**Visual Analytics Assignment 2**

**Multidimensional Data Visualisation**

**By:**

**Manan Bhatia**

**(22035587)**

**Akhilessh Sridass**

**(20104649)**

**DECLARATION**

I hold a copy of this assignment that I can produce if the original is lost or damaged. I hereby certify that no part of this assignment/product has been copied from any other student’s work or from any other source except where due acknowledgement is made in the assignment. No part of this assignment/product has been written/produced for me by another person except where such collaboration has been authorised by the subject lecturer/tutor concerned.

**Introduction:**

In a complicated world, there are many factors and causes for a country to fall into poverty like war, famine, pandemic, economic fall down, etc. Poverty is a complex issue which has affected millions of citizens around the world. Poverty is a state of deprivation where either an individual, a community or a region lacks the resources needed to meet the basic needs like food, shelter, clothes, healthcare, education and clean water to drink. Poverty is measured in many ways like the level of income, living conditions, or access to necessary services.Poverty is caused in many factors, like an economic recession where a country either goes broke, inflation and unemployment is rising, war, famine, pandemic and a lack of law and order leads to poverty.

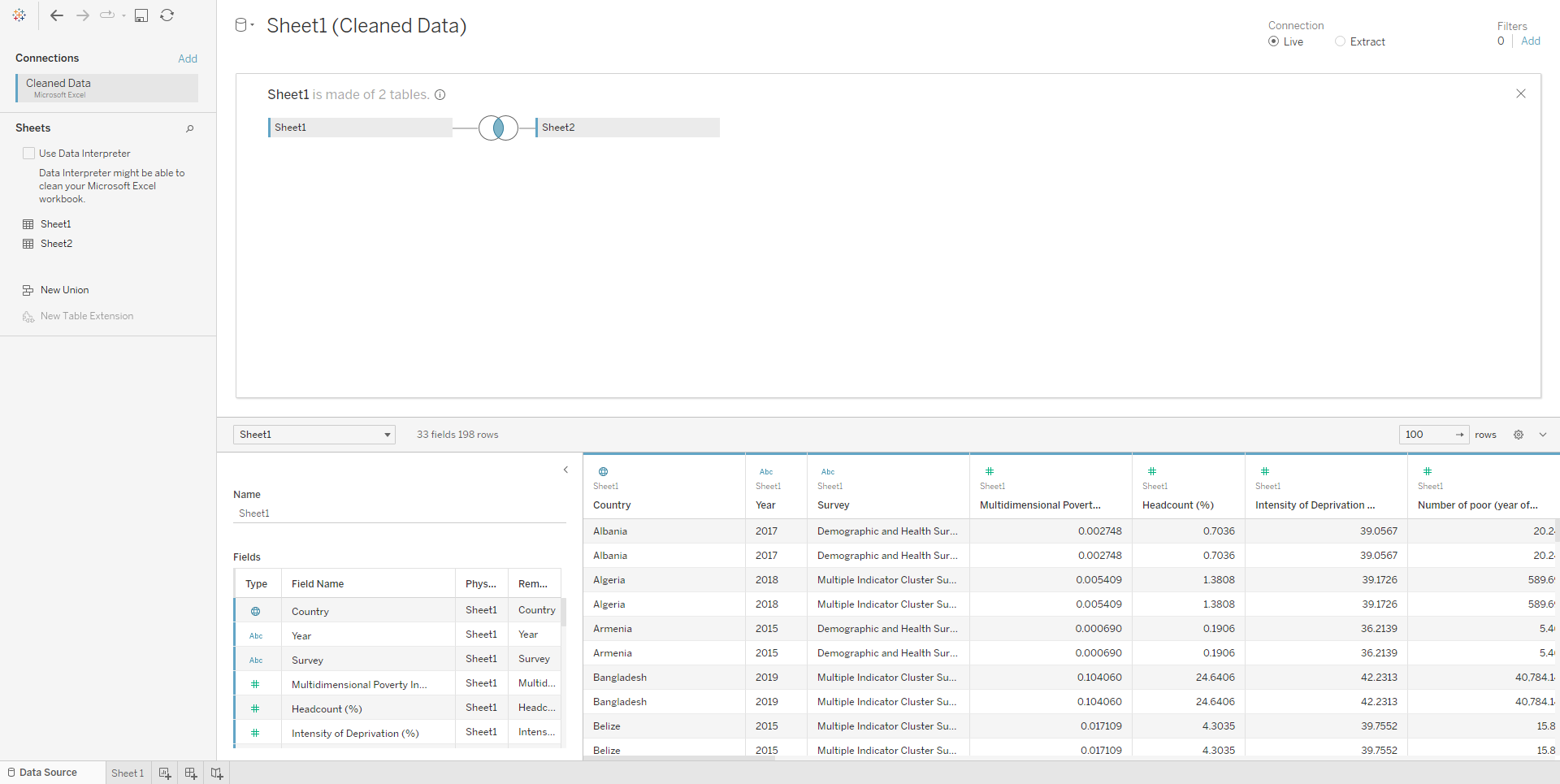
Using Tableau, we will be creating various visualisations to identify which country has been hit hard by poverty. Tableau is a visual analytics platform that is revolutionising how we use data to address issues by enabling individuals and organisations to maximise their data. The dataset that has been given is a multidimensional data which contains multiple tables. From these tables a number of visualisation data can be created. For this report, there will be a geospatial data, bar-plot, tree-map and a scatterplot.

**Implementing the Visualisations:**

Multidimensional Data is a data set which contains multiple different columns and when there are more columns, it is more likely to find out more insights. To produce a graphical representation from a multidimensional data given are:

* Data Gathering, Formatting and Storing
* Data Mining and Enrichment
* Data Mapping and Layout
* Rendering

In data gathering, we loaded up the data file by opening up the given csv file in Tableau. Then we formatted it by joining the two sheets.

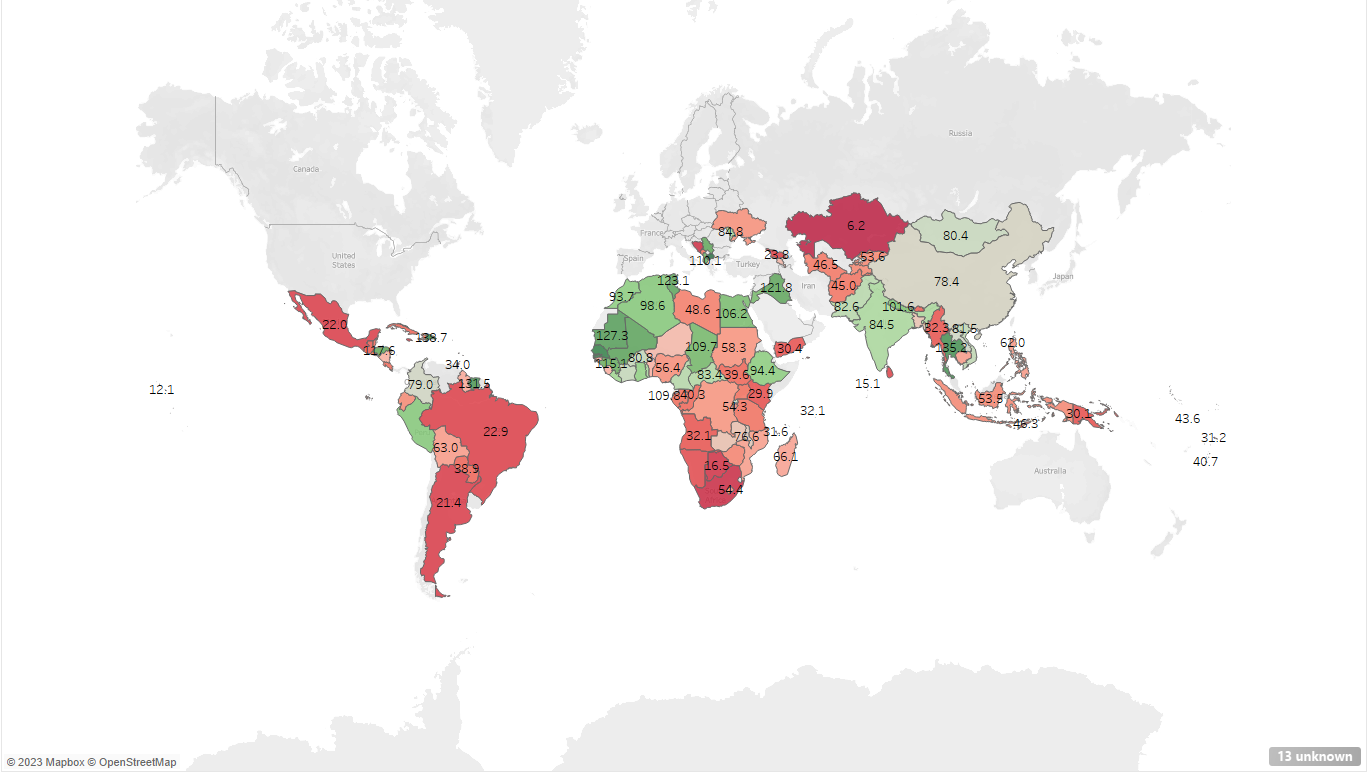


From there, we can also store all the tables that have been formatted and joined inside the data source.

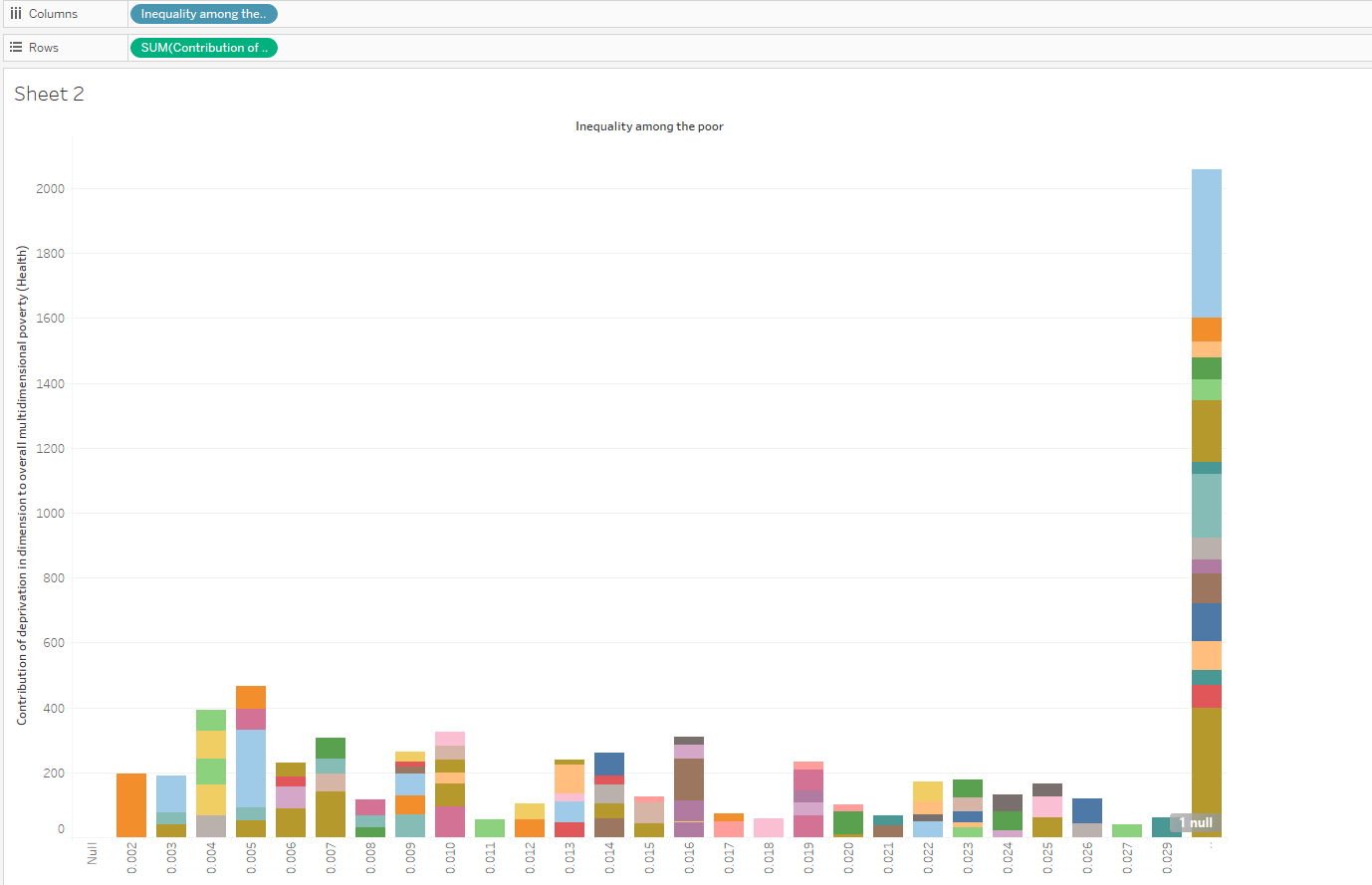
Data mining is the process of sorting through large data sets in order to identify patterns and relationships that can aid in the resolution of business problems through data analysis, whereas data enrichment is the process of fusing first-party data from internal sources with sporadic data from other internal systems. While loading up the files, from there, the software was able to recognise which visualisations to be used and what is the best tools to make the best visualisation.

While mapping out the data, there are several visualisations that have been implemented. These are, choropleth map, bar plot, tree-map, bubble chart, and scatter plot. Bar plots are a type of data visualisation that uses rectangular bars to represent data. Each bar's height represents the value of a data point, and its width represents the data category. The Choropleth Map displays data as an aggregate sum of geographic regions. Colour scales are used to assign colours to categorical or numerical data. TreeMap is a map implementation that keeps its entries organised in accordance with the keys' natural ordering or, if one was supplied by the user during construction, by using a comparator. A bubble chart is a type of data visualisation in which multiple circles are displayed in a two-dimensional plot where it can be labelled, coloured or sized.

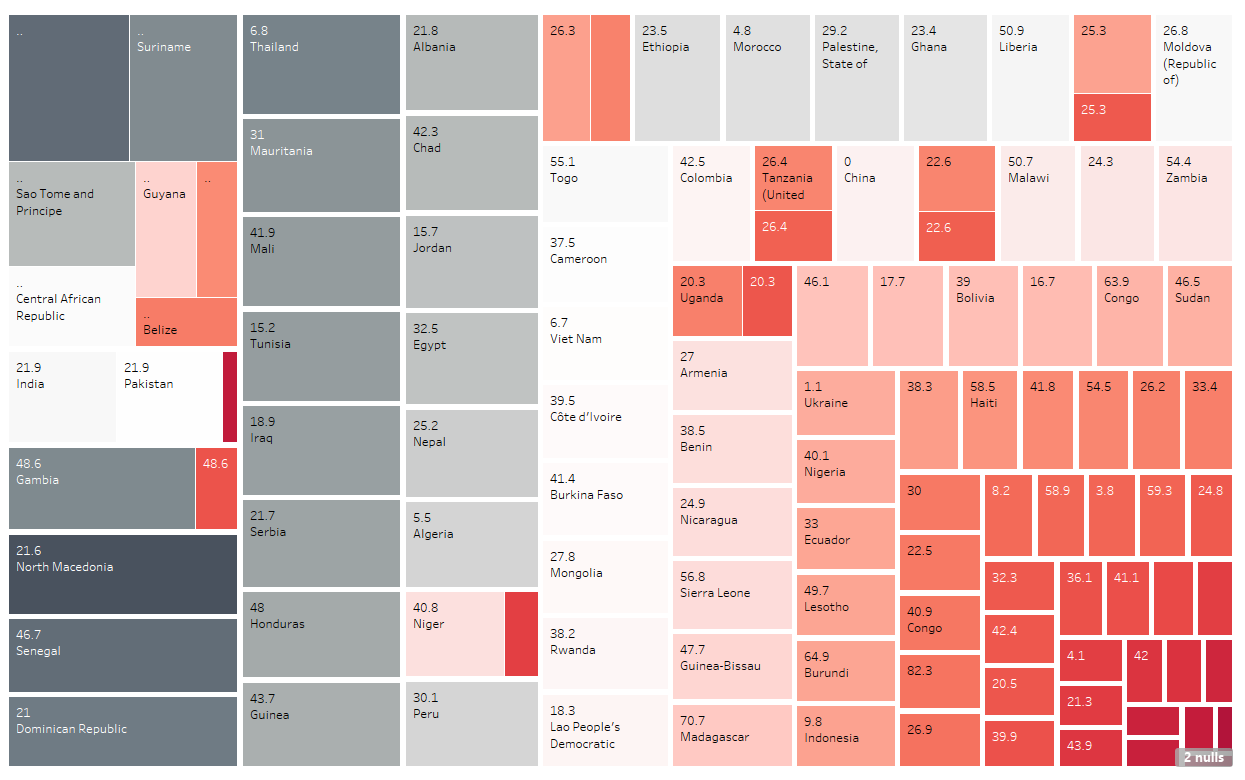
Below shows us all the visualisations done in Tableau using the multidimensional data given.

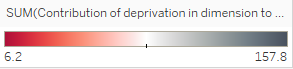


This map is based on the contribution of deprivation overall property, in relation to education. It clearly shows that the majority of the world suffers from educational issues, when it comes to providing children with the necessary requirements to give them a good education. Based on the findings from the map, it shows that Australia, the majority of North America, half of Europe and Russia, don’t seem to be affected by this. Countries in Northern Africa have high numbers of low education, showing how poor some parts of the world are. Countries near India and one country near Italy also have a high rate of low education. All of these countries have been given these contributions due to the poor education systems.

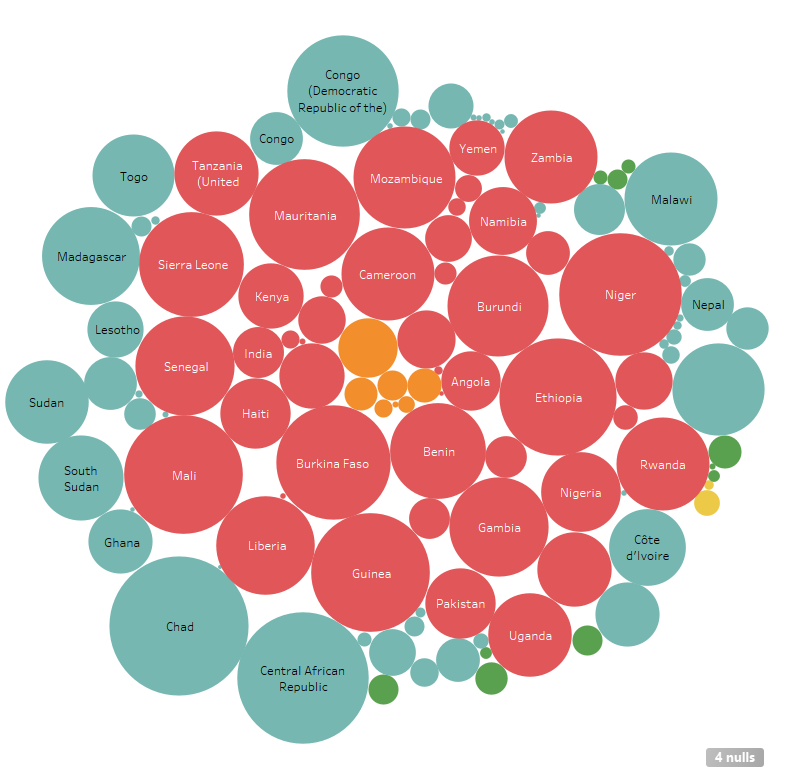


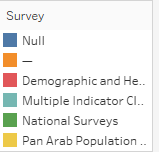
This chart focuses on the contribution of deprivation overall property, in relation to health and the inequality among the poor. The highest inequality among the poor matches with the highest health contribution which is almost 1200. Some of the lowest inequality among the poor almost reaches 0, for the contribution of health, with two of them being 0.011 and 0.018 for inequality. Majority of the inequality seems to be between the 100 and 200 range, when it comes to the contribution for health. Both inequalities of 0.004 and 0.005 have almost an equal contribution to health as they both pass the 200 mark being roughly around 250. Another trend that can be found is that most of the inequality numbers from 0.002 to 0.010, have higher numbers against the contribution of health.



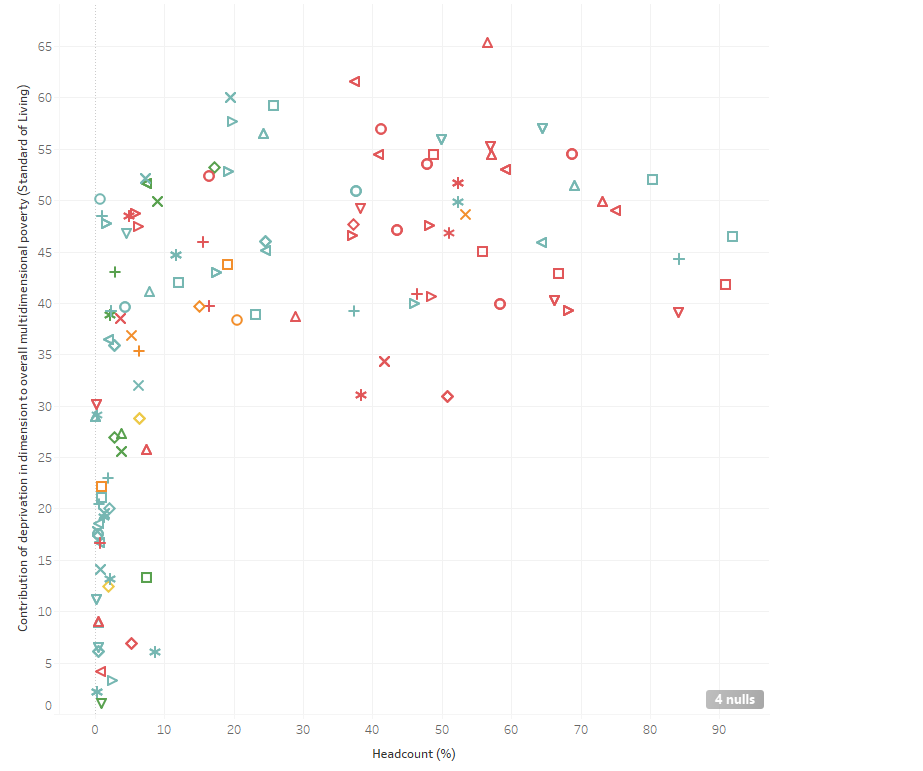


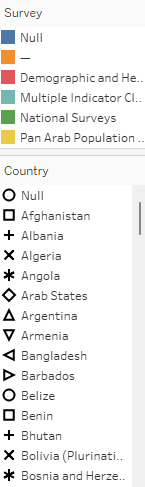
This treemap is based on the national poverty from 2009-2020 as well as showing the largest contribution for deprivation relating to education. Based on this chart, it is clear that there is a lot of variety when it comes to the different numbers being displayed. The numbers range from as low as 0 to a high of 70. While the number is hidden, the number in the black area represents the highest contribution for education which is 313.7. The closer the numbers get to the black area, based on their respective column, the higher the contribution for education is. Since the first column is the one that houses the black area, that means that the numbers in that column all have a higher contribution of education, with even a lower poverty as well.





This chart is based on the population in severe multidimensional poverty. The bigger the circle is, the bigger the poverty rate is. South Sudan and Niger both have the highest rate of poverty, due to them both being African countries. Niger actually has the highest poverty rate with a number of 76.27, with South Sudan, holding the second highest poverty rate of 74.34. While not all the countries are shown, Kenya, is the country with the lowest poverty of 12.44. Based on this chart, it is clear that mostly African countries have the lowest multidimensional poverty. Yemen and Haiti seem to be the only two worded countries that join the list of multidimensional poverty.





This scatter plot based on the deprivation in dimension to overall multidimensional poverty relating to the cost of living. With the headcount (%) as the row, we can see a logarithmic type of scatter plot with different shapes and colours. Looking at the plot, we can see that there is a difference between different countries and its relations to the cost of living. The upper part of the graph shows that the cost of living is really high and the countries with high and low headcounts are spread while having similar cost of living numbers. While hovering on the scatter plot, the countries tend to be mostly in Sub-Saharan Africa. Whilst, the bottom part of the graph tends to have low cost of living numbers and low headcount, there are many factors such as the fact that like the citizens tend to be under the poverty line is really low and that these countries are not suffering from economic disasters or war.

**Conclusion:**

In conclusion, the visualisation that has been used in Tableau has helped us in identifying and analysing the problems and the data given to us about poverty by visualising many tables together. Using visualisations like choropleth map, bar plot, tree map, bubble chart and scatter plot we were able to identify which country or region is suffering from poverty. More could have been done especially by filtering and creating parameters and also having multiple views, however, with large data given, it would have been difficult to manage is it may have changed the dataset.

**References:**

[1] Virtualitics (n.d.), What Is Multidimensional Data?, blog.virtualitics.com, viewed 10 June 2023, <<https://blog.virtualitics.com/what-is-multidimensional-data>>.

[2] SafetyCulture. (n.d.), A Guide to Data Collection Process. [online] Available at: <<https://safetyculture.com/topics/data-collection/>>.

[3] Tableau (2022). What is Tableau? [online] Tableau. Available at: <<https://www.tableau.com/why-tableau/what-is-tableau>>.

[4] Stedman, C. (2021). What is Data Mining? [online] SearchBusinessAnalytics. Available at: <<https://www.techtarget.com/searchbusinessanalytics/definition/data-mining>>.

[5] Alteryx. (n.d.). What is Data Enrichment? [online] Available at: <<https://www.alteryx.com/glossary/data-enrichment>>.

[6] Displayr. (2018). What is a Bubble Chart? [online] Available at: <<https://www.displayr.com/what-is-a-bubble-chart/>>.

[7] baeldung (2017). A Guide to TreeMap in Java | Baeldung. [online] www.baeldung.com. Available at: <<https://www.baeldung.com/java-treemap>>.

[8] www.xenonstack.com. (n.d.). Geospatial Visualization Tools and its Techniques. [online] Available at: <<https://www.xenonstack.com/insights/what-is-geospatial-visualization>>.

[9] Analytics Vidhya. (2021). Bar Plots in Python | Beginner’s Guide to Data Visualization using Bar Plots. [online] Available at: <<https://www.analyticsvidhya.com/blog/2021/08/understanding-bar-plots-in-python-beginners-guide-to-data-visualization/>>.

‌

‌

‌