

Alex Mutebe
MSc. Data Science
Dashboard Report
24th –July- 2023

Report on a data visualization dashboard for a study on
Immunization and vaccination for Medical Research Council
Research Uganda

1 CONTENTS

2	Introduction	3
2.1	Purpose	3
2.2	Objectives.....	3
3	Methodology.....	4
3.1	PREPARING DATA.....	4
3.2	DEFINING USER REQUIREMENTS	4
3.3	Dashboard Design	4
3.4	Dashboard Development	5
3.4.1	Defined the audience	5
3.4.2	Data extraction.....	5
3.4.3	Data reprocessing:	5
3.4.4	Worksheets and Visualizations in Tableau:	5
3.4.5	Selected appropriate chart types as in below	5
3.4.6	Dashboard layout:.....	6
3.4.7	The dashboard's interactivity:.....	6
4	Conclusion.....	6
4.1	Limitation	6
4.2	Recommendation.....	6

2 INTRODUCTION

If there is insufficient oversight and reporting, the collection of data in clinical trials may be flawed, which could result in inaccurate conclusions about the efficacy or safety of a certain treatment (Sven, 2020). My present job involves creating dynamic progress reports for the project leads and main investigators in the Medical Research Council Uganda unit where I manage clinical data. The lead researchers provide scientific direction both during and after the research initiative (Ellenberg et al., 2002). However, since these reports are created manually, it is necessary to use visualization software like Tableau to automate the reporting process for users to receive crucial data from the data visualization dashboard in real time (Tableau,2023).

According to Renato et al. (2018), a picture really is worth a thousand words when seeking to understand and draw conclusions from data. Data visualization, according to Sharvari (2021), uses maps, graphs, and data analytics to help users make informed decisions.

2.1 PURPOSE

The dashboard's goal is to provide principle investigators and other stakeholders with an intuitive and thorough user interface that enables them to monitor and assess participant enrolling and follow-up data.

2.2 OBJECTIVES

- A responsive dashboard that is interactive and has clear graphics.
- Incorporate key performance indicators (KPIs) important to the principal investigators and other users.
- Give users the ability to filter and drill down into specific information subsets.
- Display the data in an aesthetically appropriate and visually appealing way.

3 METHODOLOGY

3.1 PREPARING DATA

Data from the I-VAC research project was taken from the main data collecting system, REDCap, in order to fulfill the dashboard's purpose. Data was limited to 10 variables for this study's purposes and exported to an excel file, which is compatible with Tableau.

In compliance with the General Data Protection Regulations and Ugandan data protection law participant IDs were removed from the data during data preparation (GDPR, 2016) (The Republic of Uganda, 2019). To ensure consistency and correctness, the data was then cleaned i.e. duplicate data were deleted and transformed accordingly (Sven, 2020).

3.2 DEFINING USER REQUIREMENTS

The following techniques were utilized to gather user requirements and analyze them in order to improve the functionality of our dashboard:

- Domain specific requirements i.e. clinical research standards (Good Clinical Practice, General Data Protection Regulation)
- User experience in data management and reporting
- Interviewing users

3.3 DASHBOARD DESIGN

Based on the user requirements, we designed the dashboard layout and selected appropriate data visualization techniques including:

- Map visualization to show distribution of participant in the study sites. Lewis (2019) asserts that maps offer an eye-catching visual aesthetic that increases the user attractiveness of the dashboard.
- Bar charts to display distribution in categorical data, such as sex
- Boxplots to show the distribution of continuous data, such as age, sex groups, and location
- Line graphs to show trends in participants' follow-up over the course of the study.
- Data on the enrolment and retention rates of study participants.

3.4 DASHBOARD DEVELOPMENT

The dashboard was developed using tableau software, a visual analytics platform and is widely used by the business intelligence industry (Tableau, 2023). In order to make the dashboard interactive and visually appealing, the following was done.

3.4.1 Defined the audience

- Principle investigator. He/she oversees study procedures, including enrolling and obtaining consent from participants.
- Research assistants. They collect and submit data on research participants.

3.4.2 Data extraction

- Excel files are one of the many data sources that Tableau supports. Before importing the primary dataset into Tableau, ten variables were found and extracted from it in excel format. See appendix A
- The variables chosen indicate information needed to achieve the goals of the I-VAC dashboard. The following information is what was extracted.
 - demographic information (age, sex)
 - Status of participant enrollment (eligibility, consent)
 - visit status (visit period, visit date),
 - study status (study completion status, lost to follow-up)
 - geographical information (site location)

3.4.3 Data reprocessing:

- Data cleaning i.e. deleted duplicates
- Transformation i.e. appended country codes for generation the site coordinates

3.4.4 Worksheets and Visualizations in Tableau:

- Created worksheets for the following
 - Location of project sites
 - Participant Follow-up rates
 - Demographics data
 - Age distribution
 - Summary statistics

3.4.5 Selected appropriate chart types as in below

- Map for the site location
- Bar charts that display distribution in categorical data, such as gender. Marc and Nils (2014) suggest using bar charts to represent counts by length since they are a very accurate visual encoding.
- Boxplots showing distribution in age, grouped by sex, site. Boxplots are useful in comparing different distribution groups (Mike ,2021).
- Line graphs to visualize the trend in participant follow-up. Paul & Ghulam (2020) argue that line charts are used to visualize data while examining trends and patterns in time series.

3.4.6 Dashboard layout:

- Containers were used to organize and control the layout of the charts from the worksheets. Enrolment data, or summary statistics, have been made available in the top left corner for quick visualization. See appendix B

3.4.7 The dashboard's interactivity:

filters, parameters, and actions have been added enable users to explore the data and change visualizations dynamically. Users can filter the data based on location and date of visit criteria. Additionally, drill-down capabilities allow users to explore specific data subsets in detail. See appendix C & D

4 CONCLUSION

The dashboard for data visualization in this study is intended to be a powerful tool for analyzing and interpreting complex data. By incorporating interactive visualizations, user-friendly features, the dashboard empowers stakeholders to make informed decisions.

4.1 LIMITATION

- limited access to the testing data
- The time for refining the dashboard design i.e. 14 days' trail period for Tableau.

4.2 RECOMMENDATION

- Principal investigators and project managers ought to look into adopting tableau as a tool for the study's data visualization.
- Additional data should be made accessible to add more functionality.

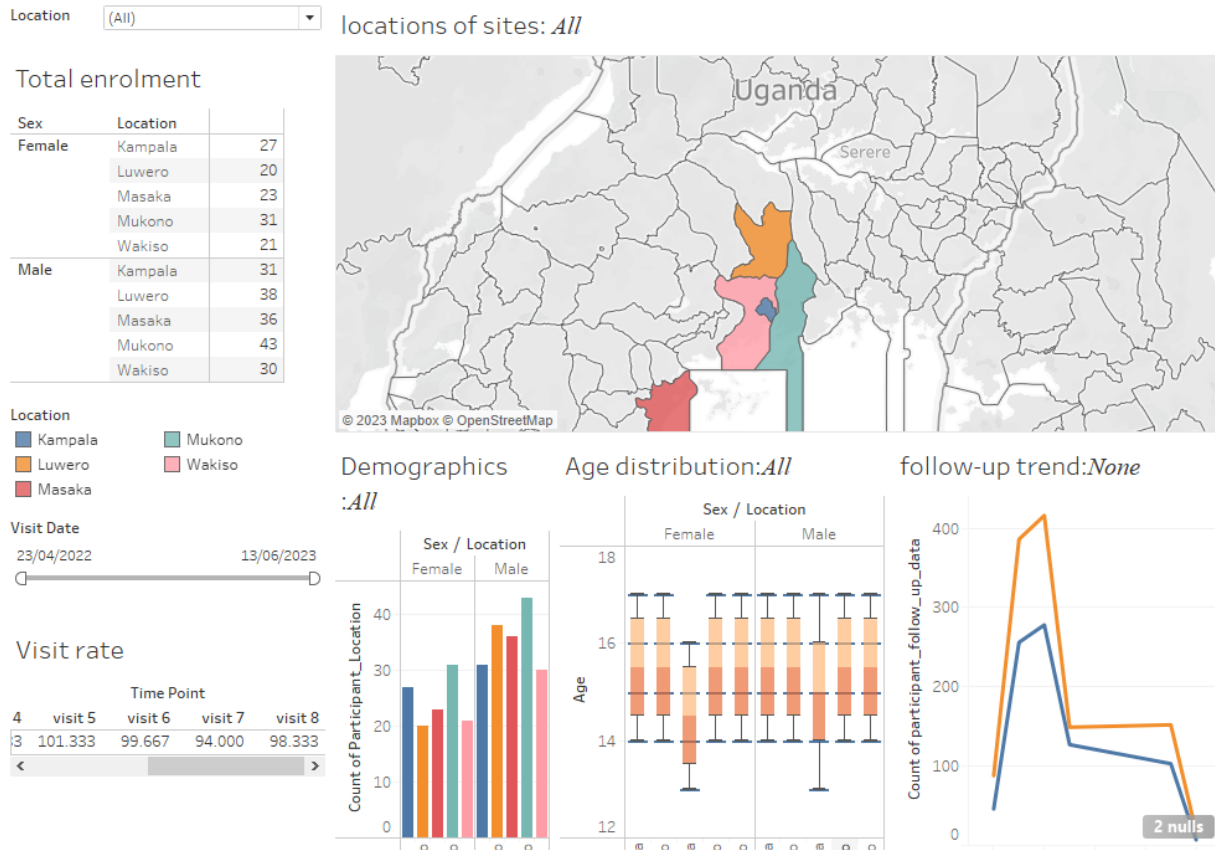
Appendix:

- Appendix A shows list of variables which were considered for our I-VAC dashboard

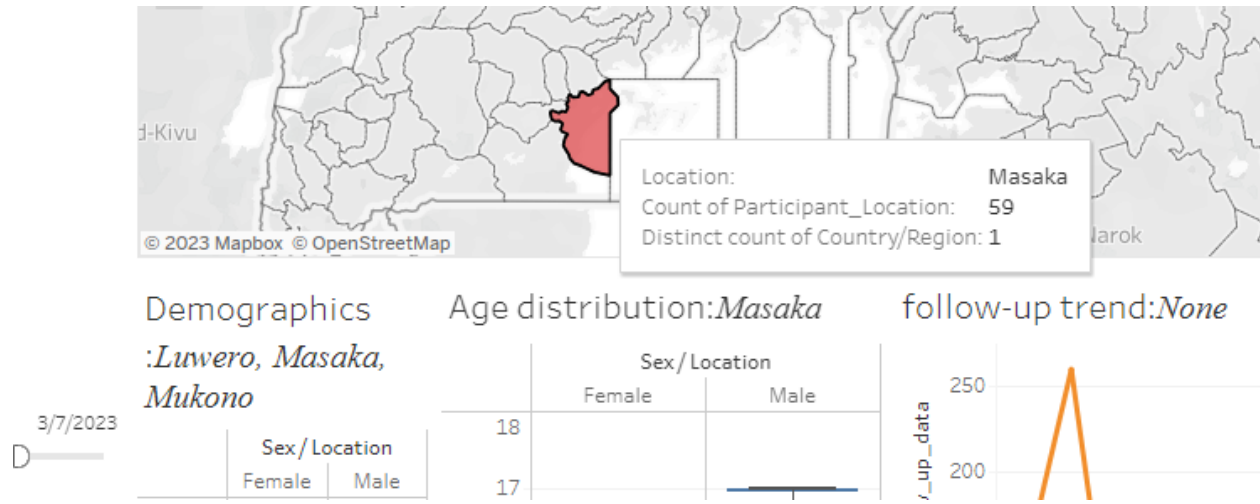
Variable Name	Question
participantID	Participant identification number
sex	Sex of participant
age	Participant's age
consented	Participant's consent status
eligible	Participant's eligibility status
Time_Point	Visit number
visit_date	Date of the visit/follow-up period
completed_study	Participant's study completion status
location	Study site
country	country
drop_out_reason	Lost to follow-up reason
entered_by	Initials of the person entering data

- Appendix B shows the I-VAC Dashboard layout

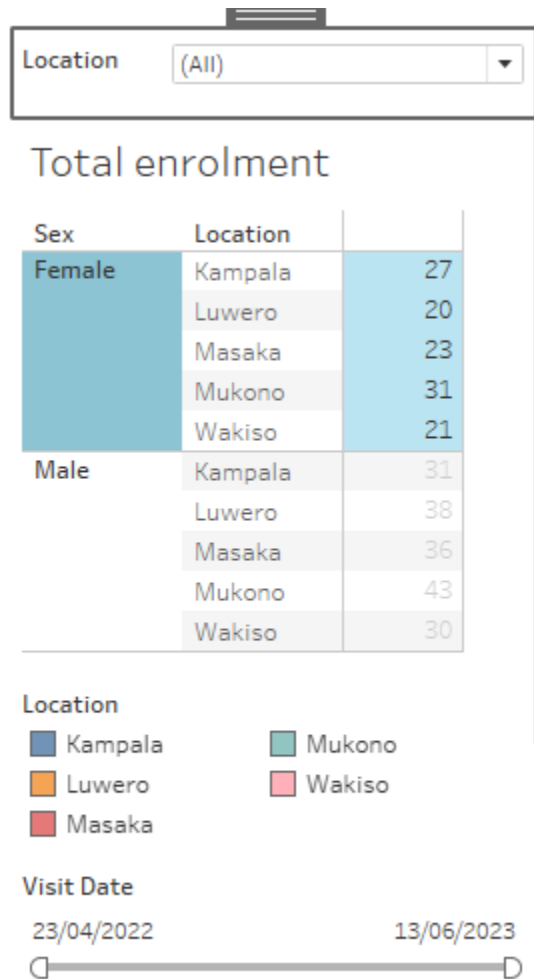
A data visualization dashboard for a study on immunization and vaccination for Medical Research Council Research Uganda.



- Appendix C shows drilled site and enrolment details



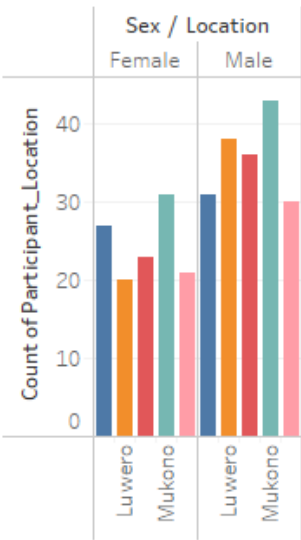
- Appendix D shows filtered data based on visit Date and geographical location



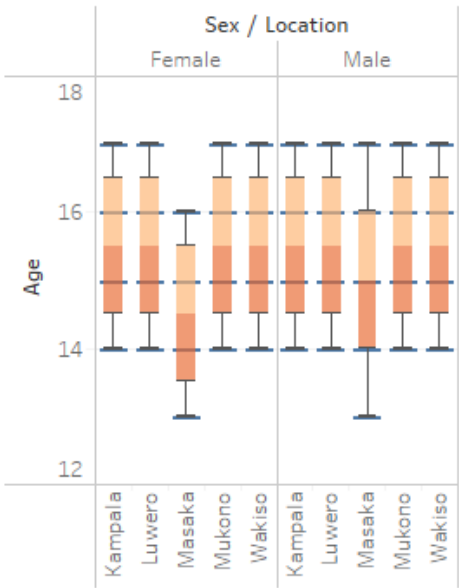
- Appendix E show bar and boxplot with distribution in the data, line chart shows participant follow-up rates

Demographics

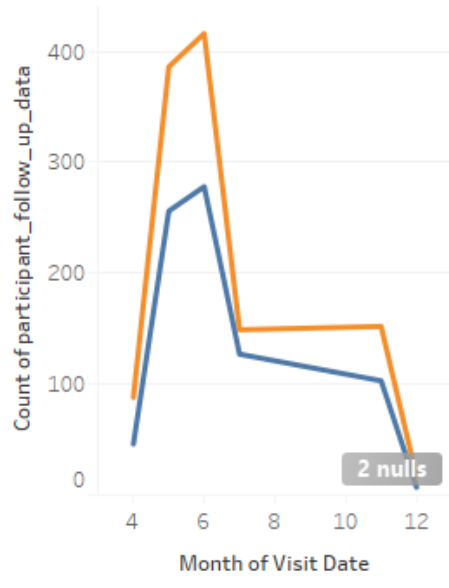
:All



Age distribution:All



follow-up trend:None



References:

Santosh. S. (2021) A Comprehensive Guide On Data Visualization In Python, Analytics Vidhya. Available from: <https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-data-visualization-in-python/> [Accessed on 12 July 2023]

Renato, T., Marias, M., Catarina, R., David, G. (2018) Data visualization techniques for real-time information — A custom and dynamic dashboard for analyzing surveys' results. Available from: <https://ieeexplore.ieee.org/abstract/document/8398641> [Accessed on 15 July 2023]

European Union General Data Protection Regulation (GDPR). (2016). Available from: <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1532348683434&uri=CELEX:02016R0679-20160504>

The Republic of Uganda. (2019). The Data Protection and Privacy Act, 2019. Available from: <https://ict.go.ug/wp-content/uploads/2019/03/Data-Protection-and-Privacy-Act-2019.pdf> [Accessed 10 July 2023]

Sven, B. (2020). Data Strategy: Good Data vs. Bad Data. Towards Data Science from <https://www.datasciencecentral.com/profiles/blogs/no-data-is-better-than-bad-data>

Tableau (2023). Tableau Desktop. Available from <https://www.tableau.com/why-tableau/what-is-tableau>

Mike, Y. (2021) A Complete Guide to Box Plots. CHARTIO. Available from: <https://chartio.com/learn/charts/box-plot-complete-guide/> [Accessed on 27 June 2023]

Marc, S., Nils, G. (2014) Bar charts and box plots: creating a simple yet effective plot requires an understanding of data and tasks. NATURE METHODS. Available from <https://go.gale.com/ps/i.do?id=GALE%7CA361242512&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=15487091&p=HRCA&sw=w&userGroupName=anon%7Ed539b7cd&aty=open+web+entry>

[Accessed 20 Jul 2023]

Paul, R, Ghulam, J. (2020) LineSmooth: An Analytical Framework for Evaluating the Effectiveness of Smoothing Techniques on Line Charts. IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 2

Lewis, C. (2019) Top 10 Map Types in Data Visualization. Toward Data Science. Available from: <https://towardsdatascience.com/top-10-map-types-in-data-visualization-b3a80898ea70>

[Accessed on 20 July 2023]

Ellenberg, S. S., Fleming, T. R., & DeMets, D. L. (2002). Data monitoring committees in clinical trials: a practical perspective. John Wiley & Sons.