HIV analysis Mini project

Mutwoki Dennis Mureti

95003

Motivation

- It is everyone's best interest to have a healthy child and reduce the chances of them ever having to fall ill.

- The aim of my analysis is to seek the trends that can be used to determine the chances of having new HIV infected child by observing the variable around the growth of this children.

- The problem I seeks to solve is the reduction of newly infected children by controlling the variable that may lead to this like mother to child infections and the poverty rate of the household in that they can afford the prescribed drugs for their children and whether the environment increases their chances of getting HIV.

Dataset(s)

- The dataset that will be of use to me will be the adult HIV prevalence rate by county where I seek to explore more on the children HIV rate and other related matters.

- Sample dataset will be on children, house holds with orphans, poor hose holds with orphans, total population and poe prevention of mother to child transmission.

- I have worked with the first 7 counties to reduce the scope and make it easier in reading the data.

- The mean and the Standard deviation relatively vary in the either can be used in the statistical analysis of the data.

Research Question(s)

- Does the number of poor homes with orphans signify the new HIV infected children while considering the county total population.

- Does the poe prevention of mother to child transmission signify the child risk level of individual counties.

Methodology

- For my regression analysis I used linear regression. This is due to the fact that I was comparing two variable at a times. New HIV infections children 0 14 being of main focus.

- There was sampling of the first 7 counties.

- There was the use of scatter plots.

- Use of bar graphs.

- Use of line graphs.

- All this was done for the relevant datasets to compare the graphs.

Findings
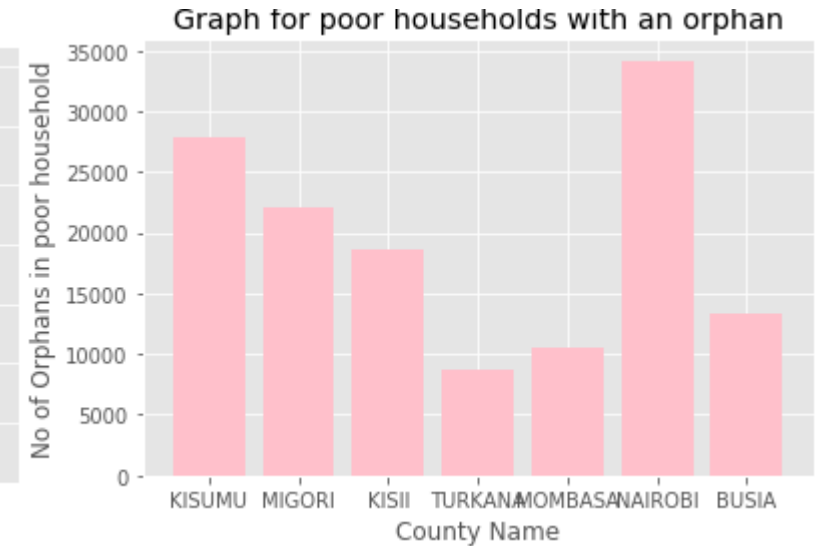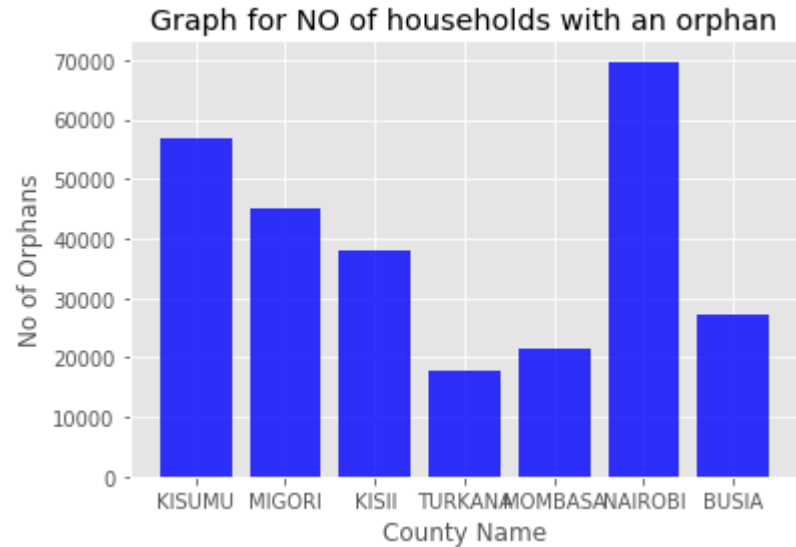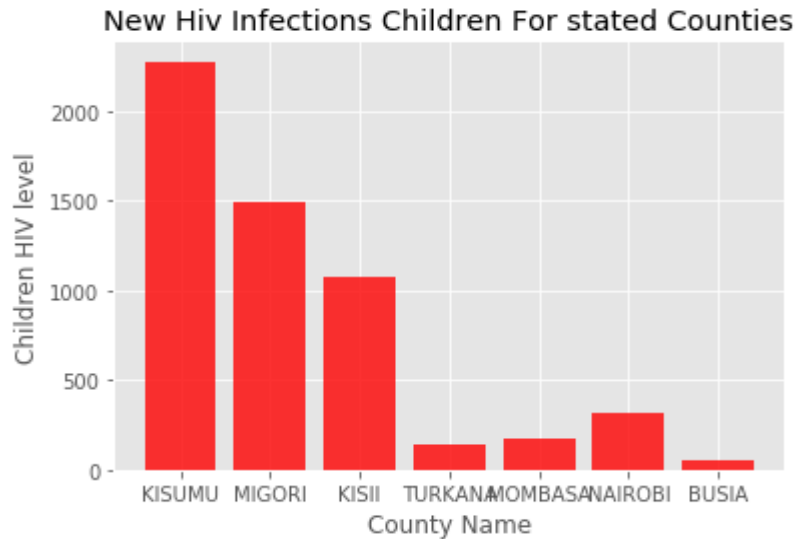
- From the three bar graphs above we try to find out thru the representation f the graphs on whether variable in the three graphs are related though not directly.
- From the observation we can see that for example by looking at the graph we see the level of poor households with orphans in Kisumu we can also see the similarity in the graph hight for the new children with HIV infections for Kisumu as well.
- There is similarity in the length of the bars for the three sample datasets.

Findings
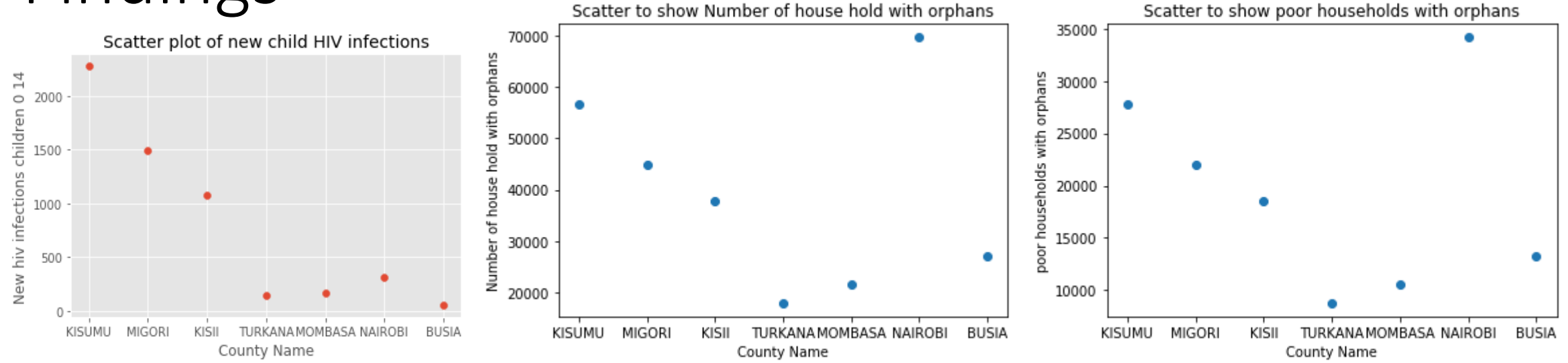
- From the scatter plots above we can identify the similarity in the curves after plotting the data.
- The similarity can be used to conclude that as the number of orphans go high so do the number of a cases for newly infected HIV Children.
- Getting prior information for orphans should give the awareness of cases of HIV infected children and can be acted upon.

Findings

Line Graph showing Children HIV Statistics for seven counties
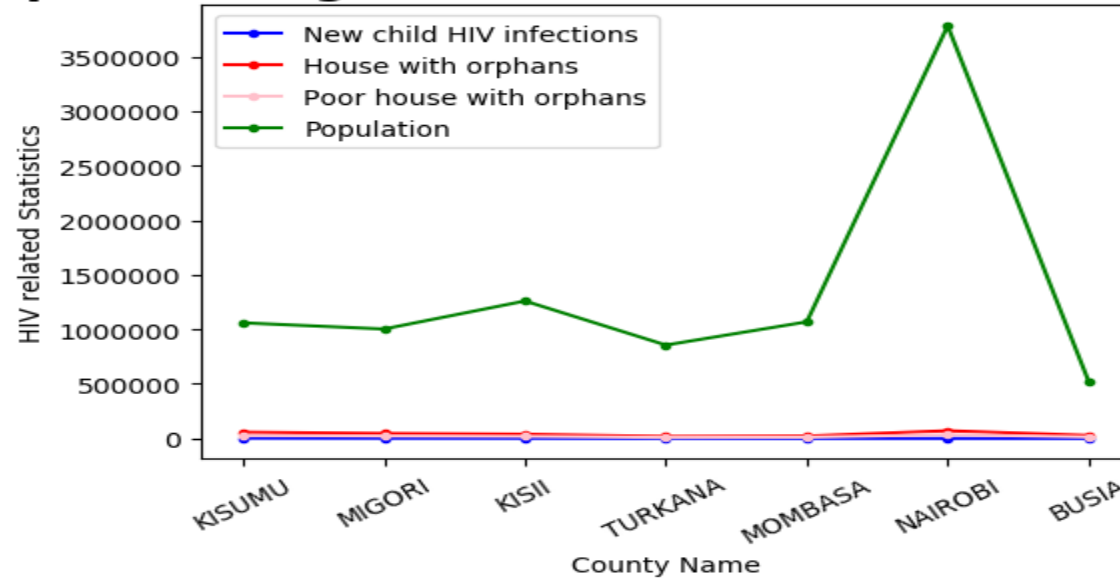
- From the line graph we can evidently see the similarity in the curves for the three datasets i.e. new child HIV infections, house with orphans and poor house with orphans.

- The points of curvature for the three line graphs are almost the same that proving the notion that was earlier mentioned.

Limitations and Challenges

- Lack of continuous consistency in the graphs.
- Dataset provided was rather large for clear analysis and thus the reason for sampling of the first 7 counties.

- Limitation was lack of accurate data due the overpopulation by some of the counties.