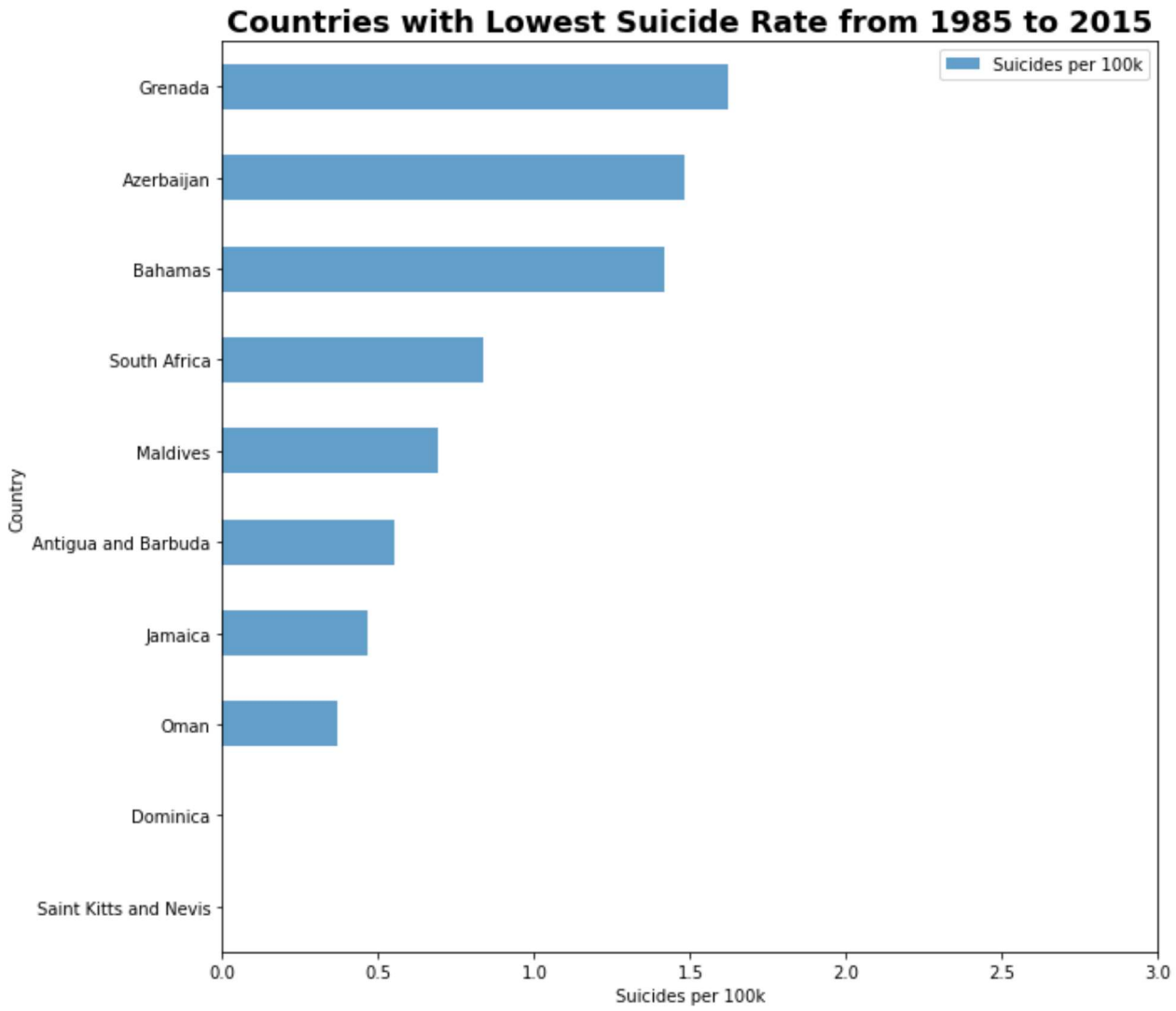
In [14]:
```
topcountry_df = country_df.head(10) #dataframe with 10 countries with highest suicide r

fig, ax = plt.subplots(figsize = (10,10))
ax.set_title("Countries with Highest Suicide Rate from 1985 to 2015", size = 18, fontwe
topcountry_df = topcountry_df.sort_values(by = ["Suicides per 100k"], ascending = True)
topcountry_df.plot(ax = ax, kind = "barh", alpha = 0.7)
ax.set_xlabel("Suicides per 100k")
ax.set_ylabel("Country")
ax.set_xlim(0,45)
```

Out[14]:
```
(0.0, 45.0)
```



In [15]:
```
bottomcountry_df = country_df.tail(10) #dataframe with 10 countries with lowest suiide

fig, ax = plt.subplots(figsize = (10,10))
ax.set_title("Countries with Lowest Suicide Rate from 1985 to 2015", size = 18, fontwei
bottomcountry_df = bottomcountry_df.sort_values(by = ["Suicides per 100k"], ascending =
bottomcountry_df.plot(ax = ax, kind = "barh", alpha = 0.7)
ax.set_xlabel("Suicides per 100k")
ax.set_ylabel("Country")
ax.set_xlim(0,3)
```

Out[15]: (0.0, 3.0)



30 years is a long range, and averages may become misleading when considering the entire dataset. I'm curious to see how the countries on the extreme change when only looking at the 10 more recent years of data.

In [16]:
```
dfrecent = df[(df["Year"] >= 2005) & (df["Year"] <= 2015)] #creating new dataframe with
dfrecent
```

Out[16]:

	Country	Year	Sex	Age	Suicides_no	Population	Suicides/100k pop	Country-year	Hdi for year
192	Albania	2005	female	15-24 years	0	281922	0.00	Albania2005	0.695
193	Albania	2005	female	25-34 years	0	190745	0.00	Albania2005	0.695
194	Albania	2005	female	35-54 years	0	386513	0.00	Albania2005	0.695
195	Albania	2005	female	5-14 years	0	276559	0.00	Albania2005	0.695

	Country	Year	Sex	Age	Suicides_no	Population	Suicides/100k pop	Country-year	Hdi for year	
196	Albania	2005	female	55-74 years	0	210998	0.00	Albania2005	0.695	
...	...	...	...	...	...	...	...	...	...	
27815	Uzbekistan	2014	female	35-54 years	107	3620833	2.96	Uzbekistan2014	0.675	6
27816	Uzbekistan	2014	female	75+ years	9	348465	2.58	Uzbekistan2014	0.675	6
27817	Uzbekistan	2014	male	5-14 years	60	2762158	2.17	Uzbekistan2014	0.675	6
27818	Uzbekistan	2014	female	5-14 years	44	2631600	1.67	Uzbekistan2014	0.675	6
27819	Uzbekistan	2014	female	55-74 years	21	1438935	1.46	Uzbekistan2014	0.675	6

10848 rows × 12 columns

In [17]:
```
recentcountry_list = dfrecent["Country"].unique() #adding to list unique countries
recentcountry_list.sort()
recentcountry_df = pd.DataFrame(recentcountry_list)
recentcountry_df.columns = ["Country"]
recentcountry_df = recentcountry_df.set_index(["Country"])

recentcountrysuicide_list = []
for country in recentcountry_list:
    temp_value = dfrecent[dfrecent["Country"] == country]["Suicides_no"].values.sum()
    temp_value2 = dfrecent[dfrecent["Country"] == country]["Population"].values.sum()
    recentcountrysuicide_list.append(temp_value*100000/temp_value2)
recentcountry_df["Suicides per 100k"] = recentcountrysuicide_list
recentcountry_df = recentcountry_df.sort_values(by = ["Suicides per 100k"], ascending =

recentcountry_df
```

Out[17]:

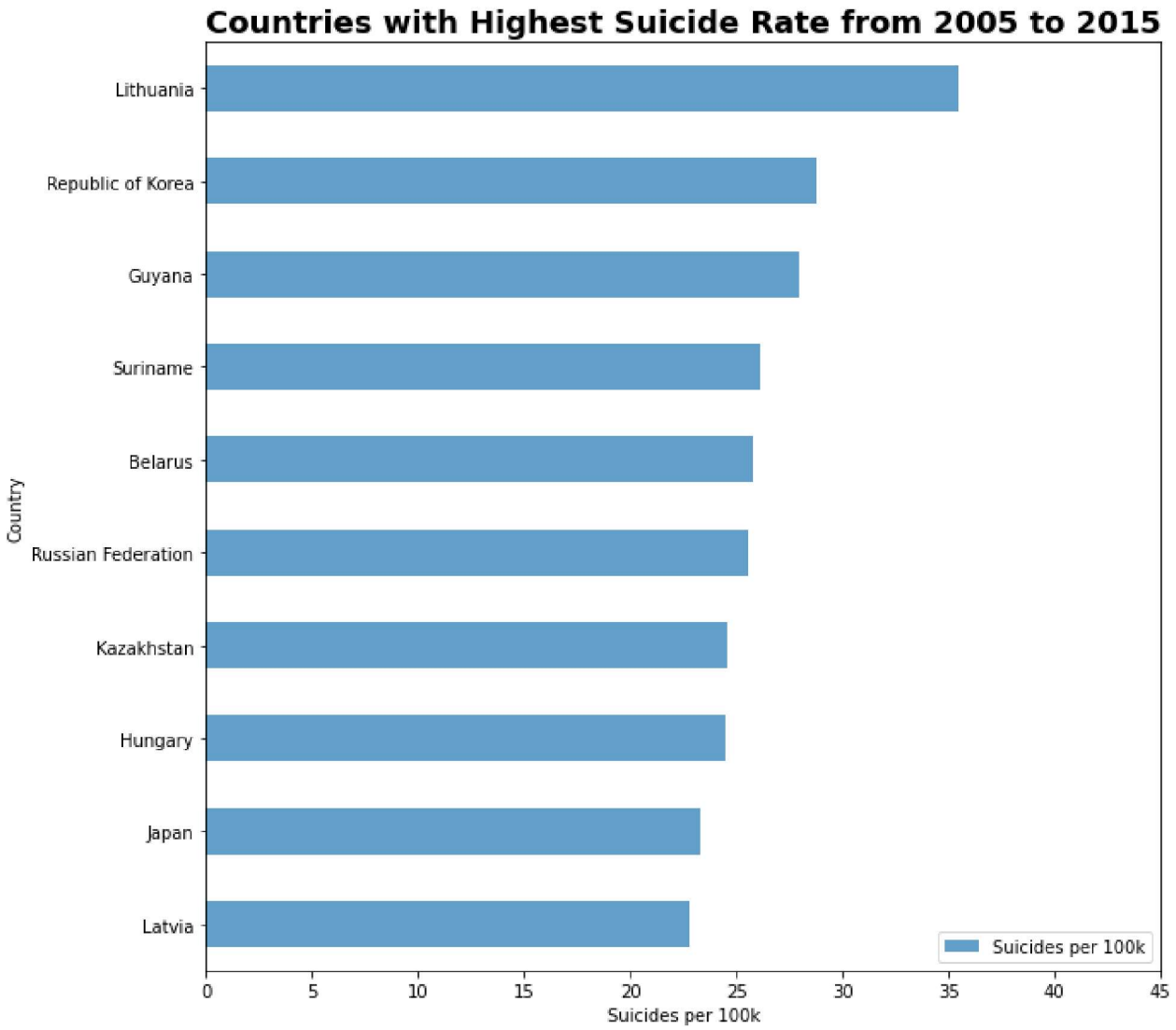
Country	Suicides per 100k
Lithuania	35.468707
Republic of Korea	28.787487
Guyana	27.945491
Suriname	26.132968
Belarus	25.777271
...	...

Country	Suicides per 100k
Barbados	0.810044
Azerbaijan	0.694348
Maldives	0.569270
Oman	0.367194
Antigua and Barbuda	0.257467

96 rows × 1 columns

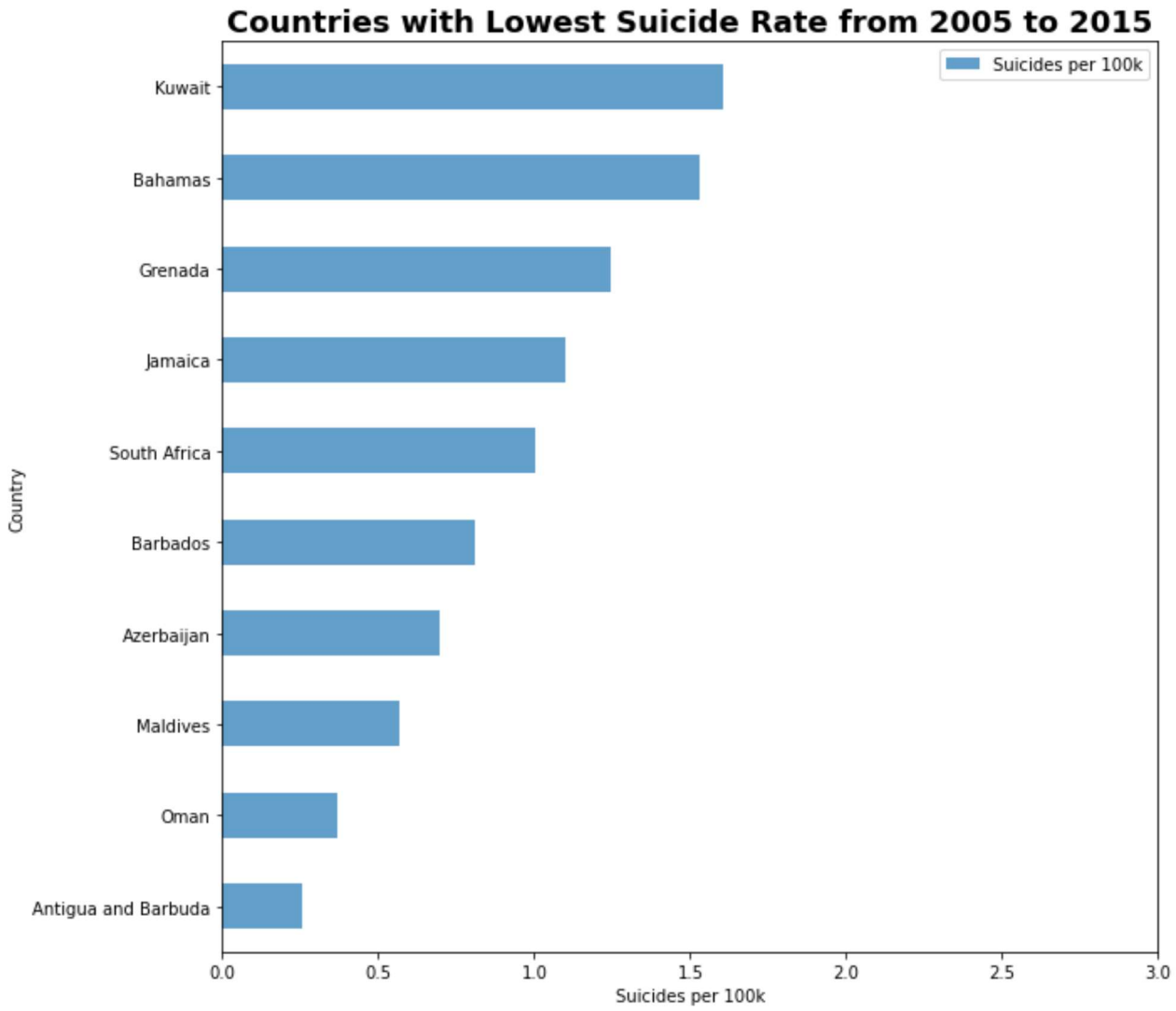
In [18]:
```
recenttopcountry_df = recentcountry_df.head(10)

fig, ax = plt.subplots(figsize = (10,10))
ax.set_title("Countries with Highest Suicide Rate from 2005 to 2015", size = 18, fontwe
recenttopcountry_df = recenttopcountry_df.sort_values(by = ["Suicides per 100k"], ascen
recenttopcountry_df.plot(ax = ax, kind = "barh", alpha = 0.7)
ax.set_xlabel("Suicides per 100k")
ax.set_ylabel("Country")
ax.set_xlim(0,45)
```

Out[18]:
```
(0.0, 45.0)
```



In [19]:
```
recentbottomcountry_df = recentcountry_df.tail(10)

fig, ax = plt.subplots(figsize = (10,10))
ax.set_title("Countries with Lowest Suicide Rate from 2005 to 2015", size = 18, fontwei
recentbottomcountry_df = recentbottomcountry_df.sort_values(by = ["Suicides per 100k"],
recentbottomcountry_df.plot(ax = ax, kind = "barh", alpha = 0.7)
ax.set_xlabel("Suicides per 100k")
ax.set_ylabel("Country")
ax.set_xlim(0,3)
```

Out[19]: (0.0, 3.0)



Lithuania remained at the top of the list, though decreasing to 35 suicides per 100k from 41 per 100k. We see that other countries that were at the top of the list also enjoying lower suicide rates, such as Russia going from 32 suicides per 100k to 25 per 100k. Korea, which wasn't in the original top 10 graph, shot up behind Lithuania at 29 suicides per 100k between 2005 and 2015.

No obvious conclusions can be made from these graphs - it appears like the most developed nations are generally neither at the top nor bottom of the spectrum.

Next, we examine the suicide statistics for four countries: United States, South Korea, Russia, and Japan.

United States

In [20]:
```
us_df = df[df["Country"] == "United States"] # creating dataframe for us data

usyear_list = us_df["Year"].unique()
usyear_list.sort()
ustotal_df = pd.DataFrame(usyear_list)
ustotal_df.columns = ["Year"]
ustotal_df = ustotal_df.set_index(["Year"]) #indexed by year

ussuicide_list = []
```

```
for year in usyear_list:
    temp_value = us_df[us_df["Year"] == year]["Suicides_no"].values.sum()
    ussuicide_list.append(temp_value)
ustotal_df["Total # Suicides"] = ussuicide_list #adding suicide data

uspopulation_list = []
for year in usyear_list:
    temp_value = us_df[us_df["Year"] == year]["Population"].values.sum()
    uspopulation_list.append(temp_value)
ustotal_df["Total Population"] = uspopulation_list #adding population data
ustotal_df["Suicide Ratio per 100k"] = (ustotal_df["Total # Suicides"]*100000)/ustotal_
ustotal_df

ustotal_df
```

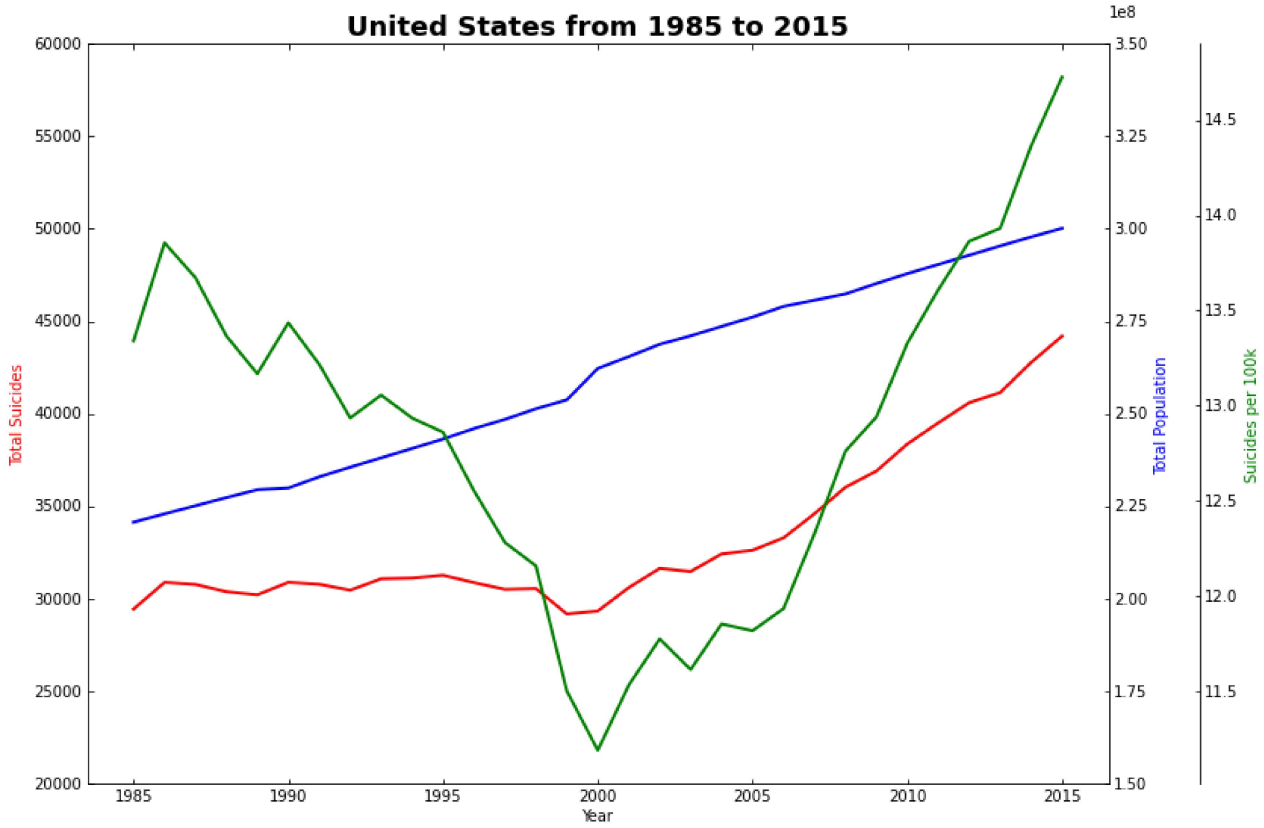
Out[20]:

Year	Total # Suicides	Total Population	Suicide Ratio per 100k
1985	29446	220723000	13.340703
1986	30892	222953000	13.855835
1987	30783	225148000	13.672340
1988	30388	227353000	13.365999
1989	30218	229487000	13.167630
1990	30895	229952200	13.435401
1991	30790	232955000	13.217145
1992	30471	235565600	12.935250
1993	31084	238092300	13.055441
1994	31123	240614100	12.934820
1995	31272	243164200	12.860446
1996	30879	245997800	12.552551
1997	30517	248486400	12.281155
1998	30558	251332500	12.158396
1999	29183	253748671	11.500750
2000	29343	262246108	11.189108
2001	30607	265427546	11.531207
2002	31645	268759559	11.774465
2003	31477	271041510	11.613350
2004	32428	273584136	11.853026
2005	32629	276106680	11.817534
2006	33292	278980848	11.933436
2007	34596	280658606	12.326720
2008	36030	282355565	12.760506

Year	Total # Suicides	Total Population	Suicide Ratio per 100k
2009	36900	285162662	12.939983
2010	38362	287839149	13.327582
2011	39508	290313825	13.608722
2012	40596	292827128	13.863470
2013	41143	295322862	13.931532
2014	42769	297749735	14.364077
2015	44189	300078511	14.725813

In [21]:
```
plt.figure(figsize=[15,9])
host = host_subplot(111, axes_class=axisartist.Axes)
plt.subplots_adjust(right = 0.75)

par1 = host.twinx()
par2 = host.twinx()

par2.axis["right"] = par2.new_fixed_axis(loc = "right", offset = (60,0))

par1.axis["right"].toggle(all = True)
par2.axis["left"].toggle(all = True)

p1, = host.plot(ustotal_df.index, ustotal_df["Total # Suicides"], color = "red", label
p2, = par1.plot(ustotal_df.index, ustotal_df["Total Population"], color = "blue", label
p3, = par2.plot(ustotal_df.index, ustotal_df["Suicide Ratio per 100k"], color = "green"

host.set_title("United States from 1985 to 2015", size = 18, fontweight = "bold")
host.set_xlabel("Year")
host.set_ylabel("Total Suicides", color = "red")
host.set_ylim(20000, 60000)

par1.set_ylabel("Total Population", color = "blue")
par1.set_ylim(150000000, 350000000)

par2.set_ylabel("Suicides per 100k", color = "green")
```

Out[21]:
```
Text(0, 0.5, 'Suicides per 100k')
```



Total suicides in the US increase by over 10,000 per year over the span of the range. We see that even though the population grew to 300,000 in 2015, the suicide rate has risen over recent years and reaches a high of approximately 14 per 100k in 2015. Unlike the global trend where suicide rates peaked around 1995, the US suicide rate actually troughed in 2000.

In [22]:
```
usheatdata = {"Gender": ["male", "male", "male", "male", "male", "male", "female", "fem
              "Age": ["5-14 years", "15-24 years", "25-34 years", "35-54 years", "55-74
usheat_df = pd.DataFrame(usheatdata) #creating dataframe for US heat map
usheat_df
```

Out[22]:

	Gender	Age
0	male	5-14 years
1	male	15-24 years
2	male	25-34 years
3	male	35-54 years
4	male	55-74 years
5	male	75+ years
6	female	5-14 years
7	female	15-24 years
8	female	25-34 years
9	female	35-54 years
10	female	55-74 years
11	female	75+ years

In [23]:
```
usheat_list = []
for index, row in usheat_df.iterrows():
    ussuicide_value = df[(df["Sex"] == row["Gender"]) & (df["Age"] == row["Age"]) & (df
    uspopulation_value = df[(df["Sex"] == row["Gender"]) & (df["Age"] == row["Age"]) &
    us_value = ussuicide_value*100000/uspopulation_value
    usheat_list.append(us_value)
usheat_df["Suicide Rate"] = usheat_list #adding suicide rate data
usheat_df
```

Out[23]:

	Gender	Age	Suicide Rate
0	male	5-14 years	1.049541
1	male	15-24 years	18.759516
2	male	25-34 years	23.083141
3	male	35-54 years	24.624456
4	male	55-74 years	25.738658
5	male	75+ years	43.099450
6	female	5-14 years	0.415830
7	female	15-24 years	3.784173
8	female	25-34 years	5.319043
9	female	35-54 years	7.653027
10	female	55-74 years	6.260330
11	female	75+ years	4.644952

In [24]:
```
usheatdata2 = {"Gender": ["male", "male", "male", "male", "male", "male", "female", "fe
               "Age": ["5-14 years", "15-24 years", "25-34 years", "35-54 years", "55-74
usheat_df2 = pd.DataFrame(usheatdata) #creating dataframe for recent suicide rate data
usheat_df2
```

Out[24]:

	Gender	Age
0	male	5-14 years
1	male	15-24 years
2	male	25-34 years
3	male	35-54 years
4	male	55-74 years
5	male	75+ years
6	female	5-14 years
7	female	15-24 years
8	female	25-34 years
9	female	35-54 years
10	female	55-74 years

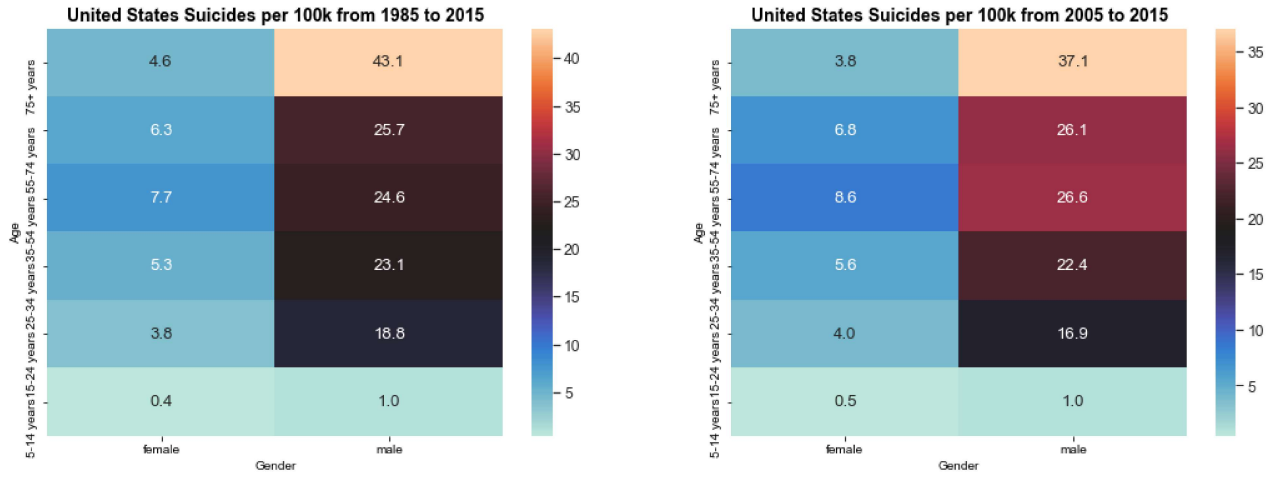
	Gender	Age
11	female	75+ years

In [25]:
```
usheat_list2 = []
for index, row in usheat_df2.iterrows():
    ussuicide_value = dfrecent[(dfrecent["Sex"] == row["Gender"]) & (dfrecent["Age"] ==
    uspopulation_value = dfrecent[(dfrecent["Sex"] == row["Gender"]) & (dfrecent["Age"]
    us_value = ussuicide_value*100000/uspopulation_value
    usheat_list2.append(us_value)
usheat_df2["Suicide Rate"] = usheat_list2
usheat_df2
```

Out[25]:

	Gender	Age	Suicide Rate
0	male	5-14 years	0.956589
1	male	15-24 years	16.943771
2	male	25-34 years	22.363404
3	male	35-54 years	26.619459
4	male	55-74 years	26.079560
5	male	75+ years	37.070393
6	female	5-14 years	0.476960
7	female	15-24 years	3.973328
8	female	25-34 years	5.594178
9	female	35-54 years	8.624128
10	female	55-74 years	6.787031
11	female	75+ years	3.843002

In [33]:
```
fig, ax = plt.subplots(1, 2, figsize = (18, 6))
sns.set(font_scale = 1.1)
ustable = usheat_df.pivot(index = "Age", columns = "Gender", values = "Suicide Rate")
ustable2 = usheat_df2.pivot(index = "Age", columns = "Gender", values = "Suicide Rate")
ustable = ustable.reindex(["75+ years", "55-74 years", "35-54 years", "25-34 years", "1
ustable2 = ustable2.reindex(["75+ years", "55-74 years", "35-54 years", "25-34 years",
sns.heatmap(data = ustable, annot = True, fmt = "1.1f", cmap = "icefire", ax = ax[0])
sns.heatmap(data = ustable2, annot = True, fmt = "1.1f", cmap = "icefire", ax = ax[1])
ax[0].set_title("United States Suicides per 100k from 1985 to 2015", size = 14, fontwei
ax[1].set_title("United States Suicides per 100k from 2005 to 2015", size = 14, fontwei
plt.show()
```



We observe that the males in the US are committing suicide at significantly higher rates than females, akin to the global gender trend. While the suicide rate has decreased for those who are 75+, the rate has increased among people aged 35-54 and 55-74 for both genders. The group with the highest suicide rate remains males who are 75+.

South Korea

In [26]:
```
korea_df = df[df["Country"] == "Republic of Korea"]

koreayear_list = korea_df["Year"].unique()
koreayear_list.sort()
koreatotal_df = pd.DataFrame(koreayear_list)
koreatotal_df.columns = ["Year"]
koreatotal_df = koreatotal_df.set_index(["Year"])

koreasuicide_list = []
for year in koreayear_list:
    temp_value = korea_df[korea_df["Year"] == year]["Suicides_no"].values.sum()
    koreasuicide_list.append(temp_value)
koreatotal_df["Total # Suicides"] = koreasuicide_list

koreapopulation_list = []
for year in koreayear_list:
    temp_value = korea_df[korea_df["Year"] == year]["Population"].values.sum()
    koreapopulation_list.append(temp_value)
koreatotal_df["Total Population"] = koreapopulation_list
koreatotal_df["Suicide Ratio per 100k"] = (koreatotal_df["Total # Suicides"]*100000)/ko
koreatotal_df

koreatotal_df
```

Out[26]:

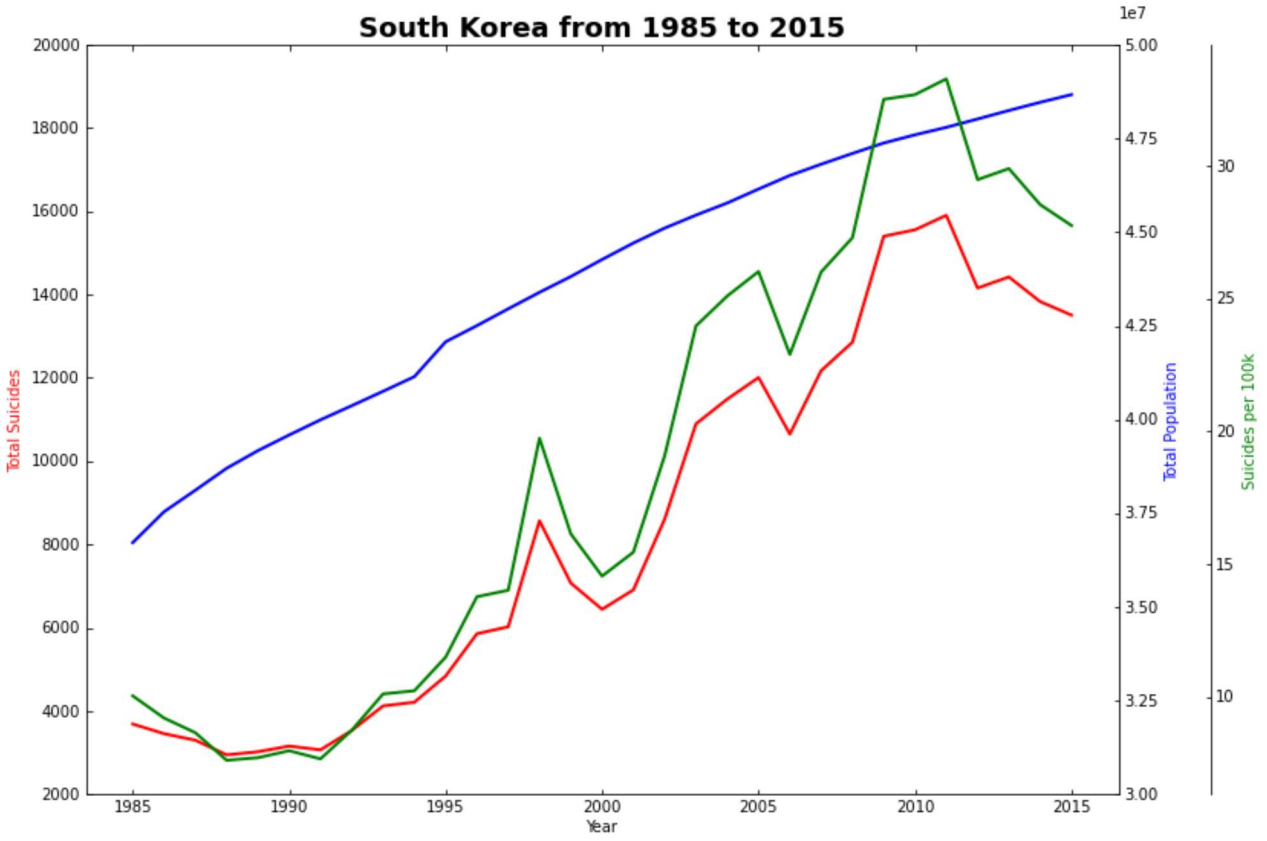
Year	Total # Suicides	Total Population	Suicide Ratio per 100k
1985	3689	36716700	10.047199
1986	3458	37540400	9.211410
1987	3301	38114900	8.660655
1988	2949	38700500	7.620057

Year	Total # Suicides	Total Population	Suicide Ratio per 100k
1989	3023	39171500	7.717346
1990	3159	39589500	7.979388
1991	3069	39993700	7.673709
1992	3533	40371200	8.751288
1993	4124	40754800	10.119053
1994	4212	41145300	10.236892
1995	4841	42076400	11.505262
1996	5857	42507100	13.778875
1997	6024	42956100	14.023619
1998	8569	43395700	19.746196
1999	7075	43817300	16.146591
2000	6444	44271743	14.555560
2001	6911	44710865	15.457093
2002	8611	45110041	19.088876
2003	10898	45455888	23.974892
2004	11492	45779970	25.102681
2005	12011	46151884	26.024940
2006	10653	46513964	22.902800
2007	12174	46814207	26.004926
2008	12858	47101747	27.298351
2009	15402	47380017	32.507375
2010	15558	47598615	32.685825
2011	15906	47798017	33.277531
2012	14159	48020985	29.485026
2013	14426	48248224	29.899546
2014	13834	48465917	28.543770
2015	13510	48671752	27.757374

In [28]:
```
plt.figure(figsize = [15,9])
host = host_subplot(111, axes_class=axisartist.Axes)
plt.subplots_adjust(right = 0.75)

par1 = host.twinx()
par2 = host.twinx()

par2.axis["right"] = par2.new_fixed_axis(loc = "right", offset = (60,0))
```

```
par1.axis["right"].toggle(all = True)
par2.axis["left"].toggle(all = True)

p1, = host.plot(koreatotal_df.index, koreatotal_df["Total # Suicides"], color = "red",
p2, = par1.plot(koreatotal_df.index, koreatotal_df["Total Population"], color = "blue",
p3, = par2.plot(koreatotal_df.index, koreatotal_df["Suicide Ratio per 100k"], color = "

host.set_title("South Korea from 1985 to 2015", size = 18, fontweight = "bold")
host.set_xlabel("Year")
host.set_ylabel("Total Suicides", color = "red")
host.set_ylim(2000, 20000)

par1.set_ylabel("Total Population", color = "blue")
par1.set_ylim(30000000, 50000000)

par2.set_ylabel("Suicides per 100k", color = "green")
```

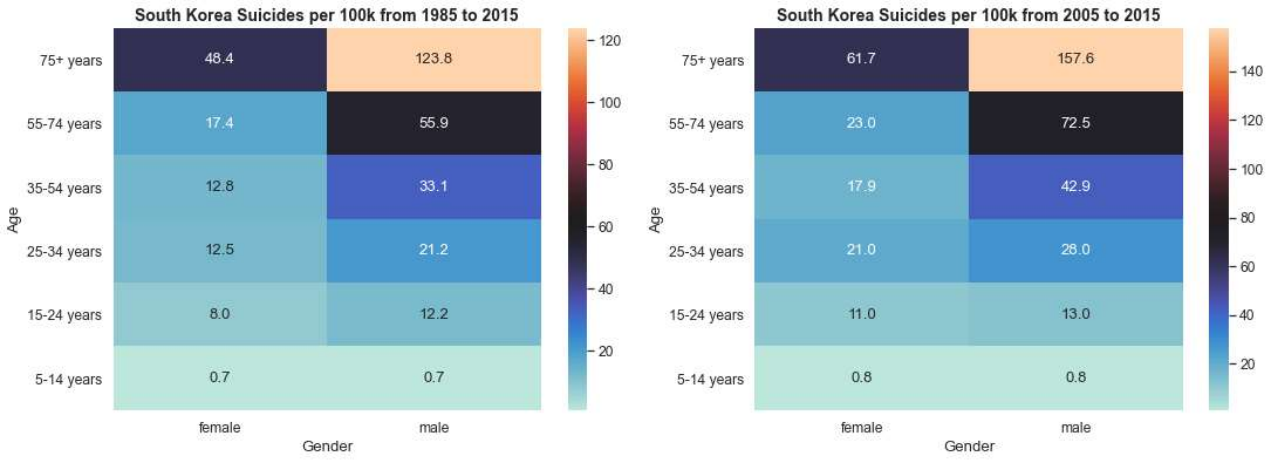
Out[28]: Text(0, 0.5, 'Suicides per 100k')



Suicide in Korea has steadily trended upwards at a concerning rate, increasing from 10 suicides per 100k in 1985 to a high of 33 per 100k in 2011. Total suicides and the suicide rate follow each other closely, suggesting that suicides are becoming more prominent despite a growing population.

In [34]:
```
koreaheatdata = {"Gender": ["male", "male", "male", "male", "male", "male", "female", "
                 "Age": ["5-14 years", "15-24 years", "25-34 years", "35-54 years", "55-74
koreaheat_df = pd.DataFrame(koreaheatdata)

koreaheat_list = []
for index, row in koreaheat_df.iterrows():
    koreasuicide_value = df[(df["Sex"] == row["Gender"]) & (df["Age"] == row["Age"]) &
    koreapopulation_value = df[(df["Sex"] == row["Gender"]) & (df["Age"] == row["Age"])
    korea_value = koreasuicide_value*100000/koreapopulation_value
```

```
    koreaheat_list.append(korea_value)
koreaheat_df["Suicide Rate"] = koreaheat_list
koreaheat_df

koreaheatdata2 = {"Gender": ["male", "male", "male", "male", "male", "male", "female",
             "Age": ["5-14 years", "15-24 years", "25-34 years", "35-54 years", "55-74
koreaheat_df2 = pd.DataFrame(koreaheatdata)
koreaheat_df2

koreaheat_list2 = []
for index, row in koreaheat_df2.iterrows():
    koreasuicide_value = dfrecent[(dfrecent["Sex"] == row["Gender"]) & (dfrecent["Age"]
    koreapopulation_value = dfrecent[(dfrecent["Sex"] == row["Gender"]) & (dfrecent["Ag
    korea_value = koreasuicide_value*100000/koreapopulation_value
    koreaheat_list2.append(korea_value)
koreaheat_df2["Suicide Rate"] = koreaheat_list2
koreaheat_df2

fig, ax = plt.subplots(1, 2, figsize = (18, 6))
sns.set(font_scale = 1.1)
koreatable = koreaheat_df.pivot(index = "Age", columns = "Gender", values = "Suicide Ra
koreatable2 = koreaheat_df2.pivot(index = "Age", columns = "Gender", values = "Suicide
koreatable = koreatable.reindex(["75+ years", "55-74 years", "35-54 years", "25-34 year
koreatable2 = koreatable2.reindex(["75+ years", "55-74 years", "35-54 years", "25-34 ye
sns.heatmap(data = koreatable, annot = True, fmt = "1.1f", cmap = "icefire", ax = ax[0]
sns.heatmap(data = koreatable2, annot = True, fmt = "1.1f", cmap = "icefire", ax = ax[1
ax[0].set_title("South Korea Suicides per 100k from 1985 to 2015", size = 14, fontweigh
ax[1].set_title("South Korea Suicides per 100k from 2005 to 2015", size = 14, fontweigh
plt.show()
```



The suicide rate in Korea has increased for all age groups, most prominently in those who are 25+. The group with the highest suicide rate remains males who are 75+, and the rate had grown to an alarming 157.6 during the most recent years.

Russia

In [29]:
```
russia_df = df[df["Country"] == "Russian Federation"]

russiayear_list = russia_df["Year"].unique()
russiayear_list.sort()
russiatotal_df = pd.DataFrame(russiayear_list)
russiatotal_df.columns = ["Year"]
russiatotal_df = russiatotal_df.set_index(["Year"])
```

```
russiasuicide_list = []
for year in russiayear_list:
    temp_value = russia_df[russia_df["Year"] == year]["Suicides_no"].values.sum()
    russiasuicide_list.append(temp_value)
russiatotal_df["Total # Suicides"] = russiasuicide_list

russiapopulation_list = []
for year in russiayear_list:
    temp_value = russia_df[russia_df["Year"] == year]["Population"].values.sum()
    russiapopulation_list.append(temp_value)
russiatotal_df["Total Population"] = russiapopulation_list
russiatotal_df["Suicide Ratio per 100k"] = (russiatotal_df["Total # Suicides"]*100000)/
russiatotal_df

russiatotal_df
```

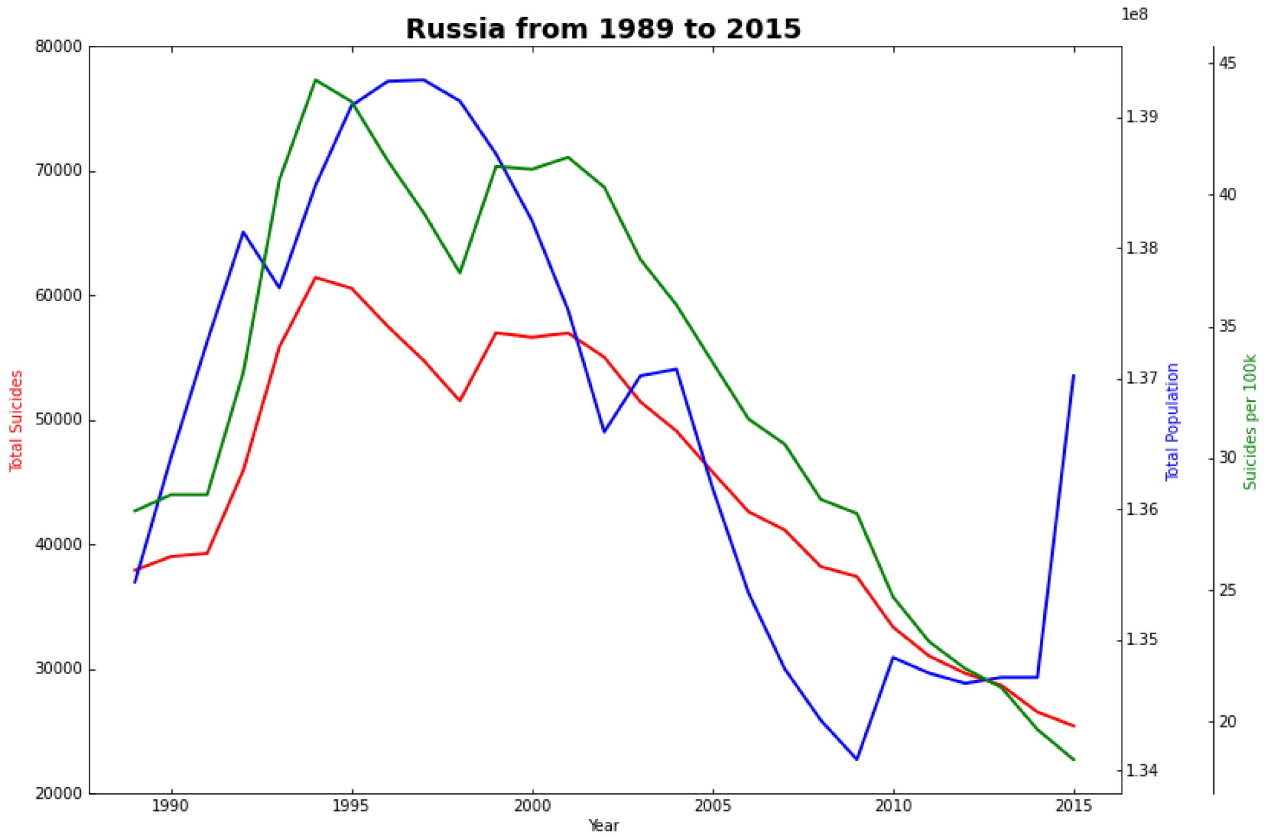
Out[29]:

Year	Total # Suicides	Total Population	Suicide Ratio per 100k
1989	37921	135443800	27.997590
1990	39028	136398200	28.613281
1991	39281	137283000	28.613157
1992	45923	138118900	33.248889
1993	55846	137692800	40.558402
1994	61420	138476200	44.354192
1995	60548	139087100	43.532434
1996	57511	139272600	41.293837
1997	54746	139283200	39.305530
1998	51518	139122100	37.030781
1999	56974	138717800	41.071874
2000	56619	138205695	40.967197
2001	56958	137521509	41.417521
2002	55024	136590900	40.283796
2003	51445	137019868	37.545650
2004	49096	137071699	35.817751
2005	45802	136162191	33.637825
2006	42614	135357462	31.482564
2007	41149	134778964	30.530729
2008	38211	134385949	28.433776
2009	37408	134085433	27.898631
2010	33356	134865987	24.732700
2011	31038	134745874	23.034472

Year	Total # Suicides	Total Population	Suicide Ratio per 100k
2012	29643	134668522	22.011825
2013	28690	134713962	21.296976
2014	26541	134713962	19.701744
2015	25432	137018943	18.560937

In [30]:
```
plt.figure(figsize = [15,9])
host = host_subplot(111, axes_class=axisartist.Axes)
plt.subplots_adjust(right = 0.75)

par1 = host.twinx()
par2 = host.twinx()

par2.axis["right"] = par2.new_fixed_axis(loc = "right", offset = (60,0))

par1.axis["right"].toggle(all = True)
par2.axis["left"].toggle(all = True)

p1, = host.plot(russiatotal_df.index, russiatotal_df["Total # Suicides"], color = "red"
p2, = par1.plot(russiatotal_df.index, russiatotal_df["Total Population"], color = "blue
p3, = par2.plot(russiatotal_df.index, russiatotal_df["Suicide Ratio per 100k"], color =

host.set_title("Russia from 1989 to 2015", size = 18, fontweight = "bold")
host.set_xlabel("Year")
host.set_ylabel("Total Suicides", color = "red")
host.set_ylim(20000,80000)

par1.set_ylabel("Total Population", color = "blue")
par1.set_ylim()

par2.set_ylabel("Suicides per 100k", color = "green")
```

Out[30]:
```
Text(0, 0.5, 'Suicides per 100k')
```

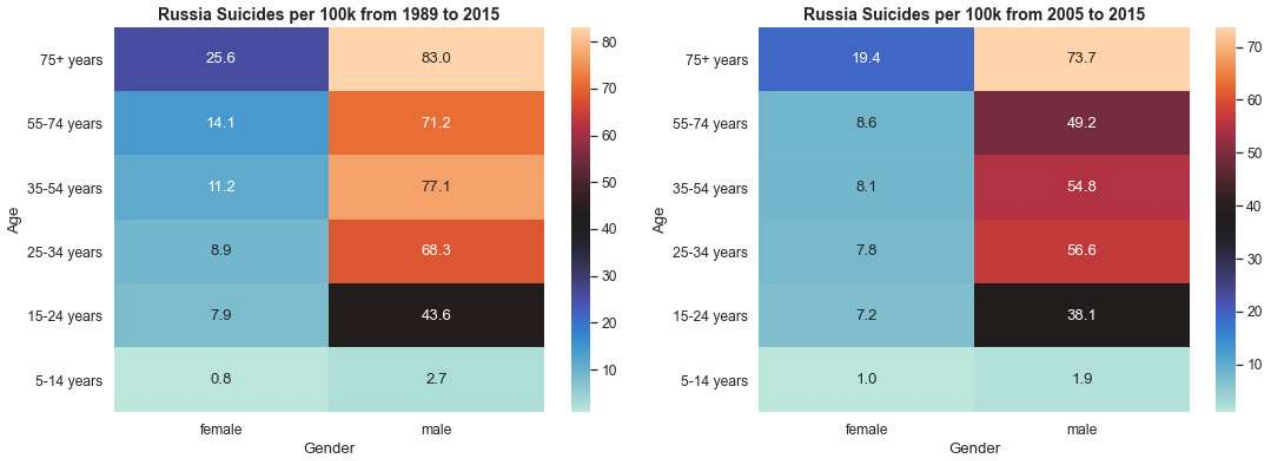


Total suicides and the suicide rate has steadily decreased since its peak in 1994. The trends seen here seem in line with the global trend of decreasing suicide rate. Because Russia has a large population and comparably high suicide rate, the country's data certainly makes a large impact on global trends. Unusually, Russia's population decreases between 1997 and 2009 - this may naturally lower the number of suicides, but the suicide rate has decreased considerably down to below 19 and 2015.

Note that unnlike previous figures, Russia's data begins four years later in 1989.

In [35]:
```
russiaheatdata = {"Gender": ["male", "male", "male", "male", "male", "male", "female",
             "Age": ["5-14 years", "15-24 years", "25-34 years", "35-54 years", "55-74
russiaheat_df = pd.DataFrame(russiaheatdata)

russiaheat_list = []
for index, row in russiaheat_df.iterrows():
    russiasuicide_value = df[(df["Sex"] == row["Gender"]) & (df["Age"] == row["Age"]) &
    russiapopulation_value = df[(df["Sex"] == row["Gender"]) & (df["Age"] == row["Age"]
    russia_value = russiasuicide_value*100000/russiapopulation_value
    russiaheat_list.append(russia_value)
russiaheat_df["Suicide Rate"] = russiaheat_list
russiaheat_df

russiaheatdata2 = {"Gender": ["male", "male", "male", "male", "male", "male", "female",
             "Age": ["5-14 years", "15-24 years", "25-34 years", "35-54 years", "55-74
russiaheat_df2 = pd.DataFrame(russiaheatdata)
russiaheat_df2

russiaheat_list2 = []
for index, row in russiaheat_df2.iterrows():
    russiasuicide_value = dfrecent[(dfrecent["Sex"] == row["Gender"]) & (dfrecent["Age"
    russiapopulation_value = dfrecent[(dfrecent["Sex"] == row["Gender"]) & (dfrecent["A
```

```
    russia_value = russiasuicide_value*100000/russiapopulation_value
    russiaheat_list2.append(russia_value)
russiaheat_df2["Suicide Rate"] = russiaheat_list2
russiaheat_df2

fig, ax = plt.subplots(1, 2, figsize = (18, 6))
sns.set(font_scale = 1.1)
russiatable = russiaheat_df.pivot(index = "Age", columns = "Gender", values = "Suicide
russiatable2 = russiaheat_df2.pivot(index = "Age", columns = "Gender", values = "Suicid
russiatable = russiatable.reindex(["75+ years", "55-74 years", "35-54 years", "25-34 ye
russiatable2 = russiatable2.reindex(["75+ years", "55-74 years", "35-54 years", "25-34
sns.heatmap(data = russiatable, annot = True, fmt = "1.1f", cmap = "icefire", ax = ax[0
sns.heatmap(data = russiatable2, annot = True, fmt = "1.1f", cmap = "icefire", ax = ax[
ax[0].set_title("Russia Suicides per 100k from 1989 to 2015", size = 14, fontweight = "
ax[1].set_title("Russia Suicides per 100k from 2005 to 2015", size = 14, fontweight = "
plt.show()
```



All of the age groups in Russia are oberserved to have lower suicide rates in recent years. Suicide amongst males are consistently higher than females.

Japan

In [31]:

```
japan_df = df[df["Country"] == "Japan"]

japanyear_list = japan_df["Year"].unique()
japanyear_list.sort()
japantotal_df = pd.DataFrame(japanyear_list)
japantotal_df.columns = ["Year"]
japantotal_df = japantotal_df.set_index(["Year"])

japansuicide_list = []
for year in japanyear_list:
    temp_value = japan_df[japan_df["Year"] == year]["Suicides_no"].values.sum()
    japansuicide_list.append(temp_value)
japantotal_df["Total # Suicides"] = japansuicide_list

japanpopulation_list = []
for year in japanyear_list:
    temp_value = japan_df[japan_df["Year"] == year]["Population"].values.sum()
    japanpopulation_list.append(temp_value)
japantotal_df["Total Population"] = japanpopulation_list
japantotal_df["Suicide Ratio per 100k"] = (japantotal_df["Total # Suicides"]*100000)/ja
japantotal_df
```

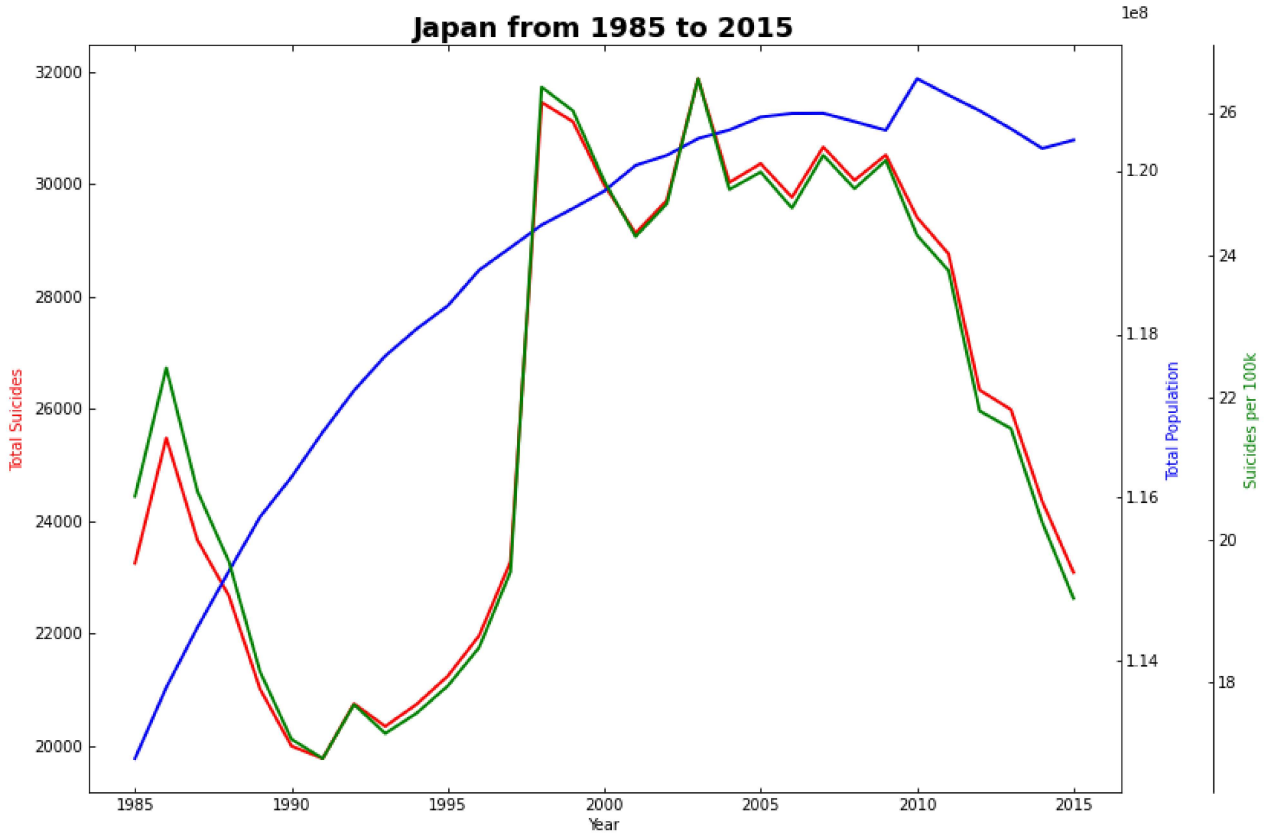
```
japantotal_df
```

Out[31]:

Year	Total # Suicides	Total Population	Suicide Ratio per 100k
1985	23257	112809200	20.616226
1986	25484	113679000	22.417509
1987	23663	114416000	20.681548
1988	22676	115101000	19.700958
1989	21013	115766000	18.151271
1990	20002	116251400	17.205814
1991	19780	116802000	16.934642
1992	20756	117311000	17.693140
1993	20353	117735000	17.287128
1994	20744	118064000	17.570131
1995	21249	118349500	17.954448
1996	21967	118785000	18.493076
1997	23280	119060000	19.553166
1998	31458	119338000	26.360422
1999	31115	119540000	26.028944
2000	29989	119752660	25.042450
2001	29132	120066000	24.263322
2002	29715	120189000	24.723560
2003	31881	120395000	26.480336
2004	30034	120500000	24.924481
2005	30369	120657507	25.169590
2006	29769	120702000	24.663220
2007	30665	120704000	25.405123
2008	30073	120602000	24.935739
2009	30523	120496000	25.331131
2010	29411	121127267	24.281073
2011	28766	120927000	23.787905
2012	26338	120735000	21.814718
2013	25991	120514000	21.566789
2014	24357	120274000	20.251260
2015	23092	120376310	19.183176

In [32]:
```
plt.figure(figsize = [15,9])
host = host_subplot(111, axes_class=axisartist.Axes)
plt.subplots_adjust(right = 0.75)

par1 = host.twinx()
par2 = host.twinx()

par2.axis["right"] = par2.new_fixed_axis(loc = "right", offset = (60,0))

par1.axis["right"].toggle(all = True)
par2.axis["left"].toggle(all = True)

p1, = host.plot(japantotal_df.index, japantotal_df["Total # Suicides"], color = "red",
p2, = par1.plot(japantotal_df.index, japantotal_df["Total Population"], color = "blue",
p3, = par2.plot(japantotal_df.index, japantotal_df["Suicide Ratio per 100k"], color = "

host.set_title("Japan from 1985 to 2015", size = 18, fontweight = "bold")
host.set_xlabel("Year")
host.set_ylabel("Total Suicides", color = "red")
host.set_ylim()

par1.set_ylabel("Total Population", color = "blue")
par1.set_ylim()

par2.set_ylabel("Suicides per 100k", color = "green")
```

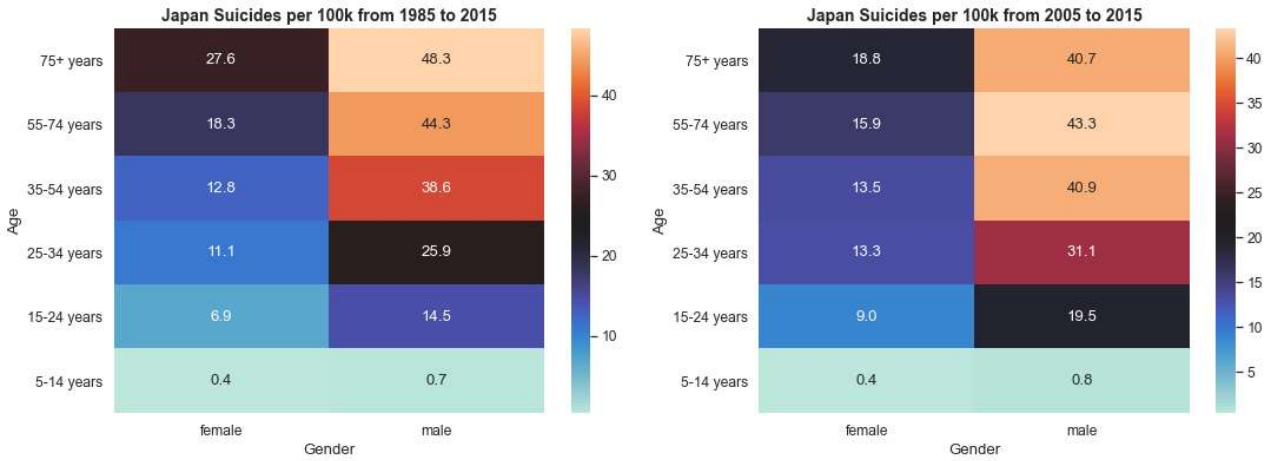
Out[32]: Text(0, 0.5, 'Suicides per 100k')



Suicide in Japan soared in 1998 to 26 suicides per 100k - the rate plateued for 10 years before decreasing in more recent years.

In [36]:
```
japanheatdata = {"Gender": ["male", "male", "male", "male", "male", "male", "female", "
```

```
            "Age": ["5-14 years", "15-24 years", "25-34 years", "35-54 years", "55-74
japanheat_df = pd.DataFrame(japanheatdata)

japanheat_list = []
for index, row in japanheat_df.iterrows():
    japansuicide_value = df[(df["Sex"] == row["Gender"]) & (df["Age"] == row["Age"]) &
    japanpopulation_value = df[(df["Sex"] == row["Gender"]) & (df["Age"] == row["Age"])
    japan_value = japansuicide_value*100000/japanpopulation_value
    japanheat_list.append(japan_value)
japanheat_df["Suicide Rate"] = japanheat_list
japanheat_df

japanheatdata2 = {"Gender": ["male", "male", "male", "male", "male", "male", "female",
            "Age": ["5-14 years", "15-24 years", "25-34 years", "35-54 years", "55-74
japanheat_df2 = pd.DataFrame(japanheatdata)
japanheat_df2

japanheat_list2 = []
for index, row in japanheat_df2.iterrows():
    japansuicide_value = dfrecent[(dfrecent["Sex"] == row["Gender"]) & (dfrecent["Age"]
    japanpopulation_value = dfrecent[(dfrecent["Sex"] == row["Gender"]) & (dfrecent["Ag
    japan_value = japansuicide_value*100000/japanpopulation_value
    japanheat_list2.append(japan_value)
japanheat_df2["Suicide Rate"] = japanheat_list2
japanheat_df2

fig, ax = plt.subplots(1, 2, figsize = (18, 6))
sns.set(font_scale = 1.1)
japantable = japanheat_df.pivot(index = "Age", columns = "Gender", values = "Suicide Ra
japantable2 = japanheat_df2.pivot(index = "Age", columns = "Gender", values = "Suicide
japantable = japantable.reindex(["75+ years", "55-74 years", "35-54 years", "25-34 year
japantable2 = japantable2.reindex(["75+ years", "55-74 years", "35-54 years", "25-34 ye
sns.heatmap(data = japantable, annot = True, fmt = "1.1f", cmap = "icefire", ax = ax[0]
sns.heatmap(data = japantable2, annot = True, fmt = "1.1f", cmap = "icefire", ax = ax[1
ax[0].set_title("Japan Suicides per 100k from 1985 to 2015", size = 14, fontweight = "b
ax[1].set_title("Japan Suicides per 100k from 2005 to 2015", size = 14, fontweight = "b
plt.show()
```



Though the suicide rate in Japan has decreased slightly in recent years, suicide had became more prominent in some groups. Suicide between the ages of 15 and 54 increased between both genders. While males 75+ had the highest rate at 48.3 suicides per 100k, the highest rate in recent years became males aged between 55-74 at 43.3 per 100k.

Conclusion

On a global level, several trends can be seen. First, the average suicide rate has been steadily decreasing since 1995. Most clearly, males are much more likely to commit suicide than females.

Yet, once we break down the data by country, it becomes apparent that the global aggregate data tells a one-dimensional story and that each country has its unique trends and problems with suicide. The four countries picked were arbritrary and certainly only capture a small portion of the world population, but the data provides much insight. While some countries, like Russia, are in-line with the global trend of lower suicide rates, other countries are trending in the opposite direction. Suicide rates in the United States and South Korea continue to increase at alarming rates, and Japan continues to have a high rate as well.

Across the four countries analyzed, some patterns stay consistent: males are shown to have higher suicide rates among all age groups, and suicide risk is generally heightened as people age.