

Application of Different Python Libraries for Visualisation of Female Genital Mutilation

Seun Adebajo^{a,1,*}, Emmanuel Banchani^{b,2}

^a Statistical training and consultation, Nigeria

^b Canadian Observatory on Homelessness, York University, Canada

¹ seunadebanjo9@gmail.com; ² eb1043@mun.ca

* corresponding author

ARTICLE INFO

Article history

Received March 15, 2023

Revised August 27, 2023

Accepted November 17, 2023

Keywords

python programming

python libraries

data visualization

female genital mutilation

ABSTRACT

Utilizing data visualization facilitates the analysis and comprehension of common data provided by the media, individuals, governments, and other sectors. Python is a well-known programming language that excels at scientific data visualization. This thesis utilizes a variety of Python modules, including Pandas, NumPy, Matplotlib, Seaborn, Plotly, and Bokeh, to illustrate female genital mutilation. The purpose of this thesis is to illustrate female genital mutilation and explain its performance pattern using a complex, interactive diagram that integrates multiple types of Python libraries. In comparison to other libraries, Plotly is the simplest, yet it performs at the highest level. NumPy and Matplotlib are combined to produce Hexbins charts. NumPy provides an N-dimensional plot, and Matplotlib allows for the plot's colours to be customized. Despite its limited customization options, the Seaborn library is suitable for both data visualization and statistical modelling. Due to this deficiency, the Seaborn library is frequently combined with Matplotlib to generate superior visualizations. As a result, this thesis will be recommended to both specialists and novices as worthwhile reading. In addition, it will assist the government in drafting legislation to end female genital mutilation. They will comprehend the significance of combining multiple Python modules to generate intricate interactive diagrams for data visualization in the field of data science. This information will be posted online to contribute to the corpus of knowledge.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



1. Introduction

On a global scale, a multitude of data is generated daily by individuals, organizations, social media platforms (such as Facebook, Instagram, etc.), governments, and other sectors like health and education. These data sets often pose challenges in terms of comprehension. Hence, there is a requirement to employ data visualization techniques to facilitate the comprehension of patterns, trends, and relationships within the data. Data visualization refers to the graphical display of data to facilitate comprehension for individuals without specialized knowledge of the subject matter. Data visualization refers to the systematic transformation of numerical data into non-numerical representations while preserving the original data's entropy (Tam & Thien, 2020). Various programming languages are available for data visualization in the field of data science, including SPSS, Minitab, MATLAB, R, Python, and others. However, each of these languages possesses distinct advantages and limitations. Python is widely regarded as a versatile programming language that is well-suited to visualize data in the field of data science. This is achieved through the utilization

of several libraries such as Matplotlib, Seaborn, Plotly, and Bokeh, among others, which offer a range of graphing capabilities for data visualization.

Moreover, data visualization holds considerable importance within the realm of healthcare as it enables healthcare facilities to effectively monitor the health development of their patients. Additionally, this technology facilitates the ability of nurses and doctors to access and analyze a patient's past medical records, enabling them to develop more effective treatment strategies. Moreover, this enhanced capability contributes to the provision of exceptional customer care and satisfaction for the patient. The utilization of data visualization techniques in primary healthcare settings might facilitate the enhanced comprehension of patient feedback rating reports, thereby aiding health centre management in their efforts to increase both operational efficiency and patient happiness.

Data visualization is widely utilized in the banking industry to effectively represent the annual financial profit and loss statement of a bank. Additionally, this feature facilitates the bank's personnel in effectively interpreting clients' transaction history in instances of discrepancies. Moreover, it provides users with a convenient means to gain a clear understanding of their bank statements, particularly in cases of perplexing transactions.

Furthermore, datasets exhibit variations in their structures, types, and sizes. For instance, structured data may be characterized as small data, whereas unstructured data may be classified as big data. Additionally, datasets can be categorized as univariate when they consist of a single data variable, bivariate when they involve two categories of data, and multivariate when they encompass more than two categories of data. The inherent complexity of the data type may provide challenges in its raw form, hence emphasizing the importance of visualization as a means to enhance comprehension. In their study, Psallidas et al. (2019) utilized various Python libraries, including Matplotlib and Seaborn, to enhance data visualization in the field of data science. The objective was to improve the clarity and comprehensibility of data presentation. The significance of data visualization cannot be overstated in various contexts, including professional settings, everyday life, and governmental policy-making. The health sector has experienced significant advantages as a result of the implementation of data visualization techniques. This thesis will primarily examine the utilization of various Python libraries, including Pandas, NumPy, Matplotlib, Seaborn, and Plotly, for data visualization in the context of female genital mutilation. The objective is to assess the performance of these libraries and determine which ones exhibit superior capabilities.

2. Review of Related Work

In recent times, the utilization of data visualization has become an essential tool not just within the realm of data science but also across various other disciplines, including medicine, statistics, agronomy, economics, and applied mathematics, among others (Raschka, 2018). Data exploration, data mining, and big data are all processes that heavily depend on the utilization of data visualization techniques. Each organization produces visual depictions of the data they collect at various intervals, such as hourly, daily, weekly, monthly, and sometimes annually, depending on the specific circumstances. This practice is implemented to facilitate stakeholders' understanding and interpretation of the information.

Moreover, a multitude of widely recognized statistical software packages, such as SPSS, STATA, EViews, and Stata, can be employed for data visualization and analysis. However, the choice of statistical tool varies among experts due to the influence of the specific application in deciding the appropriate instrument to be employed. The program often utilized in the field of social sciences is SPSS, an acronym for "Statistical Package for the Social Sciences". The SPSS software, initially developed for the social sciences, has expanded its usage to encompass several fields such as market research, data collection, survey administration, and health research, among others. Analysts extensively utilize the statistical software STATA in their research and data analysis endeavours. It is a utility commonly employed by several persons. When comparing SPSS with STATA, it becomes evident that a notable challenge associated with the latter is its relatively higher level of complexity,

particularly for individuals lacking a strong foundation in statistics. STATA is a highly adaptable software package, predominantly employed in the fields of economics, sociology, medicine, and computer science.

Python and R are widely recognized computer languages commonly used for data visualization and analytical purposes. The superiority of these languages for analytical tasks stems from their advanced level, dynamic and interactive nature, and the extensive availability of scientific literature. The utilization of Python has witnessed a notable surge across diverse fields, encompassing research and industrial domains, commencing from the early 2000s. The utilization of Python libraries for data visualization in research and business applications has witnessed a notable surge in recent years.

Sehgal et al. (2019) employed the Python programming language and the Seaborn data visualization packages to visualize data. Seaborn has been shown to possess the ability to effectively handle multivariate data and provide visually appealing representations. Nonetheless, it is important to note that Seaborn's customisation options are notably constrained, hence often requiring its utilization in conjunction with Matplotlib. This study employs an interactive visual environment to examine the patterns of data, while also offering a comprehensive elucidation of the function played by the Seaborn Python library in the visualization of datasets.

In their study, Stancin and Jovic (2019) conducted a comparative analysis of various Python modules to evaluate their suitability for data mining and large-scale data analysis purposes. The research findings indicate that the utilization of diverse Python libraries enables the customization of data visualization through the manipulation of datasets and the presentation of outcomes in a range of graphical representations, including graphs, charts, maps, and plots. One of the findings of the investigation was as follows. This observation illustrates that making modifications to different Python modules will have an impact on the pattern of data visualization. The study conducted by Stancin and Jovic (2019) suggests that for effective data visualization, it is recommended to utilize Python libraries such as Matplotlib, Seaborn, and Plotly. Additionally, for handling massive datasets, the study recommends employing Park and Hadoop.

In the domain of data science, Waskom (2018) employed the Seaborn Python package to facilitate data visualization and statistical modelling. Based on the findings of his study, Seaborn emerges as a very proficient tool for data modelling and visualization. Nonetheless, the researcher posits that augmenting Seaborn with Matplotlib would yield enhanced customisation capabilities. Furthermore, the study provides evidence that the utilization of Seaborn libraries is effective in visually depicting a significant proportion of the dataset.

In their study, Oberoi and Chauhan (2019) employed the Matplotlib and Seaborn tools to virtualize data. The research findings indicate that Matplotlib is well-suited for fundamental visualization tasks and offers customization capabilities. On the other hand, Seaborn is deemed appropriate for handling extensive datasets and producing high-quality visual representations. However, it is recommended to utilize Matplotlib in conjunction with Seaborn to access a wider range of graphical possibilities.

Moreover, Seaborn, a Python library, is built upon the foundation of Matplotlib. It offers a wide range of essential functionalities that enhance its adaptability in visualizing diverse forms of data, encompassing univariate, bivariate, multivariate, and time-series data. The Seaborn software package was developed by the Seaborn crew, as documented in the work by Odegua (2020).

Python is widely recognized as a crucial programming language within the field of data science owing to its inherent adaptability in facilitating data modelling and visualization through an extensive range of libraries. The utilization of Python in data analysis allows for the creation of visual representations, which in turn enhances the process of interpreting and predicting future trends. Furthermore, the inclusion of a visual depiction of data facilitates a more straightforward interpretation and enables the anticipation of potential future patterns (Sindhu, 2018).

In their study, Shangjia, Yuhan, and Songlin (2021) conducted research on Python data visualization technology and utilized the Matplotlib package for implementation purposes. This technique holds

considerable importance in various domains like as medical diagnosis, commercial decision-making, and engineering applications, owing to its code's simplicity and the production of visually appealing Matplotlib visuals.

In his study, McKinney (2017) investigates the Python Library for Data Analysis and its application in data visualization, specifically focusing on the use of Pandas. According to the individual, Pandas is a Python library that is built on NumPy and is utilized to alter and analyze data. The author illustrated that Python is a widely recognized library in the field, offering reliable data operations and structures for the manipulation of tabular data.

In the field of data science, Li (year) analyzes several Python libraries, including Matplotlib and Plotly. The researcher found that Matplotlib is a Python library that offers a straightforward and efficient solution for visualizing data, while Plotly is a tool that excels in generating intricate and dynamic representations.

Pandian (year) stressed the particular capabilities of Python as a data visualization tool in the field of data science. Python enables the identification of patterns, trends, correlations, associations, and outliers within datasets. The researcher employed Python tools, specifically NumPy and Matplotlib, to develop Hexbin plots. These plots were leveraged to improve the presentation of scattered plots when dealing with huge sample sizes.

In the study conducted by Mishra (2019), the focus is on exploring the application of Python in data visualization. Specifically, the researcher utilizes the Matplotlib library as a fundamental tool for individuals at different skill levels, ranging from novices to experts, to effectively represent data through a diverse range of visualizations, including histograms, bar charts, pie charts, box plots, and scatter plots. Furthermore, the researcher emphasizes the importance of customization to enhance the visual appeal of the presented data. This information can be found in the publication by Mishra (2019). It was found that the Matplotlib module offers valuable graphical components, and it was further observed that Seaborn is an auxiliary Python library that incorporates Matplotlib.

In a recent study, Rastogi (2021) explores the application of the Bokeh Library in Python for data visualization. The findings of the study reveal that Bokeh exhibits a seamless ability to build highly sophisticated interactive visualizations. The study additionally demonstrates that Bokeh possesses a diverse array of graphical functionalities and exhibits seamless compatibility with the Pandas library. The investigation further demonstrates that Bokeh's interaction is constrained and its grammar exhibits greater complexity.

This project aims to construct a comprehensive interactive diagram utilizing several Python tools to illustrate and analyze the practice of female genital mutilation and its patterns of occurrence. This will enhance the existing body of knowledge and furnish scholars, data scientists, and policymakers with valuable insights.

3. Research Method and Implementation Process

This study utilized a quantitative research design and obtained data from the United Nations Children's Fund (UNICEF) publication, which served as a secondary data source. The data was specifically gathered from eight countries, namely Burkina Faso, Egypt, Ethiopia, Gambia, Guinea, Kenya, Mauritania, and Nigeria, exclusively for this research. Various sophisticated visualizations, such as mixed subplots, scatterplots, heat maps, and 3D graphs, can be employed to effectively represent datasets. Additionally, the combination of vertical tables and charts can be utilized to enhance data visualization. These visualizations can be implemented using a range of Python libraries.

Simultaneously, an online platform will be created as a tangible representation, showcasing a collection of interactive diagrams generated through the integration of various Python libraries such as Pandas, NumPy, Matplotlib, Seaborn, Plotly, and Bokeh.

To execute this procedure, the dataset about female genital mutilation will be organized in a manner that eliminates any inconsistencies, inaccuracies, or instances of missing information, hence ensuring the integrity and dependability of the data. This particular step is commonly referred to as data cleansing. The subsequent chapter will showcase the outcome of female genital mutilation through the utilization of various Python libraries.

3.1 Justification for the use of Python for visualisation

The Python programming language was selected as the tool for visualizing the dataset on female genital mutilation due to its versatility, robustness, and extensive capabilities, which allow for enhanced flexibility and broad applicability. Python possesses a distinct edge over other data analysis and visualization applications or software, such as Stata, SPSS, MATLAB, and others. Moreover, it possesses the capability to effectively manage substantial volumes of data without experiencing system failures or crashes. Moreover, it facilitates computational processes and exhibits user-friendly features that enhance analytical capabilities.

4. Results and Discussion

4.1 Results

Heatmap For Female Genital Mutilation

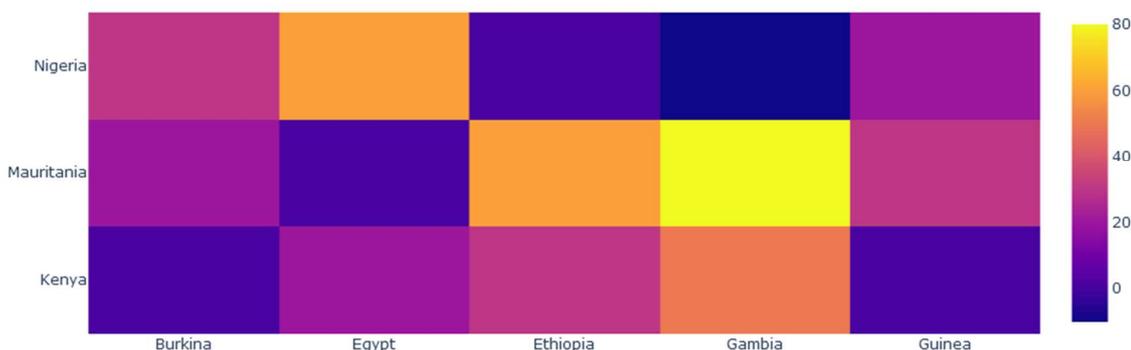


Figure 1. Interactive Heatmap with Plotly and Pandas

The heatmap presented in Figure 1 illustrates the patterns of interaction among eight African countries (Kenya, Egypt, Burkina Faso, Gambia, Guinea, Mauritania, Nigeria, and Ethiopia) that continue to engage in the practice of female genital mutilation. The data manipulation process employed the Pandas library, while the production of the visually striking and interactive diagram was facilitated by Plotly. This utilization of Plotly distinguishes it from other Python libraries employed in this research. The exceptional interactive features were facilitated by the utilization of the Plotly library, while the manipulation of the dataset was facilitated by the Pandas library.

Figure 2 depicts the aforementioned Chart, which was generated utilizing data visualization frameworks such as Plotly and Pandas. The chart presents a visual representation of the relationship between female genital mutilation and wealth quantile across different countries. It illustrates the prevalence of female genital mutilation within each wealth quantile and compares it between countries. Based on the analysis conducted using the chart, it is evident that Mauritania exhibits the greatest rate of female genital mutilation, while a

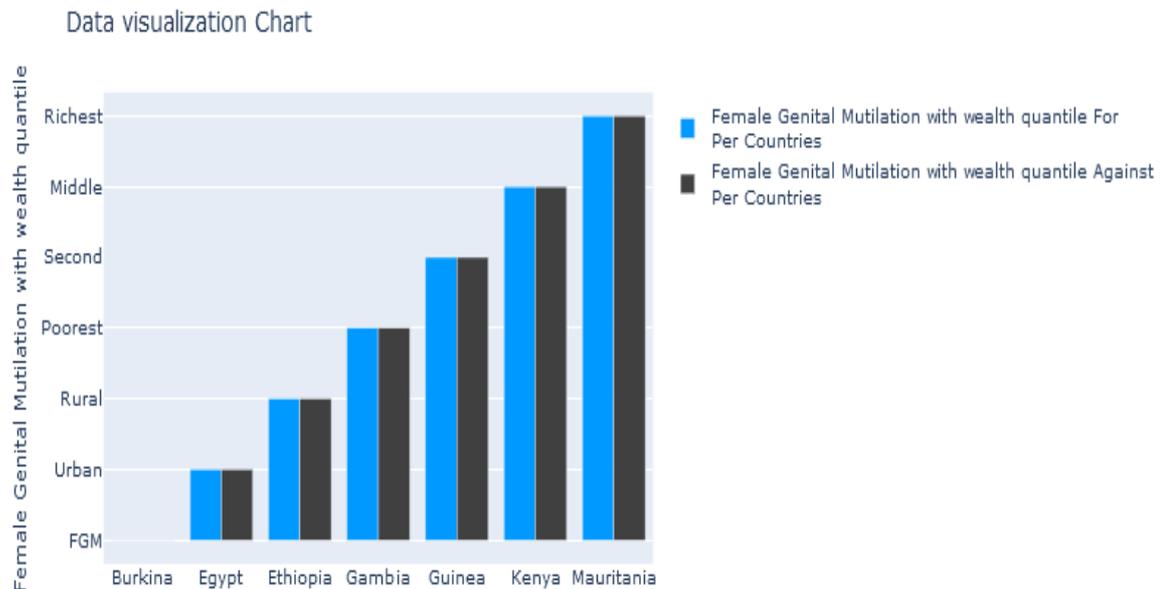


Figure 2. Chart with Plotly and Pandas

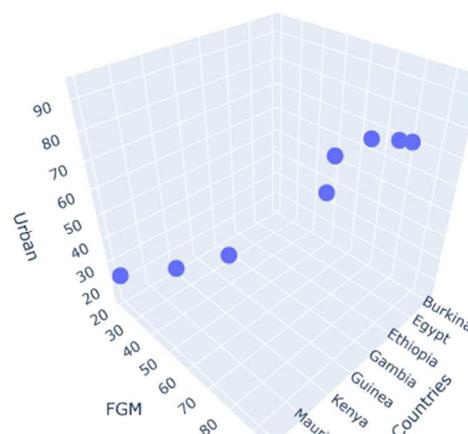


Figure 3. Plotting 3D graph with Plotly and Pandas

significant number of countries have a relatively lower prevalence of this practice. The visualization of the countries and income quantile variables was achieved through the utilization of the Chart, employing data manipulation techniques with the Pandas and Plotly libraries. It is worth noting that Pandas excels in its ability to efficiently manipulate data, whilst Plotly stands out for its exceptional visualization capabilities.

Figure 3 presents a three-dimensional graphical representation of female genital mutilation (FGM) in urban countries such as Burkina Faso, Egypt, Ethiopia, Gambia, Guinea, Kenya, and Mauritania. This visualization was created using the Plotly and Pandas Python packages.

Figure 4 presents a Heat map visualization depicting the prevalence of female genital mutilation in African countries. This visualization was created using many Python tools, including Plotly, Seaborn, and Matplotlib, resulting in a sophisticated and interactive figure. The Heat map visually represents the interconnectedness within the dataset through the utilization of libraries, while Matplotlib facilitates the distinct customization of the diagram.

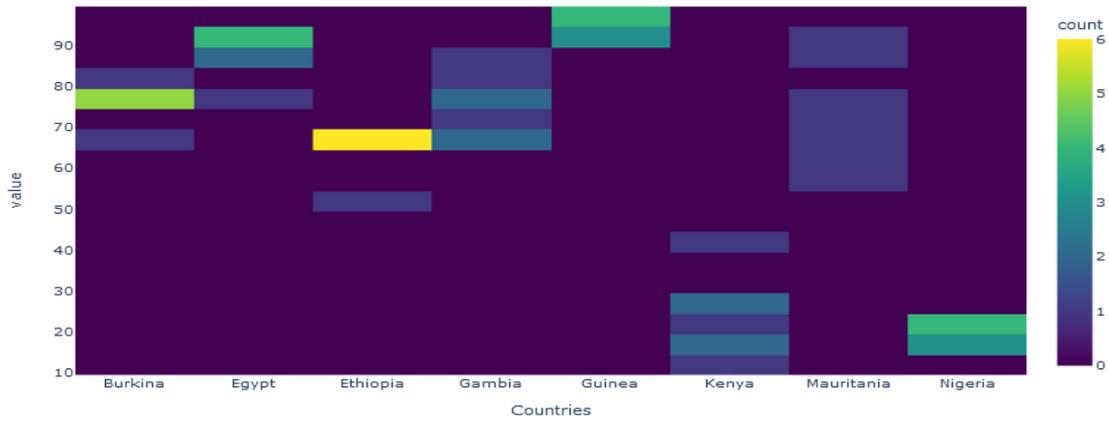


Figure 4. Heat map by combining Plotly, Seaborn and Matplotlib

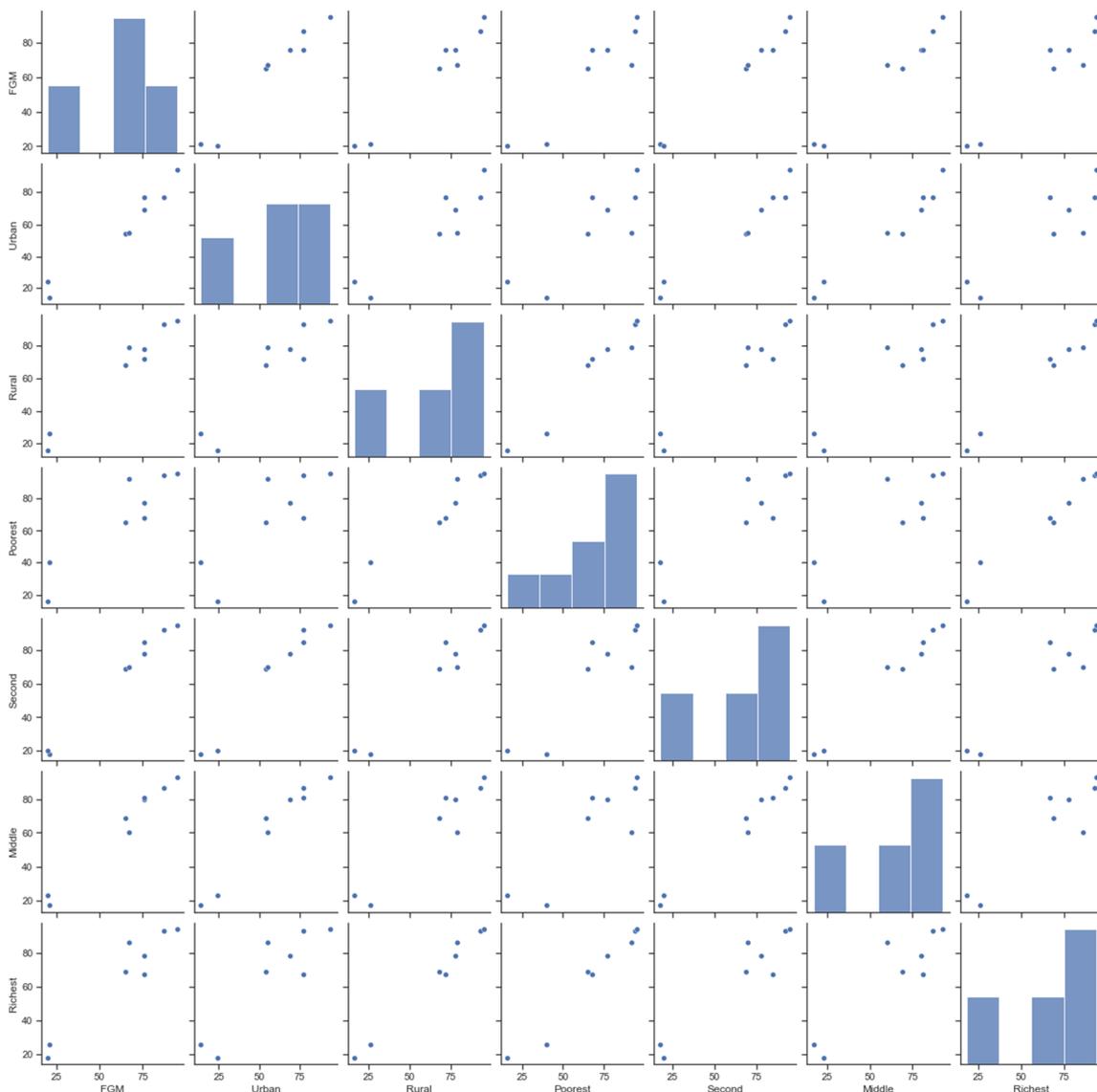


Figure 5. Scatter Plot matrix by combining Pandas, Seaborn and Matplotlib

The scatter plot matrix, as depicted in Figure 5, is generated by the integration of many Python modules, including Pandas, Seaborn, and Matplotlib. The scatter plot facilitates the examination of the association between wealth quantile variables (e.g., urban, rural, poorest, second, middle, and richest) and the prevalence of female genital mutilation (FGM). The data analysis was conducted using the Seaborn Library, which facilitated the creation of scatterplots. The Matplotlib library was employed for customizing purposes, while the Pandas library was utilized for data manipulation.



Figure 6. Plotting image with Plotly

Figure 6 depicts the visual representation of an astronaut photograph accompanied by the display of the United States flag, achieved through the utilization of a Python library such as Plotly. The presented content showcases an interactive diagram featuring a vibrant visual representation.

The graphic showcased the exceptional capabilities of Plotly as a Python library in terms of its superior visualization compared to other libraries. In contrast, the utilization of Seaborn in isolation does not suffice to provide an identical visual representation, necessitating its combination with Plotly. One significant drawback of the Seaborn library in visualization is its reliance on Plotly for generating exceptional images. In contrast, Plotly alone is capable of producing amazing images, hence demonstrating its superior performance compared to other Python libraries. For instance, by examining the provided plot, it is possible to generate a heatmap utilizing the Seaborn library without the integration of Plotly.

Figure 7 displays a visual representation of a scatter plot depicting two sets of numerical data through the utilization of Hexbins. This visualization was created utilizing two distinct Python libraries, namely NumPy and Matplotlib. The Hexbins plots provide an advancement above conventional dispersed plots, offering a more effective display pattern. The NumPy library is well-suited for handling integer data and effectively manipulates the dataset depicted in Figure 7. On the other hand, Matplotlib facilitates the creation of visualizations, offering a range of customization options.

Figure 8 depicts interactive multiple-line plots that utilize Python packages such as Pandas and Plotly to visually represent African countries with instances of female genital mutilation and their accompanying wealth quantile variables. Based on the conducted analysis, it can be inferred from the line plot that each line represents different variables. Notably, Nigeria and Kenya exhibit the lowest prevalence of female genital mutilation, while the rural

and economically disadvantaged countries, such as Mauritania, Egypt, and Guinea, demonstrate the highest rates of this practice as depicted in the provided diagram. The graph utilized the variable key to denote the representation of the lines. The data manipulation process was performed using the Pandas library, while the visualization aspect was accomplished using Plotly.

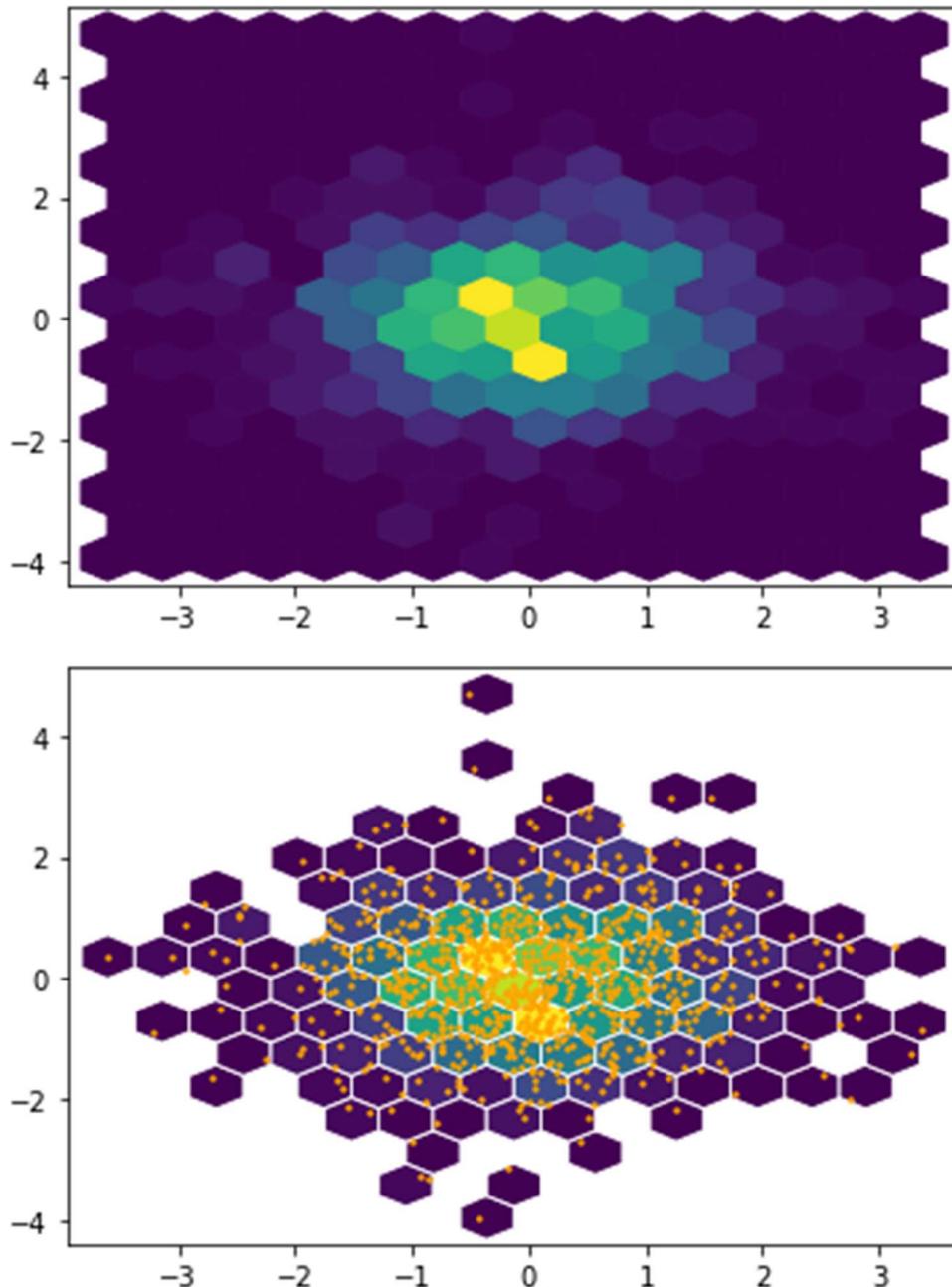


Figure 7. Hexbins plots with NumPy and Matplotlib

Figure 9 shows the interactive multiple Bar chart that visualizes eight Africa countries with cases of genital mutilation using Pandas and Plotly libraries.

Figure 10 shows the box plot that visualizes eight Africa countries with female genital mutilation cases using Python libraries like Pandas and Plotly. Box plot also helps to identify outliers in the dataset.

Data Visualization

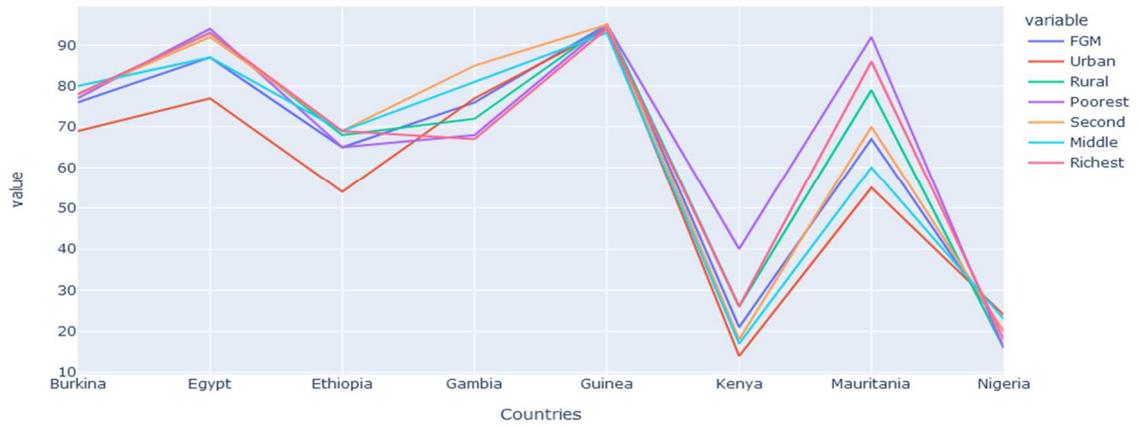


Figure 8. Interactive Multiple Line Plots with Pandas and Plotly

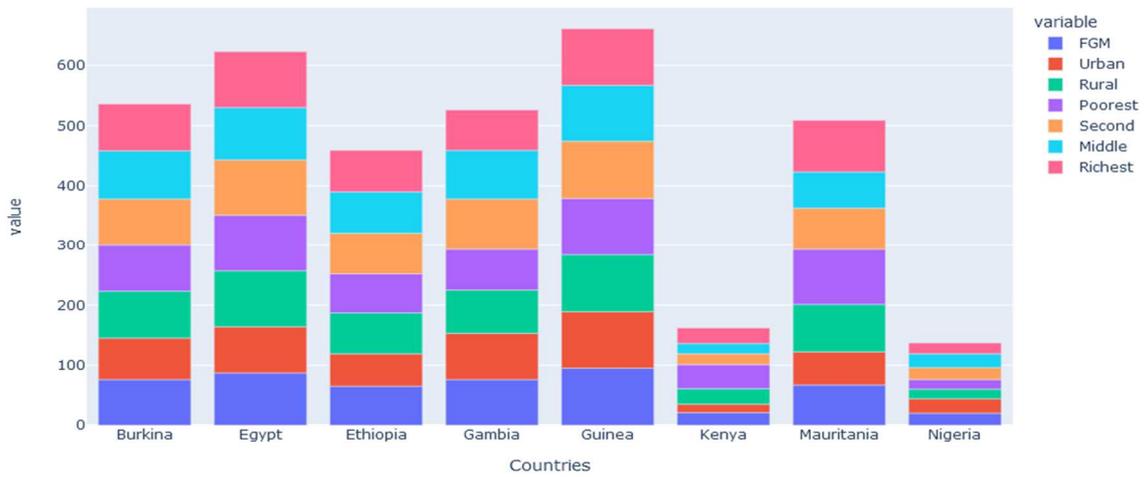


Figure 9. Interactive Multiple Bar chart with Pandas and Plotly

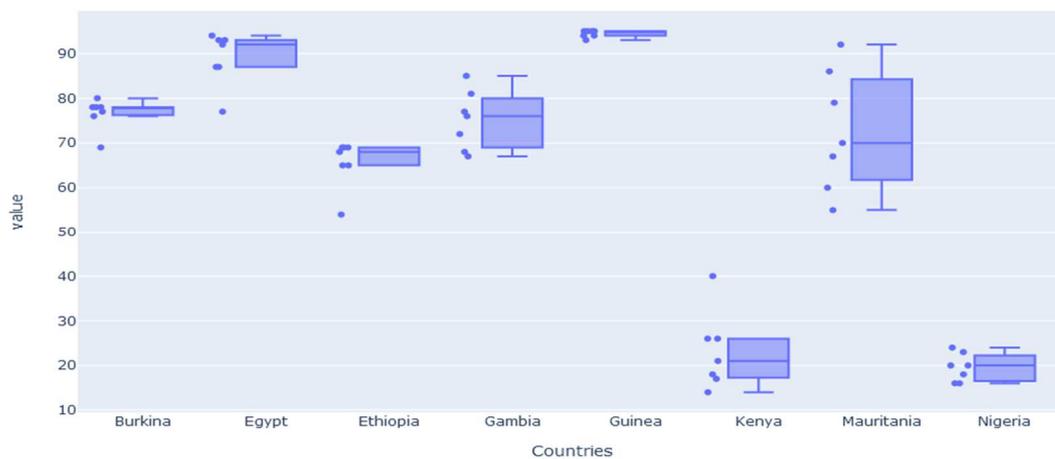


Figure 10. Interactive Box plot with Pandas and Plotly

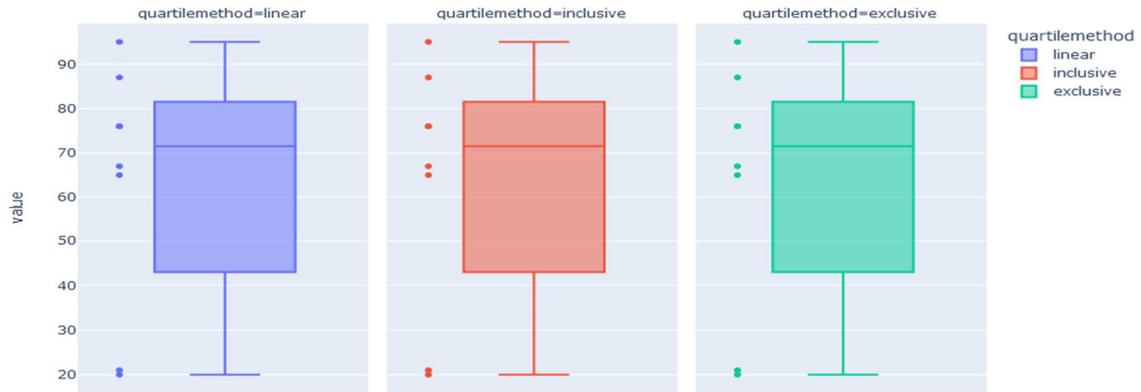


Figure 11. Female genital mutilation visualization to show three different quartile computation

Figure 11 shows the difference between three quartile computation method such as linear, inclusive and exclusive with combine python libraries of Pandas and Plotly to visualize the female genital mutilation (FGM).

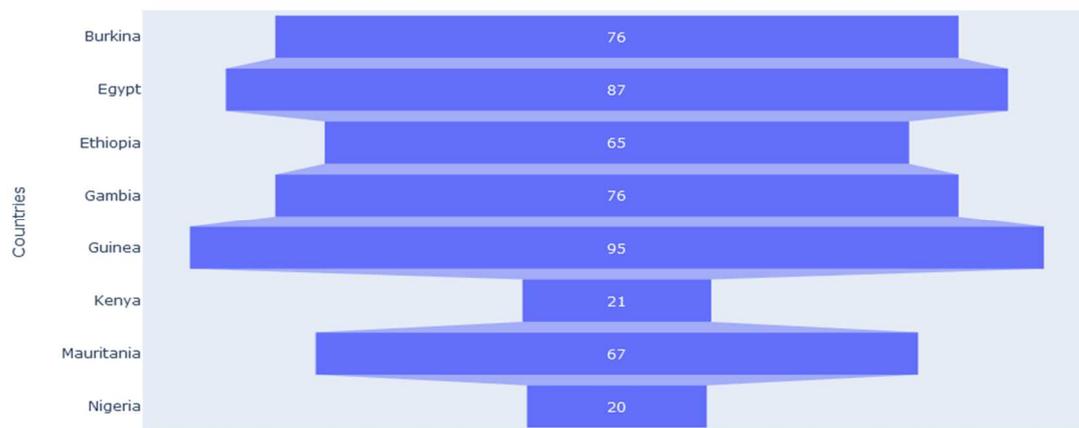


Figure 12. Basic Funnel Plot with Plotly and Pandas

Figure 12 presents a visualization depicting the distribution of nations affected by incidents of female genital mutilation. The visualization was generated using the Basic funnel plot technique, utilizing the Plotly and Pandas libraries. The Funnel plot can be utilized to identify the countries with the highest and lowest values. By employing this method, we can effectively evaluate the outcomes and determine the countries that exhibit the widest and narrowest ranges.

Figure 13 shows the visualisation of the countries with cases of female genital mutilation (FGM) with density heatmaps plot using combination of python libraries like Plotly and Pandas. Plotly produces the outstanding interactive diagram.

Figure 14 shows interactive line graph indicating the connection of each of each Africa countries with female genital mutilation. It shows the rate of female genital mutilation in each of the countries. The visualization was done by Plotly library.

Figure 15 displays an interactive scatter plot depicting the prevalence of female genital mutilation in different nations, categorized by their corresponding wealth quantiles (Urban, Rural, Poorest, Second, Middle, and Richest). The visualization was created using Plotly. It is important to note that Seaborn is unable to produce a scatter plot independently,

necessitating its combination with Matplotlib. This characteristic distinguishes Plotly as a more prominent library compared to other libraries examined in this study.

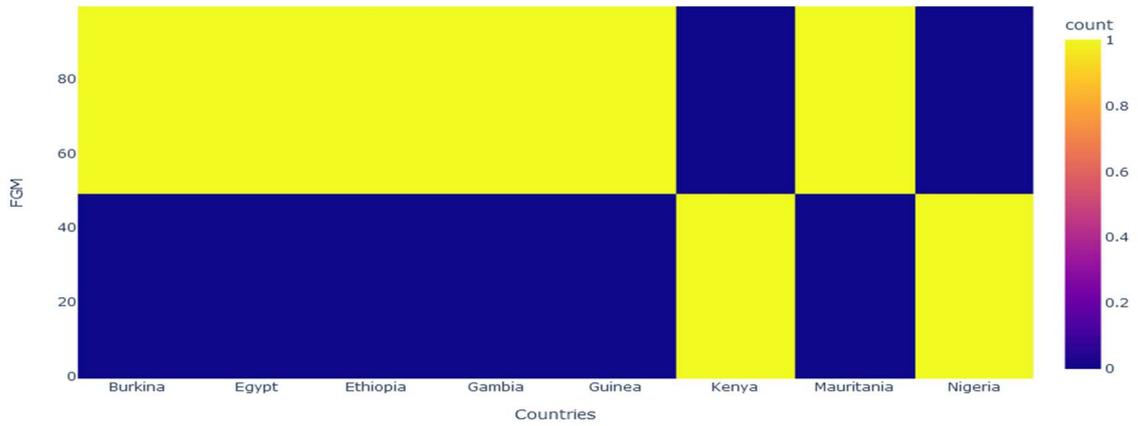


Figure 13. Density Heatmaps with Plotly and Pandas

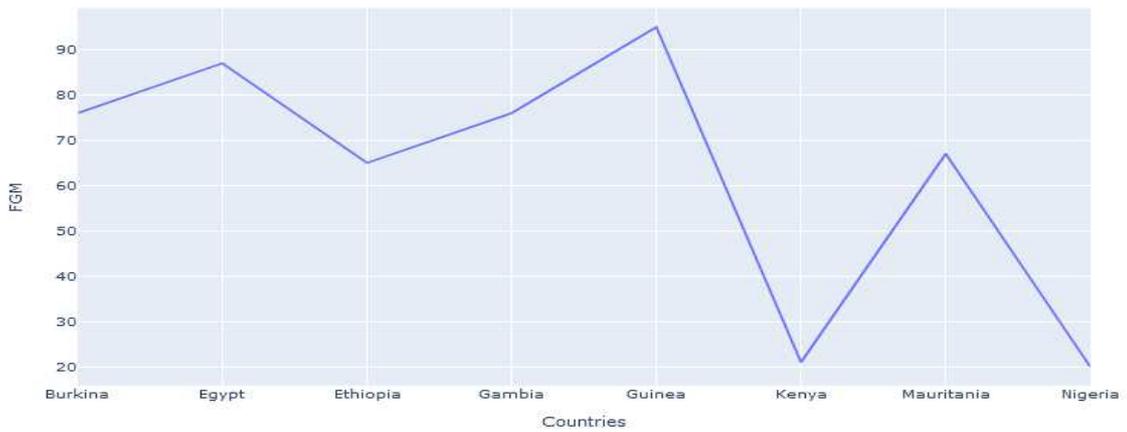


Figure 14. Interactive line graph with Plotly and Pandas

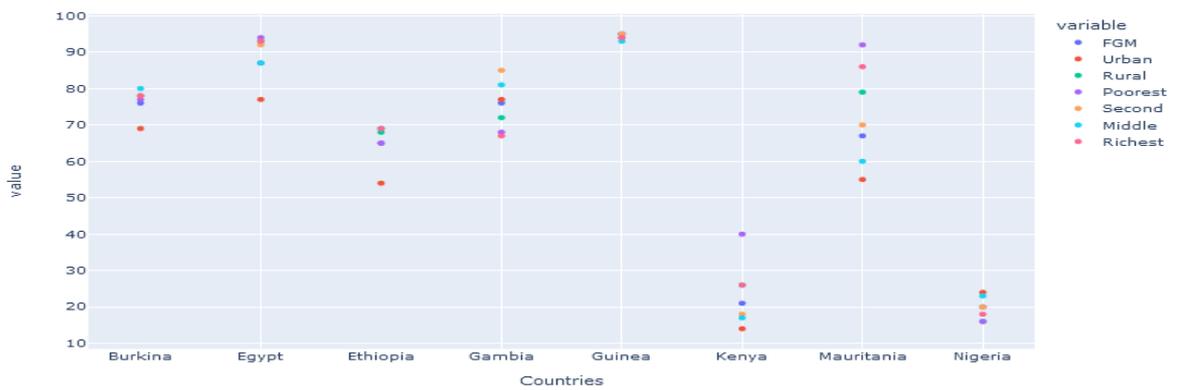


Figure 15. Interactive scattered plot with Plotly

Female Genital Mutilation

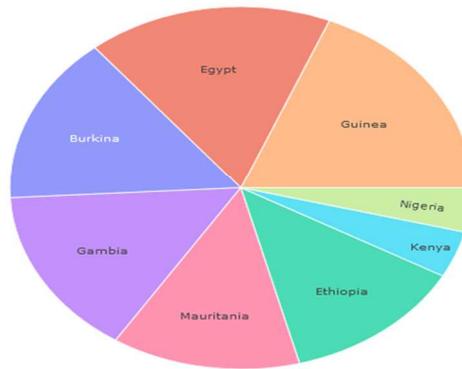


Figure 16. Advanced Interactive Pie Chart with Plotly

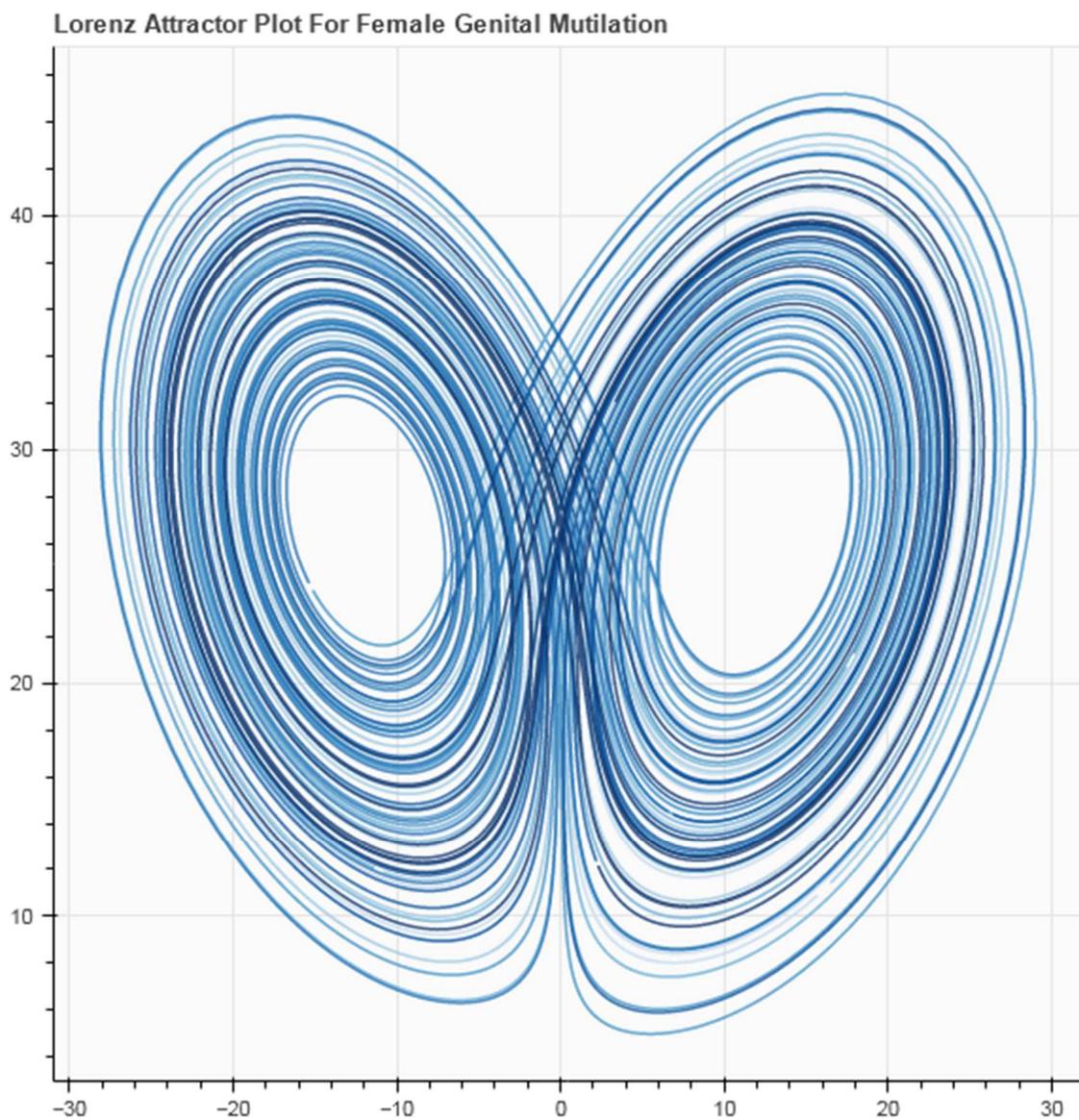


Figure 17. Lorenz Attractor plot

Female Genital Mutilation

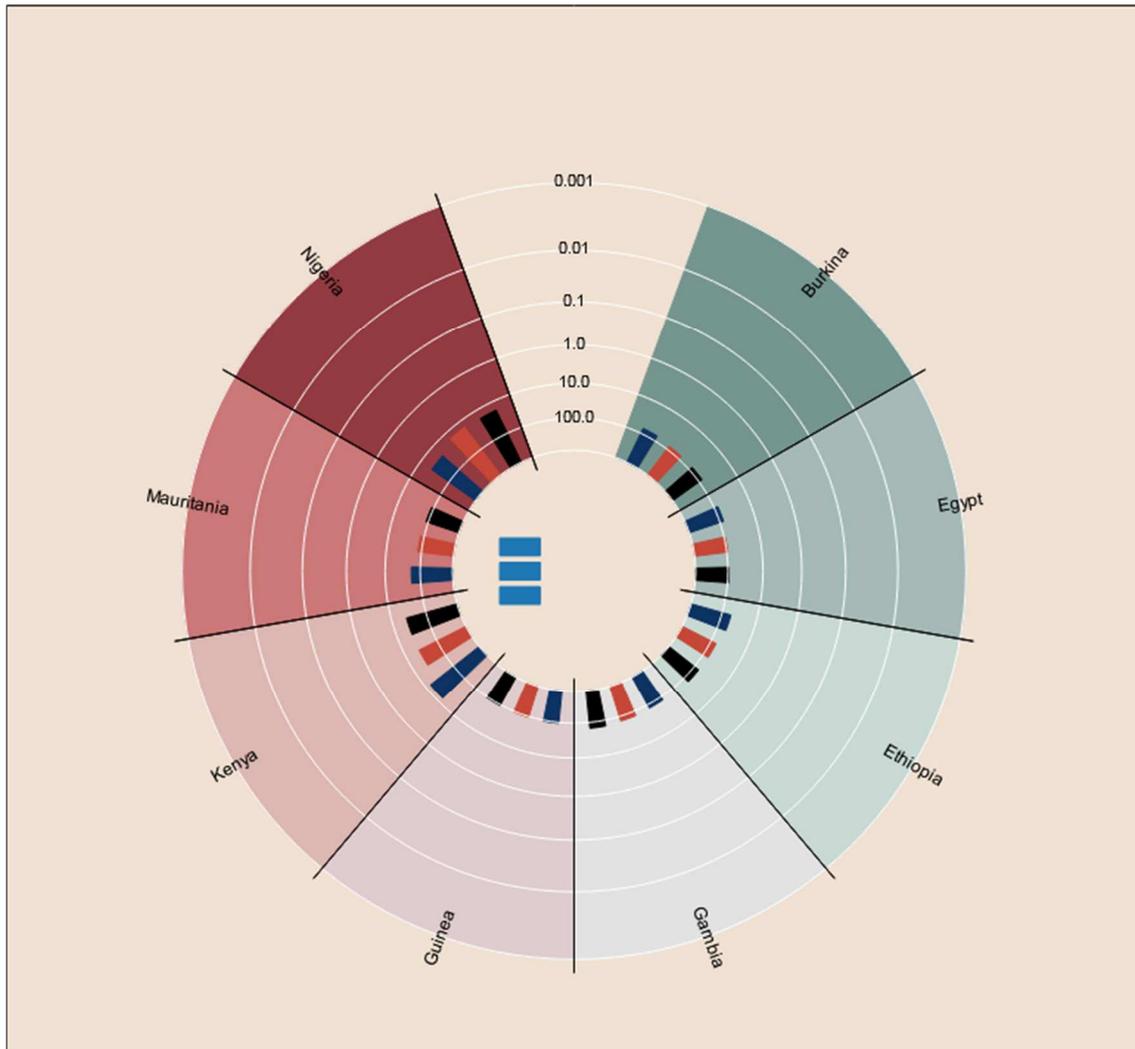


Figure 18. Circular Diagram with Bokeh

Figure 16 depicts an advanced Pie chart illustrating the interplay between female genital mutilation and the eight nations in which it is prevalent. The interactive visualization was developed using the Plotly library, incorporating distinct colour schemes.

Figure 17 displays the attractor plot proposed by Lorenz, which exhibits the complicated interaction associated with female genital mutilation. The exceptional performance characteristics of the Bokeh Library are evident in this observation. Bokeh exhibits the ability to generate sophisticated interactive diagrams with high-quality graphical attributes, hence surpassing other comparable tools. The investigation incorporates the use of Plotly, Seaborn, and Matplotlib libraries, which employ more intricate syntax.

Figure 18 presents a circular interactive diagram that showcases the eight countries engaged in the practice of female genital mutilation. The diagram is accompanied by a visually appealing graphical component. This further reinforces the adaptability of the Bokeh library in generating exceptional visualization capabilities compared to the Plotly library. Plotly is known for its ability to generate complicated diagrams, although it is generally less compatible with Pandas compared to the alternative. While it is true that the use of this particular tool enables the creation of intricate visual representations, its interactive capabilities are somewhat restricted and its usage necessitates the mastery of intricate syntax.

Figure 19 presents a Lollipop chart, which serves as an interactive visual representation of eight nations engaging in the practice of female genital mutilation. The Bokeh library exhibits a high level of flexibility and interactivity, although it does possess certain limitations in terms of the extent of interactivity it offers. Additionally, the utilization of Bokeh entails the employment of more intricate syntax.

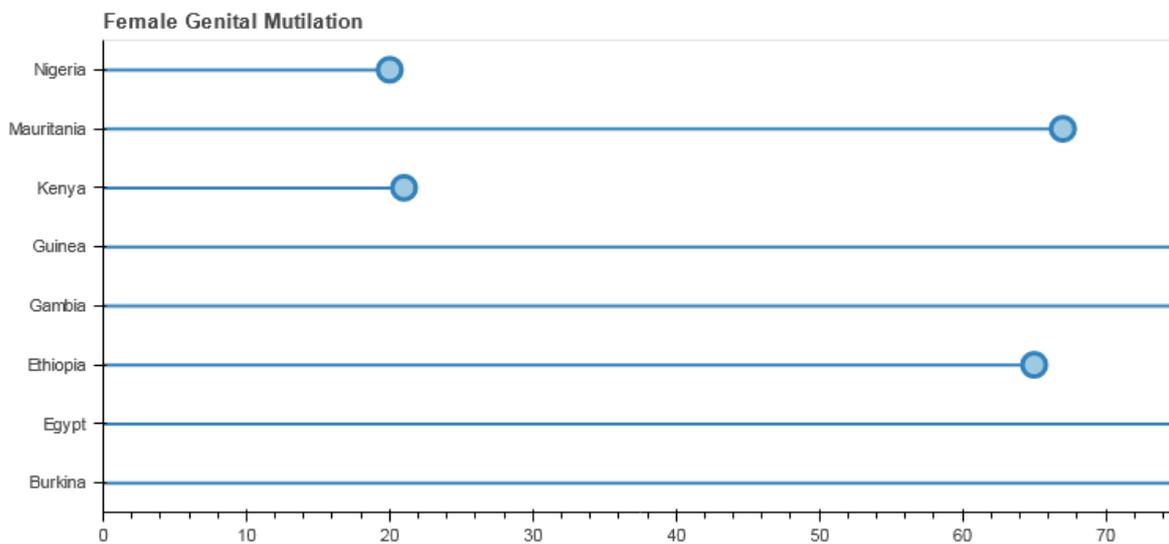


Figure 19. Plotting Lollipop Chart with Bokeh

4.2 Discussion of Findings

The major strengths and demerits of Python libraries can be inferred by conducting a comparative analysis of complicated interactive diagrams derived from the visualization output.

Figures 1, 6, and 13 depict the pivotal function of Pandas in data manipulation, whereas Plotly emerges as a highly effective tool for data visualization. Data manipulation tasks cannot be performed solely with Plotly; however, when paired with Pandas, Plotly can be utilized for such processes. This, however, demonstrates the exceptional performance of Plotly compared to other Python libraries utilized in the research study.

Figure 4 demonstrates that the utilization of many Python libraries, such as Plotly, Seaborn, and Matplotlib, can yield a proficient interactive figure. In contrast, it is worth noting that Seaborn exhibits certain limitations in terms of customisation, prompting its integration with Matplotlib to achieve enhanced customization capabilities.

Figure 5 illustrates the distinctive capabilities of Seaborn in the context of scatter plot visualization, as it enables the exploration of interactions among a larger number of variables compared to Matplotlib. Additionally, Pandas prove to be highly valuable in this interactive diagram for effectively altering the dataset.

Figure 7 demonstrates the significance of NumPy in the manipulation of an array of numbers, highlighting a constraint of Pandas. Additionally, Matplotlib is responsible for generating the distinct customization observed in Figure 7.

Nevertheless, the diagrams depicted in Figures 17 to 18 exhibit a heightened level of complexity, as they employ the utilization of Bokeh, a powerful interactive visualization library. Figures 17 and 18 illustrate the versatility of Bokeh in generating exceptional visualizations, showcasing its attractive graphic features. However, it is worth noting that Bokeh employs a more intricate vocabulary for data visualization, in contrast to Plotly, which offers a simpler syntax while still achieving remarkable visual outputs.

5. Conclusion and recommendations

The process of generating a comprehensive interaction diagram, which was visually presented, entailed the utilization of several Python libraries including Pandas, NumPy, Seaborn, Matplotlib, Plotly, and Bokeh. Pandas provides users with a wide array of unique functions and structures for data manipulation and analysis, particularly when combined with additional Python libraries like as Seaborn, Matplotlib, Plotly, and Bokeh. This enables a heightened rate of visual perception. The utilization of many Python packages has enabled the development of advanced interactive diagrams for data visualization. Various diagrams exemplify the concepts being discussed, including the Basic Funnel, Density Heatmaps, Hexbins plots, Multiple Bar charts, 3D graphs, Scatter Plot matrix, heat map, Lorenz Attractor plot, and circle diagram.

On the other hand, Matplotlib is a well-acknowledged and user-friendly Python library. The scripts of Matplotlib exhibit a clear and concise structure, making them highly advantageous for persons with diverse degrees of skill in Python programming. Furthermore, this library has remarkable flexibility, allowing users to effortlessly customize its functionalities to meet their individual needs.

It is worth noting that Plotly distinguishes itself as the most user-friendly choice among competing libraries, while also providing exceptional performance capabilities. The utilization of NumPy and Matplotlib in the construction of Hexbins plots facilitates the production of N-dimensional charts through the computational capabilities of NumPy, while Matplotlib enhances the visual aspect by offering customisable and aesthetically pleasing colour schemes for the plot. The Seaborn Library is widely recognized for its proficiency in both data visualization and statistical modelling. Nevertheless, it is important to acknowledge that the library offers a restricted range of choices for customization. The utilization of the Seaborn library in conjunction with Matplotlib is frequently employed to improve the visual quality of data representations, as previously noted.

Moreover, it has been demonstrated that Matplotlib is very appropriate for the fundamental data visualization patterns within the realm of data science, rendering it a vital instrument for both novice and proficient data scientists.

Furthermore, it is worth noting that Plotly has exhibited outstanding performance in comparison to other Python libraries employed in the present study. The system has effectively produced complex interactive diagrams, such as a heatmap that visually represents the interplay among the eight nations affected by female genital mutilation, a density heat map, and a coloured graphic portraying the flag of the United States.

The Python package known as Bokeh provides a wide range of visualization capabilities, making it a highly versatile and visually appealing tool.

In comparison to Plotly's employment of a simplified syntax, Bokeh provides enhanced flexibility in terms of its visual design elements and capacity to build complex data representations. Nevertheless, Bokeh demonstrates a reduced degree of interaction and requires the employment of a more complex vocabulary for its execution.

Considering the ramifications of this study, it is recommended that this research be regarded as a significant asset for persons across different levels of proficiency, encompassing both neophytes and seasoned practitioners. The utilization of this information by the government in the formulation of policies targeting the eradication of female genital mutilation may provide advantageous outcomes. This will enhance their comprehension of the significance of incorporating various Python modules to construct complex interactive visuals for data representation within the field of data science. The aforementioned content will be made available through online channels, thereby acting as a vital contribution to the current body of knowledge.

References

- [1] Ejoh, A., 2022. [online] View.officeapps.live.com. Available at: <<https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fmi->

- linux.wlv.ac.uk%2F~2056217%2FMy%2520Project%2520Work.docx&wdOrigin=BROWSELINK> [Accessed 15 June 2022].
- [2] Li, S. (2021). "Leap from Matplotlib to Plotly: Python libraries for beginners in data science". 43(7): 21-24.
- [3] Mishra, A. (2019). Data Visualization with Python. Machine learning in the AWS cloud: pp.51-78.
- [4] Psallidas et al, F. (2019). Data science through the looking glass and what is found there.
- [5] Stancin, I & Jovic, A. (2019). An overview and comparison of free Python libraries for data mining and big data analysis. PP 977-982.
- [6] Waskom, M. (2018). "seaborn: statistical data visualization", seaborn 0.8.1 documentation. [Online]. Available: <https://seaborn.pydata.org/>
- [7] Oberoi, A & Chauhan, R. (2019). Virtualizing data using Matplotlib and Seaborn libraries in Python for data science.
- [8] Odegua, O. (2020). "A python-based library for easy data analysis, visualization and modelling".
- [9] Pandian, S. (2021). "Effective data visualization techniques in data science using Python".
- [10] Raschka, S. (2018). "MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack", The Journal of Open Source Software, 24(3):55-58.
- [11] Rastogi, K. (2021). Guide for data visualization with Bokeh python library. Data science Blogathon, September 2021.
- [12] Shangjia, Yunhan, Z & Songlin, C. (2021). Research on python data visualization technology. Journal of physics: Conference series.1757 012122.
- [13] Sindhu, G.R. (2018). "Python as a key for data science". International Journal of computer science engineering. 6(4): 325-328.
- [14] Sehgal, O. (2019). "Visualizing data using Seaborn in python for data science". 12(9):9609.
- [15] Tam, V. N., & Thien, D. N. (2020). Real-time data visualisation using open source platform Thingsboard. Journal of Science , 14(3), 89-97.
- [16] UNICEF. (2019). Female genital mutilation/cutting: What might the future hold? New York, USA, pp. 1–6. World Health Organization. (2019). Female genital mutilation: Key facts. <https://www.who.int/news-room/fact-sheets/detail/female-genital-mutilation>.
- [17] Wes McKinney (2017). "pandas: A Foundational Python Library for Data Analysis and Statistics". Retrieved 2 August 2018.
- [18] World Health Organization. (2019). Female genital mutilation and other harmful practices: Prevalence of FGM.