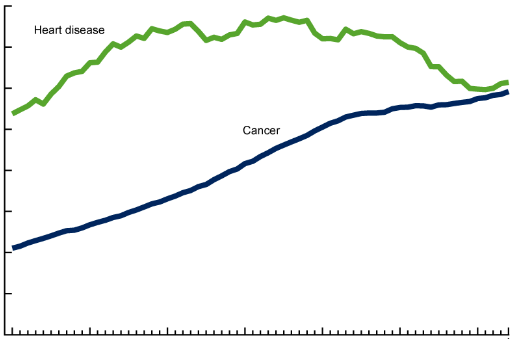
**Which health issues should people pay the most attention to? An analysis on 2014 world health.**

****

**A Brief Background First**

I downloaded a dataset of 2014 world heath from Kaggle’s website. The title of the dataset is “CAUSES OF DEATH IN THE WORLD 2014” and is in csv format. This dataset was provided by the United Nations to the public with data reported by the World Health Organization. The data consists of information for the year 2014 on various countries, gender, age groups, causes of death, and number of deaths.

I was drawn to this dataset because I interested in understanding the health of the world better, to see if I can uncover any key findings and insights from the data. One example of something I was looking forward to find out was whether there might be any significance in the results based on location, gender, and age demographics.

In a more macro perspective, measuring how many people die each year and why they died or information on infant mortality rates are important ways in assessing the effectiveness of a country’s health system.

A consideration to make note of is that high-income countries may have mature systems/processes in place for collecting information on causes of death. Many low- and middle-income countries may not have effective systems and processes resulting in data not be reported and issues with quality of data. Improvements in producing high quality cause-of-death data are crucial for improving health and reducing preventable deaths in these countries.

**More Specifics about the Data**

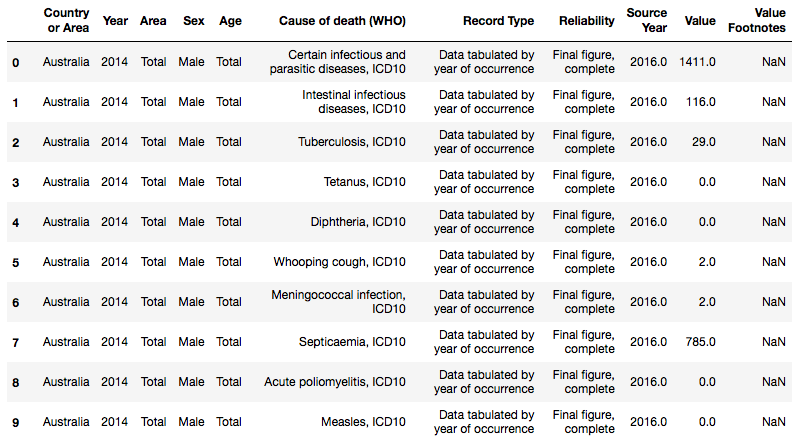
What else do we know about the data set? From initial exploration and surveying of the data, the data seemed pretty clean, not too large, would be a good start for my first Analytical Report. The following are my general observations:

The data provides:

* The number of deaths by Country or Area, Gender, Age Groups, and Cause
* The data has additional rows that represent subtotal number of Deaths by All Causes of Death, and also by each Cause of Death for each Country or Area and Gender.

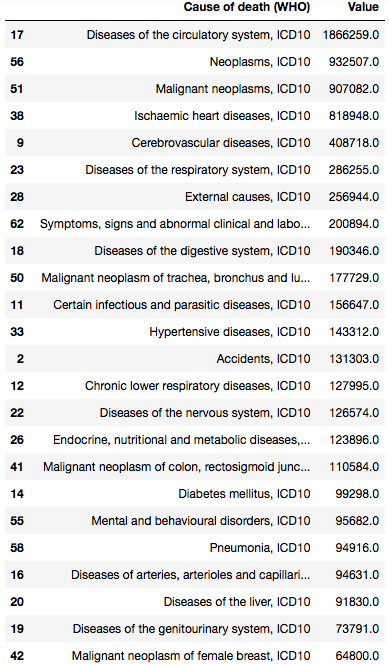
**A Deeper Look into the Data**

Exploring the first few rows of data:

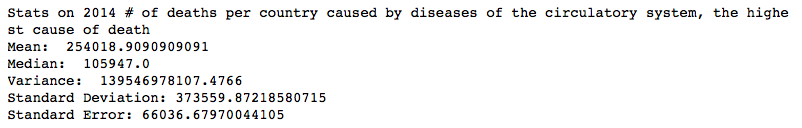


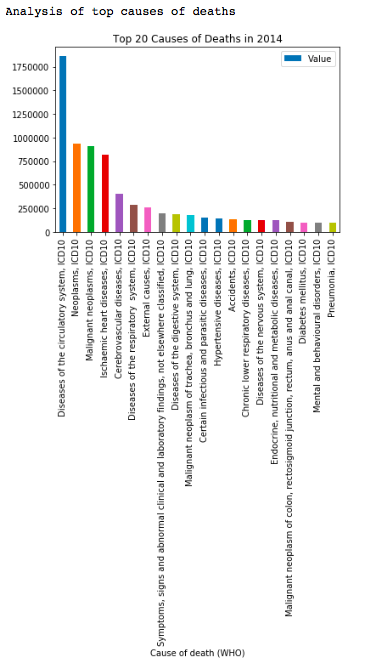
**What are the top causes of death?**

Total # of deaths per cause of death



A statistical summary on the totals for top cause of death across countries showing central tendency and variance



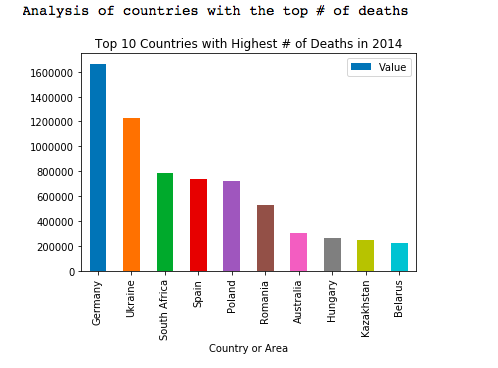


**Which country had the highest total number of deaths?**

Here are the top 10 countries with highest total numbers of deaths.

Are there any significance to the results that might imply a correlation to the types of countries whether a 1st world country or a 3rd world country…or whether a country generally with larger or smaller populations?

From the results, I don’t seem to see a correlation. I was surprised to see that Germany was at the top since based on my general knowledge, the country is a 1st would country, with advanced technologies, high standard of living, and a leader amongst the world. I would expect that it would be an overall healthier country.



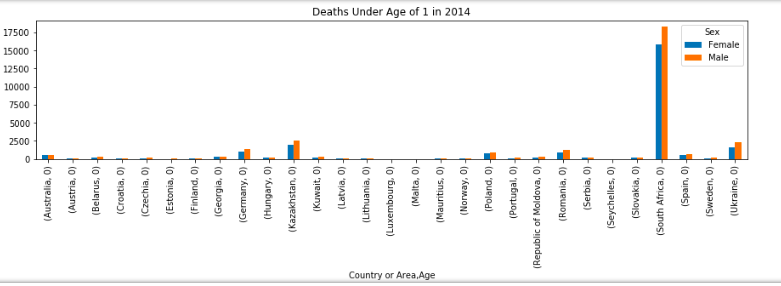
**A deeper look into infant mortality**

Why focus on infant mortality? Infant mortality information is widely used as a measure of population health and the quality of health care. Infant mortality is defined as the death of an infant before their first birthday.

Infant mortality information can provide indicators on the following:

* Community health status
* Poverty and socioeconomic status levels in a community
* Availability and quality of health services and medical technology

Here is an analysis on the data set for numbers of infant mortality (deaths ages under 1 year old) and by gender across countries.



**Conclusion**

People should pay closer attention to their health on disease of the circular system since that’s the top cause of death. It’ll be interesting to run some further research and analysis on the data to see if this is the top cause of death for most countries.

Germany has the highest numbers of deaths. People in Germany may want to pay close attention to their overall health and maybe to preventative care or to improve the countries health policies/system. We can run some future analysis on the data to see what are the top causes of death for Germany.

It is alarming to see the huge number of infant deaths in South Africa. This should be concerning to the people in the country as it is a significant indication that there are major issues with the country’s health. For future analysis, we can see what are the top causes of deaths in South Africa for ages under 1 year old.

**Further Research**

Here are some additional ideas for other general future analysis on the data that’s available in this particular data set:

* Analysis of number of deaths and cause of deaths per age group.
* Were there particular age groups that had more deaths?
* Were there more deaths for certain age groups that were attributed to particular causes of deaths?

Outside the scope of this data set, we can have a more expanded research project to tap into more robust data available from the World Health Organization. We might be able to get more comprehensive data for more effective analysis of the world’s health with the following additional data scope:

* More years of data in addition to just 2014
* Total population size per country
  + This would provide a better overall picture of death rate relative to each country by analyzing percentage stats that may be a better indicator of relative health of each country more fairly. For example, countries with a larger population may have higher amount of deaths but still a very small percentage of death given the large size of the population which may mean it still have good overall health compared to a country with a lower death amount but much small population with yields a higher percentage and suggest it being a country with worse health.

Finally, it would also be a better analysis of the world if more countries were included since the data available in the dataset looks to be only a subset of countries that are part of the UN, for ex, United States is not even included.

**Appendix**

The python code used to build the analysis and figures in this report can be found in the following Github site/link:

<https://github.com/TLThinkful/cosmo/blob/master/CapstoneAnalyticReport-Final.ipynb>