

Automation of Exploratory Report Generation using R shiny and R Markdown

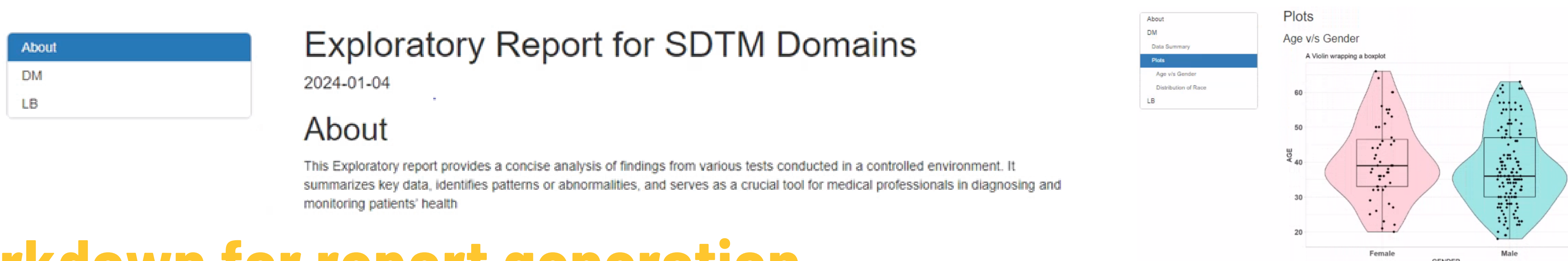
Denny Sebastian (denny.sebastian@catalystcr.com) , Vikash Jain (vikash.jain@catalystcr.com)

Introduction

Exploratory Data Analysis (EDA) in clinical research involves analysing and summarizing data to extract insights, identify patterns, and produce hypotheses—all of which are essential for guiding subsequent research directions. EDA here seeks to understand treatment responses, potential correlations, and patient demographics. Researchers investigate data variability, distributions, and correlations between variables via statistical measurements, visualizations, and summary approaches. Data distributions, outliers, and trends are shown by graphical representations like histograms, boxplots, and scatter plots, while summary statistics like mean, median, and standard deviation provide light on central tendencies and dispersion.

In R, a powerful statistical programming language, numerous visualization libraries help to exploratory data analysis (EDA). Here are some major visualization tools commonly used in EDA with R:

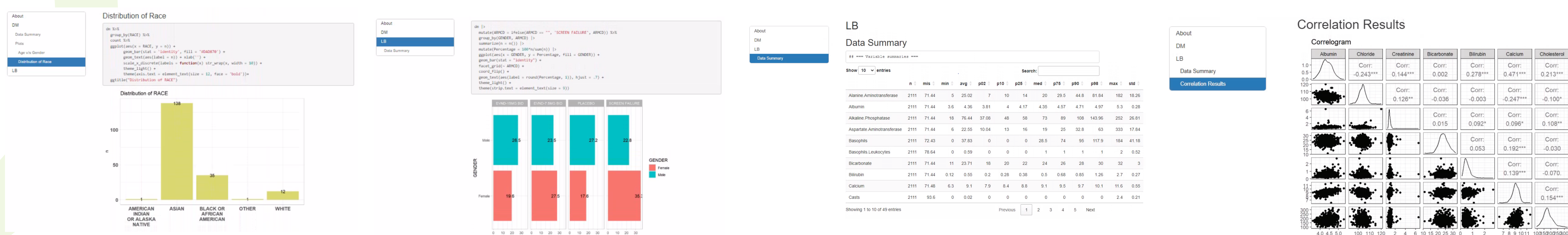
- **ggplot2**: A versatile and widely-used package for creating a variety of visualizations using a layered grammar approach.
- **Plotly**: Allows for interactive visualizations, enhancing EDA by enabling zooming, panning, and hover-over details.
- **ggvis**: Builds on ggplot2 but emphasizes interactive web-based visualizations, suitable for dynamic exploration and manipulation of data.
- **lattice**: Provides a powerful framework for creating conditioned plots like trellis plots, which are useful for visualizing relationships across multiple variables.
- **R base graphics**: The default plotting system in R offers various functions for creating basic visualizations directly from R's base package.



Rmarkdown for report generation

Rmarkdown is a strong tool in R for creating dynamic reports that incorporate text, code, and visuals. It facilitates the smooth amalgamation of text, code, and output into a unified document format, hence enabling the automatic creation of reports. The integration of narrative, code chunks, and visuals improves document readability and reproducibility by streamlining the production of dynamic, repeatable reports in several forms such as Word, PDF, HTML, or presentations.

- Integrating Rmarkdown reports into an R Shiny dashboard enables dynamic and interactive reporting.
- Using the render function in Shiny, the report generated in Rmarkdown can be embedded within the Shiny app, allowing users to generate and view customized reports based on their inputs and selections in the dashboard interface



Conclusion

Automating the generation of exploratory reports through the integration of R Shiny and R Markdown marks a significant leap in efficiency and productivity. This integration makes it possible to create dynamic, data-driven reports with flexibility that not only accelerate the reporting process but also empower users to explore data insights in greater detail.

This automated fusion meets a variety of user requirements by utilizing R Markdown's adaptable report layout and R Shiny's interactive features. It provides the freedom to quickly produce thorough reports, interactively examine data, and alter graphics.

Overall, the automation of exploratory report generation using R Shiny and R Markdown is a revolutionary method that improves accessibility, effectiveness, and the breadth of insights obtained from data research.