Statistics Project– HIV/AIDS

*Please use a computer for this project. Please paste everything into the same document and answer the prompts in order in the document. If you need to leave space for showing your work, just leave some extra space and write it in after you print it out.*

**Part 1**

Read “The State of the Epidemic” from the 2012 Global Aids Report, which is published by UNAIDS every 2 years. In the first paragraph of your project, summarize what you learned from the report, what questions you have, and what surprised you.

**Part 2**

Look through the Excel spreadsheet titled “HIV Prevalence” which lists data by country. It shows HIV prevalence rates for different countries. HIV prevalence is defined as the proportion of the population that is living with HIV or AIDS. Find at least 5 countries that you would like to focus on. Choose at least one from each category on page 11 of “State of the Epidemic” – one from the Increasing >25%, one from Stable, one from Decreasing 26-49%, one from Decreasing >50%, and at least one more.

For each country, do this in order:

* Make an individual scatterplot for each of the 5 countries you select.
* Determine a function that best fits each country’s data. Explain how you know that the function is the best fit and what other functions you considered.
* Display the function and the correlation coefficient (if available, it is not available for each type of regression) on each scatterplot. Remember that Excel does not provide all function family options for regressions. If you find a better matching function family, print out the scatter plot, then hand draw the function over the scatterplot.
* Make a prediction for HIV prevalence rate in the year 2018 for all 5 countries. Show your work. Do you think your equation will be a good predictor? Why or why not?
* Make a prediction for that statistic for each country in 2050. Do you think your equation will be a good predictor? Why or why not?

**Part 3**

After looking at individual countries, write a summary paragraph that compares the rates in the different countries. Explain what the functions tell us about each country (speed of growth or decline, what you think might cause the country to be where it is on its curve, etc).

**Part 4**

Now, choose one other question to answer. You can look at specific country data over time, focus on a particular region or subgroup of a population, or you can compare HIV/AIDS statistics to another statistic that you think might correlate with HIV/AIDS (for example, level of education, some other disease prevalence). I have provided you with statistics describing HIV related deaths in each country. You could use this, or perform additional research. Please cite any additional sources you use with a Works Cited page at the end of your assignment. A great place to find data is at [www.gapminder.org/data](http://www.gapminder.org/data). You might need to reformat the data from that website; please let me know if you need help with this.

For the question that you choose, explain why you chose to answer this question. You can create a scatterplot of that statistic over time like you did for the HIV prevalence rate, or you can create a scatterplot of that statistic and the HIV prevalence rate. Answer the question for each of the 5 countries that you selected, or you can focus in on a specific region. Explain what looking at that statistic helped you further understand.

**Part 5**

Write a conclusion with a summary of important findings, as well as any questions that you still have after completing this project.

**If you want to look into different statistics entirely, that is fine, but you must find all of the data yourself.** [**www.gapminder.org/data**](http://www.gapminder.org/data) **is a great place to start. You would need to follow the same basic outline as this project, where you look at a statistic, model it in several different instances then answer an additional question about it. Please see me if you are doing this so I can approve your plan.**