COMPARING THE ENVIRONMENTAL BURDEN OF DISEASE AND DIFFERENT HEALTH ISSUES IN SOUTH EAST ASIA AND EUROPE

DATA VISUALISATION AND THE WEB
Noshin Begum
2020

Contents

Research topic

"The environmental burden of disease quantifies the amount of disease caused by environmental risks" *(Quantifying environmental health impacts, World Health Organisation (WHO), 2018)*. I want to investigate the differences in environmental health impacts in South East Asia with Europe.
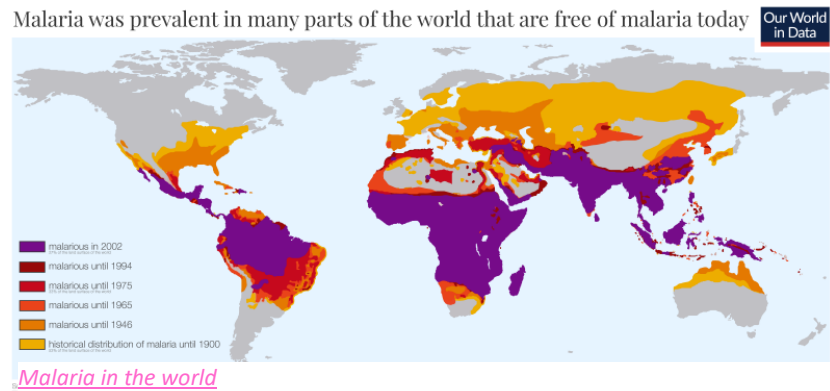

South East Asia Map

Europe Map

Questions

The questions I want to investigate is what diseases are affecting South East Asia and how those same diseases impact Europe, what are the differences in population and the rate of diseases and disease related mortality rates. I also want to focus on specific diseases, why it is it high or low in South East Asia compared with Europe.

Diseases I am investigating:
1. Diarrhoea
2. Malaria
3. Lung Cancer
4. Other cancers
5. Asthma
6. Stroke
7. Tuberculosis
8. Lower respiratory infections
9. Other respiratory diseases

Malaria was prevalent in many parts of the world that are free of malaria today

- malarious in 2002
- malarious until 1994
- malarious until 1975
- malarious until 1965
- malarious until 1946
- historical distribution of malaria until 1900

Malaria in the world

Background

I will be looking into the different issues in South East Asia and Europe to provide context to why certain diseases may be prominent, also looking into some preventive measures already in place to see if that lowers the rate or risk of certain diseases. I will be using WHO research done on quantifying environmental health impacts (*Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks, WHO, 2018*) to see different preventive measures offered.

WHO records that South East Asia consists of Bangladesh, Bhutan, Democratic People's Republic of Korea, India, Indonesia, Maldives, Myanmar, Nepal, Sri Lanka, Thailand, Timor-Leste.

WHO records Europe HIC (High Income Countries) consists of Andorra, Austria, Belgium, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Latvia, Lithuania, Luxembourg, Malta, Monaco, Netherlands, Norway, Poland, Portugal, San Marino, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom.

Concepts

Economic state of the countries could affect the rates or mortality or diseases due to how much preventive measures or treatments the country can afford (*Health and the economy: A vital relationship, OECD Observer, 2004*). Different cultures can lead to different habits like eating habits or sleeping habits which can affect the rates of specific diseases in the country (*Culture and the risk of disease, William W Dressler for British Medical Bulletin, 2004*).
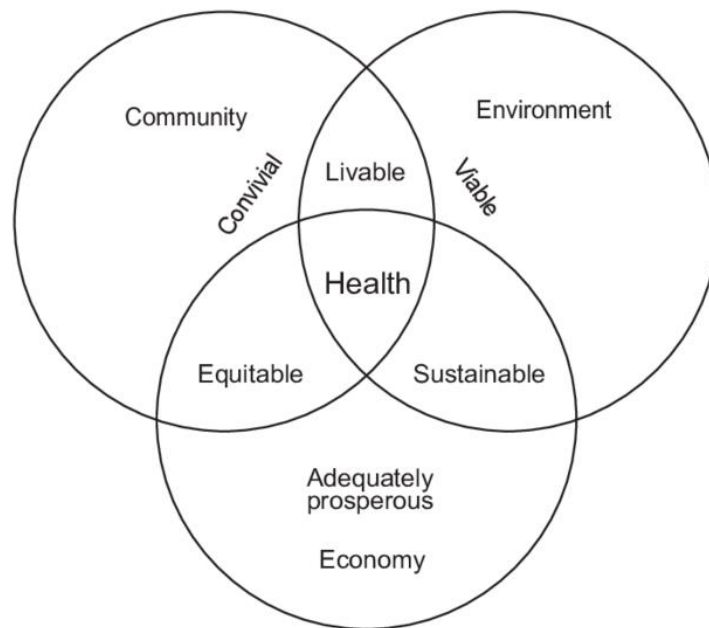
Health and economy

Diarrhoea is mainly spread through poor hygiene, contaminated food, water and organisms. Malaria is spread via infected vectors (mosquitos) and is bred in specific climates. Lung cancer and other cancers are usually due to genetics or diets or underlying health issues. Asthma is mainly caused due to the environment or genetics. Main causes of stroke are diet, habits and genetics. Tuberculosis is due to the exposure of another infected person. Respiratory infections or diseases can be spread via person to person or due to genetics and habits. These are the main causes, but diseases and health issues can happen randomly or due to underlying undiagnosed health problems.

South East Asia risk factors contributing to diseases and health issues are citizens are "economically disadvantaged, poor health care, changes in ecosystems, poor farming practices, increase global travel and tourism" (*Emerging infectious diseases in southeast Asia: regional challenges to control, Prof R Coker MD, B Hunter MSC, J Rudge PhD, M Liverani MSC, 2011*).

South East Asia consists of many low-income countries which leads to citizens being unable to afford enough hygiene practices and health care which leads to increase in communicable diseases such as tuberculosis and diarrhoea . Cultural beliefs can impact citizen diets which can lead to different non communicable diseases such as strokes and respiratory issues. South East Asia ecosystem and warm climate can breed harmful organisms which increases in diseases such as infections, diarrhoea and malaria.

Europe has increased use of "tobacco and alcohol" (*Leading risk factors for health in Europe: fact sheet, WHO Europe, 2012*) which can lead to increase in non-communicable diseases such as different types of cancers and respiratory issues. Cultural beliefs in Europe can lead to different diets which can increase diseases like stroke and diarrhoea.

South East Asia preventative methods have positively impacted Asia-pacific countries; "under age 5 mortality has declined by an average 50%. Countries such as Myanmar, China and Cambodia reported a drop of 70% or more due to effective preventive interventions, immunisations, malaria prevention and treatment" *(Health at a Glance: Asia/Pacific, measuring progress towards universal health coverage, 2018)*.

Europe has tobacco prevention plan which helps educate on the harms of tobacco use as well as provide nicotine replacement therapies (NRT) which has increased the rate of success. This plan also offers "prevention of relapse (i.e. behavioural support, counselling services, "quit lines" and routine advice on cessation of tobacco use) into national health programmes, plans and strategies, including those for primary health care, alcohol and drugs control, reproductive health, tuberculosis control*)" (European strategy for Tobacco Control, WHO Europe, 2007)*.

Data

Data source

For my data I wanted to use data from respectable sources with health as the primary focus and a source where I could find data for South East Asia and Europe. As if the data comes from the same source it may have been collected the same way so allows for accurate results for both areas.
I chose to use data collated by the World Health Organisation:

1. https://www.who.int/quantifying_ehimpacts/publications/Updated-2016-data-tables_Preventing_disease_Deaths_DALYs_PAFs_Sept_2019_rev.xlsx?ua=1
2. Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks (who.int)

Cleaning and Pre-processing

The dataset is in Excel format with both nominal and numerical data. To clean my data and for pre-processing I coded (in Jupyter Notebook) to read the specific sheet I wanted to investigate (sheet 2 – "Deaths") and removed the locations and disease that were not relevant to my investigation. I also ensured that I renamed the column headings to appropriate and clear names. I changed the data types of the 'Europe' and 'South-East Asia' columns to 'int' to allow for numerical calculations to be done using the data (percentages, differences, averages and statistics.) (Shown in the appendices).

Exploratory and explanatory data visualisation

South East Asia compared to Europe's- Population and Deaths

Shown below in Figure 1, it is clear to see that the total population in Europe is significantly smaller than the total population in South-East Asia however the total number of deaths is not as significantly differenced compared to the difference in population.
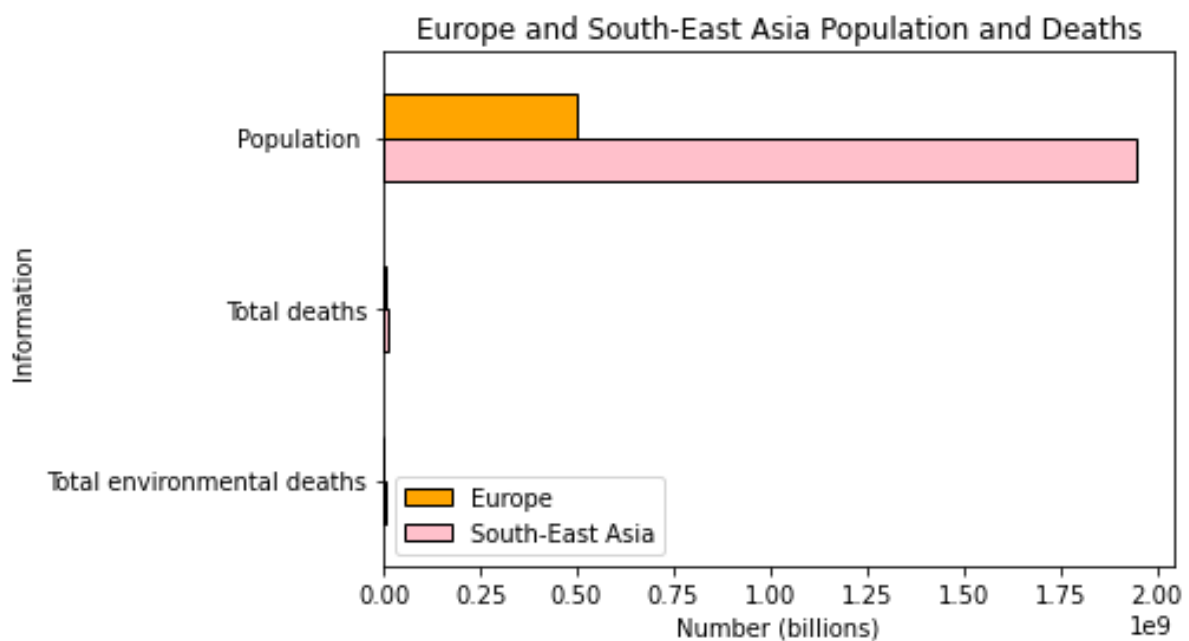

Figure 1

The key variables for this analysis are population, total number of deaths and total environmental deaths. The X-axis is in standard form (billions) to allow for clearer labels as the population is in the billions. I used a bar chart to present this information to visualise the large difference in population size.

South East Asia compared to Europe- Disease Related Deaths

From figure 2, lower respiratory infections and stroke are the leading cause of the deaths out of the nine main diseases for South-East Asia whereas in Europe the leading causes are trachea, bronchus, lung and other cancers. From this bar chart we can see that there are no malaria related deaths and small proportion of diarrhoeal disease related deaths in Europe. We can also see that only trachea, bronchus and lung cancers total of deaths surpass South-East Asia death rate whereas the remaining eight diseases have higher death rate in South-East Asia.
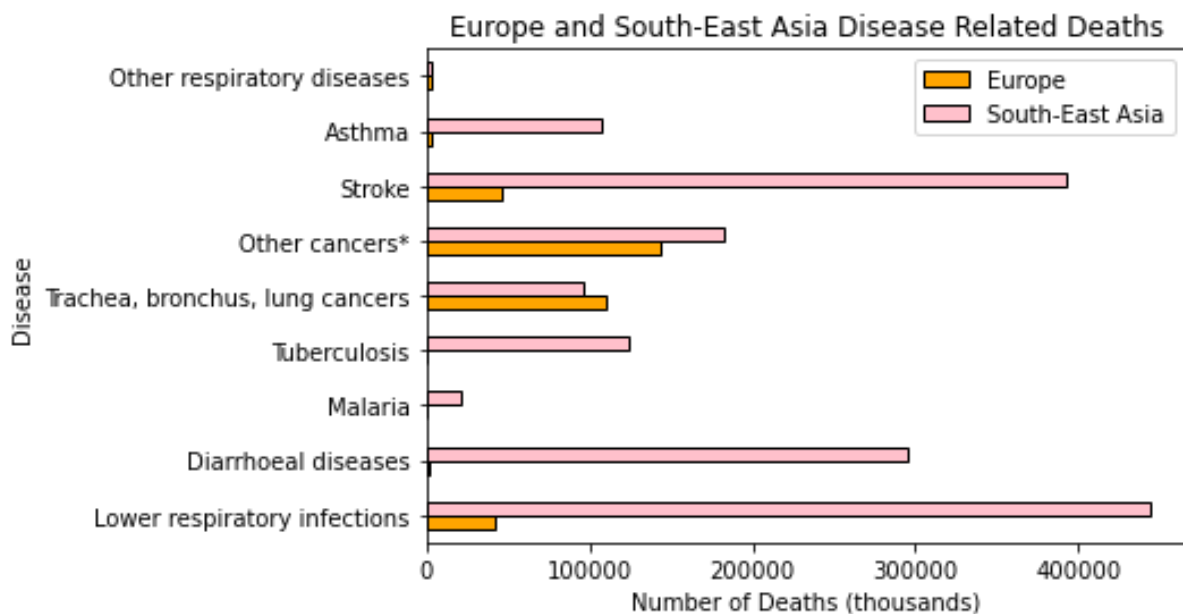

Figure 2

I used bar charts to present this information to compare what the leading cause of deaths are in each location and how different the deaths rates are.

Figure 3 is a bar chart to display the differences in South-East Asia disease related deaths compared to Europe. From this we can see the exact numeric difference for each disease death rate.
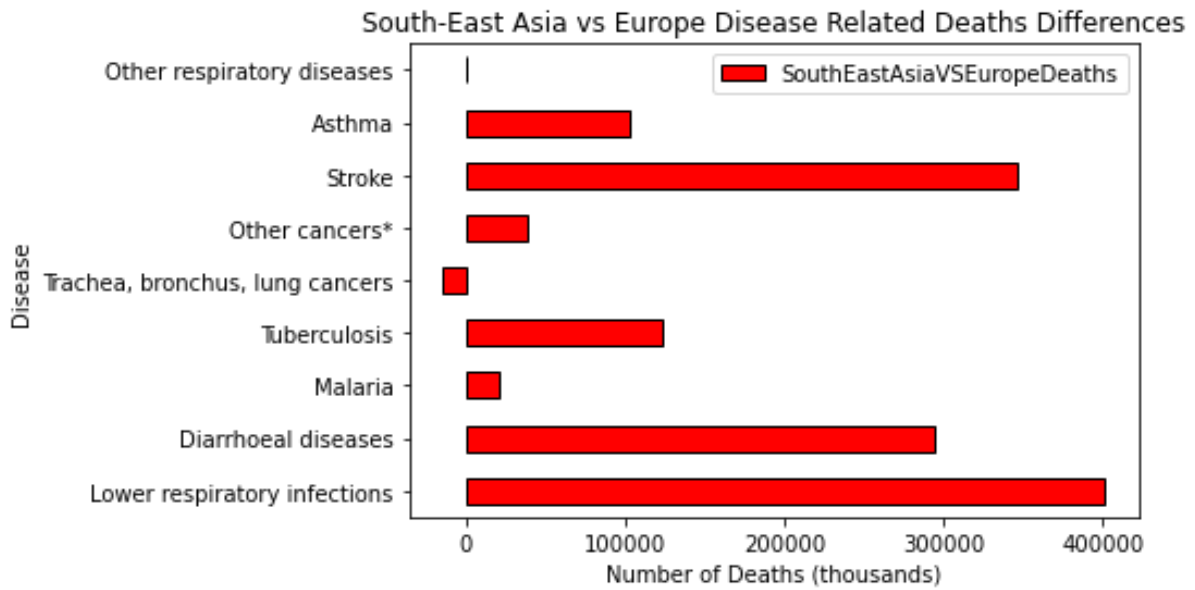
South-East Asia vs Europe Disease Related Deaths Differences

Figure 3

Figure 5 shows the total deaths for each disease as percentages of the population. As Europe and South-East Asia has different populations it is difficult to compare the true extent of the environmental burden of the disease, calculating the percentages of the deaths compared to the population shows the true environmental burden. Europe has a large issue with cancer related deaths and South-East Asia has with lower respiratory infections. Using bar charts to display the numeric differences for each disease allows for the user to see which disease South-East Asia has greater number of deaths in compared to Europe. It's very clear from this chart that Europe has a greater number of cancers related to trachea, bronchus and lung cancers compare to South-East Asia.
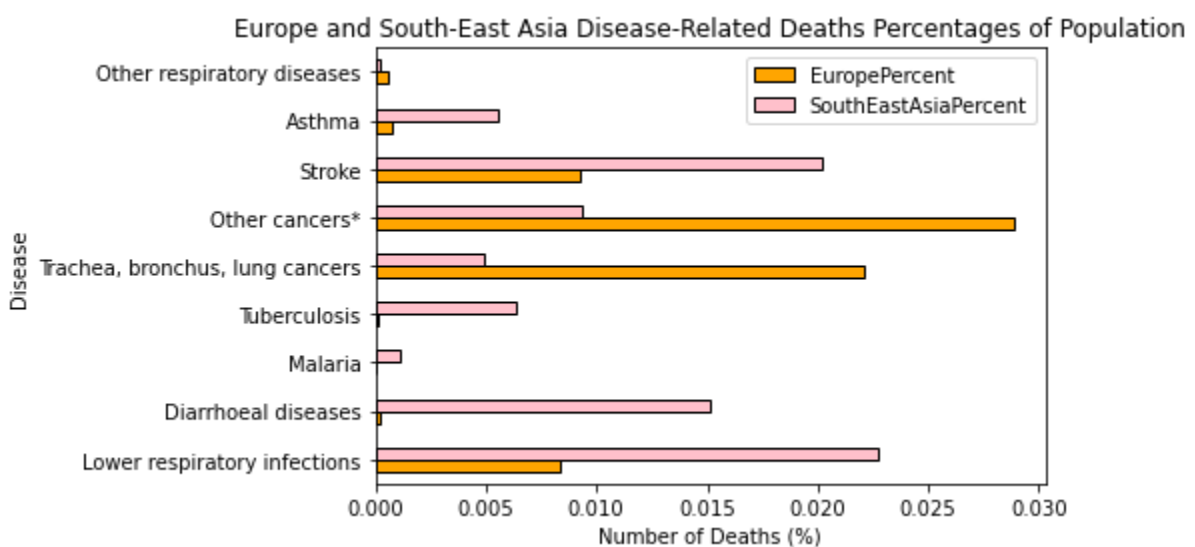

Europe and South-East Asia Disease-Related Deaths Percentages of Population

Figure 4

Figure 6 shows the differences between South-East Asia percentage against Europe. I used bar charts for figure 5 and 6 to show clearly the numeric percentages.
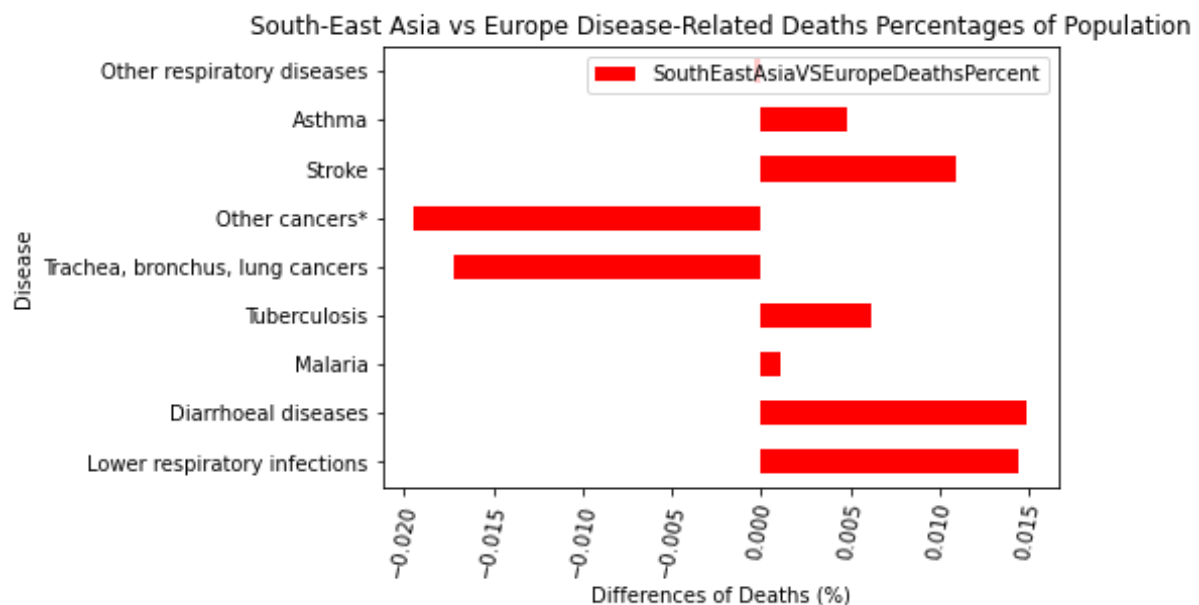
Figure 5

For figure 5 and 6 I used line graphs to display the disease related deaths for Europe and South-East Asia and the differences between the two areas. This allows users to see any correlations between the death rates and if the areas have similar death rates for any of the diseases. Malaria accounts for small proportion of deaths in South-East Asia and no deaths in Europe thus having similar environmental burdens.

Europe and South-East Asia Disease-Related Deaths

Figure 5

South-East Asia vs Europe Deaths Differences
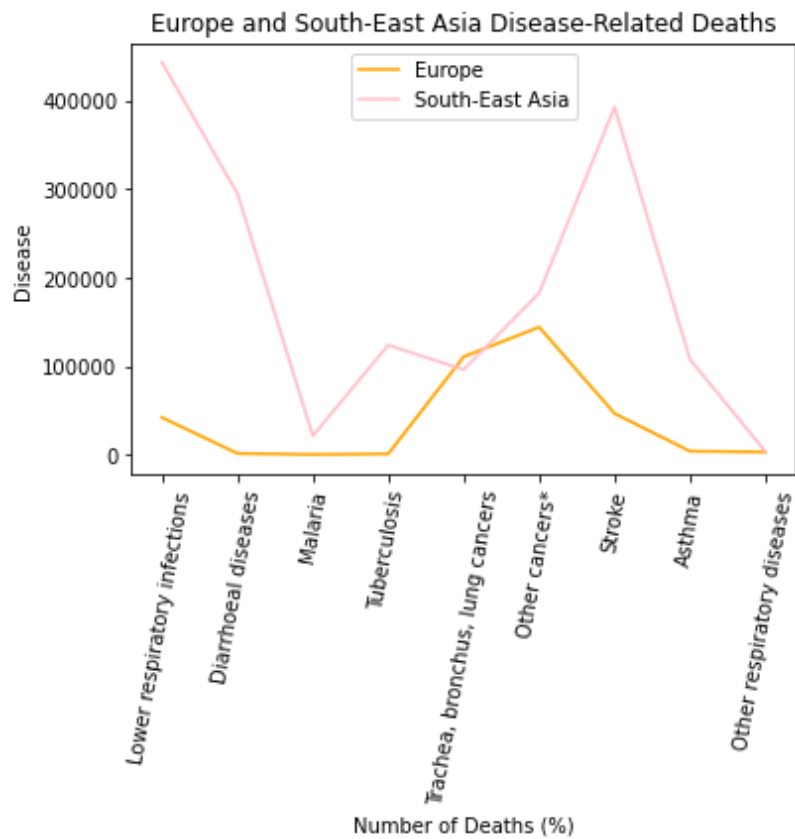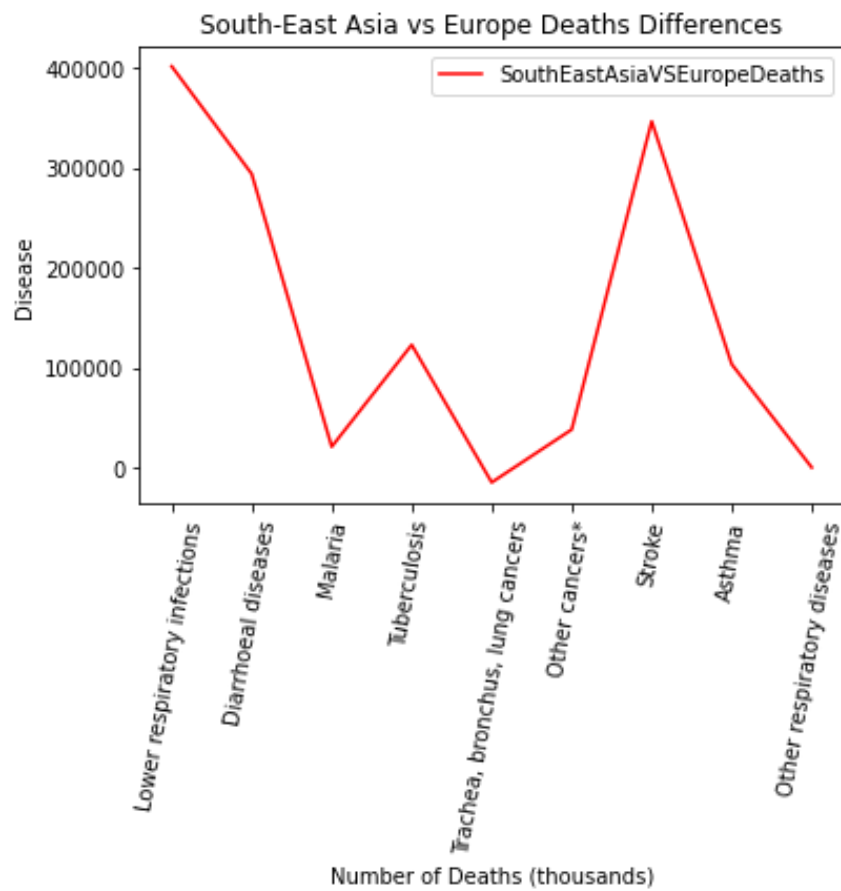
Figure 6

Figure 7 and 8 are line graphs to show the disease related deaths as percentages of population and the differences. This allows us to see for any correlations on the environmental burden of the diseases.
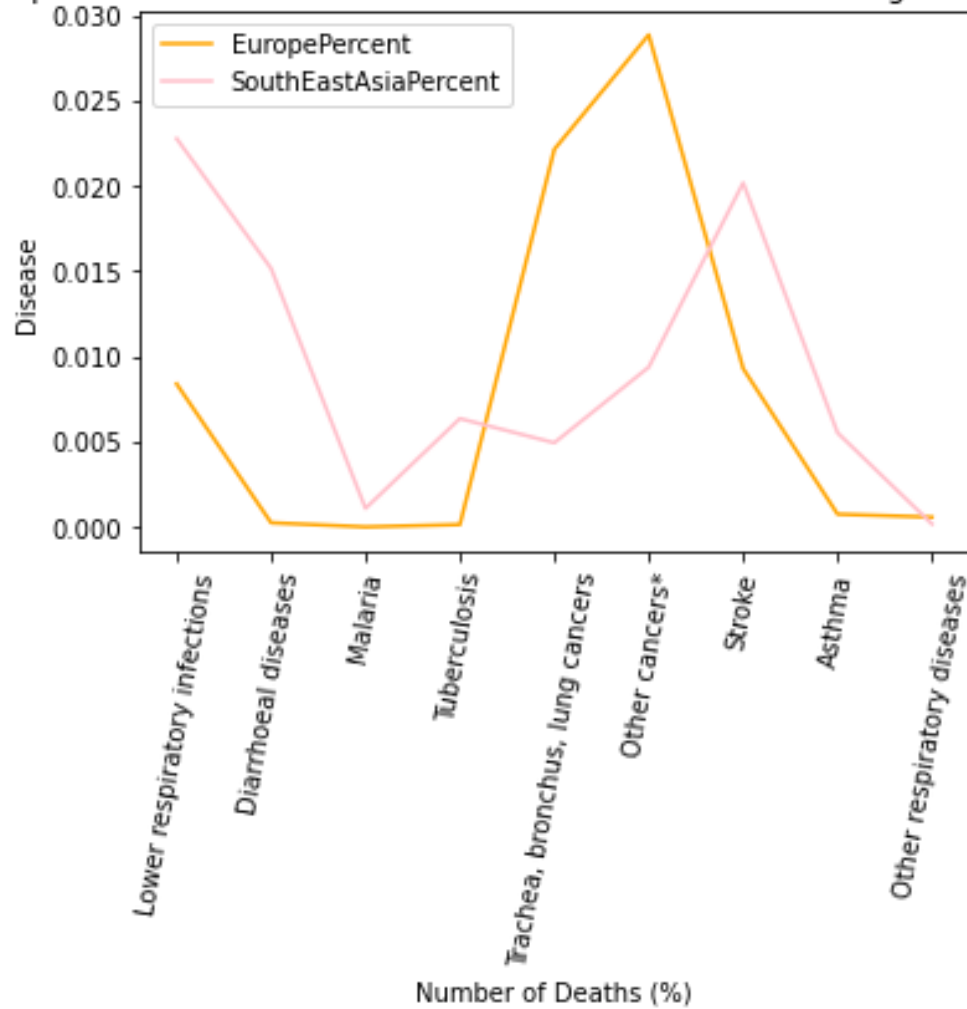
Figure 7

South-East Asia vs Europe Disease-Related Deaths Percentages of Population
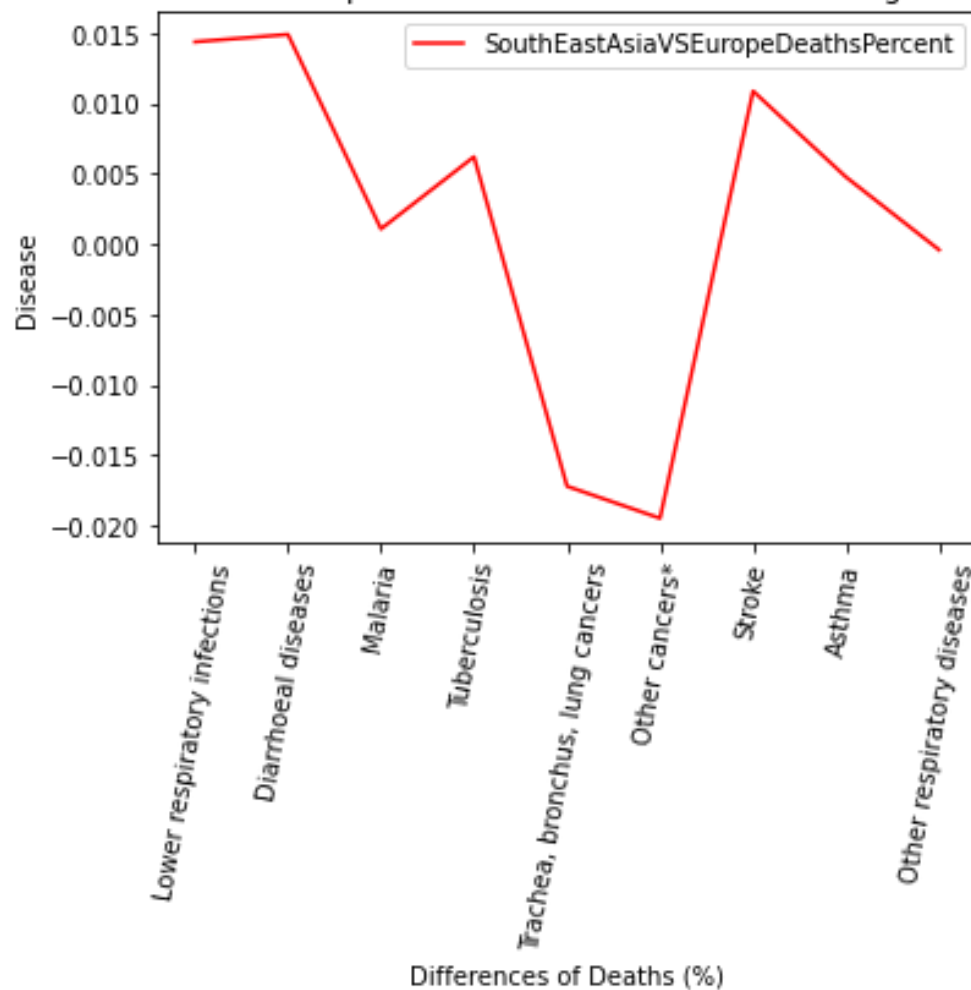
Figure 8

Figure 9 (South-East Asia) and 10 (Europe) are pie charts presenting the percentages of the disease related deaths out of the total nine diseases I am investigating. This allows us to see clearly which disease out of the nine diseases have the greatest impact on the death toll.
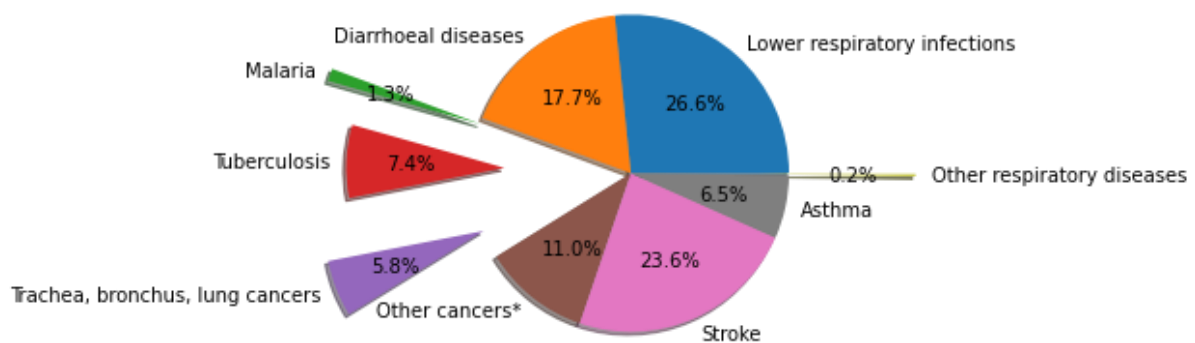
Figure 9
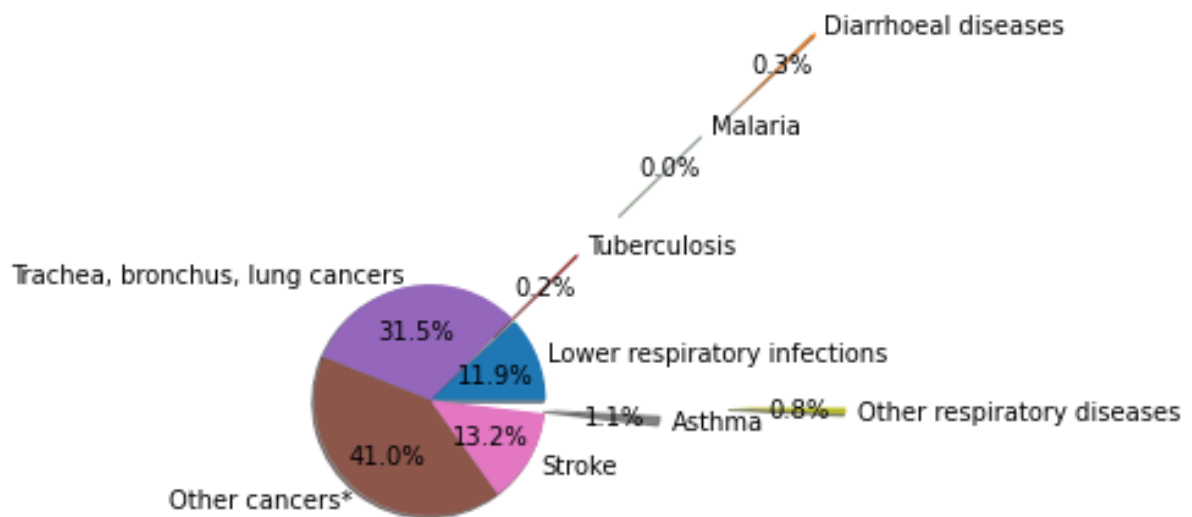
Figure 10

Figure 11 and 12 shows the distribution of the disease related deaths for Europe and South East Asia and the distributions for the percentages of population. From these boxplots we can clearly see the large distribution of deaths in South-East Asia and deaths as percentages of population in Europe.
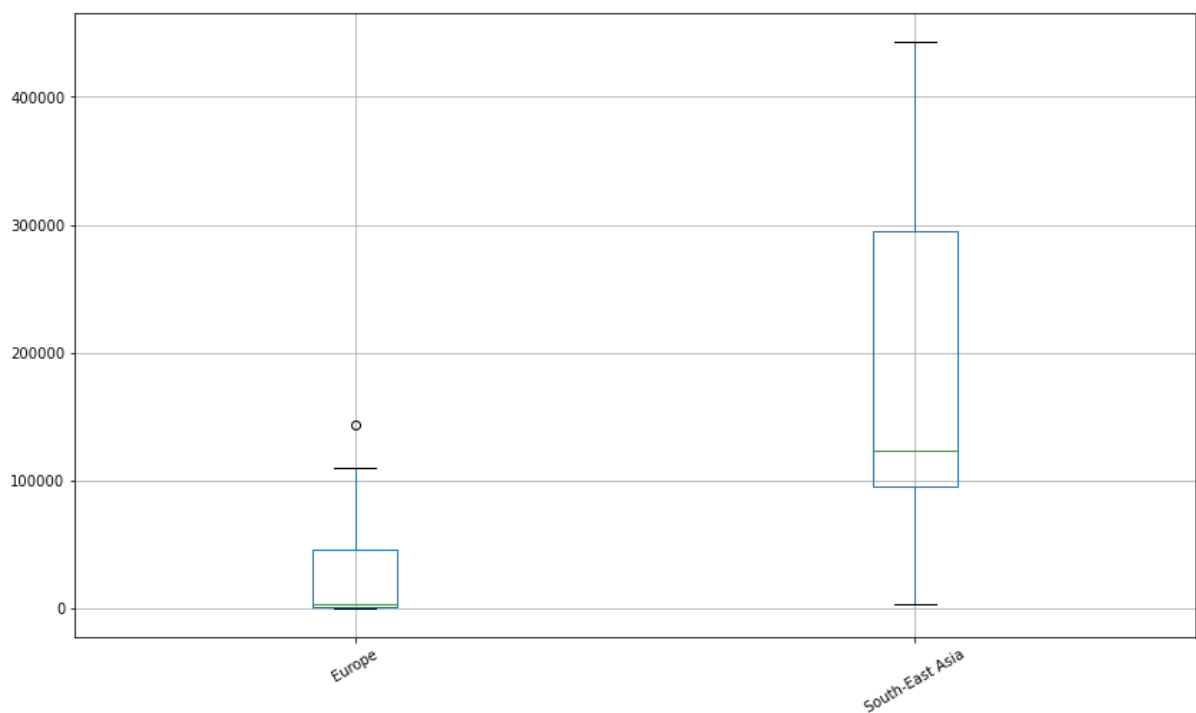
Figure 11
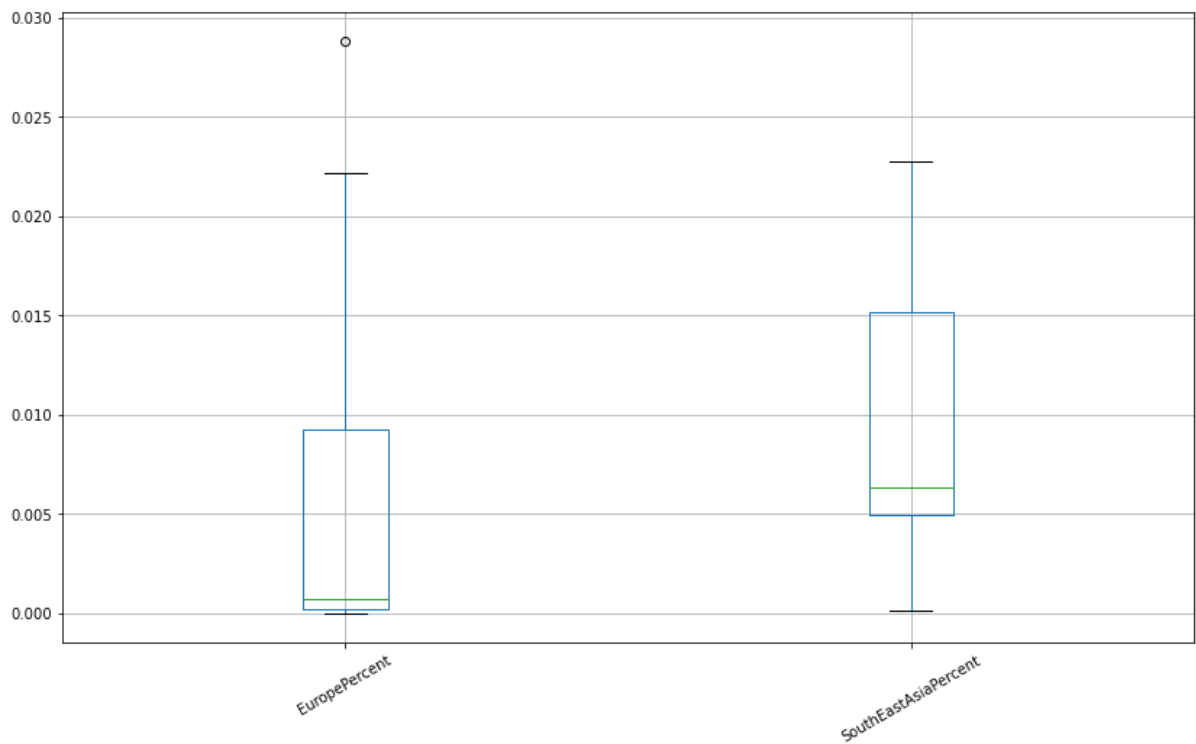
Figure 12

Figure 13 is a pairwise plot to summarise differences between Europe alone and South-East Asia deaths alone with the comparisons between the two areas.
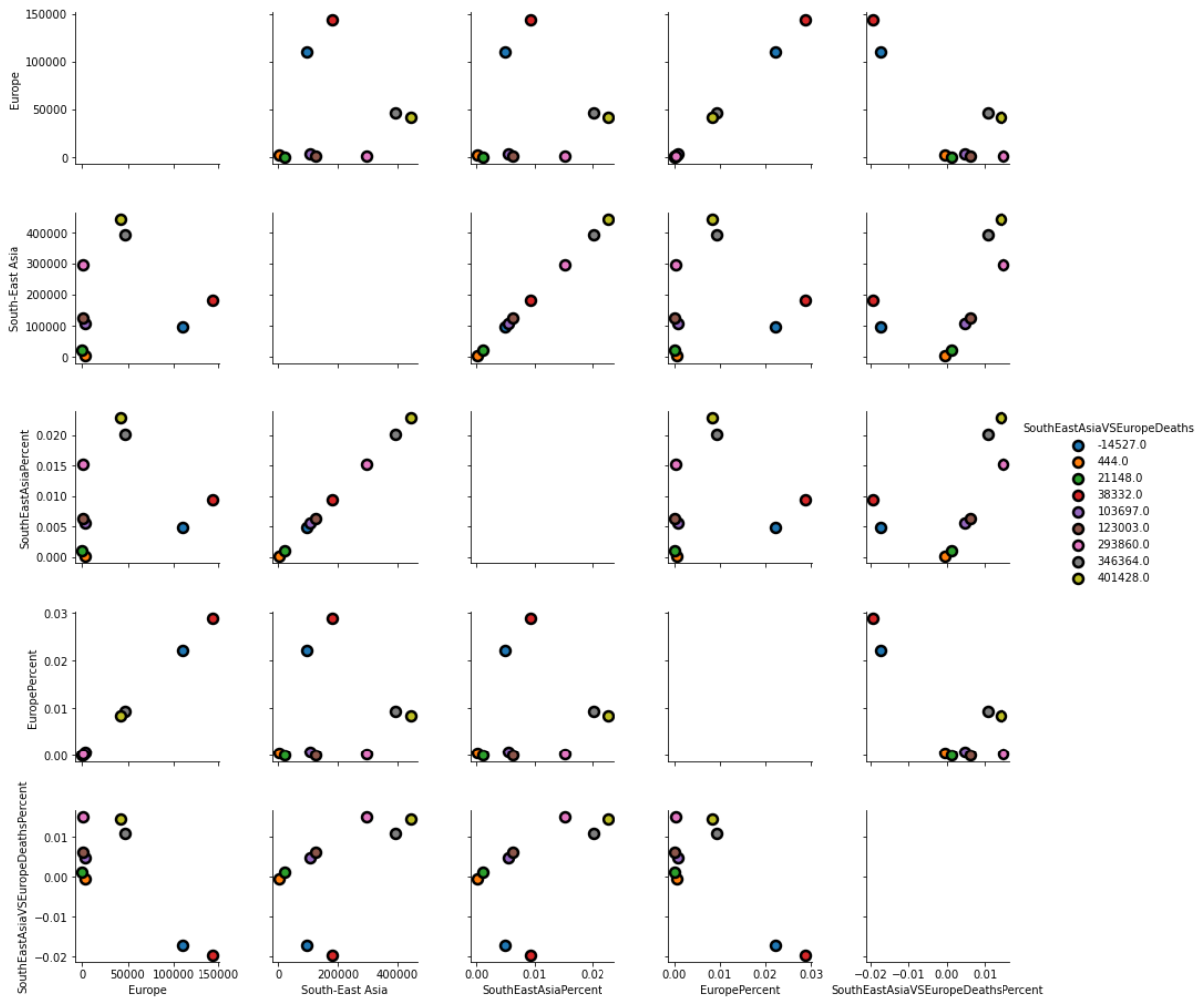
Figure 14 is a heatmap to visualise a summary between Europe and South-East Asia disease related deaths and the percentages of population.

Disease	Europe	South-East Asia	SouthEastAsiaVSEuropeDeaths	SouthEastAsiaPercent	EuropePercent	SouthEastAsiaVSEuropeDeathsPercent
Lower respiratory infections	4.2e+04	4.4e+05	4e+05	0.023	0.0084	0.014
Diarrhoeal diseases	1.2e+03	3e+05	2.9e+05	0.015	0.00024	0.015
Malaria	0	2.1e+04	2.1e+04	0.0011	0	0.0011
Tuberculosis	6.8e+02	1.2e+05	1.2e+05	0.0064	0.00014	0.0062
Trachea, bronchus, lung cancers	1.1e+05	9.6e+04	-1.5e+04	0.0049	0.022	-0.017
Other cancers*	1.4e+05	1.8e+05	3.8e+04	0.0094	0.029	-0.019
Stroke	4.6e+04	3.9e+05	3.5e+05	0.02	0.0093	0.011
Asthma	3.7e+03	1.1e+05	1e+05	0.0055	0.00075	0.0048
Other respiratory diseases	2.9e+03	3.3e+03	4.4e+02	0.00017	0.00057	-0.0004

Figure 14

Conclusion

Key findings

South-East Asia leading causes of total number of deaths compared to Europe are:
1. Diarrhoea
2. Malaria
3. Other cancers
4. Asthma
5. Stroke
6. Tuberculosis
7. Lower respiratory infections
8. Other respiratory diseases

Whereas for Europe compared to South-East Asia is:
1. Lung Cancer

In terms of percentages of population, the leading diseases for South-East Asia compared to Europe are:
1. Diarrhoea
2. Malaria
3. Stroke
4. Asthma
5. Tuberculosis
6. Lower respiratory infections

Whereas for Europe compared to South-East Asia is:
1. Trachea, bronchus, lung cancers

2. Other cancers
3. Other respiratory diseases

Evaluations

Visualisations are clear and users can see differences between Europe and South-East Asia quickly without having to perform any calculations. The heatmap shows different values thus is using standard form for the numeric values which makes it difficult to see the large numbers easily. The pie chart has exploded slices to show the very small percentages clearly.

I cleaned and pre-processed the data well which allowed for clear columns and data and visualisations. To improve I would remove the row 'total attributable deaths' as I calculate this in terms of populations to allow for accurate analysis and comparisons, the row did not provide extra analysis. I would also look for yearly data to compare the death rates across many years to also see if the preventative measures are having positive impacts on the diseases death rates.

I have learnt more about what visualisations work, ways to make visualisations clearer and what are appropriate visualisation plots to use for different types of data.

References

Who.int. 2018. Preventing Disease Through Healthy Environments: A Global Assessment of The Burden Of Disease From Environmental Risks. [online] Available at: Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks (who.int) *[Accessed 1 December 2020].*

World Health Organization. 2016. Preventing Disease Through Healthy Environments: A Global Assessment of The Burden Of Disease From Environmental Risks. [online] Available at: WHO | Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks *[Accessed 2 December 2020].*

Who.int. 2016. Quantifying Environmental Health Impacts. [online] Available at: Quantifying environmental health impacts (who.int) *[Accessed 1 December 2020].*

Oecdobserver.org. 2004. Health and The Economy: A Vital Relationship - OECD Observer. [online] Available at: Health and the economy: A vital relationship - OECD Observer *[Accessed 6 December 2020].*

William W Dressler, Culture and the risk of disease, British Medical Bulletin, Volume 69, Issue 1, June 2004, Pages 21–31, Available at: https://doi.org/10.1093/bmb/ldh020 *[Accessed 4 December 2020].*

Euro.who.int. 2012. Leading Risk Factors for Health In Europe: Fact Sheet. [online] Available at: Leading risk factors for health in Europe Fact Sheet (who.int) *[Accessed 4 December 2020].*

Oecd-ilibrary.org. 2018. OECD Library | Health at A Glance: Asia/Pacific 2018: Measuring Progress Towards Universal Health Coverage. [online] Available at: OECD iLibrary | Health at a Glance: Asia/Pacific 2018: Measuring Progress towards Universal Health Coverage (oecd-ilibrary.org) [Accessed 6 December 2020]

Euro.who.int. 2016. European Strategy for Tobacco Control. [online] Available at: EUR/02/50 (who.int) [Accessed 6 December 2020].

Coker, R., Hunter, B., Rudge, J., Liverani, M, 2011. Emerging infectious diseases in southeast Asia: regional challenges to control. The Lancet, 377(9765), pp.599-609, Available at: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(10)62004-1/fulltext?rss=yes

Appendices

Reading data

Out[3]:

Updated Table A2.3. Deaths attributable to the environment, by WHO region and income status, 2016§	Unnamed: 1	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9	Unnamed: 10	Unnamed: 11	Unnam 12
NaN	World	World	Africa	Americas	NaN	Eastern Mediterranean	NaN	Europe	NaN	South-East Asia	Western Pacific	NaN
NaN	Total \n(all ages)	\n0-4 years	All	HIC	LMIC	HIC	LMIC	HIC	LMIC	All	HIC	LMIC
Population	7430261888	673904396	1019920205	382020167	610136393	54017611	610317786	499400125	416914475	1947631149	213384103	167651
Total deaths	56188823	5599415	8715561	3162224	3603281	155238	3934487	4830599	4317888	13754870	1794181	119204
Total environmental deaths	13668365	1573296	2534165	421463	684808	37812	945743	696167	659833	4068942	268195	334944
...	...	...	...	...	...	...	...	...	...	...	...	...
WHO regional groupings: World Health Sta...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
see "Country grouping" worksheet of this...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
# Malnutrition and consequences	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
HIC: High-income countries	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
LMIC: Low- and middle-income countries	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

ws × 13 columns

Out[4]:

	Europe	South-East Asia
Disease		
Population	499400125	1947631149
Total deaths	4830599	13754870
Total environmental deaths	696167	4068942
Lower respiratory infections	41849	443276
Diarrhoeal diseases	1210.79	295070
Malaria	0	21148
Tuberculosis	683.498	123686
Trachea, bronchus, lung cancers	110544	96016
Other cancers*	144055	182387
Stroke	46356.6	392720
Asthma	3730.96	107427
Other respiratory diseases	2853.58	3297

Cleaning the data

```
In [5]:   # How many rows and columns
          select_rows.shape
```

Out[5]: (12, 2)

```
In [6]:   # Changing data type to int
          select_rows['Europe'] = select_rows['Europe'].astype(int)
```

```
In [7]:   # Changing data type to int
          select_rows['South-East Asia'] = select_rows['South-East Asia'].astype(int)
```

```
In [8]:   # Checking data types
          select_rows.dtypes
```

```
Out[8]: Europe             int32
        South-East Asia    int32
        dtype: object
```

```
In [9]:   # Checking for missing data
          select_rows.isnull().sum()
```

```
Out[9]: Europe             0
        South-East Asia    0
        dtype: int64
```

```
In [10]:  # Checking for duplicated data
          select_rows.duplicated().sum()
```

Out[10]: 0

Data manipulations

Statistics

In [11]: `# Statistics for disease related deaths and total population and total deaths`
`select_rows.describe()`

Out[11]:

	Europe	South-East Asia
count	1.200000e+01	1.200000e+01
mean	4.210651e+07	1.639267e+08
std	1.440167e+08	5.617351e+08
min	0.000000e+00	3.297000e+03
25%	2.442250e+03	1.045742e+05
50%	4.410200e+04	2.387285e+05
75%	2.820830e+05	1.349692e+06
max	4.994001e+08	1.947631e+09

In [12]: `# Statistics for disease related deaths`
`select_rows.iloc[3:12].describe()`

Out[12]:

	Europe	South-East Asia
count	9.000000	9.000000
mean	39030.888889	185003.000000
std	53834.262121	157929.933221
min	0.000000	3297.000000
25%	1210.000000	96016.000000
50%	3730.000000	123686.000000
75%	46356.000000	295070.000000
max	144055.000000	443276.000000

Out[14]:

Disease	Europe	South-East Asia	SouthEastAsiaVSEuropeDeaths
Lower respiratory infections	41848	443276	401428.0
Diarrhoeal diseases	1210	295070	293860.0
Malaria	0	21148	21148.0
Tuberculosis	683	123686	123003.0
Trachea, bronchus, lung cancers	110543	96016	-14527.0
Other cancers*	144055	182387	38332.0
Stroke	46356	392720	346364.0
Asthma	3730	107427	103697.0
Other respiratory diseases	2853	3297	444.0

Out[16]:

Disease	Europe	South-East Asia	SouthEastAsiaVSEuropeDeaths	SouthEastAsiaPercent
Population	499400125	1947631149	NaN	NaN
Total deaths	4830599	13754870	NaN	NaN
Total environmental deaths	696167	4068942	NaN	NaN
Lower respiratory infections	41848	443276	401428.0	0.022760
Diarrhoeal diseases	1210	295070	293860.0	0.015150
Malaria	0	21148	21148.0	0.001086
Tuberculosis	683	123686	123003.0	0.006351
Trachea, bronchus, lung cancers	110543	96016	-14527.0	0.004930
Other cancers*	144055	182387	38332.0	0.009365
Stroke	46356	392720	346364.0	0.020164
Asthma	3730	107427	103697.0	0.005516
Other respiratory diseases	2853	3297	444.0	0.000169

Out[18]:

Disease	Europe	South-East Asia	SouthEastAsiaVSEuropeDeaths	SouthEastAsiaPercent	EuropePercent
Population	499400125	1947631149	NaN	NaN	NaN
Total deaths	4830599	13754870	NaN	NaN	NaN
Total environmental deaths	696167	4068942	NaN	NaN	NaN
Lower respiratory infections	41848	443276	401428.0	0.022760	0.008380
Diarrhoeal diseases	1210	295070	293860.0	0.015150	0.000242
Malaria	0	21148	21148.0	0.001086	0.000000
Tuberculosis	683	123686	123003.0	0.006351	0.000137
Trachea, bronchus, lung cancers	110543	96016	-14527.0	0.004930	0.022135
Other cancers*	144055	182387	38332.0	0.009365	0.028846
Stroke	46356	392720	346364.0	0.020164	0.009282
Asthma	3730	107427	103697.0	0.005516	0.000747
Other respiratory diseases	2853	3297	444.0	0.000169	0.000571

Differences between South-East Asia vs Europe diseases as percentages

In [19]:
```python
# Europe- disease related deaths percentages
select_rows['SouthEastAsiaVSEuropeDeathsPercent'] = (select_rows['SouthEastAsiaPercent'].iloc[3:12] - select_rows['EuropePercent'].iloc[3:12])
select_rows['SouthEastAsiaVSEuropeDeathsPercent'].round(6)
```

Out[19]:
```
Disease
Population                            NaN
Total deaths                          NaN
Total environmental deaths            NaN
Lower respiratory infections       0.014380
Diarrhoeal diseases                0.014908
Malaria                            0.001086
Tuberculosis                       0.006214
Trachea, bronchus, lung cancers   -0.017205
Other cancers*                    -0.019481
Stroke                             0.010882
Asthma                             0.004769
Other respiratory diseases        -0.000402
Name: SouthEastAsiaVSEuropeDeathsPercent, dtype: float64
```