AssumpSure: a user-friendly R Shiny package for automated validation of statistical assumptions and appropriate test selection

Ahmed Bargheet [1]¶

1 Host-Microbe Interactions Research Group, Department of Medical Biology, UiT The Arctic University of Norway, Tromsø, Norway ¶ Corresponding author

Summary

AssumpSure is an open-source R Shiny package that enforces mandatory assumption checks to improve statistical rigor and guide appropriate test selection. It supports a wide array of statistical tests. The app automatically evaluates key assumptions such as normality (Shapiro–Wilk) and homogeneity of variance (Levene's test), using color-coded indicators (green for passed, red for violated) with clear interpretive summaries. When assumptions are not met, it recommends alternatives like Kruskal–Wallis for non-normal ANOVA data. *AssumpSure* also prevents invalid variable types, provides preprocessing tools (log, Box–Cox, centered log-ratio transformations), and generates downloadable diagnostic plots and publication-ready outputs.

Statement of Need

Rigorous statistical analysis depends on checking underlying assumptions and choosing appropriate tests. Yet in practice, these steps are often skipped, misunderstood, or incorrectly applied. Reviews of social science literature report that fewer than 25% of studies explicitly mention assumption checks, and even fewer report them correctly (Shatz, 2024). This contributes to widespread misuse of statistical tests and undermines reproducibility. One common issue is the misinterpretation of assumption tests. For example, many users mistakenly interpret a significant Shapiro–Wilk test result ($p < 0.05$) as evidence for normality rather than deviation from it (Ghasemi & Zahediasl, 2012; Gosselin, 2024; Midway & White, 2025). Moreover, in small samples, where the Shapiro–Wilk test lacks power, borderline results are easily overlooked or misjudged.

AssumpSure directly addresses these problems by making assumption checking a required part of the analysis workflow. Through an intuitive web interface, users are guided step-by-step: assumptions are tested automatically, results are color-coded (green for met, red for violated, orange for borderline), and clear recommendations are given. For example, if ANOVA assumptions are not met, Kruskal–Wallis is suggested. By complementing color-coded squares with tooltips and clear messages, *AssumpSure* helps reduce misinterpretation of common assumption tests and guides the user on how to proceed.

Comparison to other Packages

Tools like *Jamovi* (Şahin & Aybek, 2019) and *JASP* (Love et al., 2019) offer user-friendly interfaces but lack integrated workflows that enforce mandatory assumption diagnostics, appropriate test selection, and actionable recommendations, leaving users to manually interpret

complex outputs. This increases errors and undermines reproducibility. Additionally, non-R applications such as assumption-sheriff (Python) ([Root, 2024](#)) also aim to simplify statistical assumption checking and test selection. However, these alternatives lack fully integrated, interactive diagnostic workflows or do not provide enforced guidance on test appropriateness for each step.

Software Description

AssumpSure's R Shiny architecture organizes statistical analyses into continuous variable tests, categorical tests, correlations, and regression models. Users upload CSV data, and the interface automatically detects variable types, preventing invalid selections with clear error messages (e.g., suggesting one-way ANOVA if a t-test is chosen for a categorical variable with three levels). Mandatory assumption diagnostics (normality via Shapiro-Wilk, variance homogeneity via Levene's test) are visualized through diagnostic plots (QQ plots, histograms, scatter plots), with tooltips to guide the user how to interpret the plots.

For correlations, a tooltip explains the centered log-ratio (CLR) transformation's importance, conducted by compositions R package ([Boogaart et al., 2013](#)), in reducing spurious correlations in compositional data (e.g., microbiome profiles). Prevalence filtering is available for correlations to enhance reliability. Regression models leverage the performance package ([Lüdecke et al., 2021](#)), providing transformations like log or Box-Cox that are offered to improve residual normality and reduce heteroscedasticity, with guidance on their role in ensuring model validity.

The package supports independent/paired t-tests, Welch test, one-way ANOVA, Mann-Whitney U, Wilcoxon signed-rank, Kruskal-Wallis, chi-square, Fisher's exact, linear, logistic, multinomial, Poisson, negative binomial, zero-inflated negative binomial, and mixed-effects models, plus Pearson, Spearman, Kendall, and biweight midcorrelation methods. Post-hoc comparisons (e.g., Tukey HSD, Dunn's test) and multiple testing corrections (e.g., Benjamini-Hochberg) are included where applicable. Outputs, including p-values, confidence intervals, effect sizes, and publication-ready plots, are exportable in PNG, PDF, or CSV formats for transparent reporting as suggested by researchers ([Shatz, 2024](#)). *AssumpSure* integrates R libraries like *rstatix* ([Kassambara, 2019](#)), *performance* ([Lüdecke et al., 2021](#)), *lmerTest* ([Kuznetsova et al., 2015](#)), *MASS* ([Ripley et al., 2013](#)), and *bestNormalize* ([Peterson & Peterson, 2020](#)) for robust functionality, with all used packages listed in the DESCRIPTION file.
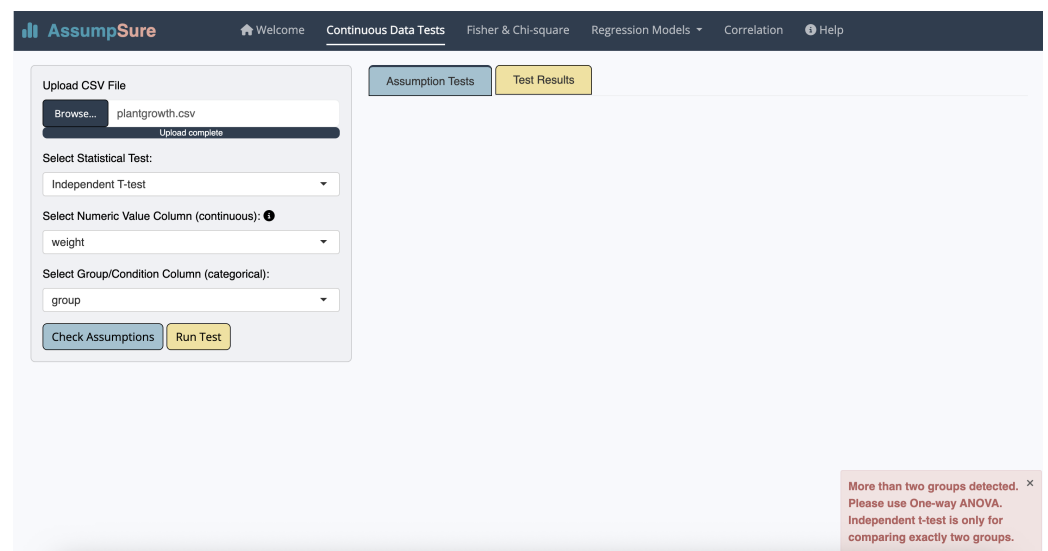

Figure 1: The workflow from data upload to test selection, guiding users to switch to the appropriate statistical test.

Illustrative Example

A researcher uploads a CSV dataset, imported using the readr package (Wickham et al., 2024) with bacterial alpha diversity (Shannon index) across three countries. This example uses the `infants.csv` dataset included with *AssumpSure*. *AssumpSure* detects variable types, restricting invalid selections (e.g., excluding count variables for ANOVA). Selecting one-way ANOVA triggers normality (Shapiro-Wilk) and homogeneity (Levene's) checks using the rstatix package (Kassambara, 2019), visualized via QQ plots and histograms with ggplot2 (Wickham, 2011) with tooltips implemented using the shiny package (Chang et al., 2015) (which provides the icon() function and HTML integration for Bootstrap tooltips). A red indicator signals non-normality and heterogeneous variances, recommending Kruskal-Wallis (Figure 2).

Assumption Test Results

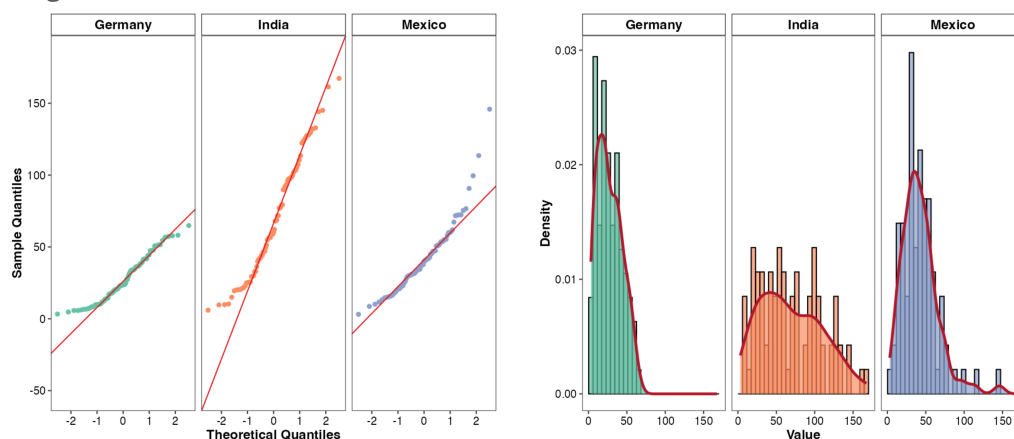### Levene's Test (Equal Variance)

df1	df2	statistic	p
2	247	32.29271	0

Variances are significantly different (Heterogeneous).

Shapiro-Wilk Test (Normality)

group	variable	statistic	p
Germany	value	0.9515254	0.0031726
India	value	0.9589311	0.0095929
Mexico	value	0.9067978	0.0000173

One or more groups deviate significantly from normality.

Diagnostic Plots

Assumption Check Summary

✗ Assumptions Not Met
Consider a non-parametric alternative: **Kruskal–Wallis test.**

Figure 2: The assumption checking report.

Bargheet. (2025). AssumpSure: a user-friendly R Shiny package for automated validation of statistical assumptions and appropriate test selection. *Journal of Open Source Software*, *10*(115), 9286. https://doi.org/10.21105/joss.09286.

Switching to this test, the user obtains significant results ($p < 0.05$) and effect size using Kruskal-Wallis eta-squared (Figure 3), also based on rstatix (Kassambara, 2019).

| Assumption Tests | Test Results | | | | |

.y.	n	statistic	df	p	method
value	250	59.77617	2	0	Kruskal-Wallis

Statistically Significant Difference

Effect size (Kruskal-Wallis eta-squared (η^2[H]))

.y.	n	effsize	method	magnitude
value	250	0.2339116	eta2[H]	large

Figure 3: Kruskal-Wallis test and effect sizes estimation using Kruskal-Wallis eta-squared.

The user can run Dunn's with p-value correction, such as Benjamini-Hochberg, based on rstatix R package (Kassambara, 2019), and download a publication-ready boxplot (Figure 4). This workflow, requiring no coding, ensures valid test selection and transparent reporting for researchers with minimal statistical expertise.
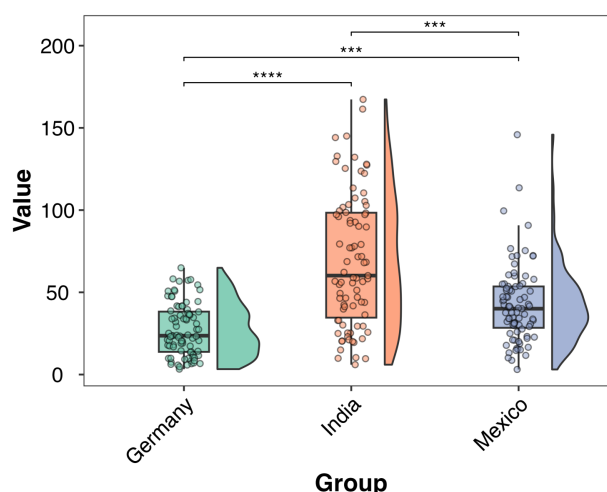
Figure 4: Dunn test and Wilcoxon r unpaired test for effect size, and publication-ready boxplot.

Example Usage and Reproducibility

AssumpSure provides example datasets for all modules. infants.csv is downloadable within the app (Help tab), and the remaining datasets are available in the GitHub repository (inst/extdata/) for manual testing.

- Continuous data tests: t-test.csv, plantgrowth.csv, and infants.csv (for independent/paired t-tests, one-way ANOVA, and other continuous analyses).
- Fisher and Chi-square tests: chi_fisher.csv and infants.csv for categorical comparisons.
- Regression models: infants.csv for linear, logistic, and other regression analyses.
- Correlation analyses: correlation.csv and bacteria_for_correlation.csv for standard and compositional correlation methods.

Limitations

AssumpSure accepts UTF-8, comma-delimited CSV files with a header row (parsed by readr package (Wickham et al., 2024). Unusual delimiters or encodings are not supported. There is no strict size limit, but large files may slow performance. *AssumpSure* visualizes assumption checks using diagnostic plots (e.g., from the *performance* package (Lüdecke et al., 2021)) and provides general guidance via tooltips, but does not generate automated, context-specific recommendations. Users should interpret plots with the provided guidance.

Licensing and Availability

AssumpSure is licensed under the GPL-3 license, with all source code stored at GitHub (https://github.com/Ahmedbargheet/AssumpSure). In the spirit of honest and open science, we encourage requests, tips for fixes, feature updates, as well as general questions and concerns via direct interaction with the developer.

Acknowledgments

AssumpSure acknowledges the open-source R community for maintaining the packages that support the application's statistical and graphical functionality.

References

Boogaart, K. G. van den, Tolosana, R., Bren, M., & Boogaart, M. K. G. van den. (2013). Package "compositions." *Compositional Data Analysis Ver*, *1*, 40–41. https://doi.org/10.32614/CRAN.package.compositions

Chang, W., Cheng, J., Allaire, J., Xie, Y., & McPherson, J. (2015). Package "shiny." *See Http://Citeseerx. Ist. Psu. Edu/Viewdoc/Download*. https://doi.org/10.32614/CRAN.package.shiny

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, *10*(2), 486. https://doi.org/10.5812/ijem.3505

Gosselin, R.-D. (2024). Testing for normality: A user's (cautionary) guide. *Laboratory Animals*, *58*(5), 433–437. https://doi.org/10.1177/00236772241276808

Kassambara, A. (2019). Rstatix: Pipe-friendly framework for basic statistical tests. *CRAN: Contributed Packages*. https://doi.org/10.32614/CRAN.package.rstatix

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B., & others. (2015). Package "lmertest." *R Package Version*, *2*(0), 734. https://doi.org/10.32614/CRAN.package.lmerTest

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Šmíra, M., Epskamp, S., & others. (2019). JASP: Graphical statistical software for common statistical designs. *Journal of Statistical Software*, *88*, 1–17. https://doi.org/10.18637/jss.v088.i02

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60). https://doi.org/10.21105/joss.03139

Midway, S., & White, J. W. (2025). Testing for normality in regression models: Mistakes abound (but may not matter). *Royal Society Open Science*, *12*(4), 241904. https://doi.org/10.1098/rsos.241904

Peterson, R. A., & Peterson, M. R. A. (2020). Package "bestNormalize." *Normalizing Transformation Functions. R Package Version*, *1*. https://doi.org/10.32614/CRAN.package.bestNormalize

Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., & Ripley, M. B. (2013). Package "mass." *Cran r*, *538*(113-120), 822. https://doi.org/10.32614/CRAN.package.MASS

Root, B. (2024). Assumption-sheriff: Automated assumption checking for statistical tests. In *Python Package Index*. PyPI. https://pypi.org/project/assumption-sheriff/

Şahin, M., & Aybek, E. (2019). Jamovi: An easy to use statistical software for the social scientists. *International Journal of Assessment Tools in Education*, *6*(4), 670–692. https://doi.org/10.21449/ijate.661803

Shatz, I. (2024). Assumption-checking rather than (just) testing: The importance of visualization and effect size in statistical diagnostics. *Behavior Research Methods*, *56*(2), 826–845. https://doi.org/10.3758/s13428-023-02072-x

Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*, *3*(2), 180–185. https://doi.org/10.1002/wics.147

Wickham, H., Hester, J., Francois, R., Bryan, J., Bearrows, S., Jylänki, J., & Jørgensen, M. (2024). Package "readr." *Read Rectangular Text Data. Available Online: Https://Cran. R-Project. Org/Web/Packages/Readr/Readr. Pdf (Accessed on 23 August 2023)*. https://doi.org/10.32614/CRAN.package.readr