D(i)VA

BY

Keyvan Fatehi
Brett Walker
Sania Bishnoi
Sam Applebaum

SRS Requirements Document

Table of Contents

1. Introduction

There is a vast amount of data in society today, increasingly driven by the business world. Companies have access to this data but not all know what to do it with, this data-centered problem has created a lot of excitement within technical disciplines and academic environments within the field of "data science". This project is centered primarily in the field of data science and data visualization. There is space to innovate and build better tools, such as the one we are trying to build.

The following document is divided into ten main sections. The first two sections are the introduction and Overview. Following that is the Application Context section and the User Stories. The System Requirement Section is next with Software Qualities following right after. Other Requirements is also a section and after it is Assumptions/Risks. Lastly, a section is included regarding Priorities and Future Directions.

2. Overview/ Executive Summary

To build a software appliance that integrates the D3.js JavaScript library for manipulating documents based on data. As well as to provide an environment in which users without deep technical knowledge of JavaScript or D3 can try different visualization quickly, rapidly iterating over different ideas without getting caught up in the details of programming.

3. Application Context /Environmental Constraints

The software system will be a web application that is used to help users visualize their data. Users will be able to login and upload data sets of their own to the system. The system in turn will then allow users to choose what visualization to apply to their data sets as well as the dimensions and colors of the visualization. The system will integrate the D3.js JavaScript library that contains a lot of predefined and pre-coded visualization methods.

4. User Stories/Scenarios

Story 1**: Authentication**

As a User
I want to login
Because I want to protect my content

Story 2**: Upload data set**

As a User
I want to upload a data set to my account

Because I want to use it for visualizations

Story 3**: Load previously stored data sets**

As a User
I want to load data sets I have previously uploaded on my account
Because I want to use it for visualizations

Story 4**: Apply a visualization style to a data set**

As a User
I want to apply an existing visualization style to my data set
Because I am trying to quickly scan for the most effective style

Story 5**: Customize visualization aesthetics**

As a User
I want to be able to customize aesthetic parameters (colors, line widths, etc) of a visualization
Because I am trying to create the most effective aesthetic

Story 6**: Customize visualization dimensionality**

As a User
I want to be able to customize the dimensionality (i.e. which columns to use) of a visualization
Because I am trying to focus the visualization on the most relevant dimensions

Story 7**: Data transformation pipeline**

As a User
I want to be able to create one or more filters and data mutations via JavaScript code blocks
Because I am trying to change the data, and thus, the visualization in a particular way without altering the raw data

Scenario 1**: Non-technical User**

A User, that works for large company, is tasked to create a "(insert a type of visualization)". This User is an administrative assistant with excellent knowledge of the Microsoft Suite, but that is the extent of technical abilities. The data and a list of all of the necessary display parameters are given to the user. The User's supervisor expects them to be able to input the data correctly, select all of the correct features to alter the visualization, and export the visualization.

Scenario 2**: Technical User**

A User is a Big Data analyst working as a consultant to a large corporation that sells products. The User is attempting to determine the environmental factors that lead to increases and decreases in sales. The User has a somewhat large dataset with various attributes. The User wants to find a correlation between attributes and trends and wants to see which visualization well show that the best.

5. System Requirements Specification

The system shall allow users to be able to provide their own sets of data.

The system shall allow users the ability to load a previously loaded and stored data set.

The system shall provide a set of predetermined data visualizations for users to choose from.

The system shall allow users to apply their choosing of visualization to their data sets.

6. Software Qualities and Non-functional Requirements

The system shall provide a login system in order to provide security and protect users content.

The system shall allow users to customize their chosen visualization such as the size or color of the visualization.

The system shall allow users to customize the dimensionality of the visualization.

7. Other Requirements

(To be determined...)

8. Assumptions / Risks

We are assuming that the data sets being used in our software are not your typical big sets of data, but rather more medium sized and manageable (i.e. less than 1 million data points).

9. Priorities / Implementation Phases

(To be determined...)

10. Future Directions and Expected Changes

(To be determined...)