**MAKERERE UNIVERSITY**

**COLLEGE OF COMPUTING AND INFORMATION SCIENCES**

**DEPARTMENT OF NETWORKS**

**BACHELOR OF SCIENCE IN SOFTWARE ENGINEERING (YEAR 2)**

**RECESS TERM 2**

**COURSE CODE: BSE 2301**

**PROPOSAL FOR YOUTUBE TRENDING STATISTICS**

**GROUP NUMBER: G-09**

|  |  |  |
| --- | --- | --- |
| NAME | REG.NO | STUDENT NO. |
| SEMPA ETHAN IAN | 16/U/11403/PS | 216004900 |
| KWESIGA PETER | 16/U/6411/EVE | 216002578 |
| NABAGIDDE GRACE DIANA | 16/U/8005/PS | 216012180 |
| BWIRE IVAN | 16/U/4522/PS |  |

**SUPERVISOR: NOAH KANGE**

**PROJECT LEADER: KWESIGA PETER**

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE SOFTWARE ENGINEERING RECESS PROJECT BSE 2301

**PROJECT REPORT**

**For**

**YouTube Trending Statistical System**

Version 1.0 approved

Prepared by Nabagidde Grace Diana, Kwesiga Peter, Sempa Ethan Ian, Bwire Ivan

23rd JULY 2018

Table of Contents

1. Introduction 4

1.1 Purpose of this document 4

1.2 Intended Audience 4

1.3 Scope 4

1.4 Definitions and acronyms 4

2. Background and Objectives 4

2.1 Project Manager 4

2.2 Project Group 5

2.3 Steering Group 5

2.4 Customer 5

3. Milestones 5

4. Project Result 6

4.1 Requirement Compliance Matrix 6

4.2 Requirements Compliance Summary 6

4.3 Work Products and Deliverables 7

5. Project Experiences 7

5.1 Positive Experiences 7

5.2 Improvement Possibilities 7

6. Financials 7

6.1 Project Cost Summary 7

LIST OF TABLES

[Table 1: Definitions and acronyms 4](#_Toc520173917)

[Table 2: Project group 5](#_Toc520173918)

[Table 3: Milestones 5](#_Toc520173919)

[Table 4:Requirements Compliance Matrix 6](#_Toc520173920)

[Table 5:Requirements Compliance Summary 6](#_Toc520173921)

[Table 6: Work products and deliverables 7](#_Toc520173922)

[Table 7:Cost summary 7](#_Toc520173923)

# 1. Introduction

In this project we are tasked to use the R programming language to carry out data analysis. In this case we were given a dataset containing the top trending YouTube videos from five regions hence come up with a YouTube Trending Statistics System.

## Purpose of this document

In this document there is overall description of the YouTube Trending Statistics System. It includes what we did. There is also description of the work experience gain from this project. Whole description of the designing face and cost is included in this document.

## Intended Audience

The intended audiences in this case were the YouTube management team and the channel owners on YouTube though our system is open and can be accessed by anyone interested in this particular data set.

## Scope

Based on given datasets for the USA, Great Britain, Germany, Canada and France regions with up to 200 trending videos daily, we are meant to develop a system. The purpose of this system is to classify opinions through the number of comments, likes and dislikes hence providing feedback to the intended audiences and also carry out statistical analysis on the feedback.

## Definitions and acronyms

|  |  |
| --- | --- |
| **Keyword** | **Definitions** |
| Data set | A set of data to be analyzed |
| Requirement compliance matrix | Table that verifies all requirements have been met fully. |

Table 1: Definitions and acronyms

# Background and Objectives

The project is aimed at developing a system based on the given data set pertaining YouTube trending video statistics for regions France, Germany, Canada, USA and Great Britain. Our system analyzes and visualizes the data provided in order to provide better statistical feedback to the viewers as well as the YouTube management team.

Here we have three main actors in the project i.e. Supervisor, Project leader and Project members who have played different roles in making this project successful.

## Project Manager

Kwesiga Peter is the Manager of the group.

## Project Group

|  |  |
| --- | --- |
| **Name** | **Responsibility** |
| Kwesiga Peter | Project manager, Analysis, Implementation, Designing |
| Sempa Ethan Ian | Implementation, Documentation, Analysis |
| Nabagidde Grace Diana | Implementation, Documentation, Designing, Analysis |
| Bwire Ivan | Analysis, Documentation |

Table 2: Project group

## Steering Group

 Mbabazi Isaac & Nkange Noah

## Customer

YouTube Management Team

Channel owners

# Milestones

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Id** | **Milestone Description** | **Finished week** | | | | **Metr** | **Remarks** |
| **Plan** | **Forecast** | | **Actual** |
| **Week** | **+/-** |
| M-001 | Analysis of the project | 19-05-18 | 0 | 0 | 19-05-18 | Y | Good |
| M-002 | Coming up with a concept paper | 25-05-18 | 1 | 0 | 25-05-18 | Y | Good |
| M-003 | Coming up with a System Requirement specification (SRS) document | 03-06-18 | 2 | 0 | 03-06-18 | Y | Good |
| M-004 | Coming up with a Software Design Document (SDD) | 10-06-18 | 3 | 0 | 10-06-18 | Y | Good |
| M-005 | Project Implementation | 23-07-18 | 4 | 0 | 23-07-18 | Y | Good |
| M-006 | Final Presentation & delivery | 25-07-18 | 4 | 0 | 25-07-18 | Y | Excellent |

Table 3: Milestones

# Project Result

## Requirement Compliance Matrix

|  |  |  |
| --- | --- | --- |
| **Id** | **Requirement Description** | **Completed** |
| REQ-1 | Popularity analysis | Yes |
| REQ-1.1 | System should extract videos above 5000 likes, convert to %age and plot a graph. | Yes |
| REQ-1.2 | System should be able to show that the more you post, the more your videos will gain popularity and visualize. | Yes |
| REQ-2 | Time analysis | Yes |
| REQ-2.1. | The system should be able to calculate difference between trending date an publish time in days and plot a graph. | Yes |
| REQ-2.2 | The system should be able to extract videos that have been trending less than 2 days and plot a graph. | Yes |
| REQ-3 | Sentiment Analysis | Yes |
| REQ-3.1 | The system should be able to calculate most liked and most disliked based on average likes or dislikes per category. | Yes |
| REQ-3.2 | The system should be able to determine the polarity of video opinion i.e. if a video is liked or disliked taking into consideration the ratio of likes to dislikes.. | Yes |
| REQ-4 | The system should be able to be display visualizations of statistics and relations through word cloud and scatter plots. | Yes |

Table 4:Requirements Compliance Matrix

## Requirements Compliance Summary

|  |  |
| --- | --- |
| Requirements | Total number |
| Number of Requirements implemented | 5 |
| Requirements partially fulfilled | 0 |
| Requirements not fulfilled | 0 |
| Requirements dropped | 0 |

Table 5:Requirements Compliance Summary

## Work Products and Deliverables

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **To** | **Output** | **Planned week** | **Promised week** | **Late +/-** | **Delivered week** | **Rem** |
| Noah | Concept Paper | **W25** | **W25** | No | **W25** |  |
| Noah | System Requirement Specifications Document | **W26** | **W26** | No | **W26** |  |
| Noah | Software Design Document | **W27** | **W27** | No | **W27** |  |
| Noah | Progress with project implementation | **W29** | **W29** | No | **W29** |  |
| Noah | Final Presentation & delivery | **W30** | **W30** | No | **W30** |  |

Table 6: Work products and deliverables

# Project Experiences

It has exposed us to working with given data sets and performing analysis on them. It also enabled us to familiarize ourselves with the use of version control system in this case **Git.**

## Positive Experiences

The main experience is that we learnt how to use the R programming language in order to carry out data analysis. In addition, it has broadened our horizons particularly documentation and encouraged team work.

## Improvement Possibilities

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

# Financials

## Project Cost Summary

|  |  |
| --- | --- |
| Planned cost of research and printing | 20000 UGSHS |
| Actual cost | 50000 UGSHS |

Table 7:Cost summary