Abstract

Mining of big data brings out hidden knowledge that medium size and sample data cannot reveal. This research analyzed Nigeria Population Census data in order to bring forth knowledge that can aid Government in social-economic decision-making. Thus, k-means algorithm, which is an unsupervised learning technique, was implemented on MapReduce with the aim of discovering knowledge from Priority Table IX of Nigeria Census Data of 2005. MapReduce was used to aid k-means computational challenges such as Euclidean distance computation, minimum sum of square error (MSSE) computation and global objective computation effectively. The big data analytics revealed local government areas that need Government Intervention in terms of low cost housing and those local governments that need urban restructuring for good distribution of population. Further work can be done by implementing other data such as malaria data of children to reveal hidden pattern and knowledge.

References

**Index Terms**

Computer Science

Algorithms

**Keywords**

k-means, MapReduce, Euclidean distance, MSSE, Global Objective Function.