PROJECT REPORT

For

PRODUCT REVIEW ANALYSIS PROJECT

Version 1.0

Prepared by AYESIGA TONY NSUBUGA, NUWASIIMA BRENDA, ABUBAKARI SIMBA AND KIGGUNDU ISMAIL SSALI

GROUP 17

July 25, 2018

**Table of Contents**

**[1.](#_Toc520303647)****[Introduction](#_Toc520303647)** [3](#_Toc520303647)

[1.1 **Purpose** 3](#_Toc520303648)

[1.2 **Intended Audience** 3](#_Toc520303649)

[1.3 **Scope** 3](#_Toc520303650)

[1.4 **Definitions and acronyms** 4](#_Toc520303651)

[1.4.1 Keyword Definition 4](#_Toc520303652)

[1.4.2 Acronyms and Abbreviations 4](#_Toc520303653)

[1.5 **References** 4](#_Toc520303654)

[**2.** **Background and Objectives** 4](#_Toc520303655)

[**3.** **Organization** 4](#_Toc520303656)

[3.1 **Project Manager** 4](#_Toc520303657)

[3.2 **Project Group** 5](#_Toc520303658)

[3.3 **Supervisor** 5](#_Toc520303659)

[**4.** **Milestones** 5](#_Toc520303660)

[**5.** **Project Results** 5](#_Toc520303661)

[5.1 Requirements 5](#_Toc520303662)

[5.1.1 Requirements Compliance Matrix 5](#_Toc520303663)

[5.1.2 Requirements Compliance Summary 6](#_Toc520303664)

[**6.** **Project Experiences** 6](#_Toc520303665)

[6.1 **Positive Experiences** 6](#_Toc520303666)

[6.2 **Improvement Possibilities** 6](#_Toc520303667)

[**7.** **Financials** 6](#_Toc520303668)

[7.1 **Project Cost summary** 6](#_Toc520303669)

# **Introduction**

R Project is the project undertaken as a mandatory requirement for the course “**Professional Software Engineering Mini Practical Project II**” that is being conducted mutually by College of Computing and Information Sciences at Makerere University. The aim of the course is to provide a holistic hands-on experience in building software products using R programming language in a team environment. In this course we have to develop an R Product Review Analysis system. We are working as a team to provide this project in time.

## **Purpose**

The document contains the overall project description. It includes where, when and what we did and the respective experiences gained throughout the Recess Term. Whole description of the designing phase and cost is included in the document.

## **Intended Audience**

The intended audiences for this report include;

Mostly this document is intended for the development team of PRAS which include Programmers, Architects, System testers, System Analysts etc.

Other groups of people that might need to read the document include;

* Customer representatives
* User Support Agents.
* Marketing staff
* Makerere Teaching staff

## **Scope**

The project is aimed to provide an environment for analyzing, visualizing text data. In analyzing, the software provides a sentimental analysis of the text provided from a csv file, it also carries out text mining which shows a word cloud and word frequency. Also analyses Grammatical correctness of the text e.g. Punctuation and Spelling.

It includes visual forms of the analyzed data for example Bar plots, Histograms, ggplots etc. We have incorporated this in a user friendly interface for Data scientists.

## **Definitions and acronyms**

### Keyword Definition

|  |  |
| --- | --- |
| **Key words** | **Definition** |
| Histogram | A graph that shows the frequency distribution of a variable. |
| Bar plot | A graph is a chart or graph that presents [categorical data](https://en.wikipedia.org/wiki/Categorical_variable) with [rectangular](https://en.wikipedia.org/wiki/Rectangle) bars with [heights](https://en.wikipedia.org/wiki/Height) or [lengths](https://en.wikipedia.org/wiki/Length) proportional to the values that they represent. |
| ggplot | A system for declaratively creating graphics, based on The Grammar of Graphics. |
| Word cloud | An image composed of words used in a particular text or subject, in which the size of each word indicates its frequency or importance. |
| Sentiment Analysis | The process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine whether the writer's attitude towards a particular topic, product, etc. is positive, negative, or neutral. |
| Text Mining | Process of deriving high-quality information from text. |
| Data scientist | New breed of analytical data expert who have the technical skills to solve complex problems |

|  |  |
| --- | --- |
| **Acronym** | **Definition** |
| PRAS | Product Review Analysis System |

### Acronyms and Abbreviations

## **References**

# **Background and Objectives**

This is a mandatory school project in which we were given a dataset in form of a csv file containing over 71045 reviews from 1000 different products provided by the Datafiniti’s Product Database and tasked to analyze how writing quality impacts Positive and Negative online product Reviews.

We have developed a software which analyses the data and provides a Computational Analytical solution.

# **Organization**

## **Project Manager**

KIGGUNDU ISMAIL SSALI is the manager of Group 17

## **Project Group**

|  |  |
| --- | --- |
| **Name** | **Responsibility** |
| NUWASIIMA BRENDA | Documentation, Analysis and Implementation, Designing |
| ABUBAKARI SIMBA | Documentation, Analysis and Implementation, Designing |
| KIGGUNDU ISMAIL SSALI | Project Manager, Documentation, Implementation, Designing |
| AYESIGA TONY NSUBUGA | Documentation, Implementation, Designing, Testing |

## **Supervisor**

NOAH KANGE was our supervisor.

# **Milestones**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID** | **Milestone Description** | **Finished Week** | | **Remarks** |
| **Start Date** | **End Date** |
| M-001 | Project Description, Plan and Documentation | 11-06-2018 | 20-06-2018 | Good |
| M-002 | Requirement, Specification Definitions and Documentation | 21-06-2018 | 27-06-2018 | Good |
| M-003 | Project Design and Documentation | 28-06-2018 | 04-07-2018 | Good |
| M-004 | Implementation | 05-07-2018 | 25-07-2018 | Excellent |
| M-005 | Final presentation and Delivery | 26-07-2018 | 26-07-2018 | Excellent |

# **Project Results**

## Requirements

### Requirements Compliance Matrix

|  |  |  |
| --- | --- | --- |
| **Id** | **Requirement Description** | **Completed** |
| PRA-1 | Upload file | Yes |
| PRA-1.1 | System user should be able to upload a csv file containing data to analyse. | Yes |
| PRA-2 | Data Analysis | Yes |
| PRA-2.1 | The system should be able to do a Grammar Analysis on the uploaded csv file. | Yes |
| PRA-2.1.1 | The system should be able to identify whether a review has Punctuation end mark e.g. full stop, question mark and Exclamation mark. | Yes |
| PRA-2.1.2 | The system should be able to identify misspelled words in a review and displays the results in table form. | Yes |
| PRA-2.2 | The system should be able to do word count on the data file and return the results in a nicely formatted table. | Yes |
| PRA-3 | Sentiment Analysis | Yes |
| PRA-3.1 | The system should be able to calculate eight types of emotions for example Anger, Joy, Disgust, Anticipation, Sadness, Surprise, Fear and Trust present within the uploaded data file. | Yes |
| PRA-3.2 | The system should be able to deduce the polarity of a given review i.e. if a review is Negative or Positive. | Yes |
| PRA-4 | The system should be able to be display results in Visual form. | Yes |
| PRA-5 | Rating Analysis | Yes |
| PRA-5.1 | The system should able to show the distribution of rating across products. | Yes |
| PRA-5.2 | The system should able to show a visual of frequency of words with spelling errors against rating. | Yes |

### Requirements Compliance Summary

|  |  |
| --- | --- |
| Total number of Requirements | 10 |
| Number of Requirements implemented | 10 |
| Requirements partially fulfilled | 0 |
| Requirements not fulfilled | 0 |
| Requirements dropped | 0 |

# **Project Experiences**

## **Positive Experiences**

The main experience we learn from this Recess Term is to work in group. We also learnt how to use Version control software such as GitHub and other new software such RStudio as the Integrated Development Environment.

## **Improvement Possibilities**

We have experience that if we organize our resources according to the requirements then we can make the project more successful.

# **Financials**

## **Project Cost summary**

|  |  |
| --- | --- |
| **Item** | **Cost** |
| Printing | Shs.25,000 |
| Research | Shs.50,000 |