Project Title: HIV Awareness Application

Project design and description

HIV/AIDS is one of the most fatal infectious diseases that has an increasing rate of new infections and death rates since it can not be treated by medical procedures with the current medical procedures. This is one of the main reasons why we choose this topic because by providing enough information for the public we believe that we can raise enough awareness to avoid it in the first place. As we included in our data set, education and funding plays a high role when it comes to both avoiding HIV and increasing the life expectancy and well-being of infected people.

Steps

We researched many topics and landed on HIV because we thought this project could be useful to raise awareness. By making research on Our World in Data we found useful resources about our topic. Within further research on https://ourworldindata.org/hiv-aids we have selected four of the data sets which we believe will be useful to make comparisons between different data on the topic and also visualize the problem with graphs and tables. From all the data searched, we obtained entity sets their relationship sets with all their attributes. By separating data which we found useful from the flat data, we fully obtained our data set clearly without any unnecessary information to avoid any complications in the future, so that we can enlarge our data if ever needed by keeping our data clean and useful. After clearing all the duplicates, we were able build a data set that won't cause problems when we are sorting and looking up for the data in the future. From the 4 datasets, we copied and pasted the columns that we are interested in a new sheet in Excel. In the annual-number-of-deaths-by-cause dataset, we separated entity, code, year, and death number. In HIV-expenditure dataset, we selected entity, code, year, and funding source columns. In young-people-with-knowledge-on-hiv-prevention dataset, we used entity, code, year, and knowledge columns. Lastly in share-of-women-among-the-population-living-with-hiv dataset, we copied entity, code, year, and woman with HIV columns. Furthermore, we formed an additional column which includes isocode + year to be able to use VLOOKUP function. Starting with the classification of our entity sets, we moved on by declaring weak entities and their attributes. Finally, by analyzing our entities we were able to clearly make connections between them that resulted in being our relationship sets.

ER Diagram: