grapesAgri1: Collection of Shiny Apps for Data Analysis in Agriculture

Pratheesh P. Gopinath[1], **Rajender Parsad**[2], **Brigit Joseph**[1], **and Adarsh V. S.**[3]

1 Department of Agricultural Statistics, College of Agriculture, Vellayani, Kerala Agricultural University, Kerala, India. **2** ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India. **3** Department of Agricultural Statistics, BCKV, West Bengal, India

Summary

Agricultural experiments demand a wide range of statistical tools for analysis, which includes exploratory analysis, design of experiments, and statistical genetics. It is a challenge for scientists and students to find a suitable platform for data analysis and publish the research outputs in quality journals. Most of the software available for data analysis are proprietary or lack a simple user interface, for example SAS® is available in ICAR (Indian Council of Agricultural Research) for data analysis, though it is a highly advanced statistical analysis platform, and its complexity holds back students and researchers from using it. Some web applications like WASP (https://ccari.res.in/waspnew.html) and OPSTAT (http://14.139.232.166/opstat/) used by the agricultural research community are user friendly but these applications don't provide options to generate plots and graphs.

The open source programming language R and associated ecosystem of packages, provides an excellent platform for data analysis but as of yet, is not heavily utilised by researchers in agricultural disciplines. Insufficient programming and computational knowledge are the primary challenges for agricultural researchers using R for analysis, as well as a preferences for researchers in agriculture to prefer a graphical user interface.

Efforts should therefore be made to develop a high quality, reliable open-source platform with a simple interactive user interface for data analysis in agriculture. Utilising the features of shiny package in R, we have developed a collection of shiny apps for agricultural research called grapesAgri1 (General R shiny based Analysis Platform Empowered by Statistics for data analysis in Agriculture-part1). grapesAgri1 is hosted on the web (http://www.kaugrapes.com), downloadable as a standalone application, and now released as an R package.

By calling the functions in the grapesAgri1 package, a graphical user interface will open. Apps are self-explanatory and work by uploading files in CSV format. Results can be downloaded in a HTML format. Plots and graphs can be generated, which are also downloadable as .png files.

Statement of need

India has one of the world's largest agricultural education systems. According to the Indian Council of Agricultural Research (ICAR), the main regulator of agricultural education in India, there are 63 State Agricultural Universities, 4 deemed universities, and 3 Central Agricultural Universities. These institutions enroll annually about 15,000 students in undergraduate programs and over 7,000 students in post graduate programs and more than 2000 at PhD

level. At any point, there are over 75,000 students studying in these institutions. Research activities are performed actively in all these universities. `grapesAgri1` not only serves as a platform for data analysis but also can be used as a teaching tool in agricultural statistics. `grapesAgri1` includes some basic statistical tools which were covered in the syllabi of undergraduate programs as well as in post graduate programs.

Information for Users

`grapesAgri1` is a collection of six `shiny` applications. Details of six applications are given below. Model dataset for testing can be downloaded from the main window of the application. Details for preparing CSV files are also included in the main window. Users just need to click on the browse button and upload the CSV file for analysis.

Apps included in the package

Sl. No.	App Title	Function to call	Utility
1	Descriptive Statistics and Visualization	descApp()	Summary statistics, summary statistics by group, box plot, histogram, Q-Q plot and Shapiro-Wilk's test
2	Correlation Analysis	corrApp()	Simple correlation, correlation Matrix, correlogram and scatterplot
3	Compare Means: Small samle tests	ttApp()	One sample t-test, two sample unpaired t-test, paired t-test, two sample unpaired Welch t-test, F test, box plot, paired Plot
4	Completely Randomized Design	crdApp()	One-way ANOVA (equal or unequal replications), multiple comparison test, boxplot and barchart with confidence interval
5	Field layout of experiments	layoutApp()	Field layouts of following designs can be obtained: completely randomized design (CRD), randomized complete block design (RCBD), split-plot design, strip-plot design, augmented RCBD
6	Randomized Block Design	rbdApp()	Two-way ANOVA, multiple comparison test, boxplot and barchart with confidence interval

The package can be installed from CRAN as follows:

```
# Install from CRAN
install.packages('grapesAgri1', dependencies=TRUE)
```

The development version can be installed from GitHub as follows:

```
# Install grapesAgri1 development version from Github using the code below:
if (!require('devtools')) install.packages('devtools')
devtools::install_github("pratheesh3780/grapesAgri1")
```

Package dependencies and details of functions used

DescApp() function uses `descr` and `stby` functions of `summarytools` package (Comtois, 2021) to calculate summary statistics and summary statistics by group. `knitr` (Xie, 2021) and `kableExtra`(Zhu, 2021) packages were used to produce HTML tables. `shapiro.test`, `qqnorm` and `qqline` functions of `stats` package were used for the Test of Homogeneity of variance and obtaining Q-Q plot. `hist` and `boxplot` of package `graphics` were used to obtain histogram and boxplot respectively. `ggqqplot` of package `ggpubr`(Kassambara, 2020) is also used to plot Q-Q plot in the app.

CorrApp() function uses `cor.test` to calculate correlation. Correlation matrix is calculated using `rcorr` function in `Hmisc` package (Harrell Jr et al., 2021). Correlogram is obtained using `corrplot` function in `corrplot`(Wei & Simko, 2021) package.

ttApp()function uses `t.test` function to calculate t statistic. Descriptive statistics were calculated using `stat.desc` function of `pastecs` package. `var.test` function is used for F-test. `ggboxplot` function of `ggpubr` (Kassambara, 2020) package is used to draw boxplot. Paired plot is obtained using `paired` function of package `PairedData`(Champely, 2018).

crdApp() uses `anova` function of `stats` package to obtain one-way ANOVA. `LSD.test`,`duncan.test` and `HSD.test` functions of `agricolae` (de Mendiburu, 2020) package is used for multiple comparison test like LSD,DMRT and Tukey respectively. `ggboxplot` function of `ggpubr` (Kassambara, 2020) package is used for boxplot. `ggplot` function of `ggplot2`(Wickham, 2016) is used for barchart with confidence interval.

layoutApp() uses `design.crd`, `design.rcbd`, `design.dau`, `design.strip`, `design.split` functions of package `agricolae` (de Mendiburu, 2020) to generate random layout of designs. Field layout were plotted using `desplot` function in `desplot` package (Wright, 2020).

rbdApp() uses `anova` function of `stats` package to obtain two-way ANOVA. `LSD.test`,`duncan.test` and `HSD.test` functions of `agricolae` package (de Mendiburu, 2020) is used for multiple comparison test like LSD,DMRT and Tukey respectively. `ggboxplot` function of `ggpubr` package (Kassambara, 2020) is used for boxplot. `ggplot` function of `ggplot2` (Wickham, 2016) is used for barchart with confidence interval.

Usage

```
grapesAgri1::descApp() # descriptive Statistics and Visualization
grapesAgri1::corrApp() # Correlation Analysis
grapesAgri1::ttApp() # Compare Means
grapesAgri1::crdApp() # Completely Randomized Design
grapesAgri1::layoutApp() # Field layout of experiments
grapesAgri1::rbdApp() # Randomized Block Design
```

Acknowledgements

References

Champely, S. (2018). *PairedData: Paired data analysis*. https://CRAN.R-project.org/package=PairedData

Comtois, D. (2021). *Summarytools: Tools to quickly and neatly summarize data*. https://CRAN.R-project.org/package=summarytools

de Mendiburu, F. (2020). *Agricolae: Statistical procedures for agricultural research*. https://CRAN.R-project.org/package=agricolae

Harrell Jr, F. E., Charles Dupont, with contributions from, & others., many. (2021). *Hmisc: Harrell miscellaneous*. https://CRAN.R-project.org/package=Hmisc

Kassambara, A. (2020). *Ggpubr: 'ggplot2' based publication ready plots*. https://CRAN.R-project.org/package=ggpubr

Wei, T., & Simko, V. (2021). *R package "corrplot": Visualization of a correlation matrix*. https://github.com/taiyun/corrplot

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. ISBN: 978-3-319-24277-4

Wright, K. (2020). *Desplot: Plotting field plans for agricultural experiments*. https://CRAN.R-project.org/package=desplot

Xie, Y. (2021). *Knitr: A general-purpose package for dynamic report generation in r*. https://yihui.org/knitr/

Zhu, H. (2021). *kableExtra: Construct complex table with 'kable' and pipe syntax*. https://CRAN.R-project.org/package=kableExtra