

INTRODUCTION

The cloud-based KIWI™ platform is an enterprise-scale application supporting the execution, management, and communication of modeling and simulation activities.¹ Exploration of analysis datasets and model outputs is a key component of these activities as it allows the identification of data patterns and guides the creation and execution of modeling plans and supports communication of modeling results within the development team or with regulatory agencies who rely on exploratory displays to convey the rationale of decision-making or to illustrate problems and knowledge gaps that should be further investigated. The creation of meaningful exploratory plots and tables is a time-consuming task even for skilled programmers.

KIWI includes multiple modules providing user-friendly graphical interfaces for efficient and flexible data exploration and reporting:

- Visualize was presented at a previous PAGE meeting¹ and provides a library focused on the automated creation of model diagnostic plots.

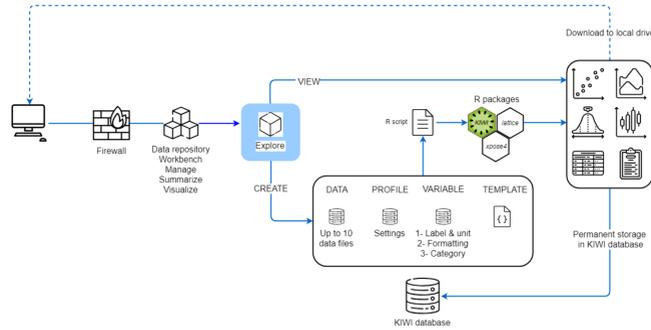
- Explore and ExploreLive are 2 new modules providing libraries of generic tools to automate the creation of summary statistics tables and plots. It was designed to explore the content of analysis datasets and model output tables.

All displays created in KIWI are generated with templated R code² and do not require any programming experience on the part of the users.

DESIGN OVERVIEW

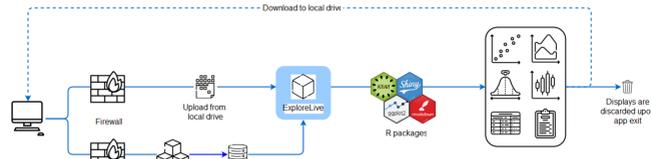
Explore

Explore is a module intended for the creation and permanent storage of report-quality tables and lattice-based plots.³ R scripts are automatically generated and executed using reusable profiles of settings selected in the graphical interface. All outputs are stored in a database and tied to a unique dataset identifier assuring traceability and reproducibility. Execution of these custom profiles typically generates multiple displays which are automatically organized in a hierarchical tree format for convenient navigation and comparison across multiple data sources. Explore leverages the graphical standards and data formatting features established for model diagnostic plotting in the Visualize module, offering seamless visual integration in technical reports.



ExploreLive

ExploreLive is a Shiny application intended as a sandbox environment for the live and nonpermanent manipulation of data, creation of tables and interactive ggplot2-based plots, and execution of small-scale data analysis.^{4,5} ExploreLive is remotely hosted and rendered in the user's web browser. All displays are generated on the server side based upon the user-selected options and rendered in the application. The Shiny architecture implements swift reactivity between user input changes and the update of the display rendered on the screen. Interactive features allow users to click on plots and obtain relevant information about the data source. Displays can be downloaded to a local drive. Upon exiting the application, all settings and displays are discarded on the server and web browser sides.



Quick Comparison of Functionality

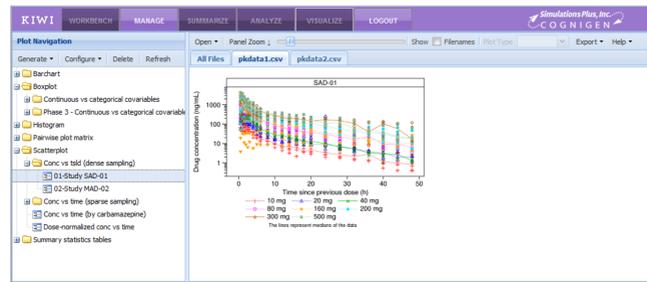
Feature	Explore	ExploreLive
Secure access	✓	✓
Data source	KIWI database	KIWI database Upload from local drive
Number of datafiles that can be explored simultaneously	Up to 10	1 (3 data files loaded)
Data manipulation	✗	✓
Data subset	✓	✓
Plot engine	Lattice	ggplot2
Scatter and line plots	✓	✓
Box-and-whisker plot	✓	✓
Barchart	✓	✓
Histogram	✓	✓
Pairwise plot matrix	✓	✓
Kaplan-Meier plot	✗	✓
Summary statistics table	✓	✓
Contingency table	✓	✓
Data binning plot	✗	✓
Linear regression	✗	✓
Interactive plot	✗	✓
Permanent setting storage	✓	✗
Permanent display storage	✓	✗
Permanent R script storage	✓	✗
Display organization	✓	✗
Multi-format export (png, jpg, pdf, docx, xlsx, html)	✓	✗

REFERENCES

- Bihorel S, et al. KIWI: a collaborative platform for modeling and simulation. PAGE 23 (2014) Abstr 3124 [www.page-meeting.org/?abstract=3124].
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Sarkar D. (2008) Lattice: Multivariate Data Visualization with R. Springer, New York. ISBN 978-0-387-75968-5.
- Wickham H. (2009) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.
- Chang W, et al. Shiny: Web Application Framework for R. <https://CRAN.R-project.org/package=shiny>.

*Some icons used in this poster were made by Freepik from www.flaticon.com.

EXPLORE FUNCTIONALITY



Shared Functionality

Explore shares many elements of functionality with the Visualize module, offering seamless visual integration of displays created by both modules:

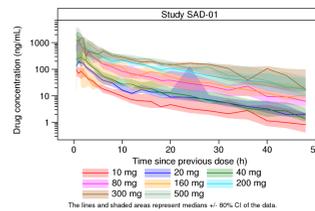
- Library of reusable settings profiles within and across projects
- R code libraries for data extraction, processing, and plotting
- Library of variable labels and units (eg, AGE: age [y])
- Library of variable format collections (eg, 0: male; 1: female)
- Variable categorization (eg, AGE is continuous, RFCAT is categorical)

Shared functionality across exploratory types in Explore:

- Data subset using custom R expressions or predefined conditions (eg, first record per subject)
- Data grouping using different colors, symbols, and/or line styles
- Multivariable paneling and stratification of data
- Creation of additional combined plots when paneling and grouping is used
- Five methods of handling below the lower limit of quantitation data
- Advanced layout settings
- Reference and spline lines
- Custom title and footnote
- Single or multiple display downloads

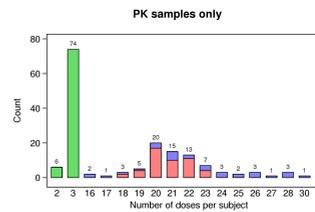
Scatter and Line Plots

- Overlay of multiple Y axis variables
- Secondary Y axis
- Axis zoom, log scale, and transformation
- Central tendency (median/mean) ± variability (SD/SE/CI)



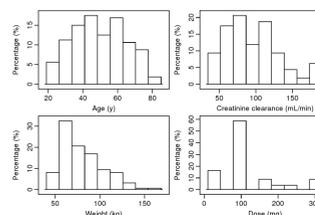
Barchart

- Single-variable count
- Multiple binning methods for continuous X-axis variable
- Count of record number per individual
- Y versus X variable



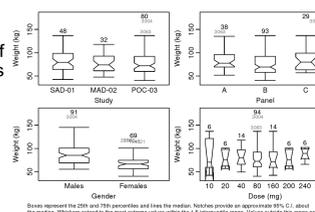
Distribution Plot

- Density, count, or percentage
- Histogram created for multiple variables at once



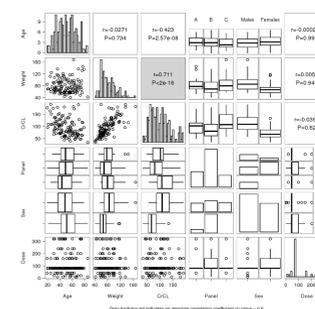
Boxplot

- Boxplots created for multiple pairs of continuous and categorical variables at once
- Mini statistics report associated with each boxplot



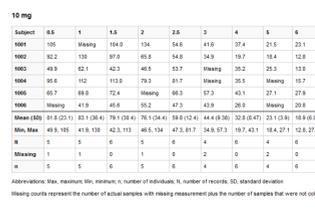
Pairwise Plot Matrix

- Co-distribution between multiple continuous or categorical variables
- Automated detection of high correlation between continuous variables



Summary Statistics and Contingency Tables

- Custom selection of statistics
- Coercion of continuous variables into categorical variables
- Four-table layout for summary statistics tables, including 1 listing individual variable value + summary statistics

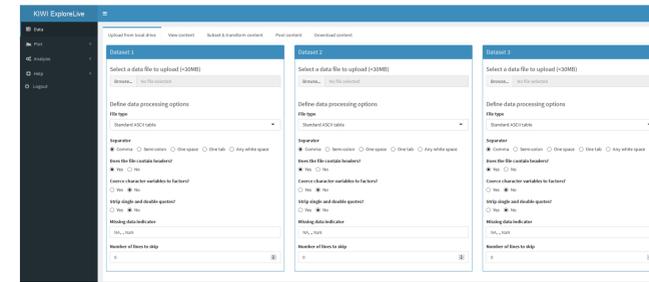


Superprofiles are coming...

Bihorel S, Fox D, Kumar P, Samuel S, Manunta S, Rokitka A, Walawander C, Grasela T. Data exploration libraries in KIWI. Poster presented at: PAGE 2018. 27th Annual Meeting of the Population Approach Group in Europe; 2018 May 29-June 1; Montreux, Switzerland.

For additional information, please contact
 Sebastian Bihorel, PharmD, PhD
 Cognigen Corporation, a *SimulationsPlus* Company
 1750 Wehrle Drive, Suite 110, Buffalo, NY 14221
 (716) 633-3463, ext. 323 or sebastian.bihorel@cognigencorp.com

EXPLORELIVE FUNCTIONALITY



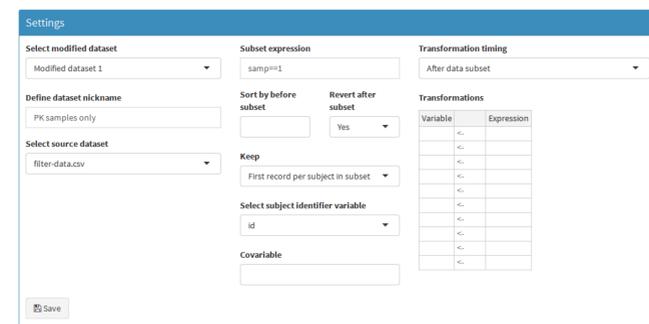
Shared Functionality

Shared functionality across ExploreLive sub-modules:

- Data subset
- Data grouping using different colors, symbols, and/or line styles
- Multivariable paneling and stratification of data
- Reference lines
- Download

Data

- Provide dataset for use in other sub-modules
- Subset using custom R expressions or predefined conditions (eg, first record per subject)
- Sort
- Add and transform variables
- Pool and merge datasets



Scatter and Line Plots

- Interactive plot with data highlight and query
- Axis zoom, log scale, and transformation
- Reference and spline lines
- Central tendency (median/mean) ± variability (SD/SE/CI)

Barchart

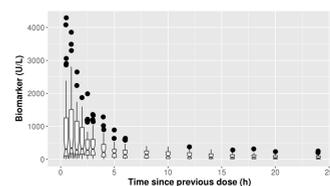
- Single-variable count or Y versus X variable (with or without error bar)

Distribution Plot

- Histogram and QQ plots
- Density, count, or percentage
- Choice of distribution and kernel functions

Boxplot

- Interactive plot with data highlight and query
- Box and whiskers for continuous or categorical X-axis variable



Linked Plots

- Up to 4 scatter or box-and-whisker plots
- Cross-interactivity between plots with data highlight
- Axis zoom, log scale, and transformation

Pairwise Plot Matrix

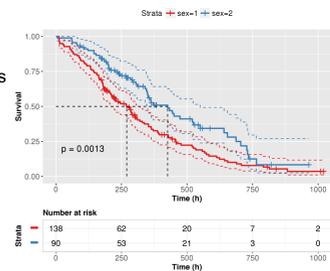
- Co-distribution between multiple continuous or categorical variables

Summary Statistics and Contingency Table

- Custom selection of statistics
- Three different layouts for summary statistics tables

Survival Plots

- Multiple estimators
- Number at risk and censoring panels
- Stratification by multiple continuous and categorical predictors with calculation of p-value



Regression, data binning, and more to come...

SUMMARY

Explore and ExploreLive naturally extend the functionality of the KIWI platform and provide powerful and user-friendly tools for data exploration. The library of displays provided by Explore enables busy scientists and those with no programming experience to quickly generate meaningful tables and plots to support their pharmacometric needs. By leveraging the infrastructure and features previously established, Explore maintains consistency and reproducibility in the design of tables and plots across projects and between exploratory and model-diagnostic displays. ExploreLive offers a convenient environment for quick data exploration and investigation of data patterns.