CS 171 - Visualization

Adolescent AIDS

Prevalence and Death Burden

Shannon Milroy, Baishi Wu, Sabina Prajapati

5/5/2015

Introduction

HIV/AIDS is the second leading cause of death amongst adolescents (people 10-19 yrs old). Adolescence is a critical age to intervene in the fight against HIV/AIDS. Proper intervention at these ages can prevent future infection and improve life outcomes for those already living with HIV.

Overview and Motivation

Adolescents and young people represent an increasing share of the people living with HIV/AIDS today. Risky behaviours in adolescence and young adulthood such as unprotected sex and experimentation with drugs leave these young people vulnerable to HIV infection. Furthermore, adolescents who were perinatally infected (contracted HIV from their mothers) have the poorest rates of adherence to antiretroviral medication. In 2014, the WHO ranked HIV/AIDS as the second leading cause of death amongst adolescents. Despite concerns over the increasing numbers of adolescents living with HIV/AIDS very little data is collected on this group.

Shannon is the founder of a non-profit organization called YBank which aims to increase investment in adolescent health. A project being developed at YBank involves improving adherence to HIV/AIDS medication amongst HIV/AIDS positive youth. Inspired by this research, we decided to focus our final project on exploring how HIV/AIDS has evolved over time in the adolescent population.

Related Work

The Institute for Health Metrics and Evaluation has created multiple visualizations for the data that they collect. The Millenium Development Goal (MDG) visualization shows the progress made by multiple countries on MDG #6 which relates to the spread of HIV/AIDS and universal access to antiretroviral therapy. While this shows progress made at an aggregate level this visualization does not allow for an adolescent specific view. The UN has written a lot about adolescent HIV/AIDS and collects data but does not provide comprehensive visualizations to display the data. Given the recent interest in adolescent HIV/AIDS expressed by the creation of the All in HIV campaign we felt this was an excellent opportunity to contribute to this area by creating a visualization tool of the available adolescent HIV/AIDS data.

Data

Our data came from two main data sources: The Global Burden of Disease Study and the United Nations. The HIV/AIDS mortality, HIV/AIDS prevalence and all-cause mortality data was from the Global Burden of Disease Study. The total population by country over time was collected from the United Nations. The data can be retrieved from the following links:

- **HIV Mortality and Prevalence by sex and age**
- **All-cause mortality**

- **Population Data**

Tasks

With this visualization we wanted to create a tool for public health professionals to understand the historical development and current state of adolescent HIV/AIDS across the world. We wanted our visualization to accomplish the following tasks:

1. Display the geographic distribution of the burden adolescent HIV/AIDS
2. Allow for trend analysis of adolescent HIV/AIDS cases and mortality
3. Show country level data on adolescent HIV/AIDS relative to overall HIV/AIDS burden
4. Highlight gender based disease burden between sexes
5. Show trends over time by continent

Users

The target audience is public health professionals who want to understand the context of adolescent HIV globally. This project provides a tool to understand how HIV/AIDS has evolved since the beginning of the epidemic in each country as well as allows for comparison of age groups and sex disparities. As the discussion around the importance of services for adolescents grows this tool can provide a way to contextualize the current state of HIV/AIDS prevalence and mortality and highlight areas of focus regionally and within specific countries.

Process

The process of creating our visualization involved research, data collection, data cleaning, and exploratory analysis. Within this process, each of us separately sought out information that would be valuable for understanding the HIV/AIDS amongst the adolescent population and analyzed the best way in which to communicate this information.

Data Collection and Cleaning

The main HIV/AIDS data was retrieved from the Institute for Health Metrics website from the Global Burden of Disease study. The data was very clean so it required very little processing beyond manipulating it into a csv file format that could be easily parsed by D3. The initial data exploration revealed that the prevalence data was only available every 5 years whereas the mortality data was available every year so this informed our user control implementation. We also collected data on the total population by sex, age and country from the UN and all-cause mortality data from the Global Burden of Disease study. Adding in the additional data required converting all measures to equivalent count units, cleaning the country names to match across dataset and paring down the list of countries to create identical lists.

Exploratory Data Analysis

Initially we used the total population data and all-cause mortality data for each country to understand the prevalence rates and percentage of deaths due to HIV. While these were valuable we found in the course of our visualization development that there was a lot of information already included in the Global Burden of Disease data on HIV/AIDS. Since HIV/AIDS prevalence rates are frequently published we thought it would be more interesting to explore the detailed differences within the HIV positive population. As a result, we decided to focus on comprehensively displaying the HIV specific data and did not incorporate the total population or all-cause mortality data into the visualization.

Once we decided to focus on the Global Burden of Disease study HIV/AIDS data alone, we explored the data by taking cross-sectional views at different years for adolescents and seeing the variation geographically. It was evident that there is a wide range of prevalence and mortality values across the world, with only a few countries contributing to the majority of adolescent cases globally. We also explored the distribution of cases across age groups and sex. The sex disparities were striking so we decided that it would be important to highlight this in the visualization. Trends over time also showed increasing numbers of adolescents living and dying from HIV/AIDS across the globe.

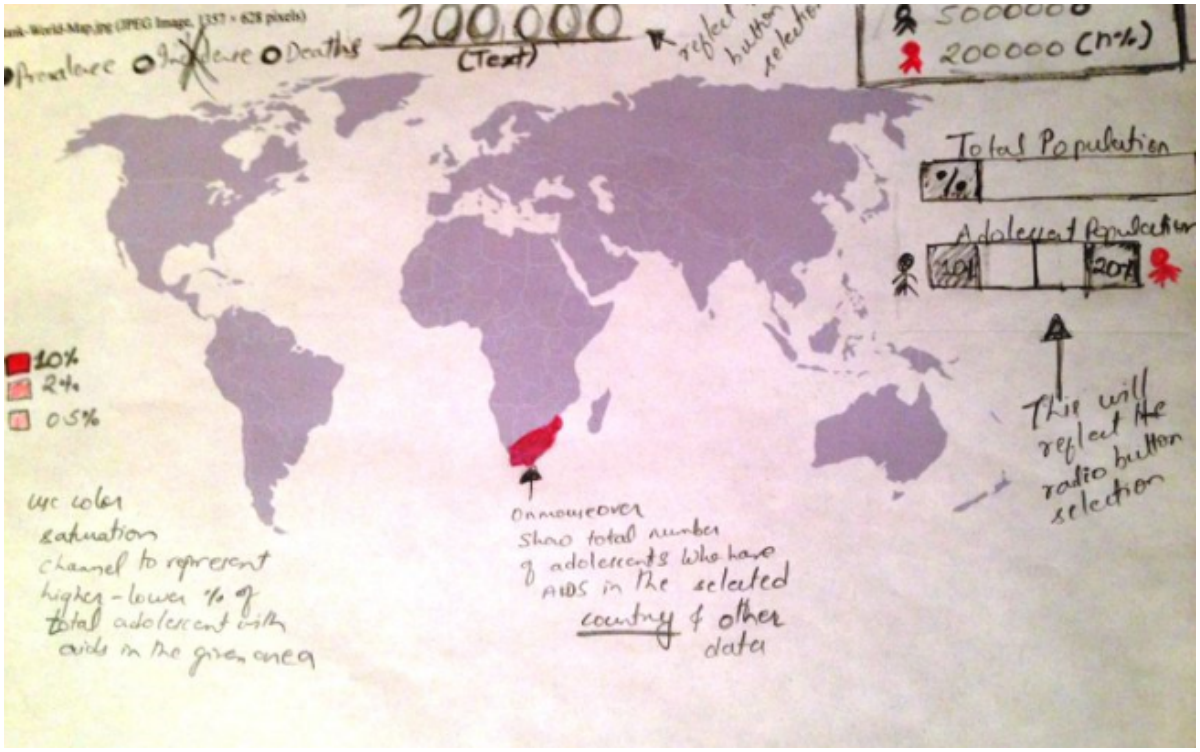
Design Evolution

After collecting a sufficient amount of information and analyzing the best way in which to communicate this, we began to sketch and then implement multiple ways to communicate this information across different dimensions.

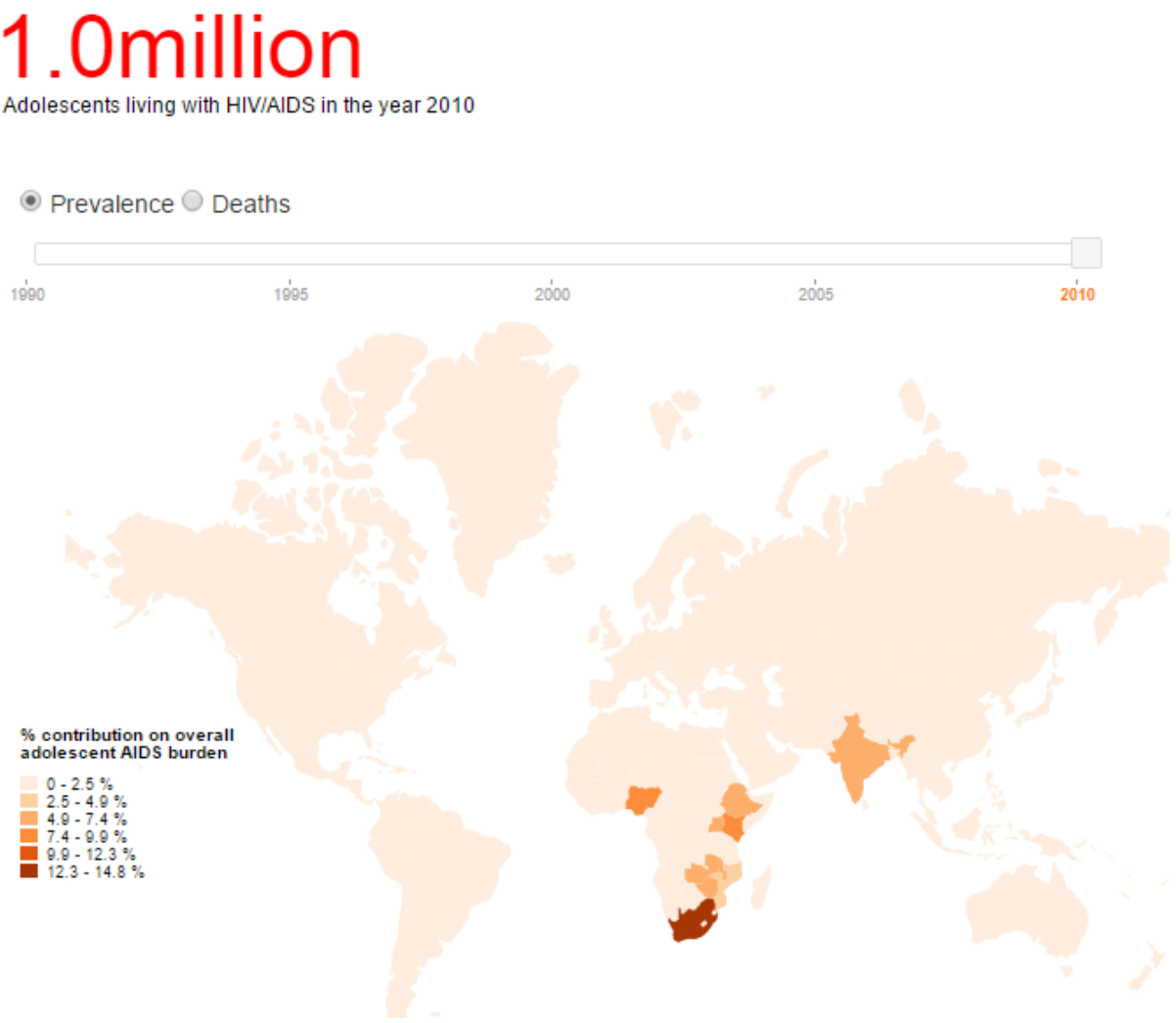
Map

Based on our exploratory data analysis, we knew that we wanted to show data on a world map, covering 20 years from 1990 to 2010. The original map had radio buttons for prevalence, incidence and deaths but for simplicity we decided to eliminate the incidence data. The incidence data was less frequently recorded for each country and for fewer years so including it reduced the quality of the visual. We also included bars above the graph showing the adolescent population relative to the total population but decided to replace them with the population distribution chart.

Sketch Phase:



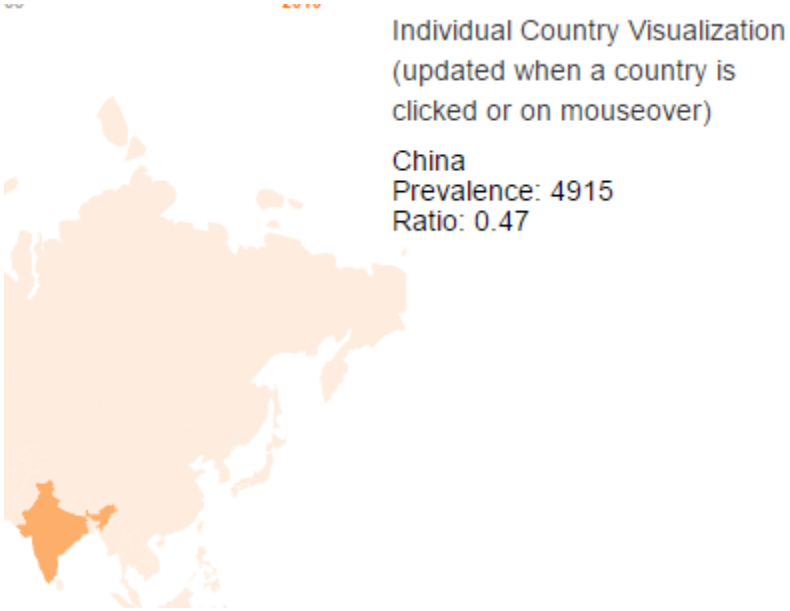
Final Design:



The map shows the percentage of burden (defined as either the total number of adolescents living with HIV or the total number of adolescent deaths from HIV in a selected year) each country contributes towards the total global burden. From the map we can clearly see that the burden is concentrated in Africa and India. We debated whether we should show the percentage of HIV/AIDS in an individual country instead of comparing it to the world total but we felt this makes a more compelling story because it shows that HIV/AIDS is very much concentrated in Africa and India only, with little burden in the rest of the world. Though this approach makes the majority of the map pretty colorless we think that this helps to emphasize that the burden is very concentrated.

The legend on the left hand side of the map shows the percentage of global cases that each color represents.

We had a side panel on the right that updated when a user clicked on a country to display important information about the selected country, such as the name and total number of cases. However, we removed it and added a tooltip instead to allow for more space for the other charts and to make the data easier to read.



We tested our map with couple of projection options but we decided to use Mercator because it gave the best view for the kind of visualization we wanted to show. Since the purpose of our map was to display disease distribution across space and within regions the Mercator was the best option.

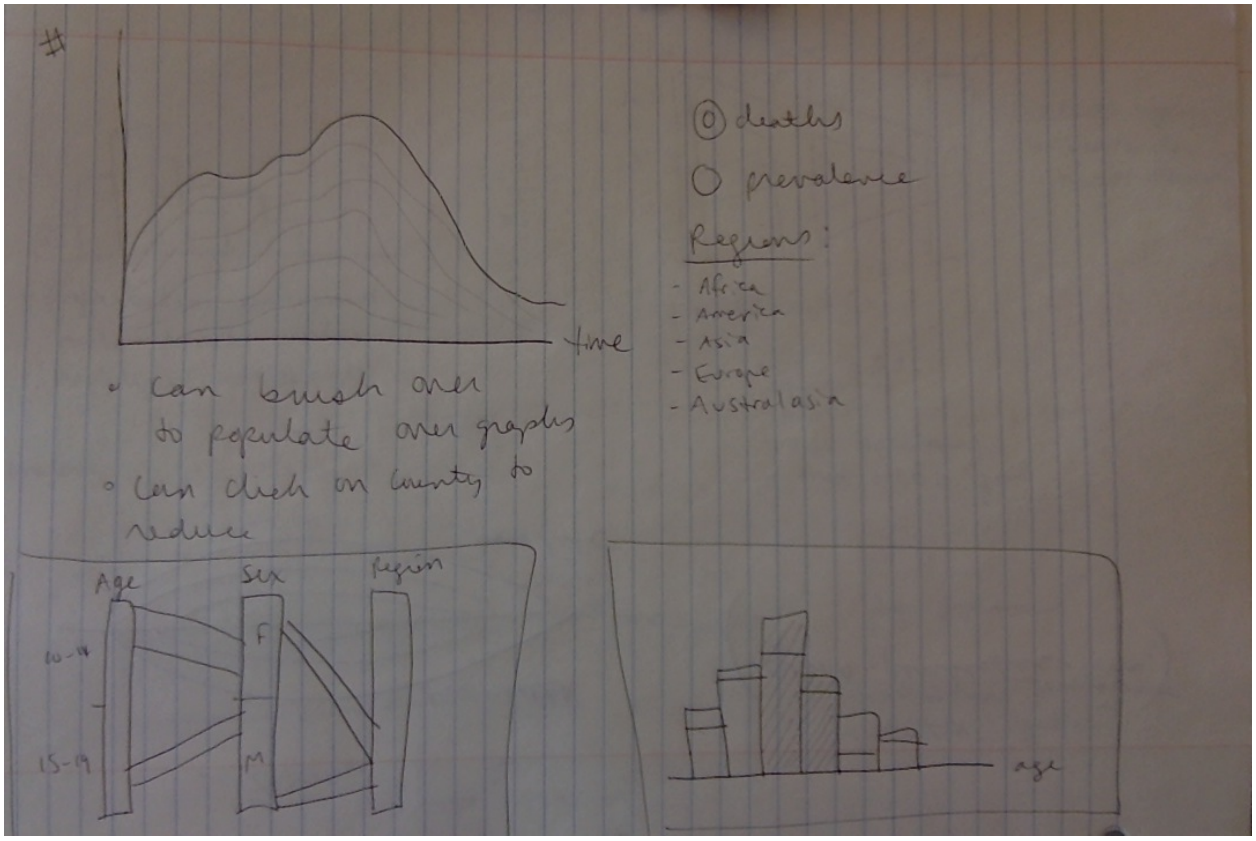
The map colours use Colorbrewer's six category orange scale. We chose this scale because the values of high burden countries are very obvious and it is close to the red color scheme that is usually associated with HIV/AIDS without being too literal.

Population Distribution

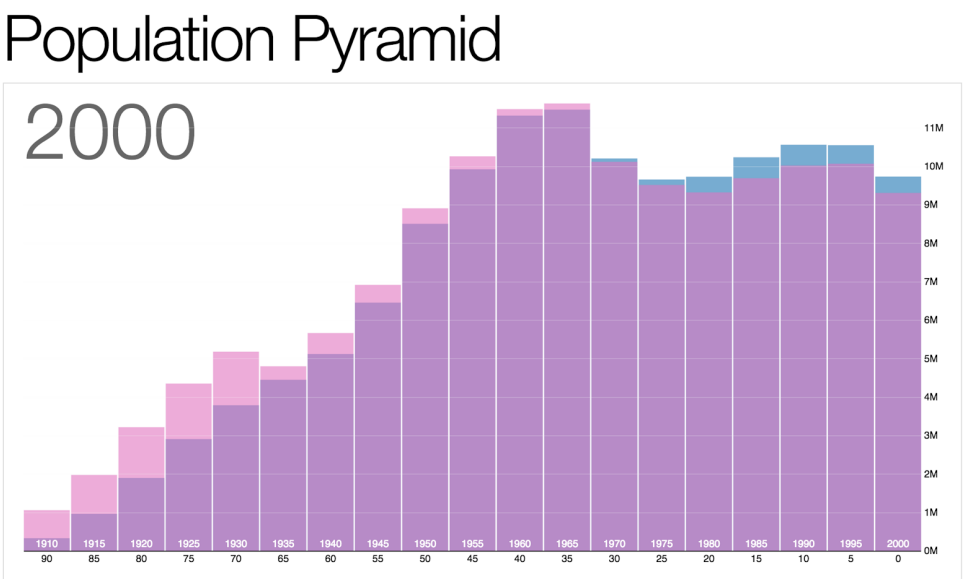
To support the map visualization we wanted to add some charts which gave detailed country data. Based on our data analysis we wanted to highlight the disparities of burden between the sexes as well as put some context around the size of the HIV positive adolescent population relative to the total HIV positive population. One of the reasons that adolescent HIV/AIDS may not have received much focused attention in past years is that the population is relatively small.

To display the disparity in cases between male and female we originally had thought to use a parallel sets chart, breaking down cases by continent, sex and the two adolescent health age groups (10-14 and 15-19).

Sketch Phase:



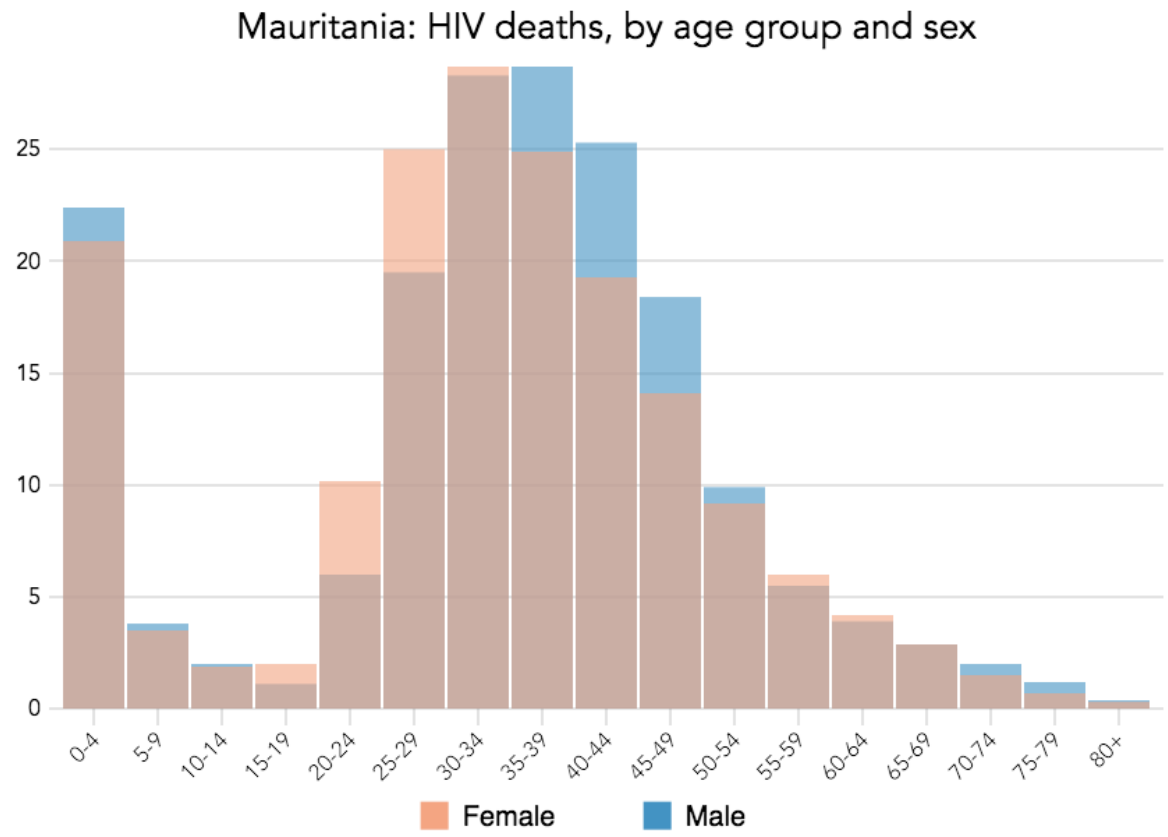
However, we decided that this did not add enough new information to the visualization so we decided that the adolescent health burden should be directly compared to the burden at other ages. This led us to the use of an overlapping bar chart so that the disparities between males and female at each age group could be easily seen and the adolescent health cases would have some context. As inspiration we used the [population pyramid chart] developed by Mike Bostock.



This is an elegant chart as it allows for very quick assessments of disease disparity within an age group and across age groups while showing the magnitude of the population in each group.

After the overlapping bar chart was created we added interaction effects so that the bars would update with the selection of prevalence or deaths as well as the year.

Final Design:

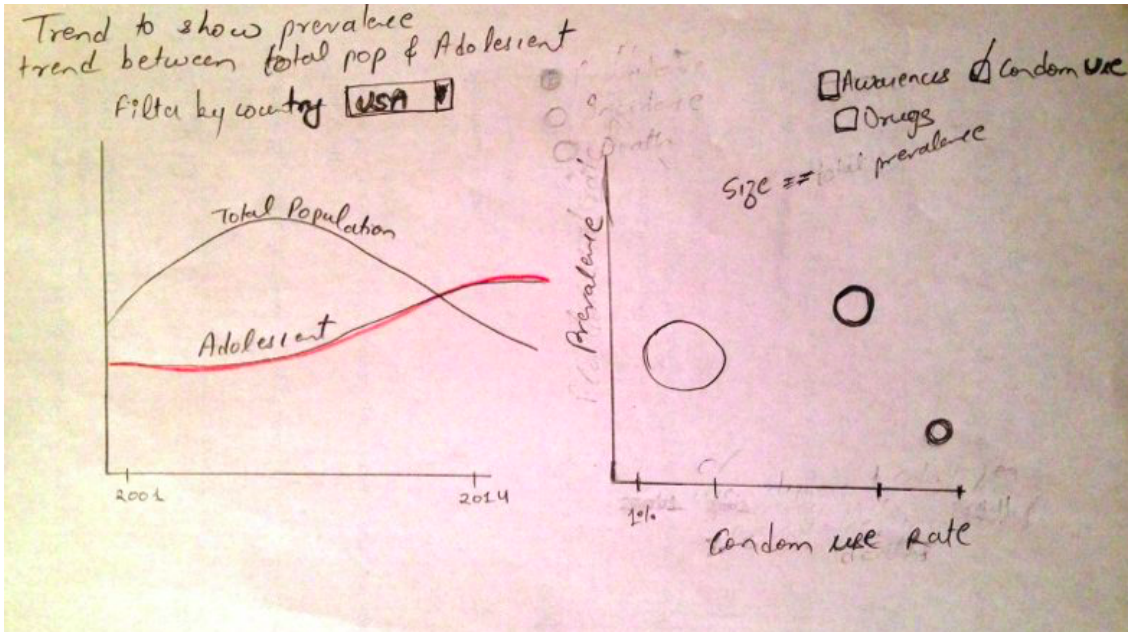


The addition of this chart also led us to add a drop down menu with the list of country names. Originally we had conceived that a new country could be selected by clicking on it. However, we realised that some countries were so small on our map that it was nearly impossible to select them. As a result, we added the drop down menu so that the bar chart and trend chart could be generated for any country. The drop down menu also led to the idea to highlight the selected country so that users could understand where the country was geographically located to get a better sense of regional trends.

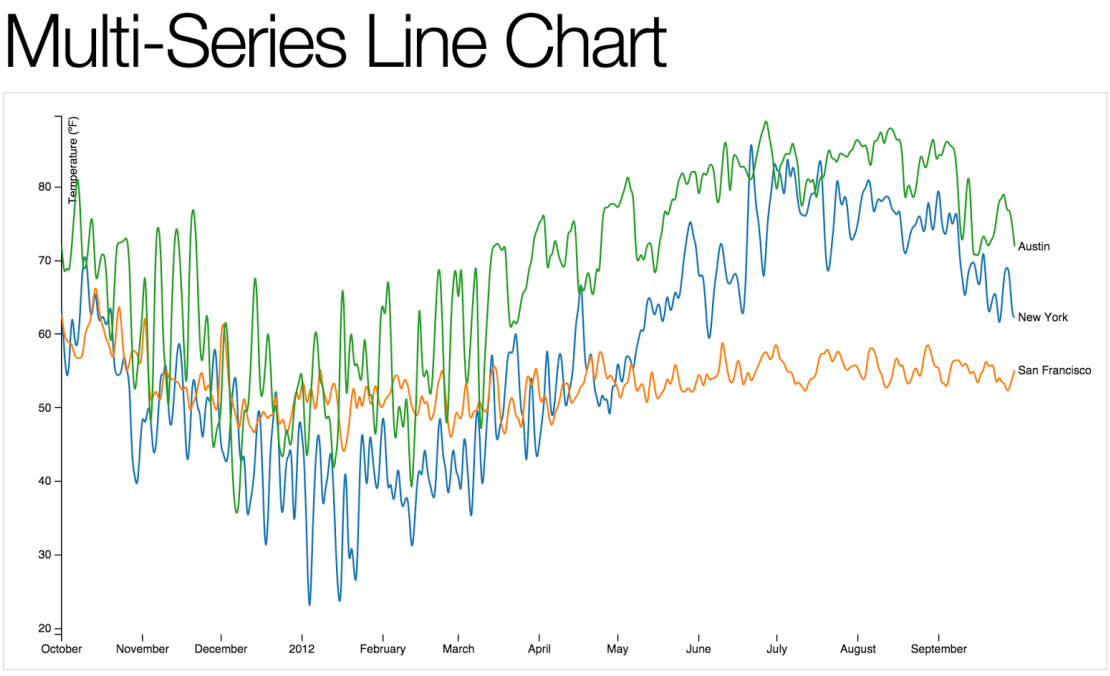
Trend Chart

We had originally conceived of a simple way to show trends over time between the total population and the adolescent population. This was to be controlled through a dropdown menu with adolescent population highlighted separately. The initial sketch is shown below:

Sketch Phase:



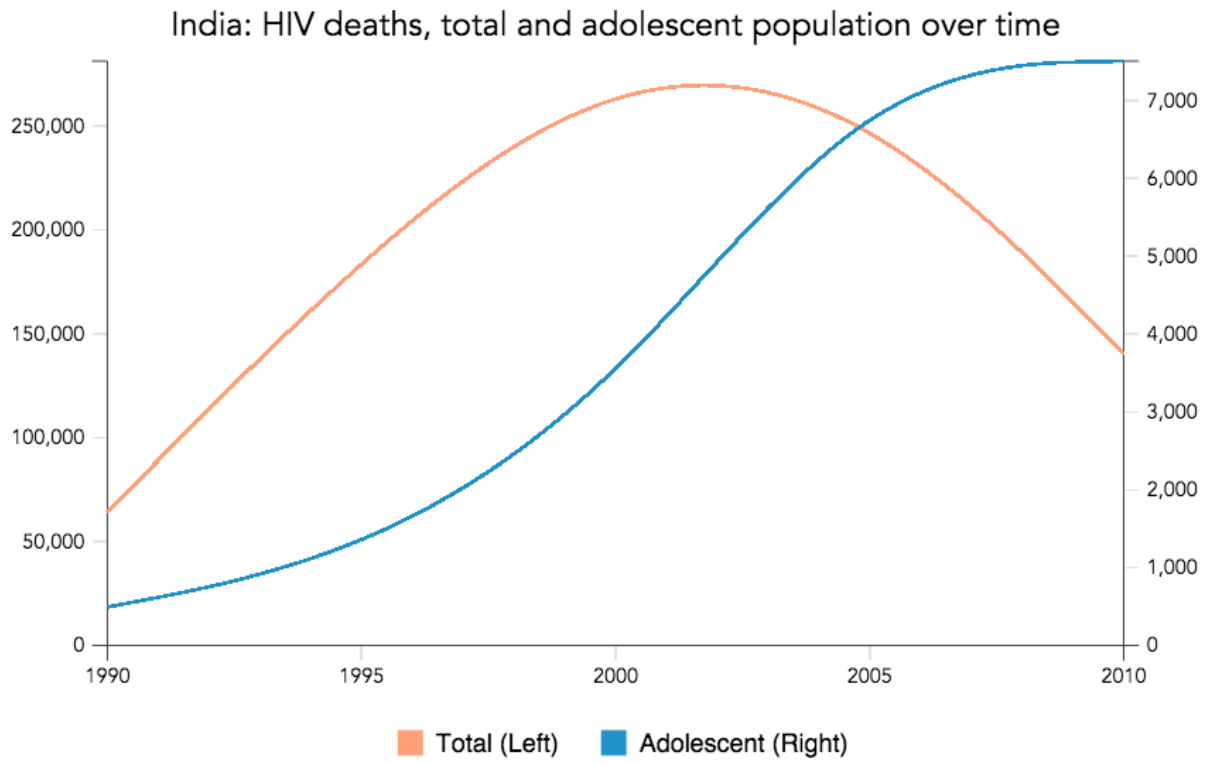
In addition to this concept, we were seeking another way to visualize trends in a manner that highlighted the most important areas. We were initially inspired by a standard, [multi-series line chart](#) from Mike Bostock. The thought would be to visualize trends across the different continents but allow a further level of interaction after clicking-through continent trends to find top areas of prevalence.



However, we realized that this type of visualization focused on the different continents would be insufficient in a few ways. First, while it would be beneficial to display trends based on the different continents, it was difficult to put this in-sync with the rest of the visualization. For example, we did

not think it was intuitive to navigate in too many different ways and felt there would be a disconnect between navigation through the map, the drop-down menu, and the side-charts. Second, we felt there was some inconsistency in displaying information from six different continents but then clicking through to find a number of different countries within those continents, both in terms of how the navigation would be structured and what colors would be used to convey what information. Finally, we thought this idea lacked the focus on the adolescent population that we believed was important. It was with this that we decided to stick with our original idea where we would compare overall death and prevalence from HIV with that of the adolescent population. To do so, we needed to display information on two different axes given the vast difference in scale that existed in the data. This allowed us to implement the following:

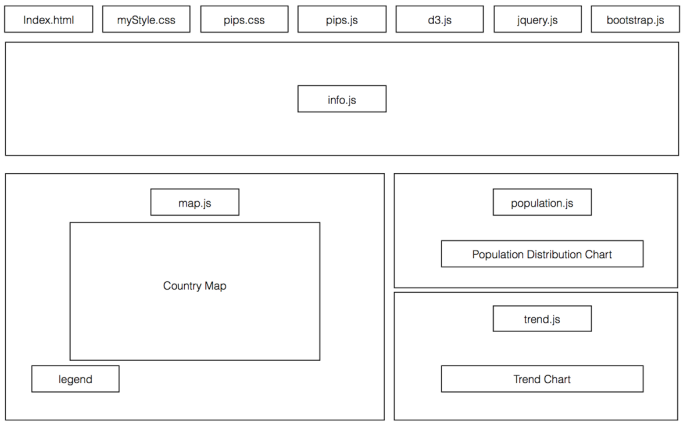
Final Design:



In the above visualization, we find it valuable to compare trends across populations but with vastly different scales. It was important for us to highlight the epidemic that is affecting the youth population and how it was changing over time. In certain countries like the Indian example above, we noticed that while overall deaths from HIV were decreasing, this was not true in all populations - and adolescent HIV deaths have actually been increasing in recent years.

Implementation

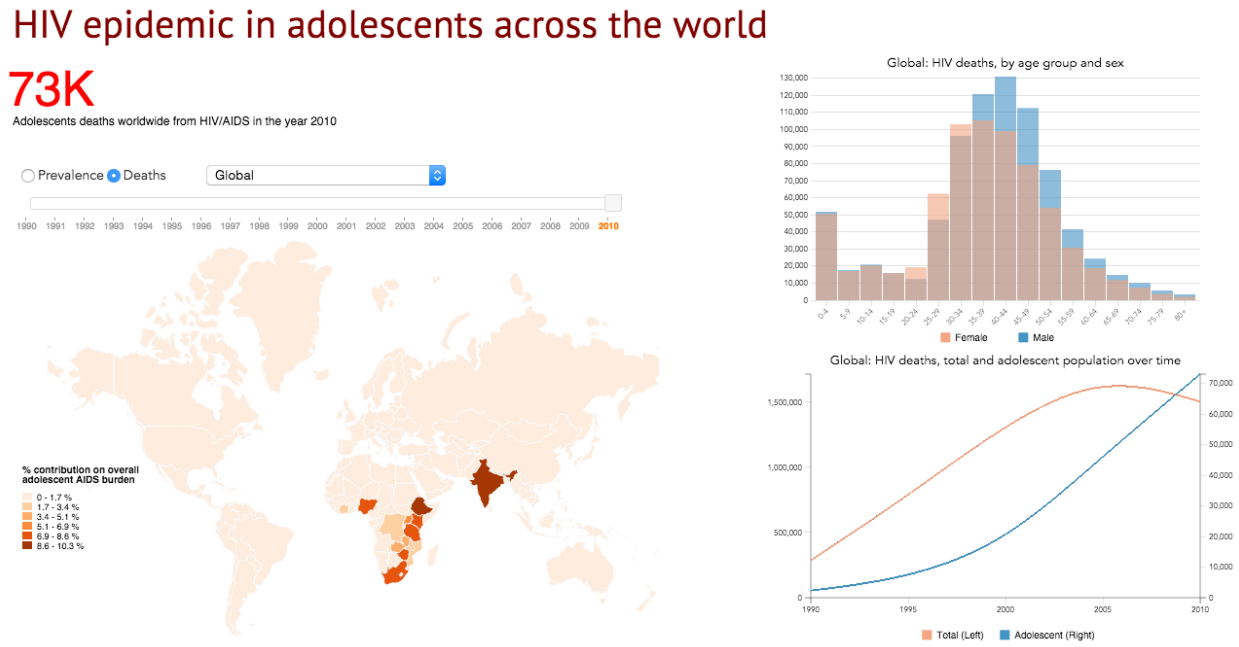
The following diagram shows all the files and visual components of the project:



The final visualization is a one screen display with four main components. The top title shows the count of either the number of adolescents living with HIV or the number of adolescent deaths due to HIV in the selected year (the HIV/AIDS metrics). We wanted the number to pop so that when users come to the page, they know the total burden right away. Since we had data collected for prevalence and death metrics, we provided radio button controls so that the users can easily filter the metric they wanted to see the data for. The time slider allows users to choose the year for which they want to explore the data. The time slider was implemented using the pips slider library. Since we didn't have yearly data for prevalence metric, the slider also gets reset for the step based on which metric is selected. If prevalence is selected the slider shows 5 step slider else it shows 1 step slider. Every interaction on the different maps are synchronized based on the selection of the year and metric, so when user changes either the year or metric, all the graphs and information on the page get updated to reflect the data belonging to the selection.

The main component of the visualization is the map which displays the percentage that each country contributes to the selected HIV/AIDS metric. The country drop down menu allows the user to select a country which is then filled in blue and the country specific data is generated on the two charts to the right of the map. Within the map the user can also scroll over countries to see their name and the relevant metric. The user can also click on the country to highlight it in blue and generate it's data in the side charts. The first side chart shows the distribution of HIV/AIDS metric by age group and sex for the country in the selected year. The user can change the metric of interest or update the year by using the same controls as used for the map. The third chart shows a trend over time for the

adolescent HIV/AIDS metric versus the same metric for the entire HIV positive population. Again this data can be updated by using the controls above the map.



Evaluation

Overall, we are very pleased with our implementation. It presents a comprehensive view of the state of adolescent HIV through 2010 and the evolution since the beginning of the HIV epidemic in 1990. This contributes to the field of adolescent health by making HIV data on this age group easily accessible. This can be used as a tool to explore areas where further research and intervention should be done to help reduce the burden of HIV. One of the most powerful insights from the visualization is that adolescents is an age group in nearly every country where HIV burden is at its lowest. This indicates that this age group is essential for decreasing the spread of HIV. Sexual and reproductive health services that are targeted at this age group are essential to reduce the amount of sexual transmission. Additionally, interventions to increase adherence in this group can reduce the death rate before it reaches higher age groups.

Many countries, such as Brunei, have predominantly male HIV cases. Though many countries have more cases amongst the female population no country has predominantly female HIV cases. By comparing the adolescent HIV cases to the rest of the population it was also evident that adolescents make up a small proportion of those living with HIV. This is due to the small number of children who are perinatally infected. However, when we looked at the number of adolescent HIV cases over time it is evident that in most places they are increasing despite overall HIV cases decreasing. Furthermore, the distribution of cases by age showed that the number of HIV cases in every country spikes significantly in late adolescence and early adulthood once adolescents become sexually active.

Map visualization: Aside from providing a comprehensive view of the burden of HIV cases around the world, the map visualization also acts as an excellent form of navigation and quickly highlights the cases where the incidence is highest.

Population chart: In our implementation this chart allows the user to see the effect of high infant mortality in those who are infected perinatally as well the surge of cases that begins in late adolescence as people start becoming sexually active.

Trend chart: This chart allows a user to compare the trends of adolescent HIV cases over time with trends in the overall population. Similar to comparing cases amongst males and females, this trend was meant to highlight areas where the adolescent trends differed relative to overall population trends, which is the case globally as of 2010.