

ANNEX 1. BASES DEL CONCURS PER LA IMPLEMENTACIÓ D'UNA INTERFÍCIE GRÀFICA PER DIDÀCTICA DE R A LES CIÈNCIES SOCIALS

LLISTA DE REQUERIMENTS QUE HAURIA DE TENIR EL PROGRAMARI

- Software lliure, d'accés obert i gratuït.
- Funcionament multiplataforma (Windows, Mac i Linux).
- Funcionament a través d'un sistema de gestió gràfica (amb menús, icones, finestres).
- Especialment didàctic pels estudiants de grau que s'inicien en l'ús d'eines d'anàlisi estadística.
- Menús i missatges disponibles com a mínim en català i/o en castellà. (Els participants en el premi poden presentar una primera versió en anglès sempre que sigui possible la seva localització i facin constar el seu compromís a proveir una versió posterior en català i/o en castellà).
- Possibilitat d'executar directament instruccions en codi R (consola).
- Visualització de la matriu de dades, de la llista de variables amb les seves etiquetes i altres característiques).
- Visualització dels resultats de les anàlisis que vagi realitzant.
- Possibilitat d'obrir i guardar matrius de dades en formats diversos (csv, xls, sav, Rdata, etc).
- Exportació de taules de resultats i de gràfics com a arxius publicables (pdf, png,...).
- Possibilitat de copiar/enganxar taules i gràfics correctament a altres aplicacions.
- Possibilitat d'afegir i editar anotacions als resultats generats.
- Creació d'una matriu de dades nova i introducció de dades.
- Etiquetar variables i definir-ne les característiques mètriques.
- Definir valors perduts.
- Insertar i eliminar variables i casos.
- Creació de noves variables via transformacions i càlculs a partir de variables existents.
- Recodificació del valors de variables (reagrupació de categories).
- Fusionar dues matrius de dades, per variables i per casos.
- Filtratge de casos, i segmentació de resultats, en funció de variables.

- Taules de Freqüències (absoluta, relativa, acumulada, %,...)
- Descriptius bàsics (mitjana, mediana, moda, desviació típica, rang,...).
- Gràfiques descriptives corresponents (sectors, barres/columnes, histograma, caixes,...).
- Taules de contingència (observades, esperades, condicional, i indicadors d'associació).
- Estimació d'interval de confiança (proporcions i mitjanes).
- Diferència de mitjanes bivariant (ANOVA) i multivariant, per cas paramètric i no-paramètric.
- Tests per a comprovació de condicions d'aplicació de test d'hipòtesis (normalitat, homoscedasticitat,...).
- Matriu de Correlació (amb coeficients i significativitat).
- Regressió lineal simple i múltiple.
- Regressió logística.
- Anàlisi factorial de components principals.
- Anàlisi log-linear.
- Anàlisi de conglomerats.
- Possibilitat d'expansió de capacitats mitjançant addició de mòduls o paquets.

ANNEX 1.

LIST OF REQUIREMENTS

- Free of charge, open source software.
- Cross-platform (Windows, Mac and Linux).
- Operation through a graphic management system (menus, icons, windows).
- Particularly didactic for undergraduate students who are beginners in the use of statistical analysis tools.
- Menus and messages available, at least, in Catalan and / or Spanish. (*Participants in the contest may submit a first version in English provided that other languages location is possible, and must show its commitment to provide a later version in Catalan and / or Spanish*).
- Possibility of running instructions in R code (console mode).
- Data matrix view, and list of variables view with labels and other attributes.
- Separate results view.
- Ability to open and save data files in a variety of usual formats (csv, xls, sav, Rdata, etc).
- Export of the result tables and plots as publishable files (pdf, png, ...).
- Availability for tables and plots copy/paste operation into other applications.
- Capability to add and edit annotations to the generated results.
- Creating a new data matrix and entering data.
- Labeling variables, and fine control of its values and metric.
- Defining NA values.
- Inserting and deleting variables and cases.
- Creating new variables through transformations and calculations from existing variables.
- Recoding values of variables (by grouping categories).
- Merging data matrix, by variables and by cases.
- Case filtering, and segmentation of results based on variables.
- Frequency tables (relative, cumulate, %, ...).
- Basic descriptives (mean, sd, median, mode, min, max, quartiles, range, ...).
- Descriptive plotting (sectors, bars/columns, histogram, boxes, ...).
- Contingency tables (observed, expected, conditional, and association statistics).

- Confidence intervals estimation (for percentages and means).
- Bivariate and multivariate mean comparisons (ANOVA), for parametric and non-parametric cases.
- Checking the conditions of hypothesis test (normality, homoscedasticity, ...).
- Correlation matrix (with coefficients and significance).
- Simple and multiple linear regression modelling.
- Logistic regression modelling.
- Principal components factor analysis.
- Log-linear analysis.
- Cluster analysis.
- Expanding capabilities by adding modules or packages.