

death_rate_information

May 1, 2022

Death rate information from World bank

```
[2]: #WorldBank_DeathRate.csv
import pandas as pd
import numpy as np

df_death_rate = pd.read_csv ('../Data/DeathRate/WorldBank_DeathRate.csv')

df_death_rate.head()
```

```
[2]:
```

		Series Name	Series Code	Country Name \
0	Death rate, crude (per 1,000 people)	SP.DYN.CDRT.IN		Afghanistan
1	Death rate, crude (per 1,000 people)	SP.DYN.CDRT.IN		Albania
2	Death rate, crude (per 1,000 people)	SP.DYN.CDRT.IN		Algeria
3	Death rate, crude (per 1,000 people)	SP.DYN.CDRT.IN		American Samoa
4	Death rate, crude (per 1,000 people)	SP.DYN.CDRT.IN		Andorra

	Country Code	1960 [YR1960]	1961 [YR1961]	1962 [YR1962]	1963 [YR1963]	\
0	AFG	32.219	31.649	31.093	30.551	
1	ALB	11.326	10.719	10.175	9.707	
2	DZA	20.355	19.967	19.574	19.176	
3	ASM	
4	AND	

	1964 [YR1964]	1965 [YR1965]	...	2011 [YR2011]	2012 [YR2012]	2013 [YR2013]	\
0	30.022	29.501	...	7.936	7.645	7.38	
1	9.32	9.009	...	6.915	6.996	7.096	
2	18.776	18.383	...	4.67	4.685	4.699	
3	
4	3.9	..	

	2014 [YR2014]	2015 [YR2015]	2016 [YR2016]	2017 [YR2017]	2018 [YR2018]	\
0	7.141	6.929	6.742	6.575	6.423	
1	7.219	7.367	7.534	7.714	7.898	
2	4.709	4.715	4.717	4.717	4.716	
3	4.2	5.1	..	
4	4.3	..	4.4	

	2019 [YR2019]	2020 [YR2020]
--	---------------	---------------

0	6.285	..
1	8.082	..
2	4.716	..
3
4	3.9	..

[5 rows x 65 columns]

```
[3]: df_death_rate = df_death_rate.melt(id_vars=["Series Name", "Series Code",
→ "Country Name", "Country Code"],
    var_name="Year",
    value_name="death_per_1000")
```

```
[4]: #Change year to 4 digits only.
df_death_rate['Year'] = df_death_rate['Year'].str[:4]

df_death_rate = df_death_rate.replace('..',np.NaN)

df_death_rate.to_csv("../Data/Cleaned_Data/death_rate_cleaned.csv")
```

```
[5]: #Death rate is per 1000 population, hence we cann't aggregate it.
df_death_rate_wld = df_death_rate[df_death_rate["Country Code"] == 'WLD']
```

```
[ ]:
```