MIDDLE EAST TECHNICAL UNIVERSITY

NORTHERN CYPRUS CAMPUS

CNG 491:First Report

Eye Tracking Statistical Analysis Web App

Prepared By:

Table of Contents

Table of Figures

Introduction

This section of the report discusses three main points about the topic and the project, these are what is Eye Tracking Data Analysis (EDA)? the rationale behind choosing such a project, and finally how we think we will achieve this project. The sections following this in the report will focus more about the technical design of the project.

What is Eye-Tracking Data Analysis (EDA)?

Eye tracking data analysis is a research field that has been around for a long time, precisely since the 1800s. However, in the beginning the eye tracking was done using direct observation. As technology evolved eye tracking analysis methods varied and in the 1900s the first eye tracking device was built [1]. Furthermore, one of the first questions asked on what might be the benefit of eye tracking data analysis was in the pattern of reading, how people read, previously it was thought that a person smoothly scan over the text, but it was found that this was not the case and that there is an actual pattern to how people read which is fixating on specific letters and words along the lines. The latter raised the question of why and how does this pattern appear when people read and how understanding this can help improve technology and production. Moreover, this further gave rise to many research fields such as how to improve web sites and online applications, and also improve the experience of various classes of people with disabilities [1].

Eye Tracking data analysis revolves around the idea of understanding and making sense of a set of data which represents the eye movement of a number of people [6]. Nevertheless, this idea of studying the data is not a static procedure that can be always followed, one needs to consider the end result first before thinking of how to start, since the study might revolve around pattern recognition of eye movement, or difference in eye movements for different groups and so on [2]. In contrast, in this project we are trying to provide a platform for researchers to use regardless of their end goal as long as certain characteristics about the research are valid.

Why Eye Tracking Data Analysis (EDA) ?

As mentioned above one of the fields of eye tracking analysis is to investigate how users interact with web pages, in order to improve a website's usability and accessibility. The aim of this project is to create a web-based application for comparative statistical analysis of eye-tracking data between different groups of people . For instance, a study conducted by Eraslan et al. [1] to find out whether users with high-functioning autism experience barriers while using the web compared to users without autism. The main aim behind this project is to provide a platform for researchers that will ease the process of data analysis to a level and reduce the time it takes to finish a specific job in this case preliminary data analysis.

How will this be achieved?

Developing web interface using ASP.net, Which is a framework for building a web app based on HTML, CSS, JavaScript and C#, in order to allow the users to load eye tracking data and image of a web page, and to draw their areas of interest on this web page [7]. Then, processing this data to produce specific features for comparative statistical analysis. In order to process the features we will first check the assumptions of the relevant statistical tests in order to apply the appropriate statistical test. Furthermore, we will use an R based package in C# RdotNet that will help in the process of producing the statistical analysis. R.NET allows the .NET framework to operate alongside the R programming language [5].

We separated the system into two main modules (refer to figure 2 page 12), backend and frontend, the main reason for this is to separate the concerns of the system. Moreover, the backend will mainly focus on handling all the processing that is in the project it will be directly working with R.NET, while the frontend will focus mainly on the interactions with the user and it will be the intermediary between the backend and the user.

The following parts of the report will discuss various technical sides about the project and explain clearly the design of the project.

System Requirements and Specifications

Functional Requirements

	Register
Importance	High
Description	Since the application targets academics mostly, privacy is important and also an environment where users can come back to, and access previous saved data and analysis results. Hence, an account system is implemented to allow users to have their tests online available all the time.
Input	str* Email, str * password
Output	Void

	login
Importance	High
Description	The system should not allow the user to use it unless logged in for privacy purposes
Input	str* Email, str * password
Output	Str * response

	Upload Dataset
Importance	High
Description	This is basically to allow the user to provide the data set at which the application will work on
Input	Text File
Output	Void

	Upload Dataset Visuals
Importance	High
Description	This is to allow the user to provide visuals, from local disks, that test volunteers were tested on and link them to the data automatically.
Input	Images[]
Output	Void

	Link web pages to Dataset Visuals
Importance	High
Description	The system should automatically be able to link the web pages from the dataset to the visuals uploaded by the users.
Input	Images[], WebSiteList[]
Output	Void

	Highlight square segment
Importance	High
Description	The system should allow the user to segment a part of the Dataset visuals in order to base the analysis on them
Input	Selected areas of the screen
Output	segments

	Select Dataset Parameters
Importance	High
Description	The system should allow the user to decide on the set of features he wants to use from the data set to generate them and base the analysis on them. The user will select from a set of predefined features
Input	vector<str> features
Output	void

	Generate features
Importance	High
Description	The system should be able to generate a list of user pre-selected features about the dataset
Input	void
Output	Table of features

	Generate Descriptive study
Importance	High
Description	The system should allow the user to decide when to generate the statistical analysis data and check it as a raw file on the web page
Input	void
Output	Raw text file

	Visualise statistical data
Importance	High
Description	The system should allow the user to view the results of the study in a user friendly way i.e. a table , the system shall also allow the user to download the results as a pdf file
Input	void
Output	pdf file

	Modify data
Importance	High
Description	The system should allow the user to have access to his previous and new data i.e. data sets, segments, and feature list and modify it by changing their values.
Input	void
Output	pdf file

	Add new test to existing data
Importance	High
Description	The system should allow the user to be able to create new test cases for a single set of data by adding the new set of features and segments to the original dataset
Input	void
Output	pdf file

Non-functional Requirements

Category	Non-functional requirements
Usability	● All the features provided by the system shall be easy to understand and use. ● User must be able to have easy access to the elements of his previous tests as soon as he logs in. ● User must be able to work and modify his data with ease in a user friendly interface after login
Accessibility	● The system should be accessible by all types of users i.e. companies, researchers, students … etc.. ● The system should be accessible to users with poor internet connections.
Reliability	● The system should detect input faults in the data i.e. too many NAs to pass the statistical test, or wrong type of data . ● The system should not allow for wrong credentials in login and registration. ● The system should guarantee privacy for all users. ● The system should guarantee the safety of the data that was uploaded and that it can be accessed again.
Performance	● The access time to the system should be short. ● The response time of feature generating should be relatively short. ● The response time of data processing for the user to generate the descriptive study should be short relative to the size of the data. ● The system should be able to handle a large set of users simultaneously. ● The system should be capable of restoring itself to its previous state in the event of a failure.
Supportability	● The system should support all sorts of image file types. ● The system should also be able to recognize different formats for text files of datasets.
Capacity	● The system should provide the user with an initial memory size to store his data.

System Models

Use Case Diagram

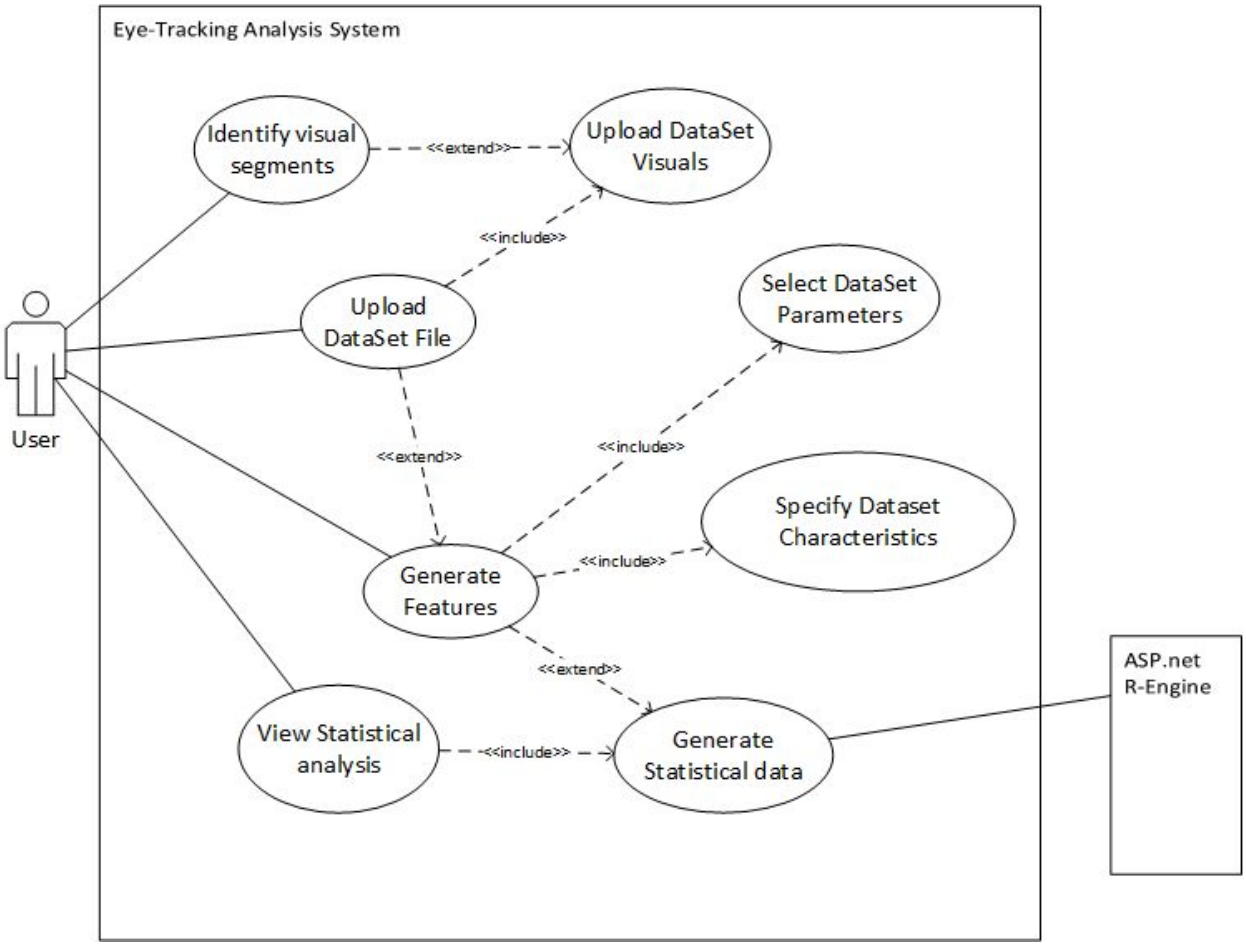


Figure 1: Use Case Diagram

Architecture Diagram

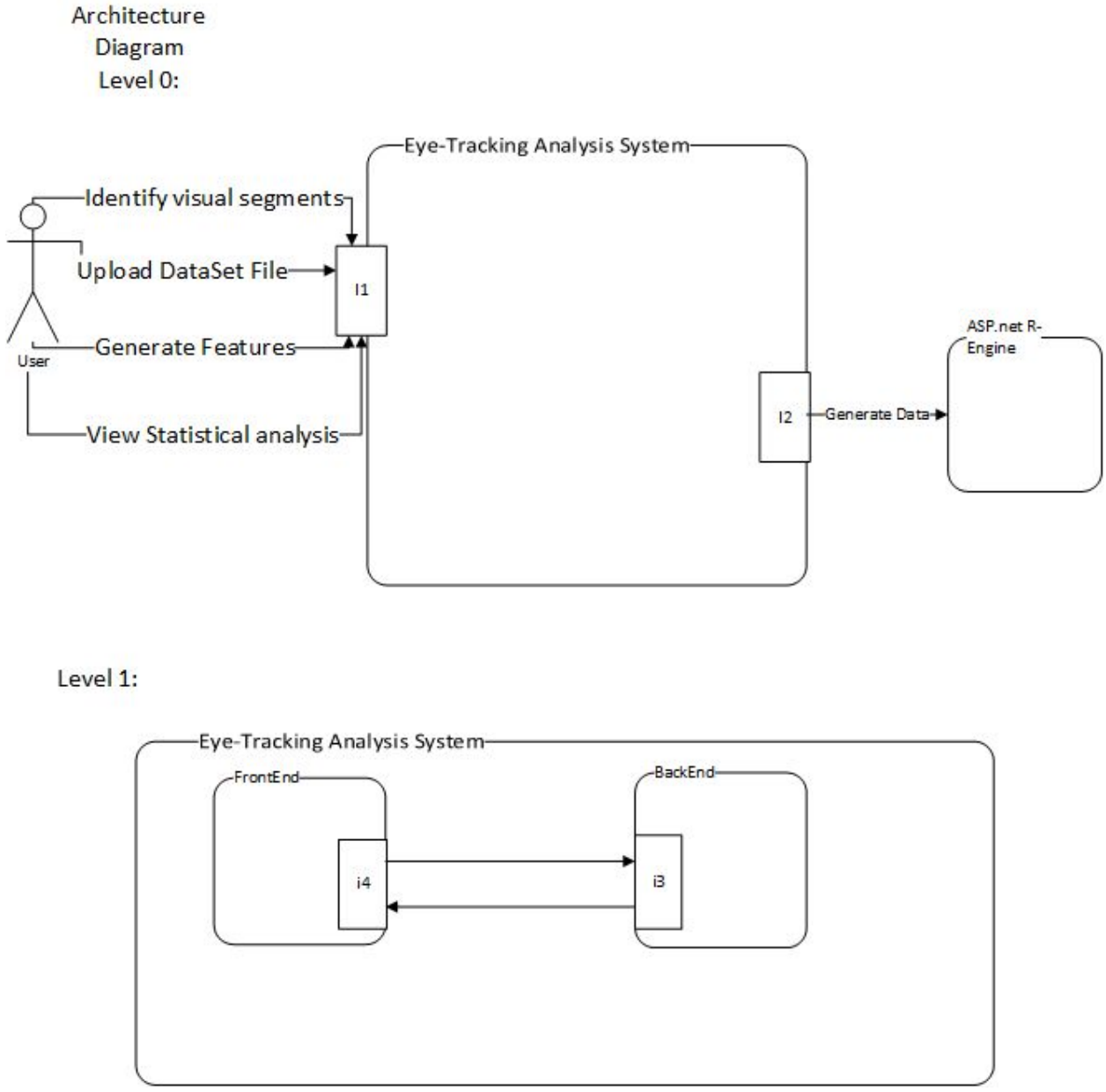


Figure 2: Architecture Diagram Level 0 and 1

Level 2:

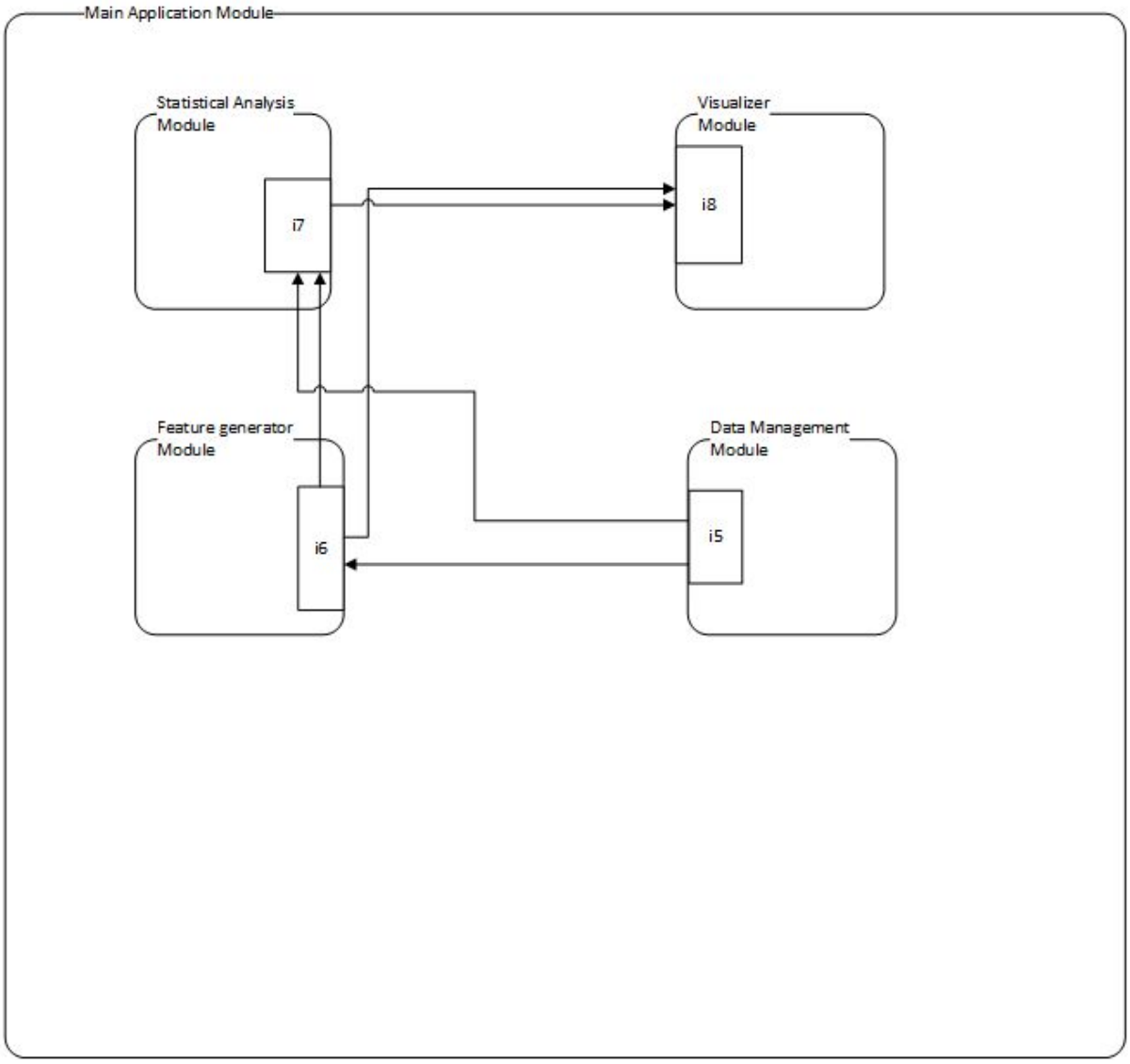


Figure 3: Architecture Diagram Level 2

Level 2 MVC

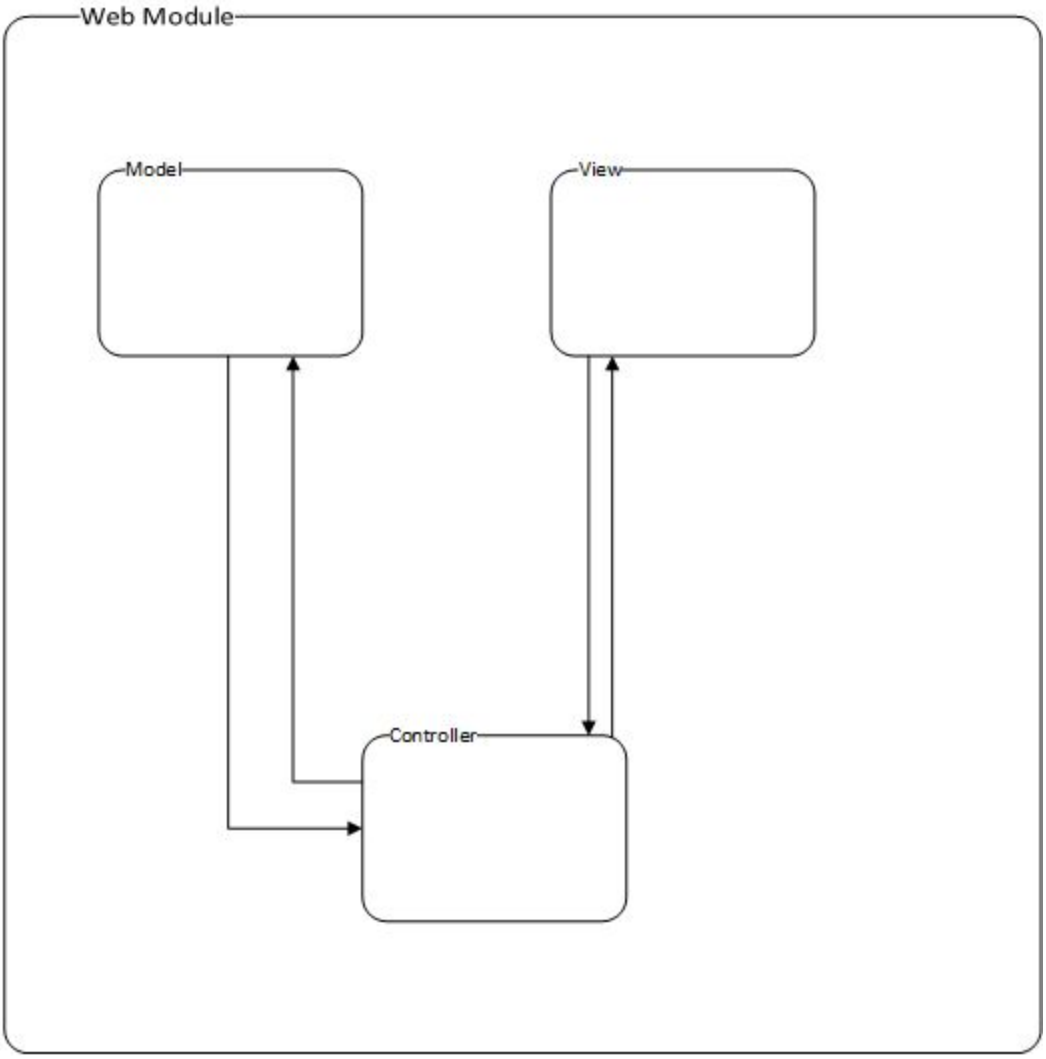


Figure 4:Web Module Level 2 Architecture Diagram

Level 3:

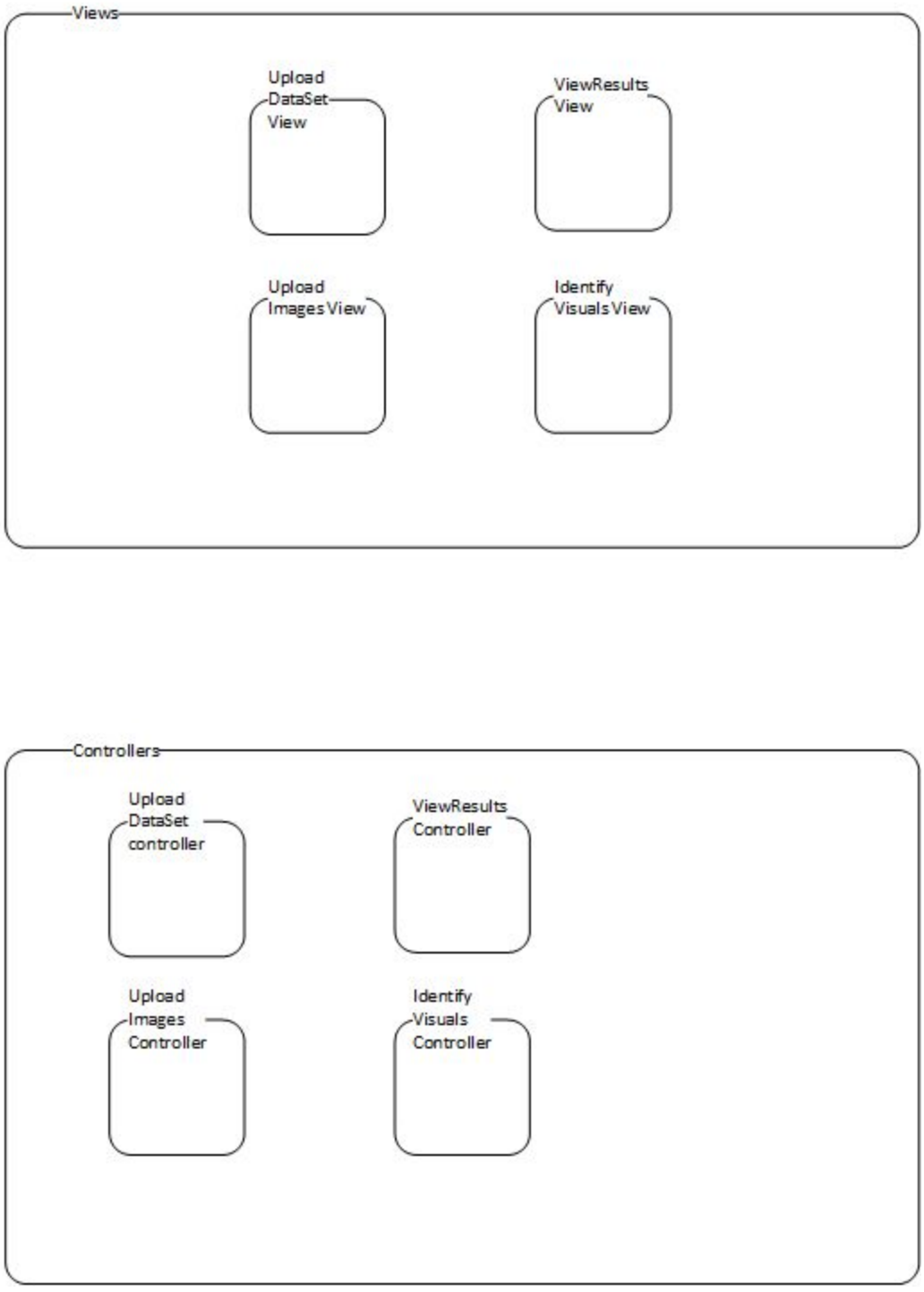


Figure 5: Architecture Diagram Level 3

Sequence Diagram

Upload Dataset

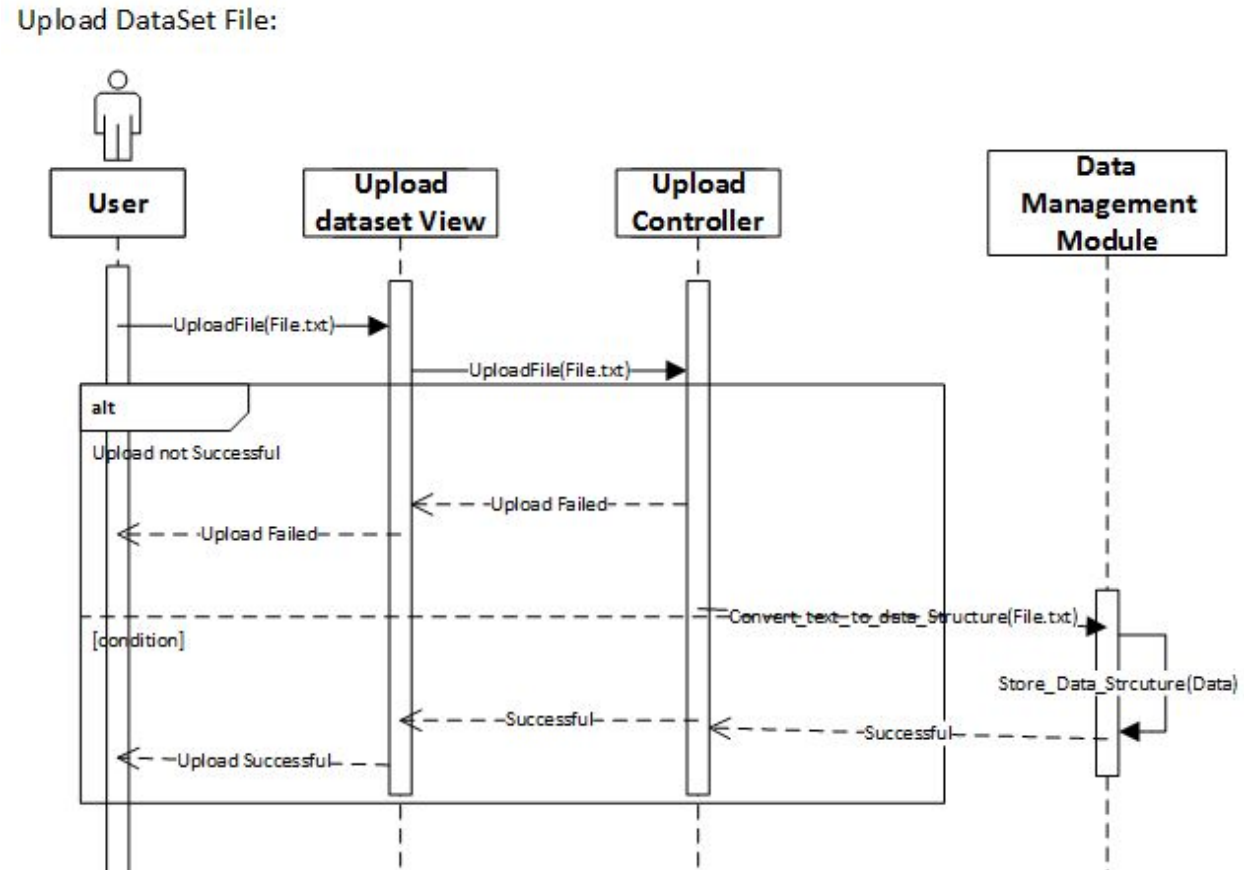


Figure 6: Upload Dataset File Sequence Diagram

Upload Dataset Visual

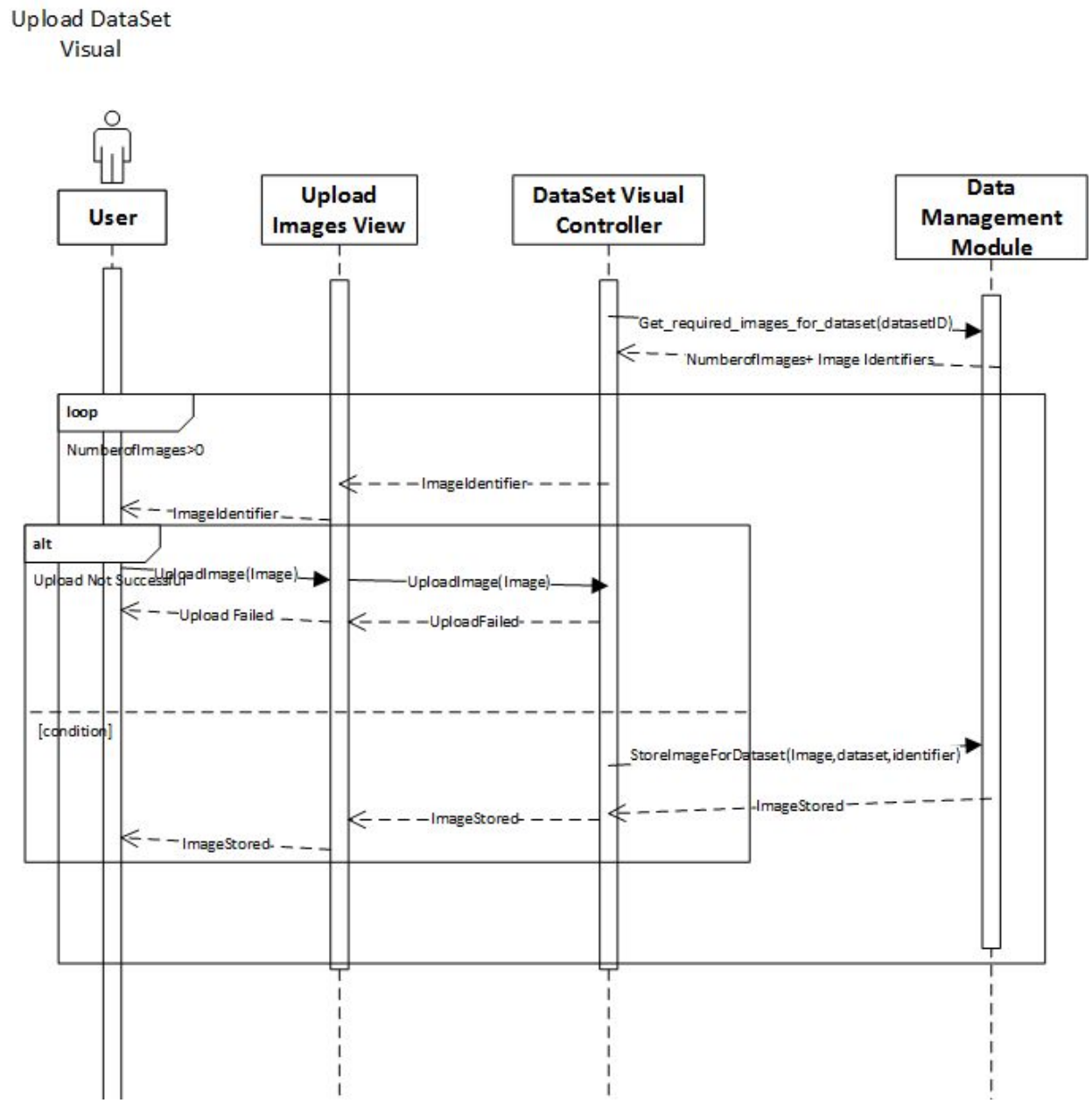


Figure 7: Upload Dataset Data Sequence Diagram

Identify Visual Segments

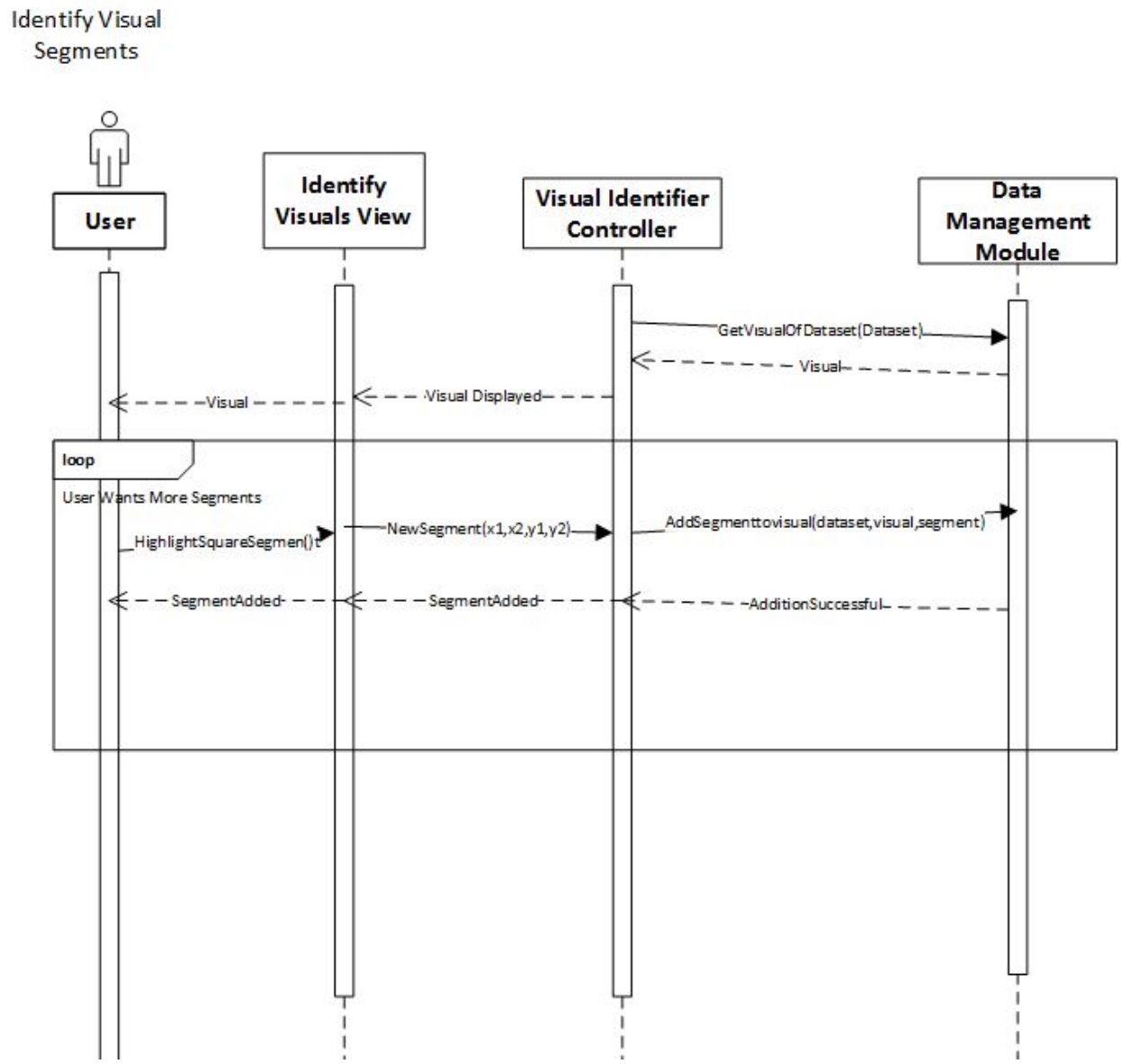


Figure 8: Identify Visual Sequence Diagram

Generate Statistical Analysis

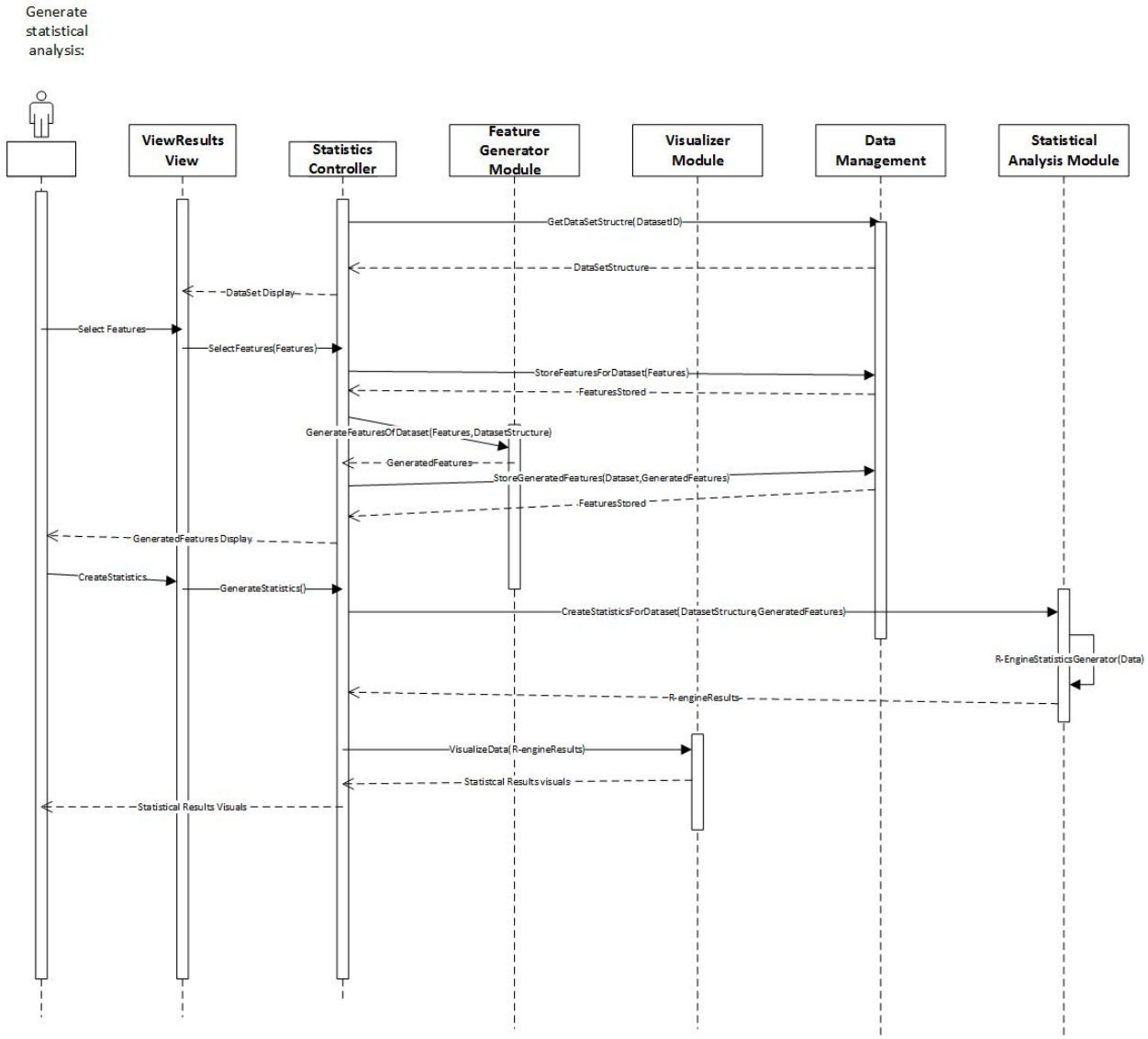


Figure 9: Generate Statistical Analysis Sequence Diagram

Scrum Details

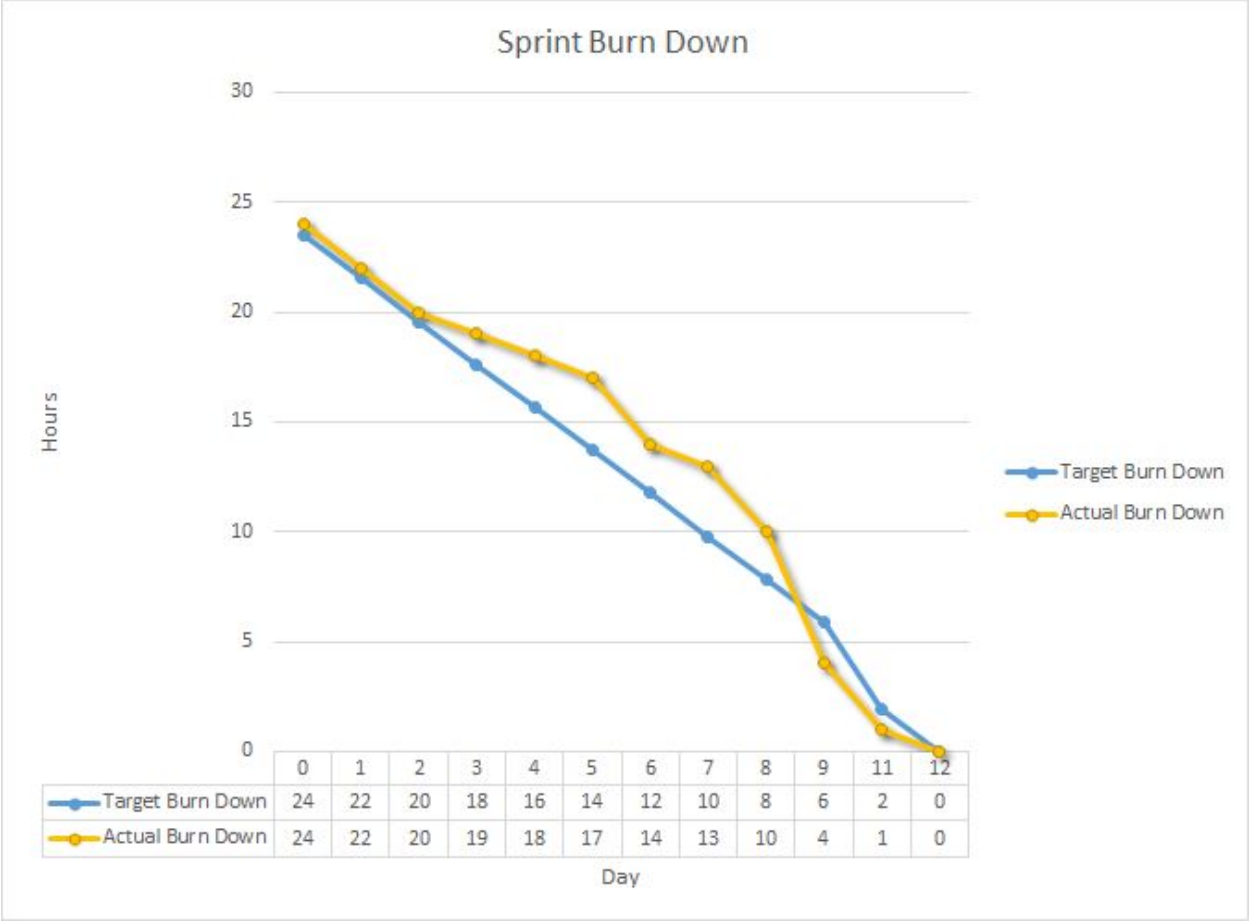
Since the project was based on a topic we are not familiar with , eye-tracking, the main focus of the first sprint was literature review on the topic to get familiar with the web application's domain which was done individually for a period of one week. After literature review we started meeting twice a week to set the functional requirements of the web application and then meet once every two weeks with our supervisors to ensure the correctness of the requirements and design. After that, we started the process of designing the architecture of the web application by creating architecture and sequence diagrams. Finally we created an initial sketch of the interface of the web application.

Sprint backlog

Sprint	Item ID	Estimated Hours	Task Name	Assigned To	Remaining Hours	Status
1	1	10	Literature review and reading eye-tracking research papers	Besher,Abdulmalik, Abdulrahman	0	Completed
1	2	2	Web application Platform literature review	Abdulrahman,Besher,Abdulmalik	0	Completed
1	3	3	Finding out functional/Non-functional requirements	Abdulmalik	0	Completed
1	4	4	Architecture Design	Besher,Abdulmalik, Abdulrahman	0	Completed
1	5	2.5	Setting up visual studio and ASP.net	Abdulrahman	0	Completed
1	6	2	Initial interface sketch	Besher	0	Completed
Total		23.5	SPRINT 1 (01.10.17 – 31.10.17)		0	

Table 1: Sprint backlog

Sprint Burndown Chart



	0	1	2	3	4	5	6	7	8	9	11	12
Target Burn Down	24	22	20	18	16	14	12	10	8	6	2	0
Actual Burn Down	24	22	20	19	18	17	14	13	10	4	1	0

Figure 10: Sprint burndown Chart

Sprint Review

As we reached the end of the sprint we believe that we have done enough literature review and have a good design so that we can start the first sprint of the project. The initial design we have is sufficient to start with the main components of the web application in the next sprint. More work could have been done in the previous sprint and we will add more hours of work starting with the next sprint since more progress should be seen.

Sprint Retrospective

What went well?

- ➔ Everybody was on time during each meeting.
- ➔ The tasks were divided equally and each put a lot of focus on his task.
- ➔ The tasks were done in a synchronized manner and no one outpaced the other.

What went wrong?

- ➔ Literature review took longer than expected whose time could have been used for implementation.
- ➔ More work was done towards the end of the sprint and less throughout the sprint.
- ➔ Time was wasted during meetings to find a gathering point.

What could we do differently to improve?

- ➔ The tasks should be better organized and worked on throughout the sprint.
- ➔ The meetings should be more centered towards design issues and task division.

Project Estimation

Inputs

In the system there are input GUI that we considered and each counted as 1 functional point, we considered some as having a medium complexity depending on the GUI.

Inputs	Complexity	Multiplier	Functional points
Register user	low	3	1
Login	low	3	1
Upload Dataset	Medium	4	1
Upload images	Medium	4	1
Features	low	3	1
Area of interests	High	6	1

Table 2: Inputs

Output

In the system there are output GUI that we considered and each counted as 1 functional point, we considered some as having a medium complexity depending on the GUI.

Output	Complexity	Multiplier	Functional points
Statistical Data	high	7	1

Table 3: Outputs

Query

Query	Complexity	Multiplier	Functional points
Get Dataset	High	6	1

Table 4: Queries

Logical internal files

In the system there are 3 main tables that have transactions that each is counted as one functional point.

Logical internal files	Complexity	Multiplier	Functional points
Data Table	Medium	10	1
Feature Table	low	7	1
User Account table	low	7	1

Table 5: Logical Internal files

External interface files

R.NET is an in-process interoperability bridge to R from the .NET Framework.

External interface files	Complexity	Multiplier	Functional points
R.net	high	10	1

Table 6: External Interface files

Total unadjusted function calls

Total unadjusted function calls	91

Influence Multiplier

TDI	1- data communications	1
	2- Distributed Functions	1
	3- Performance	3
	4- Heavily used configurations	3
	5- Transaction Rate	3
	6- online data entry	4
	7- end user efficiency	5
	8- online updates	5
	9- complex processing	4
	10- reusability	4
	11- installation ease	4
	12- operations ease	5
	13- Multiple sites	1
	14- Facilitate change	3
	TDI	46

Table 7: TDI

Adjusted Total Function Points (ATFP)

IM/VAF= TDI*0.01+0.65	1.11
ATPF= UTPF*IM	77.7

Lines of codes

To calculate the number of code lines we chose Java as the language to use and had a unit size of 55 which we multiplied by ATFP

	C# Unit Size	LOC
LOC = ATFP * Language Unit size	55	4273.5

Estimate the efforts

Using COCOMO I model

a	b	c	Man Month(MM) = a*KDSI^(b)	Total Development Time (TDEV) = 2.5*MM^(c)
2.4	1.05	0.38	11.02894918	6.224478628
			~11	~6

KDSI = ATFP* Language unit size/1000	4.2735

To calculate the effort of Man Month and TDEV we used the formulas for COCOMO with values of a b and c of organic development and the MM we got was 11 while TDEV 6

Estimate the schedule:

Total development time = 6 months according to COCOMO Model

Estimate Team size

Team Size = estimated effort/estimated schedule = ~3 persons

Estimate the ranges

Effort:

Optimistic: 0.25 * 11 = 2.75
Pessimistic: 4 * 11 = 44
Range = 2 to 44 Man Month

Schedule:

Optimistic: 0.6 * 6 = 3.6
Pessimistic: 1.6 * 6 = 9.6
Range = 3 to 9 Months

Team Size:

Optimistic: 2.75 / 3.6 = 1
Pessimistic: 44 / 9.6 = 4
Range = 1 to 4 person

GUI

Register Page

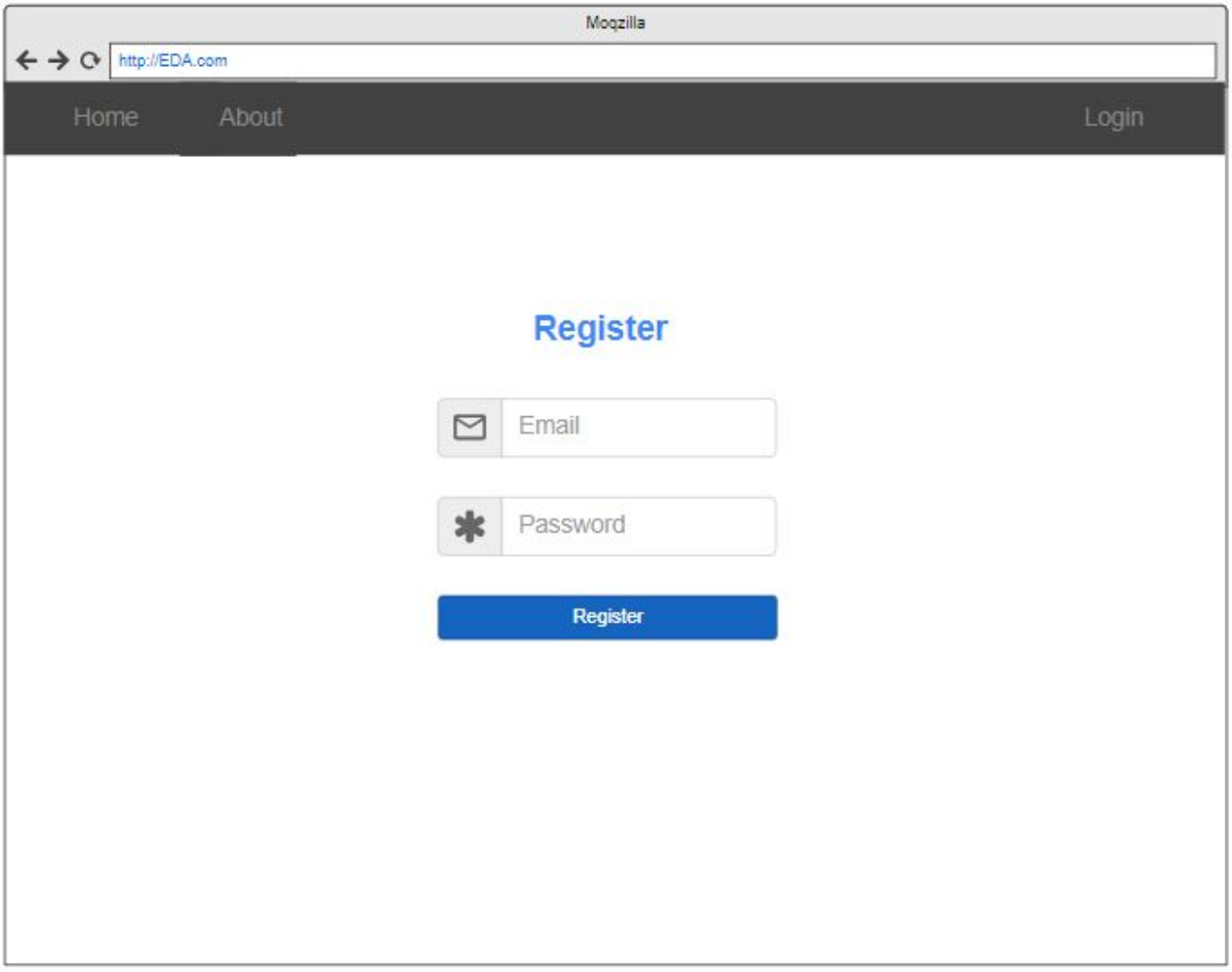


Figure 11: Register Page

Login Page

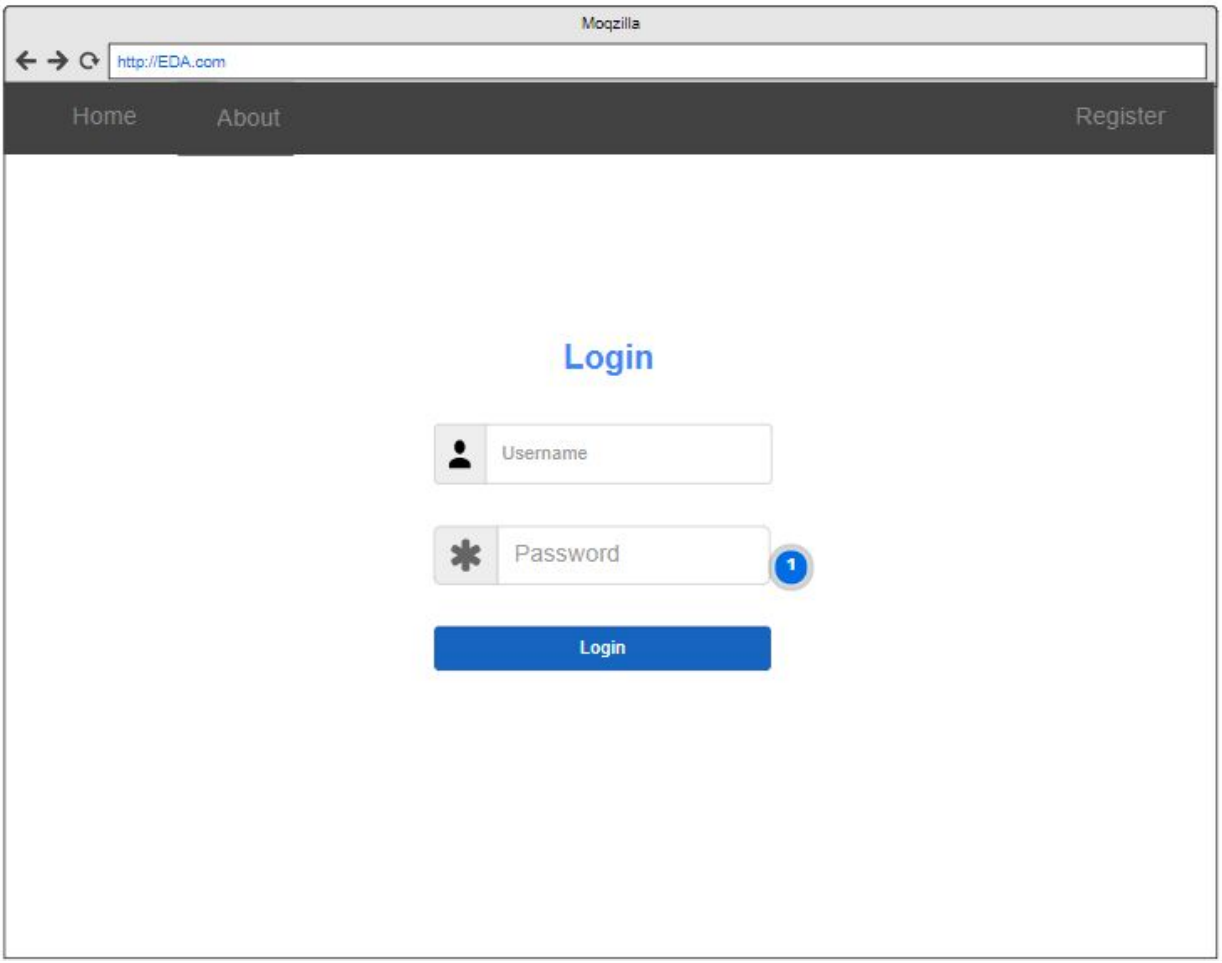


Figure 12: Login Page

Home Page

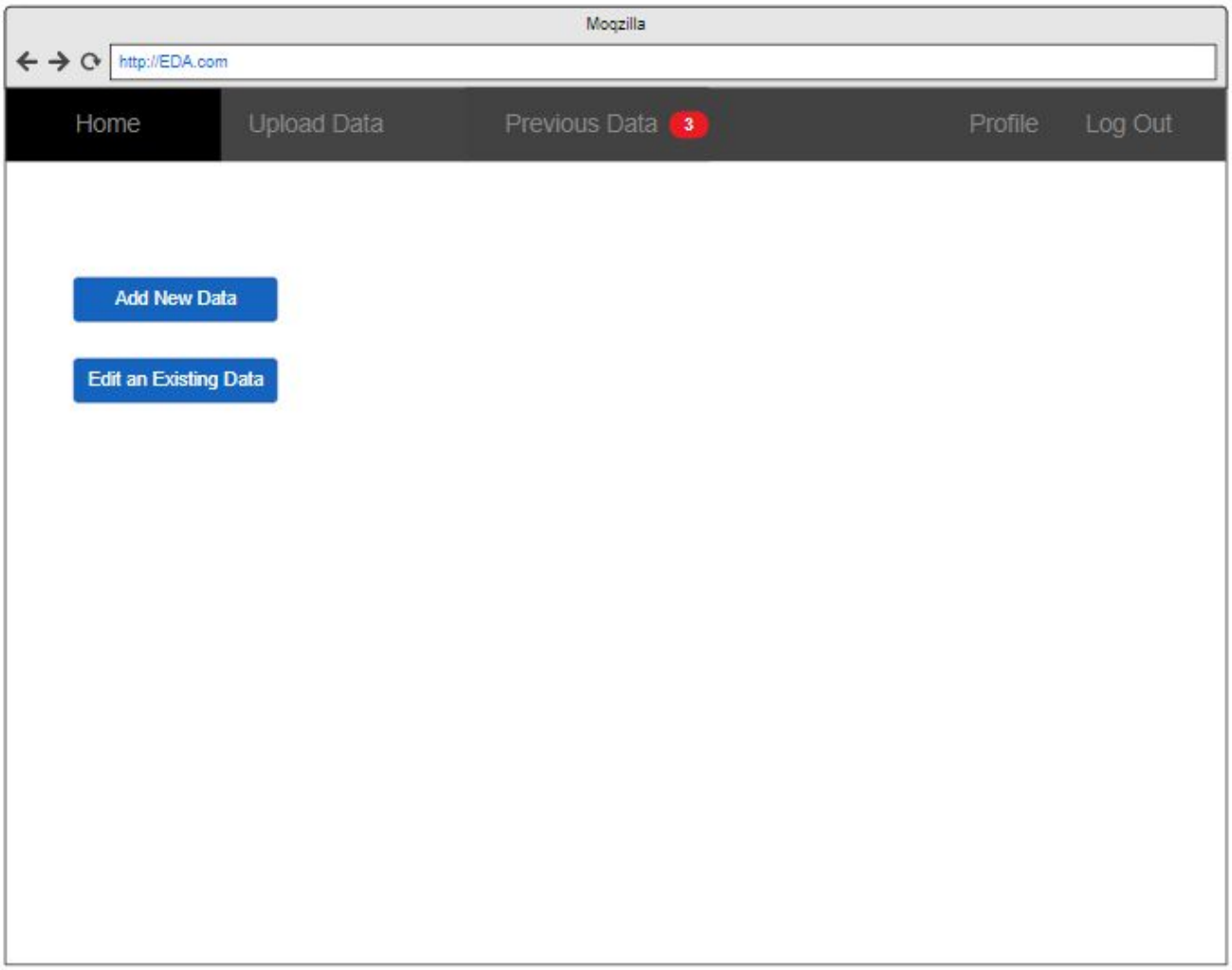
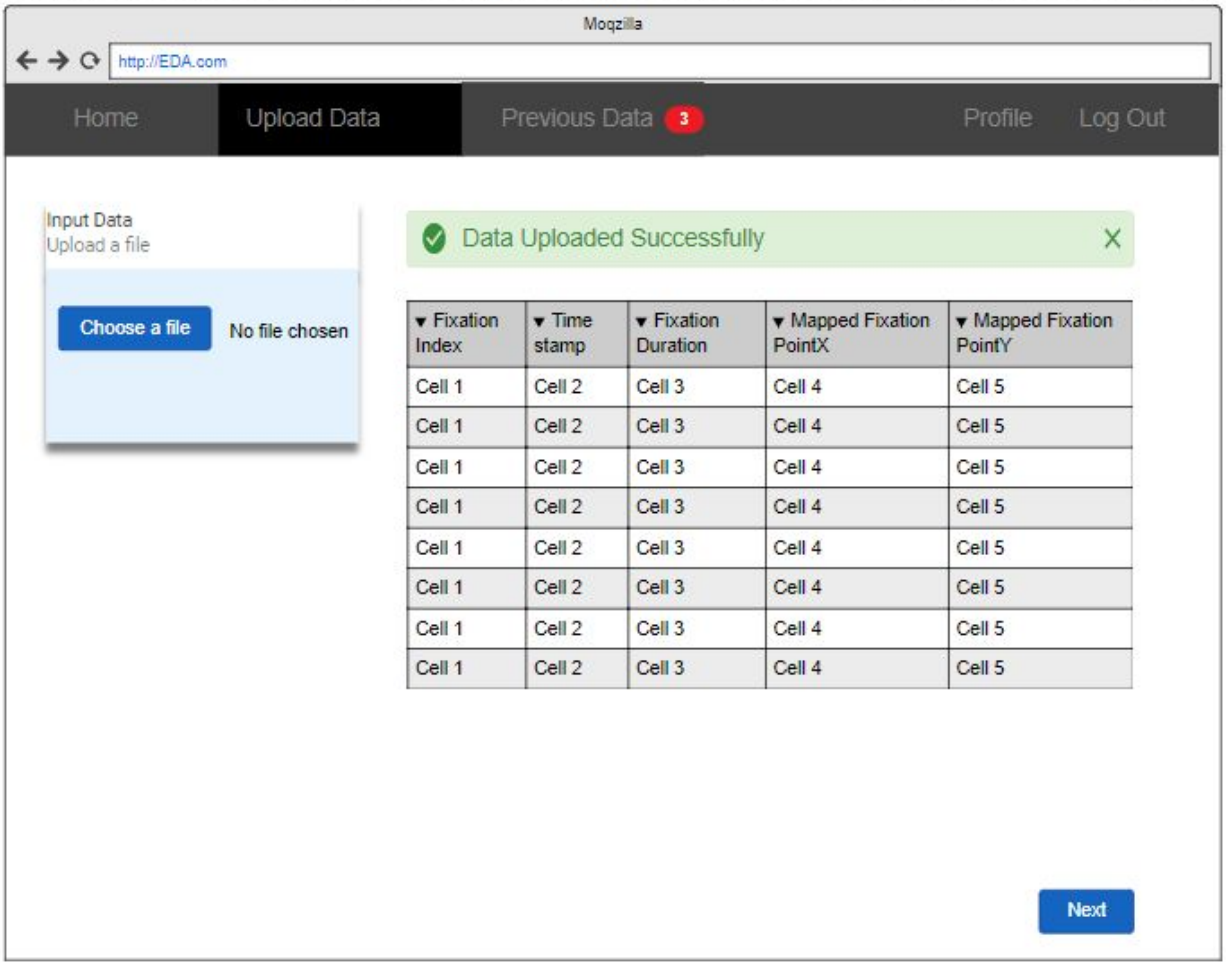


Figure 13: Home Page

Upload Dataset Page



▾ Fixation Index	▾ Time stamp	▾ Fixation Duration	▾ Mapped Fixation PointX	▾ Mapped Fixation PointY
Cell 1	Cell 2	Cell 3	Cell 4	Cell 5
Cell 1	Cell 2	Cell 3	Cell 4	Cell 5
Cell 1	Cell 2	Cell 3	Cell 4	Cell 5
Cell 1	Cell 2	Cell 3	Cell 4	Cell 5
Cell 1	Cell 2	Cell 3	Cell 4	Cell 5
Cell 1	Cell 2	Cell 3	Cell 4	Cell 5
Cell 1	Cell 2	Cell 3	Cell 4	Cell 5
Cell 1	Cell 2	Cell 3	Cell 4	Cell 5

Figure 14: Upload Dataset Page

Upload images Page

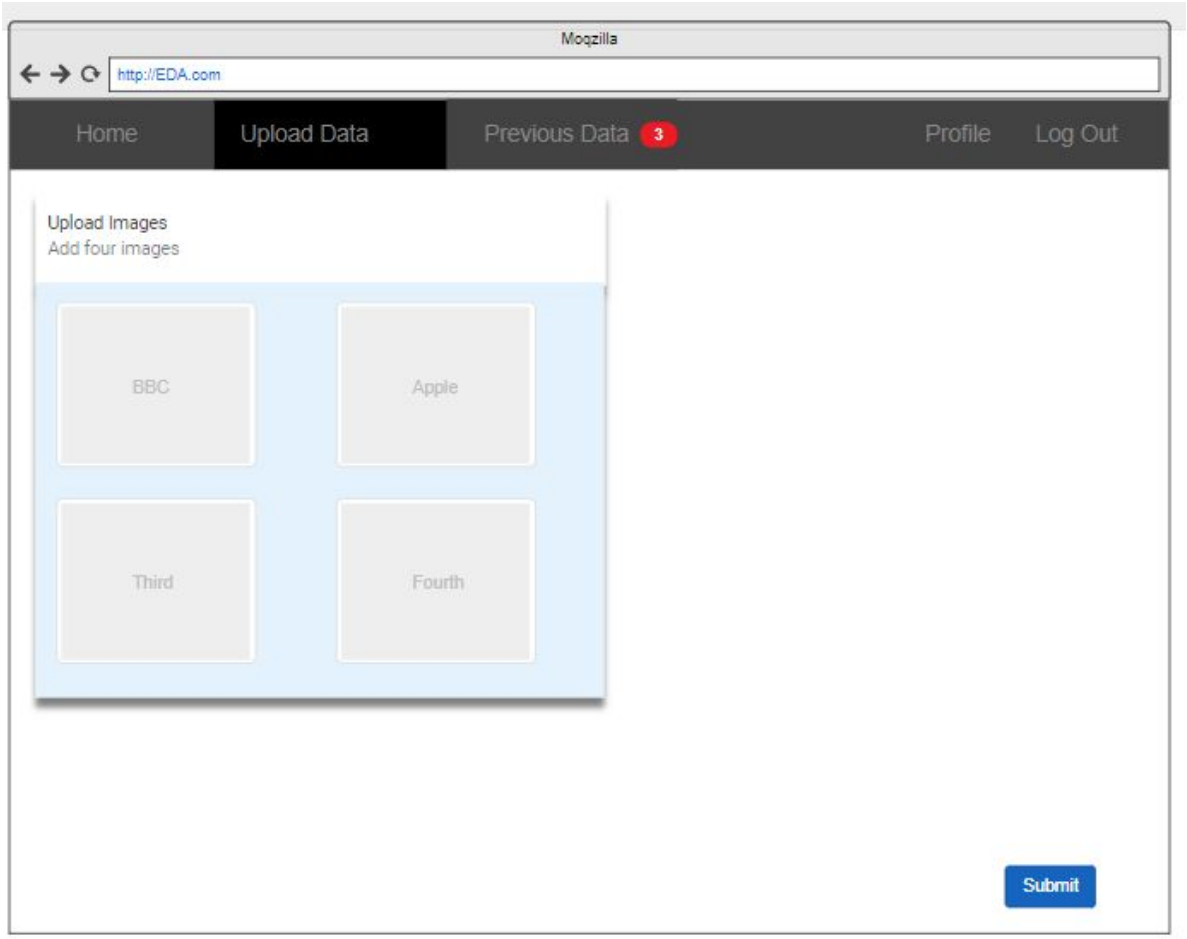


Figure 15: Upload Images Page

Identify visuals Page

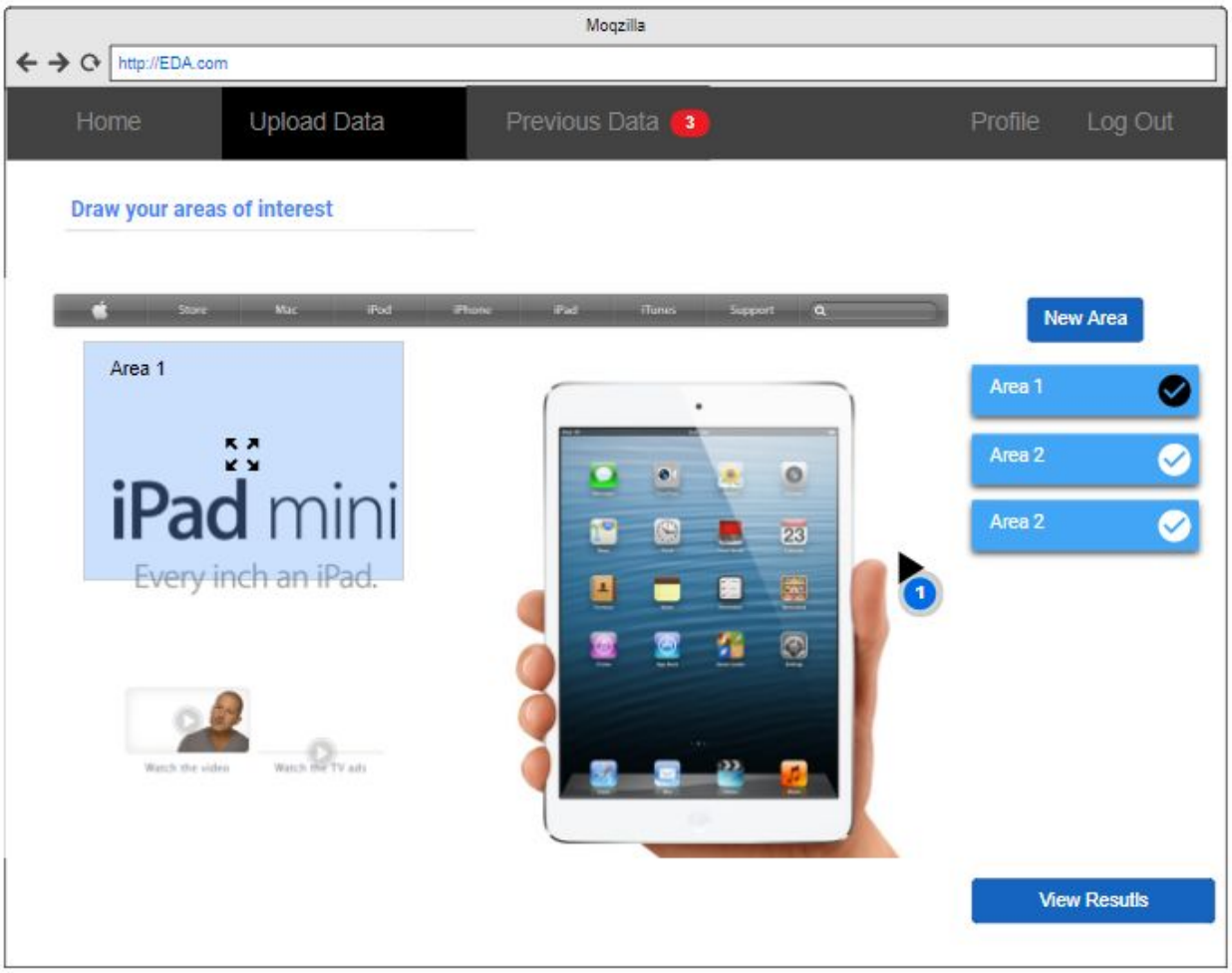


Figure 16: Identify Visuals Page

View Results Page

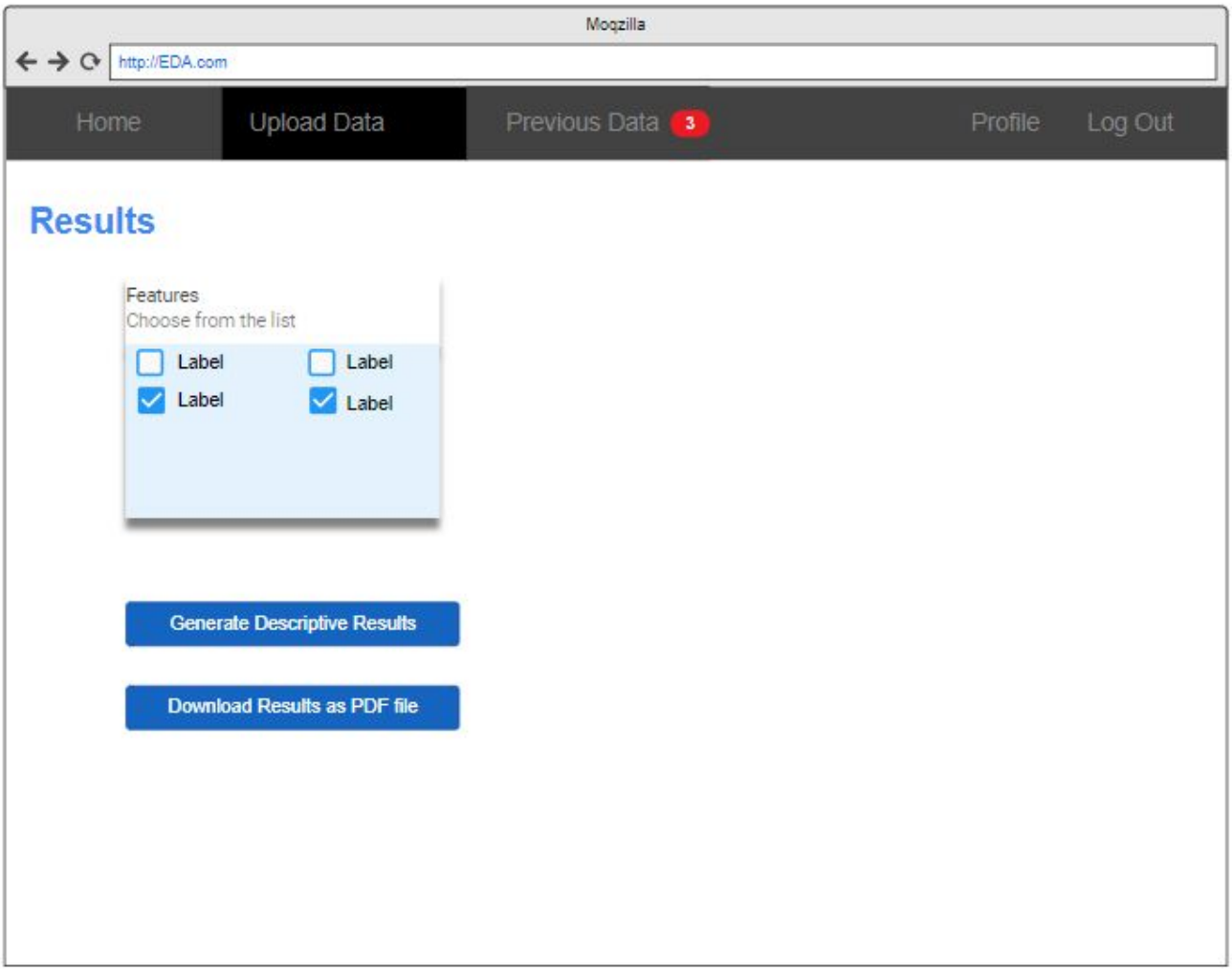


Figure 17: View Results Page

References

[1] Sukru Eraslan, Victoria Yaneva, Yeliz Yesilada, and Simon Harper. Do Web Users with Autism Experience Barriers When Searching for Information Within Web Pages?. In Proceedings of the 14th Web for All Conference on The Future of Accessible Work, W4A 17, Article 20, New York, NY, USA, 2017. ACM.

[2] Claudia Ehmke and Stephanie Wilson. 2007. Identifying web usability problems from eye-tracking data. In Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...but not as we know it - Volume 1 (BCS-HCI '07), Vol. 1. British Computer Society, Swinton, UK, UK, 119-128.

[3] Robert J.K. Jacob, Keith S. Karn, Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. 2002 from http://www.cs.tufts.edu/~jacob/papers/ecem.pdf

[4] R.Net related information from https://rdotnet.codeplex.com

[5] R.Net Documentation from http://jmp75.github.io/rdotnet/

[6] What do we study when we use Eye tracking data article from tobii website https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/what-do-we-study-when-we-use-eye-tracking-data/

[7] asp.net related information from https://www.asp.net

[8] GUI tool https://moqups.com/