PhUSE 2017

Paper AD03

EXPLORT, A Shiny App to EXPLORE data interactively and EXPORT the customized graphics as an R script

Claire Petry, F. Hoffman-La Roche Ltd, Basel, Switzerland

ABSTRACT

The *explort* R Shiny application was implemented with two main objectives. Firstly, a graphical interface was designed to create graphics and explore data interactively without any programming skills. Secondly, the R code of those graphics can be exported and used to easily reproduce the plots of interest.

Concretely, the user uploads a ready to use data file; then the application proposes several tabs to create different types of plots. In each tab, the selected plot can be customized, by choosing the variables to display, editing the labels, positioning the legend, using a log scale, adding a loess curve etc. The user can also select specific points from the plot (for example outliers, extreme values…) and visualize the corresponding data in a table underneath. When the plot fits the user needs, the corresponding R code can be exported and the graphic reproduced at any time.

INTRODUCTION

explort is a R Shiny application designed to interactively create and customize common plots (using the ggplot2 library), and export the corresponding R code. With this double feature, the user can explore and review a data set, for example to identify potential data issues or outliers, and also get a ready to run R code to reproduce the plots of interest.

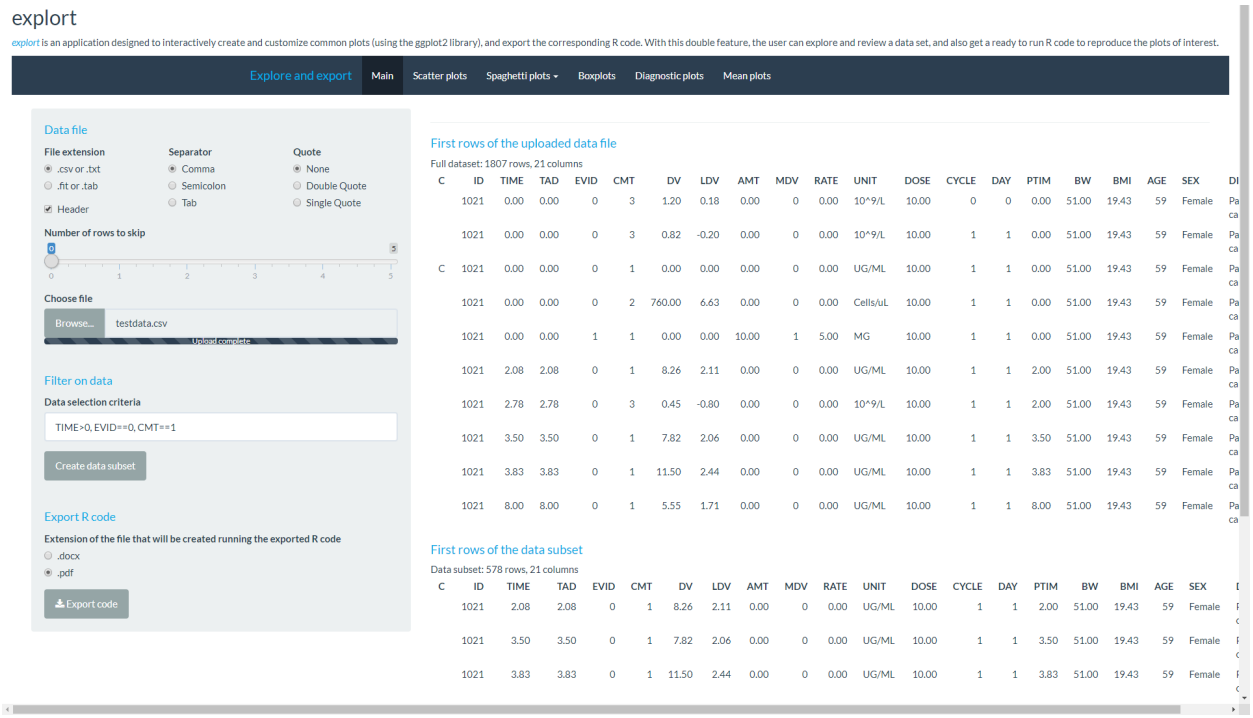
DATA EXPLORATION

When launching the application via a web browser, a page opens and proposes different tabs.

MAIN TAB

This tab allows the user to upload a data file (either a delimited text file or a NONMEM® fit file) and filter on it to retrieve the data needed for the plots. Once the data file is uploaded and the subset is created, the user can see and check the first rows of the full dataset and its subset.

This is also where the user will be able to export the R code to create the plots of interest, with the possibility to select the extension of the file that will be created running the exported R code (.pdf or .docx).



OTHER TABS

Five other tabs allow the user to create and customize different kinds of plots, using some specific features which are proposed in the left panel of each tab.

The graphic being created is simultaneously updated in the right panel according to the parameters defined by the user.

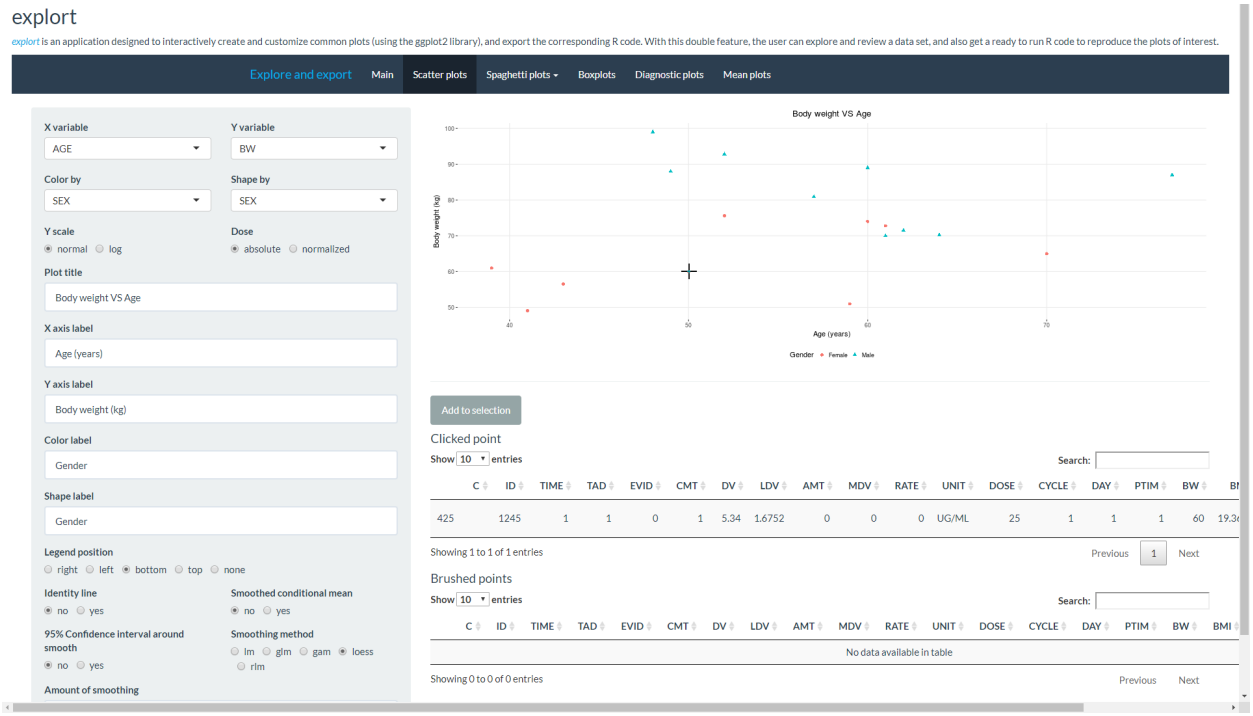
The user can also interact with some of the plots by clicking on them, in order to identify some specific points (like outliers for example). Then the corresponding data are displayed in a table below the plot in the right panel.

In the following sections, an overview of each of those tabs is presented.

SCATTER PLOTS TAB

This tab allows the user to create and customize scatter plots.

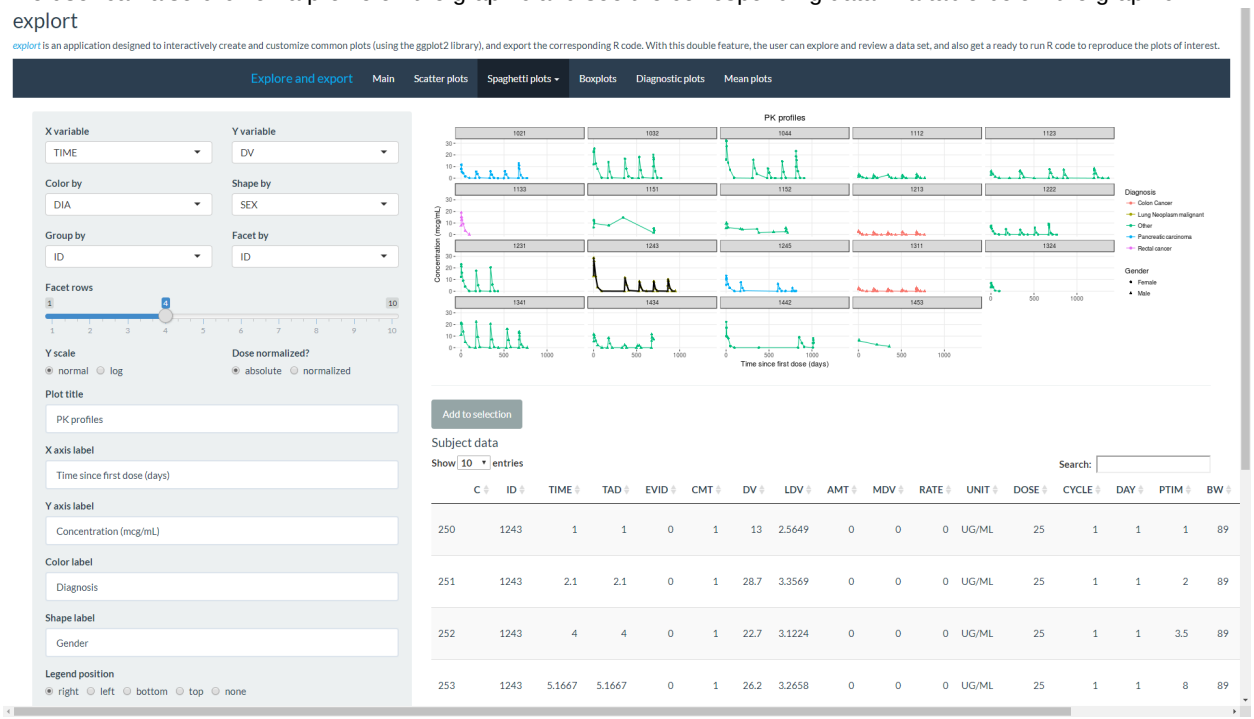
The user can also click on a specific point or select a group of points on the graphic and see the corresponding data in a table below the graphic.



SPAGHETTI PLOTS TAB

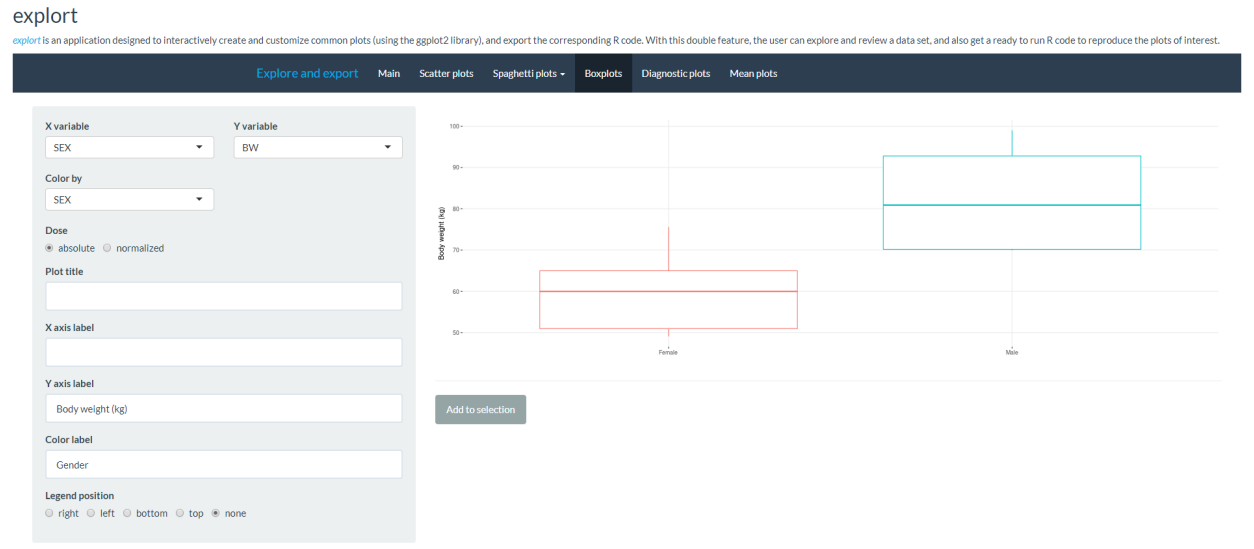
This tab allows the user to create and customize spaghetti plots.

The user can also click on a profile on the graphic and see the corresponding data in a table below the graphic.



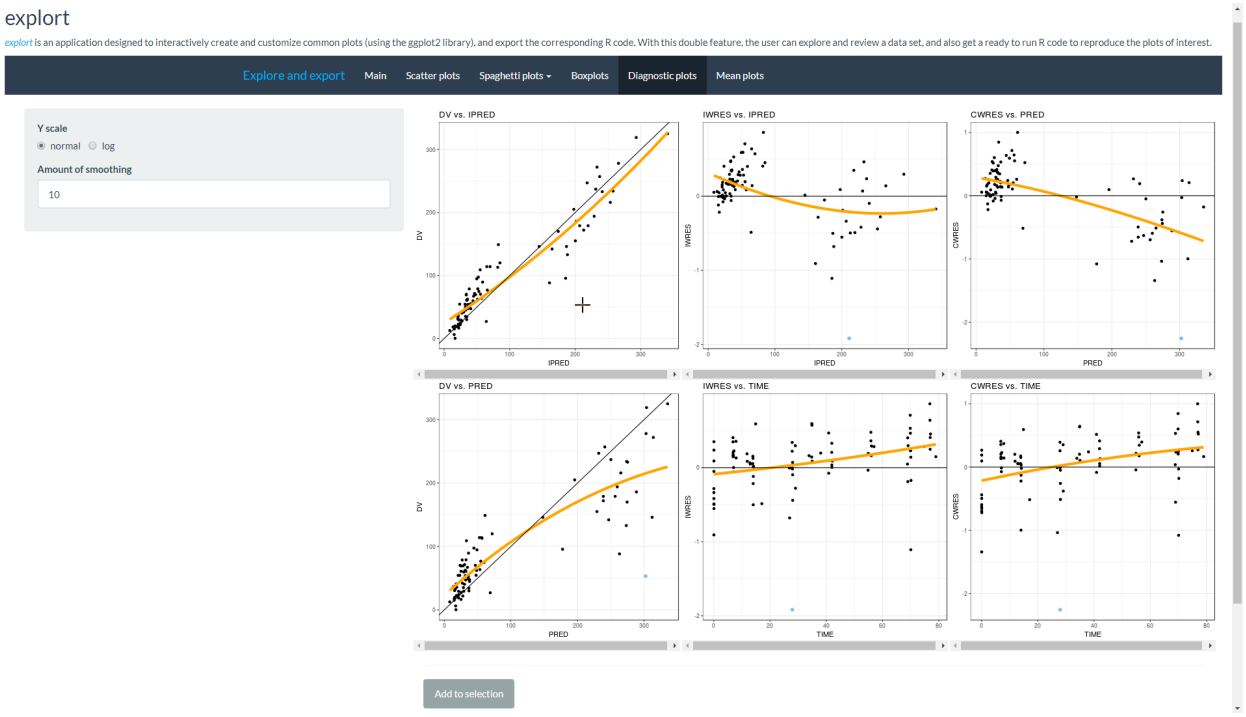
BOXPLOTS TAB

This tab allows the user to create and customize boxplots.



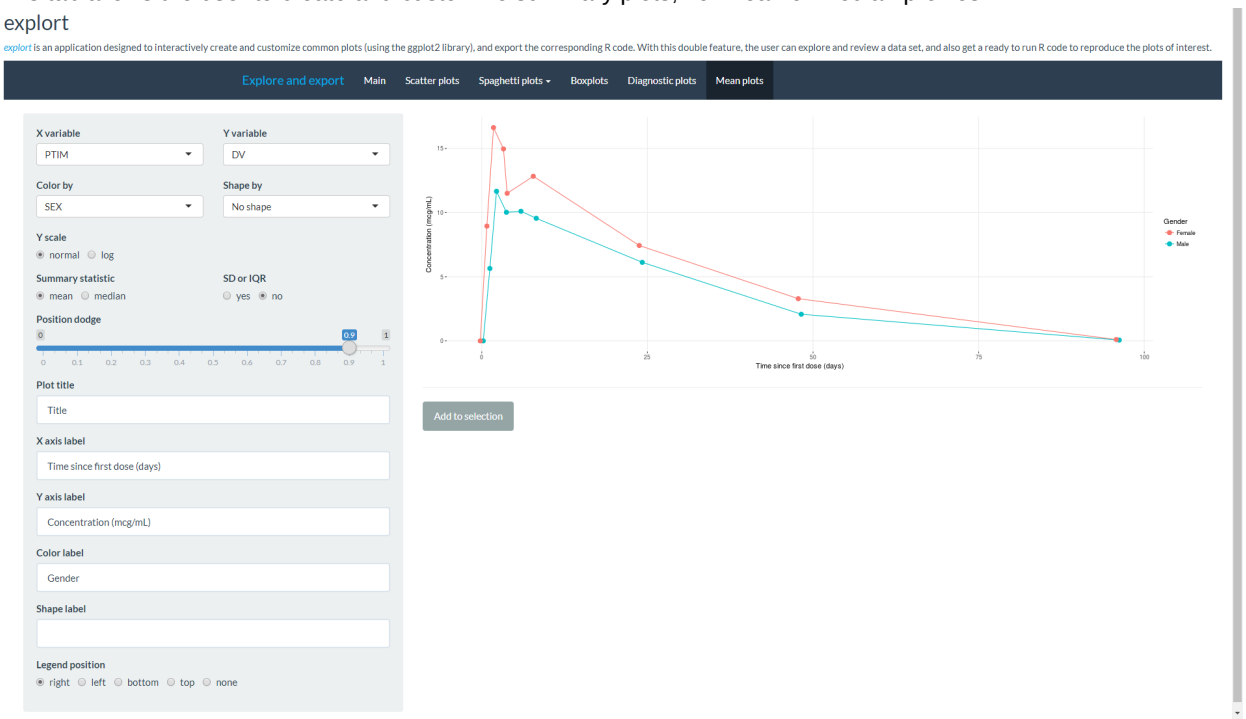
DIAGNOSTIC PLOTS TAB

This tab allows the user to create and customize diagnostic plots based on a NONMEM fit file.
The user can also click on a point on one of the diagnostic plots and see the corresponding data in a table below the graphic. The same point is highlighted in the other diagnostic plots.



MEAN PLOTS TAB

This tab allows the user to create and customize summary plots, i.e. mean or median profiles.



AVAILABLE FEATURES

As shown in the previous sections, in each tab, the user can customize the plot being created by playing with different parameters. The table below summarizes the available features for each type of plots.

	Scatter plots	Spaghetti plots	Boxplots	Diagnostic plots	Mean plots
X and Y variables	Y	Y	Y		Y
Color / shape variables	Y	Y	Y		Y
Faceting variable(s)		Y			
Y scale (normal or log)	Y	Y		Y	Y
Dose normalization	Y	Y	Y		
Plot title and labels	Y	Y	Y		Y
Legend position	Y	Y	Y		Y
Identity line	Y				
Smooth line	Y			Y	
Mean / median line					Y
Interactivity	Y	Y		Y	

CODE EXPORTATION

When the plots of interest have been identified, the user can export the corresponding R code and reproduce the graphics at any time.

PLOTS SELECTION

While creating exploratory plots with the application, the user can, at any time, select the graphics that need to be saved (using the “Add to selection” button below the plot) and continue the graphical exploration.

Once the data exploration phase is complete, the user can export the R code of all selected graphics, using the “Export R code” section in the main tab.

PLOTS REPRODUCTION

Reproducibility of the selected graphics is possible by simply running the exported R script. The plots are then created in the suitable format (either .pdf or .docx), as defined in the main tab before the R code exportation.

CONCLUSION

The *explort* Shiny application eases interactive data exploration and review for the scientists. The feature to export R code allows the scientists with no specific programming skills to create efficient plots using ggplot2. This code can be used to, either reproduce the graphics, or as a template, to create and customize additional plots. This functionality of exporting corresponding R scripts guaranties the traceability and reproducibility of the graphics produced.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Claire Petry
Roche Pharma Research & Early Development
F. Hoffman-La Roche Ltd
Grenzacherstrasse 124
CH-4070 Basel, Switzerland
+41 61 68 22 870
claire.petry@roche.com