

EZ-R4Excel: How to do powerful data analysis easily in an EXCEL Environment

Jyotsna Kasturi, Davit Sargsyan, Mariusz Lubomirski, Bill Pikounis and Dhammika Amaratunga

Non-Clinical Statistics & Computing, Janssen R&D, United States, Raritan NJ



Abstract

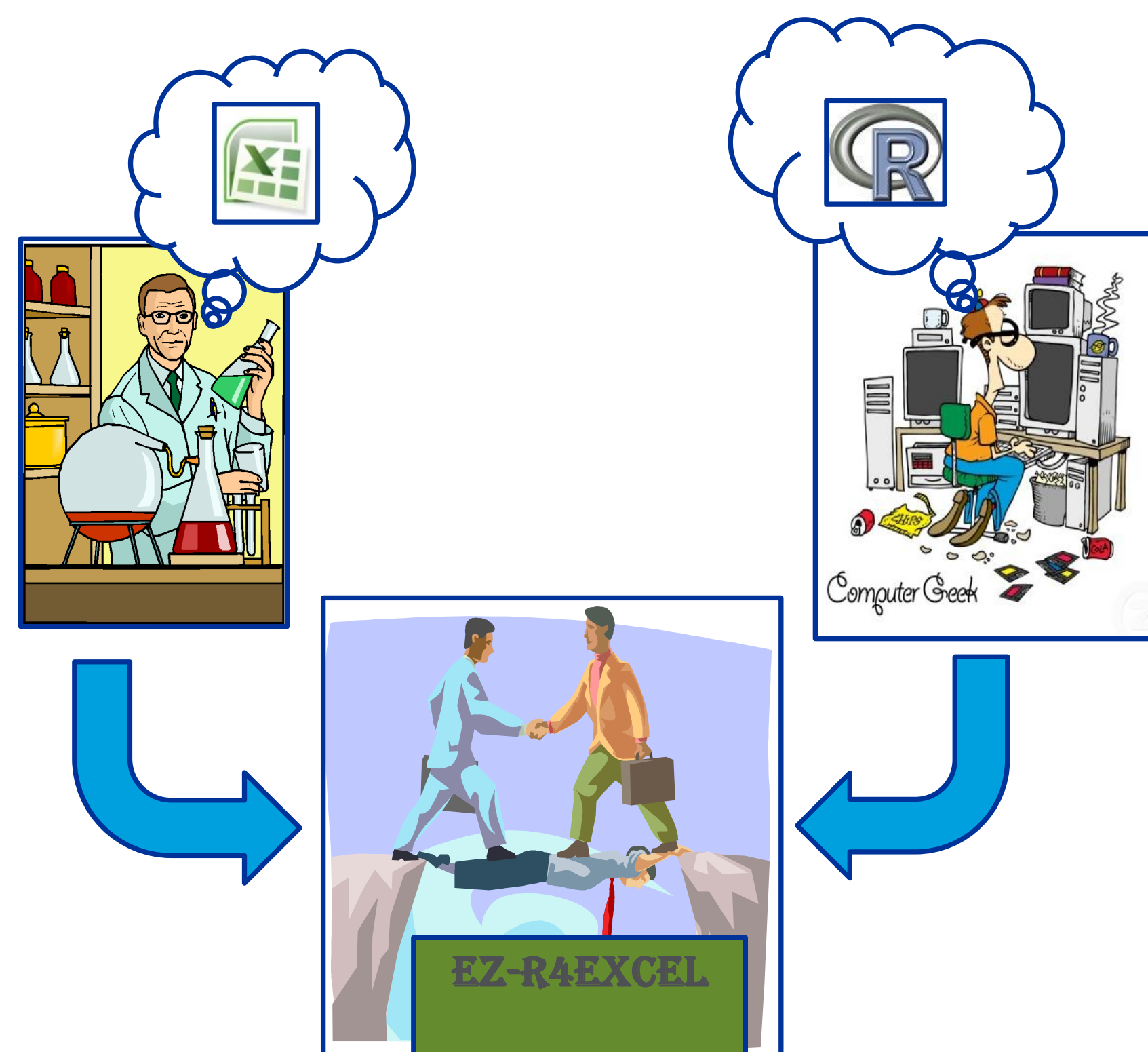
Microsoft Excel has become the industry-wide standard spreadsheet software for data storage, management and analysis used by individuals and large corporations. In particular it is used by the pharmaceutical industry for statistical analysis, especially for non-GXP applications.

We present EZ-R4Excel, a novel user-friendly software that extends Excel's statistical functionality with the power of the statistical programming language R. It provides an extensible framework incorporating powerful statistical tools, graphical capabilities and customized data analysis methodologies with the easy-to-use familiar Excel user interface.

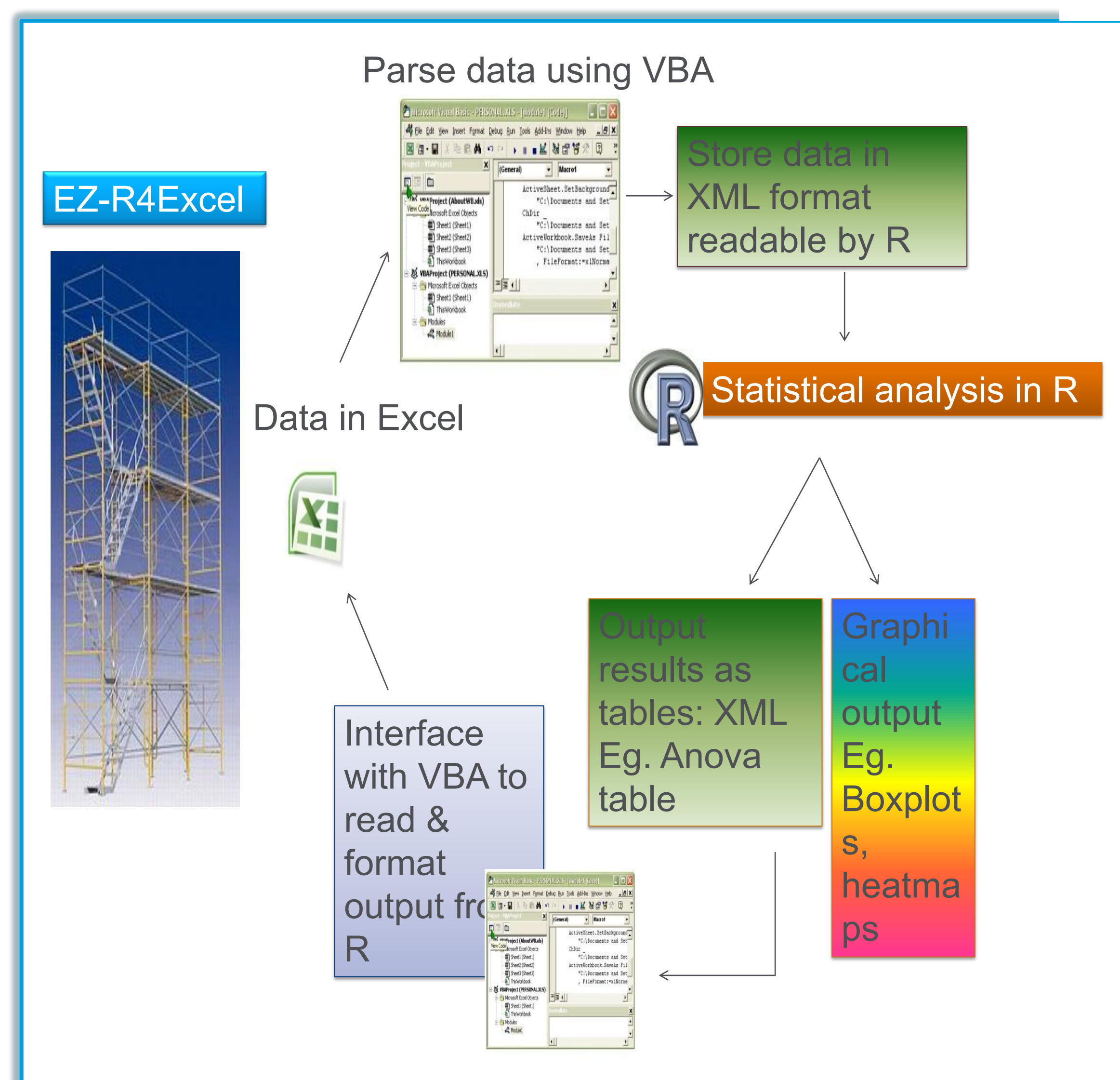
EZ-R4Excel provides methods for exploratory analyses and advanced graphical features. It has been successfully adopted by various teams within J&J with applications in *in-vivo* pharmacology, *in-vitro* screening, IC50 calculations and modeling, etc.

Data analysis in Excel

- ✓ Excel is the most widely used software application for data storage and data manipulation.
- ✓ Excel is a convenient environment to use.
- ✗ Major drawback – poor visualization and lack of comprehensive statistical analysis tools.



EZ-R4EXCEL: A PLUG & PLAY FRAMEWORK FOR DATA ANALYSIS IN EXCEL

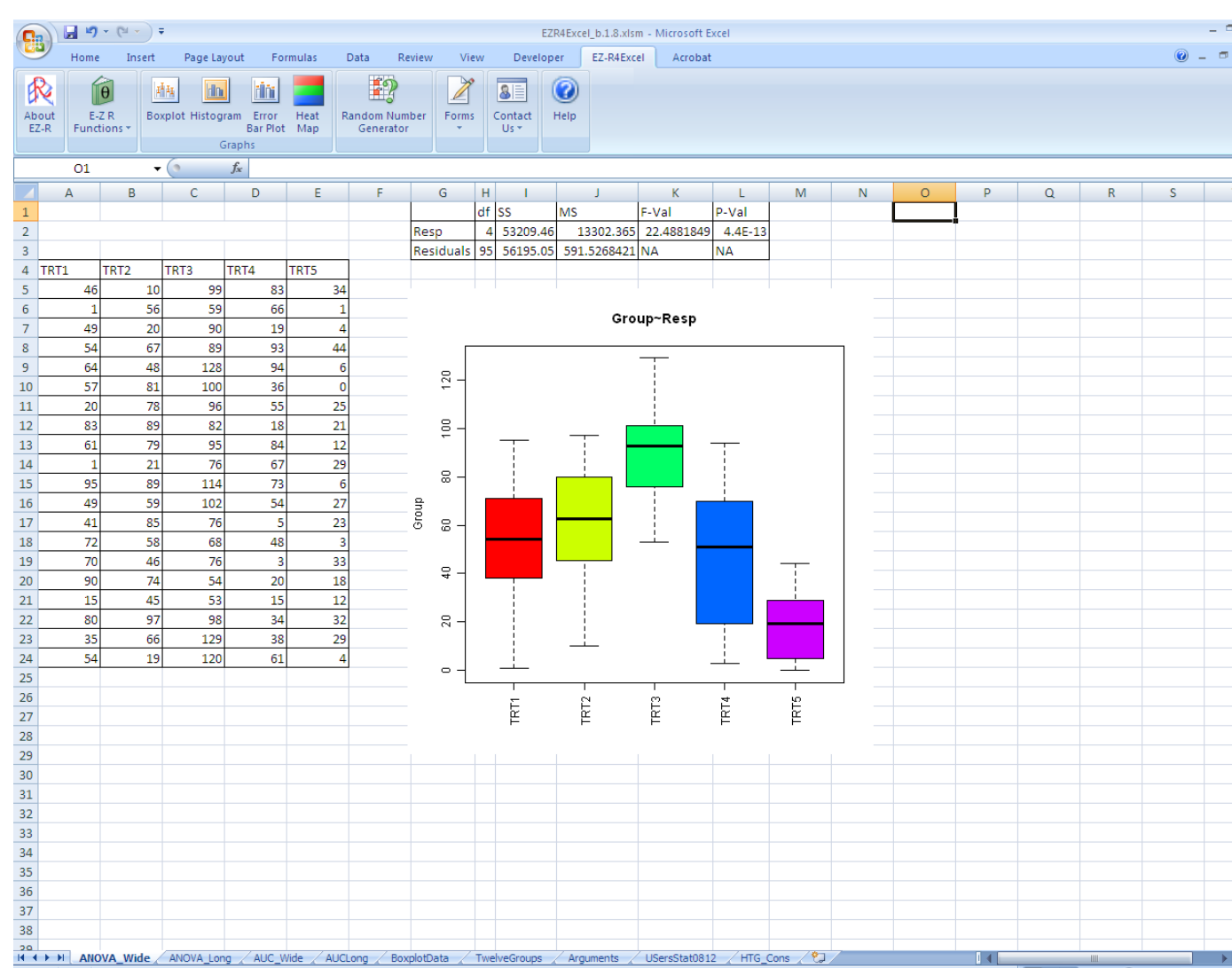


Advantages of EZ-R4Excel

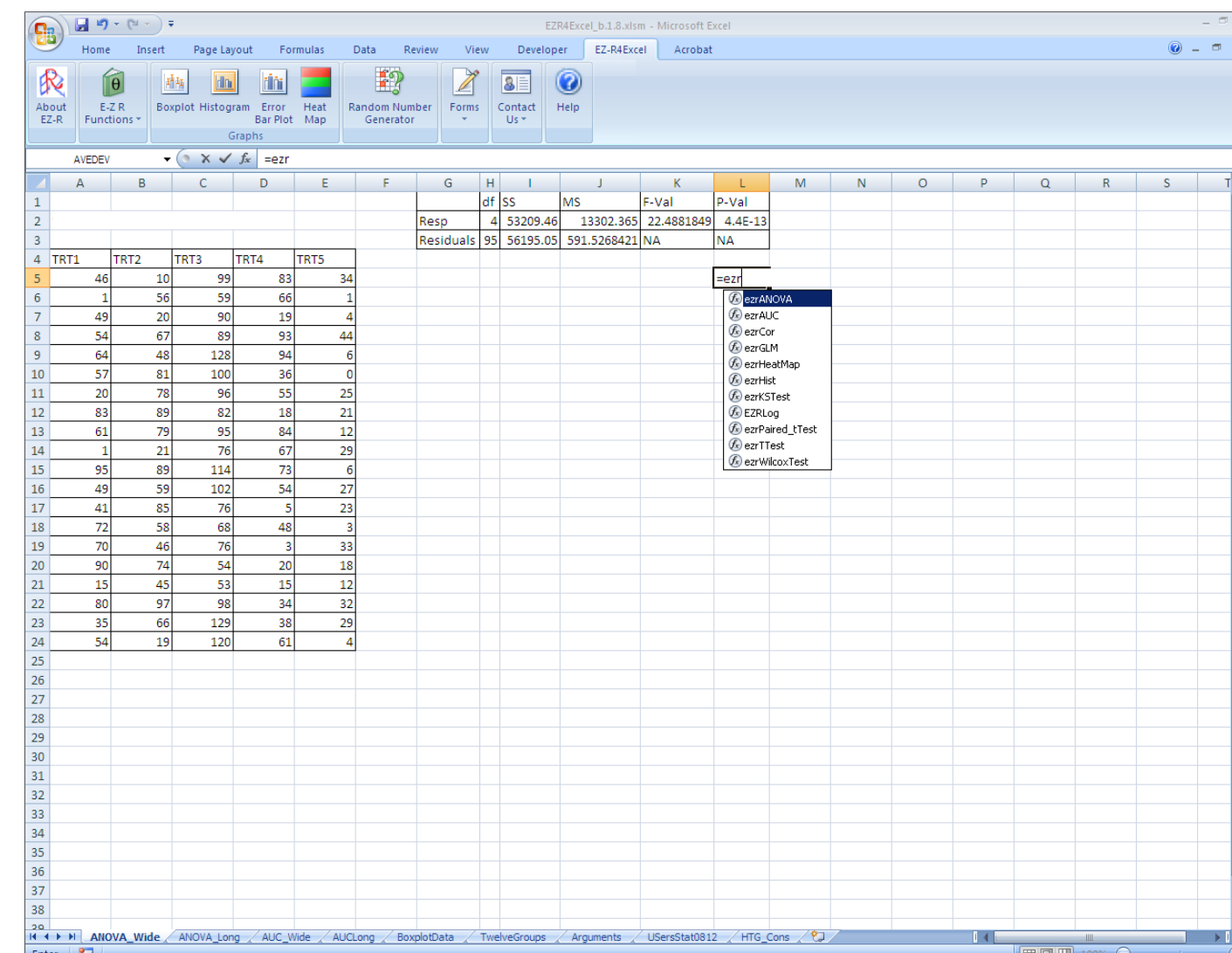
- ✓ User is the owner of locally installed application.
- ✓ Seamlessly integrated within Excel.
- ✓ Looks and feels exactly like Excel.
- ✓ User friendly and quick exploratory analysis.
- ✓ Create customized menus for special applications.
- ✓ Eliminate manual data formatting and copy/pasting.
- ✓ Short learning curve.
- ✓ Methods accessible as menus and formulae.
- ✓ Standardize analysis methods across experiments.
- ✓ Improve Efficiency and Quality of analysis.
- ✓ Traceability of results is easy.

General & Customized Statistical Methods and Graphical Capabilities

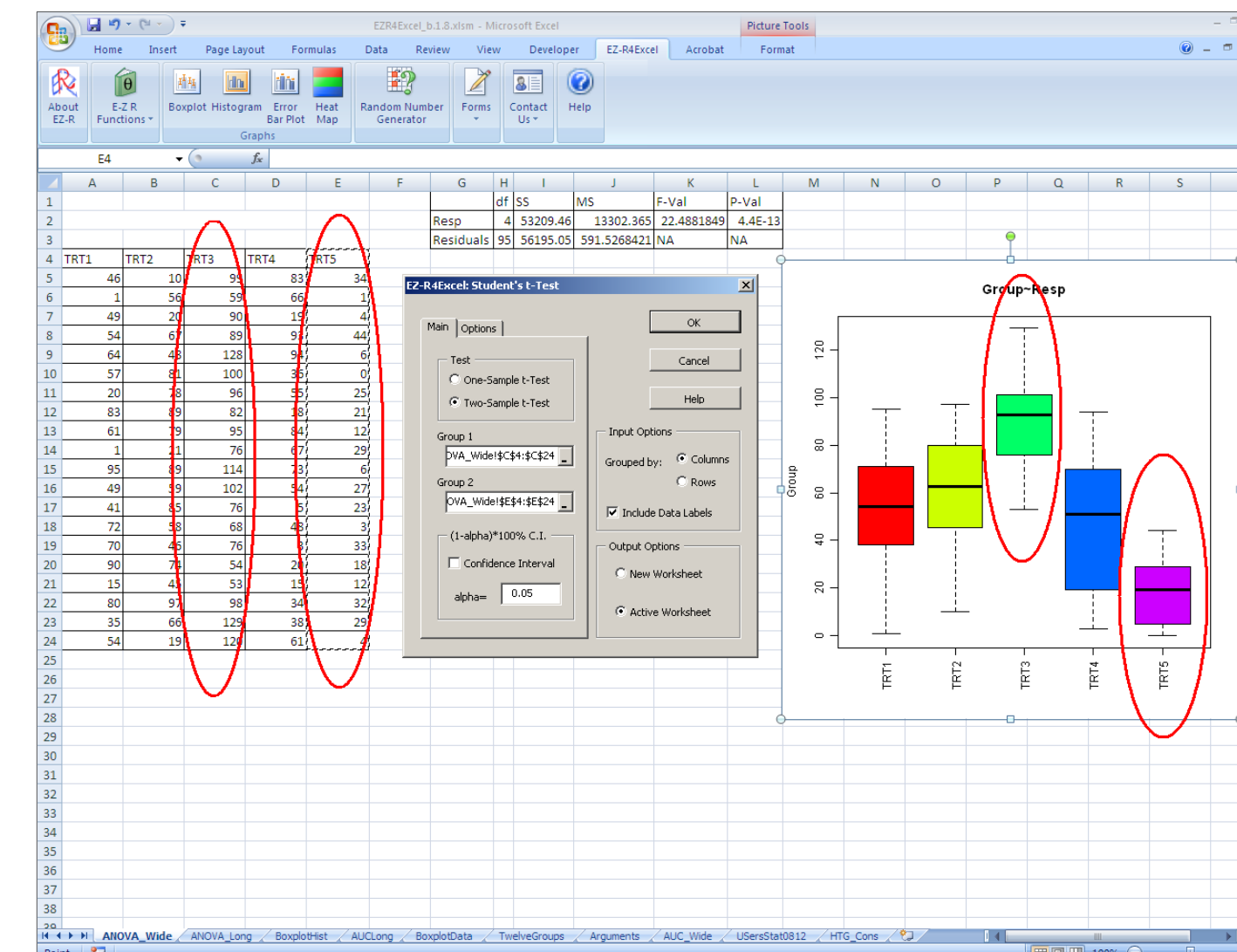
Anova and Boxplots



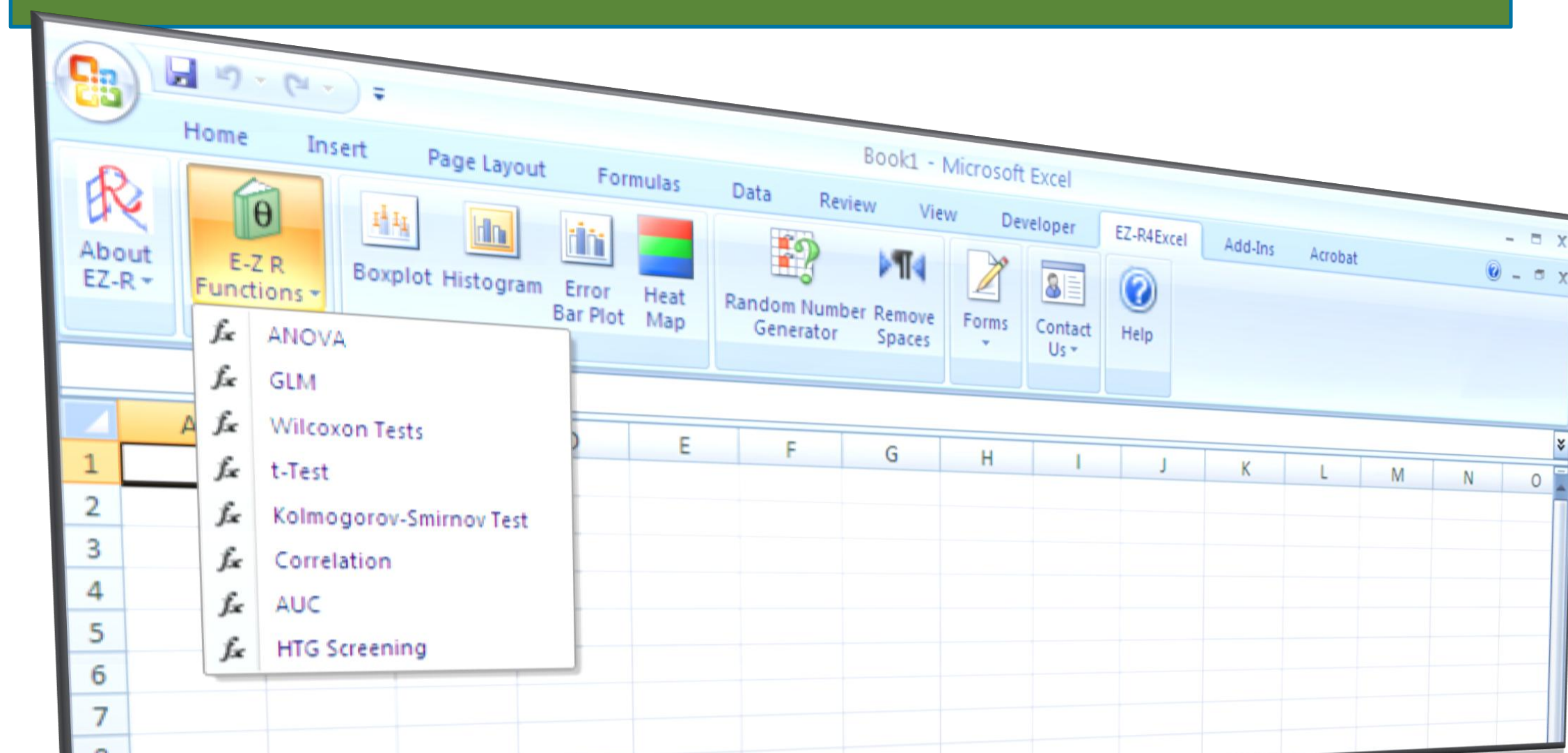
Anova from Cell



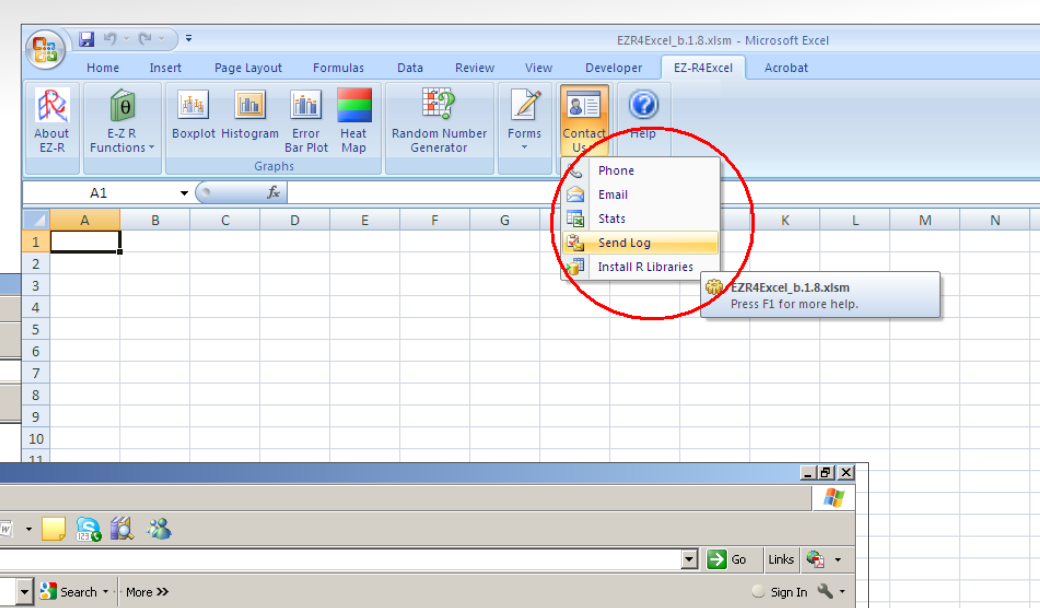
T-tests and Boxplots



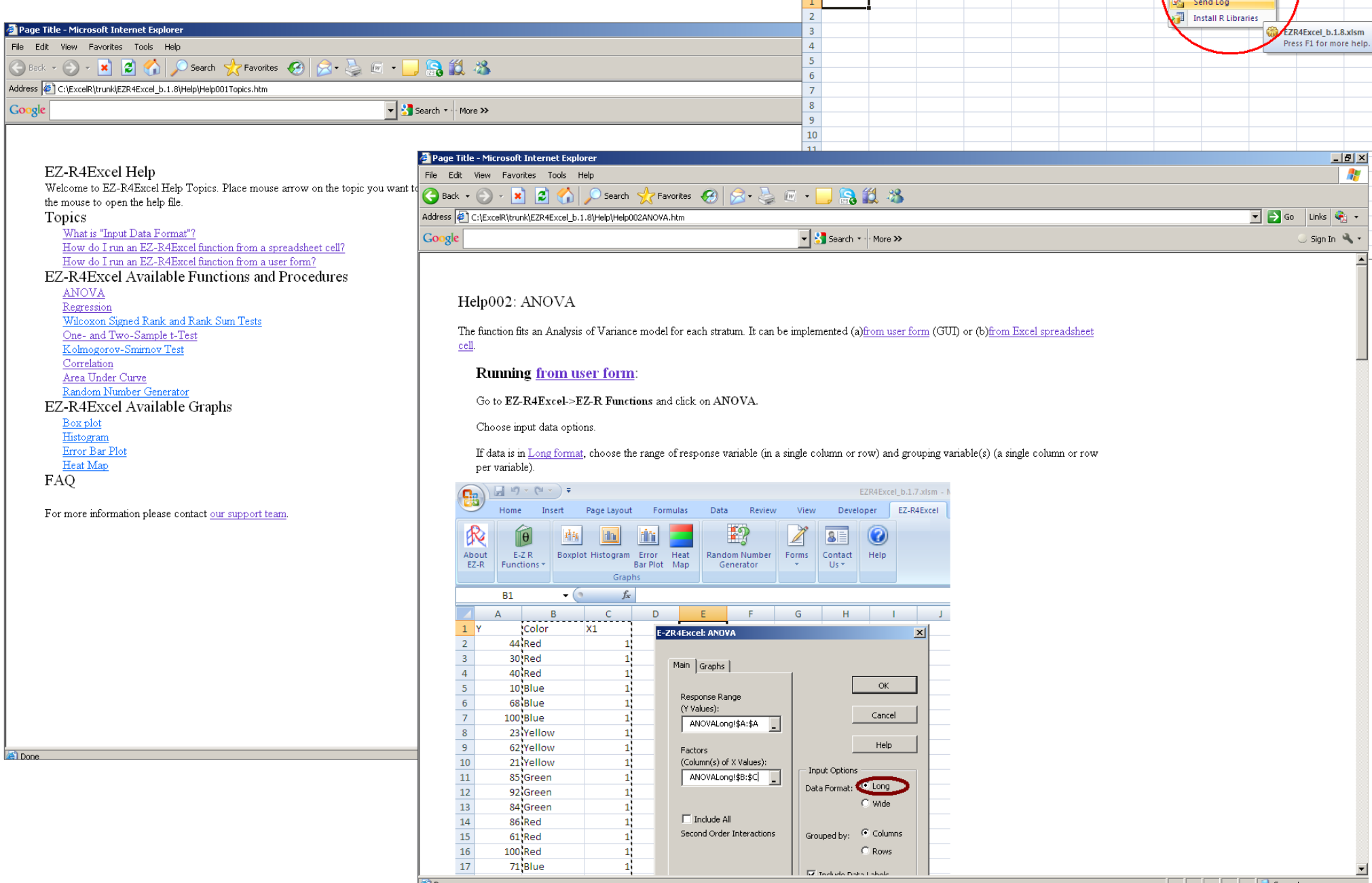
EZ-R4Excel AddIn Features



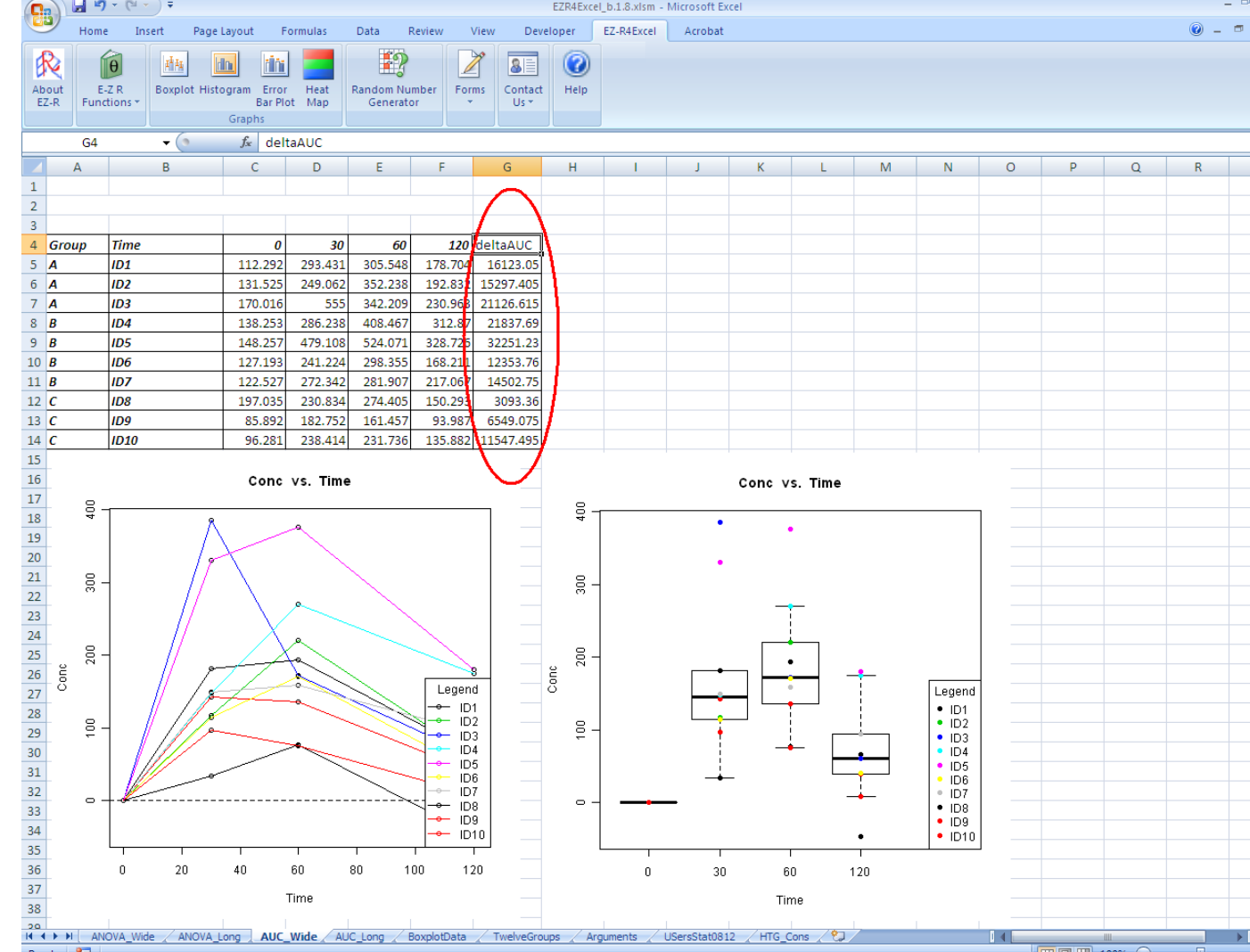
Troubleshooting



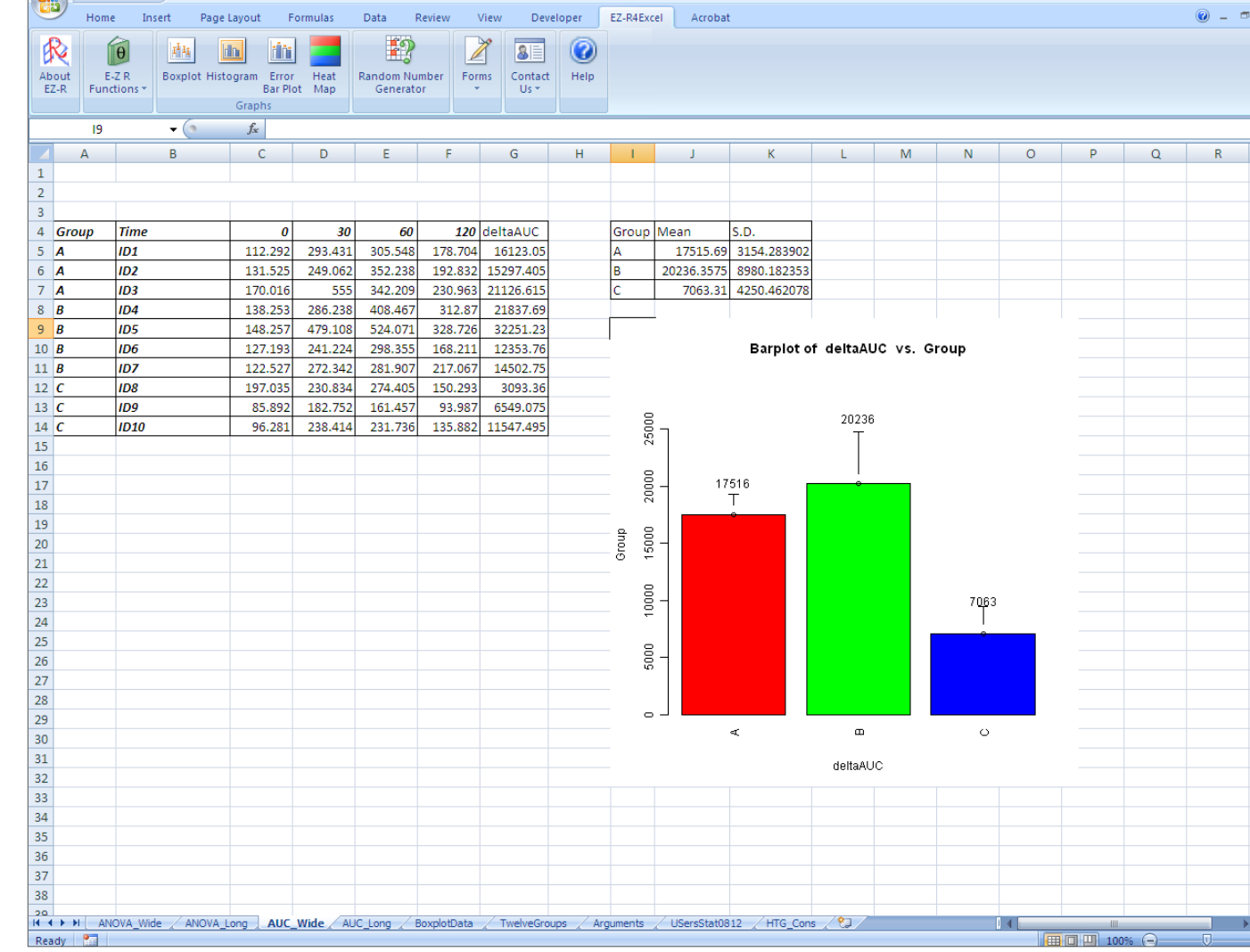
Help Page Utility



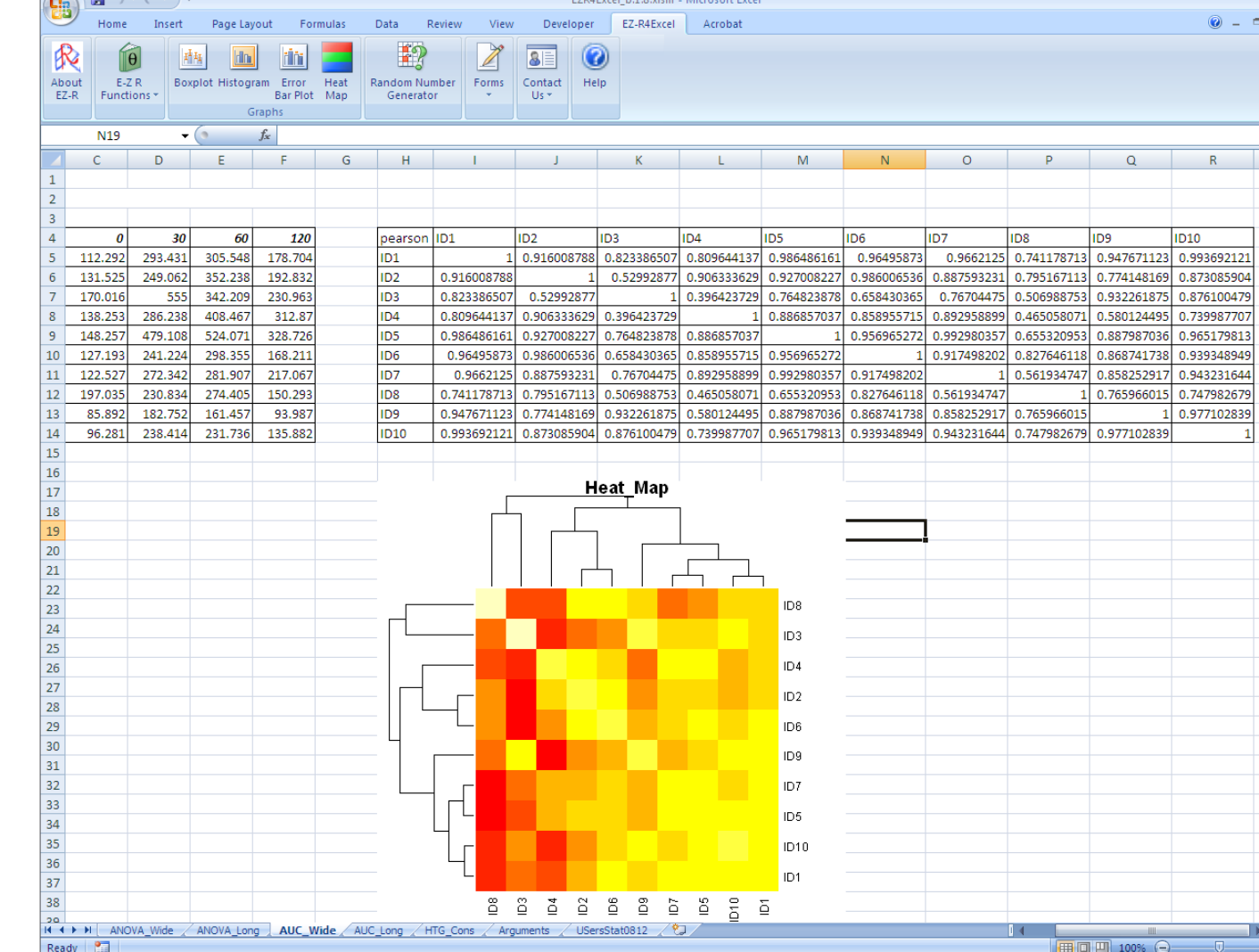
Area under the Curve



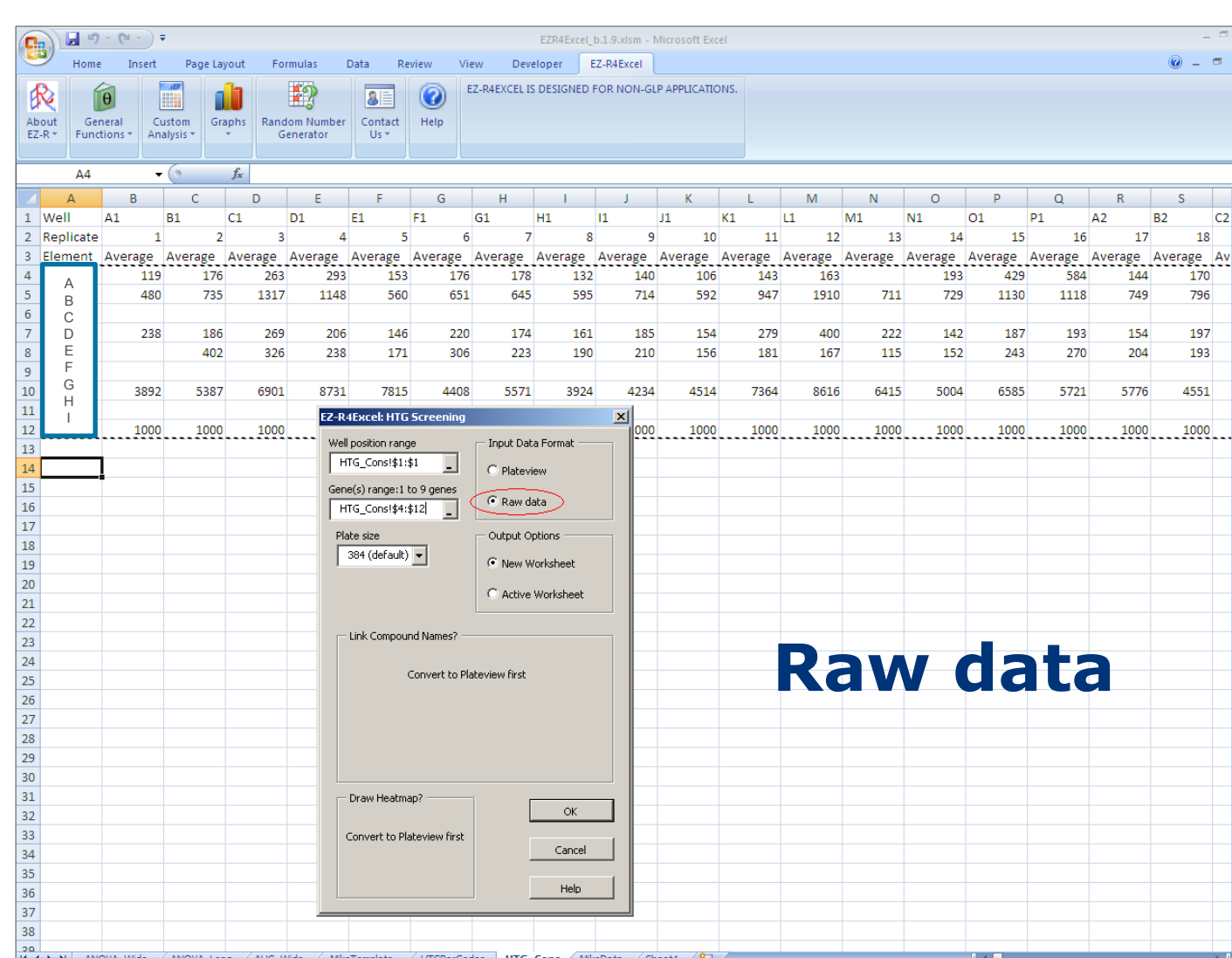
Error Bar Plot



Correlation and Heatmaps



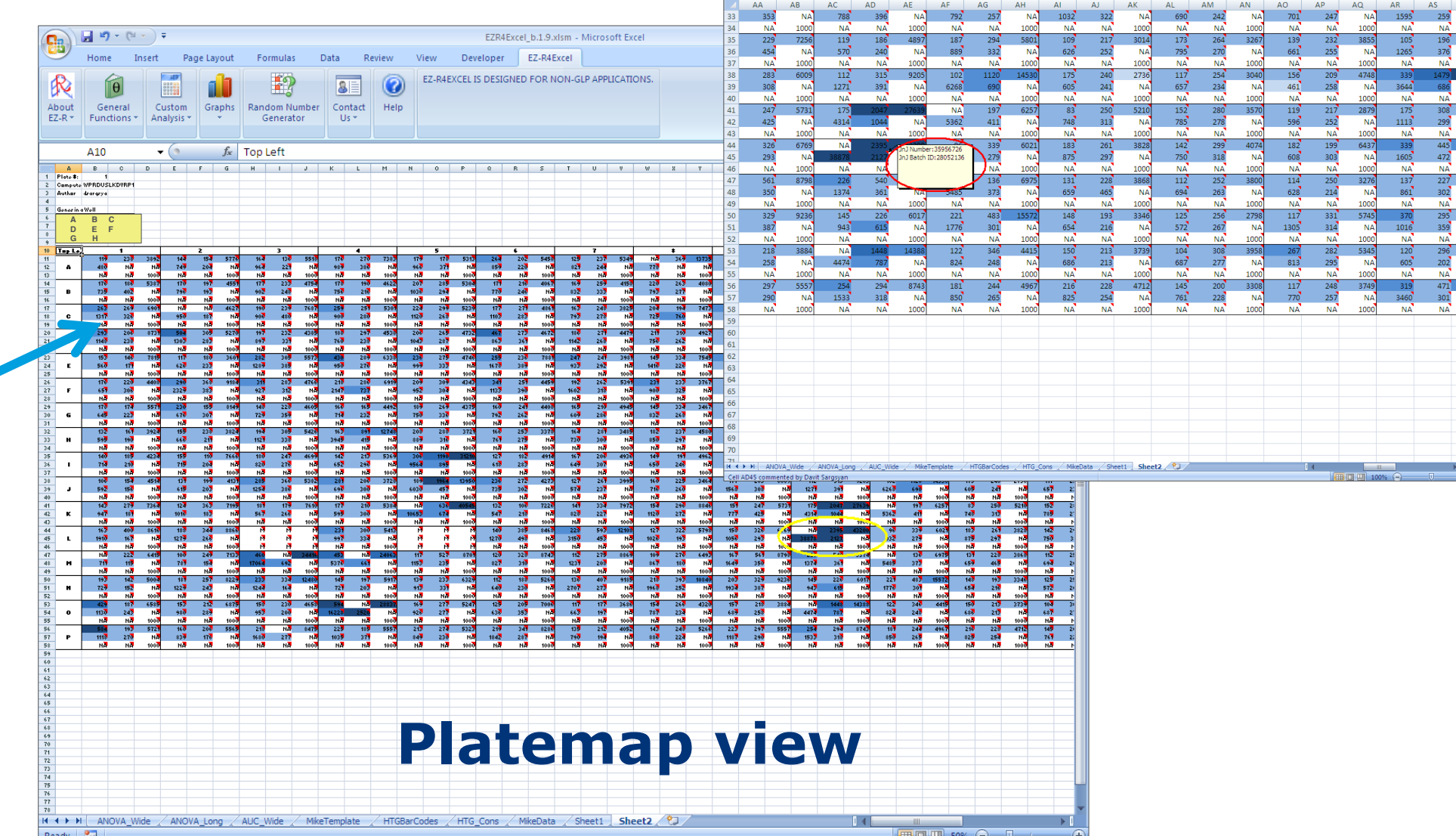
Custom PlateView Visualization for High-Throughput Screening



Raw data

Compound Info

Sample ID	Conc	Time	AUC	Group
000001	0.0001	0.0001	0.0001	0.0001
000002	0.0001	0.0001	0.0001	0.0001
000003	0.0001	0.0001	0.0001	0.0001
000004	0.0001	0.0001	0.0001	0.0001
000005	0.0001	0.0001	0.0001	0.0001
000006	0.0001	0.0001	0.0001	0.0001
000007	0.0001	0.0001	0.0001	0.0001
000008	0.0001	0.0001	0.0001	0.0001
000009	0.0001	0.0001	0.0001	0.0001
000010	0.0001	0.0001	0.0001	0.0001



Platemap view