Cluster Analysis of the Incidence of Hiv/Aids in Nigeria

W. B. Yahya[1]; M. O. Olaniyi[2]; M. K. Garba[3]; M. A. Raheem[4], R. B. Afolayan[5]; K. Jimoh[6]

[1-3, 5]Department of Statistics,
University of Ilorin. Ilorin, Nigeria.

[4]Material and Engineering Research Institute,
Sheffield Hallam University, Sheffield, United Kingdom.

[6]Department of Physical Sciences,
Al-Hikmah University, Adeta, Ilorin, Nigeria
E-mail: wbyahya@unilorin.edu.ng[1]

***Abstract* — HIV/AIDS is a global crisis with Nigeria currently ranked among the countries with a high number of people living with HIV/AIDS. Undoubtedly, this disease poses a serious threat to Nigeria as a nation. Thus, it becomes imperative to examine the incidence of HIV/AIDS across Nigeria for stakeholders to develop a necessary interventions to arrest it. This work, therefore, investigates the distributions of HIV/AIDS as currently being experienced across the 36 states including the Federal Capital Territory (FCT) and the six geopolitical zones in Nigeria. Results from hierarchical clustering with a single linkage search revealed that Akwa-Ibom, Cross-River, Enugu, Benue states as well as FCT, Abuja have similar experience in terms of incidence of HIV/AIDs in Nigeria. According to the results, the incidence of HIV/AIDS in these states including FCT was more pronounced than in any other state in Nigeria. Further results showed that the incidence of HIV/AIDS is largely dependent on the geopolitical zones in Nigeria ($p < 0.05$) with North East, North Central, and South-South being more vulnerable than other zones. National data set obtained from the National Bureau of Statistics was employed for this study.**

Keywords: *Hierarchical Clustering, HIV/AIDS, Dendrogramm, Hit-map, Boxplot.*

I. Introduction

Records show that HIV/AIDS was first discovered in Los Angeles and New York in 1981. About four years later, the first two cases of HIV/AIDS were detected in Nigeria which was reported at an international AIDS conference in 1986. In a swift reaction, the Nigerian government established National AIDS Advisory Committee through the Federal Ministry of Health. This was closely followed by the establishment of the National Expert Advisory Committee on AIDS (NEACA), the National Action Committee on AIDS (NACA), and a host of other Non-Governmental agencies on HIV/AIDS.

A report from avert.org which has Sub-Sahara African HIV/AIDS statistics (2010) shows that Nigeria has 3.3 million people living with HIV/AIDS, out of which 1.7million are women, 360,000 are children. About 220,000 deaths ensued from AIDS-related diseases with not less than 2.5 million orphans. The foregoing shows that the prevalence of HIV/AIDS in Nigeria is becoming alarming. Hence, the need to examine the incidence of HIV/AIDS across Nigeria is not an overemphatic exercise. Nigeria has a population of over 140 million people with thirty-six states and federal capital territory (FCT).

The nation is divided into six geo-political zones viz; North-Central (*Benue, Kogi, Kwara, Nassarawa, Niger & Plateau*), North-East (*Adamawa, Bauchi, Borno, Gombe, Taraba & Yobe*), North-West (*Kaduna, Kano, Katsina, Kebbi, Jigawa, Sokoto & Zamfara*), South-East (*Abia, Anambra, Ebonyi, Enugu & Imo*), South-South (*Akwa-Ibom, Bayelsa, Cross-River, Delta, Edo & Rivers*) and South-West *(Ekiti, Lagos, Ogun, Ondo, Osun & Oyo*).

This study investigates the distribution of HIV/AIDS as being experienced across Nigeria, to determine those states and geopolitical zones having similar HIV/AIDS incidence, and to identify the critical states and geopolitical zones that require necessary and immediate intervention based on their reported cases of HIV relative to other states and geopolitical zones.

II. Materials and Methods

The dataset used for this work was extracted from the publication of the Nigerian National Bureau of Statistics. It consists of reported cases of HIV/AIDS for the thirty-six states and FCT for a period of five years (2003-2007). A cluster analysis procedure was employed to identify cases that are similar to one another but different from those in another group. This procedure attempts to identify subgroups whose elements are as homogeneous as possible based on the information inherent in the data. Having considered the features of the dataset at hand, the ***agglomerative hierarchical clustering technique*** was used.

One of the main intents of cluster analysis is to form similar groups. Similarities are a set of rules that serve as criteria for grouping or separating items. **Distance** is a measure of how far apart two objects (cases) are while **similarity** is a measure of how similar they are. For cases that are alike, distance measures are small and similarity measures are large. Then we established the similarity or distance matrix and used Squared Euclidean Distance (SED) criterion defined below for clusters x and y:

Squared Euclidean distance (x, y) $=\sum_i (x_i - y_i)^2$

After several objects are linked together, it is necessary to determine the distance between these new clusters. That is, a linkage or amalgamation rule is required to determine when two clusters are sufficiently similar to be linked together. This work utilized the **furthest neighbour** (complete linkage method) which determines the distance between clusters by the method of greatest distance between any two objects in the different clusters. The analysis of data used for this work was performed using partly R statistical package and SPSS.

III. Results And Discussion

The results obtained from the analysis of the extracted data are discussed here. Relevant tables and figures are presented accordingly and are appropriately referred to while discussing our results. For instance, Tables 1 and 2 give the Squared Euclidean Distances (SED) between pairs of States and pairs of Geo-Political Zones. It should be noted that Table 1 is of dimension 37 *by* 37, but the abridged form is produced here due to space.

Table 1: The Squared Euclidean Distance between Pairs of States

Case	Abia	Adamawa	Akwa Ibom	Anambra	. . .	Taraba	Yobe	Zamfara	Abuja
Abia	0	9.224E9	1.768E10	2.538E9	. . .	8.491E9	5.130E7	2.127E9	1.686E10
Adamawa	9.224E9	0	7.103E9	1.563E10	. . .	2.435E9	9.462E9	1.723E10	2.890E9
Akwa Ibom	1.768E10	7.103E9	0	2.213E10	. . .	3.355E9	1.892E10	2.868E10	1.741E9
Anambra	2.538E9	1.563E10	2.213E10	0	. . .	1.556E10	2.351E9	1.100E9	2.229E10
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
Taraba	8.491E9	2.435E9	3.355E9	1.556E10	. . .	0	9.356E9	1.829E10	2.971E9
Yobe	5.130E7	9.462E9	1.892E10	2.351E9	. . .	9.356E9	0	1.692E9	1.763E10
Zamfara	2.127E9	1.723E10	2.868E10	1.100E9	. . .	1.829E10	1.692E9	0	2.731E10
Abuja	1.686E10	2.890E9	1.741E9	2.229E10	. . .	2.971E9	1.763E10	2.731E10	0

Table 2: The Squared Euclidean Distance between Pairs of Geopolitical Zones in Nigeria

Case	SS	SE	SW	NW	NE	NC
SS	0	3.259E11	7.839E11	1.719E11	5.303E10	1.162E11
SE	3.259E11	0	1.205E11	2.892E10	1.253E11	8.030E11
SW	7.839E11	1.205E11	0	2.510E11	4.527E11	1.481E12
NW	1.719E11	2.892E10	2.510E11	0	3.597E10	5.535E11
NE	5.303E10	1.253E11	4.527E11	3.597E10	0	3.201E11
NC	1.162E11	8.030E11	1.481E12	5.535E11	3.201E11	0

The hit map presented in Figure 1 displays the reported cases of HIV/AIDS in the 36 states and FCT. It was discovered that Cross River and Benue states are having high reported cases, while Ekiti, Ondo, Osun, and Jigawa states have relatively low cases reported for the years under review. Meanwhile, the hierarchical tree, also known as dendrogram, presented in Figure 2 displays how clusters are merged at each of the stages of analysis. It can be visualized how states are merged with respect to the cases of HIV/AIDS reported.

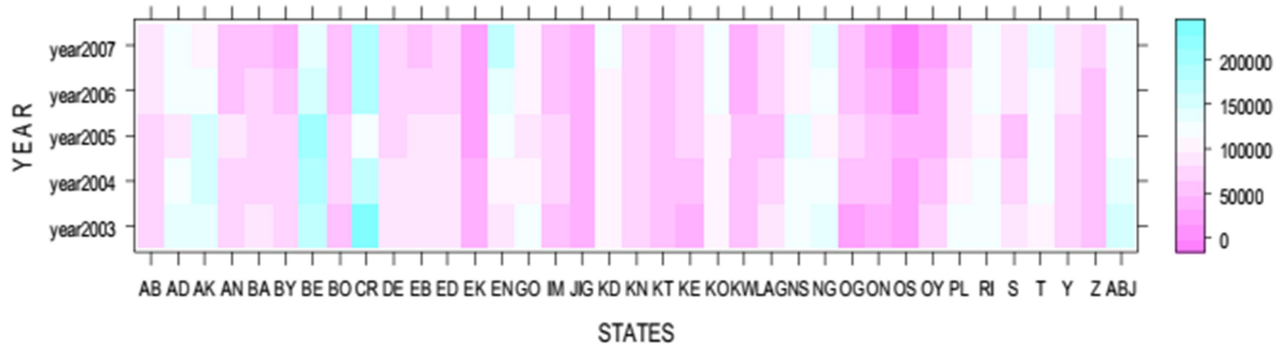

Figure 1: Hit Map of HIV cases in Nigeria by States and Years.

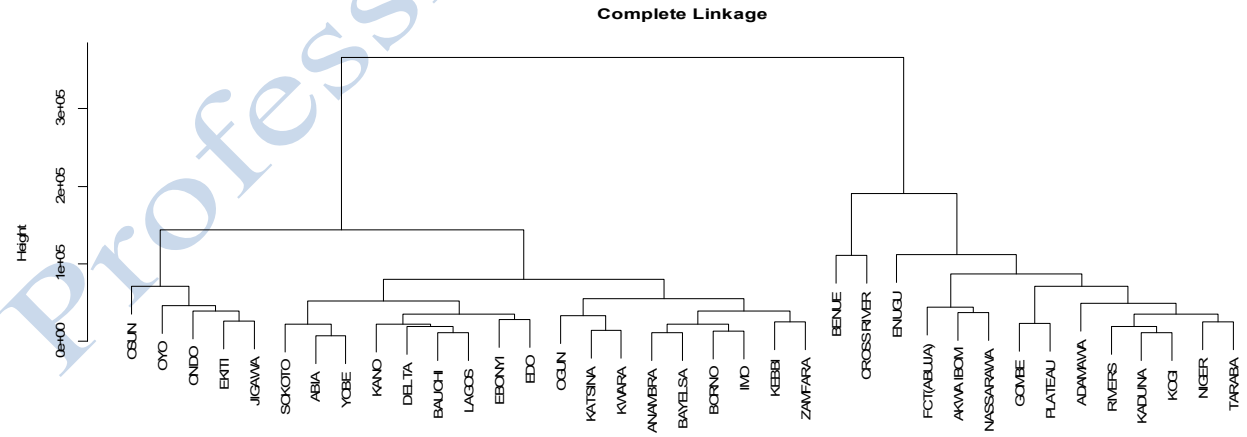

Figure 2: Dendrogram of HIV cases in Nigeria by States

Results from the cluster analysis as shown in the debrogram in Fig 2, it can be found that Osun, Oyo, Ondo, Ekiti and Jigawa states have similar reported cases of HIV/AIDS and are relatively the lowest. Sokoto, Abia, and Yobe states also have similar reported cases. Reported cases in Kano, Delta, Bauchi, Lagos, Ebonyi, and Edo states are not significantly different from one another.

In the same vein, cases of HIV/AIDS in Ogun, Katsina, Kwara, Anambra, Bayelsa, Borno, Imo, Kebbi, and Zamfara states are similar. While FCT (Abuja), Akwa Ibom, Nassarawa, Gombe, Plateau, Adamawa, Rivers, Kaduna, Kogi, Niger, and Taraba states experienced similar cases which are different from those in other groups earlier mentioned. Benue, Cross River, and Enugu states are the same categories, though Benue and Cross River have the highest number of cases of HIV/AIDS in Nigeria for the period been considered.

Boxplot of HIV/AIDS cases for the states and FCT was obtained as presented in Figure 3 to corroborate the aforementioned claims. It was revealed that 13 states (Adamawa, Akwa Ibom, Benue, Cross River, Enugu, Gombe, Kaduna, Kogi, Nassarawa, Niger, Plateau, Rivers, and Taraba) and FCT have reported cases higher than the combined median (i.e. average). 11 states (Borno, Ekiti, Imo, Jigawa, Katsina, Kwara, Ogun, Ondo, Osun, Oyo, and Zamfara) have cases below the median while reported cases in the remaining 12 states (Abia, Anambra, Bauchi, Bayelsa, Delta, Ebonyi, Edo, Kano, Kebbi, Lagos, Sokoto, and Yobe) are about the median of reported cases in Nigeria.

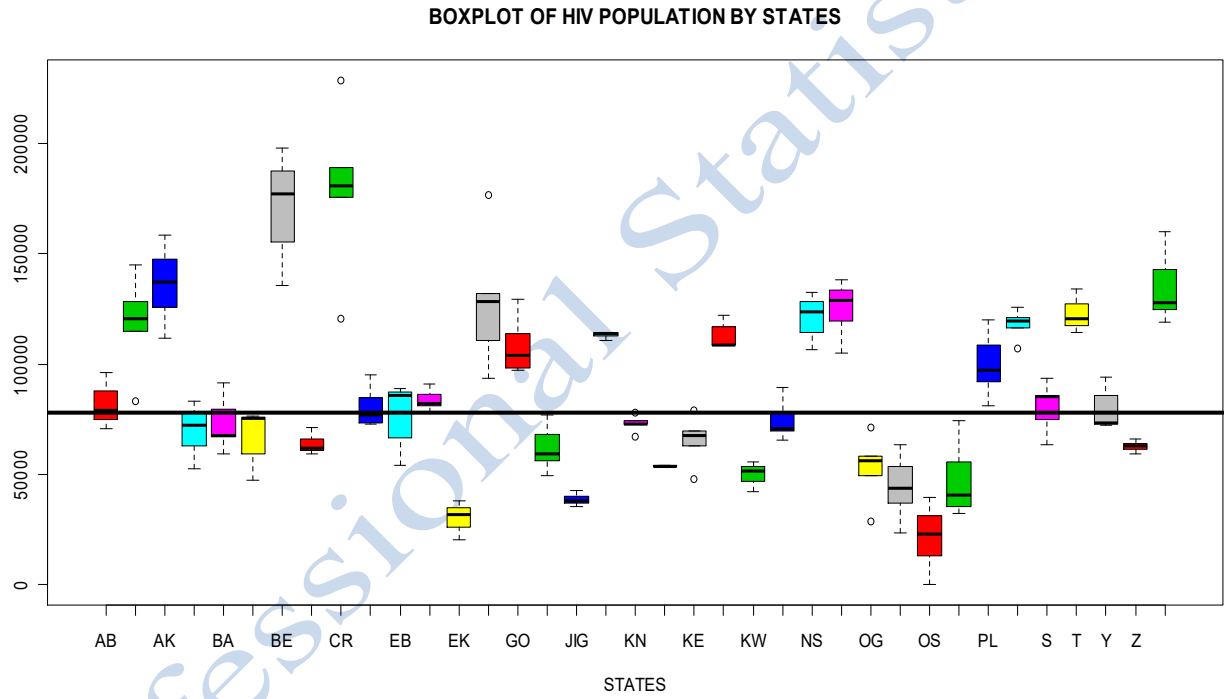

Figure 3: Box-Plot of HIV cases by States.

Studying the incidence of HIV/AIDS in the six geo-political zones of the country, the data were combined according to the zone the states belong. An agglomerative hierarchical cluster analysis was carried out on this new dataset. The resulting SED is presented in Table 2. Figures 4, 5, and 6 are the corresponding Hit Map, Dendrogram and Boxplot respectively.

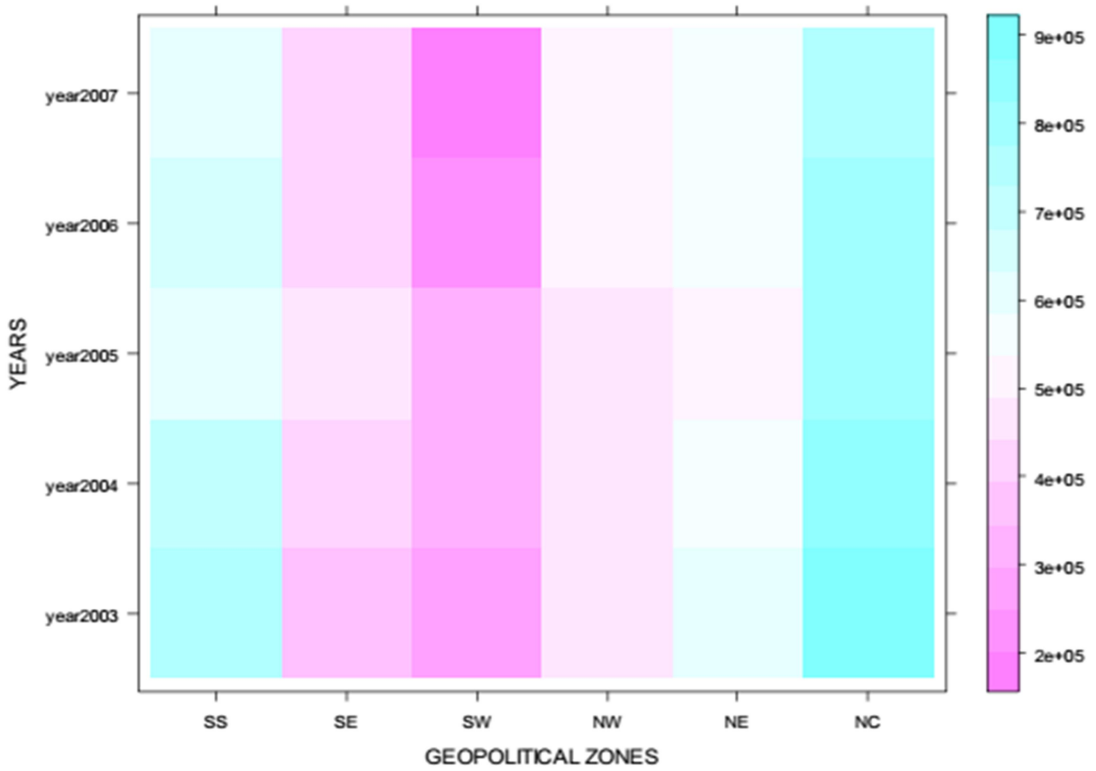

Figure 4: Hit Map of HIV cases by Geopolitical Zones over the years considered

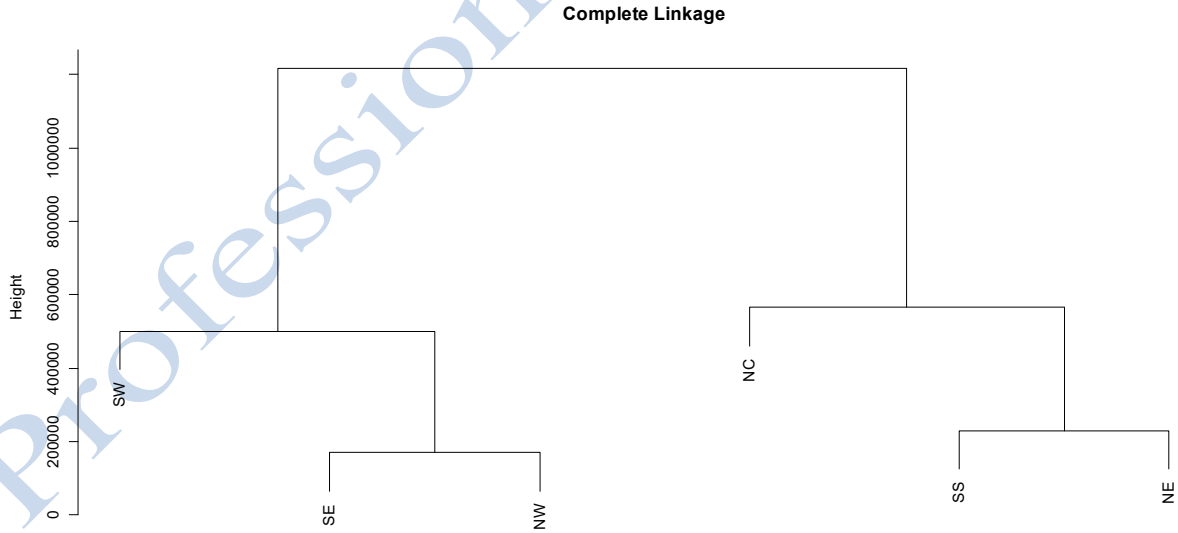

Figure 5: Dendrogram of HIV Cases by Geo-Political Zones

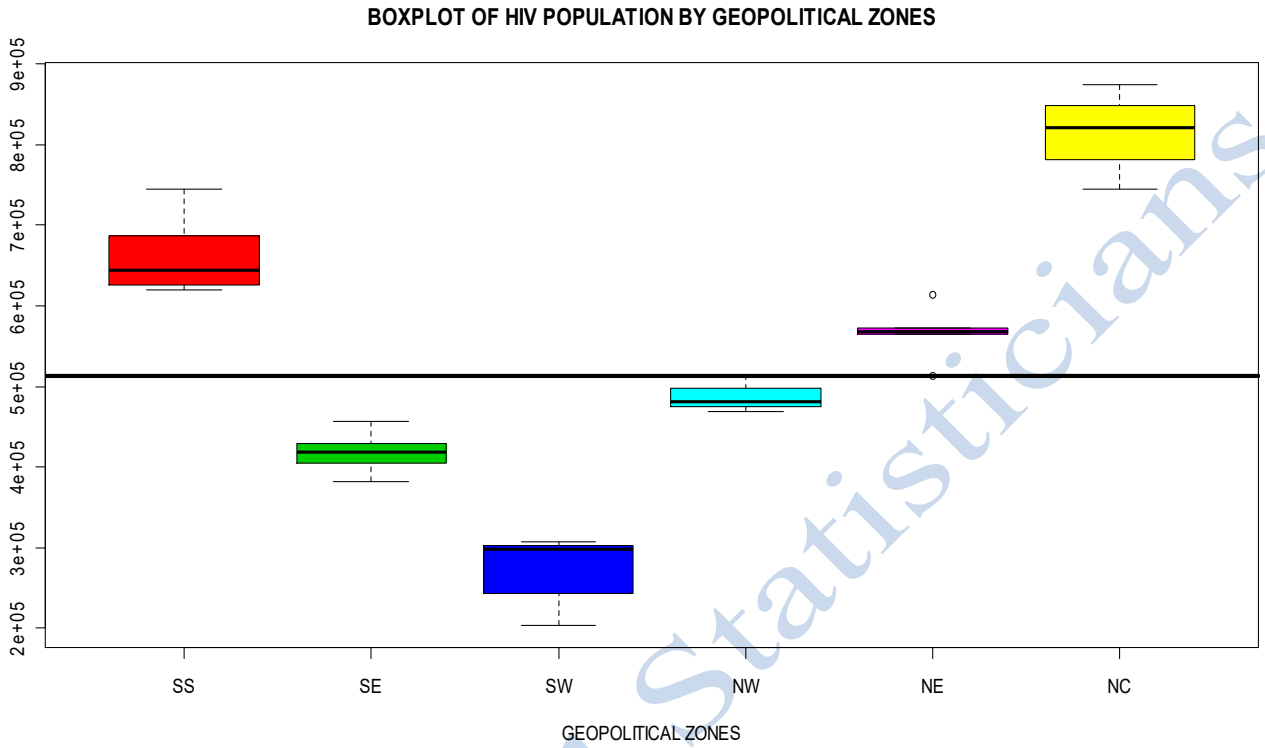

Figure 6: Box-Plot of HIV Population in Nigeria by Geo-Political Zones

It was discovered that North-Central and South-South have higher reported cases of HIV/AIDS, whereas South-West has lowest reported cases for the years being reviewed. Three zones, South-South, North-East, and North-Central, are above the combined median, while South-East, South-West, and North-West are below the median.

IV. Conclusion

Based on the results of the analyses of the data collected on the reported cases of HIV/AIDS in Nigerian across the 36 states and the FCT, the following conclusions ensued.

With respect to the prevalence of HIV/AIDS, Nigeria can be grouped into six clusters as follows:

i. Benue, Cross River, and Enugu States.
ii. FCT, Akwa Ibom, Nassawara, Gombe, Plateau, Adamawa, Rivers, Kaduna, Kogi, Niger and Taraba states.
iii. Kano, Delta, Bauchi, Lagos, Ebonyi and Edo states.
iv. Ogun, Katsina, Kwara, Anambra, Bayelsa, Borno, Imo, Kebbi and Zamfara states.
v. Sokoto, Abia and Yobe states.
vi. Osun, Oyo, Ondo, Ekiti and Jigawa states.

That Cross River and Benue states have the highest HIV/AIDS reported cases in Nigeria for the period considered. Followed by FCT, Akwa Ibom, Niger, Enugu, Nassarawa, Taraba, Rivers, Adamawa, Kaduna, Kogi, Gombe, and Plateau states. Sokoto, Ebonyi, and Edo states are a bit above the national average reported cases. Abia and Delta states almost the same as the average. However, Bayelsa, Anambra, Kano, Yobe, Lagos, Bauchi, Kebbi, Zamfara, Borno, Imo, Ogun, Katsina, and Kwara states are below average. While Ondo, Oyo, Jigawa, Ekiti, and Osun states are having the lowest reported cases.

That the geopolitical zones can be grouped into two clusters with South-West, South-East, and North-West as a cluster while South-South, North-East, and North-Central as another cluster

That among the geopolitical zones, North-Central had the highest HIV/AIDS cases closely followed by South-South and then North-East. Reported cases of HIV/AIDS in the North-West are high but a bit below the national average. South-East and South-West are both having the lowest reported cases of HIV/AIDS.

Arising from the above, it can be recommended that states identified with high reported cases of HIV/HIDS should be given utmost priority and immediate intervention be taken to avert a surge in cases of HIV/AIDS. Awareness campaigns should be intensified to significantly reduce the risks of contracting HIV/AIDS.

References

Adeokun, L. (2006). Social and cultural Factors Affecting the HIV Epidemic. In Olusoji, A. et al., AIDS in Nigeria. A nation on the Threshold. Harvard University Press, Cambridge, 151-173.

Chauhan, A. S., Hussain, M. A., Pati, S., Nallala, S., and Mishra, J. (2011). Knowledge and attitudes related to HIV/AIDS among Medical and Non-Allied Health Sciences Students. *Indian Journal of Community Health* 23.2: 96-98.

Everitt, B. S. and Hothorn, T. (2006). A Handbook of Statistical Analyses Using R. Chapman and Hall / CRC Press.

Everitt, B. S., Landau, S. and Leese, M. (2001). Cluster Analysis, 4th edition. Taylor and Francis.

Federal Ministry of Health of Nigeria (2003). National HIV/AIDS and Reproductive Health Survey (NARHS), Nigeria.

Mafeni, J. O. and Fajemisin, O. A. (2003). HIV/AIDS in Nigeria Situation, Response, and Prospects. Policy Projects, Nigeria.

National Action Committee on AIDS (2003). Capacity Review of SACAs and LACAs Nigeria; Baseline Report.

Odutolu, O., Mafeni, J. O., Okonkwo, P. and Fajemisin O. A. (2006). Monitoring and Evaluation of HIV/AIDS in Nigeria. In Olusoji, A. et al., AIDS in Nigeria. A nation on the Threshold. Harvard University Press, Cambridge, 537-558.

WHO/AIDS (2006): AIDS epidemic update: Special report on HIV/AIDS WHO/UNAIDS