WhatsApp Data Analysis with R

By

CODE NAME: WDAR
DEPARTMENT OF NETWORKS SCHOOL OF COMPUTING AND INFORMATICS TECHNOLOGY

A Concept Paper submitted to the School of Computing and Informatics Technology For the Study Leading to a Project Proposal in Partial Fulfillment of the Requirements for the Award of the Degree of Bachelor of Of Makerere University.

Supervisor:

Mr. Noah
Department of Networks School of Computing and Informatics Technology, Makerere University , +256-758-462841

Date: June 11, 2018

Group Members:

Name	**Name**	**Registration number**	**Student number**
1	Aaku Moses	13/U/2601/EVE	213005267
2	Mutwalib Bob	14/U/10203/EVE	214008146
3	Nyombi Nicholas	14/u/13679/EVE	214004411

1 Introduction

Due to various factors such as ease of use, essential features, the usage of WhatsApp accelerates. WhatsApp user base increases by day, due to the users ability to interact with others through audio and video calling, texting and transferring media as well as group chat. The objective of this paper is to analyze WhatsApp group and chat data between two individuals.

2 Background to the Problem

WhatsApp is a chat platform that supports one to one communication and one to many communication (this is also known as a group based communication.) This communication only occurs among individuals that have joined the chat platform. The platform can allow both synchronous and asynchronous communication. Communication on this platform can occur through a number of ways and these include video and audio calling, texting (accompanied by visual facial and other representations that are known as emojis), voice notes.

3 Problem Statement

The problem this project will address is WhatsApp Data Analysis thereby predicting the level of addiction of an individual to the WhatsApp group as per the age group and gender with the help of R statistics software programme.

4 Objectives

4.1 Main Objective

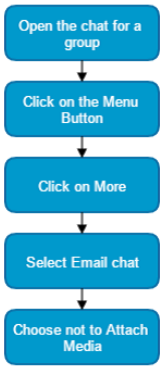
- The main aim is to analyze emotions of people with in a WhatsApp group and as well as emotions in a chat between two particular individuals and this is done using R statistics software programme.

4.2 Other Objectives

- To find what type of communication medium people prefer most in a WhatsApp group chat.
- To find the most active day or time of the week
- To find which group participants are more active and the number of messages they send per month, day and hour.
- The total number of messages sent per Timestamp.

5 Methodology

- **Data collection**
  Accurate data collection is essential in order to ensure integrity of the research. In this case the data that will be used is a copy of the history of a group chat that will be extracted using the Email chat feature



  A .txt document of your chat history will thus be attached and an email will be composed to the specified sender

- **Setting up the machine**
  An integrated development environment (IDE) R Studio will be used since it is the most popular IDE for R and supports debugging, workspace management, plotting and much more. Rstudio has windows such as the source pane where you can write and edit your R programs and documents, the Console pane where results are displayed, the workspace pane that allows quick access to additional tools and the plots pane that allows access to browse files plots and packages.

- **Data loading**
  Data loading stage includes importing resultant csv file into Rstudio. [1]

- **Exploratory Data Analysis (EDA) and Visualization**
  EDA is an approach to data analysis for summarizing and visualizing[2] the important characteristics of a data set in order to come up with a conclusion.[3]

6 Outcomes

A generated form of quantitative data from a group chat that will be used for a quantitative analysis of peoples emotions with a particular WhatsApp group based on the history chat. This quantitative analysis will be in form of graphical representations that will be used to view the different objectives such as individual emotions, the number of messages sent per given time and the type of communication mostly used in the group chat based on the extracted chat history.

References

[1] K. Willems, "This r data import tutorial is everything you need," jul 2015.

[2] Wikipedia, "Data analysis," jun 2018.

[3] Jovial, "Exploratory data analysis with r," jul 2014.