

**COLLEGE OF COMPUTING AND INFORMATION SCIENCES**

**DEPARTMENT OF NETWORKS**

**BACHELOR OF SCIENCE IN SOFTWARE ENGINEERING (YEAR 2) RECESS TERM 2 (BSE2301)**

**GROUP 16**

**SYSTEM DESIGN DOCUMENT**

**PROJECT MEMBERS**

|  |  |  |  |
| --- | --- | --- | --- |
| **NAME** | **REG. NO** | **STD. NO** | **SIGNATURE** |
| KAGANZI TIMOTHY | 16/U/5134/EVE | 216007287 |  |
| KASOZI BENJAMIN | 16/U/ 5538/EVE | 216015602 |  |
| OKURUT JULIUS | 16/U/10950/PS | 216014096 |  |
| KUNYA NATHERN | 16/U/6372/EVE | 216006158 |  |

**PROJECT LEADER**

KAGANZI TIMOTHY

**SUPERVISOR**​: ISAAC MBABAZI

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

19​TH​ JULY, 2018

**TABLE OF CONTENTS**

[1. INTRODUCTION 2](#_Toc519776813)

[1.1 Purpose 2](#_Toc519776814)

[1.2 Scope 2](#_Toc519776815)

[1.3 Overview 3](#_Toc519776816)

[1.4 Definitions and Acronyms 4](#_Toc519776817)

[2. SYSTEM OVERVIEW 4](#_Toc519776818)

[3. SYSTEM ARCHITECTURE 5](#_Toc519776819)

[3.1 Architectural Design 5](#_Toc519776820)

[3.2 Decomposition Description 6](#_Toc519776821)

[3.3 Design Rationale 9](#_Toc519776822)

[4. DATA DESIGN 10](#_Toc519776823)

[4.1 Data Description 10](#_Toc519776824)

[5. COMPONENT DESIGN 10](#_Toc519776825)

[6. HUMAN INTERFACE DESIGN 11](#_Toc519776826)

[6.1 Overview of User Interface 11](#_Toc519776827)

[6.2 Screen Images 11](#_Toc519776828)

[6.1 Reference Material 16](#_Toc519776829)

# INTRODUCTION

## Purpose

This SDD describes the architecture and design of YSA application. The design description defined in this document is significant in the following ways;

* It will be used to assess the impact of Youtube Statistical analysis application on the Youtube channel owners.
* It describes the modular structure, data and diagrams involved with in the application.
* It identifies the required system resources.
* It will be used in case of carrying out maintenance activities on the application.

### **1.1.1 Its intended audience include:**

* YouTube data analysists and these have the highest priority over the others.
* Top level managers of an organization who are interested in layout of the project (Youtube channel owners).
* Software developers who are interested in understanding the design of the YSA application as well as make some changes in the future

.

## Scope

This Software Design Description (SDD) describes the detailed structure of the components of YouTube Statistics Analysis (YSA) and the precise implementation details required to satisfy the requirements as speciﬁed in the Software Requirements Speciﬁcation (SRS). It is assumed that the reader has read the SRS, since this document also deﬁnes the implementation details of the desired behavior given the requirements within it. This document will build heavily on the YSA and so knowledge of the general system architecture is recommended prior to commencing this document

YouTube data analysis software is one that people can use to analyze YouTube data. It contains shiny app which feature is R based and it will be helpful during the visualization stage using many different techniques. The software also has a data visualization feature that will be of use to the analyst to understand the relationships between the different variables under consideration.

This document covers the shiny web feature, which is the interface in this case, all users will have access to this app but after authentication, and the authorized user shall have full rights to the app and will be able to manipulate the data the way he/she wants

Upon completion of this software, these are some of the benefits that this software will provide to the end users.

### **1.2.1 Some of the benefits of this product are**

* Proper Market research.
* Easy Decision making.
* Provide a company with an edge over their competitors.
* Provide an insight of the user’s suggestions and reactions about the product or video.

### **1.2.2 Some of the goals of this software**

* The main goal of this software is to ease the analysis process for the YouTube analysts and those who are interested in data science particularly in the field of YouTube.
* The other goal is limiting the hardships involved during data analysis thereby saving time when visualizing the relationship between variables of interest.

## Overview

In this document, detailed design of the system with user interfaces will be described. In section 3; decomposition of the system with module. Decomposition, concurrent process decomposition and data decomposition is given, in section 4; Data design with data descriptions is given, in section 5; there is description of the component design. In section 6, there are user interfaces; screen objects, images and actions. In section 8; Requirements matrix is provided and finally in section 8; there is the appendix which has the abbreviations and their in full and the reference materials.

## Definitions and Acronyms

|  |
| --- |
| **SDD** – System Design Description/document. |
| **YSA** – YouTube Statistics Analysis. |
| **YouTube Analyst** – a person with prior knowledge about data analysis and the functionalities of the YouTube analysis. |
| **Developer** – a person who designs software for either commercial or individual purposes. |
| **Component** – a module that is part of the system. |

.

# SYSTEM OVERVIEW

YSA is a software being designed to ease the work of the person doing the analysis by bringing all the separate components into one application where he/she can easily load the required dataset that he wants to do analysis on and then choose the best visualization technique to use. These different components have been merged in one place and that is in the shiny app.

The user will have to choose from among the datasets with in the app or extract and load his own. This app has three basic components, which are going to be explained in detail.

### **2.1. Development Methods**

A software development methodology is a way of managing a software development project. This typically address issues like selecting features for inclusion in the current version, when software will be released, who works on what, and what testing is done.

For YouTube statistics analysis software, we shall use scrum, because, this method enables us (developers) to add features to our system in form of short sprints (usually 7-30 days), during our short frequent meetings thus keeping people focused. Tasks are usually tracked on a scrum board. The group is self-organizing and collaboratively managed, although there is a scrum master tasked with enforcing the rules and buffering the team from outside distractions.

We shall as well use the Unified Modeling Language (UML) for visualizing and documenting the systems design.

We considered using UML because it uses object-oriented design concepts, and it is independent of any specific programming language and can be used to describe business processes and requirements generally. This enables developers to use the same design to implement the system using various programming languages rather than being constrained to one language. [2]

# SYSTEM ARCHITECTURE

## Architectural Design

The design of this YSA application will follow the client/server architecture where by the client is represented by user Interface which is used to send requests to the server, which then services the requests of the client.



**Figure 3.1.1. Showing the architecture design of the system**

In this YSA application, the UI.r gets data from the user and sends it to the server. The server manipulates the data and sends the results in form of visual diagrams such as bar graphs, pie charts and word cloud which are displayed to the user on the user interface.

## Decomposition Description

This section decomposes each use-case feature into its data flow processes by examining its data flow diagram and process. These assist in determining the preliminary members and methods of the modules that need to be implemented, or the modifications to existing modules to implement the feature. This section includes the description of the intended design to meet the requirements. When appropriate, the sequence diagram will be expanded to include system requirements. This section also incorporates a decomposition diagram providing the segments involved in each process.

### 3.2.1 Context Diagram for YSA application



**Figure 3.2.1. Showing the context diagram**



**Figure 3.2.2. Showing the level 1 DFD for the system**

### **3.3.3 User Interface component**

The user interface enables the users to access different functions of the system. The user will be able to load any required dataset into the system which is to be visualized and analyzed to come up with what is trending and the different categories of videos.



**Figure 3.3.1. Showing the sequence diagram for the system**

1. **Analyze data**

First, the user loads the application and the home page is displayed,

The user then selects the analytical tool to use so as to analyze the data, then the output is displayed on the browser which the user can be ale to analyze and make predictions.

1. **Visualize data**

For visualizing data, the user loads the application, and then selects the analysis tool they may wish to use when the home page is displayed on the left-hand tab.

The visualized data is then displayed on the browser for analysis and predictions.

**A use case of the User Interface component.**



**Figure 3.3.2. Showing use case of the user interface component.**

## Design Rationale

Client/server architecture is preferred in designing YSA application because of the need to provide the user with an interactive interface. Most of the manipulations will be done in the server and this makes the user/operator’s interaction with the application easier since he/she will only have to send requests to which the server will respond to that request.

# DATA DESIGN

## Data Description

This YSA application uses Youtube data about viewers’ opinions about the Youtube videos. This data is obtained from the interaction between Youtube platform and the viewers which are stored in the Youtube database and later extracted using the Youtube api, saved as a csv file(s) for analysis using RStudio.

# COMPONENT DESIGN

Algorithms used in the different system modules

**Analyze data**

**Algorithm**

Select an analysis tool

Analyze data

**Visualize data**

**Algorithm**

Select an analysis tool

Analyze data

Visualize analyzed data

1. Choose the x and y variable in case of the need to create a bar graph

**Extract data component.**

The project utilizes the YouTube Data API that allows the shiny app to incorporate functions that are used by YouTube application to fetch and view information. To retrieve information from YouTube using their Data API our application needs to be authenticated. Once the application is authorized we can fetch data like comments that can be used to analyze and represent.



**Figure 5.1. Showing a use case, with sub components of the extraction subsystem.**

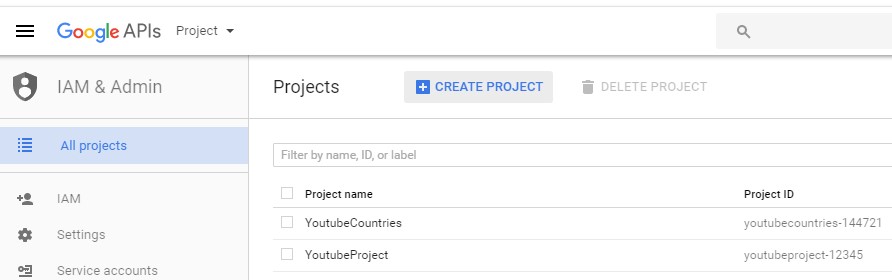
**Steps taken to get a YouTube API.**

**Step 1** Log into [https://developers.Google.com/](https://developers.google.com/) with existing credentials.

**Step 2** To create the unique API key for retrieving data, a new project needs to be created from the Google provided developer’s console.

**Step 3** Go to [https://console.developers.Google.com/project](https://console.developers.google.com/project)

**Step 4** Click create project.



**Step 5** A new project needs to be created. Provide a name for the project

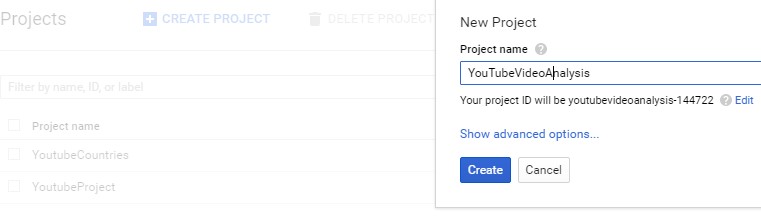


Figure 6 Providing a Name for Project

**Step 6**  To create a new API key Google provides the YouTube Data API that is available under the developer tools.

#### Library > YouTube APIs > YouTube Data API

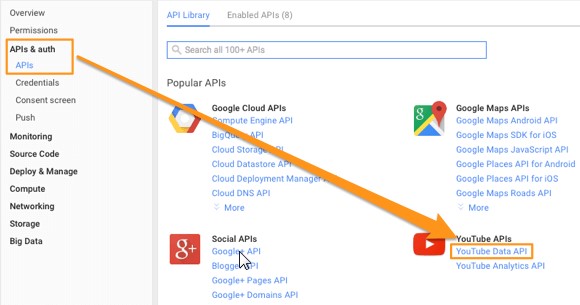


Figure 7 Getting the YouTube API Key [14]

**Step 7** To utilize the YouTube Data API, it needs to be enabled under the logged in credentials. Click “Enable” under the YouTube Data API v3. **Dashboard > YouTube Data API v3: Enable**

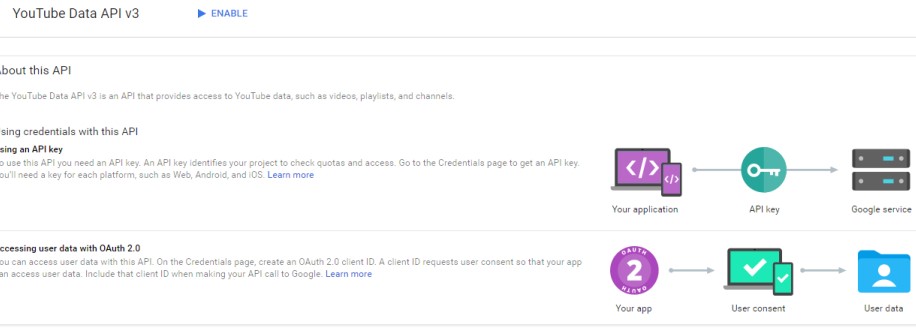


Figure 8 Enabling YouTube API Key

**Step 8** Once the YouTube data API is enabled, create credentials in order to utilize the API. To create credentials

#### Dashboard > Go to Credentials Button

**Step 9** Add credentials to the project. YouTube provides three options for creating an API Key.

* API key  client ID
* service account

After finishing the steps, you can be able to use the API key to extract comments of any video using its video id in the shiny app

# HUMAN INTERFACE DESIGN

## Overview of User Interface

The user interface is necessary for the users as it enables the use of the different functions of the system. The user will be able to interact with the system. The user will be able to load any required dataset into the system, which are to be visualized and analyzed to come up with various trends of videos and categories of videos.

## Screen Images

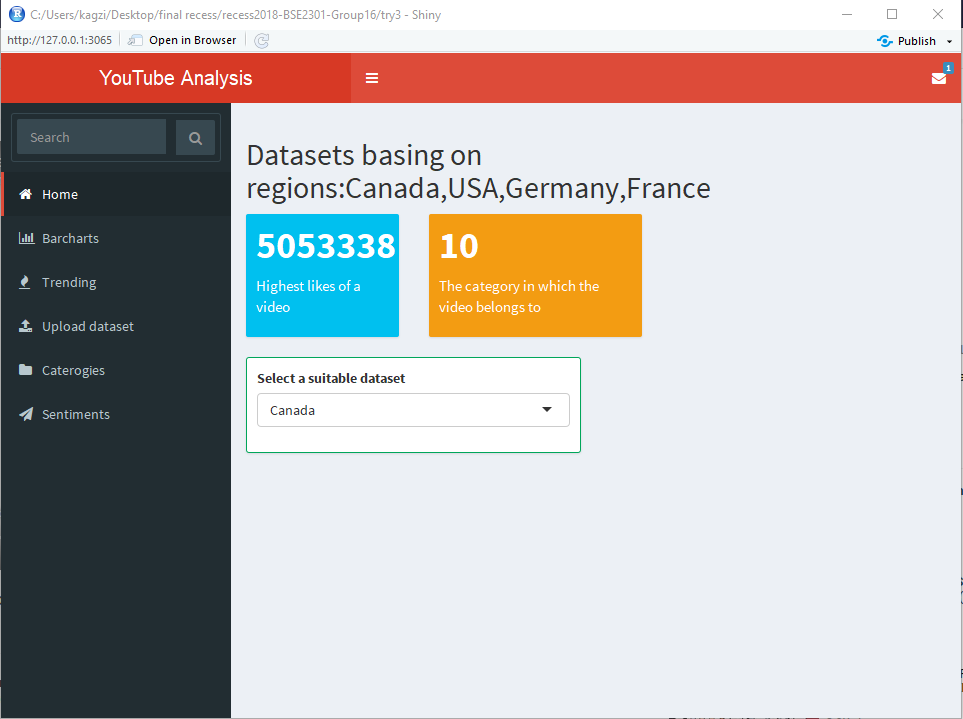
These screen images are user interfaces that provides a user a platform for them to perform different tasks with YSA application.

6.2.1 **Home page**

When the user loads YSA application, the page below will be shown on their browser.

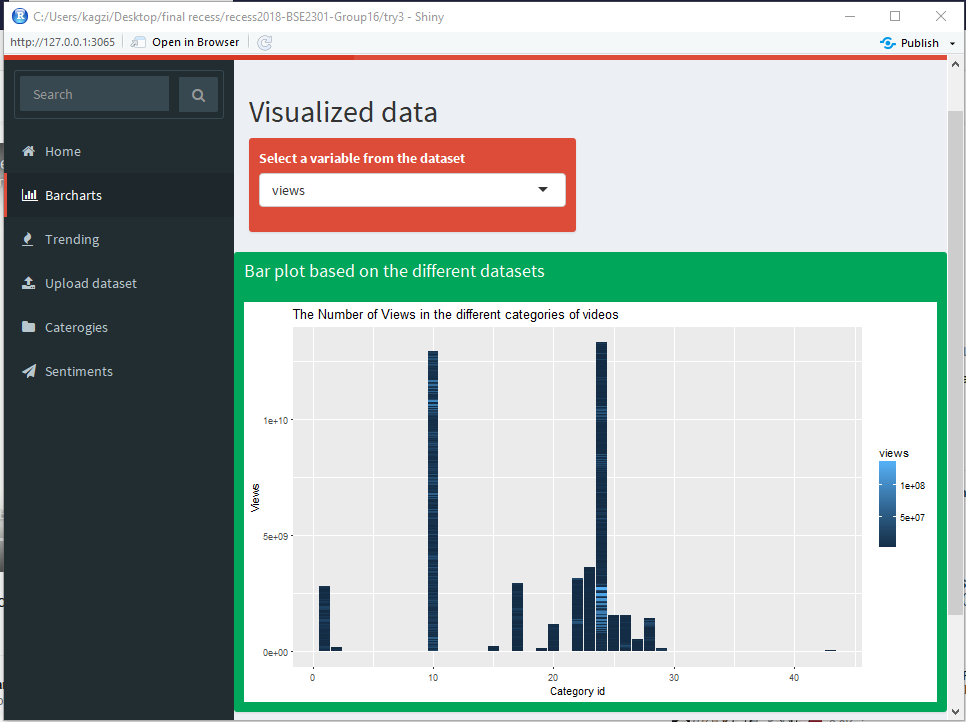
The left-hand tabs include the home page, bar charts for analysis, trending for determining trending videos, upload, for uploading personal dataset, category for the different categories of the videos and the sentiment, for identifying users’ reactions on the video

The right-hand has a select suitable dataset option, these include Canada, Great Britain, USA, Germany and France



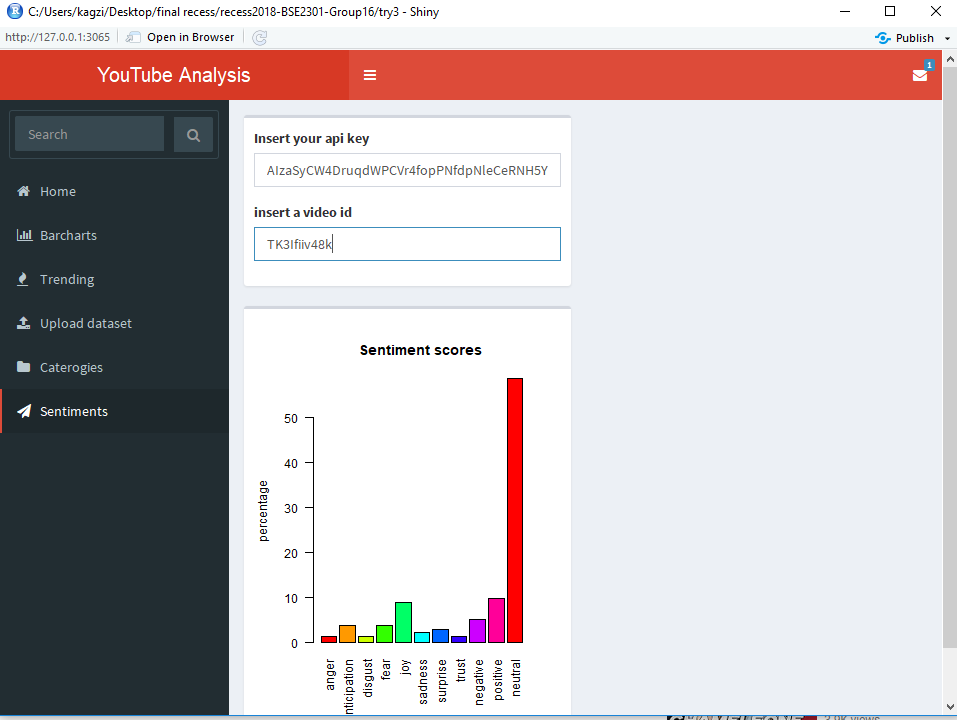
### 6.2.2. **Bar chats**

This tab is used for visualizing data by selecting the x and y variables, for this case, a bar plot of views against category ID as shown



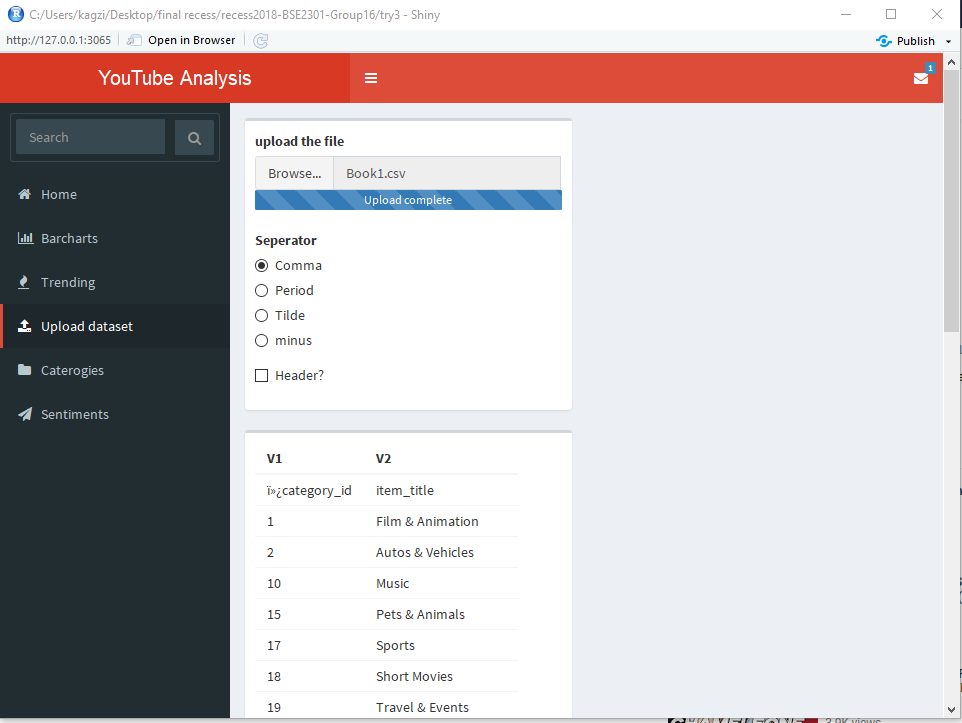
6.2.3. **Sentiment analysis**

In this section, the user enters the api key and the video they intend to carry sentiment analysis for. Sentiment analysis helps the analysis to determine the viewers’ reaction about the video as shown



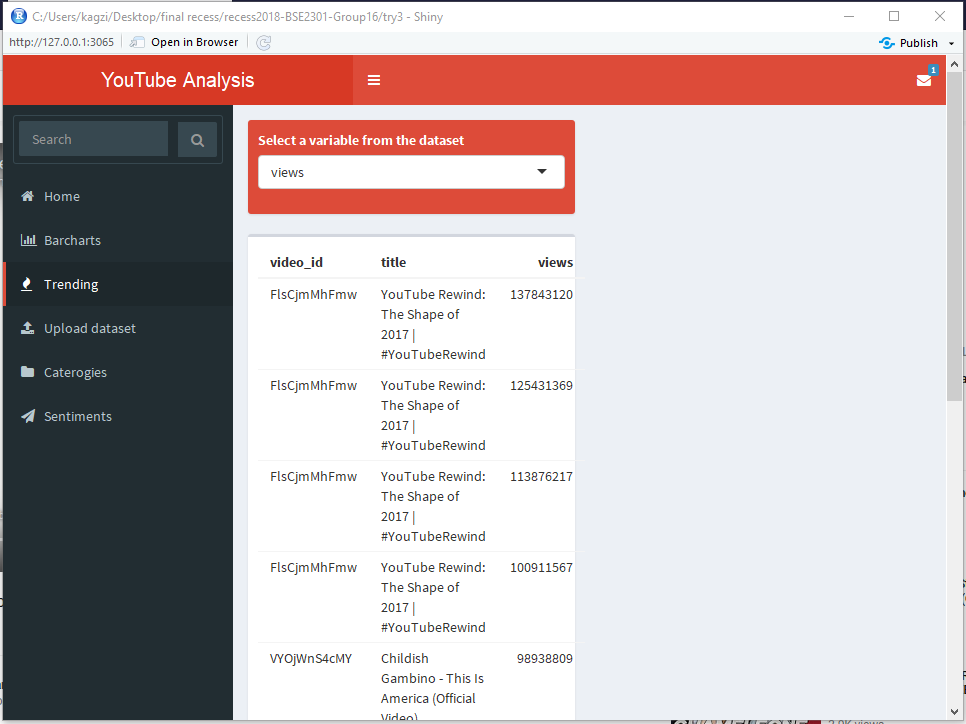
### 6.3.4 **Upload dataset**

Here, the user after selecting the upload option, they then click on browse which will ask them to select the location where their dataset is, after which its complete and its ready for analysis. The image below shows Book1.csv dataset and the data in contains.



### **6.3.5 Trending**

This is used to determine the most trending videos by using the video ID against views, likes or comments as shown below



## Reference Material

|  |  |
| --- | --- |
| [1] | [Online]. Available: <www.nada.kth.se/~karlm/prutt05/lectures/prutt05_lec7.pdf.> |
| [2] | [Online]. Available: <https://www.asc.edu/sites/default/files/org_sections/HPC/.../sw_devel_methods.pdf.> |
| [3] | [Online]. Available: <https://developers.google.com/youtube/v3/getting-started.> |
| [4] | [Online]. Goals of software design: [www.nada.kth.se/~karlm/prutt05/lectures/prutt05\_lec7.pdf](http://www.nada.kth.se/~karlm/prutt05/lectures/prutt05_lec7.pdf) |

.