

# Enhancing 'ggplot2' : The 'ggstatsplot' approach

Indrajeet Patil<sup>1</sup>

DOI:

<sup>1</sup> Center for Humans and Machines, Max Planck Institute for Human Development, Berlin, Germany

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted:

Published:

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

## Summary

**ggstatsplot** is an extension of **ggplot2** package for creating graphics with details from statistical tests included in the plots themselves and targeted primarily at behavioral sciences community to provide a one-line code to produce information-rich plots. In a typical exploratory data analysis workflow, data visualization and statistical modeling are two different phases: visualization informs modeling, and modeling in its turn can suggest a different visualization method, and so on and so forth. The central idea of **ggstatsplot** is simple: combine these two phases into one in the form of graphics with statistical details, which makes data exploration simpler and faster.

## Statement of Need

### ggstatsplot at a glance

#### Summary of types of plots included

It produces a limited kinds of ready-made plots for the supported analyses:

Function	Plot	Description
<b>ggbetweenstats</b> , <b>ggwithinstats</b>	<b>boxviolin plots</b>	for comparisons <i>between</i> and <i>within</i> groups/conditions
<b>gghistostats</b>	<b>histograms</b>	for distribution of a numeric variable
<b>ggdotplotstats</b>	<b>dot plots/charts</b>	for distribution about labeled numeric variable
<b>ggscatterstats</b>	<b>scatterplots</b>	for correlation between two variables
<b>ggcorrmatrix</b>	<b>correlation matrices</b>	for correlations between multiple variables
<b>ggpiestats</b>	<b>pie charts</b>	for categorical data
<b>ggbarstats</b>	<b>bar charts</b>	for categorical data
<b>ggcoefstats</b>	<b>dot-and-whisker plots</b>	for regression models and meta-analysis

In addition to these basic plots, **ggstatsplot** also provides **grouped\_** versions (see below) that makes it easy to repeat the same analysis for any grouping variable.

#### Summary of types of statistical analyses

Most functions provide a **type** (of test) argument that is helpful to specify the type of statistical approaches:

- "parametric" (for **parametric** statistics)
- "nonparametric" (for **non-parametric** statistics)

- "robust" (for **robust** statistics)
- "bayes" (for **Bayesian** statistics)

In the following sections, we will discuss at depth justification for why the plots have been designed in certain ways and what principles were followed to report statistical details on the plots.

## Licensing and Availability

`ggstatsplot` is licensed under the GNU General Public License (v3.0), with all source code stored at [GitHub](#), and with a corresponding issue tracker for bug reporting and feature enhancements. 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 contributors and developers, by [filing an issue](#). See the package's [Contribution Guidelines](#).

## Acknowledgements

I would like to acknowledge the support of Mina Cikara, Fiery Cushman, and Iyad Rahwan during the development of this project. `ggstatsplot` relies heavily on the [easystats](#) ecosystem, a collaborative project created to facilitate the usage of R for statistical analyses. Thus, I would like to thank the [members of easystats](#) as well as the users.

## References