Second Evaluation Documentation

Overview

Documentation for work done in the project “Statistical Analysis in Labplot” during GSoC from First Evaluation till Second Evaluation (June 28, 2019 - July 22, 2019).

List of Added Features

- **ANOVA:** Two Way Anova
- **Correlation Coefficient:**
  - Pearson’s R
  - Kendall’s Tau
  - Spearman Rank
- **Unit Testing:**
  - T-Test
    - Two Sample Independent
    - Two Sample Paired
    - One Sample
  - ANOVA
    - One Way
    - Two Way
  - Correlation Coefficient
    - Pearson’ R
    - Kendall’s Tau
    - Spearman Rank

Minor Improvements on the work done till First Evaluation:

- Improved syntax and coding styles
- Fixed Crash issues for boundary cases on Levene's Test, T-Test and ANOVA
- Improved Hypothesis Test Dock Initialization
- Removed HTML Tags from i18n() function calls
- Passing column pointers along with names in variables' combo-boxes in Hypothesis Test and Correlation Coefficient Docks. This increases efficiency as well as the robustness

Base Class GeneralTest

Base class GeneralTest is created for HypothesisTest and CorrelationCoefficient to avoid repetition of code, simplify maintenance of code and to make the code look cleaner.
Similarly, GeneralTestView class is created as a base class for HypothesisTestView and CorrelationCoefficientView.

Summary Table for Two Way ANOVA

The summary table needs to show data with hierarchical headers. The new helper function, "getHtmlTable3", is created for this purpose: It takes arguments, row-span, column-span and data as its arguments and intelligently creates the table using HTML and CSS formatting. This function will also help in creating a contingency table for many upcoming features like the Chi-Square Test for Independence, Phi and Cramer's V.

Here is one example of a contingency table created using this function: Summary Table for Two Way ANOVA
The summary tables in GeneralTestView lacks tooltips. There was no direct way to set separate tooltips for different words in the summary table. The “getHtmlTable3” takes the tooltip as an optional argument and then maps the words with tooltip passed. This mapping is then used to show the tooltip whenever the user clicks on that word.
Here is one example of this function in action showing a tooltip: Tool Tip Example

Backend For Correlation Coefficient and Two-Way Anova

Backends for Two-Way ANOVA, Pearson’s R Correlation Coefficient and Spearman Rank Correlation Coefficient Tests are coded to perform with O(n) complexity (where n is the total number of data values, irrespective of the number of rows and columns selected).
For Kendall's Tau Correlation Coefficient, the backend was coded to perform in O(nlogn) using Knight’s Algorithm described here. The straight forward implementation would have taken $O(n^2)$ complexity.

Unit Testing

The data and results for unit testing were taken from various sources available on the internet. The link for each source is given as a comment in the unit testing framework. Some of the results were calculated using online calculators to get higher precision and to decrease the relative error. The results were

rechecked manually. Currently, the relative error used is at least 1.e-5.

TODO

All minor TODOs are added as comments in source files itself.

Screenshots

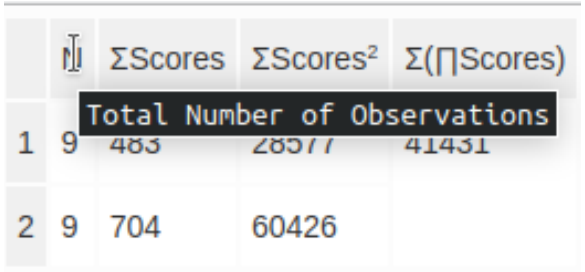
Summary Table for Two Way ANOVA

Two Way Anova

Contingency Table

	cold		warm		hot		Mean
	Mean	Replicate	Mean	Replicate	Mean	Replicate	
Super	5	4	9	4	10.5	4	8.167
Best	5	4	13	4	12	4	10
Mean	5		11		11.25		9.083

Tool Tip Example



Dock Widget of Two Way ANOVA

Summary and Results Table: Two Way ANOVA

Two Way Anova

Contingency Table

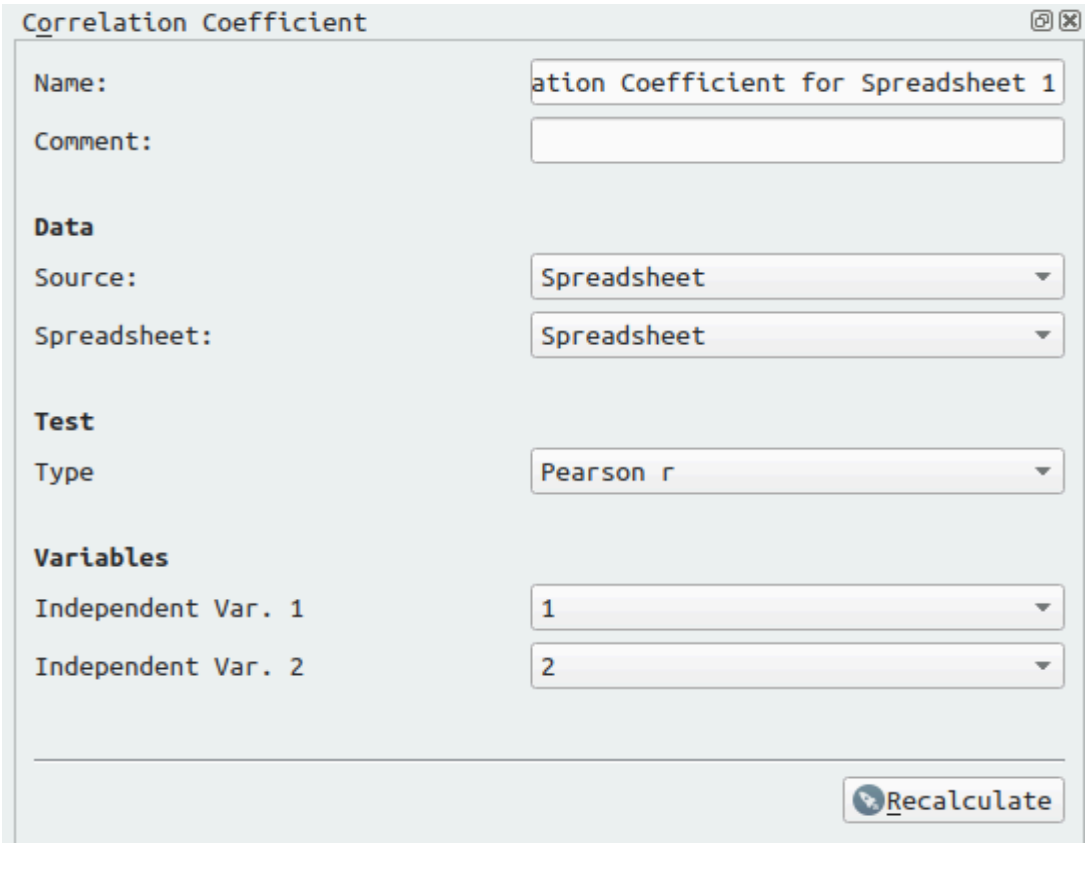
	cold		warm		hot		Mean
	Mean	Replicate	Mean	Replicate	Mean	Replicate	
Super	5	4	9	4	10.5	4	8.167
Best	5	4	13	4	12	4	10
Mean	5		11		11.25		9.083

results table

	SS	DF	MS
Detertgent	20.167	1	20.167

$F(df_{Detertgent}, df_{within})$ is 9.811
$F(df_{Temperature}, df_{within})$ is 48.73
$F(df_{interaction}, df_{within})$ is 3.973

Correlation Coefficient Dock Widget



Summary and Results View: Pearson’s R Correlation Test

Pearson's r Correlation Test

	N	ΣScores	ΣScores²	Σ(∏Scores)
1	9	483	28577	41431
2	9	704	60426	

Correlation Value is 0.968

Result View: Kendall’s Tau

Kendall's Rank Correlation Test

Number of Discordants are 11
Number of Concordant are 34
Tau a is 0.511
Z Value is 2.057