Software Requirements Specification (SRS) Document

1. Introduction

1.1 Purpose

The purpose of this Software Requirements Specification (SRS) document is to define the requirements and functionalities for a data visualization project. This project will enable users to upload various file types (CSV, TXT, XLSX), perform different types of data analysis (univariate, bivariate, multivariate), and generate multiple suggested plots. The users will also be able to download these plots in various formats such as PNG, JPEG, and PDF.

1.2 Scope

This project will serve as a tool for students, researchers, and professionals who need to perform data analysis and generate visual representations of their data. The key functionalities include file upload, data selection, analysis type selection, plot suggestion, and plot download.

1.3 Definitions, Acronyms, and Abbreviations

- **CSV**: Comma-Separated Values
- **TXT**: Text file
- **XLSX**: Excel Spreadsheet
- **SRS**: Software Requirements Specification
- **PNG**: Portable Network Graphics
- **JPEG**: Joint Photographic Experts Group
- **PDF**: Portable Document Format

1.4 References

- IEEE Std 830-1998, IEEE Recommended Practice for Software Requirements Specifications
- Matplotlib Documentation: https://matplotlib.org/stable/contents.html
- Pandas Documentation: https://pandas.pydata.org/docs/

2. Overall Description

2.1 Product Perspective

The data visualization project is an independent tool designed to facilitate data analysis and visualization. It will provide an intuitive interface for users to upload data files, select specific data for analysis, choose the type of analysis, and generate appropriate plots which can be downloaded in various formats.

2.2 Product Functions

- File upload: Users can upload CSV, TXT, and XLSX files.
- Data selection:
    - For CSV files: Users can select columns.
    - For TXT files: Users can select the length of data to be analyzed.
    - For XLSX files: Users can select sheets and then columns.
- Analysis type selection: Users can choose univariate, bivariate, or multivariate analysis.
- Plot suggestion: Users can choose from multiple suggested plots.
- Plot download: Users can download plots in PNG, JPEG, and PDF formats.

2.3 User Classes and Characteristics

- **Students**: Users who need to visualize data for assignments or projects.
- **Researchers**: Users who analyze data for research purposes.
- **Professionals**: Users who need to create data visualizations for business or other professional purposes.

2.4 Operating Environment

The application will be web-based and should be compatible with modern web browsers (Chrome, Firefox, Safari, Edge).

2.5 Design and Implementation Constraints

- The application should handle large files efficiently.
- The user interface should be responsive and user-friendly.
- Data security and privacy should be ensured for uploaded files.

2.6 Assumptions and Dependencies

- Users have access to a stable internet connection.

- Users have a basic understanding of data files and formats.
- The application will utilize libraries such as Pandas for data manipulation and Matplotlib for plotting.

3. Specific Requirements

3.1 Functional Requirements

3.1.1 File Upload

- **REQ-1**: The system shall allow users to upload CSV, TXT, and XLSX files.
- **REQ-2**: The system shall display a confirmation message upon successful file upload.

3.1.2 Data Selection

- **REQ-3**: For CSV files, the system shall allow users to select specific columns for analysis.
- **REQ-4**: For TXT files, the system shall allow users to specify the length of data to be analyzed.
- **REQ-5**: For XLSX files, the system shall allow users to select sheets and then columns for analysis.

3.1.3 Analysis Type Selection

- **REQ-6**: The system shall allow users to select univariate, bivariate, or multivariate analysis.

3.1.4 Plot Suggestion and Generation

- **REQ-7**: The system shall suggest appropriate plots based on the selected data and analysis type.
- **REQ-8**: The system shall generate plots for the selected data and analysis type.

3.1.5 Plot Download

- **REQ-9**: The system shall allow users to download generated plots in PNG, JPEG, and PDF formats.

3.2 Non-Functional Requirements

3.2.1 Performance Requirements

- **REQ-10**: The system shall process and analyze data files efficiently.
- **REQ-11**: The system shall generate plots within a reasonable time frame (e.g., within 5 seconds for average file sizes).

3.2.2 Usability Requirements

- **REQ-12**: The system shall have an intuitive and user-friendly interface.
- **REQ-13**: The system shall provide clear instructions and feedback to users.

3.2.3 Security Requirements

- **REQ-14**: The system shall ensure the security and privacy of uploaded data files.
- **REQ-15**: The system shall delete uploaded files after a certain period (e.g., 24 hours) to ensure data privacy.

3.2.4 Compatibility Requirements

- **REQ-16**: The system shall be compatible with modern web browsers (Chrome, Firefox, Safari, Edge).

4. Additional Considerations

4.1 Future Enhancements

- Integration with cloud storage services for file upload and download.
- Advanced data cleaning and preprocessing options.
- More plot types and customization options.
- Collaboration features allowing multiple users to work on the same data set.

5. Appendices

5.1 Glossary

- **Univariate Analysis**: Analysis of one variable at a time.
- **Bivariate Analysis**: Analysis of two variables simultaneously to explore the relationship between them.
- **Multivariate Analysis**: Analysis of more than two variables to understand complex relationships and patterns.