Exploring Life Expectancy, Aid, Health Services, and Access

Dominic Delmolino
Sarah Kelley
Andrea Pope

Project 2 – Final Presentation
INFO W18 – Python Bridge 2

Introduction:

Life expectancy – and recently more sophisticated measures including healthy life expectancy and life expectancy at age 60 – has long been an important measure of development and human progress. Across the globe, life expectancy rose steadily and country after country climbed out of poverty. Public health advances reduced infant mortality, although in recent decades, the HIV epidemic slashed life expectancy in Sub-Saharan Africa. In this context of widespread but not universal progress, exploring factors which may influence life expectancy is both academically and practically important. Fortunately, increasingly accurate data collection across the world allows us deeper insights into these factors, which both provides deeper understanding and suggests areas in which efforts might be made to improve life expectancy figures.

We focus on the relationship between life expectancy and access to clean water/sanitation, aid spending, HIV prevalence and government spending on health care, using data from the World Health Organization (WHO). We hypothesize that clean water and sanitation access and health care spending with be positively related with life expectancy, while HIV prevalence will have a strong negative relationship. The relationship between receipt of aid and life expectancy is complex, because countries that are struggling (and so have low life expectancies) receive more aid but we would also hope that aid itself improves outcomes in the countries that receive it. With the data we have, we cannot demonstrate a causal relationship between any of the factors explored, but simply focus on exploring the correlations within these factors.

Methodology:

Data sets were downloaded from the World Health Organization ("WHO") website. The core set of data for our analysis was the Life Expectancy data set. We concentrated on three (3) key statistics, specifically:

- Life Expectancy at Birth
- Life Expectancy at Age 60
- Healthy Life Expectancy (HALE) at birth

The data was provided for the years 1990, 2000, 2012, and 2013 for 194 different countries. The HALE data set takes into account degraded health due to disease or injury.

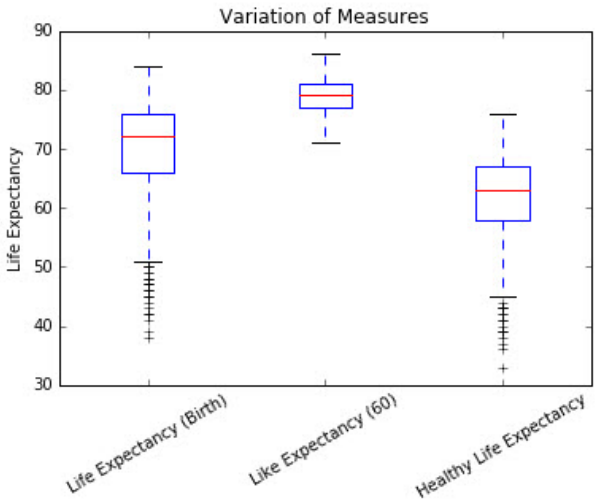
Against the life expectancy data, we compared 4 different factors that could be related to life expectancy, specifically environmental factors (access to clean drinking water and sanitation capabilities), HIV prevalence, spending on health services and aid provided to each country. In each case, we have conducted descriptive analysis of the individual data set before moving to multivariate analysis exploring the relationship between each of these factors and life expectancy.

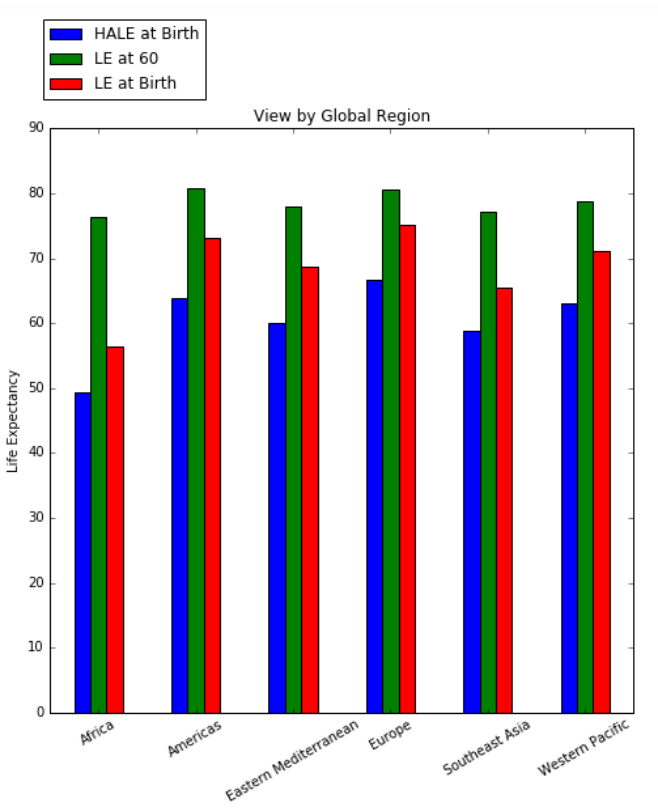
For all of the data sets, initial data cleaning was completed to structure the data such that is allowed easier analysis, and to remove unnecessary information. For instance, in several of the data frames a "Country;Date" field was separated into 2 separate fields, column names were simplified, null values were managed, and years of data that did not exist for other variable were lost. The appendix contains a complete listing of all variable used and information about accessing the data.

Descriptive Results:

Life Expectancy:

This boxplot shows both the variation between the various Life Expectancy measures and the variation of the measures across the dataset of 194 countries. As anticipated, life expectancy at age 60 has a higher mean, as the starting point takes into account they have reached age 60. Similarly anticipated, when you take into account expected years of Healthy Life (via the HALE measure), the mean and upper quartile drop. This would naturally be lower, as those that face disease or illness, may not die, but be unwell.



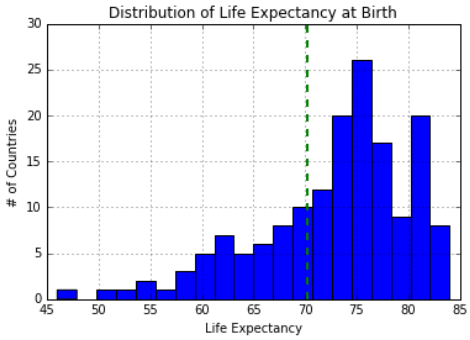


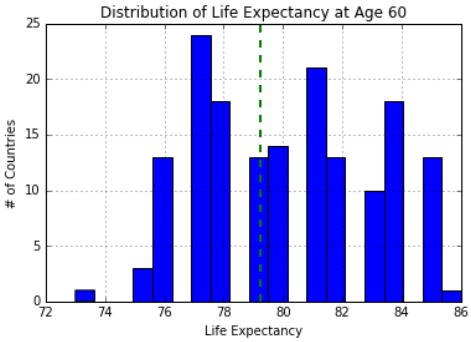
Further delving into these measures, this chart evaluates across each region of the world. Africa has the lowest Life Expectancy at Birth and Healthy Life Expectancy, but once at the age of 60 in Africa, your life expectancy is not significantly different than other regions of the world.

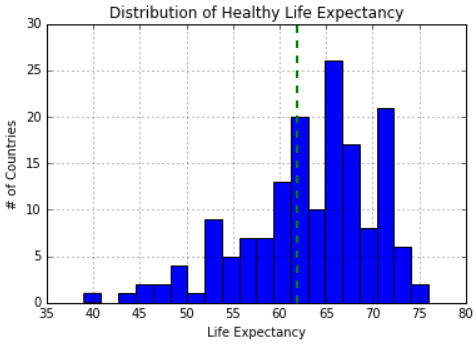
As anticipated, Europe and the Americas share honors for the longest life expectancy.

Exploring Life Expectancy, Aid, Health Services, and Access

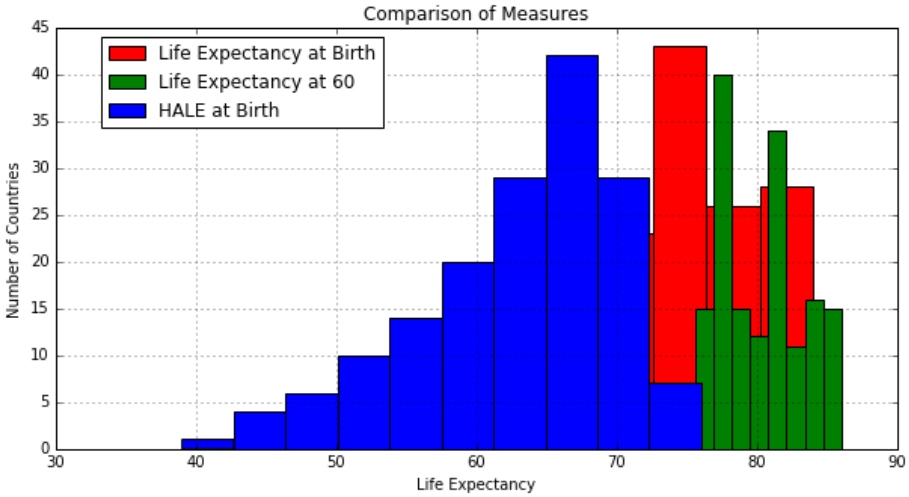
The next 3 charts show the distribution of these 3 measures, including an indicator of the mean (green line) across all of the countries.



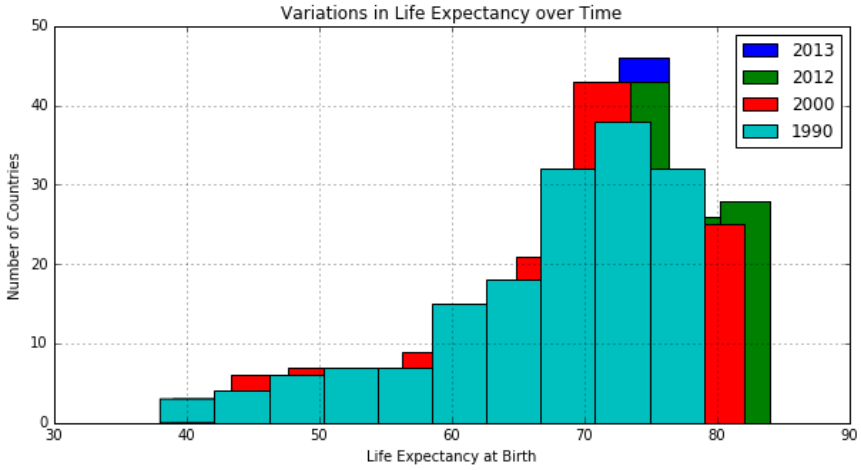




Combined together, as below, the chart supports our original assumptions. Healthy life expectancy trends lower than the life expectancy measures, and once the age of 60 is reached, the numbers that reach over 80 is significantly high.

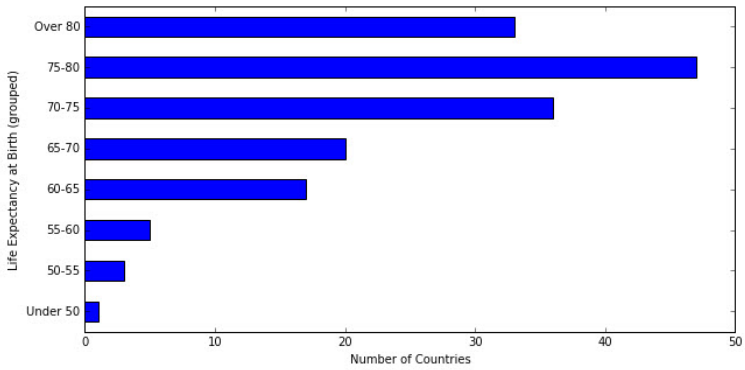


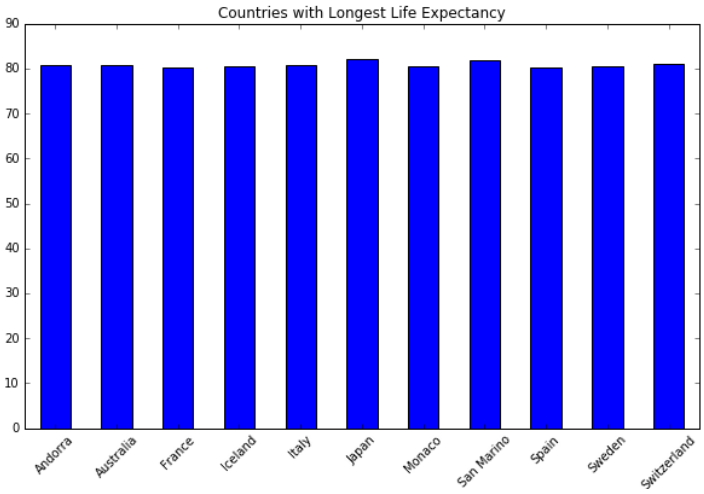
For reference purposes, the chart below shows Life Expectancy at Birth for each year of data provided. The trending remains consistent, across the years of data.



Exploring Life Expectancy, Aid, Health Services, and Access

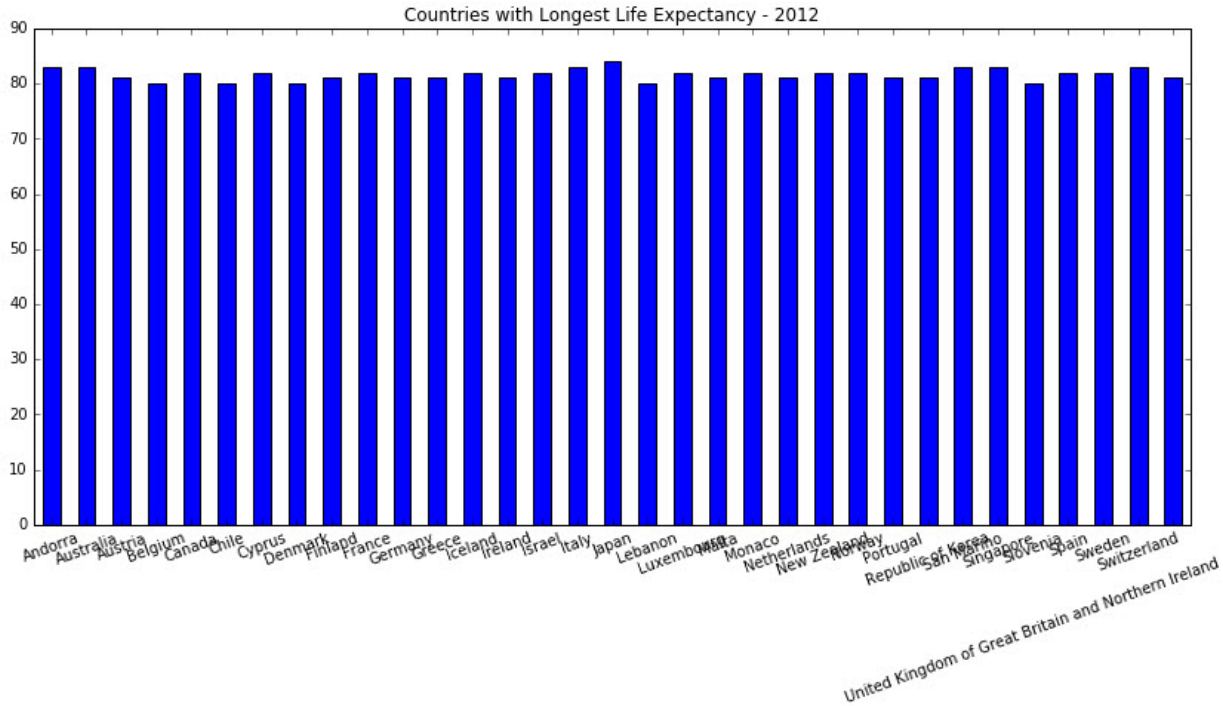
Looking a slightly different way, this chart shows the number of countries with Life Expectancy within specific ranges for 2012. Amazingly, the largest is the 75-80 range, but still a large number of countries have Life Expectancies below 65.





Diving into country specific details for Life Expectancy at Birth (mean against all years), there are only 2 countries that have a life expectancies under 50: Zambia and Sierra Leone.

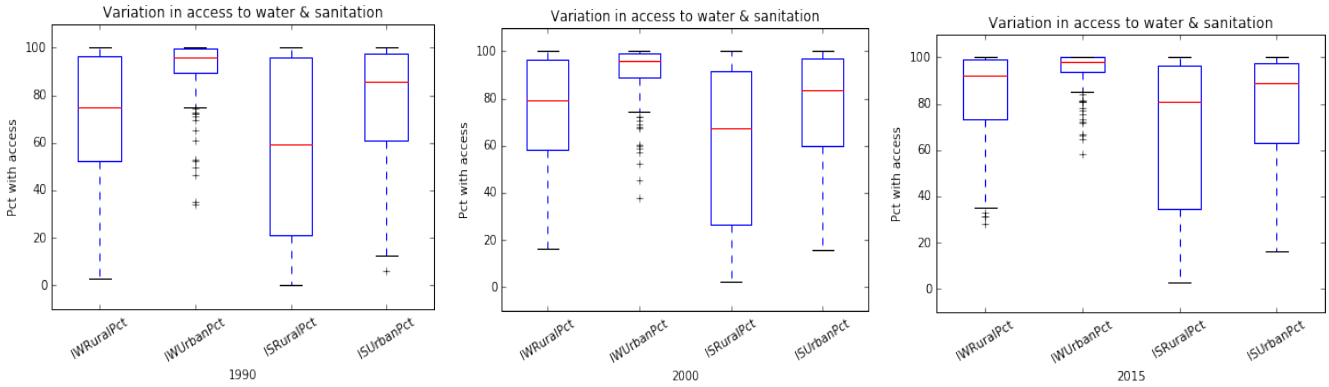
However, there are eleven countries that have maintained an average life expectancy (at Birth) over 80, as shown on the left. If you look at only 2012, that numbers jumps to 33 countries!



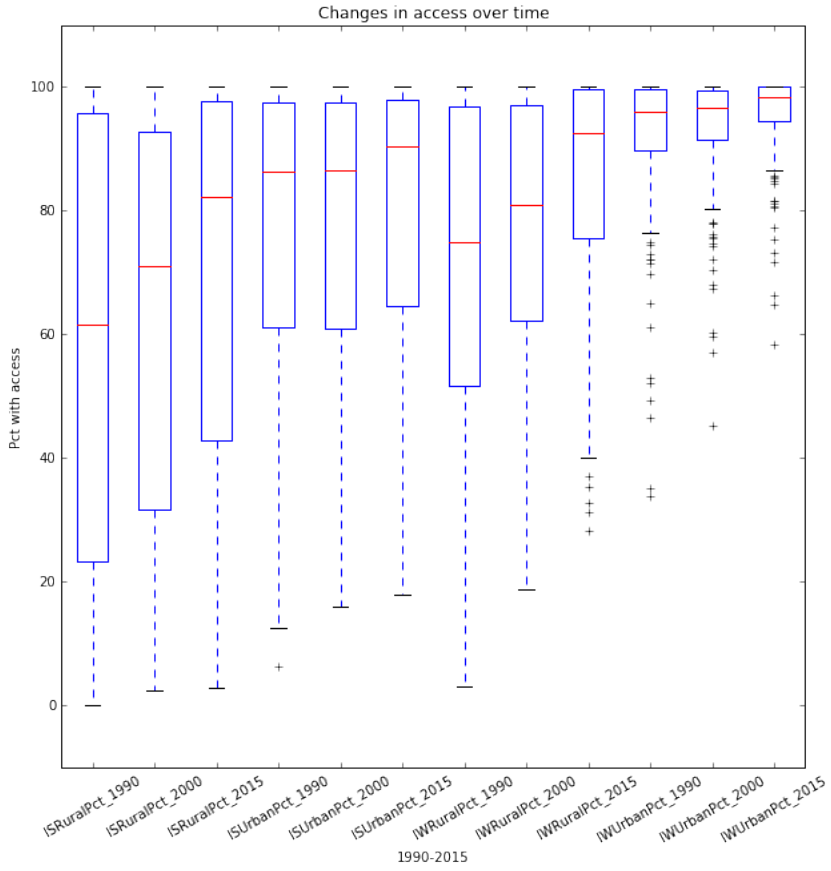
Exploring Life Expectancy, Aid, Health Services, and Access

Environmental Factors (Hygiene):

We examined changes in access to clean water and sanitation over time:



And noticed that rural access to clean drinking water appeared to have the greatest change over time from a distribution and average perspective. Plotting all elements on the same graph, we can see that rural access to water had both the highest change in size of inter-quartiles as well as the highest percentage. Size of range decreased by 21 points and 46.% compared to rural sanitation (decreased by 17 points, but only 24%), urban water access (4 points and 44%), and urban sanitation (3 points and only 8%)



Exploring Life Expectancy, Aid, Health Services, and Access

Correlations of Life Expectancy against Environmental Hygiene factors and cross correlation between Hygiene factors:

	IWRuralPct	IWUrbanPct	IWPct	ISRuralPct	ISUrbanPct	ISPct
HALEatBirth	0.778963	0.630919	0.803787	0.831852	0.844142	0.861756
LEat60	0.595615	0.514409	0.634142	0.654427	0.638081	0.679017
LEat60-Adj	0.595615	0.514409	0.634142	0.654427	0.638081	0.679017
LEatBirth	0.772890	0.638578	0.811423	0.820131	0.834542	0.856440
IWRuralPct	1.000000	0.740516	0.956718	0.826171	0.783679	0.819344
IWUrbanPct	0.740516	1.000000	0.824008	0.632543	0.643147	0.658210
IWPct	0.956718	0.824008	1.000000	0.802975	0.809191	0.840759
ISRuralPct	0.826171	0.632543	0.802975	1.000000	0.903372	0.973203
ISUrbanPct	0.783679	0.643147	0.809191	0.903372	1.000000	0.959781
ISPct	0.819344	0.658210	0.840759	0.973203	0.959781	1.000000

While access to sanitation facilities had the highest correlation to Life Expectancies (and among Life Expectancies, most highly correlated with “At Birth” expectancies), both rural ***and*** urban sanitation access is most highly correlated with rural access to clean drinking water. In fact, even urban sanitation access showed higher correlation with rural access to clean drinking water than to urban access to clean drinking water.

Comparing rural access to clean drinking water with urban access to clean drinking water, across all life expectancies the rural access was more highly correlated than the urban access.

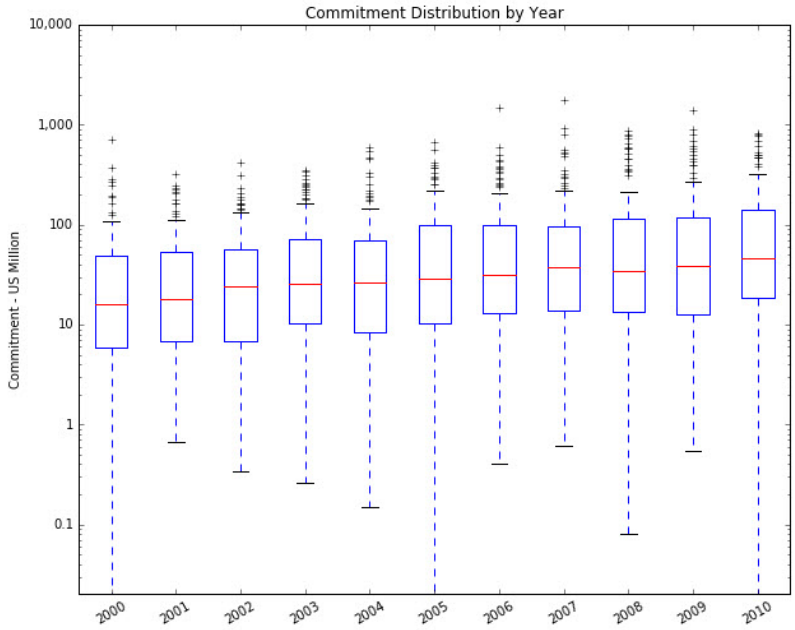
Given rural access to clean drinking water having such high correlations to sanitation and life expectancy, it appears that efforts to improve rural access to clean drinking water have the largest environmental / hygiene beneficial impact on life expectancy.

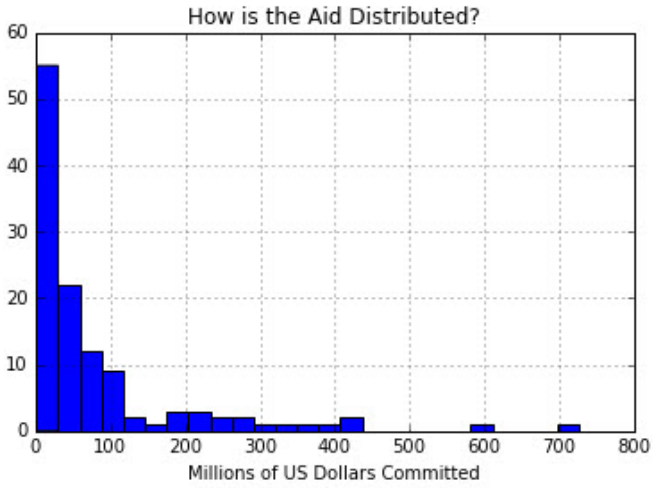
Exploring Life Expectancy, Aid, Health Services, and Access

AID:

Analysis of aid committed and disbursed to 119 different countries between the years 2000 and 2010 shows a fairly consistent distribution, with a total increase over time.

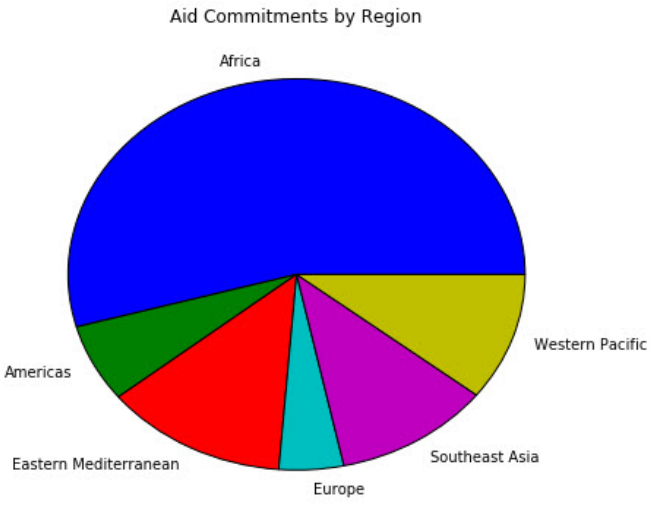
Please note the y-axis below is logarithmic.





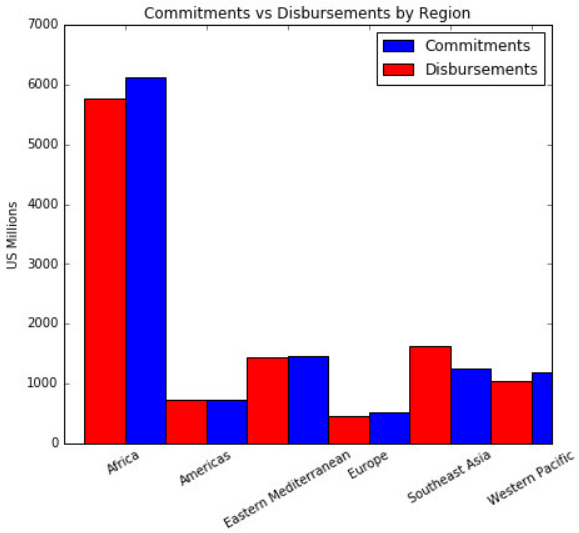
Most countries that receive aid, receive under $25 Million, as shown in this histogram, with a few significant outliers receiving over $300 million.

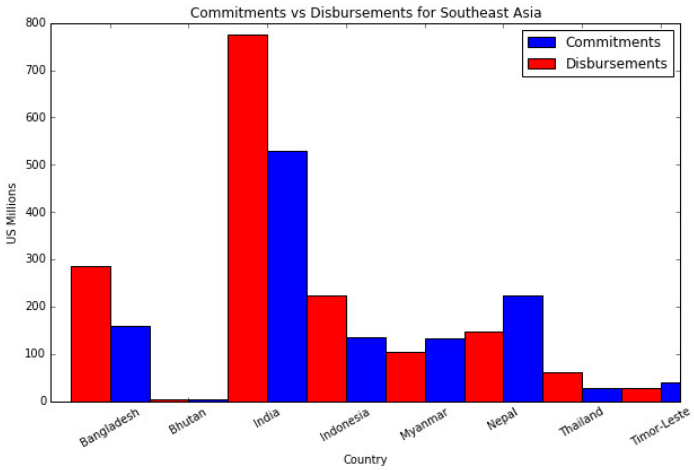
The chart to the right shows the distribution of aid by region of the world. As anticipated, Africa is the greatest recipient, with Europe and the Americas receiving the least.



Exploring Life Expectancy, Aid, Health Services, and Access

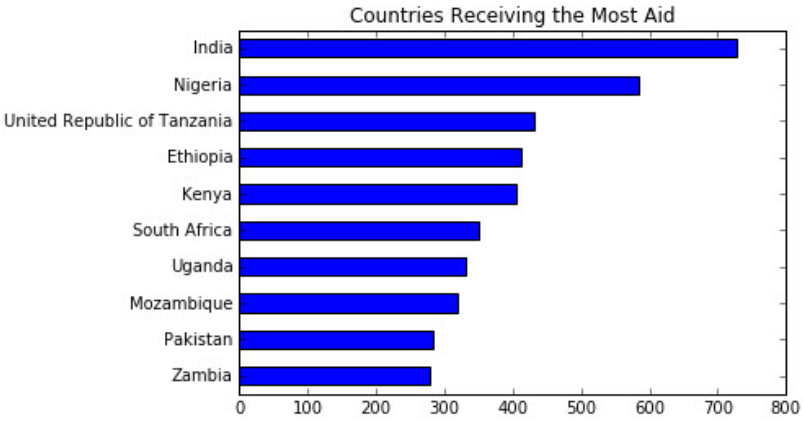
A quick review of commitments vs disbursements, for 2010 (latest year with available data) for each region, shows they are fairly consistent.



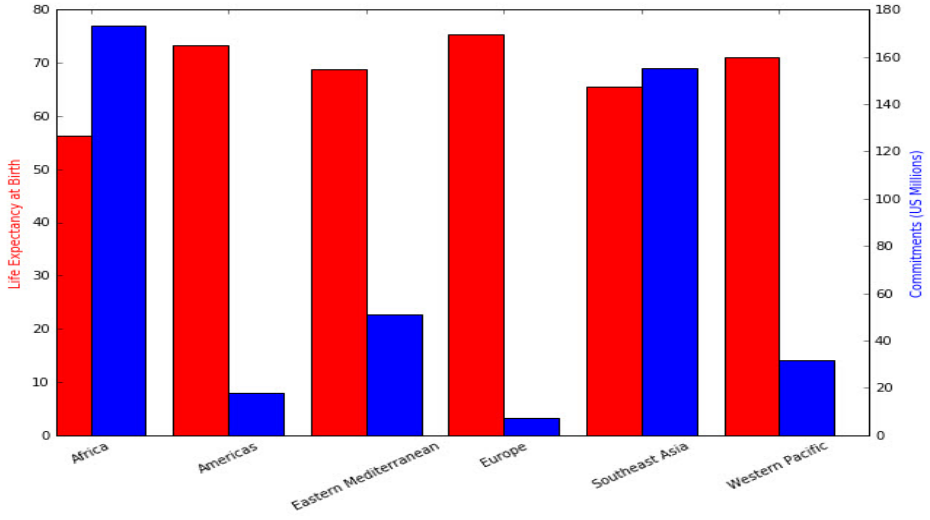


The biggest variation, by percentage, is in Southeast Asia. The chart to the left shows that broken by country. India has the greatest disparity between committed and disbursed.

India, in fact, has the most amount of aid committed, as illustrated below in the chart showing the top aid recipients; those receiving on average over $250 million per year.

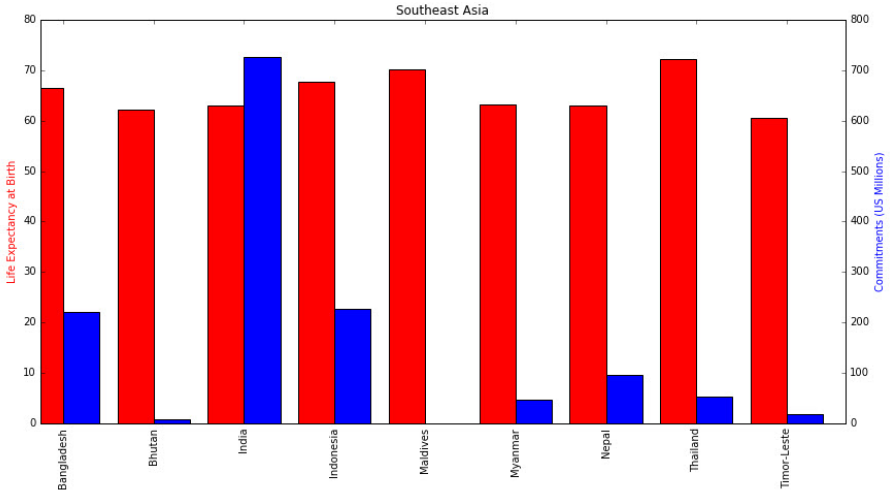


There is not a strong positive or negative correlation between life expectancy within a country, and the amount of aid received, with the exception of each end of the life expectancy range. The region with the highest life expectancy receives the least aid, and the region with the lowest life expectancy receives the most aid. This negative correlation was anticipated. *Please note the dual y-axis on the chart.*

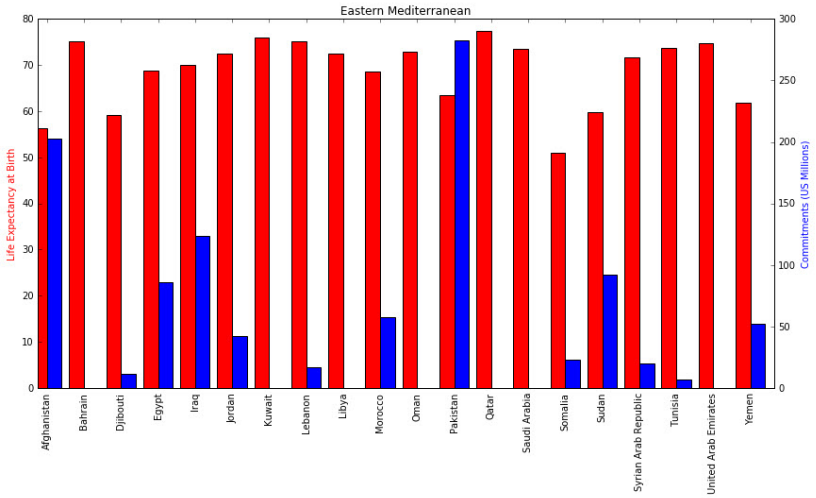


Exploring Life Expectancy, Aid, Health Services, and Access

One surprising result was Southeast Asia, given their fairly high life expectancy but also very large aid commitments, especially in comparison to the Eastern Mediterranean region, that has a fairly similar life expectancy, but receives much less aid. Each region is broken down by country.



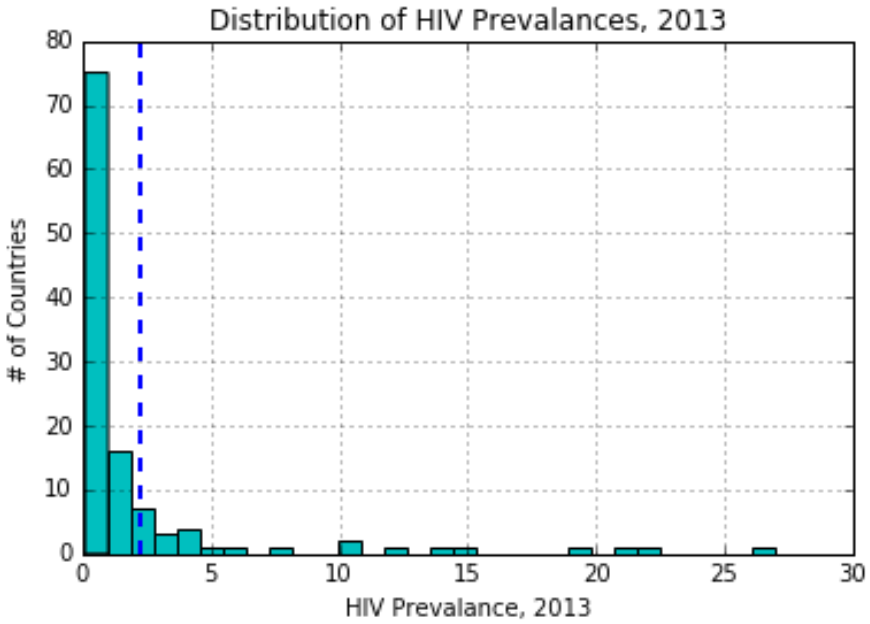
Again we see India with a very high aid commitment, and only one Southeast Asia country that receive no aid. Separately, in the Eastern Mediterranean region, a smaller percentage of countries within the region are receiving aid, however, for those that are, the life expectancies are lower, and more in line with Southeast Asia. This again supports the negative correlation between the amount of aid committed and Life Expectancy.



Exploring Life Expectancy, Aid, Health Services, and Access

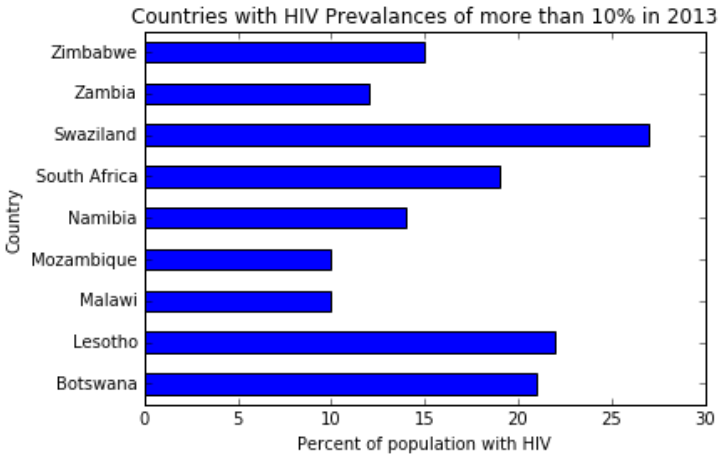
HIV Prevalence:

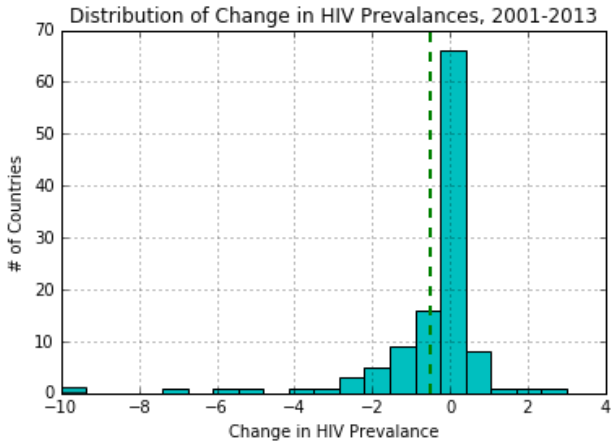
Over recent decades, the HIV epidemic in Sub-Saharan Africa has rapidly undermined hard-won gains in life expectancy across the region. Public health efforts have stemmed the tide slightly in many countries, but HIV remains an entrenched problem and continues to substantially impact health and quality of life across the region.



The distribution of HIV prevalence – measured as the percentage of the adult population living with HIV – in 2013 is heavily right skewed, with a median HIV prevalence of just .6% and a mean prevalence of 2.2% (shown by the dashed blue line in the graph to the left). In more than 75% of countries the mean HIV prevalence was less than 1.6%, but the most heavily affected country (Swaziland) had an HIV prevalence of more than 27%. This means that while the vast majority of nations have largely escaped the HIV epidemic, a small group of countries remains heavily impacted by it.

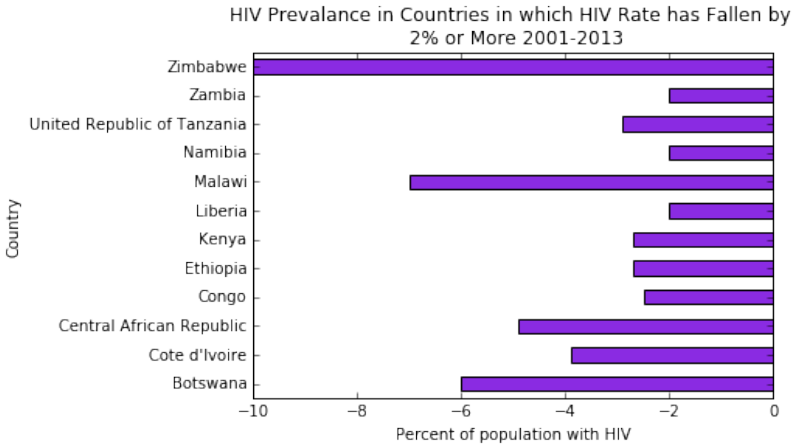
Given the long tail on the distribution, it is worth exploring in more detail the countries with particularly high rates of HIV. Looking at nations with an HIV prevalence of 10% or above (roughly two standard deviations above the mean), we find a range of Sub-Saharan African countries with HIV rates up to 27%. Swaziland has the highest HIV rate but Lesotho, Botswana and South Africa are not far behind, with HIV prevalences close to 20%. This is clearly a human tragedy, both for the victims of the disease themselves and for the countless children left orphaned by their parent's deaths. However, reason for hope remain with many countries starting to control the epidemic and reduce HIV rates. The average country saw its HIV rate fall by .5 % between 2001 and 2013, although in some countries the epidemic continued to swell.

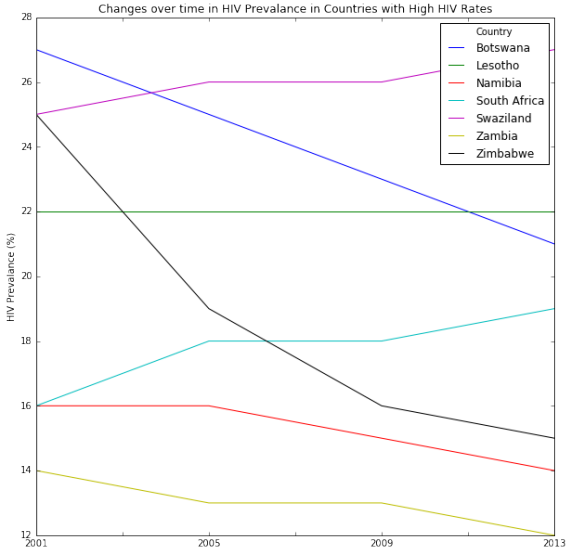




Exploring Life Expectancy, Aid, Health Services, and Access

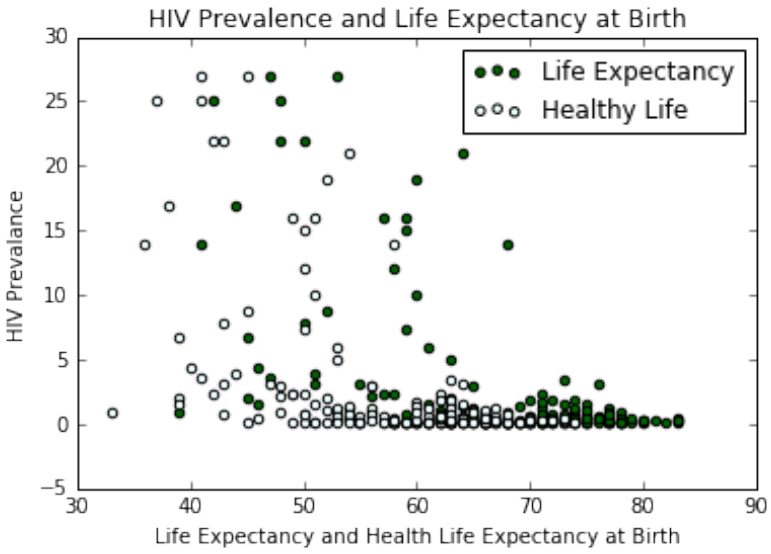
Looking individually at countries which have made a substantial impact on their HIV rates – reducing rates by 2% of the population or more (that is one standard deviation more reduction than the average nation) – we find mainly African nations, some with extremely substantial reductions in HIV prevalence over the 12 years between 2001 and 2013. In this graph, the reduction in rate refers to the decrease in





the percentage of the population with HIV (i.e. a change from 3% to 1% is a 2% fall in the HIV rate). We find the greatest reduction in Zimbabwe, where 10% fewer of the population had HIV in 2013 than in 2001, with Malawi and Botswana close behind with reductions of around 7% and 6 percent respectively. All of the countries which reduced their HIV rate by 2% or more were in Africa, which makes sense given that HIV rates are substantially higher in this region and other regions have less potential for reduction due to their lower starting rates. Both the fact that HIV rates generally fell between 2001 and 2013 and the fact that some countries achieved dramatic reductions are cause for hope. However, looking at the countries with the highest HIV prevalence rates today, we see mixed patterns over time, with some countries experiencing steep reductions in the HIV rate and other countries seeing their HIV rates stay steady or even increase.

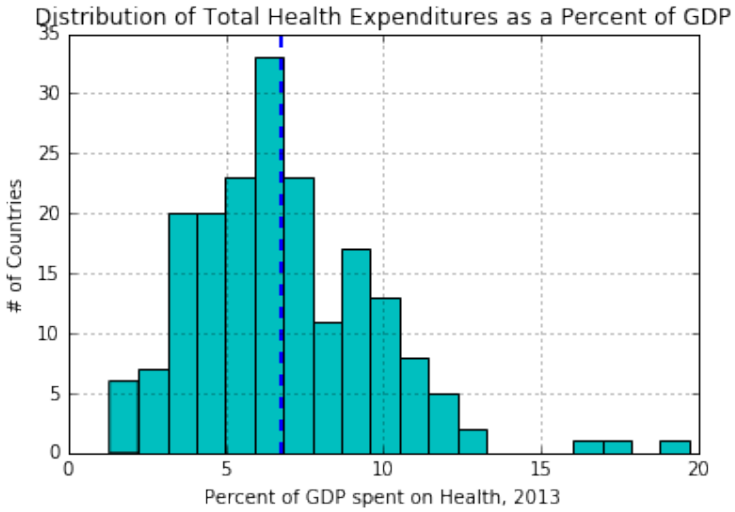
Turning to the relationship between HIV prevalence and life expectancy, we find a clear pattern with countries with high HIV prevalences generally having generally slightly lower life expectancies at birth and similarly lower healthy life expectancies at birth. However, there is not a clear linear relationship between the two variables although there is a noticeable trend. Correlations confirm this impression. Since HIV prevalence is markedly non-normally distributed (it is heavily right skewed), it is likely more appropriate (and conservative) to use Spearman's rho instead of a Pearson's correlation. Using this technique, we find that HIV Prevalence is correlated -.45 with life expectancy at birth and -.47 with healthy life expectancy at birth. Thus, we find a moderately strong negative relationship between the two variables and conclude that HIV levels are associated with a lower life expectancy and lower healthy life expectancy. While we cannot definitively make causal claims using this data, since HIV is caused by a viral infection (not by poor sanitation or



weakened health) we can reasonably assume that a higher HIV prevalence might cause a reduction in life expectancy rather than the other way around. While this does not control for other factors which may be related to both HIV and life expectancy, we can at least see a relationship between these two variables.

Health Services:

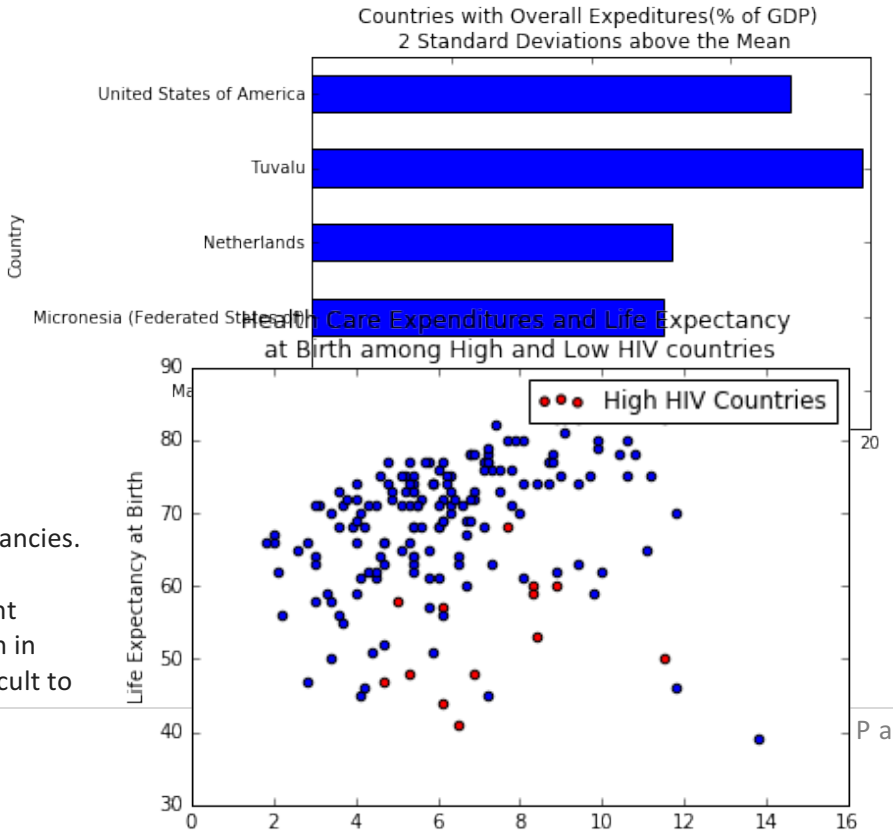
In addition to the relationship between life expectancy and Aid, sanitation and HIV prevalence, the relationship between health spending and life expectancy is also worth exploring. We find that health spending is positively related to life expectancy, although this could be because of confounding variables rather than that the spending itself is improving life expectancy.



Considering total spending – that is, spending by government, citizens and private insurers combined – on healthcare as a percentage of GDP, we find total spending as a percentage of GDP to be approximately normally distributed with the average nation spending 6.8% of their GDP annually on health services in 2013. A small group of outlier nations exist, spending more than 15% of their GDP on health services annually.

Focusing on these nations with higher spending as a percentage of GDP (more than two standard deviations above the mean), we find an interesting set of countries: the US and the Netherlands (both relatively wealthy developed nations) and Tuvalu, Micronesia and Marshall Islands (all very small island nations). It is possible (although not demonstrated by our data) that the increased percentage of GDP spent on health care in both the United States and the Netherlands reflects the fact that these two nations, unlike most developed economies in Europe, provide healthcare though a network of private insurers rather than through a more regulated system. This is, however, purely speculation and slightly beyond the scope of this paper.

Turning now to the relationship between, health expenditures and life expectancy we find slightly complicated picture. A causal relationship cannot be easily established because different scenarios are entirely plausible. It is plausible that spending more on health service could increase life expectancy. But it is also plausible that nations with unhealthy populations (whether due to HIV, poor public health, malnutrition, communicable diseases or obesity) would tend to both spend more on healthcare services and experience worse life expectancies. Alternatively, countries with bad health outcomes (including life expectancy) might spend more in an effort to improve health in their nation. In this context, it is very difficult to



interpret the relationship between life expectancy and the percentage of GDP spent on health services in a meaningful way. However, we do find that increased spending on health services is associated with increased life expectancy, as we can see in the scatter plot to the right. Countries with particularly high HIV rates are no exception – they tend to have slightly higher expenditure and lower life expectancies, but do not deviate from the general pattern (nor are these particular cases driving the correlation between expenditures and life expectancy). Looking at the correlation between the two variables we find a slight positive correlation of .21 between life expectancy and health care spending. However, as discussed above it is difficult to interpret this relationship in a meaningful way, beyond saying that we see a slight relationship between higher spending and and higher life expectancies.

Conclusion:

In this project we have sought to untangle a number of important factors which related to life expectancy, particularly in the developing world. We found that the average life expectancy cross-nationally was around 70, with the healthy life expectancy somewhat lower and the life expectancy at 60 substantially higher. Life expectancy was higher in Europe and the Americas and lowest in Africa.

Looking at the relationship between life expectancy and other factors that might be related to it, we find strong relationships between two of the most important public health issues in the developing world (sanitation and HIV) and life expectancy. We find that access to sanitation and clean water in both rural and urban areas is very strongly correlated with life expectancy. It is most strongly related with life expectancy at birth and less so with life expectancy at 60, suggesting that lack of sanitation and clean water plays a particularly important role infant mortality rather than mature-aged deaths. While we cannot demonstrate a causal relationship, it is certainly plausible to suggest that improving water and sanitation may improve life expectancy. Life expectancy is, unsurprisingly, negatively related to HIV prevalence, with countries with higher HIV prevalences experiencing generally lower life expectancies. However, there is room for hope as the average percent of nations population infected with HIV has fallen by 2 between 2001 and 2013 and many African nations with extraordinarily high HIV rates have made even bigger strides towards reducing the problem. While we cannot demonstrate causality, we can reasonably infer that falling HIV prevalences are likely to improve life expectancies, especially across Sub-Saharan Africa (which has unusually low life expectancies). Overall, we can conclude that both sanitation and water access and HIV prevalence is related to life expectancy, and likely that improving sanitation and reducing HIV would result in increased life expectancies.

When it comes to the relationships between finances and life expectancy, things are less clear. Regions with low life expectancies – particularly Africa – tend to receive large amounts of aid, and regions with higher life expectancies generally receive less aid. A notable exception is South East Asia which has a high average life expectancy but also high amount of aid. This may be because it has a mix of highly developed countries (e.g. Singapore) and larger, less developed countries that are major aid recipients (e.g. India). Overall, it is impossible with this data to establish the role aid might play in increasing life expectancies because countries that are struggling receive more aid and also have lower life expectancies. The relationship between life expectancy and health expenditure is similarly complex, and it is difficult to interpret the relationship meaningfully. However, we did find that spending a greater percentage of GDP on health services is weakly associated with a longer life expectancy.

Overall, the relationship between public health problems – lack of sanitation and clean water and HIV – and life expectancy is clear, while the role money – either through aid or through direct expenditures on health – might play in improving life expectancy is less clear from this data. We can likely conclude, thus, that improving sanitation and access to clean water and reducing HIV rates is likely to improve life expectancies. This effect should be particularly pronounced in countries with very high HIV rates. The general trend of increasing access to sanitation and decreasing HIV rate over time is therefore a positive sign that might portent continued improvements in life expectancy.

Sources:

Aid: http://apps.who.int/gho/data/node.main.A1626?lang=en
Environmental Factors: http://apps.who.int/gho/data/node.main.167?lang=en
Healthcare Services: http://apps.who.int/gho/data/node.main.75?lang=en
HIV: http://apps.who.int/gho/data/node.main.622?lang=en
Life Expectancy: http://apps.who.int/gho/data/node.main.3?lang=en
Source files for Region/Country Mapping:
- Africa: http://www.afro.who.int/en/countries.html
- Americas: http://www.paho.org/hq/index.php?option=com_wrapper&view=wrapper&Itemid=2005&lang=en
- Eastern Mediterranean: http://www.emro.who.int/countries.html
- Europe: http://www.euro.who.int/en/countries
- Southeast Asia: http://www.searo.who.int/countries/en/
- Western Pacific: http://www.wpro.who.int/countries/en/

Appendix:

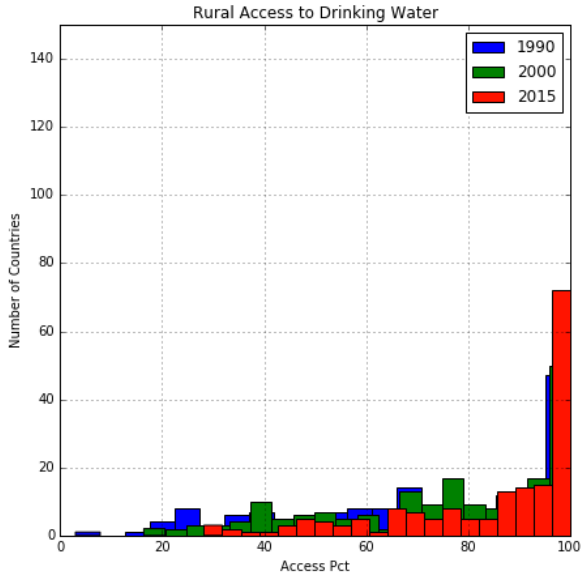
Columns from each data source can be found within this document.

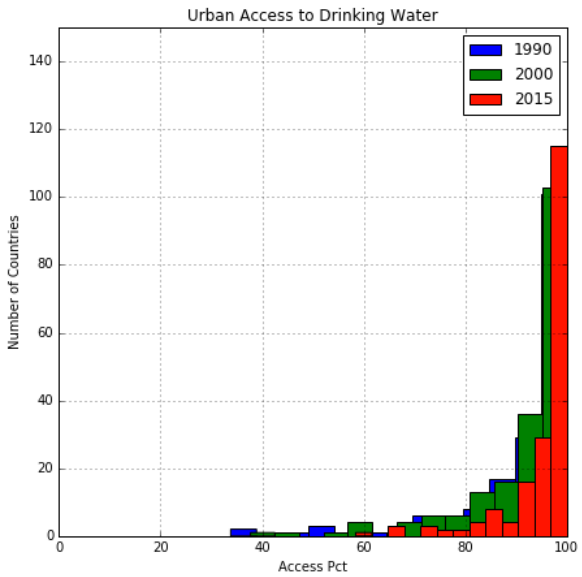

Appendix.docx

FROM ORIGINAL WRITE-UP

Environmental Factors (Hygiene):

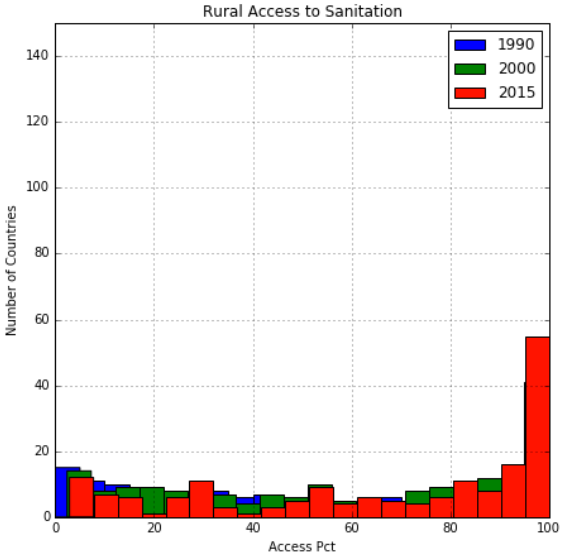
Overview: Access to improved drinking water and improved sanitation facilities has data for 192 countries and 3 sample years (1990,2000,2015) for Rural, Urban and Total populations as a percentage.
Source: http://apps.who.int/gho/data/node.main.167?lang=en
Total records: 558 ***Total Columns:*** 7
Planned Analysis: For analysis in hygiene and sanitary conditions, we will check if life expectancy of the total population has any significant correlation with sanitary access for the different populations. To do this, we will look at Country, Population (Rural, Urban, Total) and data from 1990, 2000 and 2015 (as it matches up to the Life Expectancy in the years 1990, 2000 and 2013). Our initial hypothesis is that countries with higher percentage of access will have higher life expectancy. Within the file itself, we will compare rural vs urban improvement over time.
Initial Analysis
The initial data set was fairly clean, and only required splitting up the Country;Year column into separate columns. We began our analysis looking at how rapid the changes were for increases in access to drinking water for both the rural and urban populations across all countries. Increases in rural access to drinking water showed significant improvement from 1990 to 2015, moving from having some countries to less than 30% access in 1990 to all countries being better than that in 2015.

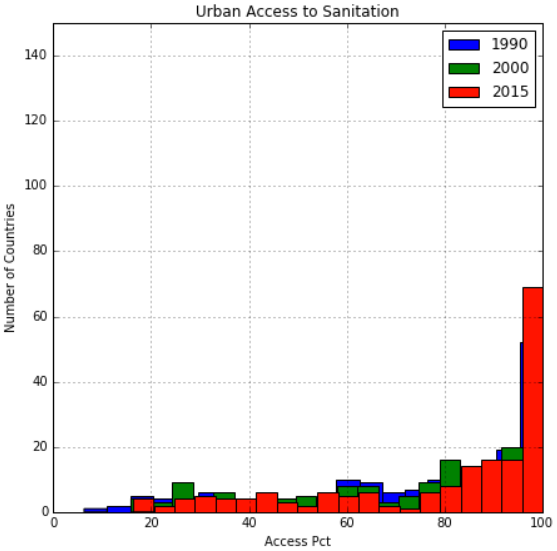




Urban access also improved, but mostly at the high end, moving from many countries with between 80 and 90% access in 1990 to most countries going to greater than 95% access in 2015.

Sanitation access did not show as much movement as drinking water access.





All of the countries with the lowest access to clean drinking water are in Africa:

	IWRuralPct	IWUrbanPct	IWPct	ISRuralPct	ISUrbanPct	ISPct
Country						
Chad	40.900000	60.266667	45.100000	5.400000	26.000000	9.866667
Congo	36.050000	95.466667	72.850000	5.600000	18.850000	13.850000
Democratic Republic of the Congo	27.566667	84.133333	47.700000	20.666667	29.333333	23.533333
Ethiopia	23.500000	87.966667	33.133333	11.433333	23.500000	13.066667
Madagascar	25.266667	76.166667	39.666667	8.100000	16.566667	10.566667
Mozambique	29.033333	75.866667	42.433333	5.666667	37.900000	14.966667
Niger	38.133333	79.700000	45.200000	2.800000	27.866667	7.133333
Papua New Guinea	28.066667	87.600000	36.300000	13.066667	59.300000	19.433333
Sierra Leone	33.133333	76.800000	48.833333	5.633333	22.066667	11.566667
Somalia	16.400000	37.700000	23.500000	10.300000	45.000000	21.800000
Togo	39.200000	84.933333	54.733333	5.033333	23.966667	11.533333