RESEARCH ARTICLE

A Developed Interactive Web Application for Statistical Analysis: Statistical Analysis Software

Seyma Yasar[1], Ahmet Kadir Arslan[1], Cemil Colak[1], Saim Yoloğlu[1]

[1]Department of Biostatistics and Medical Informatics, Faculty of Medicine, Inonu University,Malatya, Turkey

Abstract
Objective: Hypothesis testing, correlation, and regression analysis are statistical methods developed to be used in the statistical inference process which is the main purpose of all scientific studies. The purpose of this study is to develop a web-based application using the Shiny package in R software, which allows the evaluation of the results of scientific research to be made in a simpler, easier and understandable way.
Methods: In this study, the tests and techniques developed in the software were applied to the data derived from the simulation. This web tool will be updated upon the updated R software packages, including ggplot2, shiny, reshape, plotly, shinydashboard, dplyr, plyr, tinytex, DT, rhandsontable, shinyjs, tools, readxl, foreign, shinyWidgets, shinyLP, shinyjqui, stringr, olsrr, perturb, mctest, relaimpo, MASS, MKmisc, aod, caret, shinydashboardPlus, rmarkdown. Scripts were written for calculations that could not be done by these packages.
Results: In the web-based software developed to perform statistical analyses, the results and outputs of the derived dataset were interpreted.
Conclusion: The developed interactive user-friendly web application is freely accessible through http://biostatapps.inonu.edu.tr/IAY. In future studies, it is aimed to strengthen the software by adding modules that perform different multivariate statistical analyzes.
Key words: Statistical analysis software; hypothesis testing; correlation analysis; regression analysis; web-based software

Address for correspondence/reprints:

Şeyma Yaşar

Telephone number: +90 422 341 0660/1321

ORCID-ID 0000-0003-1300-3393

E-mail: seyma.yasar@inonu.edu.tr

Introduction

Scientific researches are the studies carried out in order to contribute to science through the systematic and planned collection, interpretation and evaluation of the data (Caparlar and Donmez 2016). In scientific research, the statistic has a significant place. The science of statistics enables the planning of the study, the collection of data, the evaluation, and the inference of a study in order to examine any subject from a scientific point of view. In scientific research, estimation of the populations parameters with the help of the statistics obtained from the samples and in order to generalize the findings to the universe of the research, each stage of the research plan must be made correctly. One of the important stages in research planning is the evaluation of the data collected from the individuals in the sample with the correct statistical methods. The approaches such as basic hypothesis tests, correlation, regression analysis, and data science methods are frequently used in data analysis within this context (Ozdamar, 2013).

The purpose of the hypothesis tests is to examine whether the hypotheses are realized statistically and to make a decision. For this purpose, there are many hypothesis testing methods used depending on the type of data collected (nominal/ordinal/scale), the design of the experiment (dependent/independent), whether it is independent of distribution (parametric/nonparametric) and the number of samples (one/two/more than two) (Alpar, 2010). It is very important to control whether distribution assumptions, one of the hypothesis testing assumptions, are provided. Parametric hypothesis tests used under the assumption that distribution assumptions are provided are considered to be stronger than nonparametric tests (Pituch and Stevens, 2015). Therefore, ignoring the assumption of normality may cause false tests to be applied to the data. Therefore, the results obtained can be misinterpreted (Yazici and Yolacan, 2007).

The correlation analysis is tested whether the relationship between two or more variables is statistically significant. With the correlation coefficient, information is obtained about the direction and strength of the relationship which is statistically significant. Correlation coefficients also vary depending on the type and distribution of data, as in hypothesis testing (Orhunbilge, 2017).

Regression analysis is based on the principle of determining the relationships between the dependent variable and the independent variable (s) and expressing the relationship between them with the help of a mathematical model. The main purpose of the regression analysis is to predict the dependent variable with the help of the independent variable(s). If the number of the independent variable is more than one, regression analysis is done to find the answer to the question, "Which argument is affected more by the dependent variable?" (Alpar 2013).

The aim of this study is to develop a web-based software with open source access that is independent of the operating system and user-friendly that performs basic statistical methods which are hypothesis testing, correlation analysis and regression analysis (linear /binary logistics).

Methods

In order to introduce the working principles of the developed web-based software, the data set, in which quantitative variables consisting of five variables, three quantitative and two qualitative, have the standard normal distribution ($\mu=0,\sigma=1$) and the variables contain 1000 observations, is derived from the simulation tab of IBM SPSS Statistics version 25.0 (Corp. 2017). The basic features and descriptive statistics for the data set are given in Table 1.

Table 1. The basic features and descriptive statistics for the data set

Variables	Variable Type	Descriptive Statistics	
y	Categorical	0	589 (58,9)
		1	411 (41,1)
x_1	Categorical	0	279 (27,9)
		1	402 (40,2)
		2	319 (31,9)
x_2	Continual	$197,7 \pm 18,5$	
x_3	Continual	$182,7 \pm 18,6$	
x_4	Continual	$175,2 \pm 22,3$	

Developed Web-Based Software

In the development of the web-based application, the Shiny (Chang et al. 2017) package was used on the basis of the R programming language. With web-based software, normality, hypothesis testing, correlation analysis, linear regression, and logistic regression analysis are performed. The description of the main and sub-menus of the web-based application is detailed below.

Introduction Menu

The "Introduction" menu, the first menu of this web-based software, contains information about the software. The screenshot of this menu is shown in Figure 1.

Upload File Menu

In the "Upload File" menu of the developed web-based software, MS Excel (.xls / .xlsx) and SPSS (.sav) extensions are loaded with file types containing the data set. It is possible to change the values of the loaded dataset, delete rows and add variables. In addition, in this menu, there is a sub-menu where the type and role of the variables in the loaded data are determined. The screenshot of this menu is depicted in Figure 2.

Normality Menu

In this menu of the developed web-based software, Shapiro-Wilk and Kolmogorov Smirnov tests are used to test the compliance of the variable (s) to normal distribution. Shapiro-Wilk normality test is a test used when the number of subjects for each group in independent measurement and sample size in dependent measurement is less than or equal to 50. However, in other cases, the test used for normality tests is usually the Kolmogorov-Smirnov normality test (Alpar, 2010). The image of this menu is depicted in Figure 3.
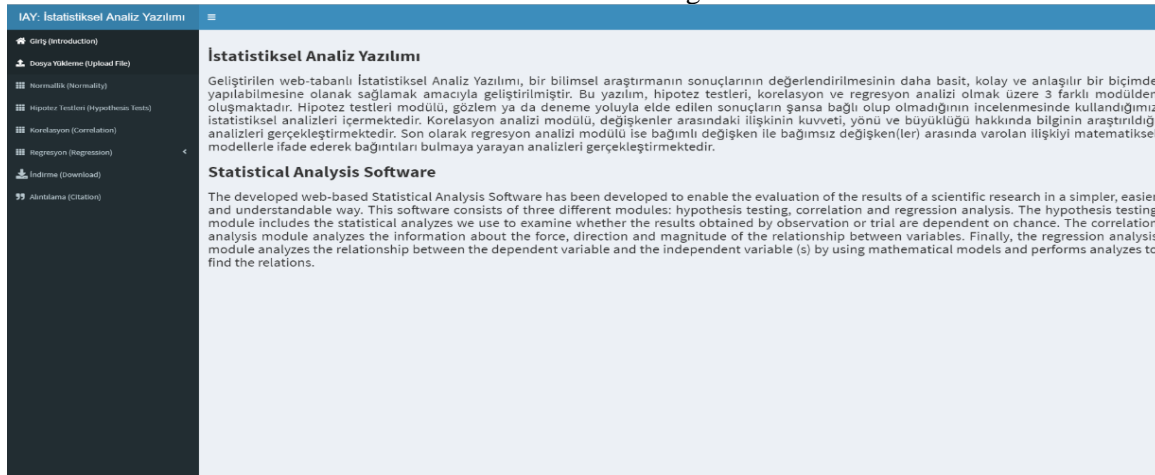

Figure 1. "Introduction" Menu

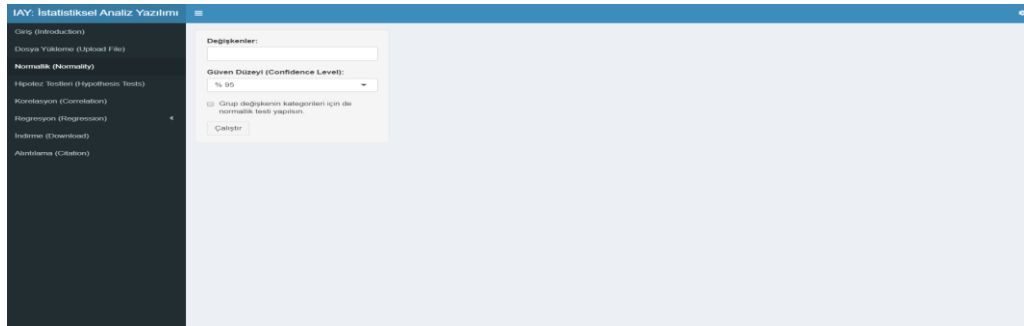**Figure 2.** "Upload File" Menu

Figure 3. "Normality" Menu

Hypothesis Tests Menu

In the related menu of the developed web-based software, there are basic hypothesis tests that can be used depending on the measurement format of the variables, the structure, and the number of groups. The hypothesis testing menu consists of two submenus, "Quantitative Variables" and "Qualitative Variables", where the measurement method of variables is determined. There are 3 options: "One", "Two" and "More than Two" in the "Quantitative Variables" sub-menu. In these sections, sample/group numbers are determined. In the "One" menu, there is the "One-Sample t-Test" option, which is used when the sample/group number is one. In the "Two" sub-menu, there are tabs that can be used depending on the sample/group structure. If the sample/group structure is independent, the "Two Independent Sample t-Tests" button is selected. However, if the sample/group structure is dependent, the "Two Paired Sample t-Test" button can be preferred. The "Two Independent Sample t-Tests" submenu includes the "Two Independent Sample t-Tests" used when the variables provide normal distribution assumptions and the "Mann-Whitney U Test" that can be applied when the normal distribution assumption is not realized. On the other hand, the "Two Dependent Sample t-Test" menu consists of the "Paired Sample t-Test" button that can be used when the normal distribution assumption of dependent measurements is provided and the "Wilcoxon Test" button that can be used when normal distribution assumptions are not provided. In the sub-menu where the number of groups is "More than 2", there are "More Than Two Independent Sample Tests" and "More than Two Dependent Sample Tests" which are preferred depending on whether the group structure is dependent and independent. The "More Than Two Independent Sample Tests" sub-menu includes the "One-Way Analysis of Variance (One-way ANOVA)" that can be applied under the normal distribution assumption of groups and the "Kruskal-Wallis Test" that can be realized when the normal distribution assumption of groups is not provided. Similarly, in "More than Two Dependent Sample Tests" submenu, there is a "Friedman Test" button that can be used in cases where normality assumptions do not occur with "One Way Analysis of Variance with Repeated Measures" button, which can be used under normality assumptions. With the "Automatic Test Selection Mode" in the hypothesis testing menu, the test to be used is performed automatically according to the normality assumption regarding the quantitative variables. In order to set an example for the developed web-based software, the image of the " One Way Analysis of Variance with Repeated Measures" menu is given in Figure 4.
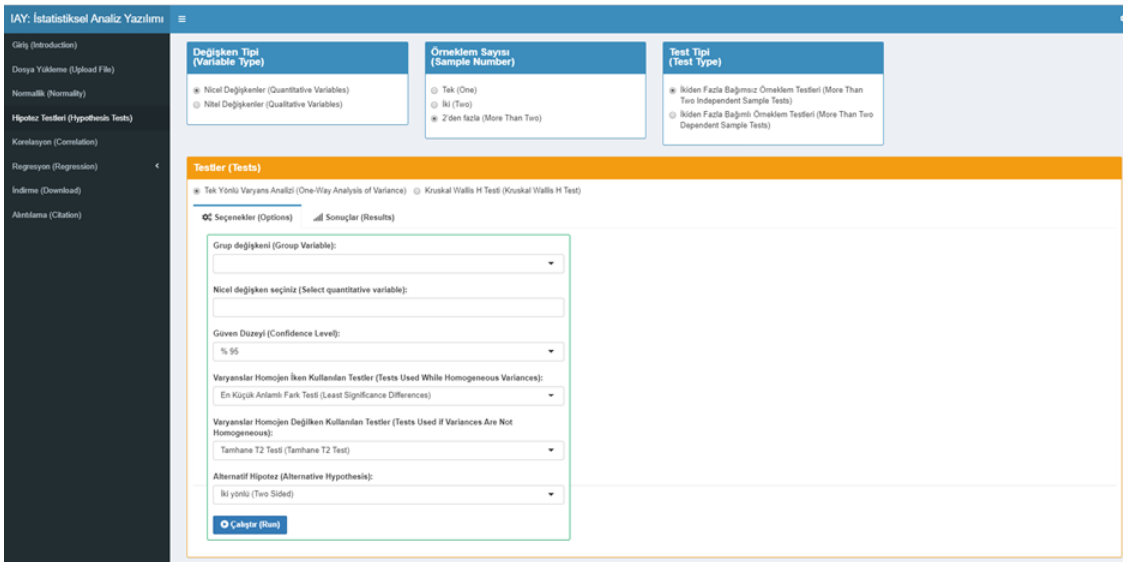

Figure 4. "One Way Analysis of Variance with Repeated Measures" Menu

In the "Qualitative Variables" sub-menu, there are 2 options, namely "One" and "Two", where the sample/group numbers are specified. There is a "Chi-Square Goodness of Fit Test" in the "Single" submenus, where the distribution of a single qualitative variable is tested for the compatibility of any theoretical distribution. On the other hand, in the "two" sub-menu, there is a "Chi-Square Independence Test", which examines whether there is a relationship between two qualitative variables with two or more groups.

Correlation Menu

Pearson Correlation, Spearman Correlation, and Kendall Tau Correlation coefficient were used in this module, where the correlation analysis used to investigate the size, direction, and significance of the relationship between variables was performed, depending on the variables' normal distribution assumptions. If two variables determined by measurement have a normal distribution, the relationship between these two variables can be determined by the Pearson correlation coefficient. On the other hand, the Spearman correlation coefficient or the Kendall Tau correlation coefficient can be used if two variables are qualitative, or if at least one of these two variables do not have a normal distribution. The value of the correlation coefficient ranges from -1 to +1. The sign for these values indicates the direction of the relationship. The strength of the relationship increases as the coefficient approaches -1 and +1. On the other hand, as the coefficient approaches 0, the strength of the relationship decreases (Alpar, 2010).

The correlation menu consists of two submenus, "Which Correlation Coefficient Should I Use?" And "Analysis Outputs". In the "Which Correlation Coefficient Should I Use?" menu, it is aimed to determine the appropriate correlation coefficient before correlation analysis. To determine the appropriate coefficient, all possible variable pairs are tested by the Shapiro-Wilk normality test. Then, the relevant p-value for each variable and the correlation coefficient suitable for use in the study of the relationship between the variable pair are suggested to the researchers. In the "Analysis Outputs" menu, the correlation coefficient value, p-value and statistical significance selected between the variable pairs are summarized with a table. As an example, the image of the "Analysis Outputs" sub-menu under the correlation menu is given in Figure 5.

Regression Menu

The "Regression" analysis, which is used to determine the relationship between two or more variables with a cause and effect relationship, may vary depending on the measurement type of the dependent variable. Similarly, when the dependent variable is a qualitative two-category variable, Binary Logistic Regression Analysis is used in studies in which the cause-effect relationship between the independent variables and the dependent variable is investigated (Orhunbilge, 2017).

The regression menu consists of two submenus: "Linear Regression" and "Logistic Regression". As an example, the image of the "Linear Regression" menu is given in Figure 6.

Download Menu

The "Download" menu in the developed web-based software allows the researchers to download and store all analysis outputs related to their work with the extension "PDF", "Word" and "HTML". The image of this menu is given in Figure 7.
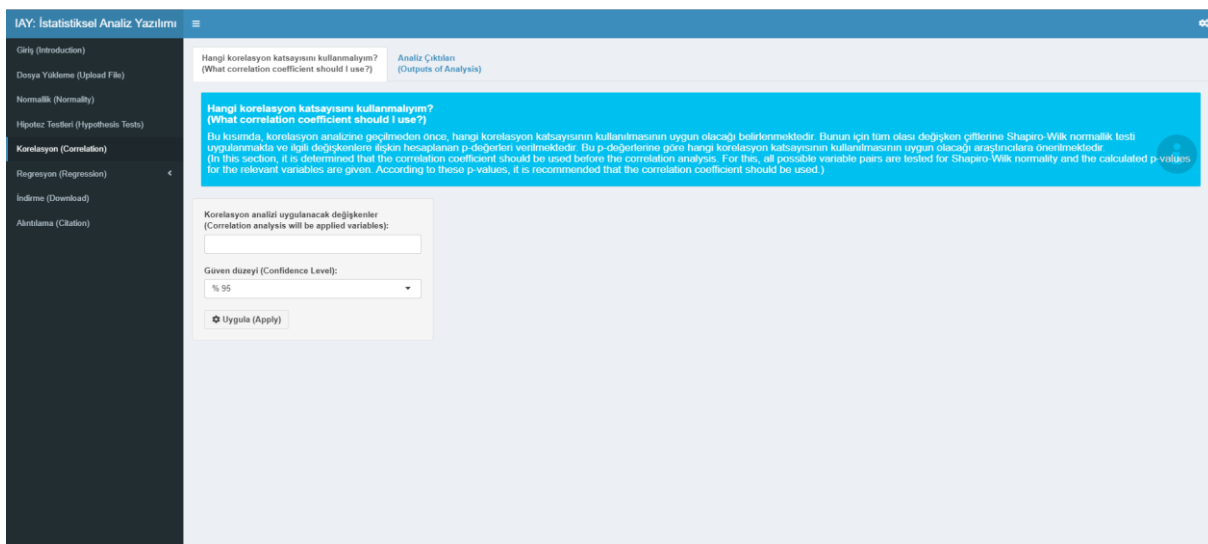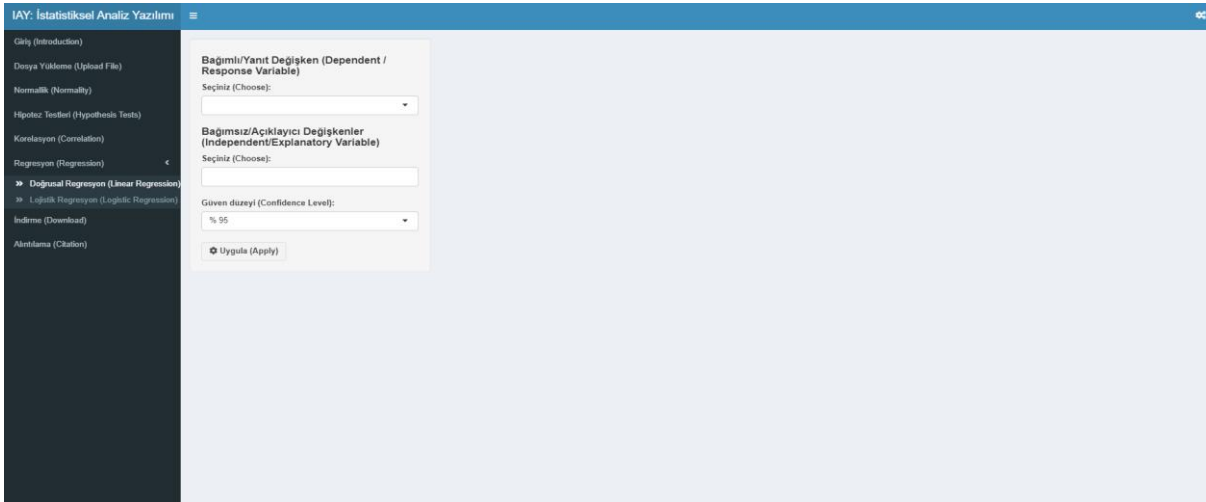

Figure 5. "Analysis Outputs" menu in correlation menu
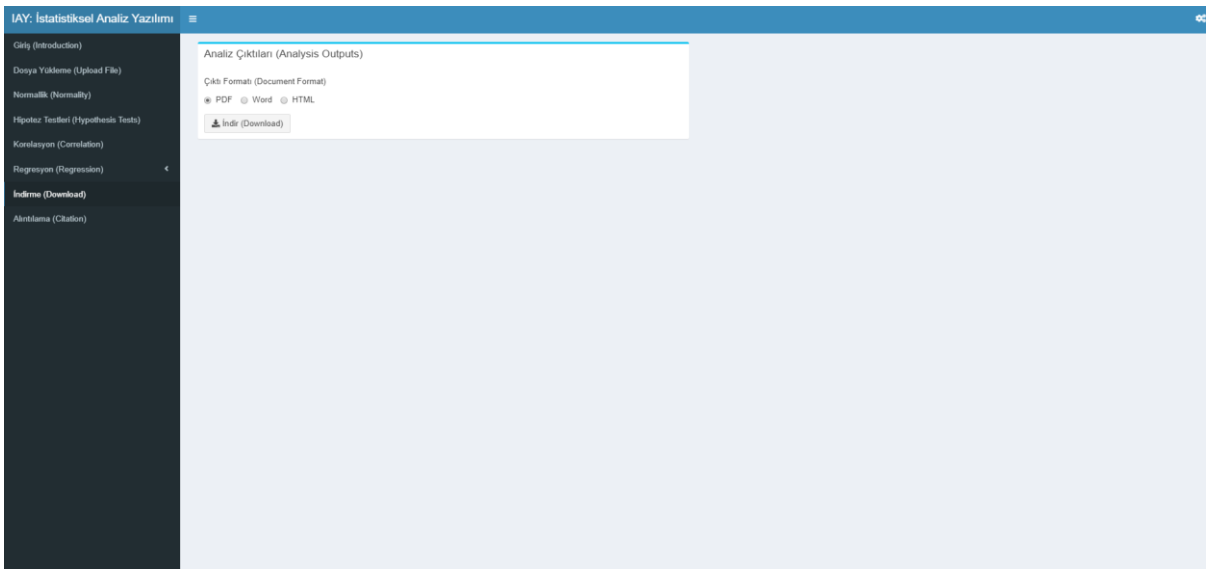
Figure 6. "Linear Regression" Menu

Figure 7. "Download" Menu"

Access to Developed Interactive Web Application
The developed interactive web application can be accessed free of charge at http://biostatapps.inonu.edu.tr/IAY/. This web-based software has been developed on updated R software packages including ggplot2 (Wickham, 2011), shiny (Chang et al. 2017), reshape (Wickham and Wickham 2017), plotly (Sievert et al. 2018), shinydashboard (Chang and Borges Ribeiro 2017), dplyr (Wickham 2018), plyr (Wickham and Wickham, 2016), tinytex (Yihui Xie et al. 2019), DT (Xie et al. 2015), rhandsontable (Jonathan Owen et al. 2018), shinyjs (Attali, 2016), tools (Hadley Wickham, 2018), readxl (Wickham and Bryan, 2017), foreign (Team et al. 2019), shinyWidgets (Perrier, 2019), shinyLP (Dumas, 2019), shinyjqui (Tang, 2018), stringr (Wickham, 2010), olsrr (Hebbali, 2018), perturb (Hendrickx and Hendrickx, 2019), mctest (Imdadullah et al. 2016), relaimpo (Groemping and Matthias, 2018), MASS (Brian Ripley et al. 2019), MKmisc (Kohl and Kohl, 2016), aod (Lesnoff et al. 2010), caret (Kuhn, 2012), shinydashboardPlus (David Granjon et al. 2019), rmarkdown (Baumer et al. 2014).

Results
The image of the "Upload File" menu formed as a result of loading the derived data set into the software to show the operating principles of the menus of the web-based software is given in Figure 8.

Normality Analysis
It is aimed to investigate the normality of the x2 variable in the derived dataset both in terms of the variable and in terms of the categories of the y variable. The image of the outputs regarding the

normality analysis that takes place when the "Run" button is pressed is given in Figure 9. Considering the results, the variable x2 showed a normal distribution according to the Kolmogorov Smirnov normality analysis. In addition, according to the subgroup analysis using the variable y, the variable x2 shows a normal distribution on the basis of each group.

Hypothesis Tests

To demonstrate the working principle of the hypothesis testing menu, consider the x4 quantitative variable and x1 categorical variable in the derived data set. Whether there is a statistically significant difference between subcategories of x1 variable in terms of x4 variable is wanted to be examined by one-way analysis of variance. The image of the one-way analysis of variance to be performed is given in Figure 10. When the assumption of homogeneity variances is provided in Table 8, it is preferred to use the "Least Significance Differences Test", and if the homogeneity assumptions are not provided, the "Tamhane T2 Test" is preferred. When the "Run" button is pressed, the table containing the descriptive statistics related to the variable on the basis of the group, the table containing the statistics obtained as a result of testing the homogeneity of the variances, table with one-way analysis of variance test statistics, a window where the statistical significance of the results is interpreted, a table with statistics on the "Least Significance Difference Test" results in which binary comparisons are made and finally, mean standard deviation graph are obtained. An image of these results is given in Figure 11.

As a result, there is a statistically significant difference between the categories of the x1 variable in terms of the x4 variable (p=0.0454). According to Levene test results, variances show homogeneous distribution. Therefore, there is a statistically significant difference between "0-2" and between "1-2" according to the binary comparisons made using the "Least Significant Difference Test".

Correlation Analysis

In order to determine the relationship between x2, x3 and x4 variables in the derived data set, it is necessary to determine the appropriate correlation coefficient to be used first. In the developed web-based software, the image of the menu where this coefficient is determined is given in Figure 12.

Since each of the variable pairs showed normal distribution, the appropriate correlation coefficient to be used in the correlation analysis was determined as "Pearson Correlation Coefficient". The image of the menu with the outputs of the correlation analysis applied according to the determined Pearson correlation coefficient is given in Figure 13.

According to the results obtained, the relationship between the x2 variable and x3 variable and x2 variable and x4 variable is statistically significant, whereas the relationship between x3 variable and x4 variable is not statistically significant.

Logistic Regression Analysis

In the derived data set, an image of the logistic regression analysis performed to model the relationship between the dependent categorical y variable and the independent x3 variable specified by continuous measurement and the categorical x1 variable is given in Figure 14. In the logistic regression analysis carried out as an example, "Entry" was selected, in which all variables were included in the model from the model selection methods. In the image of Figure 14.

In the logistic regression analysis carried out as an example, "Entry" was selected, in which all variables were included in the model from the model selection methods. In the image of Figure 14, a table of Hosmer-Lemeshow test statistics, which is a measure of the effectiveness of the best model created to explain the dependent variable of the model, table of statistics for different explanatory coefficients used as an indicator of fit to the model, table of variables coefficient in the model, classification matrix and table of statistics on classification performances are given.

Download

If it is desired to download with the .html extension of all samples made in order to explain the working principle of the developed web-based software, the resulting image is given in Figure15

Figure 8. Uploading Data to the Software

Figure 9. Normality Test Results Regarding x₂ variable

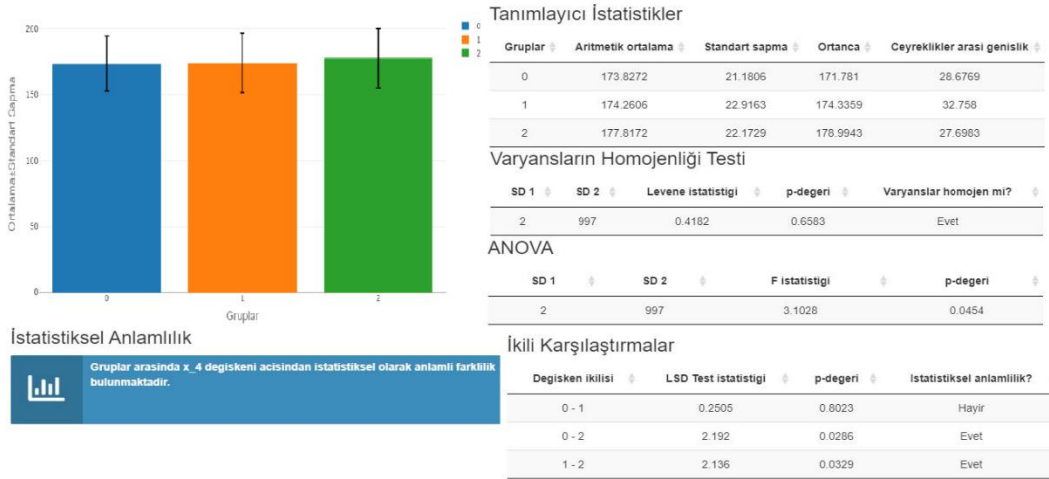**Figure 10.** "One-way Analysis of Variance" Hypothesis Testing

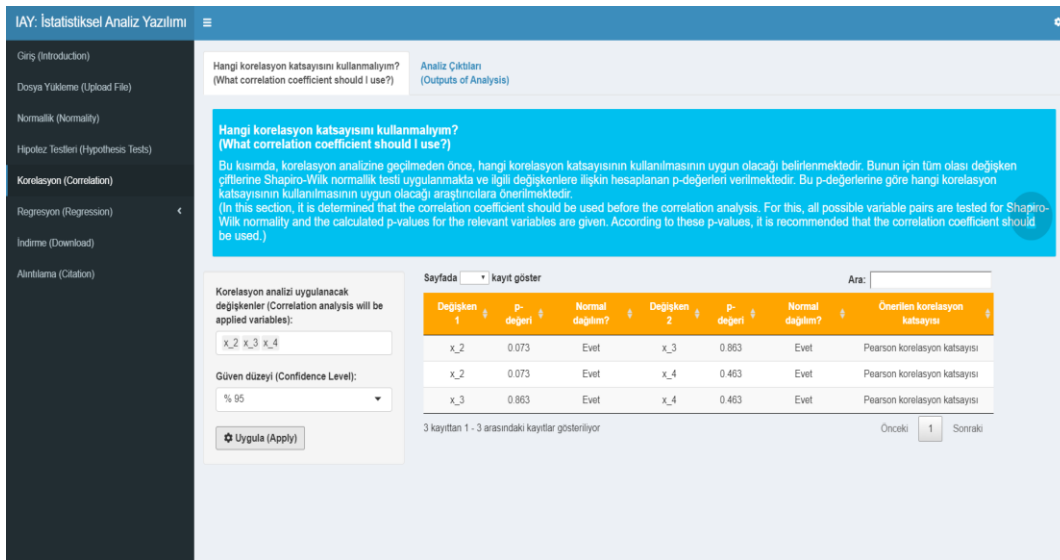**Figure 11.** Image of Results Regarding One-Way Analysis of Variance

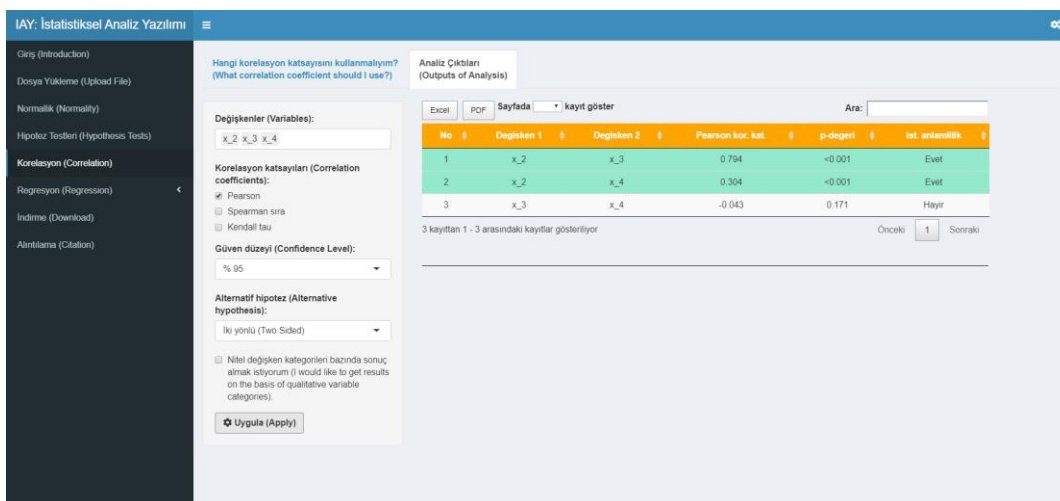**Figure 12.** Determining the Appropriate Correlation Coefficient
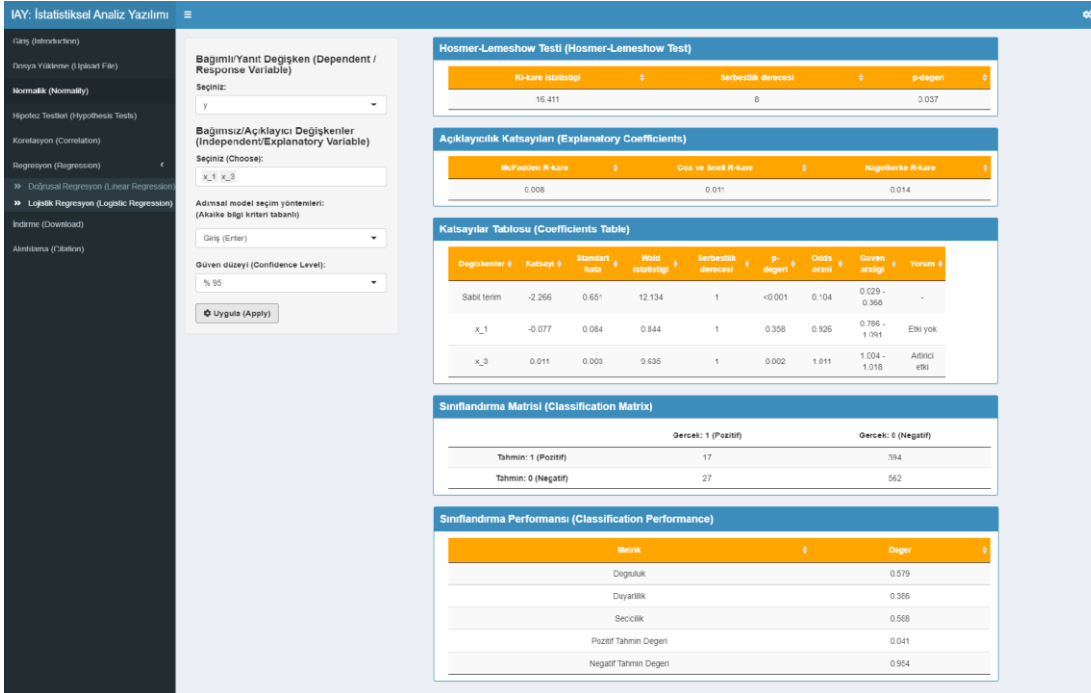

Figure 13. Image of Correlation Analysis Statistics

Figure 14. Image of statistics obtained from logistic regression analysis

Figure 15. Downloading the analysis results

Discussion

The analysis of the data collected for scientific research can be carried out by many methods. It is very important to check the assumptions of the analyses to be applied and to interpret the results correctly. At this stage, choosing the wrong statistical analysis method may cause errors in terms of the generalization of the study to the universe (external validity) (Yazici and Yolacan 2007).

Hypothesis testing varies according to the type of data used (nominal/ordinal/continuous), whether the distribution provides normality assumptions (parametric/non-parametric), the structure (dependent/independent), the number of groups (one/two/more than two) handled, etc.. There is much software used and developed to perform statistical analysis. Package programs, which are frequently used in scientific research, are; IBM SPSS Statistics (Corp. 2017), MedCalc (Schoonjans 2009), Minitab (Minitab 2000), Statistica (StatSoft 2001), etc. can be listed as. The difference of the developed web-based software from this aforementioned software is that it directs the researchers to the appropriate test that will be used according to the type of data, structure, and number of groups through consecutive menus. At the same time, the "Automatic Test Selection Mode" available in the software enables the appropriate test to be performed automatically depending on whether or not the assumption of normality on the basis of the variable is achieved. The interpretation of the results of a scientific study is done with various statistical criteria obtained as a result of the analysis. Therefore, the correct interpretation of these criteria is of great importance for the quality of the study. While none of the known statistical analysis programs have an information section that includes the interpretation of statistical criteria/values, this developed web-based software also provides the researchers with comments on the statistical criteria/values obtained. In addition, in the multi-class chi-square analysis menu, the software makes binary comparisons for two variables that have a relationship and interpret the results of the obtained statistics.

In the correlation analysis applications, the coefficient used to determine the direction and strength of the relationship between the two variables varies according to the measurement type of the variables and the assumption of normal distribution (Orhunbilge 2017). Researchers can ignore these assumptions and often choose to use the Pearson correlation coefficient. Statistics obtained by using the wrong correlation coefficient in the analysis phase may cause the results of the research to be evaluated incorrectly. The developed web-based software controls the normal distribution on a variable basis and reports the appropriate correlation coefficient to be used in the analysis. There is also an interpretation of the statistics obtained from the analysis in the correlation analysis menu.

Many assumptions must be fulfilled in order to apply the multivariate regression analysis. One of these assumptions is the multivariate normal distribution for independent variables. There are no modules in this program that control this assumption. Thanks to the developed web-based software, researchers can test the multivariate normal distribution assumption regarding the independent variables used in regression analysis with Mardia's distortion-kurtosis, Henze-Zirkler and Doornik-Hansen test and they can reach the interpretation of the statistical significance related to the obtained test statistics. The multicollinearity problem is another assumption of multivariate regression analysis. In the mentioned programs, statistics such as condition index, variance inflation factor, and tolerance value are used to detect this problem. On the other hand, in addition to these statistical values, Farrar and Glauber Chi-square hypothesis testing is used in the developed software and the test statistics are interpreted.

The significance of the odds ratio statistics obtained as a result of logistic regression analysis and the correct interpretation of the results are of great importance for the correct evaluation of the study. Therefore, in the logistic regression analysis menu of the developed web-based software, it assists the researchers in this regard as it includes the comments on the odds ratio statistics (increasing effect, reducing effect).

Conclusion

The developed web-based software controls hypothesis tests, correlation analysis and simple/multiple linear regression analysis and logistic regression analysis assumptions that will be applied to the data obtained from scientific studies and presents the results obtained in a table. It also provides support for researchers on data analysis and interpretation, as it includes statistical comments on these values.

References

Alpar R. Applied statistics and validity and reliability with examples from sports, health and education sciences. Ankara: Detay Publishing; 2010.

Alpar R. Applied multivariate statistical methods. Ankara: Detay Publishing; 2013.

Attali D. shinyjs: easily improve the user experience of your shiny apps in seconds; 2016. R package version 0.8 https://CRAN.R-project.org/package=shinyjs.

Baumer B, Cetinkaya-Rundel M, Bray A, Loi L, Horton N. J. R Markdown: Integrating a reproducible analysis tool into introductory statistics. arXiv preprint arXiv:1402.1894, 2014.

Ripley B, Venables B, Bates D M, Hornik K, Gebhardt A, Firth D, Ripley M B. MASS: support functions and datasets for Venables and Ripley's MASS. R package version, 2011, 7: 3-29.

Chang W, Ribeiro B B. shinydashboard: create dashboards with 'Shiny'; 2017. R package version 0.6.1 https://CRAN.R-project.org/package=shinydashboard.

Chang W, Cheng J, Allaire J, Xie Y, McPherson J. Shiny: web application framework for R; 2016. R Package version 0.14.2. https://CRAN.R-project.org/package=shiny.

IBM C. R. IBM SPSS Statistics for Windows, Version Q3 25.0. Armonk, NY: IBM Corporation, 2017.

Caparlar C. Ö, Dönmez A. What is Scientific Research and How Is It Done? Turk J Anaesthesiol Reanim, 2016 44: 212-8.

Granjon D. shinydashboardPlus: Add More 'AdminLTE2' Components to 'shinydashboard'; 2019. R Package Version 0.7.0. https://CRAN.R-project.org/package=shinydashboardPlus.

DUMAS J. shinyLP: Bootstrap Landing Home Pages for Shiny Applications; 2019. R package version. https://CRAN.R-project.org/package=shinyLP.

Groemping U, Matthias L. Package 'relaimpo'; (2018). R package https://cran.r-project.org/web/packages/relaimpo/relaimpo.

Wickham H, Chang W. devtools: tools to make developing R packages easier; 2018. R package version 1.13. 4. https://CRAN.R-project.org/package=devtools.

Hebbali A. olsrr: Tools for building OLS Regression Models; 2018. R version 0.5. 1. https://cran.r-project.org/web/packages/olsrr/olsrr.pdf.

Hendrickx J, Hendrickx M J. Package 'perturb'; 2019. http://cran.r-project.org/web/packages/perturb/perturb.pdf .

Imdadullah M, Aslam M, Imdadullah M. M, LazyData T. R. U. E. Package 'mctest'; 2016. https://CRAN.R-project.org/package=mctest .

Owen J. rhandsontable: Interface to the 'Handsontable. js' library; 2018. R package version 0.3, 7. https://CRAN.R-project.org/package=rhandsontable.

Kohl M, Kohl M. M. Package 'MKmisc'; 2019. R package version 0.3. https://CRAN.R-project.org/package=MKmisc.

Kuhn M. The caret package; 2012. R Foundation for Statistical Computing, Vienna, Austria. https://cran.r-project.org/package=caret.

Lesnoff M, Lancelot R, Lancelot M R, Suggests M. Package 'aod'; 2010. R package version 1.3.1. https://CRAN.R-project.org/package=aod.

Minitab, I (2000) MINITAB statistical software. Minitab Release, 13.

Orhunbilge N, Applied regression and correlation analysis. Ankara: NOBEL Academic Publishing; 2017.

Ozdamar, K. 2013. Biostatistics with SPSS. Eskişehir: Nisan Bookstore Publications; 2013.

Perrier V, shinyWidgets: Custom Inputs Widgets for Shiny; 2019. R package version, 0.4.8. https://CRAN.R-project.org/package=shinyWidgets.

Pituch K A, Stevens J P. Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS. Routledge: NewYork; 2016.

MedCalc Statistical Software. Version 10.1.6.0, bvba, Ostend, Belgium.

Sievert C, Parmer C, Hocking T, Chamberlain S, Ram K, Corvellec M, Despouy P. plotly: Create Interactive Web Graphics via'plotly. js'; 2016. R package version 3.6. 0. ed. https://CRAN.R-project.org/package=plotly.

StatSoft I, (2001) STATISTICA (data analysis software system), version 6. Tulsa, USA, 150, 91-94.

Tang Y. shinyjqui: 'jQuery UI' Interactions and Effects for Shiny; 2018. https://CRAN.R-project.org/package=shinyjqui.

Team R C, Bivand R, Carey V J, DebRoy S, Eglen S, Guha R, Herbrandt S, et al. Package 'foreign'; 2019. https://CRAN.R-project.org/package=foreign.

Wickham, H. Stringr: modern, consistent string processing. The R Journal 2010; 2(2), 38-40.

Wickham H. ggplot2. Wiley Interdisciplinary Reviews: Computational Statistics 2011; 3(2), 180-185.

Wickham H, Francois R. dplyr: A Grammar of Data Manipulation; 2015.. R package version 0.4. 3. https://CRAN.R-project.org/package=dplyr.

Wickham H, Bryan J. readxl: Read Excel Files; 2017. R package version 1.0. 0. URL https://CRAN. R-project. org/package= readxl.

Wickham H, Wickham M. H. Package 'plyr'; 2016. Obtenido de. https://CRAN.R-project.org/package=plyr.

Wickham H, Wickham M H, Rcpp L. Package 'reshape'; 2018. R package version 0.8.8. https://CRAN.R-project.org/package=reshape.

Xie Y, Cheng J, Allaire J, Reavis B, Gersen L, Szopka B. DT: a wrapper of the JavaScript library 'DataTables; 2015. R package version 0.1. http://CRAN.R-project.org/package=DT.

Yazici, B & Yolacan, S (2007) A comparison of various tests of normality. J Stat Comput Simul., 77, 175-183.

XIE Y. tinytex: Helper Functions to Install and Maintain TeX Live, and Compile LaTeX Documents; 2018. R package version 5.8. https://CRAN.R-project.org/package=tinytex.