In [192…
```
cd Desktop
```

```
[WinError 2] The system cannot find the file specified: 'Desktop'
C:\Users\zeelt\Desktop
```

In [193…
```
import pandas as pd

# Replace 'your_file_path.xlsx' with the actual path to your Excel file
excel_file_path = 'Data Set.xlsx'

# Read the Excel file into a pandas DataFrame
df = pd.read_excel(excel_file_path, header=None)

# Display the DataFrame to see its structure
print(df)
```

	0	1	2
0		NaN NaN	NaN
1		Week_number NaN	1
2		Week ended NaN	2019-01-04 00:00:00
3		NaN NaN	NaN
4		NaN NaN	NaN
5		Total_deaths NaN	10955
6	Total deaths: average of corresponding	NaN	12273

	3	4	5
0	NaN	NaN	NaN
1	2	3	4
2	2019-01-11 00:00:00	2019-01-18 00:00:00	2019-01-25 00:00:00
3	NaN	NaN	NaN
4	NaN	NaN	NaN
5	12609	11860	11740
6	13670	13056	12486

	6	7	8
0	NaN	NaN	NaN
1	5	6	7
2	2019-02-01 00:00:00	2019-02-08 00:00:00	2019-02-15 00:00:00
3	NaN	NaN	NaN
4	NaN	NaN	NaN
5	11297	11660	11824
6	11998	11623	11301

	9	...	44	45
0	NaN	...	NaN	NaN
1	8	...	43	44
2	2019-02-22 00:00:00	...	2019-10-25 00:00:00	2019-11-01 00:00:00
3	NaN	...	NaN	NaN
4	NaN	...	NaN	NaN
5	11295	...	10021	10164
6	11383	...	9674	9777

	46	47	48
0	NaN	NaN	NaN
1	45	46	47
2	2019-11-08 00:00:00	2019-11-15 00:00:00	2019-11-22 00:00:00
3	NaN	NaN	NaN
4	NaN	NaN	NaN
5	10697	10650	10882
6	10142	10226	10124

	49	50	51
0	NaN	NaN	NaN
1	48	49	50
2	2019-11-29 00:00:00	2019-12-06 00:00:00	2019-12-13 00:00:00
3	NaN	NaN	NaN
4	NaN	NaN	NaN
5	10958	10816	11188
6	10164	10585	10622

	52	53
0	NaN	NaN

```
1                    51                    52
2   2019-12-20 00:00:00   2019-12-27 00:00:00
3                   NaN                   NaN
4                   NaN                   NaN
5                 11926                  7533
6                 11499                  8014

[7 rows x 54 columns]
```

In [194…
```
df.head(10)
```

Out[194]:

	0	1	2	3	4	5	6	7	8
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	Week_number	NaN	1	2	3	4	5	6	7
2	Week ended	NaN	2019-01-04 00:00:00	2019-01-11 00:00:00	2019-01-18 00:00:00	2019-01-25 00:00:00	2019-02-01 00:00:00	2019-02-08 00:00:00	2019-02-15 00:00:00
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5	Total_deaths	NaN	10955	12609	11860	11740	11297	11660	11824
6	Total deaths: average of corresponding	NaN	12273	13670	13056	12486	11998	11623	11301

7 rows × 54 columns

In [195…
```
rows_to_remove = [0,2,3,4,6]
# Use drop to remove the specified rows
df_filtered = df.drop(rows_to_remove)

# Reset the index if after removing rows
df_filtered.reset_index(drop=True, inplace=True)

column_to_remove = df.columns[1]
df_final = df_filtered.drop(column_to_remove, axis=1)

# Display the DataFrame after removing column
print("\nDataFrame after removing rows:")
df_final.head()
```

```
DataFrame after removing rows:
```

Out[195]:

	0	2	3	4	5	6	7	8	9	10	…
0	Week_number	1	2	3	4	5	6	7	8	9	…
1	Total_deaths	10955	12609	11860	11740	11297	11660	11824	11295	11044	…

2 rows × 53 columns

In [196…
```
# Calculate the sum of values in the first row of df_final
sum_of_values_2019 = df_final.iloc[1,1:].sum()

# Print the sum of values in the first row
print("\nSum of values:", sum_of_values_2019)
```

```
Sum of values: 527234
```

SUM OF TOTAL DEATHS OF YEAR 2019

In [197…
```
print(sum_of_values_2019)
```

```
527234
```

DEATH COUNT FOR 2020

In [198…
```
cd Desktop
```

```
[WinError 2] The system cannot find the file specified: 'Desktop'
C:\Users\zeelt\Desktop
```

In [199…
```
import pandas as pd

# Specify the file path
file_path = '2020.xlsx'

# Read the Excel file into a DataFrame
df = pd.read_excel(file_path, header = None)

# Display the DataFrame
df.head(10)
```

Out[199]:

	0	1	2	3	4	5	6	7	8	
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
1	Week number	NaN	1	2	3	4	5	6	7	
2	Week ended	NaN	2020-01-03 00:00:00	2020-01-10 00:00:00	2020-01-17 00:00:00	2020-01-24 00:00:00	2020-01-31 00:00:00	2020-02-07 00:00:00	2020-02-14 00:00:00	2 0 00:(
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
5	Estimated total death occurrences	NaN	12431	12139	11746	10914	11094	10710	10877	1
6	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
7	Upper 95% confidence interval of estimate	NaN	12505	12212	11816	10979	11161	10774	10942	1
8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
9	Lower 95% confidence interval of estimate	NaN	12373	12083	11691	10863	11043	10660	10826	1

10 rows × 55 columns

In [200…

```
rows_to_remove = [0,2,3,4,6,7,8,9,10,11]
df_dropped = df.drop(rows_to_remove)
df_dropped.head()
```

Out[200]:

	0	1	2	3	4	5	6	7	8	9	…	45
1	Week number	NaN	1	2	3	4	5	6	7	8	…	44
5	Estimated total death occurrences	NaN	12431	12139	11746	10914	11094	10710	10877	10795	…	1143

2 rows × 55 columns

In [201…

```
# Calculate the sum of values in the first row of df_final
sum_of_values_2020 = df_dropped.iloc[1,1:].sum()

# Print the sum of values in the first row
print("\nSum of values:", sum_of_values_2020)
```

```
Sum of values: 603077
```

SUM TOTAL OF DEATHS FOR YEAR 2020

In [202…
```
print(sum_of_values_2020)
```

```
603077
```

DEATH COUNT FOR 2021

In [203…
```
cd Desktop
```

```
[WinError 2] The system cannot find the file specified: 'Desktop'
C:\Users\zeelt\Desktop
```

In [204…
```
import pandas as pd

# Specify the file path
file_path = '20211.xlsx'

# Read the Excel file into a DataFrame
df = pd.read_excel(file_path, header = None)

# Display the DataFrame
df.head(10)
```

Out[204]:

	0	1	2	3	4	5	6	7	8	
0	Contents	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
1	Weekly provisional figures on deaths registere...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
3	Footnotes	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
4	Week number	NaN	1	2	3	4	5	6	7	
5	Week ended	NaN	2021-01-08 00:00:00	2021-01-15 00:00:00	2021-01-22 00:00:00	2021-01-29 00:00:00	2021-02-05 00:00:00	2021-02-12 00:00:00	2021-02-19 00:00:00	20 02 00:0(
6	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
8	Total deaths, all ages (2021)	NaN	17751	18042	18676	18448	17192	15354	13809	12
9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N

10 rows × 55 columns

In [205…
```
row_to_remove = [0,1,2,3,5,6,7,9]
df_dropped = df.drop(row_to_remove)
```

In [206…
```
df_dropped = df_dropped.reset_index(drop=True)
df_dropped.head()
```

Out[206]:

	0	1	2	3	4	5	6	7	8	9	...	
0	Week number	NaN	1	2	3	4	5	6	7	8	...	
1	Total deaths, all ages (2021)	NaN	17751	18042	18676	18448	17192	15354	13809	12614	...	11
2	Total deaths, all ages (2020)	NaN	12254	14058	12990	11856	11612	10986	10944	10841	...	10
3	Total deaths: average of corresponding week in...	NaN	12175	13822	13216	12760	12206	11925	11627	11548	...	9
4	Total deaths: average of corresponding week in...	NaN	11412	12933	12370	11933	11419	11154	10876	10790	...	9

5 rows × 55 columns

In [207…

```
row_to_drop = [2,3,4,5,6,7,8,9,10]
df_final = df_dropped.drop(row_to_drop)
# Reset the index after dropping rows
df_final = df_final.reset_index(drop=True)
df_final.head(2)
```

Out[207]:

	0	1	2	3	4	5	6	7	8	9	...	45	
0	Week number	NaN	1	2	3	4	5	6	7	8	...	44	
1	Total deaths, all ages (2021)	NaN	17751	18042	18676	18448	17192	15354	13809	12614	...	11550	1

2 rows × 55 columns

In [208…

```
df_final = df_final.drop(df_final.columns[1], axis=1)
df_final.head(2)
```

Out[208]:

	0	2	3	4	5	6	7	8	9	10	...	45
0	Week number	1	2	3	4	5	6	7	8	9	...	44
1	Total deaths, all ages (2021)	17751	18042	18676	18448	17192	15354	13809	12614	11592	...	11550

2 rows × 54 columns

In [209…
```
# Convert columns to numeric
df_final.iloc[:, 1:] = df_final.iloc[:, 1:].apply(pd.to_numeric, errors='coerce')

# Calculate the sum of values
sum_of_values_2021 = df_final.iloc[1, 1:].sum()
print("Sum of values:", sum_of_values_2021)
```

```
Sum of values: 585899.0
```

SUM TOTAL OF DEATHS FOR YEAR 2021

In [210…
```
print(sum_of_values_2021)
```

```
585899.0
```

DEATH COUNT FOR 2022

In [211…
```
cd Desktop
```

```
[WinError 2] The system cannot find the file specified: 'Desktop'
C:\Users\zeelt\Desktop
```

In [212…
```
import pandas as pd

# Specify the file path
file_path = '20222.xlsx'

# Read the Excel file into a DataFrame
df = pd.read_excel(file_path, header = None)

# Display the DataFrame
df.head()
```

Out[212]:

	0	1	2	3	4	5	6	7	8
0	Sheet 14: Weekly provisional figures on death ...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	[note 2][note 4][note 7][note 9][note 10][note...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	This worksheet contains three tables presented...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	Some cells refer to notes which can be found o...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	Source: Office for National Statistics	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

In [213…
```
rows_to_remove = [0,1,2,3,4,5]
df_dropped = df.drop(rows_to_remove)
df_dropped = df_dropped.reset_index(drop=True)
df_dropped.head(120)
```

Out[213]:

	0	1	2	3	4	5	6	7	8
0	Week number	Week ending	Total deaths	Home	Hospital	Hospice	Care home	Other communal establishment	Elsewhere
1	Week 1 2020 to Current Week	2022-12-30 00:00:00	193778	12666	137016	3109	39489	654	844
2	1	2022-01-07 00:00:00	922	77	719	17	102	0	7
3	2	2022-01-14 00:00:00	1382	125	1048	13	189	3	4
4	3	2022-01-21 00:00:00	1484	98	1059	32	282	6	7
...	...	...	...	...	...	...	...	...	...
115	2	2022-01-14 00:00:00	69	8	50	1	10	0	0
116	3	2022-01-21 00:00:00	102	6	77	0	19	0	0
117	4	2022-01-28 00:00:00	60	2	43	0	15	0	0
118	5	2022-02-04 00:00:00	76	4	51	3	18	0	0
119	6	2022-02-11 00:00:00	40	4	27	0	9	0	0

120 rows × 9 columns

In [214…

```
df_dropped.columns
Total_death_values = df.iloc[:, 2].tolist()
print("Values in the second column:\n", Total_death_values)
Sum_of_values_2022 = Total_death_values[7]
print("\nTotal death values:", Sum_of_values_2022)
```

```
Values in the second column:
 [nan, nan, nan, nan, nan, nan, 'Total deaths', 193778, 922, 1382, 1484, 1385, 1242,
1066, 863, 766, 670, 671, 683, 780, 853, 960, 1003, 1042, 1125, 735, 719, 547, 410,
186, 284, 264, 285, 332, 423, 585, 745, 810, 723, 592, 551, 453, 314, 365, 301, 235,
287, 400, 565, 687, 651, 650, 518, 423, 348, 317, 326, 390, 429, 393, nan, nan, 'Tot
al deaths', 182204, 857, 1308, 1378, 1323, 1162, 1025, 814, 722, 621, 628, 650, 734,
789, 903, 951, 970, 1056, 690, 674, 528, 377, 174, 271, 246, 270, 309, 399, 542, 69
7, 745, 682, 561, 520, 419, 297, 339, 280, 217, 274, 372, 523, 647, 620, 605, 478, 4
01, 333, 299, 295, 364, 397, 367, nan, nan, 'Total deaths', 11247, 61, 69, 102, 60,
76, 40, 47, 43, 48, 40, 30, 45, 61, 56, 51, 70, 67, 45, 45, 18, 33, 12, 12, 16, 14,
20, 22, 41, 46, 62, 39, 30, 29, 32, 17, 24, 20, 17, 13, 28, 40, 40, 29, 44, 40, 21,
15, 17, 31, 26, 32, 25]

Total death values: 193778
```

SUM TOTAL OF DEATHS FOR YEAR 2022

In [215…
```
print(Sum_of_values_2022)
```

```
193778
```

DEATH COUNT FOR 2023

In [216…
```
cd Desktop
```

```
[WinError 2] The system cannot find the file specified: 'Desktop'
C:\Users\zeelt\Desktop
```

In [217…
```
import pandas as pd

# Specify the file path
file_path = '2023.xlsx'

# Read the Excel file into a DataFrame
df = pd.read_excel(file_path, header = None)

# Display the DataFrame
df.head()
```

Out[217]:

	0	1	2	3	4	5	6	7	8
0	Table 14b: Weekly provisional figures on death...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	Week number	Week ending	Total deaths	Home	Hospital	Hospice	Care home	Other communal establishment	Elsewhere
2	Week 1 2020 to Current Week	2023-10-27 00:00:00	196724	12894	139574	3453	39367	630	806
3	1	2023-01-06 00:00:00	679	29	509	13	125	2	1
4	2	2023-01-13 00:00:00	849	54	621	23	147	3	1

In [218…

```
rows_to_remove = [0]
df_dropped = df.drop(rows_to_remove)
df_dropped.head()
```

Out[218]:

	0	1	2	3	4	5	6	7	8
1	Week number	Week ending	Total deaths	Home	Hospital	Hospice	Care home	Other communal establishment	Elsewhere
2	Week 1 2020 to Current Week	2023-10-27 00:00:00	196724	12894	139574	3453	39367	630	806
3	1	2023-01-06 00:00:00	679	29	509	13	125	2	1
4	2	2023-01-13 00:00:00	849	54	621	23	147	3	1
5	3	2023-01-20 00:00:00	711	50	506	16	138	1	0

In [219…

```
df_dropped.columns
Total_death_values = df.iloc[:, 2].tolist()
print("Values in the second column:\n", Total_death_values)
```

```
Sum_of_values_2023 = Total_death_values[2]
print("\nTotal death values:", Sum_of_values_2023)
```

```
Values in the second column:
 [nan, 'Total deaths', 196724, 679, 849, 711, 545, 469, 432, 401, 399, 499, 521, 51
2, 584, 599, 484, 439, 508, 436, 286, 283, 304, 251, 175, 195, 148, 141, 122, 91, 7
3, 61, 55, 65, 101, 115, 178, 136, 197, 229, 241, 203, 233, 318, 287, 297]

Total death values: 196724
```

SUM TOTAL OF DEATHS FOR YEAR 2023

In [220…
```
print(Sum_of_values_2023)
```

```
196724
```

GRAPHS

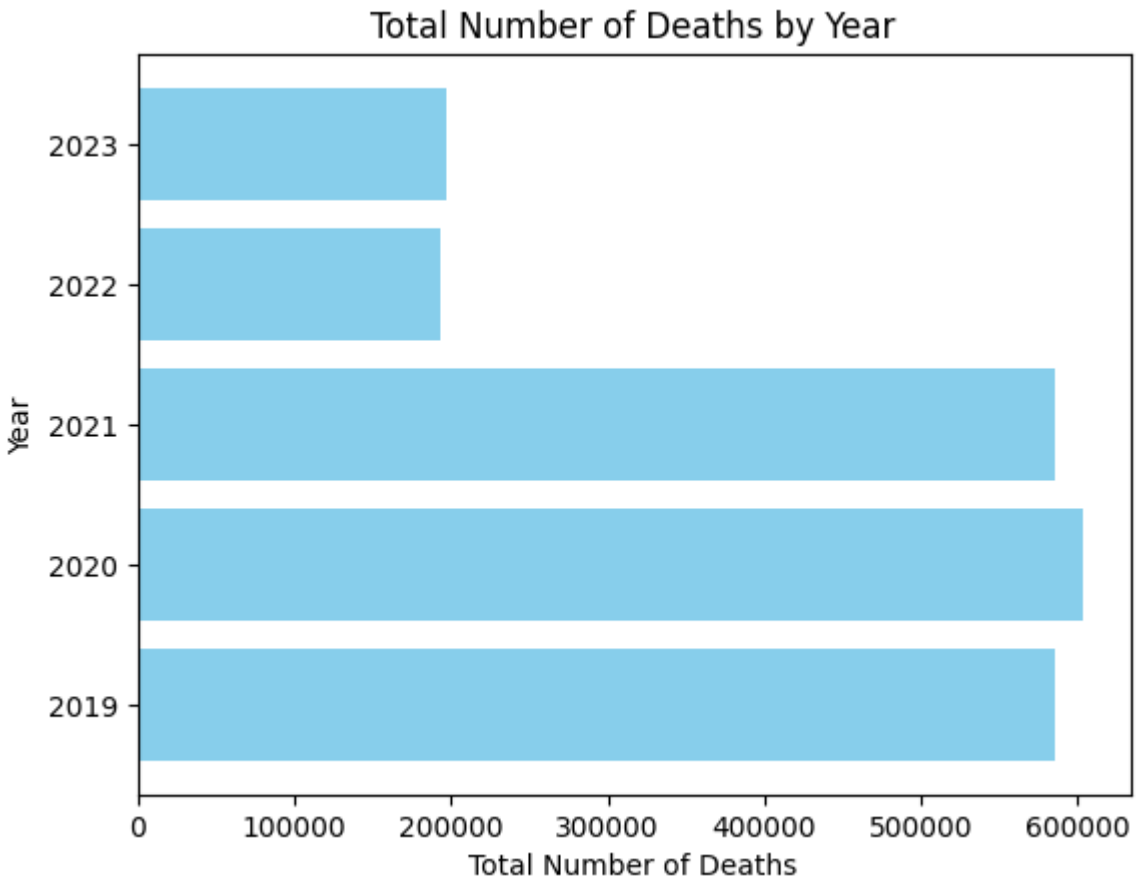
BAR GRAPH

In [221…
```
import matplotlib.pyplot as plt
import pandas as pd

data = {'Year': [2019, 2020, 2021, 2022, 2023],
        'Total Number of Deaths': [585899, 603077, 585899, 193778, 196724]}

df = pd.DataFrame(data)

plt.barh(df['Year'], df['Total Number of Deaths'], color='skyblue')
plt.xlabel('Total Number of Deaths')
plt.ylabel('Year')
plt.title('Total Number of Deaths by Year')
plt.show()
```



LINE GRAPH

In [222…

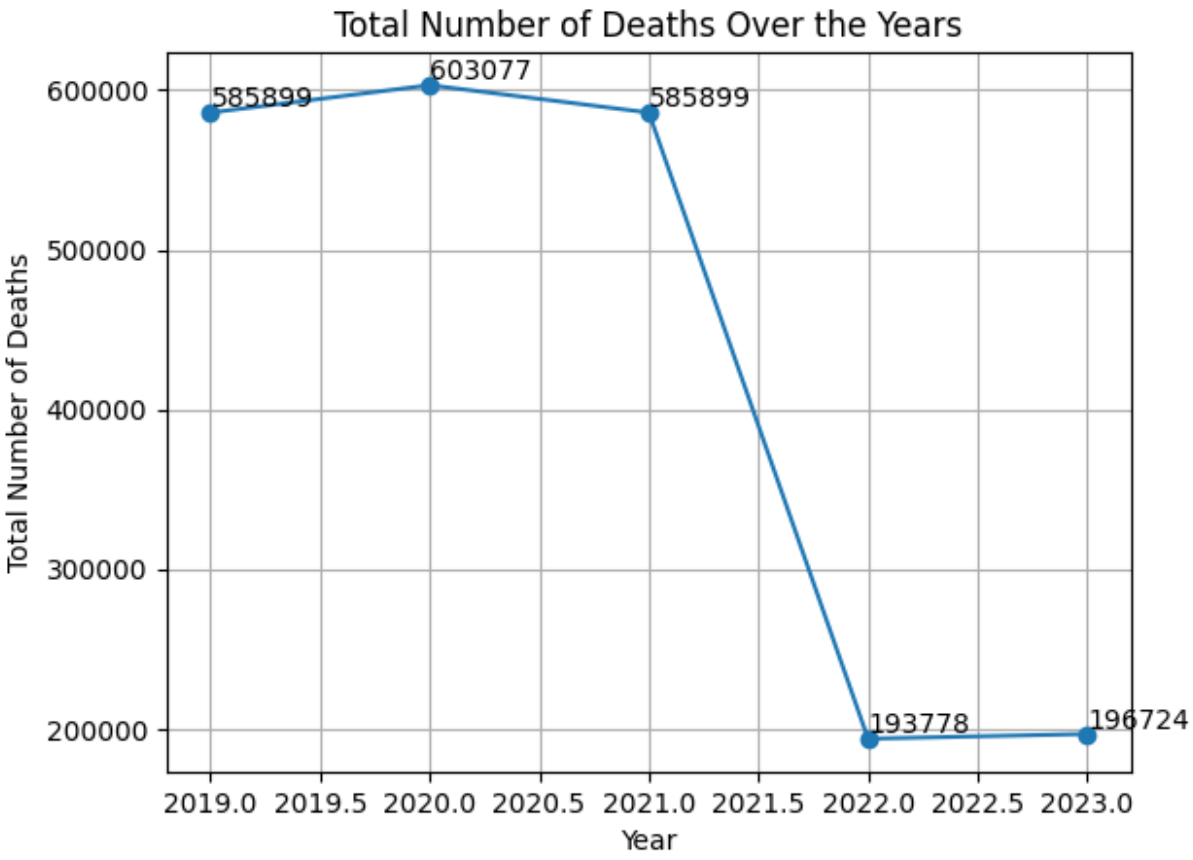
```
import matplotlib.pyplot as plt

data = {'Year': [2019, 2020, 2021, 2022, 2023],
        'Total Number of Deaths': [585899, 603077, 585899, 193778, 196724]}

plt.plot(data['Year'], data['Total Number of Deaths'], marker='o', linestyle='-')
plt.xlabel('Year')
plt.ylabel('Total Number of Deaths')
plt.title('Total Number of Deaths Over the Years')
plt.grid(True)


for year, deaths in zip(data['Year'], data['Total Number of Deaths']):
    plt.text(year, deaths, f'{deaths}', ha='left', va='bottom')

plt.show()
```



PIE CHART

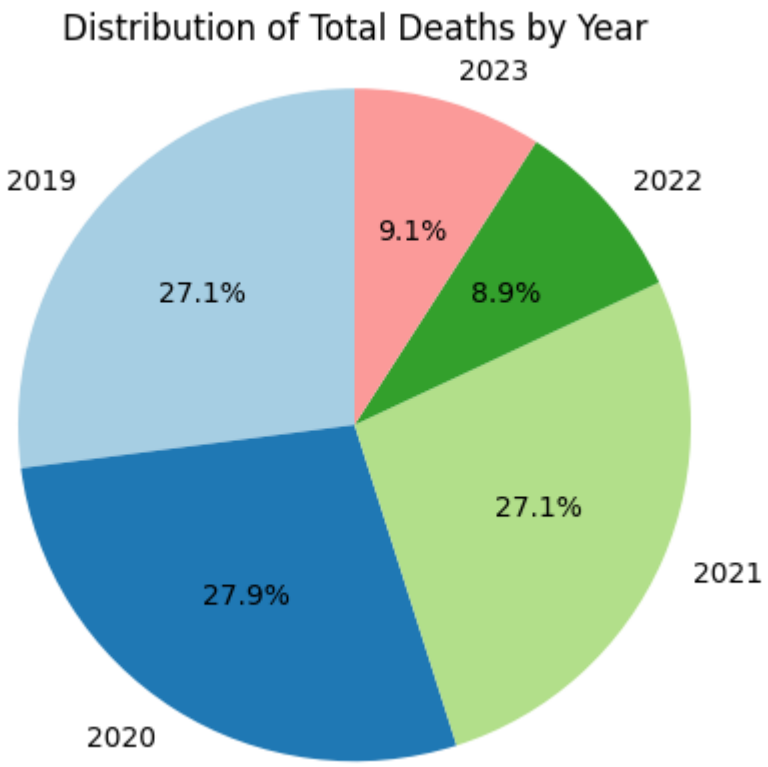
In [223…

```
import matplotlib.pyplot as plt

data = {'Year': [2019, 2020, 2021, 2022, 2023],
        'Total Number of Deaths': [585899, 603077, 585899, 193778, 196724]}

plt.pie(data['Total Number of Deaths'], labels=data['Year'], autopct='%1.1f%%', sta
plt.axis('equal')
plt.title('Distribution of Total Deaths by Year')

plt.show()
```


Distribution of Total Deaths by Year
2023
2022
2019
9.1%
8.9%
27.1%
27.1%
2021
27.9%
2020