ASSIGNMENT - 1

Architecture and Implementation of Reactive Web Applications with R Shiny

Data science has transcended the mere execution of statistical algorithms in isolated development environments to focus on the delivery of actionable knowledge through interactive interfaces. Within this ecosystem, the Shiny package for the R language allows analysts and researchers to translate complex analytical scripts into dynamic web applications without the need for deep training in front-end technologies like HTML5, CSS3, or JavaScript.

This architecture is grounded in the paradigm of reactive programming, a model where dependency relationships between input data and output results are automatically managed through a directed acyclic graph, ensuring that any modification in the user state is immediately and efficiently reflected in the visual interface.

This report details the structuring of a comprehensive pedagogical practice, designed to guide the user from the genesis of a simplified monovariate application to the construction of multivariate visualization systems and high-level dashboards, using the classic iris dataset from R's base library as the backbone.

Foundations of Data Selection in Learning and Prototyping Environments

The choice of an initial dataset is a critical decision that determines the learning curve in reactive software development. In the R environment, "toy datasets" offer a controlled environment where data cleaning is minimal, allowing the focus to remain strictly on the application's architecture.

The Analytical Rationale of the Iris Dataset and Modern Alternatives

The iris dataset, introduced by biologist and statistician Ronald Fisher in 1936, remains the gold standard for Shiny practice due to its clean structure and the clear separation of its classes. This set contains 150 observations equally divided among three species of flowers (Setosa, Versicolor, and Virginica), with four numerical variables describing the physical dimensions of sepals and petals. This distribution facilitates the creation of histograms, scatter plots, and classification models without the complications derived from missing data or outliers often found in real-world data.

While iris has historically dominated, the current landscape offers alternatives:

Dataset	Observations	Variables	Predominant Types
iris	150	5	Numeric, Factor
mtcars	32	11	Numeric
penguins	344	8	Numeric, Factor
faithful	272	2	Numeric
Titanic	2201	4	Factor

Objective: The goal of this assignment is to apply the principles of reactive programming to create a functional web application using R Shiny. You will practice data selection, UI/Server architecture, and version control by publishing your project to a public repository.

Task:

1. Choose one of the suggested datasets discussed in the research material to serve as the basis for your application.

2. Build a web application that includes: sidebar layout, reactive inputs and dynamic outputs.

3. Create a Github repository for your project.

4. You must submit a written PDF document with a maximum of 5 pages explaining the functionality of your application.