Cause of Deaths Analysis

Exploring Trends and Insights in Causes of Death from 1990 to 2019

Python Analysis

Content Map:

- Data Preparation using Pandas
- Statistical analysis and identification of key trends
- Data visualizations with Matplotlib and Seaborn
- Observations

Data Preparation using Pandas

Data Loading

Loading our dataset to our project.

```
df = pd.read_csv('cause_of_deaths.csv')
```

Data Exploration

Previewing parts of our dataset to make sure it is correctly uploaded.

```
df.head()
df.describe()
```

Data Cleaning

Checking our data to make sure if it is clean, complete or containing any NULLs or Duplicates and all the results issued the data as clean and complete data.

```
df.isnull().sum()
df.duplicated().sum()
df['Country/Territory'].nunique()
df["Year"].value_counts()
```

```
# Convert 'meningitis_change' from float64 to int64
df['meningitis_change'] = df['meningitis_change'].astype(int)
```

Data Preparation using Pandas

Feature Extraction

Transforming raw data into numerical features that can be processed while preserving the information in the original data set.

Feature Engineering

Transforming raw data into relevant and informative features that can be used as input for machine learning algorithms.

Data Preparation using Pandas

```
# Display the data types of each column
df.dtypes
```

```
Country/Territory                                object
Code                                             object
Year                                              int64
Meningitis                                        int64
Alzheimer's Disease and Other Dementias           int64
Parkinson's Disease                               int64
Nutritional Deficiencies                          int64
Malaria                                           int64
Drowning                                          int64
Interpersonal Violence                            int64
Maternal Disorders                                int64
HIV/AIDS                                          int64
Drug Use Disorders                                int64
Tuberculosis                                      int64
Cardiovascular Diseases                           int64
Lower Respiratory Infections                      int64
Neonatal Disorders                                int64
Alcohol Use Disorders                             int64
Self-harm                                         int64
Exposure to Forces of Nature                      int64
Diarrheal Diseases                                int64
Environmental Heat and Cold Exposure              int64
Neoplasms                                         int64
Conflict and Terrorism                            int64
Diabetes Mellitus                                 int64
Chronic Kidney Disease                            int64
Poisonings                                        int64
Protein-Energy Malnutrition                       int64
Road Injuries                                     int64
Chronic Respiratory Diseases                      int64
Cirrhosis and Other Chronic Liver Diseases        int64
Digestive Diseases                                int64
Fire, Heat, and Hot Substances                    int64
Acute Hepatitis                                   int64
dtype: object
```

```
len(df)
```

```
6120
```

```
df.shape
```

```
(6120, 34)
```

```
df['Year'].nunique()
# no. of years
```

```
30
```

```
# Unique no. years
df['Year'].unique()
```

```
array([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, 2017, 2018, 2019], dtype=int64)
```

```
# Check for number of unique records present in the data

df.nunique(axis = 0)
```

```
Country/Territory                                204
Code                                             204
Year                                              30
Meningitis                                      2020
Alzheimer's Disease and Other Dementias         3037
Parkinson's Disease                             1817
Nutritional Deficiencies                        2147
Malaria                                         1723
Drowning                                        1875
Interpersonal Violence                          2142
Maternal Disorders                              1818
HIV/AIDS                                        2412
Drug Use Disorders                               876
Tuberculosis                                    2843
Cardiovascular Diseases                         5225
Lower Respiratory Infections                    4106
Neonatal Disorders                              3553
Alcohol Use Disorders                           1287
Self-harm                                       2758
Exposure to Forces of Nature                     478
Diarrheal Diseases                              2874
Environmental Heat and Cold Exposure             714
Neoplasms                                       4814
Conflict and Terrorism                           918
Diabetes Mellitus                               3366
Chronic Kidney Disease                          3246
Poisonings                                      1087
Protein-Energy Malnutrition                     2091
Road Injuries                                   3393
Chronic Respiratory Diseases                    3803
Cirrhosis and Other Chronic Liver Diseases      3443
Digestive Diseases                              4023
Fire, Heat, and Hot Substances                  1406
Acute Hepatitis                                 1059
dtype: int64
```

Data Preparation using Pandas

```
# Total no.of Countries

df['Country/Territory'].nunique()
```

```
204
```

```
# Total no.of year data provided for each country

df['Country/Territory'].value_counts()
```

```
Country/Territory
Afghanistan          30
Papua New Guinea     30
Niue                 30
North Korea          30
North Macedonia      30
                     ..
Greenland            30
Grenada              30
Guam                 30
Guatemala            30
Zimbabwe             30
Name: count, Length: 204, dtype: int64
```

30 year data is provided for Each Country

```
# Correlation of various causes of death against year

# Select only numeric columns
numeric_df = df.select_dtypes(include=[float, int])

# Compute the correlation matrix
correlation_matrix = numeric_df.corr()

# Display correlation of all numeric columns with 'Year'
correlation_with_year = correlation_matrix['Year']
print(correlation_with_year)
```

```
Year                                          1.00
Meningitis                                   -0.04
Alzheimer's Disease and Other Dementias       0.08
Parkinson's Disease                           0.07
Nutritional Deficiencies                     -0.08
Malaria                                      -0.02
Drowning                                     -0.04
Interpersonal Violence                       -0.00
Maternal Disorders                           -0.03
HIV/AIDS                                      0.02
Drug Use Disorders                            0.02
Tuberculosis                                 -0.03
Cardiovascular Diseases                       0.03
Lower Respiratory Infections                 -0.03
Neonatal Disorders                           -0.03
Alcohol Use Disorders                         0.01
Self-harm                                    -0.00
Exposure to Forces of Nature                 -0.01
Diarrheal Diseases                           -0.03
Environmental Heat and Cold Exposure         -0.02
Neoplasms                                     0.04
Conflict and Terrorism                       -0.01
Diabetes Mellitus                             0.07
Chronic Kidney Disease                        0.07
Poisonings                                   -0.01
Protein-Energy Malnutrition                  -0.09
Road Injuries                                 0.01
Chronic Respiratory Diseases                  0.01
Cirrhosis and Other Chronic Liver Diseases    0.03
Digestive Diseases                            0.03
Fire, Heat, and Hot Substances               -0.01
Acute Hepatitis                              -0.03
Name: Year, dtype: float64
```

1: positive correlation (as Year increases, the cause of death increases).

-1: negative correlation (as Year increases, the cause of death decreases).

0: No correlation (no relationship between Year and the cause of death).

Statistical analysis and identification of key trends

Which country has the highest number of deaths due to cardiovascular diseases?

```
car_disease = df.groupby("Country/Territory")["Cardiovascular Diseases"].sum().sort_values(ascending=False).head(20)
car_disease
```

```
Country/Territory
China             100505973
India              52994710
Russia             33903781
United States      26438346
Indonesia          13587011
Ukraine            13053052
Germany            10819770
Brazil              9589019
Japan               9210437
Pakistan            7745192
Italy               6614384
United Kingdom      6603062
Bangladesh          6123691
Egypt               5995471
Vietnam             5323920
Poland              5233134
France              4729313
Romania             4474916
Nigeria             4176488
Turkey              4167835
Name: Cardiovascular Diseases, dtype: int64
```

What is the percentage of deaths caused by lower respiratory infections in the total number of deaths?

```
deaths_causes = df.iloc[:, 3:].sum().sort_values(ascending = False)
deaths_causes_per = (deaths_causes.div(deaths_causes.sum()) * 100).round(2)
deaths_causes_per
```

```
Cardiovascular Diseases                       30.50
Neoplasms                                     15.65
Chronic Respiratory Diseases                   7.13
Lower Respiratory Infections                   5.71
Neonatal Disorders                             5.24
Diarrheal Diseases                             4.51
Digestive Diseases                             4.47
Tuberculosis                                   3.12
Cirrhosis and Other Chronic Liver Diseases     2.55
HIV/AIDS                                       2.48
Road Injuries                                  2.47
Diabetes Mellitus                              2.14
Alzheimer's Disease and Other Dementias        2.03
Chronic Kidney Disease                         1.97
Malaria                                        1.73
Self-harm                                      1.62
Nutritional Deficiencies                       0.94
Interpersonal Violence                         0.87
Protein-Energy Malnutrition                    0.82
Meningitis                                     0.72
Drowning                                       0.70
Maternal Disorders                             0.53
Parkinson's Disease                            0.49
Alcohol Use Disorders                          0.33
Acute Hepatitis                                0.26
Fire, Heat, and Hot Substances                 0.25
Conflict and Terrorism                         0.22
Drug Use Disorders                             0.18
Poisonings                                     0.18
Environmental Heat and Cold Exposure           0.12
Exposure to Forces of Nature                   0.10
dtype: float64
```

Statistical analysis and identification of key trends

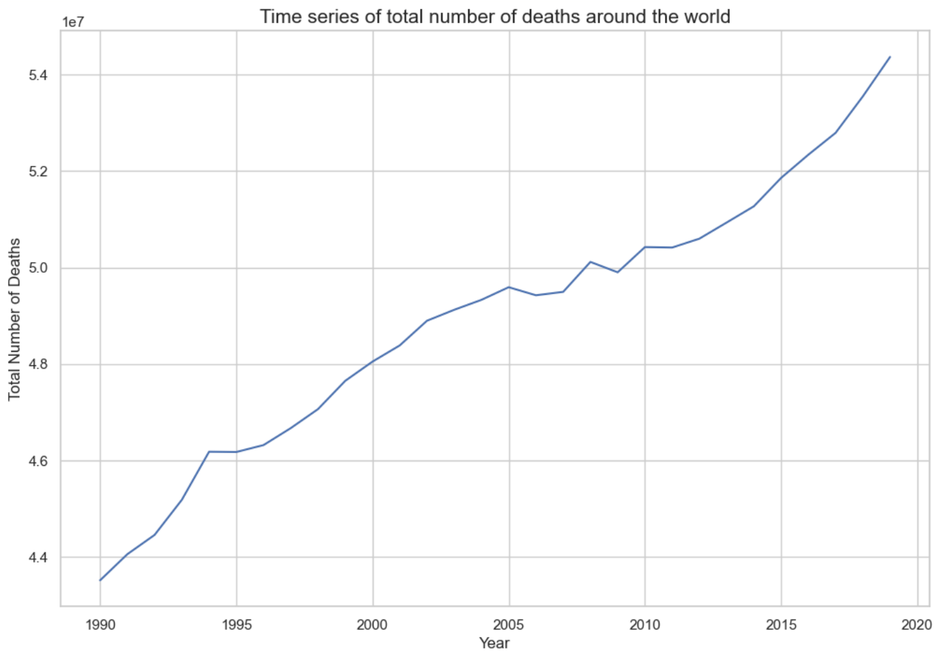
```python
# List of African countries (this can be expanded or adjusted as needed)
african_countries = [
    'Algeria', 'Angola', 'Benin', 'Botswana', 'Burkina Faso', 'Burundi',
    'Cabo Verde', 'Cameroon', 'Central African Republic', 'Chad', 'Comoros',
    'Congo', 'Cote d\'Ivoire', 'Djibouti', 'DR Congo', 'Egypt', 'Equatorial Guinea',
    'Eritrea', 'Eswatini', 'Ethiopia', 'Gabon', 'Gambia', 'Ghana', 'Guinea',
    'Guinea-Bissau', 'Kenya', 'Lesotho', 'Liberia', 'Libya', 'Madagascar', 'Malawi',
    'Mali', 'Mauritania', 'Mauritius', 'Morocco', 'Mozambique', 'Namibia',
    'Niger', 'Nigeria', 'Rwanda', 'Sao Tome and Principe', 'Senegal', 'Seychelles',
    'Sierra Leone', 'Somalia', 'South Africa', 'South Sudan', 'Sudan', 'Tanzania',
    'Togo', 'Tunisia', 'Uganda', 'Zambia', 'Zimbabwe'
]

# Filter out African countries
non_african_malaria = df[~df['Country/Territory'].isin(african_countries)]
malaria_non_african = non_african_malaria.groupby("Country/Territory")["Malaria"].sum().sort_values(ascending=False).head(10)

malaria_non_african
```

```
Country/Territory
Democratic Republic of Congo    2557219
India                           2439244
Bangladesh                       349375
Pakistan                         213590
Myanmar                          157143
Yemen                            143463
Indonesia                         74664
Brazil                            39970
Papua New Guinea                  35997
Haiti                             28833
Name: Malaria, dtype: int64
```

Statistical analysis and identification of key trends

Calculate the total number of deaths for each cause across all years and countries.

```
# Calculate total deaths for each cause
total_deaths = df.drop(columns=['Country/Territory', 'Code', 'Year']).sum().sort_values(ascending=False)

# Display the top causes of death
print(total_deaths)
```



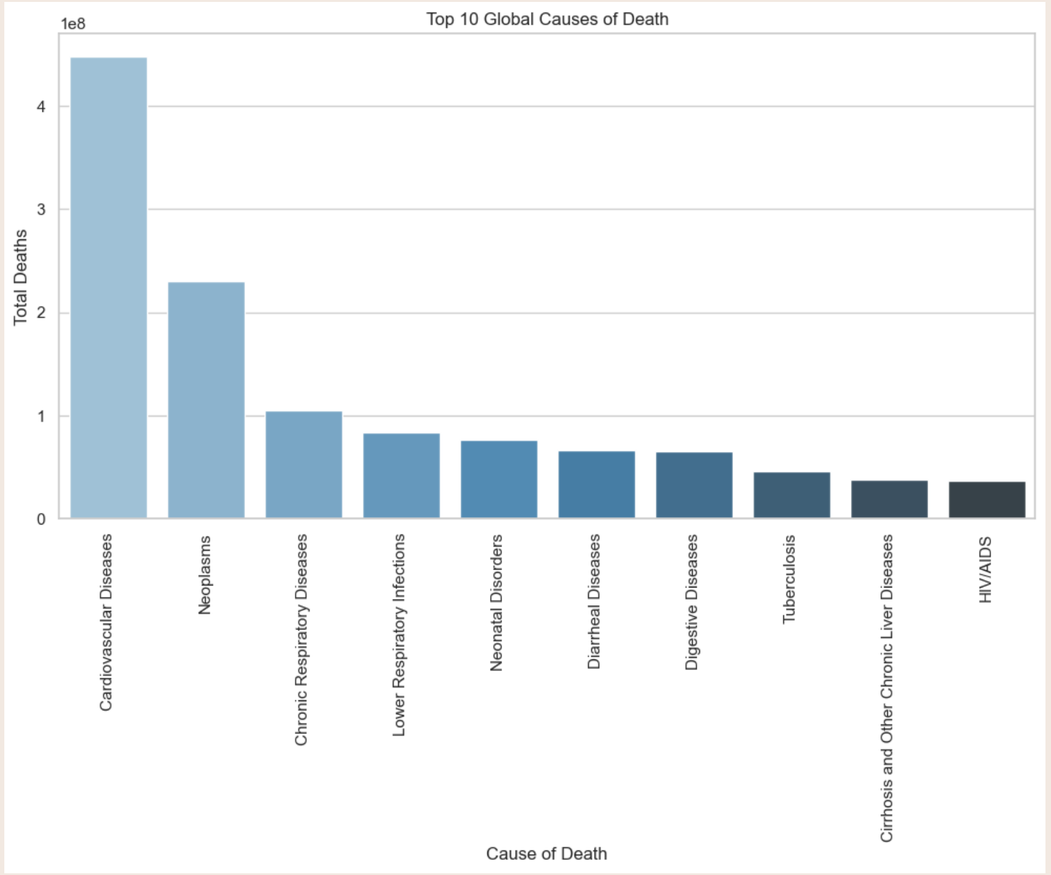
Total_no_of_Deaths	1468134716
Cardiovascular Diseases	447741982
Neoplasms	229758538
Chronic Respiratory Diseases	104605334
Lower Respiratory Infections	83770038
Neonatal Disorders	76860729
Diarrheal Diseases	66235508
Digestive Diseases	65638635
Tuberculosis	45850603
Cirrhosis and Other Chronic Liver Diseases	37479321
HIV/AIDS	36364419
Road Injuries	36296469
Diabetes Mellitus	31448872
Alzheimer's Disease and Other Dementias	29768839
Chronic Kidney Disease	28911692
Malaria	25342676
Self-harm	23713931
Nutritional Deficiencies	13792032
Interpersonal Violence	12752839
Protein-Energy Malnutrition	12031885
Meningitis	10524572
Drowning	10301999
Maternal Disorders	7727046
Parkinson's Disease	7179795
Alcohol Use Disorders	4819018
Acute Hepatitis	3784791
Fire, Heat, and Hot Substances	3602914
Conflict and Terrorism	3294053
Drug Use Disorders	2656121
Poisonings	2601082
Environmental Heat and Cold Exposure	1788851
Exposure to Forces of Nature	1490132

dtype: int64

Statistical analysis and identification of key trends

Global Causes of Death Distribution of the top causes of death.:

```
# Summing the causes of death across all countries and years
global_causes = df.drop(columns=['Country/Territory', 'Code', 'Year']).sum().sort_values(ascending=False)

# Plot the top 10 causes of death globally
plt.figure(figsize=(12,6))
sns.barplot(x=global_causes.index[:10], y=global_causes.values[:10], palette='Blues_d')
plt.xticks(rotation=90)
plt.title('Top 10 Global Causes of Death')
plt.ylabel('Total Deaths')
plt.xlabel('Cause of Death')
plt.show()
```

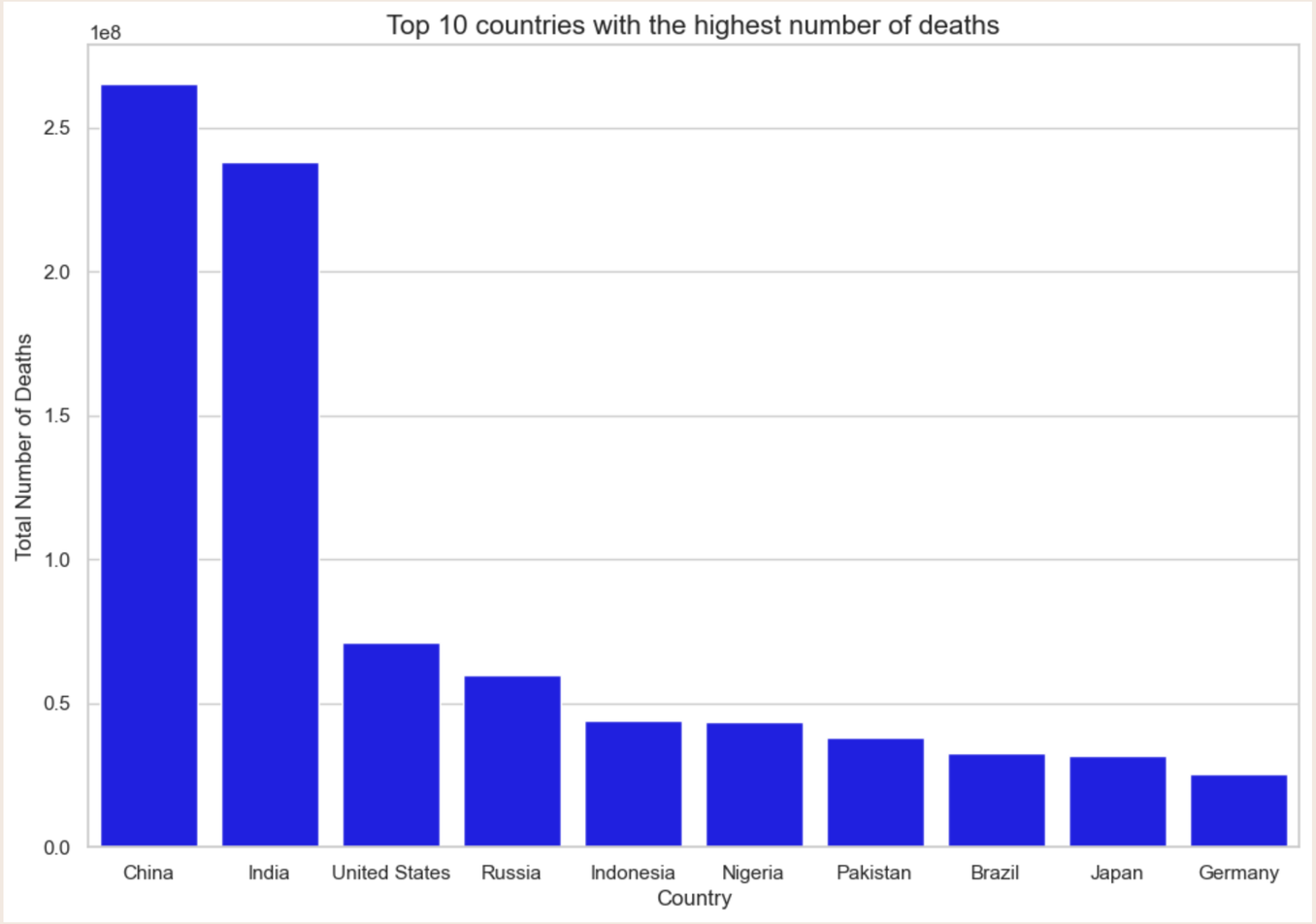


Statistical analysis and identification of key trends

Total Number of Deaths Around the World

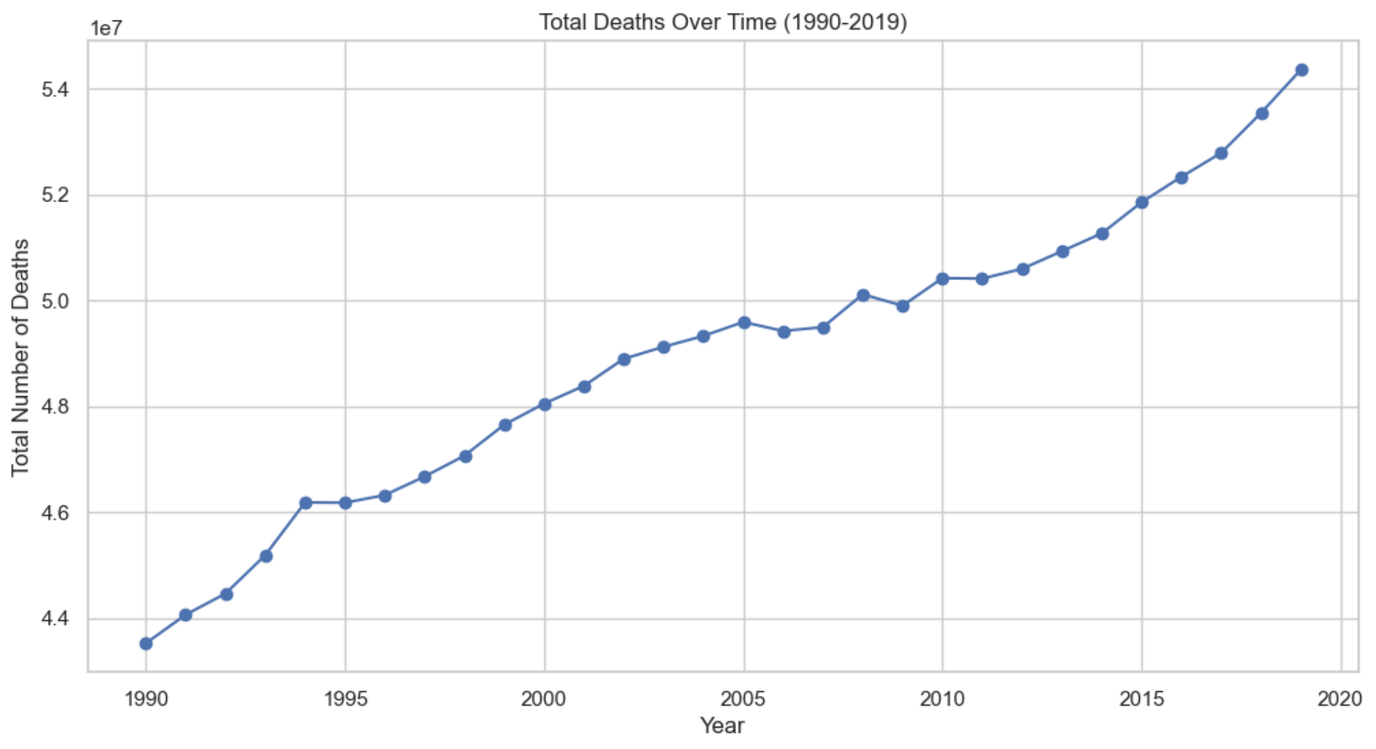
	Country/Territory	Total_no_of_Deaths
0	China	265408106
1	India	238158165
2	United States	71197802
3	Russia	59591155
4	Indonesia	44046941
...	...	...
199	Cook Islands	3999
200	Tuvalu	2962
201	Nauru	2249
202	Niue	591
203	Tokelau	299

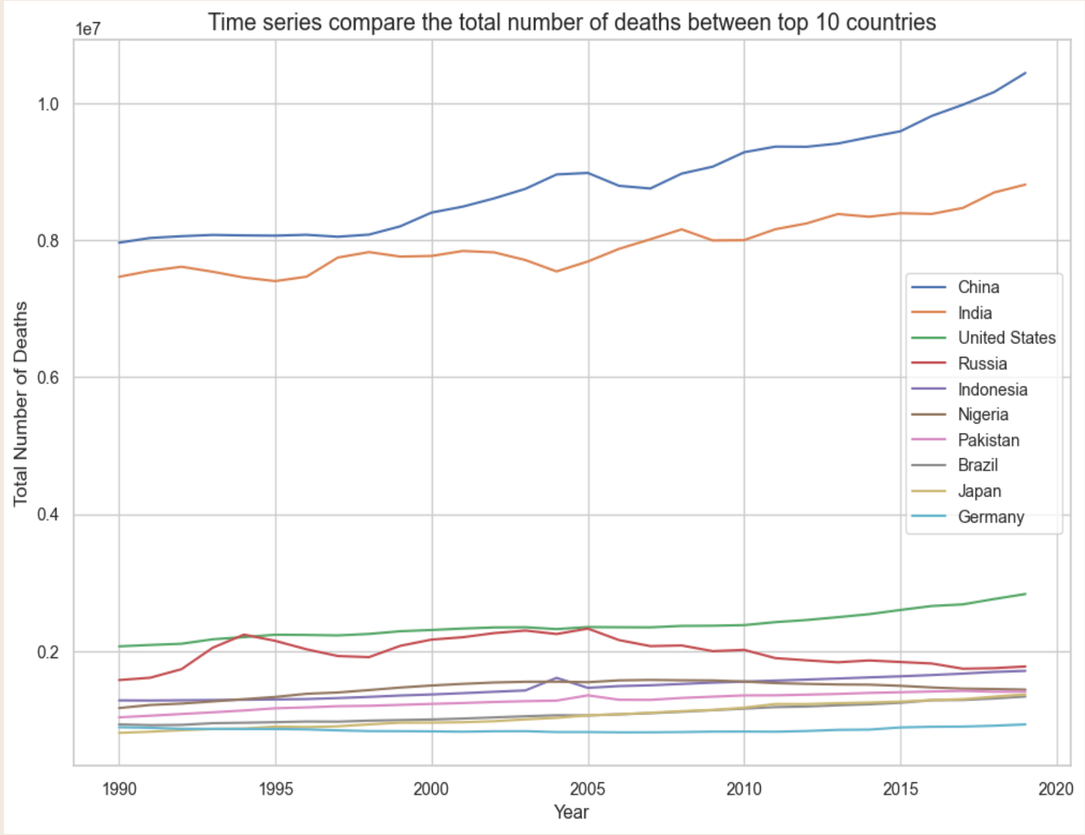
204 rows × 2 columns



Statistical analysis and identification of key trends

Evaluate Trends Over Time

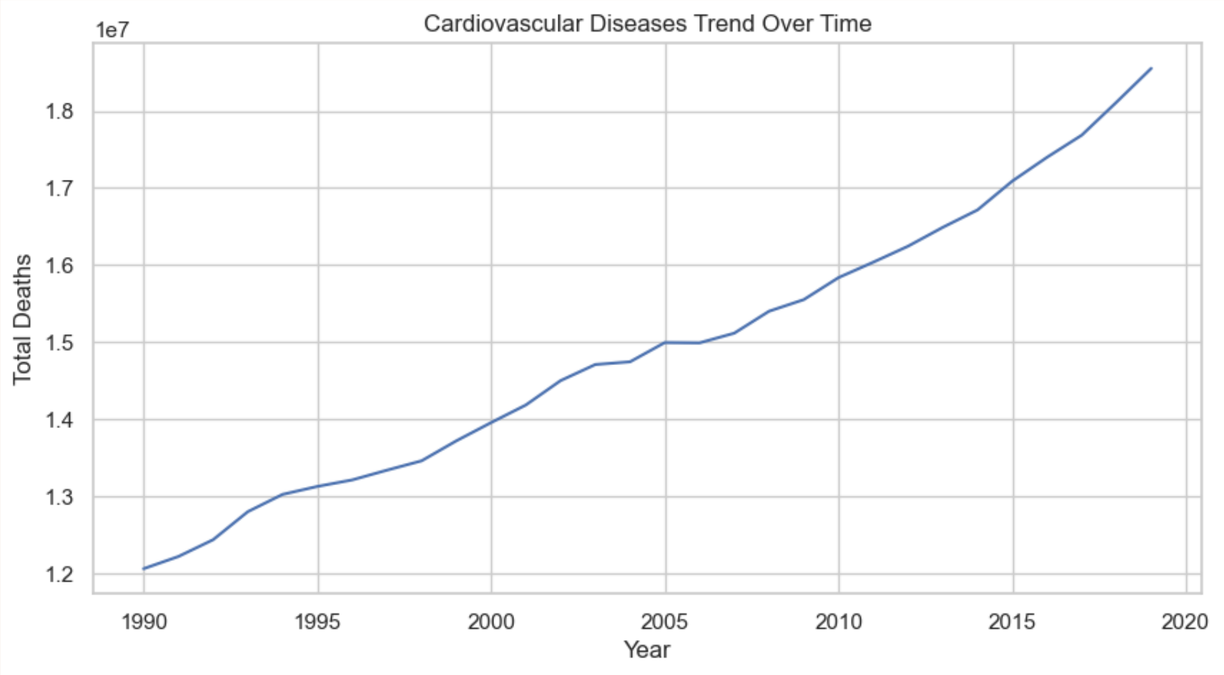




Statistical analysis and identification of key trends

Trend Analysis Over Time (how a specific cause of death (e.g., Cardiovascular Diseases) has changed over time.):

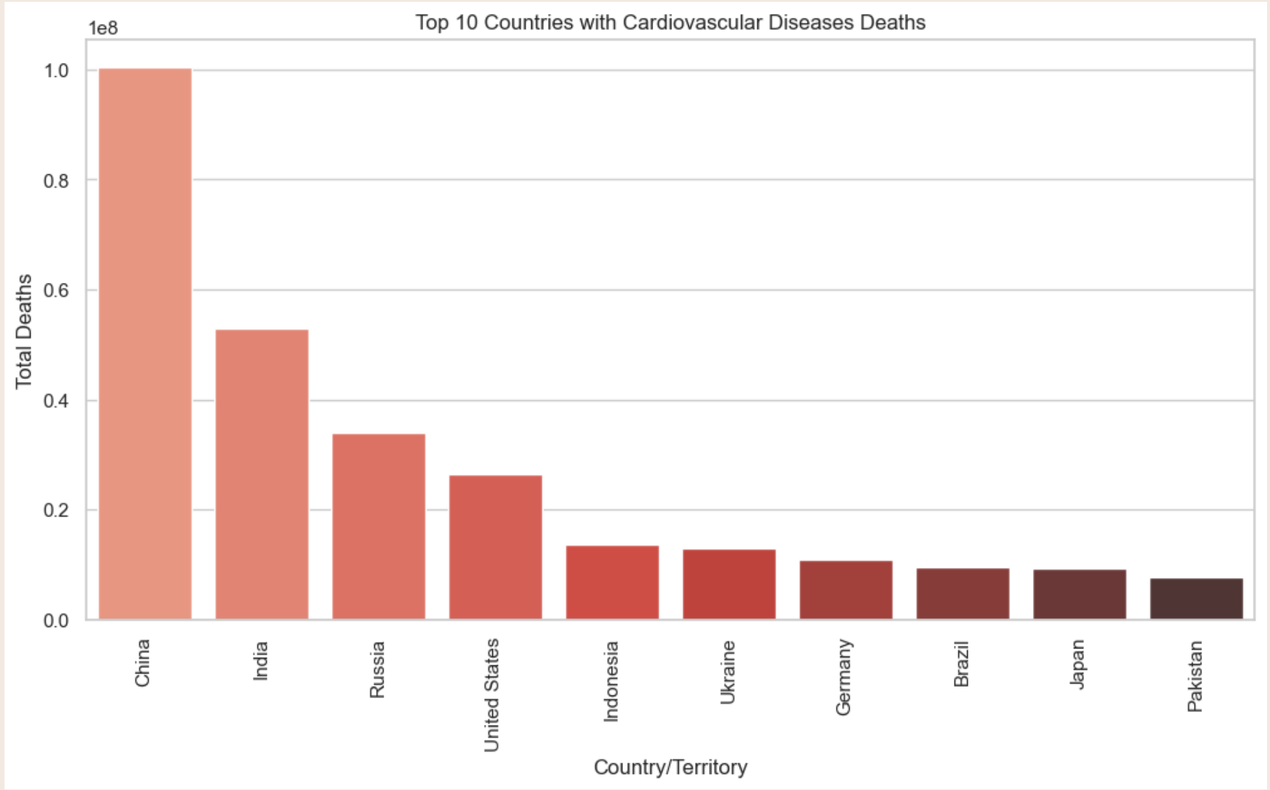
```
# Group by year and sum up deaths for Cardiovascular Diseases
cardio_trend = df.groupby('Year')['Cardiovascular Diseases'].sum()

# Plotting the trend
plt.figure(figsize=(10,5))
sns.lineplot(x=cardio_trend.index, y=cardio_trend.values)
plt.title('Cardiovascular Diseases Trend Over Time')
plt.ylabel('Total Deaths')
plt.xlabel('Year')
plt.show()
```



Statistical analysis and identification of key trends

Regional Comparison (compare regions by creating a bar chart to visualize how causes of death differ across different countries.):

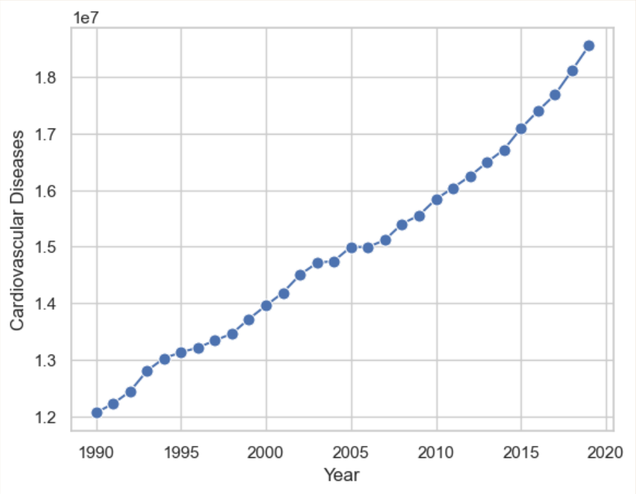
```
# Group by Country and sum up deaths for all causes
regional_comparison = df.groupby('Country/Territory').sum().drop(columns=['Year'])

# Sort by Cardiovascular Diseases
regional_comparison = regional_comparison.sort_values(by='Cardiovascular Diseases', ascending=False)

# Plotting the top 10 countries for Cardiovascular Diseases
plt.figure(figsize=(12,6))
sns.barplot(x=regional_comparison.index[:10], y=regional_comparison['Cardiovascular Diseases'][:10], palette='Reds_d')
plt.xticks(rotation=90)
plt.title('Top 10 Countries with Cardiovascular Diseases Deaths')
plt.ylabel('Total Deaths')
plt.xlabel('Country/Territory')
plt.show()
```

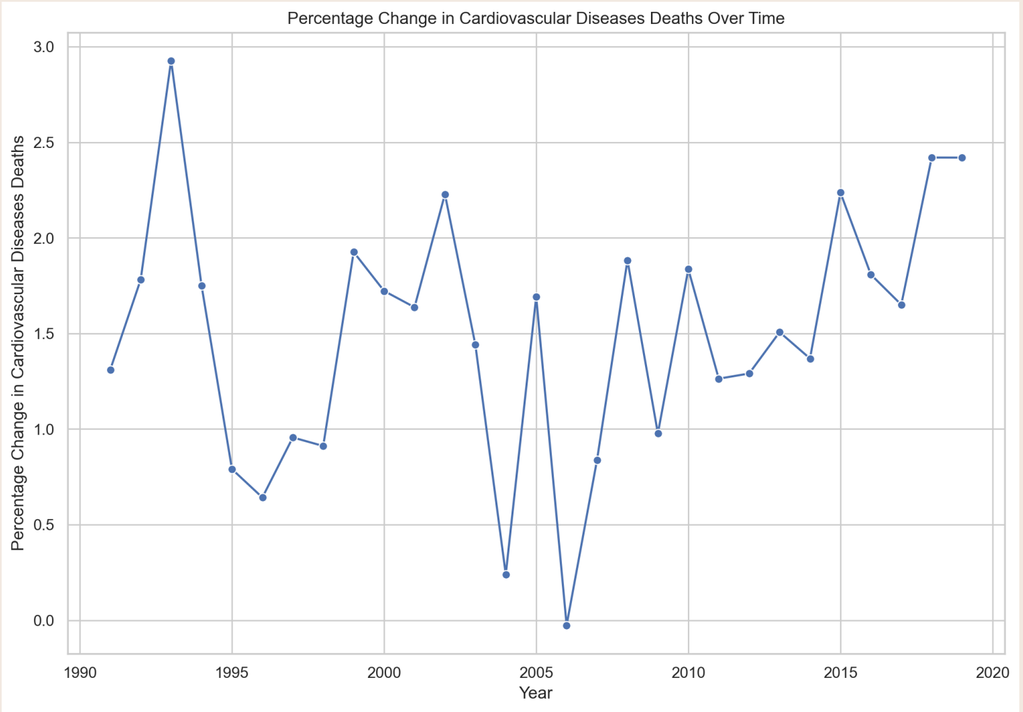


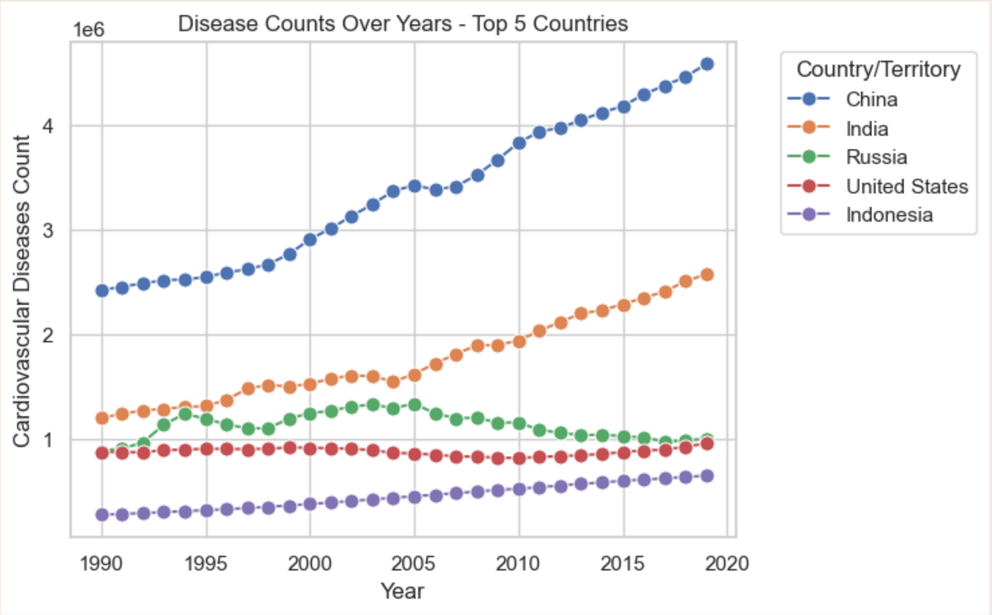
Data visualizations

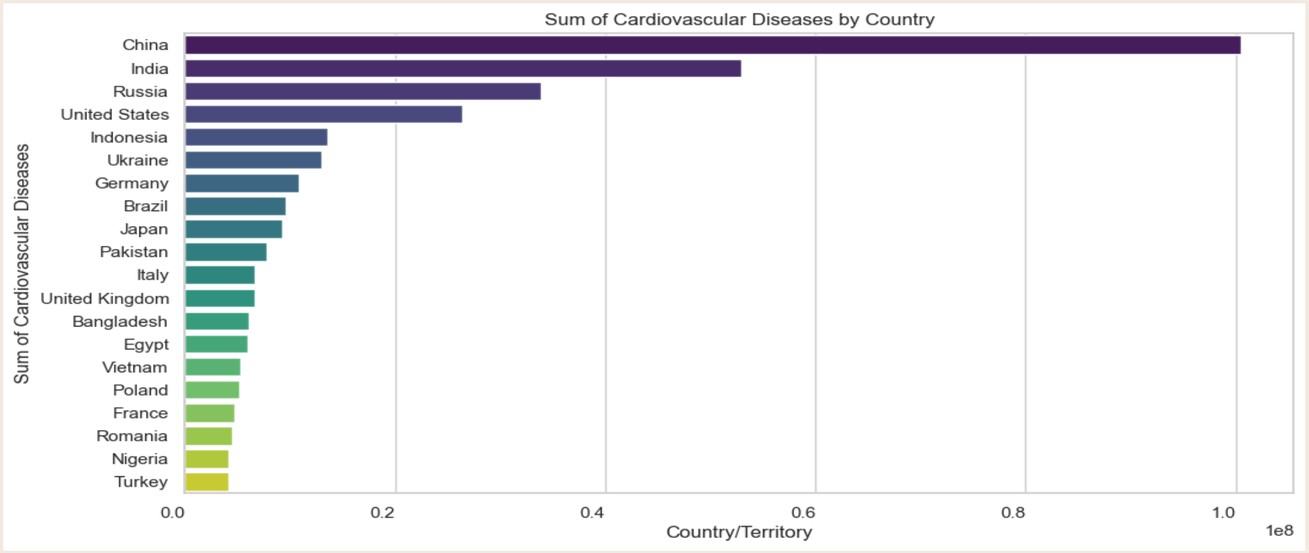
Cardiovascular Disease

Highest Cause of Death



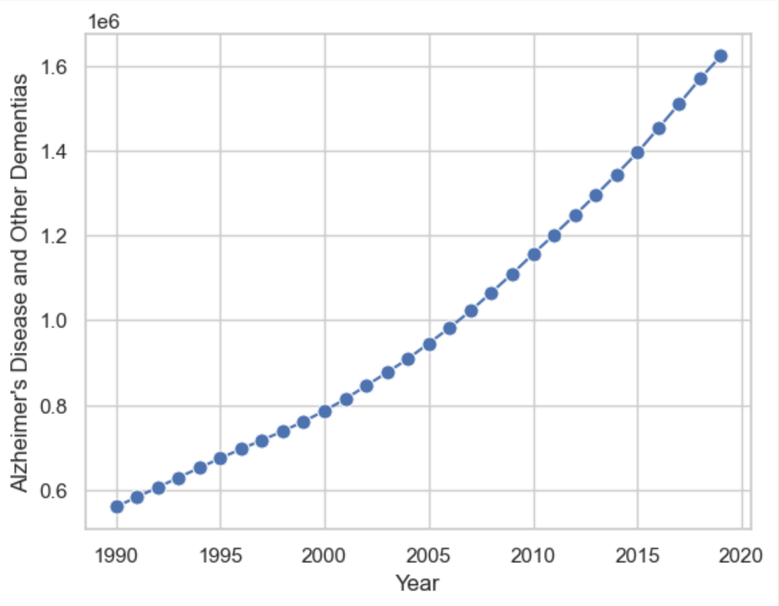






Data visualizations

Alzheimer's Disease and Other Dementias



What is the trend in the number of deaths caused by Alzheimer's disease and other dementias over the years?

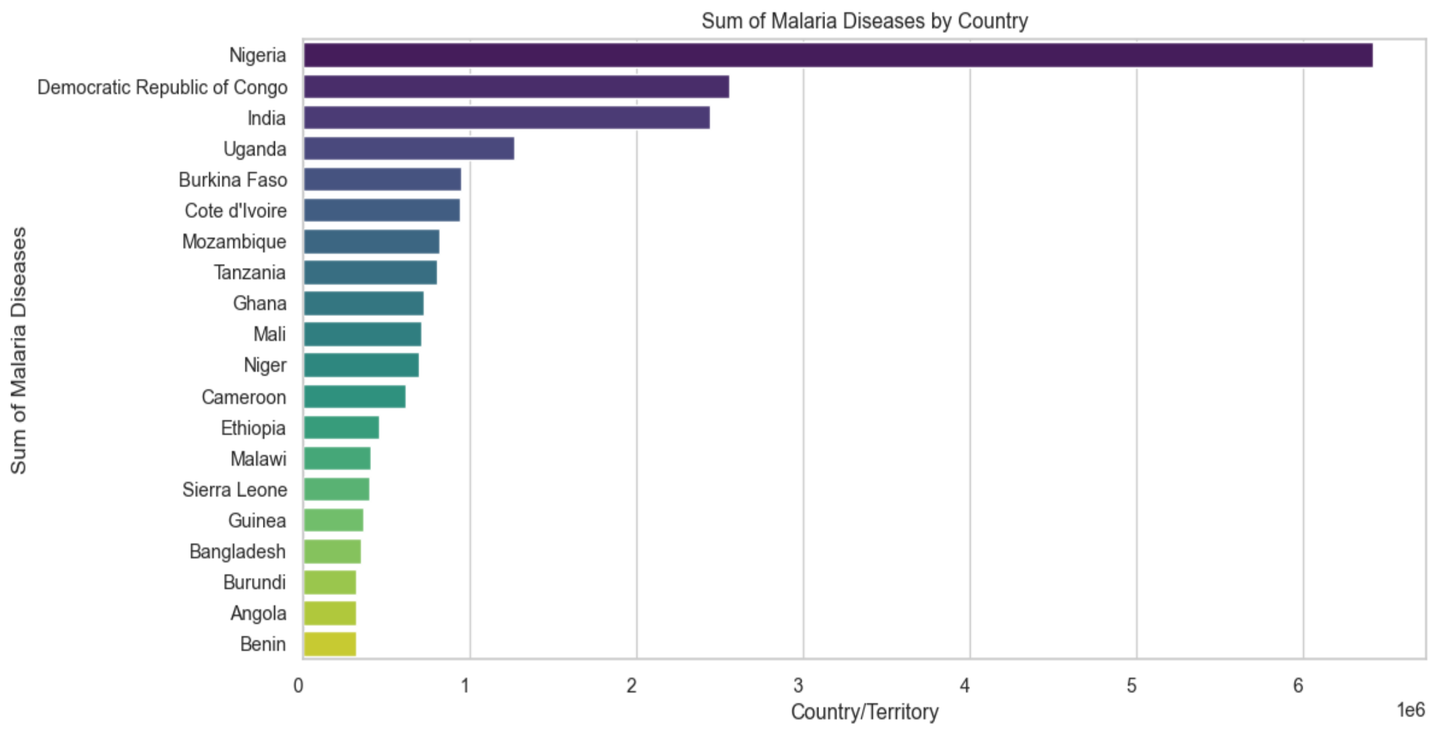
```
Dementias_trend = df.groupby("Year")["Alzheimer's Disease and Other Dementias"].sum().reset_index()
Dementias_trend
```

	Year	Alzheimer's Disease and Other Dementias
0	1990	560616
1	1991	583166
2	1992	605894
3	1993	629571
4	1994	652176
5	1995	674815
23	2013	1294701
24	2014	1343756
25	2015	1394942
26	2016	1451840
27	2017	1509646
28	2018	1568617
29	2019	1622426

Data visualizations

Malaria Disease

All coutries suffering from Malaria are in Africa except for India and Bangladesh, which makes sense



Which country has the highest number of deaths caused by malaria?

```
malaria_disease = df.groupby("Country/Territory")["Malaria"].sum().sort_values(ascending=False).head(20)
malaria_disease
```

```
Country/Territory
Nigeria                         6422063
Democratic Republic of Congo    2557219
India                           2439244
Uganda                          1265629
Burkina Faso                     950762
Cote d'Ivoire                    941597
Mozambique                       817948
Tanzania                         800490
Ghana                            721339
Mali                             711087
Niger                            693962
Cameroon                         614095
Ethiopia                         453985
Malawi                           404288
Sierra Leone                     394491
Guinea                           362660
Bangladesh                       349375
Burundi                          320767
Angola                           317069
Benin                            316834
Name: Malaria, dtype: int64
```

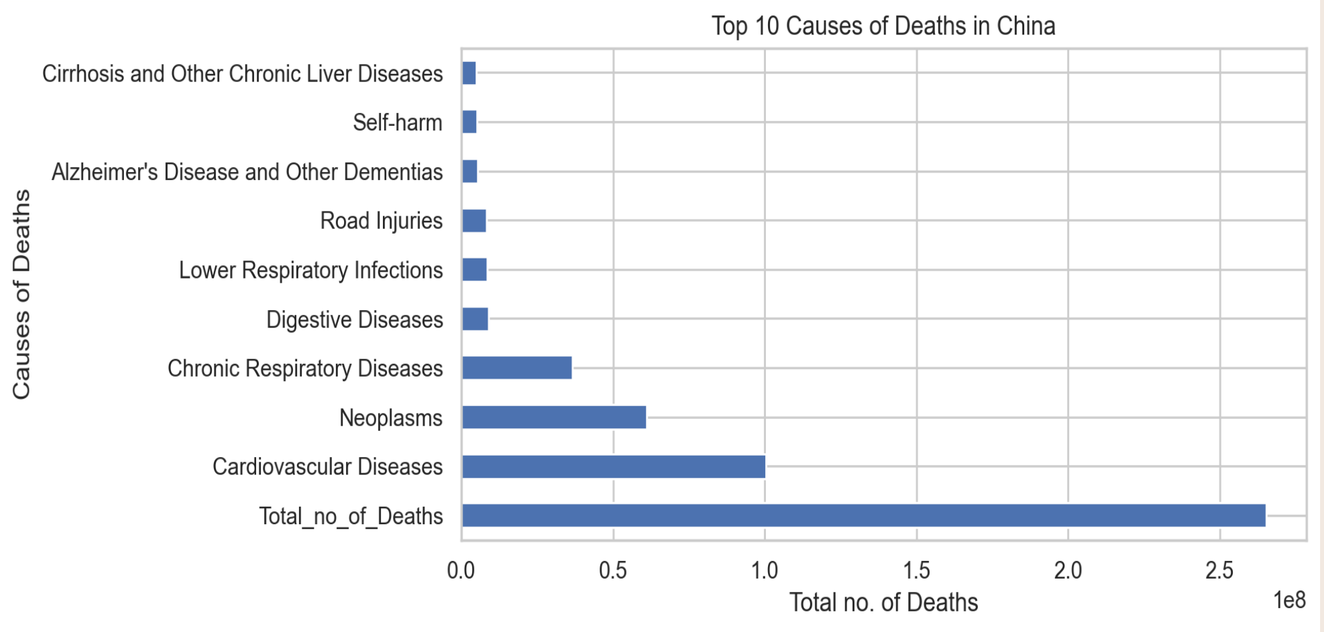
Data visualizations

China Causes of Death

China - Top 10 Causes of Deaths

Total_no_of_Deaths	265408106.00
Cardiovascular Diseases	100505973.00
Neoplasms	61060527.00
Chronic Respiratory Diseases	36676826.00
Digestive Diseases	8924906.00
Lower Respiratory Infections	8525819.00
Road Injuries	8350399.00
Alzheimer's Disease and Other Dementias	5381846.00
Self-harm	5078550.00
Cirrhosis and Other Chronic Liver Diseases	4918899.00

Name: China, dtype: float64

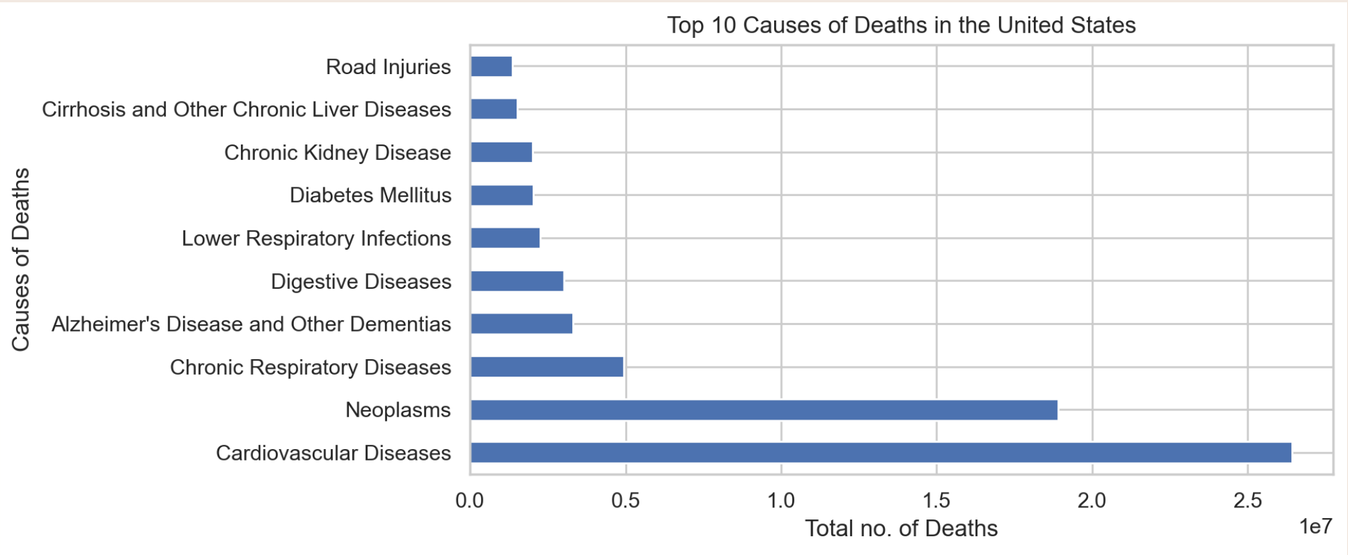


Data visualizations

US Causes of Death

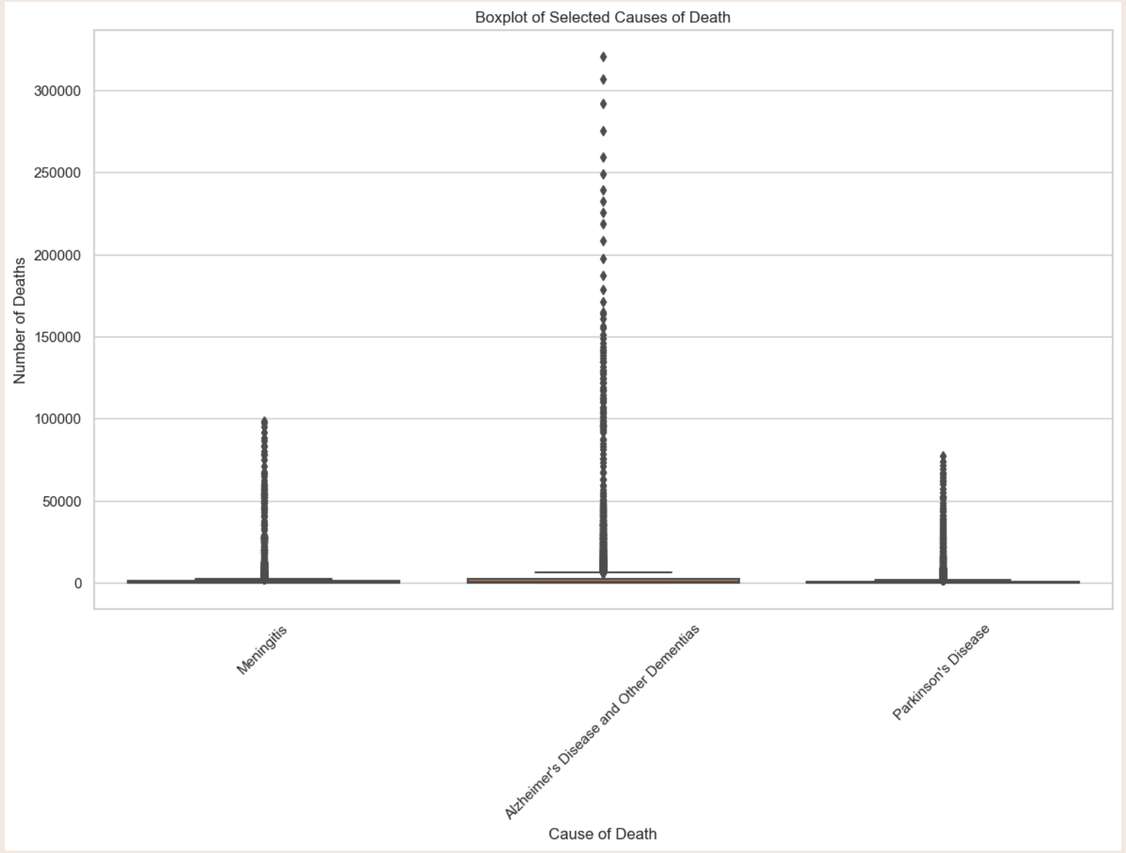
United States - Top 10 Causes of Deaths

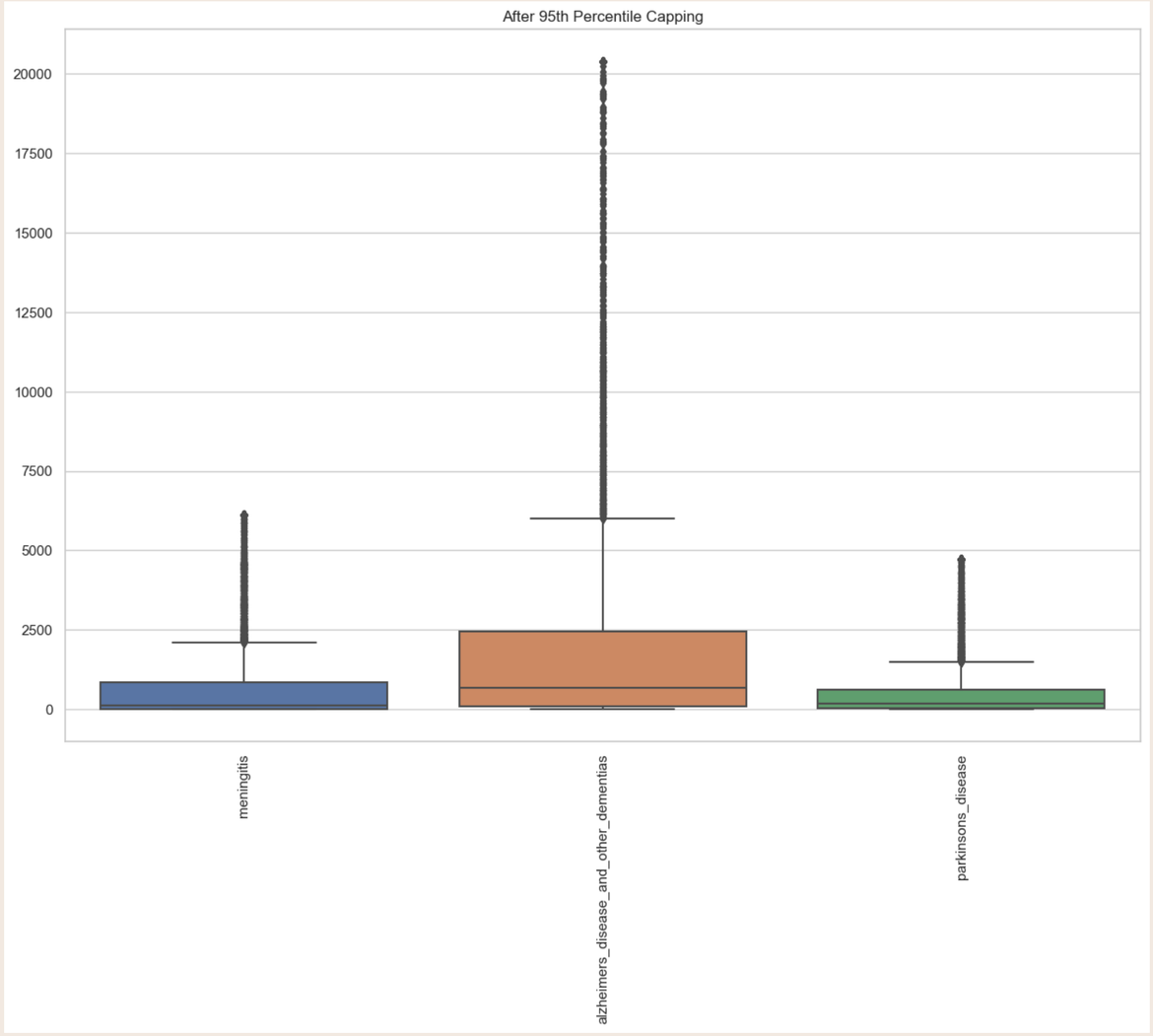
```
Cardiovascular Diseases                          26438346.00
Neoplasms                                        18905315.00
Chronic Respiratory Diseases                      4949052.00
Alzheimer's Disease and Other Dementias           3302609.00
Digestive Diseases                                3026943.00
Lower Respiratory Infections                      2248625.00
Diabetes Mellitus                                 2030631.00
Chronic Kidney Disease                            2018497.00
Cirrhosis and Other Chronic Liver Diseases        1514325.00
Road Injuries                                     1359744.00
Name: United States, dtype: float64
```



Data visualizations

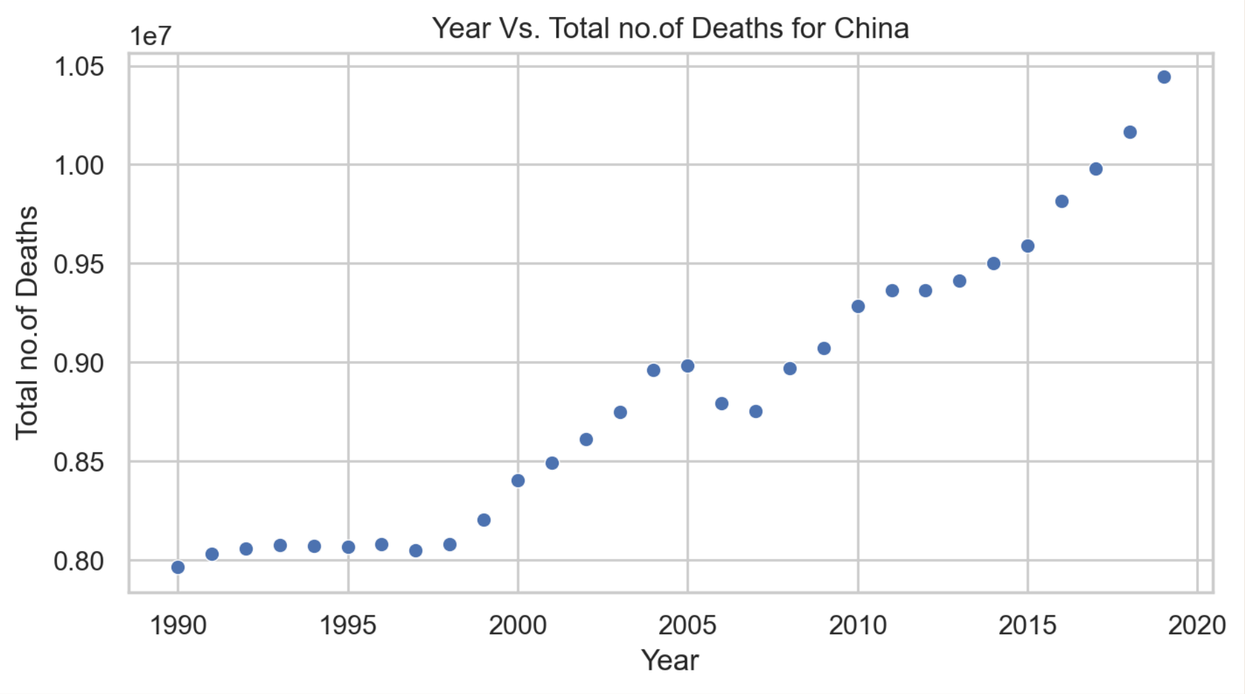
Random Visualizations

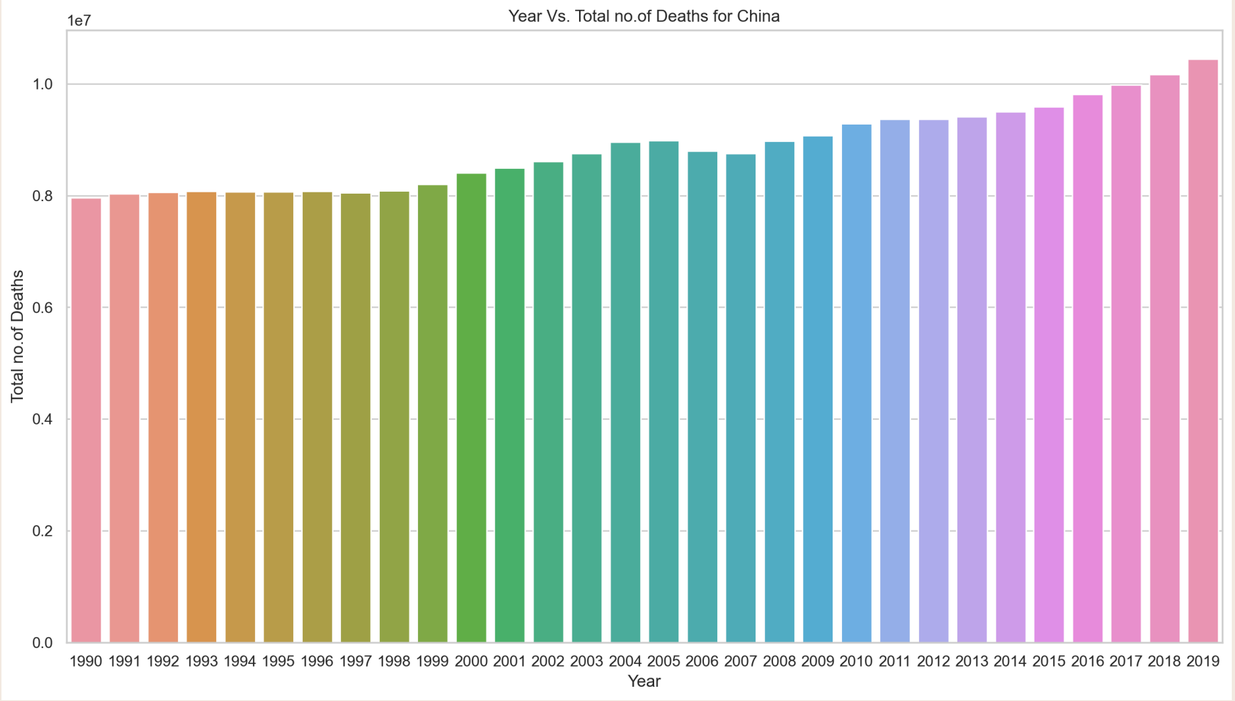




Data visualizations

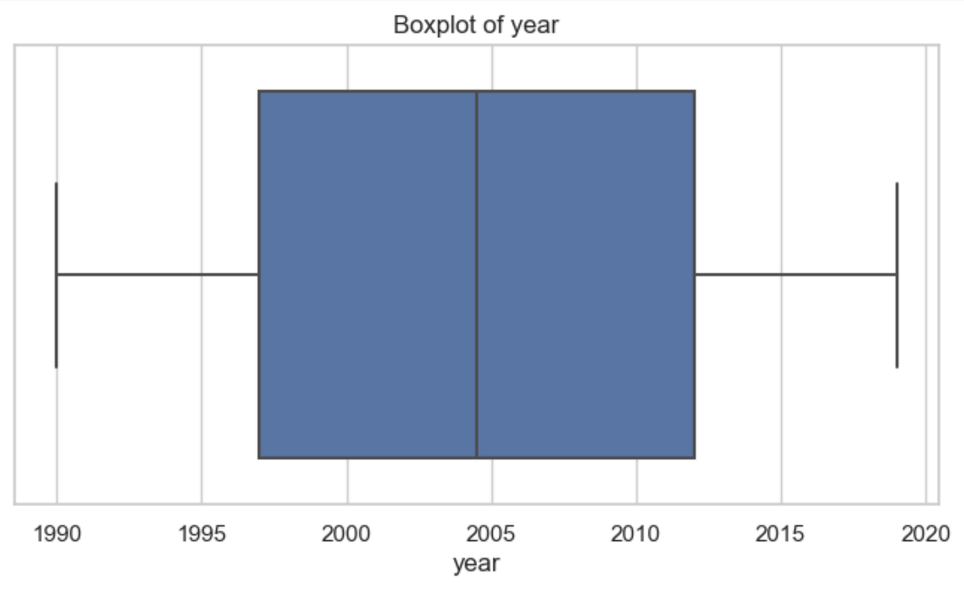
Random Visualizations

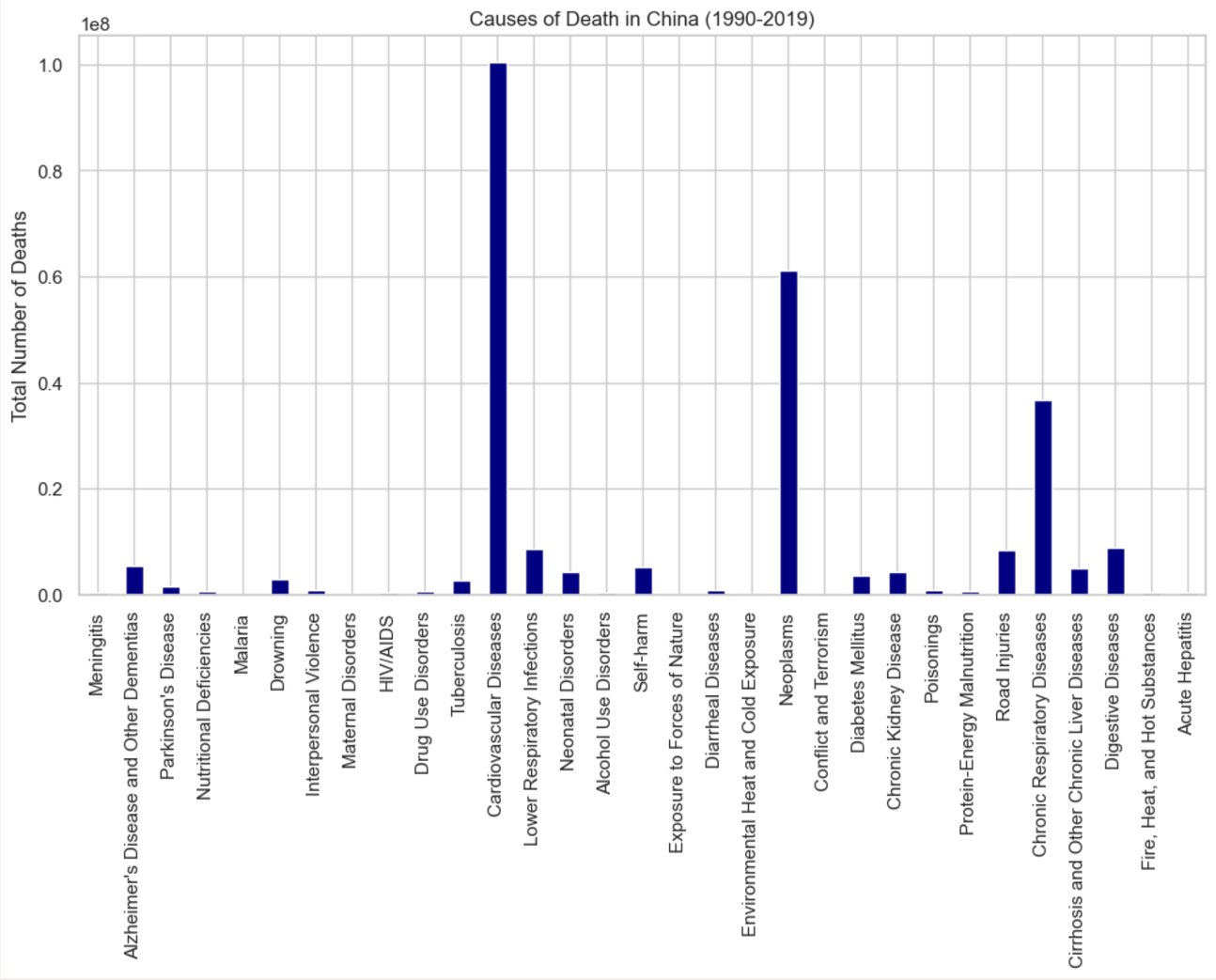




Data visualizations

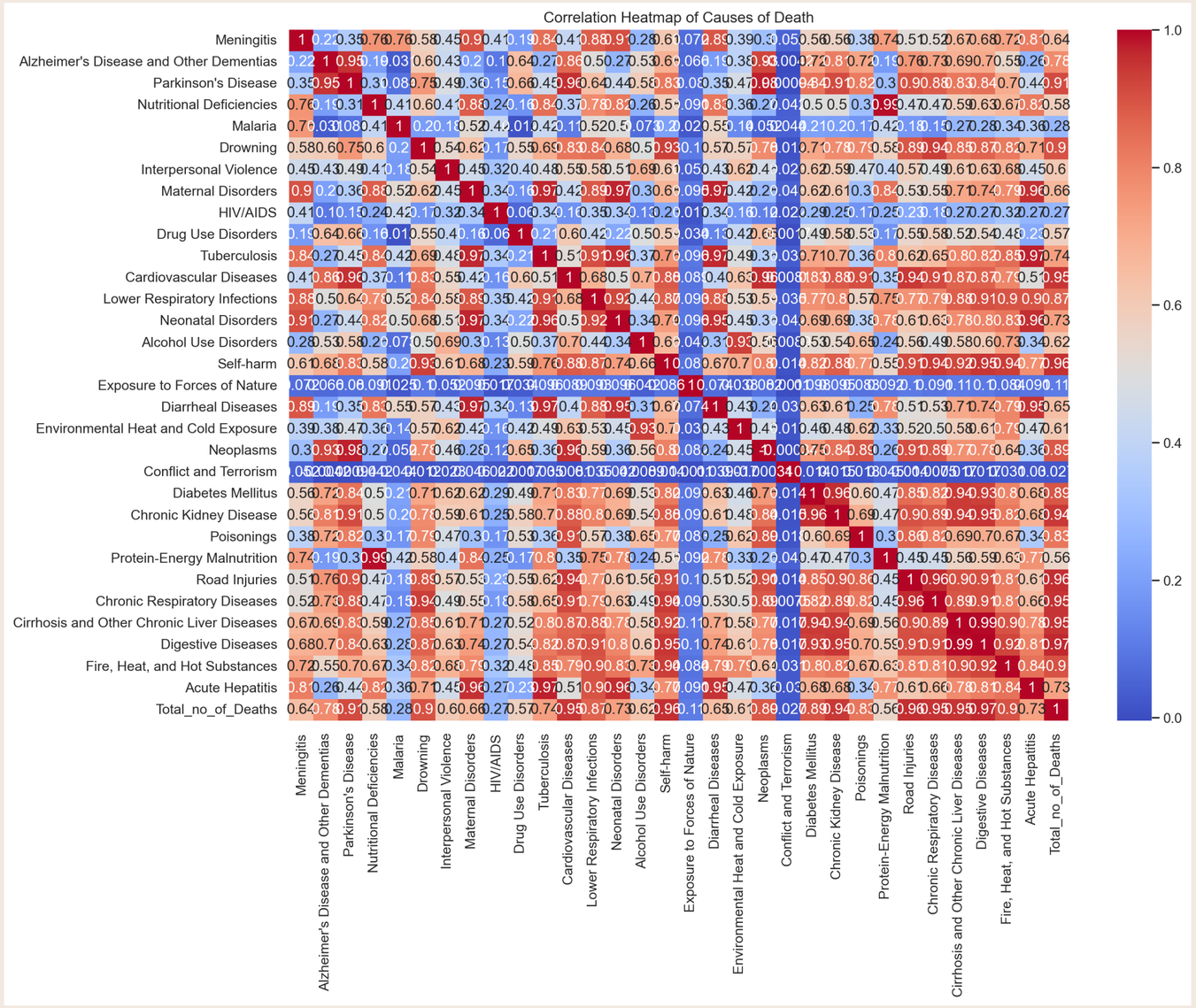
Random Visualizations





Data visualizations

Random Visualizations



Applied Machine Learning Techniques

- LinearRegression
- Ridge
- Lasso
- RandomForestRegressor
- GradientBoostingRegressor

```
Ridge Regression Mean Squared Error: 75.21693165381222
Ridge Regression R-squared: 0.9999999998717695
Lasso Regression Mean Squared Error: 2109596.3646646277
Lasso Regression R-squared: 0.9999964035408353
Random Forest Mean Squared Error: 412402826.7590321
Random Forest R-squared: 0.9992969318914813
Gradient Boosting Mean Squared Error: 607484604.8760853
Gradient Boosting R-squared: 0.9989643546930536
Cross-Validated Mean Squared Error (Ridge): 2917.6510757034894
Cross-Validated Mean Squared Error (Lasso): 24126655.184786927
```

```
Gradient Boosting Mean Squared Error: 646083812.675353
Gradient Boosting R-squared: 0.9988985504107915
```

Observations:

1.Ridge Regression and Lasso Regression both show exceptionally high R^2 values, indicating that they explain almost all the variance in the data. However, Lasso Regression has a much higher MSE, which might suggest it is less suited for this dataset compared to Ridge Regression.

2.Random Forest and Gradient Boosting have lower R^2 values compared to Ridge and Lasso, and their MSE values are significantly higher. This suggests that, while these models are capturing some patterns, they may be overfitting or may not be as well-tuned for this specific problem.

3.The cross-validated MSE values for Ridge and Lasso show a significant difference from the training MSE values, indicating that while these models perform very well on training data, their performance on unseen data might be less impressive.

Applied Machine Learning Techniques

- Hyperparameter tuning using GridSearchCV
- Feature Engineering: Remove irrelevant or redundant features to improve model performance. You can use techniques like Recursive Feature Elimination (RFE)
- Cross-Validation: Perform cross-validation to ensure that the model is not overfitting and performs well on unseen data
- Detecting Outliers using IQR

```
Tuned Random Forest Mean Squared Error: 423071262.43593276
Tuned Random Forest R-squared: 0.9992787442448273
Tuned Gradient Boosting Mean Squared Error: 351335976.6355588
Tuned Gradient Boosting R-squared: 0.9994010392157373
```

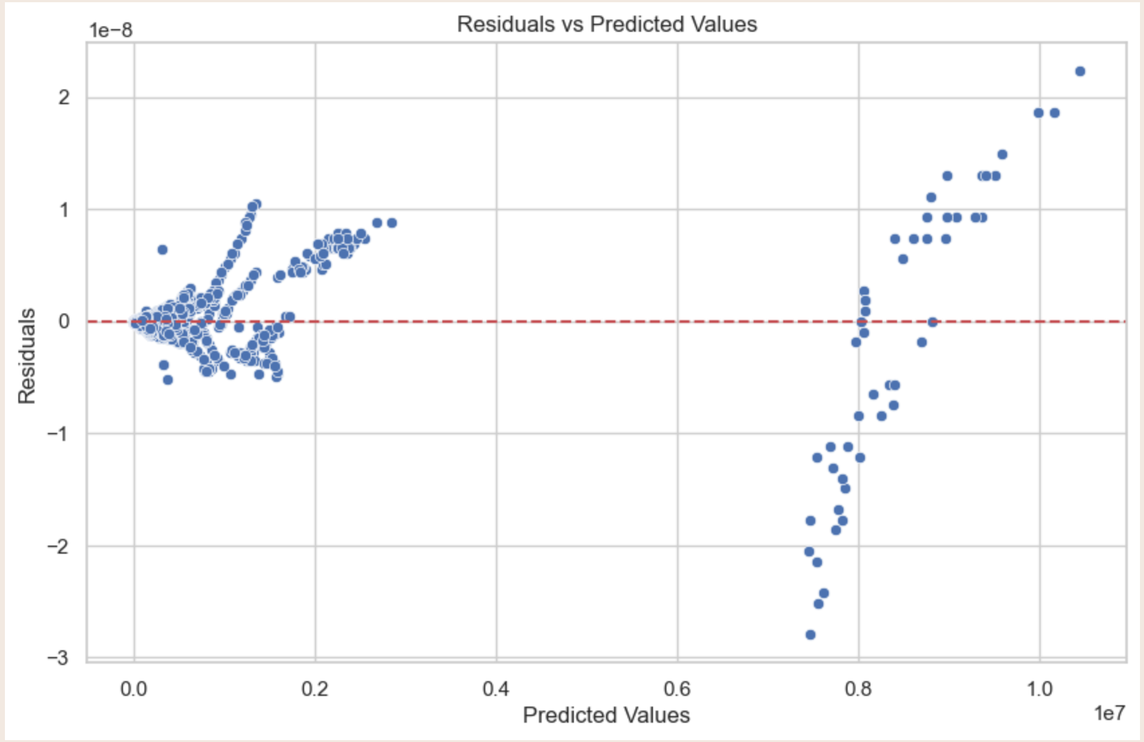
```
Ridge with RFE Mean Squared Error: 1016822564.5699499
Ridge with RFE R-squared: 0.9982665115979211
```

```
Cross-Validated Mean Squared Error (Random Forest): 491233682.1143919
Cross-Validated Mean Squared Error (Gradient Boosting): 320611482.3465633
```

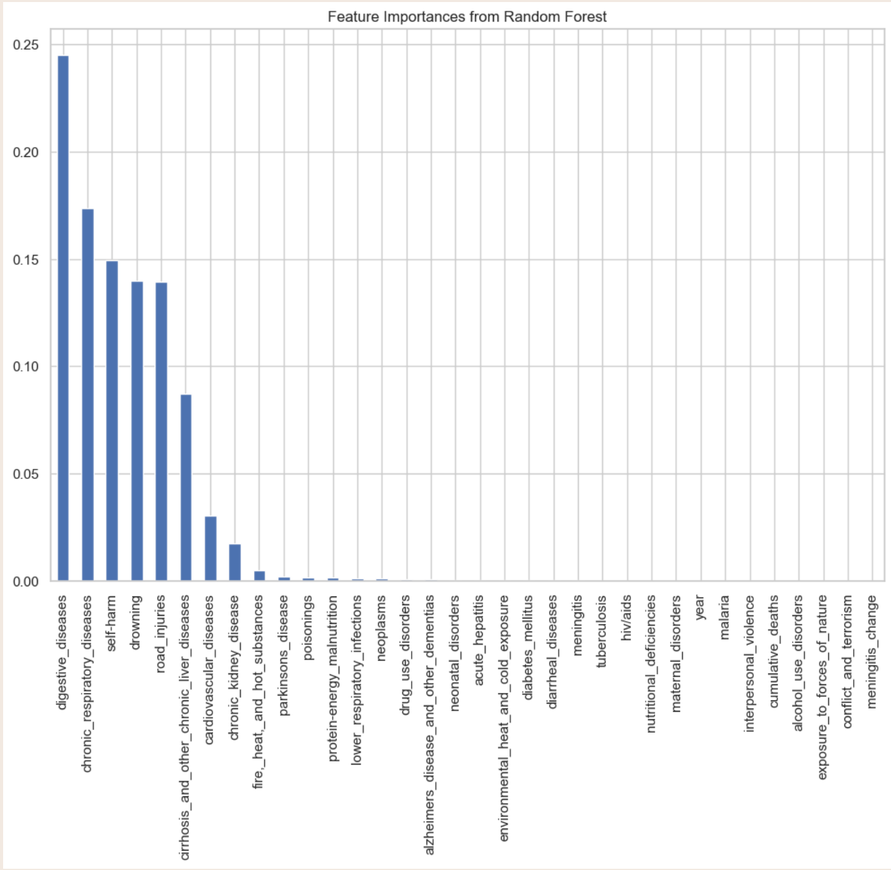
```
Linear Regression Mean Cross-Validated MSE: 6.117124717691846e-14
Random Forest Mean Cross-Validated MSE: 71124482809.29434
```

Applied Machine Learning Techniques

- Linear Regression

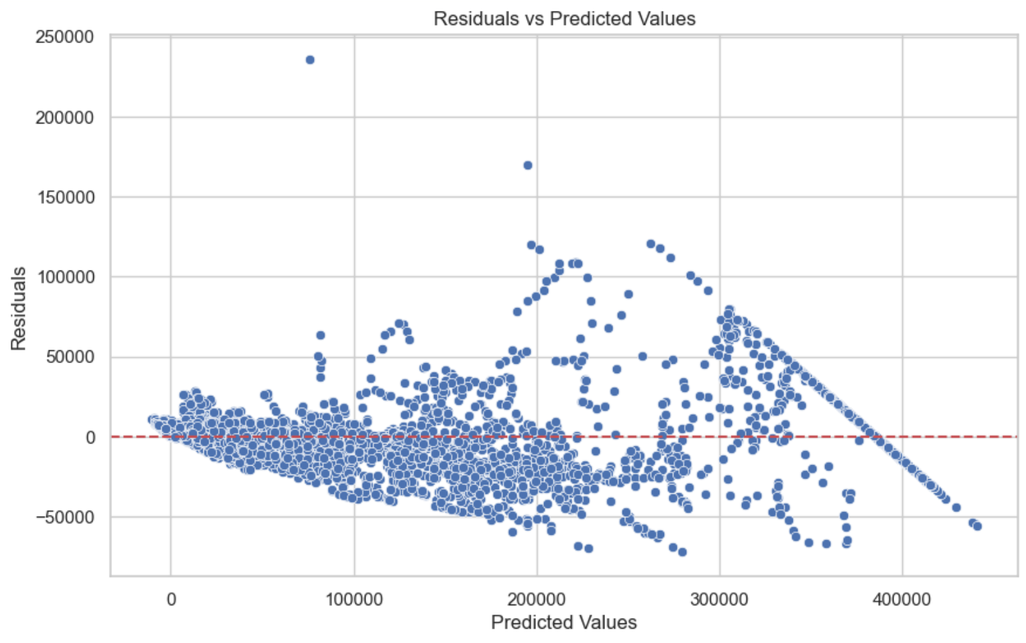


- Random Forest



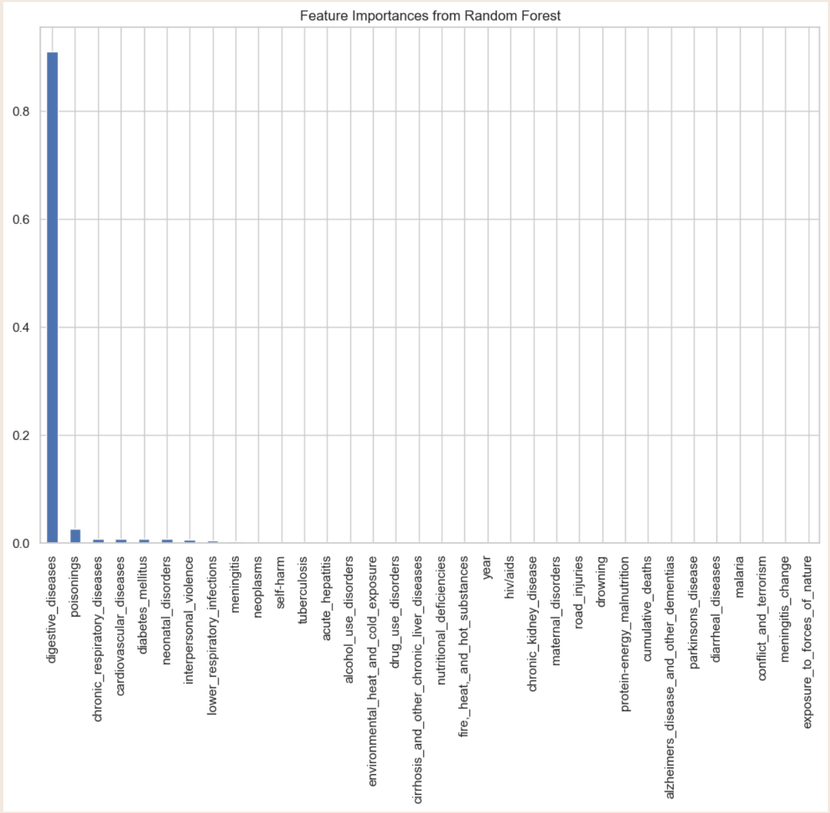
Applied Machine Learning Techniques

- Detecting Outliers using IQR
- Handling Outliers
- Re-modelling
- Applying Linear Regression to predict future trends and Random Forest
- Cross-Validation: Use cross-validation to assess model performance more reliably
- Model Diagnostics: Examine residuals for Linear Regression to ensure they are randomly dispersed



Applied Machine Learning Techniques

- Feature Importance (Random Forest): Evaluate feature importance to understand which features impact the model predictions

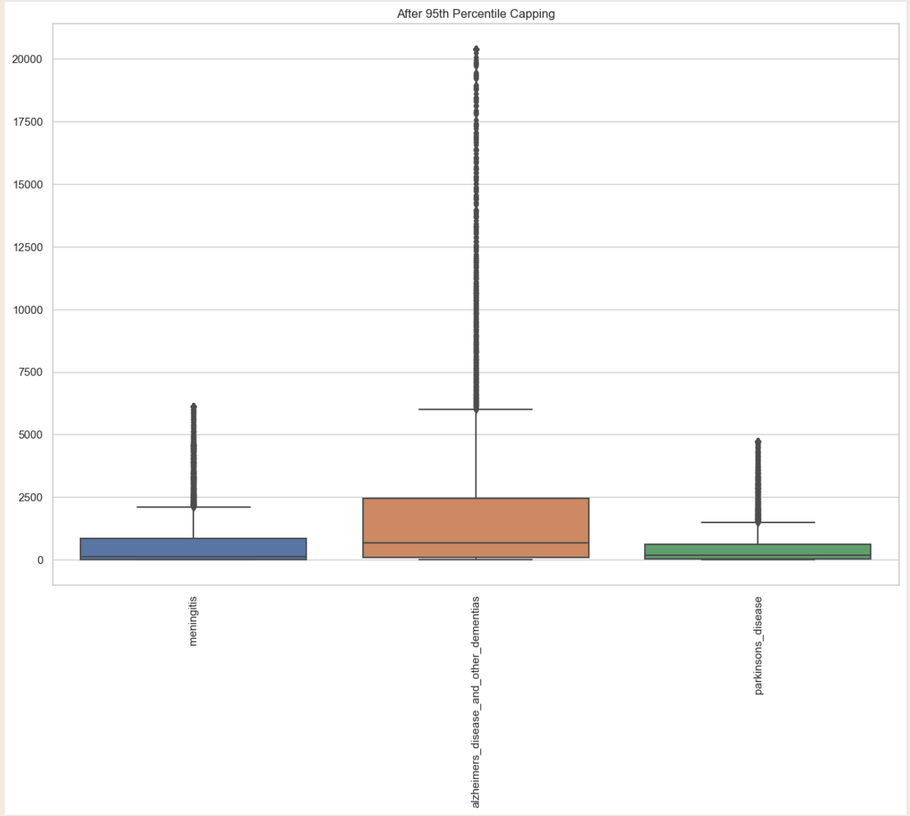


- Evaluate on Test Data

```
Linear Regression Mean Squared Error: 494300444.4112894
Linear Regression R-squared: 0.9708188257223398
Random Forest Mean Squared Error: 131796235.34221615
Random Forest R-squared: 0.9922193699072206
```

Applied Machine Learning Techniques

- Robust Scaling: Use RobustScaler to scale features in a way that's robust to outliers.
- Explore Different Thresholds: Apply capping at different percentiles (e.g., 95th percentile) to see if it’s a better fit.



Observations

- All coutries suffering from Malaria are in Africa except for India and Bangladesh, which makes sense.
- The high mortality rate in Nigeria is a result of its weak health systems and poverty.
- CHINA , INDIA and USA face the largest brunt of deaths due to diseases in the world Cardiovascular diseases , Neoplasms (Malignancy/Cancer) and Lower Respiratory Tract Infections (for example : Pneumonia) are the top 3 killer disases in the world.
- Drop in Nutritional Deficiencies Deaths recorded in China in 2007 and from 2008 the count of deaths again started to raise.
- Rapid drop in Malaria Deaths recorded in China after 1999.
- Top 10 countries in deaths: China, India, United States, Russia, Indonesia, Nigeria, Pakistan, Brazil, Japan, and Germany.
- The number of deaths is increasing over the years from 1990 - 2019.

Observations

- Top 10 Global Causes of deaths: Cardiovascular Diseases, Neoplasms, Chronic Respiratory Diseases, Lower Respiratory Infections, Neonatal Disorders, Diarrheal Diseases, Digestive Diseases, Tuberculosis, Cirrhosis and Other Chronic Liver Diseases, and HIV/AIDS.
- Heart and Circulatory Diseases, Tumors and Respiratory Diseases constitutes more than 60% of the total number of deaths around the world.
- Top 10 cause of deaths in Egypt: Cardiovascular Diseases, Digestive Diseases, Cirrhosis and Other Chronic Liver Diseases, Neoplasms, Lower Respiratory Infections, Road Injuries, Chronic Respiratory Diseases, Neonatal Disorders, Diarrheal Diseases, and Chronic Kidney Disease.
- Conflict and Terrorism has an Outlier at year 1994.

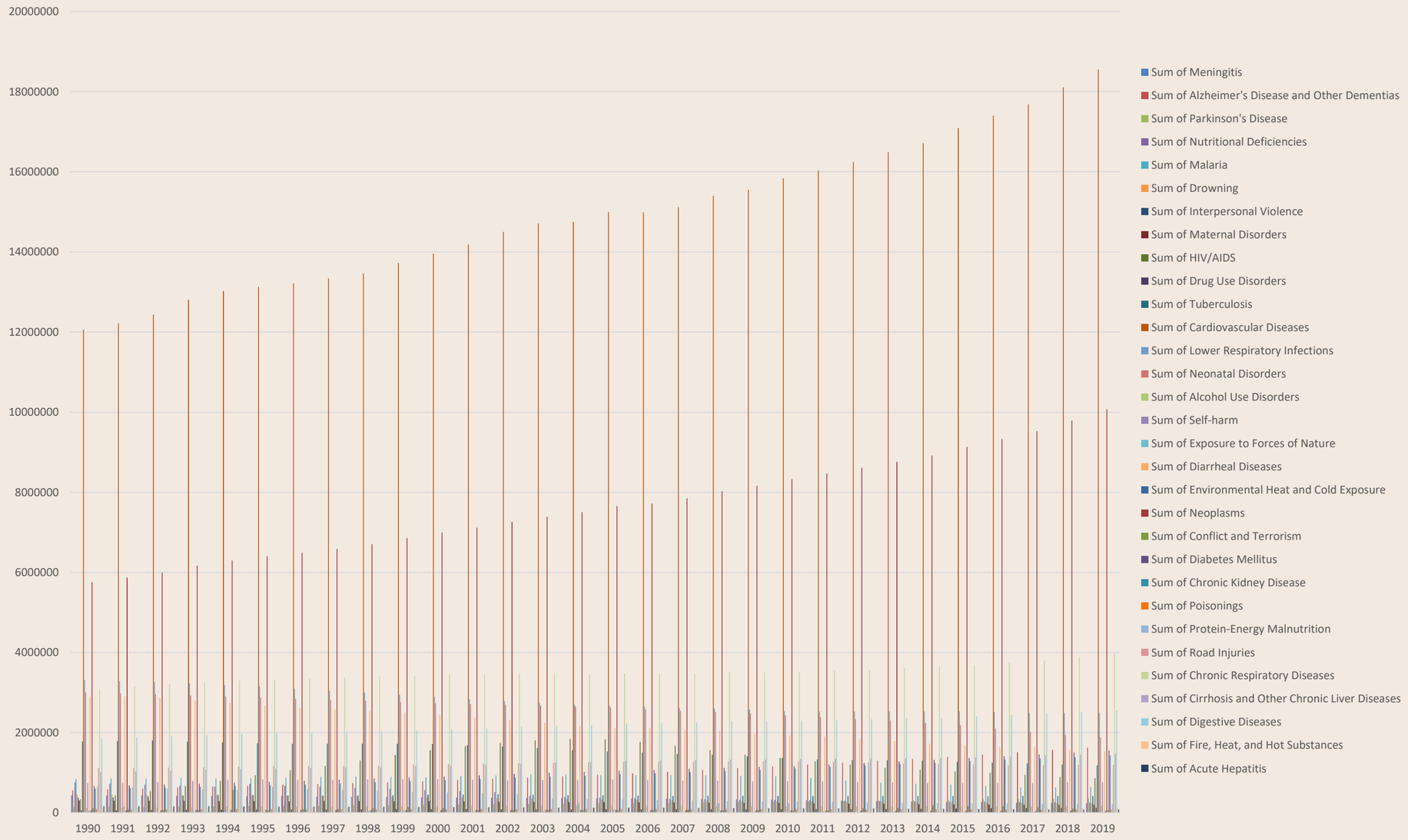
Excel Analysis

Content Map:

- Visual Charts
- Observations

Visual Charts

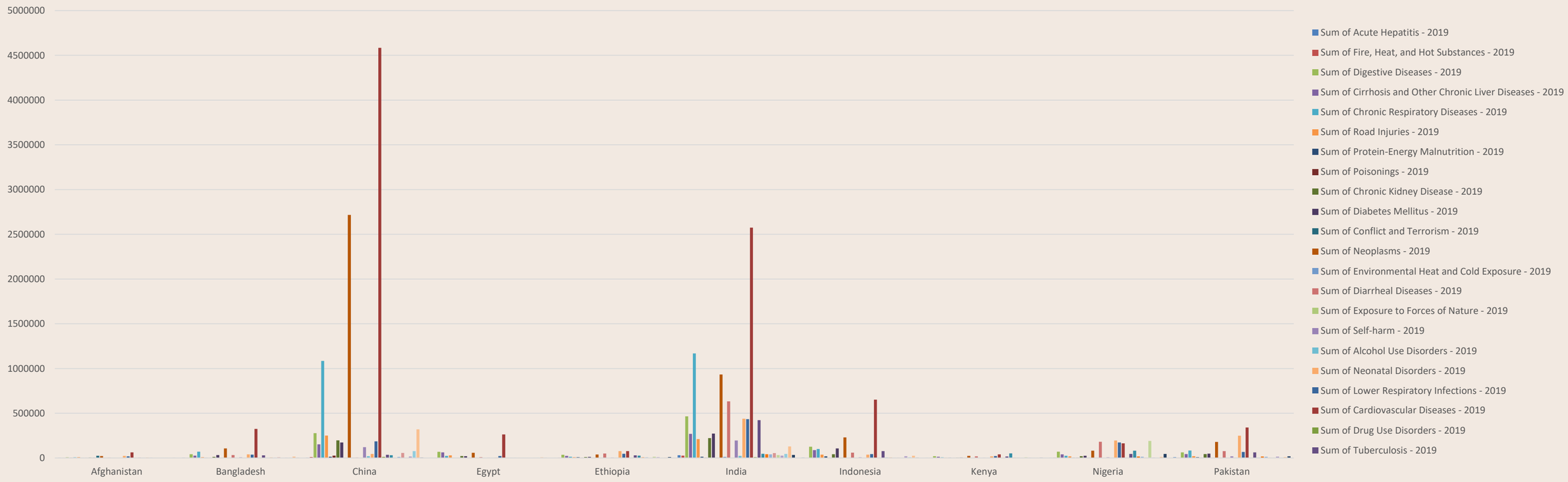
Year against Causes



Visual Charts

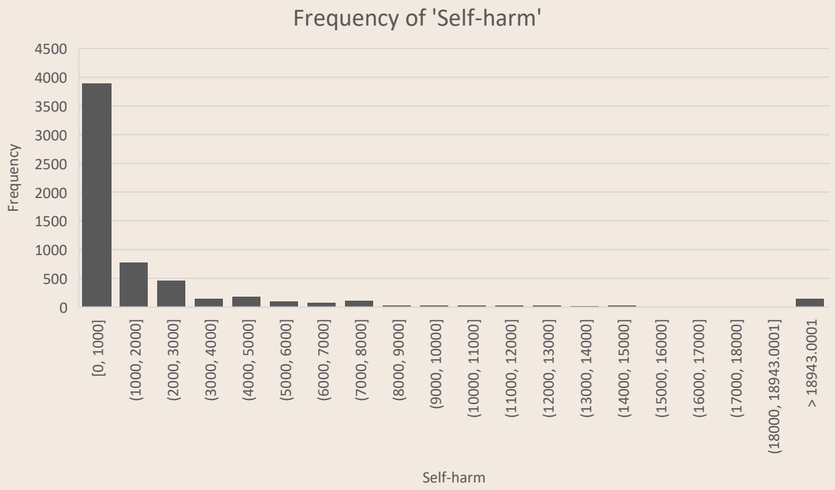
Countries against Causes in year 2019

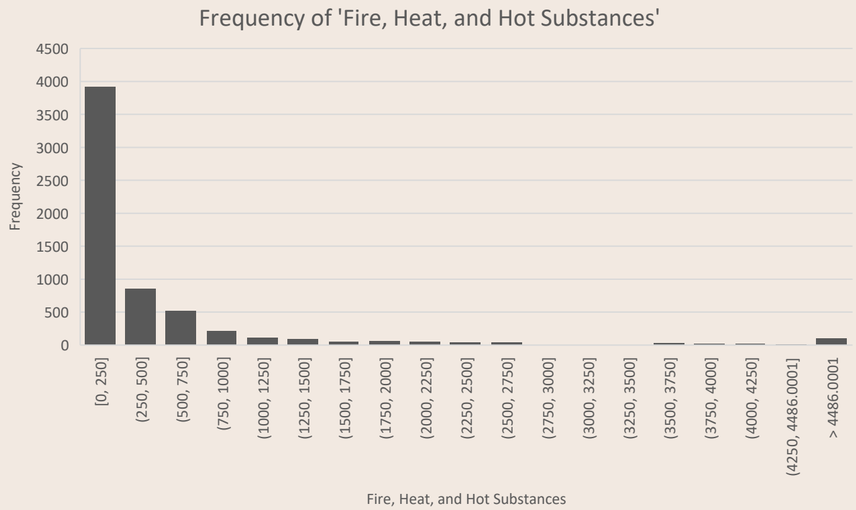
You can change the year as you like from the excel file.

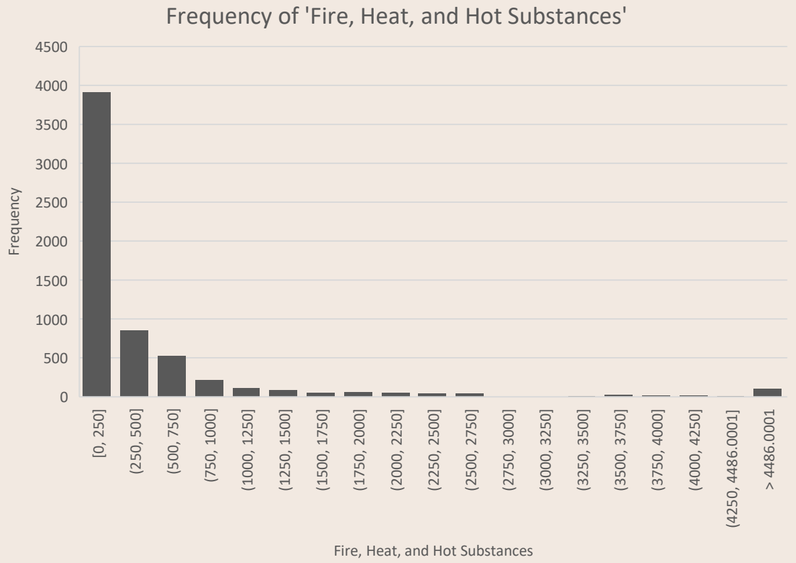


Visual Charts

Random Causes of Deaths Frequency

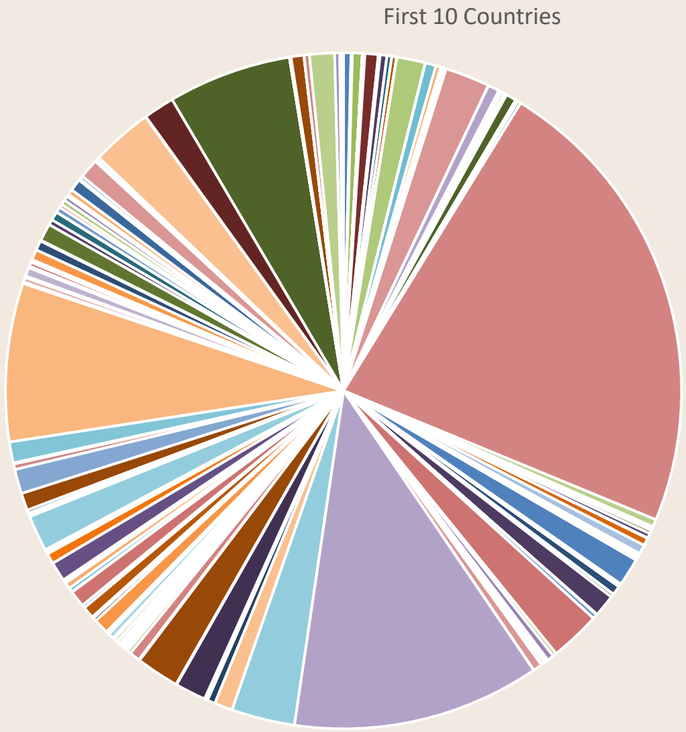


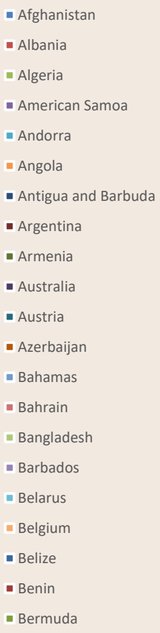


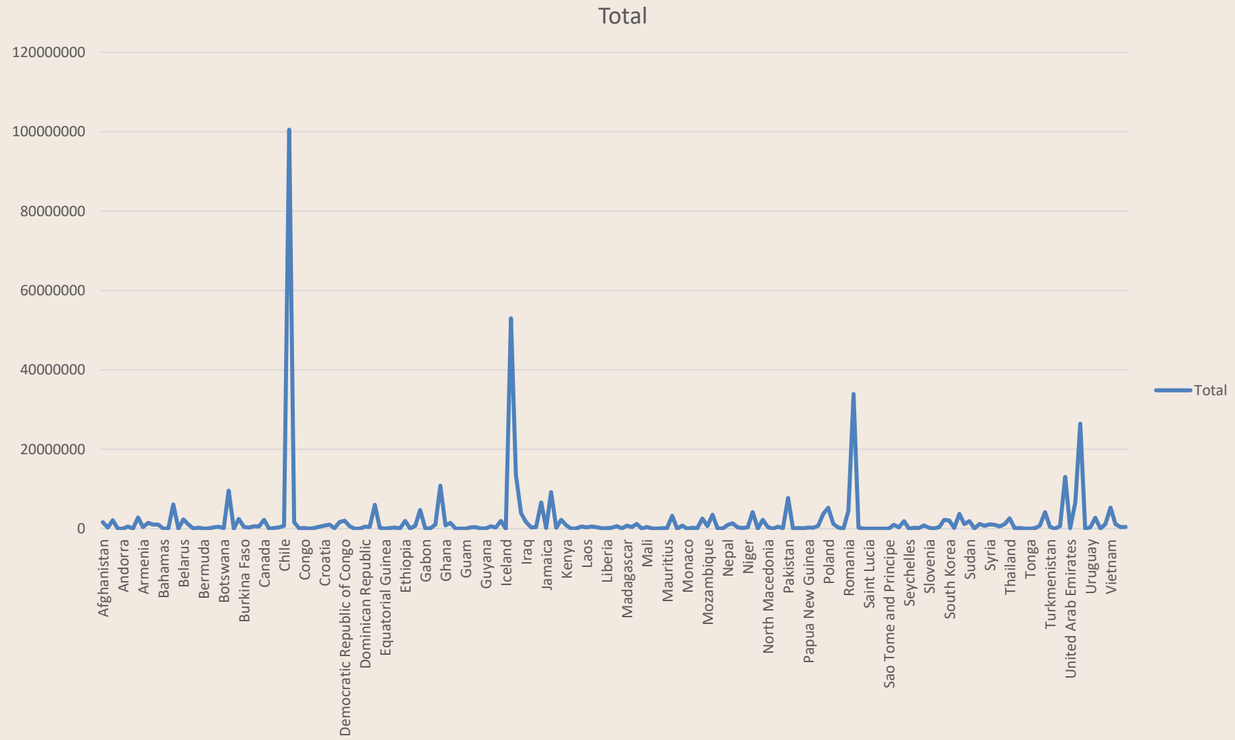


Visual Charts

First 10 Countries in the list regarding Cardiovascular disease.





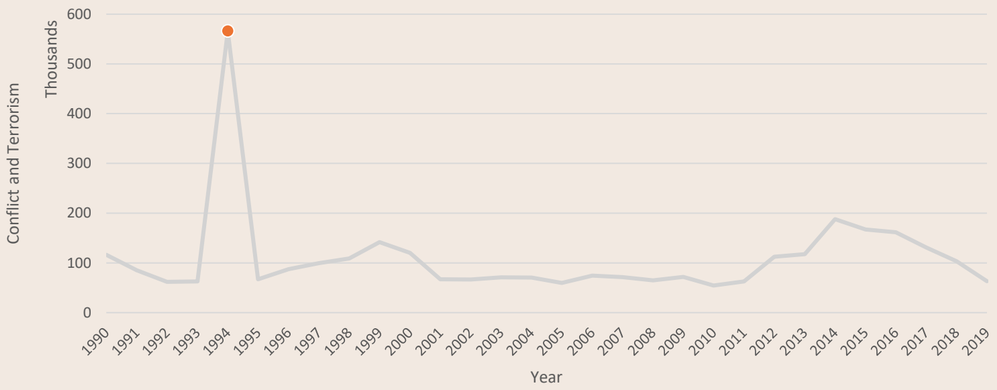


Observations

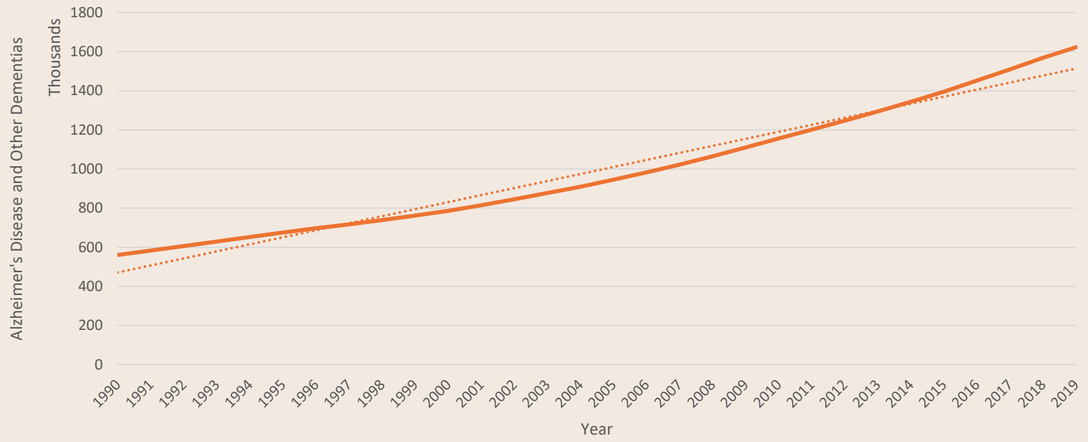
- Conflict and Terrorism has an Outlier at year 1994.
- Alzheimer's Disease and Other Dementias increases over time.
- Cardiovascular Disease increase over time.
- Meningitis decreases over time.
- Rwanda has noticeably higher Conflict and Terrorism deaths.
- India has noticeably higher Meningitis.
- China has noticeably higher Alzheimer's Disease and Other Dementias.

Observations

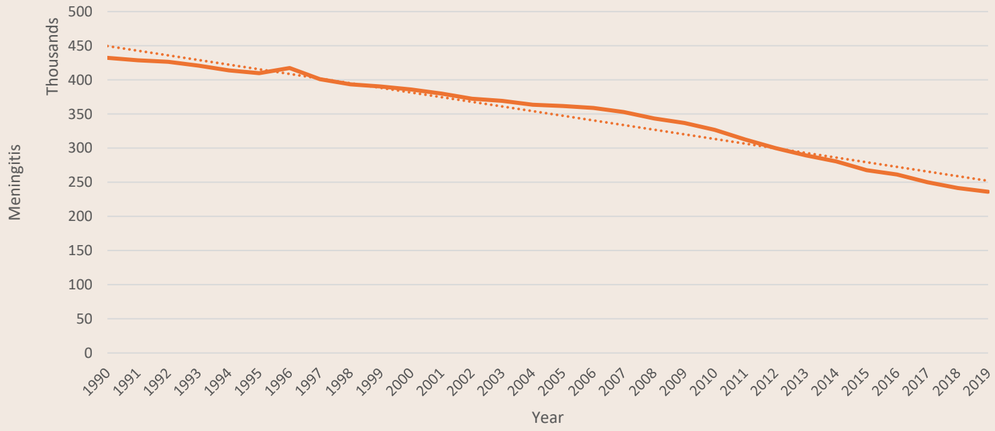
'Conflict and Terrorism' has an outlier at 'Year': 1994.



Alzheimer's Disease and Other Dementias increases over time.

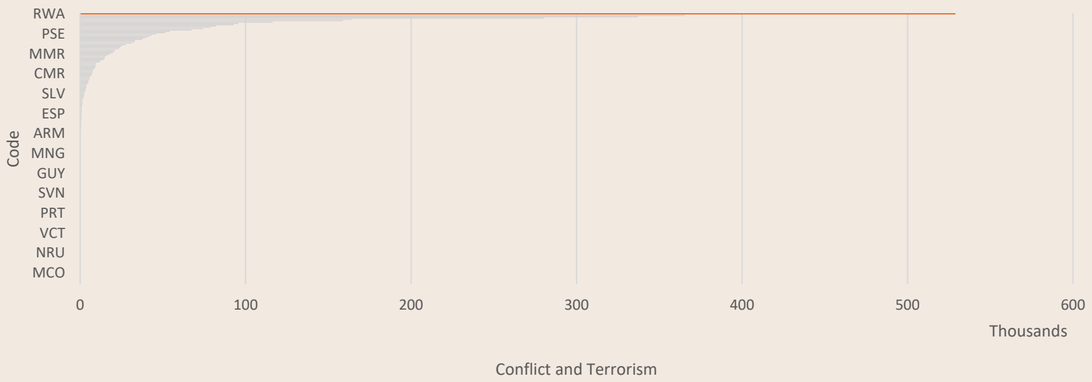


Meningitis decreases over time.

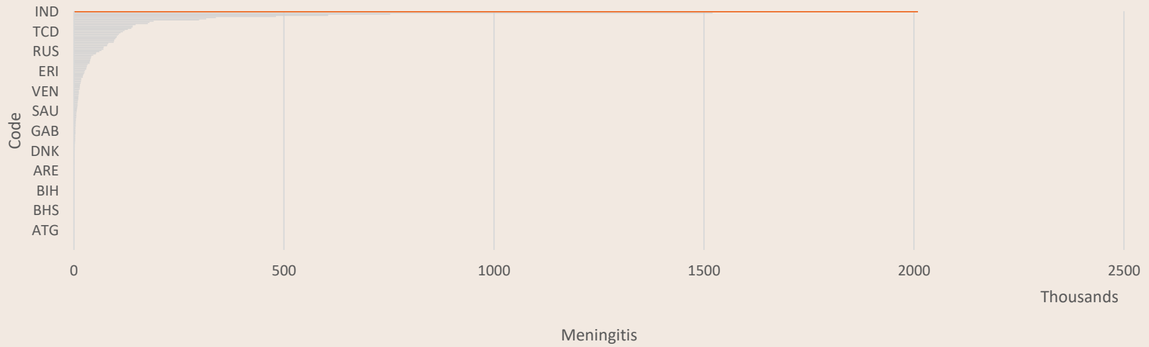


Observations

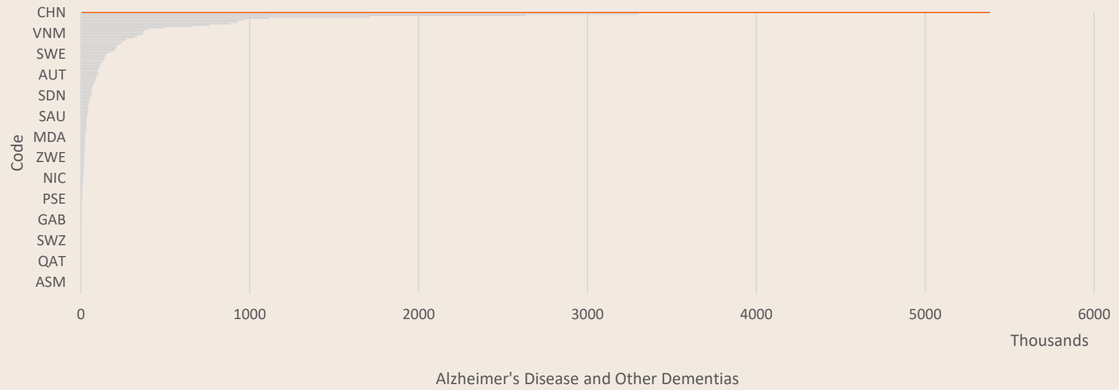
'Code': RWA has noticeably higher 'Conflict and Terrorism'.



'Code': IND has noticeably higher 'Meningitis'.



'Code': CHN has noticeably higher 'Alzheimer's Disease and Other Dementias'.

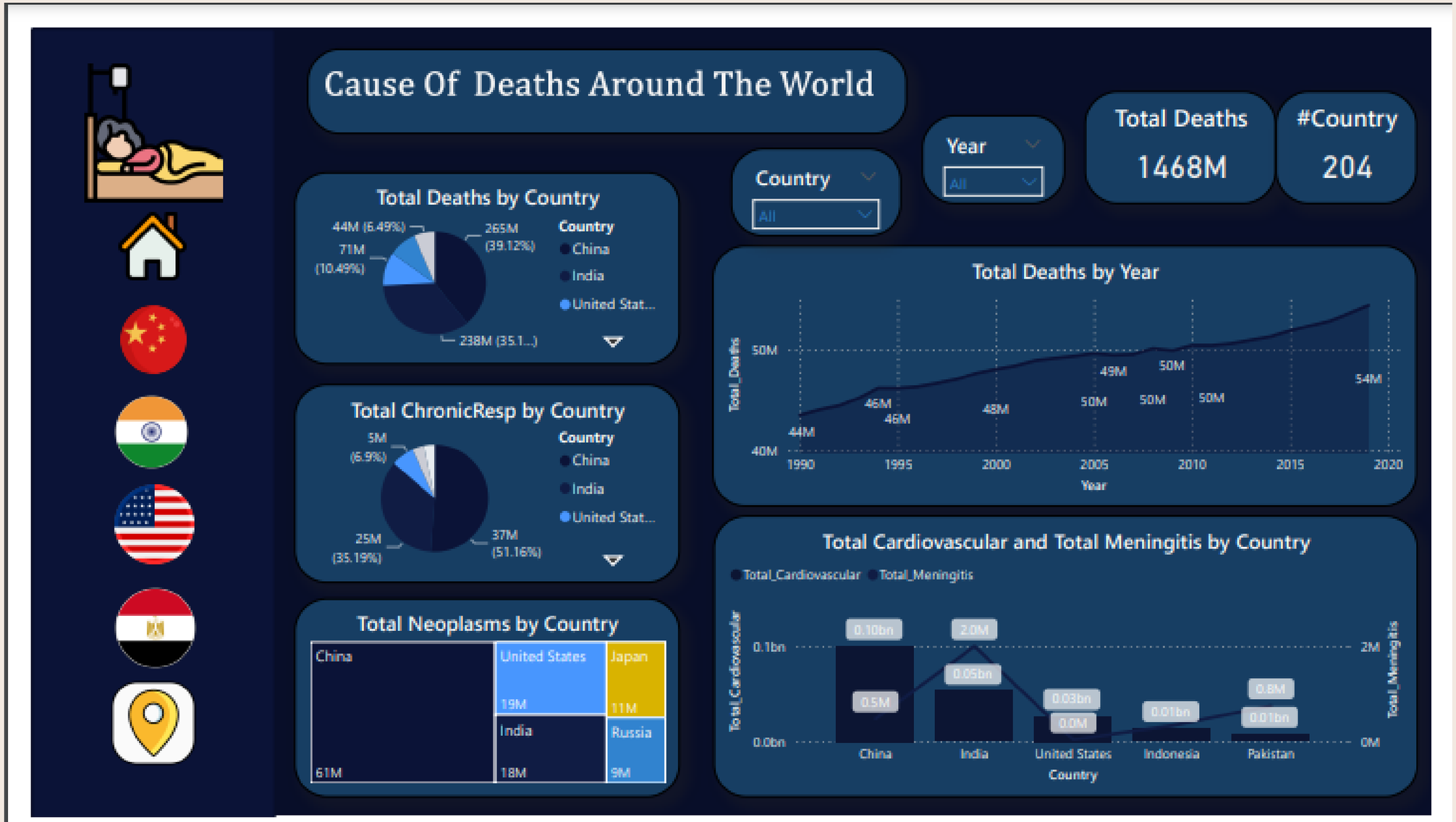


Power BI Analysis

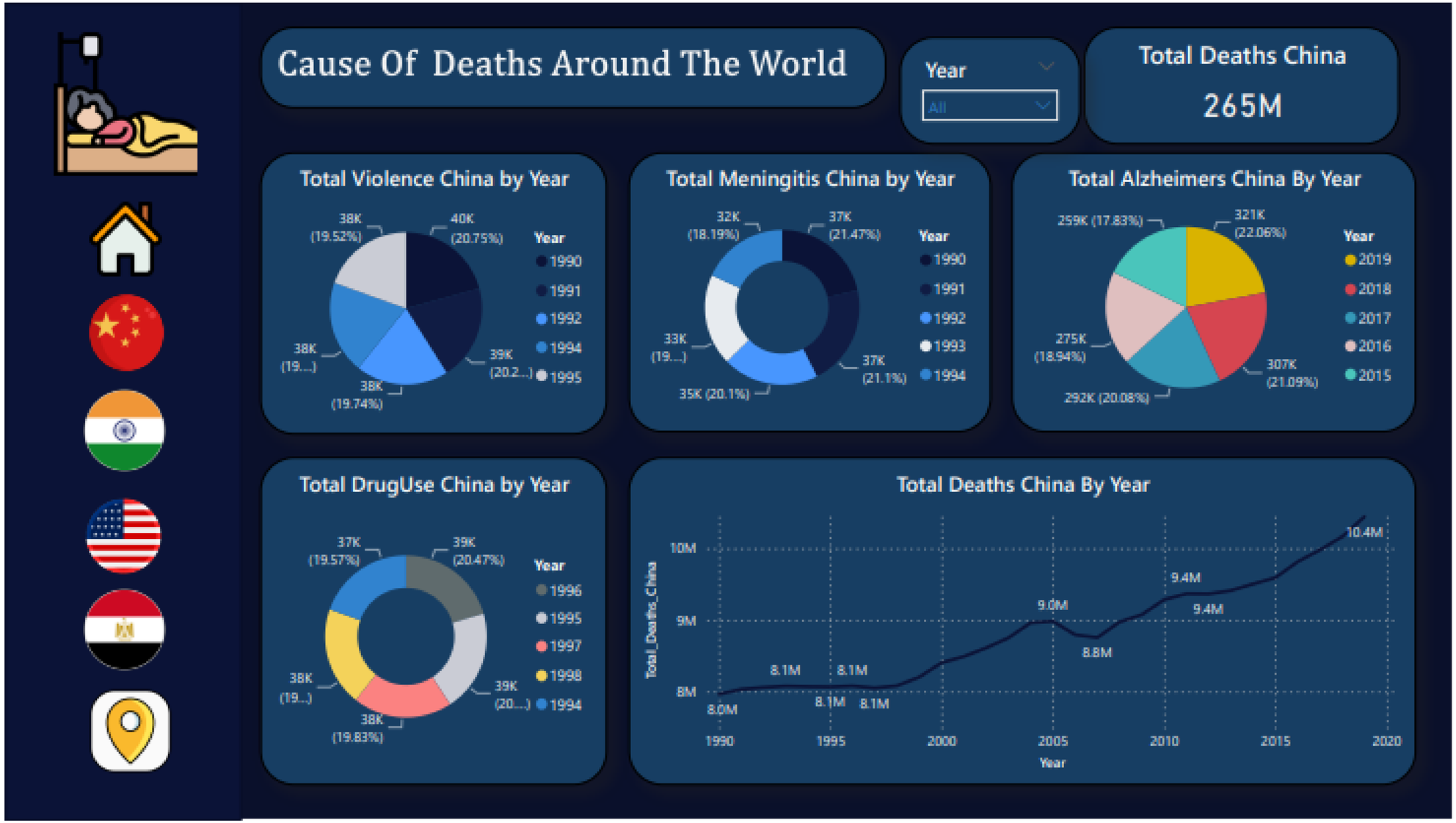
Content Map:

- Dashboards
- Geographical Maps

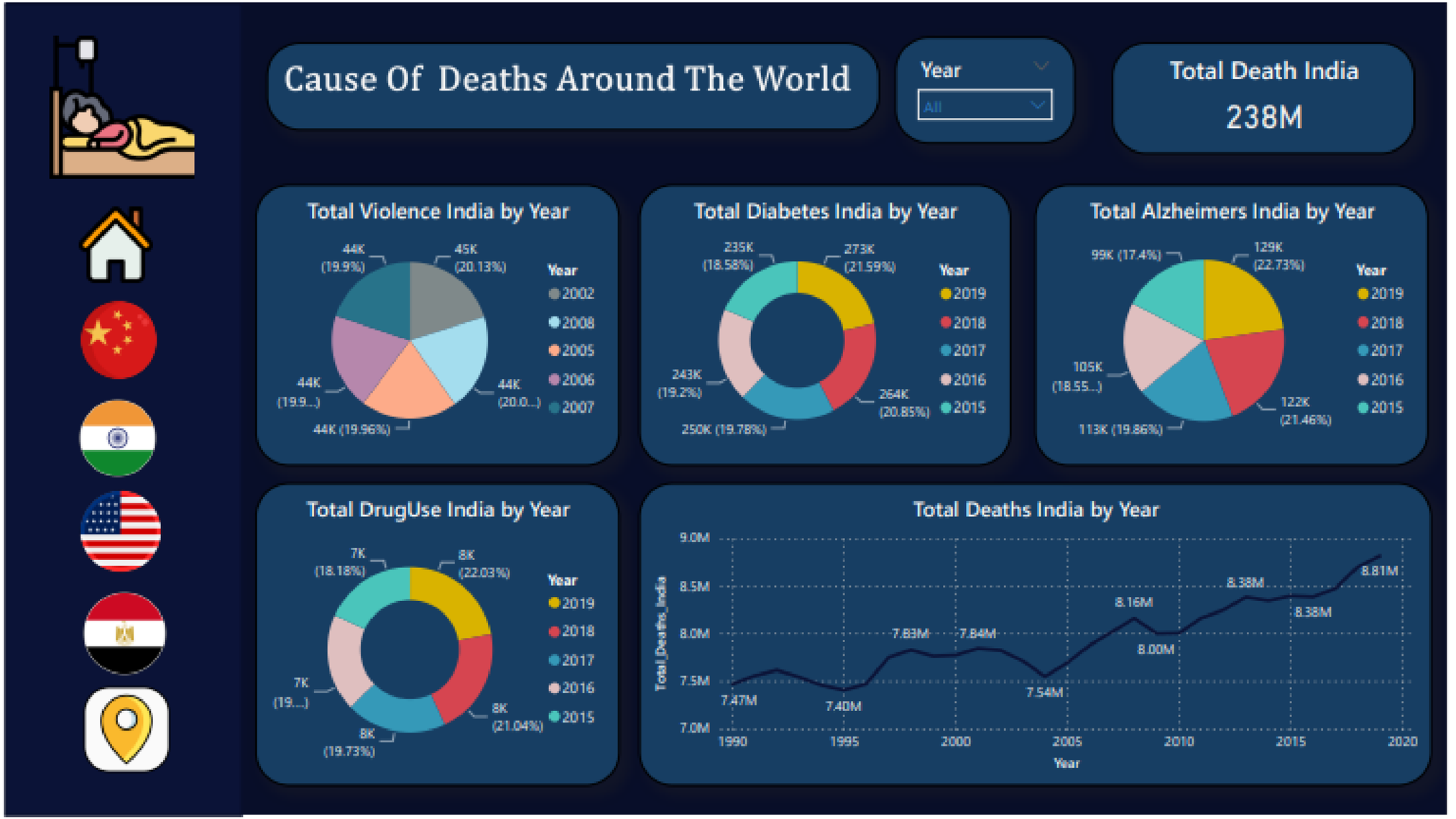
Dashboards


Cause Of Deaths Around The World
Year
All
Country
All
Total Deaths
1468M
#Country
204
Total Deaths by Country
44M (6.49%)
265M (39.12%)
71M (10.49%)
238M (35.1...)
Country
China
India
United Stat...
Total ChronicResp by Country
5M (6.9%)
25M (35.19%)
37M (51.16%)
Country
China
India
United Stat...
Total Neoplasms by Country
China
61M
United States
19M
Japan
11M
India
18M
Russia
9M
Total Deaths by Year
Total_Deaths
50M
40M
44M
46M
46M
48M
49M
50M
50M
50M
50M
54M
1990
1995
2000
2005
2010
2015
2020
Year
Total Cardiovascular and Total Meningitis by Country
Total_Cardiovascular
Total_Meningitis
0.1bn
0.0bn
0.10bn
0.5M
2.0M
0.05bn
0.03bn
0.0M
0.01bn
0.8M
0.01bn
2M
0M
China
India
United States
Indonesia
Pakistan
Country

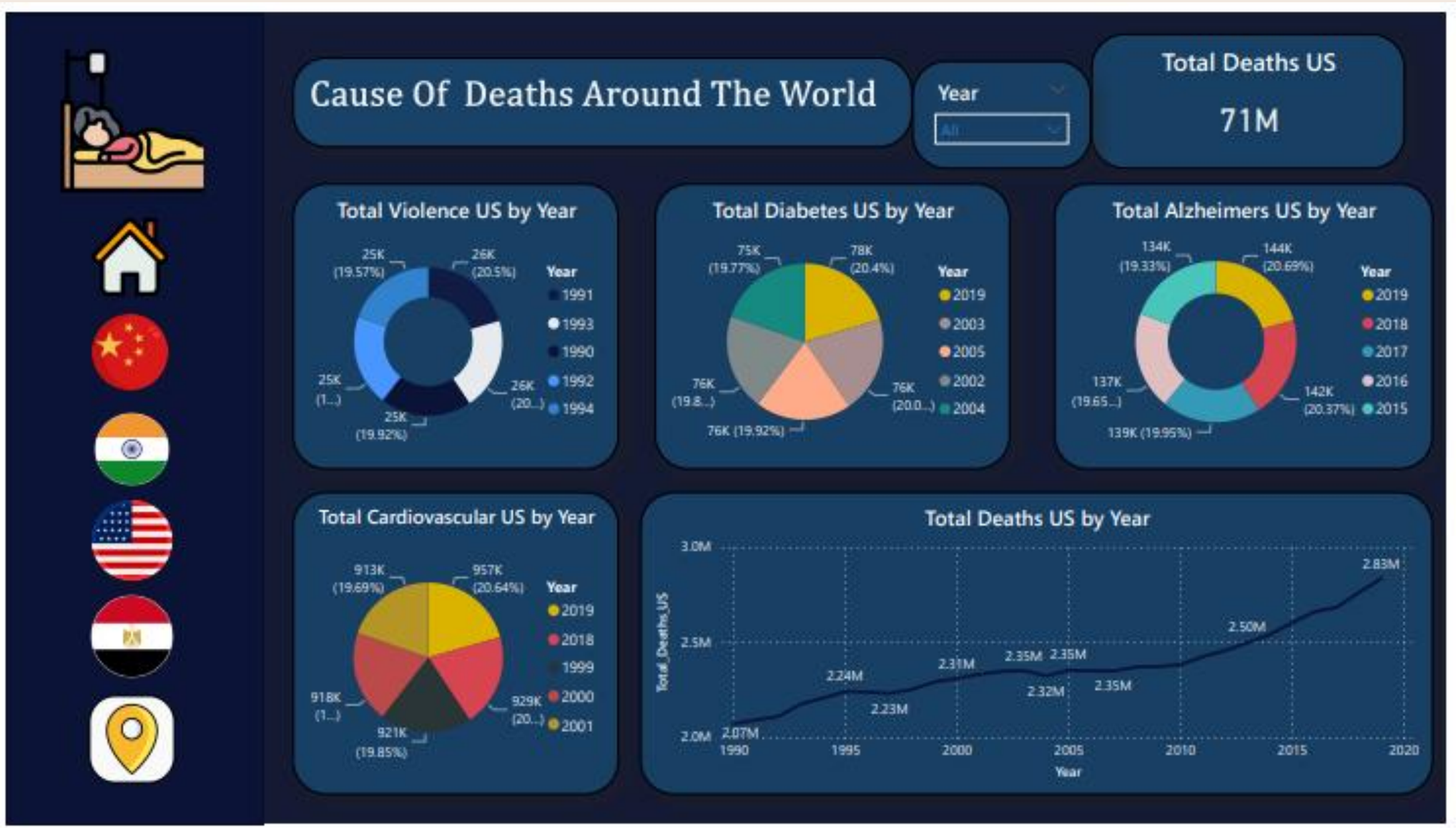
Dashboards


Cause Of Deaths Around The World
Year
All
Total Deaths China
265M
Total Violence China by Year
38K (19.52%)
40K (20.75%)
38K (19....)
39K (20.2...)
38K (19.74%)
Year
1990
1991
1992
1994
1995
Total Meningitis China by Year
32K (18.19%)
37K (21.47%)
33K (19....)
37K (21.1%)
35K (20.1%)
Year
1990
1991
1992
1993
1994
Total Alzheimers China By Year
259K (17.83%)
321K (22.06%)
275K (18.94%)
307K (21.09%)
292K (20.08%)
Year
2019
2018
2017
2016
2015
Total DrugUse China by Year
37K (19.57%)
39K (20.47%)
38K (19...)
39K (20....)
38K (19.83%)
Year
1996
1995
1997
1998
1994
Total Deaths China By Year
Total_Deaths_China
10M
9M
8M
8.0M
8.1M
8.1M
8.1M
8.8M
9.0M
9.4M
9.4M
10.4M
1990
1995
2000
2005
2010
2015
2020
Year

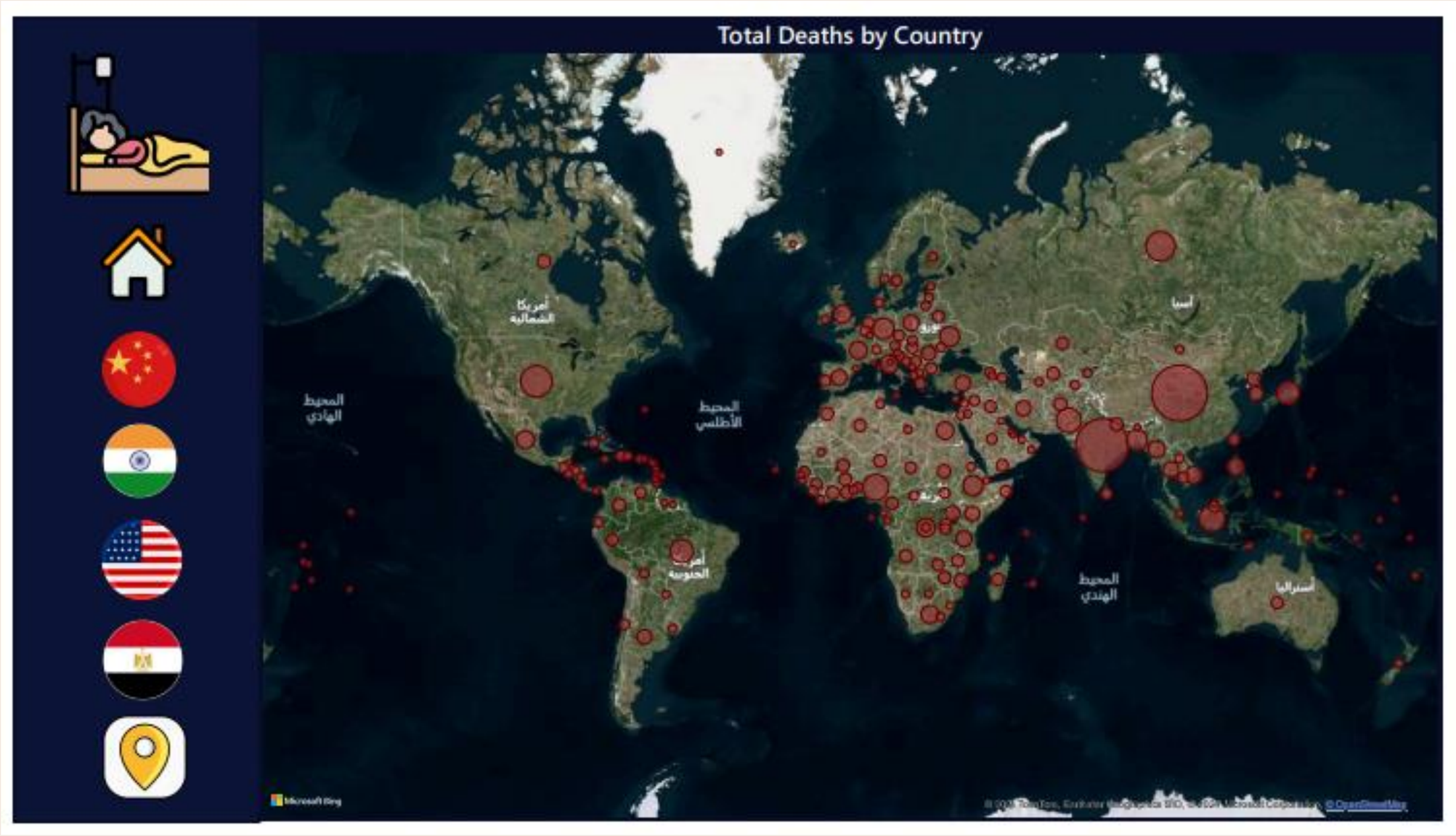
Dashboards


Cause Of Deaths Around The World
Year
All
Total Death India
238M
Total Violence India by Year
45K (20.13%)
44K (20.0...)
44K (19.96%)
44K (19.9...)
44K (19.9%)
Year
2002
2008
2005
2006
2007
Total Diabetes India by Year
273K (21.59%)
264K (20.85%)
250K (19.78%)
243K (19.2%)
235K (18.58%)
Year
2019
2018
2017
2016
2015
Total Alzheimers India by Year
129K (22.73%)
122K (21.46%)
113K (19.86%)
105K (18.55...)
99K (17.4%)
Year
2019
2018
2017
2016
2015
Total DrugUse India by Year
8K (22.03%)
8K (21.04%)
8K (19.73%)
7K (19...)
7K (18.18%)
Year
2019
2018
2017
2016
2015
Total Deaths India by Year
Total_Deaths_India
9.0M
8.5M
8.0M
7.5M
7.0M
7.47M
7.40M
7.83M
7.84M
7.54M
8.16M
8.00M
8.38M
8.38M
8.81M
1990
1995
2000
2005
2010
2015
2020
Year

Dashboards


Cause Of Deaths Around The World
Year
All
Total Deaths US
71M
Total Violence US by Year
25K (19.57%)
26K (20.5%)
25K (1...)
26K (20...)
25K (19.92%)
Year
1991
1993
1990
1992
1994
Total Diabetes US by Year
75K (19.77%)
78K (20.4%)
76K (19.8...)
76K (20.0...)
76K (19.92%)
Year
2019
2003
2005
2002
2004
Total Alzheimers US by Year
134K (19.33%)
144K (20.69%)
137K (19.65...)
142K (20.37%)
139K (19.95%)
Year
2019
2018
2017
2016
2015
Total Cardiovascular US by Year
913K (19.69%)
957K (20.64%)
918K (1...)
929K (20...)
921K (19.85%)
Year
2019
2018
1999
2000
2001
Total Deaths US by Year
3.0M
2.5M
2.0M
Total_Deaths_US
2.07M
2.24M
2.23M
2.31M
2.35M
2.32M
2.35M
2.35M
2.50M
2.83M
1990
1995
2000
2005
2010
2015
2020
Year

Geographical Maps



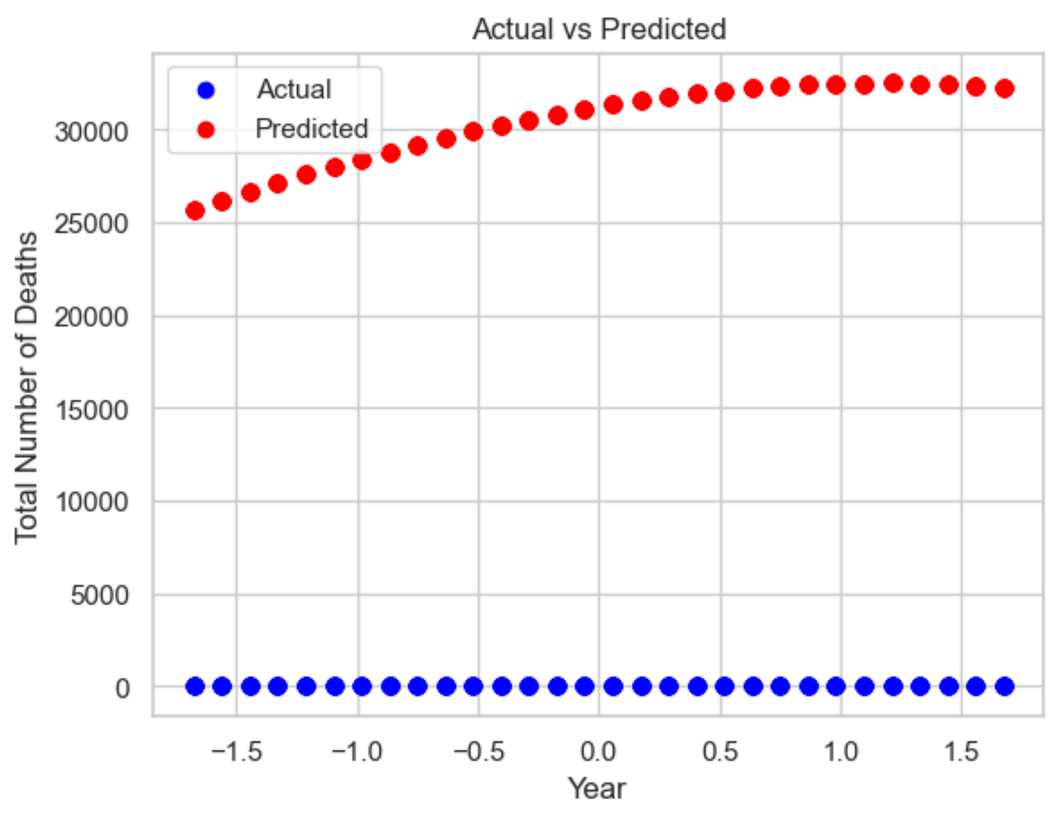
Bonus Parts

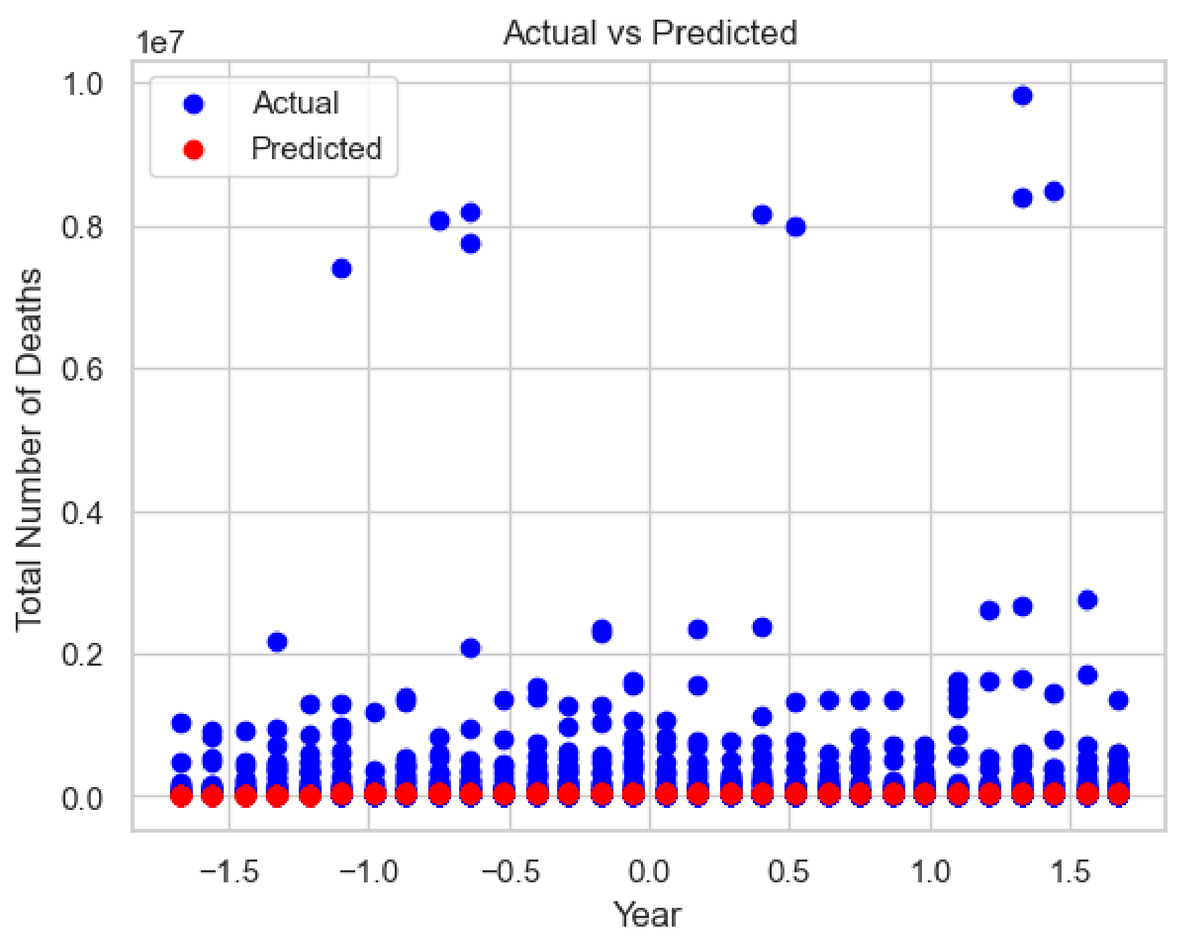
Points:

- Predictive Analysis in Python
- Egypt Analysis in Python and Power BI
- China, India and US Analysis in Python and Power BI
- Heatmaps in Python

Additional Points

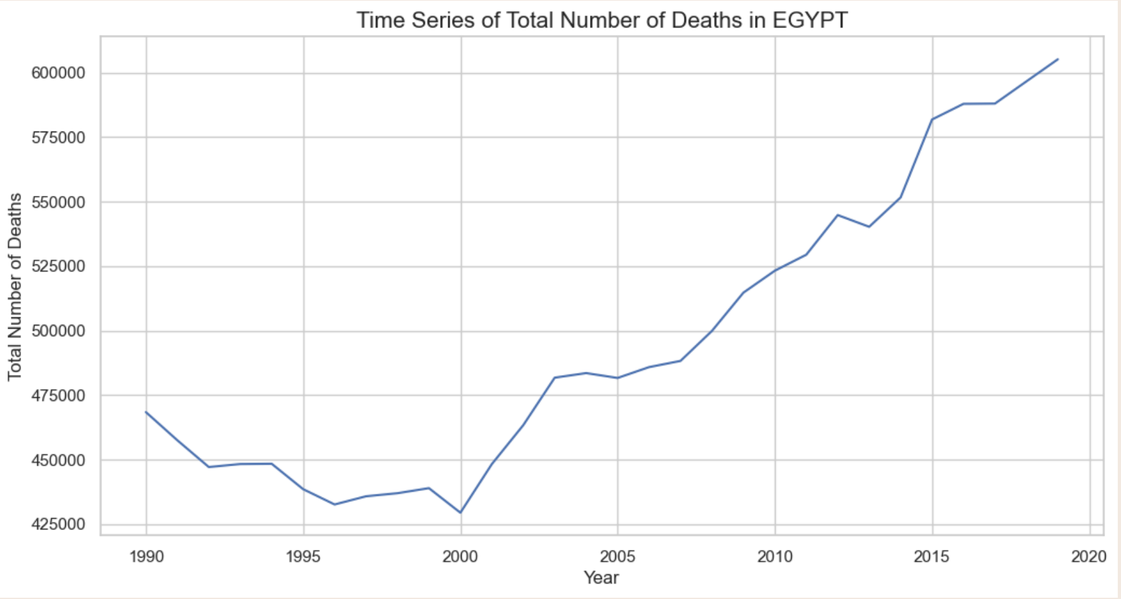
Prediction in Python

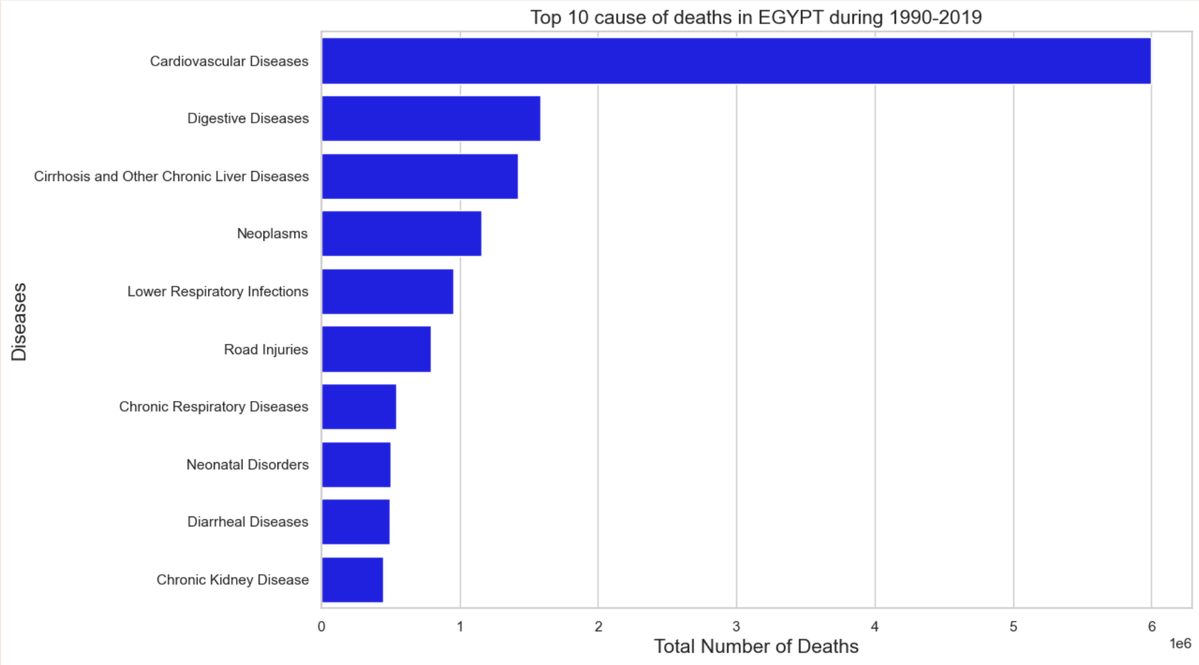


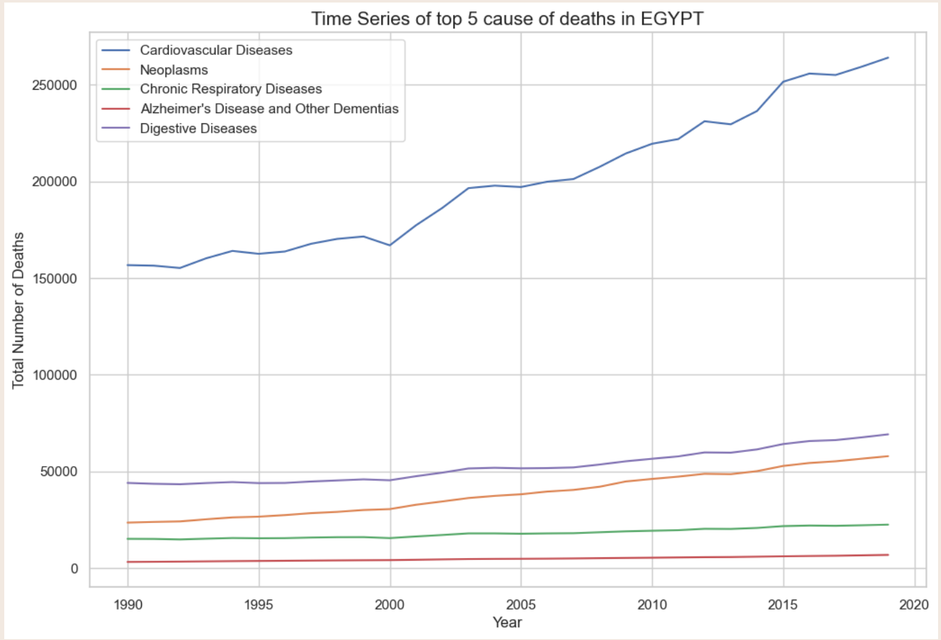


Additional Points

Egypt Analysis in Python and Power BI

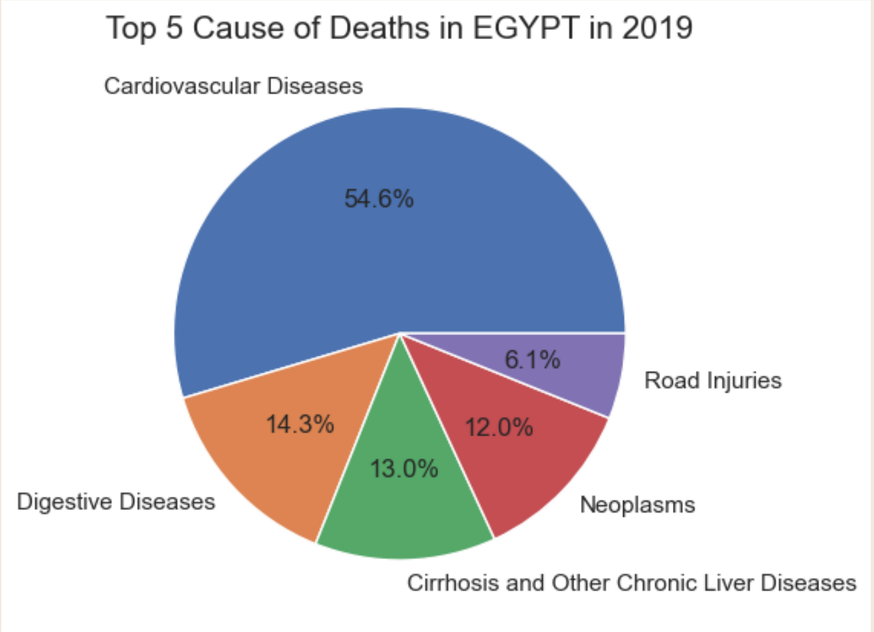


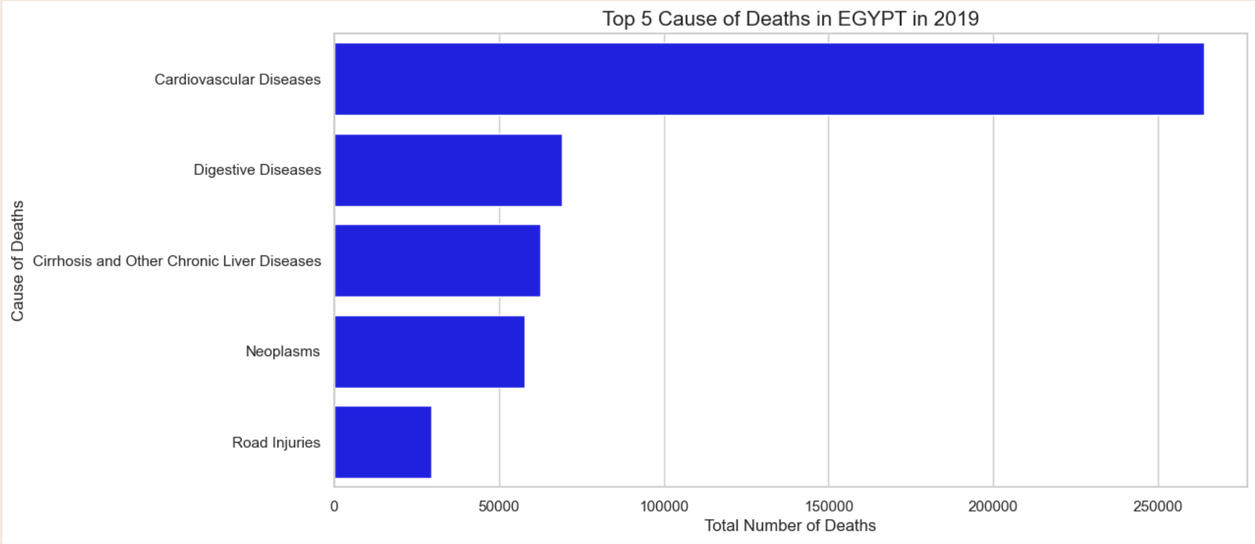


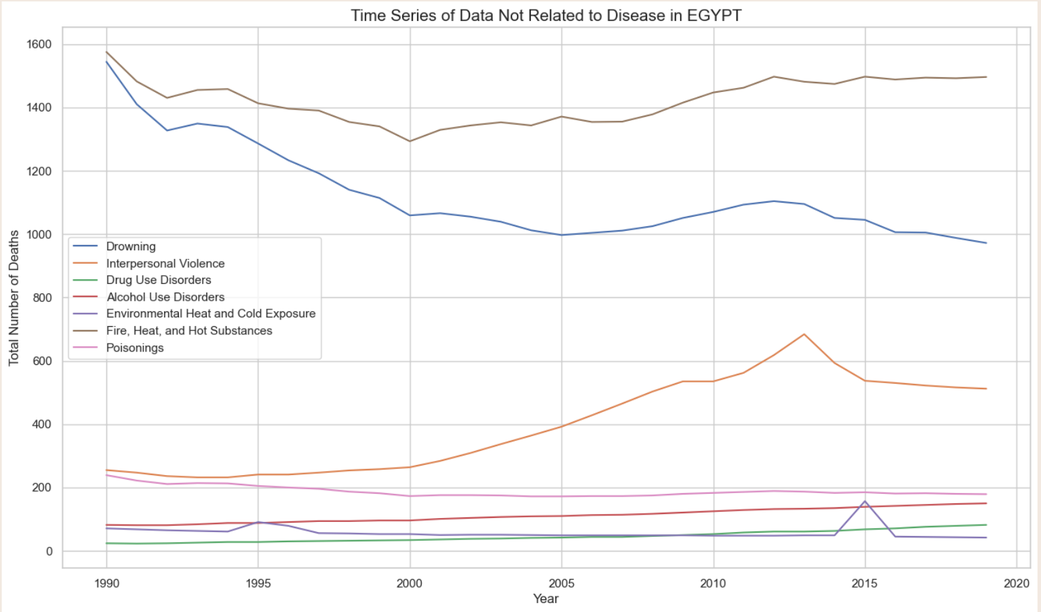


Additional Points

Egypt Analysis in Python and Power BI

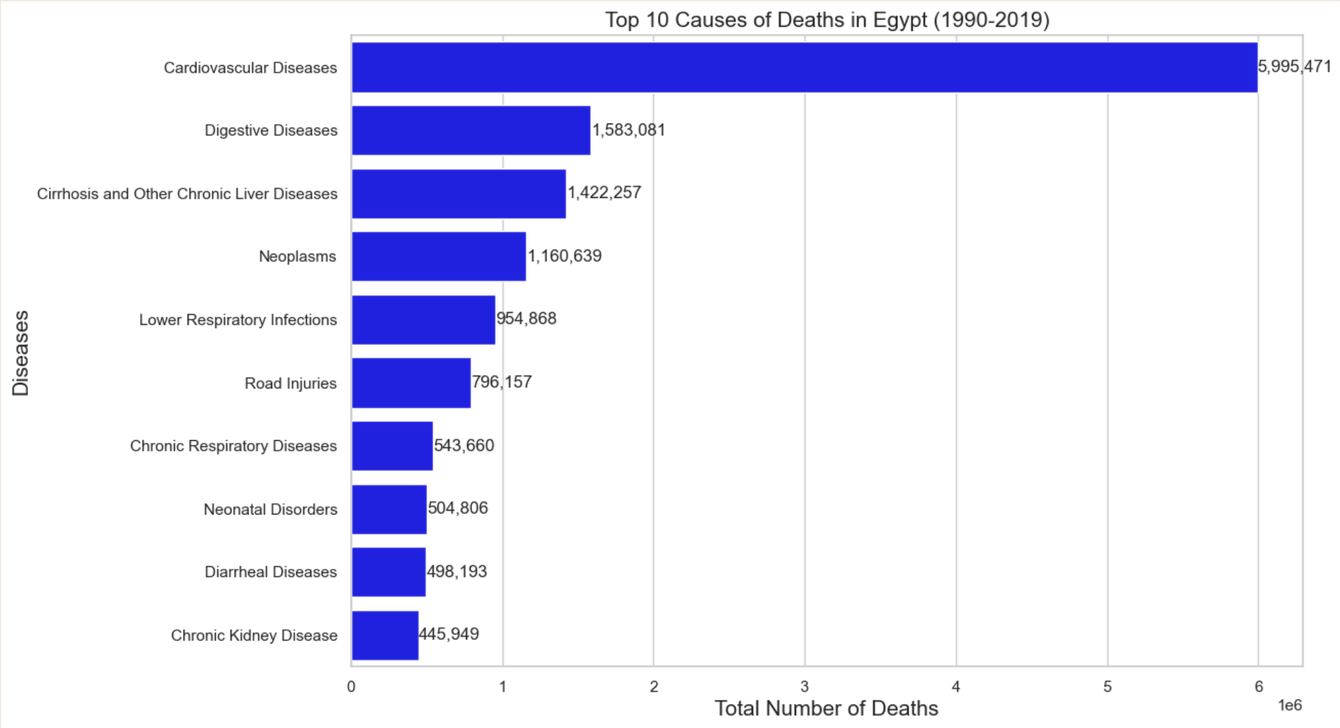


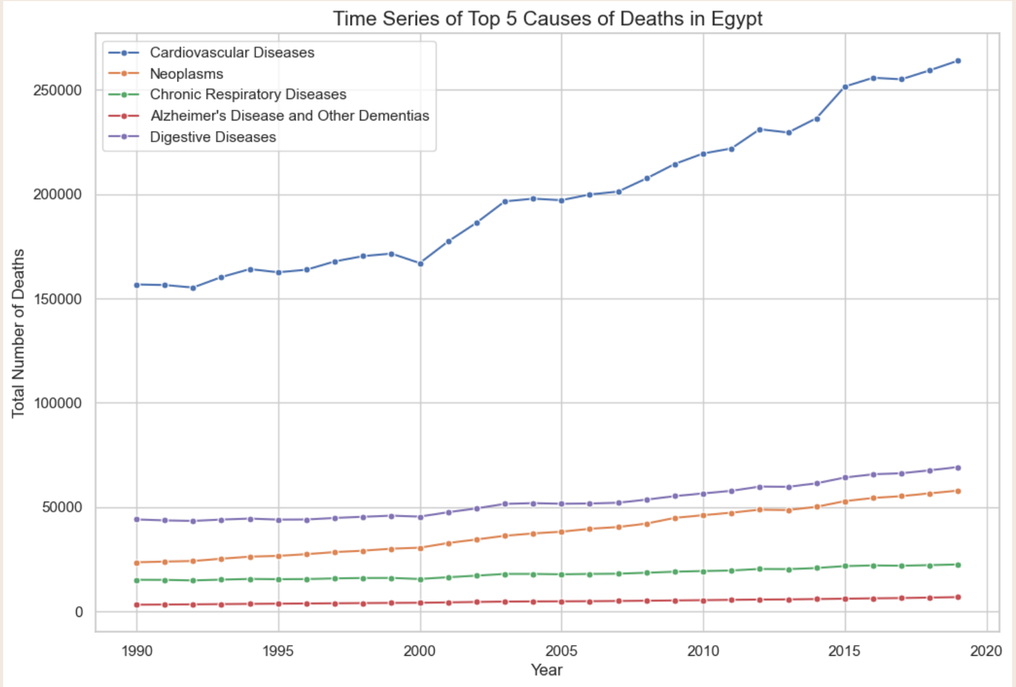




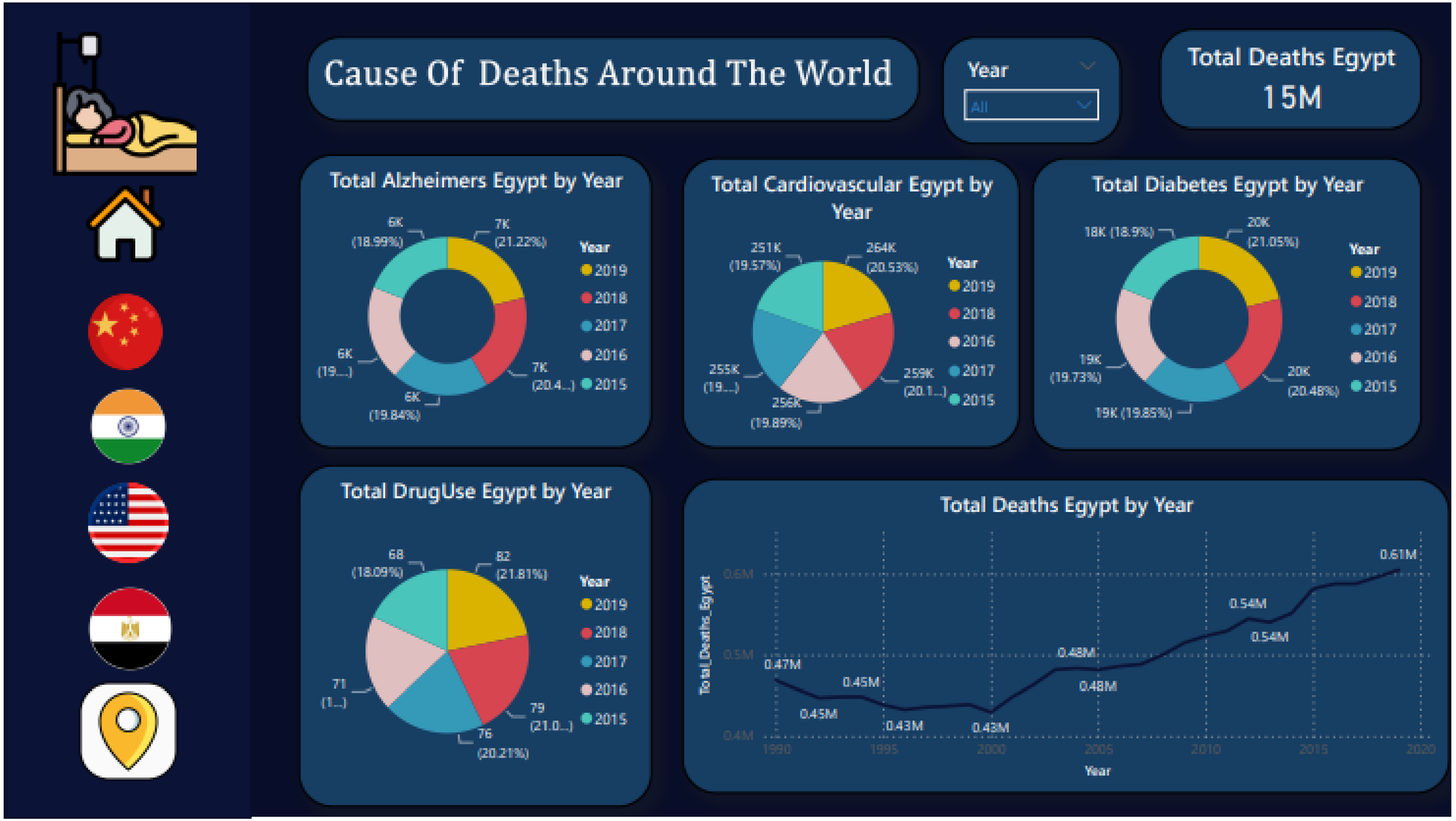
Additional Points

Egypt Analysis in Python and Power BI





Dashboard Power BI (EGYPT)



Thank You !