

SOFTWARE SELECTION THROUGH DECISION ANALYSIS AND INFORMATION SYSTEMS MANAGEMENT

Danilo Augusto Sarti

University of São Paulo, São Paulo, SP, Brazil

ABSTRACT

This paper aims to access the best alternative, from a set of five, of software for statistical analysis in a seed company in Brazil. The methodology considers the costs related to the process, R programming and its relationship with the final decision made regarding the software selection, being all these aspects framed by the tool discussed in (Howard, 1988; Howard, 2004) applied to the information technology context. The results present cost reduction in the process of statistical analysis and a change in the decision about the statistical software to be used by the enterprise. The paper consider only one specific application of analysis used by the enterprise which can be improved with the use of platforms such as Rstudio, and the packages Knitr and Shiny.

Keywords: Decision analysis; Information technology management; Open source/R programming; Software Selection; IT in agribusiness.

1. INTRODUCTION

1.1. Objectives:

The main objective of this paper was the use of the tool for decision analysis discussed in (Howard, 1988; Howard, 2004), and described in theoretical references, to select the best software application alternative, from a set of five alternatives, to proceed the Tukey test, an statistical test used by the enterprise. Four from these alternatives were programmed in R and one alternative was the use of commercial software for the statistical analysis.

The decision analysis tools presented in (Howard, 1988; Howard, 2004) was elected to minimize cognitive problems and other kind of bias related to the decision process and broadly described in (Bell & Raiffa, 1990; Hogarth, R, 1991; Plessner & Bletsch & Bletsch, 2007).

Manuscript first received/*Recebido em:* 23/07/2013 Manuscript accepted/*Aprovado em:* 09/02/2015

Address for correspondence / Endereço para correspondência

Danilo Augusto Sarti, Ex-aluno LES/USP, Eng Agrônomo/USP Msc. Applied Economics/USP, Datascientist, Aluno especial do IMECC/Unicamp, Address for correspondence: Av 33 1999 Estádio 13501440 Rio Claro SP daniloasarti@gmail.com (preferred contact).

Published by/ *Publicado por:* TECSI FEA USP – 2015 All rights reserved.

In despite of the key objective, the paper also tried to answer other questions and issues. Among them we can cite the quantification of the costs related to the statistical analysis process in each alternative and quantification of cost reduction obtained during the process of construction of the decision bases (concept to be define in theoretical references)

The paper also shows how a function in R can serve as an efficient interface and the influence of the interface level on user's preferences to the considered alternatives.

More broadly, the paper also demonstrates the possibility of using open source platforms in enterprises and how the abilities, such as R programming, made an impact on the decision regarding the most appropriate software for the enterprise.

1.2 Justifications:

The work is justified by the illustration of how to solve the problem of software selection in a formal manner. This formalization allows recording the information used at the moment of the decision making.

The example of the use of decision analysis in IT management can serve as a base for future systems choices and can also be used to decision making in other sectors.

The cost reduction obtained with the resolution of the problem was significant, showing the great potential of using open source software, in a creative context, in different economic sectors, resulting in competitiveness to organizations. The costs question is important for any kind of business, especially for new enterprises and small business, in which the initial investments are made and results are uncertain.

Results also allow to say that the use of open source alternatives in a creative way and its teaching at the universities can reduce the technological dependence of enterprises on software with high license costs.

1.3 On the demand and importance of statistical software

Enterprises that work with applied agriculture research have a big dependence on statistical analysis of the data collected in the experiments they use to conduct. This is the case of plant breeding and the seed production sector that use statistical analysis for decision making in different steps of their production and research processes.

Usually, these statistical analyses are made in software that produces reports for the decision maker or analyst involved in the process of product development.

In the plant breeding process, there are many statistical analyses used to answer questions about the performance of the products regarding productivity and other characteristics related to the vegetal material experimented.

In general, the product's tests are made in an appropriate statistical design to answer the key questions. A typical kind of statistical analysis, which comprises a big number of reports in plant breeding process, are mean comparison tests such as Tukey's test.

The production of such statistical analysis implies in costs for companies, mainly costs related to license payments for commercial software use, in addition to costs related to the statistical analysis itself. For small companies such costs can be significant and have impact on their competitiveness.

Thus, decisions related to the software used, its efficiency to solve the specific application and other aspects are strategic issues for the management of information technology and systems at the enterprise.

1.4 Potential, popularity and easy use of free and open source software

With the advent and consolidation of open-source and free platforms under GPL license, the number of alternatives for software and statistical software has increased. A good example is the R statistical software.

(Muenchen, 2012) discusses the rise in popularity of R in academic and enterprises or other institutions applications. The same author shows the advance and versatility of the software through the increase in the number of scientific papers citing R, number of messages in forums specialized in data-science, etc.

The use of free and open source platforms is interesting because of the two main aspects. First, the user can know the details of the procedures made by the software once its code is open, and in case of good familiarity with the software, the user can adapt the code to their own application. Considering R software to see the code used by a function one needs, just type in the name of the function on the software console and will see its code.

The second aspect is the reduction of bureaucracy to obtain the software by the enterprise. In general, for acquiring a system that requires license payments, the enterprise needs to start an internal process for buying the software that involves different sectors of the enterprise and that can delay the access of the software by the user, considering urgent demands this can be a bad situation. In the case of free software, the user can download it and freely use the system on their own desktop.

2. THEORETICAL REFERENCES:

2.1 Decision analysis and its implementation

The theoretical reference for choosing the best statistical application for the analysis system of the enterprise was the tool presented in (Howard, 1988; Howard, 2004).

According to this author, the search for a good decision needs to involve an efficient characterization of the alternatives related to the decision problem, its consequences, preferences of agents related to the problem and, in uncertain conditions, a complete description of the risks (Howard, 1988; Howard, 2004).

This tool allows an iterative and comprehensible modeling of the decision problem, in addition to making the process easily communicated among people of technical areas and non-technical staff related to the decision, minimizing the cognitive problems described in Bell & Raiffa, 1990; Hogarth, R, 1991; Plessner & Bletsch & Bletsch, 2007) and (Ariely, D, 2013). The main ones are:

- a) **Anchor:** Psychological phenomenon in which the mind gives a non-proportional weight to the first received information about the problem.
- b) **Status quo Bias:** A tendency that decision makers have to choose alternatives that maintain the status quo of the institutions and groups.
- c) **Protect earlier decisions:** Tendency to maintain aspects related to earlier taken decisions in despite of the fact that such maintained aspects are not reasonable.
- d) **“See what we want see”:** Bias that implies searched information that provides proof that the first assumptions about the problem are right.
- e) **Incorrect framing of the decision problem, overconfidence and neglected information.**

As previously described, challenging problems can be minimized by the use of the decision analysis tool presented in (Howard, 2004) which main objective is to obtain a set of information called “Decisions Base”. The whole process can be seen in Figure 1.

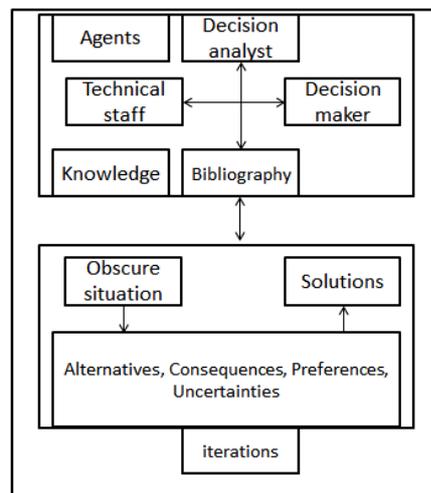


Figure 1
Decisions Base: Construction and dynamics. Source: Author

1

In Figure 1, it is possible to note that the process of decision analysis and construction of decision base starts in an “obscure” state regarding the decision problem.

From Figure 1, we can see that agents in technical areas, the decision analyst and decision maker, work in an integrated manner sharing and complementing experiences with the available literature about the problem. Through this exchange, the knowledge for the construction of decision base is born, which shows the alternatives, consequences and related preferences and uncertainties in each alternative. The process should be repeated iteratively until the solution to the problem is given and a decision is made.

In general, to make a decision we start with an initial decision base that evolves with the iterative description of the alternatives, consequences, preferences and risks until a set of information, enough to base a decision upon, is provided.

2.2 Statistical Theory related to the problem

As for the statistical topic regarding the application of a test, we consider conducting a means comparison test among the different treatments in a complete randomized blocking design. Statistical reports with this kind of test, in spite of the fact of being simple and questionable, are the base to key decisions in the enterprise, from plant breeding process to registering the products at the Agriculture Department, impacting, therefore, the legalization of the products (MAPA, 2014).

2.3 Tukey's test

The Tukey's test, reported by the software selected by the enterprise, is a classical technique used in design and analysis of experiments. (Montgomery, 2013), (Cox & Reid, 2000; Hinkelmann & Kempthorne, 2005; Mason et al. , 2003).

Such technique consists in comparing the means of all $a*(a-1)/2$ pair of treatment means used in an experiment (Montgomery, 2013). In the case considered in this paper, the experiments were conducted in a complete randomized blocking design and the compared means were productivity means of pre commercial hybrids of normal grain corn of the considered enterprise's portfolio.

The test is based on a maximum value for the difference of two given treatment means so that such treatments are considered statistically equal; in other words, they have to be in the same statistical class if the difference of such means is less than a given distance taken as maximum accepted difference. This allows them to group the developed products in classes of productivity means (Montgomery, 2013).

3. INFORMATION TECHNOLOGY

The knowledge of information technology used can be divided into programming topics, R programming topics and its availability and structure.

3.1 General theoretical programming topics

General topics regarding computer programming were used as techniques to develop algorithms for problem solving, in addition to general subjects regarding the implementation of procedures. Detailed information about such subjects can be obtained in (Cormen et al., 2009; Muller, 2006; Knuth, 1997). For a more statistical programming point of view we suggest (Givens & Hoeting, 2013).

3.2 R software: obtaining and short history

R is an open source and free software available for all operations systems at www.r-project.org. The software is a multi-objective platform and allows the analysis and visualization of statistical data at different levels of complexity. It also can be used for a new and specific application development in a completely new and unpublished context or for mixing new and pre-established codes from the software.

The software has a basic installation mode that provides a basic development environment, a console for analysis programming and a graphical interface. The basic

mode and environment are complemented by packages for specific the application in areas such as math, statistics and econometrics, ecology, finance and geospatial statistics.

R was born in 1995 and today in managed by a group of people called R-core and maintained by the R-foundation, which is supported by different enterprises and institutions.

According to (Muenchen, 2012) R is the most used software by world wide data analysis competitions such as Kaggle, more details at www.kaggle.com.

3.3 R programming

R allows the user to use a variety of programming solutions. Its base code uses well established routines made in Fortran, C e in R language itself. A great advantage of R is the production of functions to solve new and specific applications, including the programming of more user friendly applications by non-technical users.

People interested in R-programming can access the R website and its tutorial section where it is possible to find different level materials and tutorials on R, in many languages. For a structured R introduction, we suggest the reading of (Teetor, 2011), (Micheaux & Drouilhet & Lique, 2013) e (Gardener, 2012).

4. METHODOLOGY

The following methodology aims to apply the theoretical reference to the selection of statistical software problem.

4.1 Problem Framing and characterization

According to (Howard, 1988; Howard, 2004), the first step for the construction of the decision base and to decision analysis process is the framing of the decision problem.

In this paper, the problem was structured as to opt for one in a set of five alternatives for Tukey's test implementation used by the enterprise.

4.2 Considered Alternatives

Four alternatives programmed in R were considered in order to make the test required by the enterprise , with different approaches and interface levels, in addition to a commercial software solution previously known by members of the statistical analysis team of the enterprise and that is why it was suggested.

The R-programmed alternatives were called alternative1-R to alternative4-R, however, their production was made iteratively through the decision base construction process from a given start scenario, which included only alternative1-R, whcih evolved to a final scenario with the five alternatives to be considered in the decision.

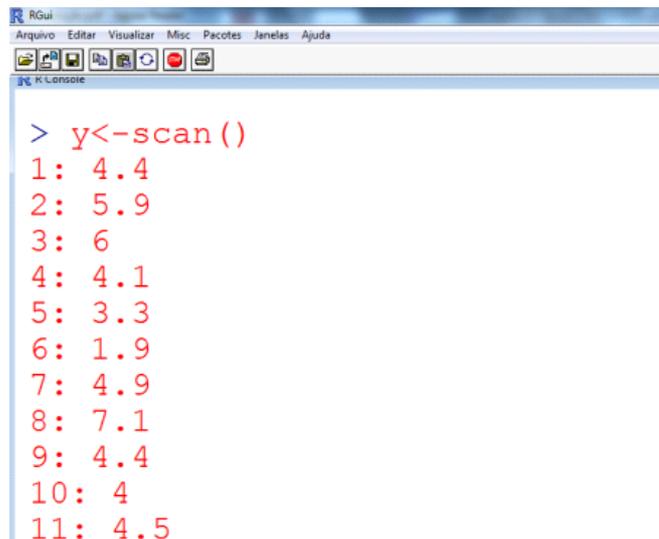
When showing the dynamics of the alternatives, we only talk about the commercial one, without figures or screenshots that could identify it and hurtto the image or market of the software and its distributors.

4.2.1 Alternative 1-R:

The first alternative programmed in R was based on the `scan()` function present in the basic installation mode. An important aspect of this alternative was the fact that it was possible to perform the manual typesetting of the data on R console and then to program the analysis and the test required by the enterprise.

Figure 2 shows the general dynamics of alternative1-R. In such a Figure we can see that for the test implementation the user needs to call the `scan()` function and typeset all the data manually.

As the data collected and analyzed were from the order of thousands, the manual typesetting was very time consuming.



```
RGui
Arquivo  Editar  Visualizar  Misc  Pacotes  Janelas  Ajuda
R Console
> y<-scan()
1: 4.4
2: 5.9
3: 6
4: 4.1
5: 3.3
6: 1.9
7: 4.9
8: 7.1
9: 4.4
10: 4
11: 4.5
```

Figure2
Alternative1-R dynamics. Source: Author

4.2.2 Commercial Alternative:

The commercial alternative was an application in the commercial software which required license payment for its use. Through this application, a file with the data previously collected was read by the software and then the analysis was made in a user friendly interface involving pre-established routines known by some members of the analysis team.

4.2.3 Alternative2-R:

Alternative2-R was based on the use of a function for `csv` (comma-separated values) importation. After the data file import, the user could produce the statistical analysis on R console without manual typesetting.

However, the data organized in `.csv` file should be in the frame shown in Figure 3 which represents the tabular structure showing the columns `bloco` (corresponds to blocks), `tratamentos` (treatments) e `produção` (production) required by the code programmed to run the analysis.

In spite of the fact that this alternative solves the problem related to manual typesetting of the great amount of data, the structure of the file .csv was different from the one generated by the data collection system shown in Figure 4, already automatized and well established in the data analysis team.

1	bloco	trat	prod
2	1	1	4.4
3	1	2	3.3
4	1	3	4.4
5	1	4	6.8
6	1	5	6.3
7	1	6	6.4
8	2	1	5.9
9	2	2	1.9
10	2	3	4
11	2	4	6.6
12	2	5	4.9

Figure 3 Data structure in the file used by Alternative2-R. Source: Author

	A	B	C	D	E
1	tratamento	b1	b2	b3	b4
2	1	4.4	5.9	6	4.1
3	2	3.3	1.9	4.9	7.1
4	3	4.4	4	4.5	3.1
5	4	6.8	6.6	7	6.4
6	5	6.3	4.9	5.9	7.1
7	6	6.4	7.3	7.7	6.7

Figure 4 Data Structure reported by the data collection system and consolidated in the enterprise. Source: Author

4.2.4 Alternative3-R:

As a consequence of the frame of the csv file consolidated in the enterprise and shown in Figure 4, alternative3-R was programmed. It is important to observe that the tabular structure is different from the one shown in Figure 3 generating consequences for the code structure that would use the tabular disposition represented in

Figure 4.

The alternative3-R included in its code procedures that allow the data analysis from the data imported in the frame shown in Figure 4. In spite of this, the mechanism to import the file was basically the same as alternative2-R.

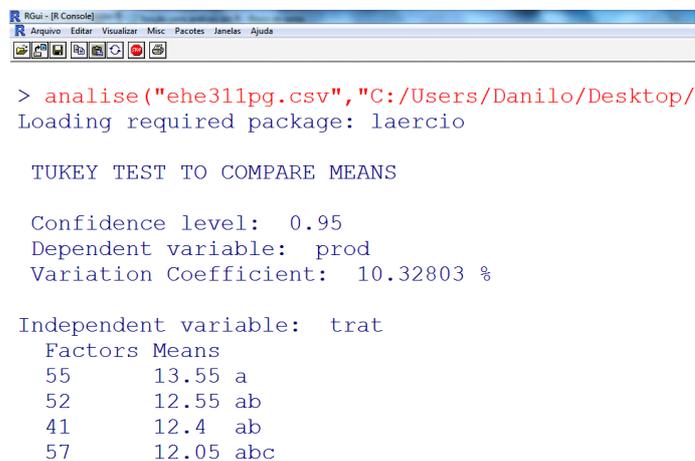
After importing the file with data, the user could conduct the programming for the data analysis and make the test to compare means.

4.2.5 Alternative4-R

The fourth alternative programmed in R was based on a completely new R programmed function. The aim of such functions was a more user friendly data analysis for Tukey's test.

The function named *analise* has as its arguments, the csv file name with the data and frame consolidated in the enterprise, in addition to the directory where the file was saved.

After calling the function and giving the names of the two arguments, the user just presses enter on the keyboard and the report is generated without the need of manual typesetting and manual code programming on the console. The general dynamics of this alternative is shown in Figure 5, where we can see that the user is only required to name the function its arguments. This solutions uses in its code the *package* called *laercio* and gives the confidence levels and variation coefficient of the experiment (Montgomery, 2013)..



```

> analise("ehe311pg.csv", "C:/Users/Danilo/Desktop/
Loading required package: laercio

TUKEY TEST TO COMPARE MEANS

Confidence level: 0.95
Dependent variable: prod
Variation Coefficient: 10.32803 %

Independent variable: trat
Factors Means
55      13.55 a
52      12.55 ab
41      12.4  ab
57      12.05 abc

```

Figure 5 Alternative4-R dynamics. Source: Author.

4.3 Consequences characterizations

To construct the decision base (Howard, 1988; Howard, 2004) it is necessary to determine the consequences of each alternative considered.

In this paper the main consequence considered was the cost related to the process of statistical analysis, modeled by the following equation:

$$CA = CL / NA + CHF * Th$$

Equation Costs modeling.

1

In this equation,

CA = Analysis cost in dollars.

CL = Cost due to license payment for software use (paid by computer on which the software is installed).

CHF = Hourly costs referred to the analyst that makes the analysis.

NA = Number of analysis made

Th = Time required to proceed the analysis.

More details about cost theory and economic theory of costs can be found in (Silberberg & Suen , 2001) and (Stackelberg, 2014).

4.4 Preference Assessment

Preference assessment can be made at different complexity levels when performing decision analysis. In this paper we adopted the following methodology considering that the work was done in an enterprise context.

For preference assessments, Tukey's test was conducted in each of the alternatives with each member of the analysis team in separate to minimize the opinion bias.

Then each member, also in secret conditions, was invited to score the solution from 0 to 10 as a way of showing their preference for the alternative considered.

In this process, the members of the team were invited to judge each alternative without considering past experiences with this kind of analysis, aiming to minimize the tendency of the user with prior commercial software alternative experience to evaluate this alternative with higher scores.

5. RESULTS

5.1.1 Decision base construction

In earlier stages of the decision regarding which software to use, the decision base was formed by two alternatives: commercial alternative and alternative1R. The composition of decision's base is shown in Table 1.

Alternative	Consequences: costs	Preferences	characteristics
Commercial	\$122,08 /analysis	8	No programming skills required, user friendly interface and ready analysis routines, requires license payments
Alternative1-R	\$7,60 /analysis	7	Manual typesetting of data on R console and programming skills required, no license payments required

Table 1
Initial Decision's Base for the problem.

Source: Author.

In this initial stage, the decision was to choose the commercial software, however, after the presentation to the enterprise staff about the possibilities of

developing applications in R, the directors made an agreement to allow the development of new alternatives, as long as these processes did not incur other costs and the decision was made by the decision analyst, who had experience with R programming.

Then the decision base was iteratively updated to a second stage and included alternative2-R, described in methodology. Thus, the decision base was as shown in Table 2.

Alternative	Consequences: costs	Preference	Other characteristics
Commercial	\$122,08 / analysis	8	No programming skills required, user friendly interface and ready analysis routines, requires license payments
Alternative1-R	\$7,60 /analysis	7	Manual typesetting of data on R console and programming skills required, no license payments required
Alternative2-R	\$4,2/analysis	8,5	Allowed data file import in csv file, however, the tabular file frame was not the one consolidated in the enterprise. No payments of license.

Table 2
Second stage of decision base. Source: Author.

Given the evidence presented in the second stage of decision base, it was decided to perform another turn for generation of new alternatives and characterization of its costs, preferences and aspects.

In a first moment alternatives2-R had its code changed to permit the use of the consolidated tabular structure of csv file used by the enterprise, in which the data to be imported were set, generating alternatives3-R. Then the process evolved to the development of alternatives4-R based in a function that makes the required analysis in a friendly user context. After this, the decision base had the structure shown in Table 3.

Alternative	Consequences: costs	Preference	Other characteristics
Commercial	\$122,08 / analysis	8	No programming skills required, user friendly interface and ready analysis routines, requires license payments
Alternatives1-R	\$7,60 /analysis	7	Manual typesetting of data on R console and programming skills required, no license payments required
Alternatives2-R	\$4,2/analysis	8,5	Allowed data file import in csv file, however, the tabular file frame was not the one consolidated in the enterprise. No payments of license.
Alternatives3-R	\$4,22	9	The same as above, however, its code allowed the use of consolidated csv file
Alternative4-R	\$0,9	9,5	Based on a function whose arguments were the name of csv file and the directory where they were saved

Table 3

Final decision base for the problem of statistical software selection. Source: Author.

5.1.2 Decision Made

Given the last decision base constructed in an iterative process represented in Table 3, the enterprise decided to choose alternatives4-R as a pattern to its analysis of means comparison.

Such an alternative was chosen because it had the highest preference score, lower cost per analysis and had user friend characteristics for users without R programming skills.

This decision opened also a set of possibilities for the use of data science and machine learning techniques to the collected data given the compatibility of R with these techniques.

Thus, in addition to the means comparison team, other kinds of more complex specific analysis were possible, a fact that had a great impact on staff of the enterprise involved in the statistics management of information system data analysis.

6. DISCUSSION

6.1 Impacts of R programming skills on the final decision

We verified that, as a result of decision base construction, R programming skills impacted the decision about the best software alternative to be used and as consequence about the software to be used.

The main impact was the change in the decision from the commercial alternative to alternative4-R and consequently no need of license payments.

Thus, we show as the skills for production of application in open and free platforms can reduce the technological dependence of organizations. This result also allows recommending the teaching of such platforms in universities and technical courses, as a strategy to build human capital that could reduce the technical dependence of organizations.

6.2 Evolution of interface level among alternatives programmed in R and its impacts on costs and preferences

In the iterative process of alternatives development for conducting the means comparison tests, there was an evolution in interface level between the alternatives and the users. This increase in interface level was related to the evolution in preference of the user for the alternatives developed, shown in Figure 6. The highest interface level was achieved with alternative4-R.

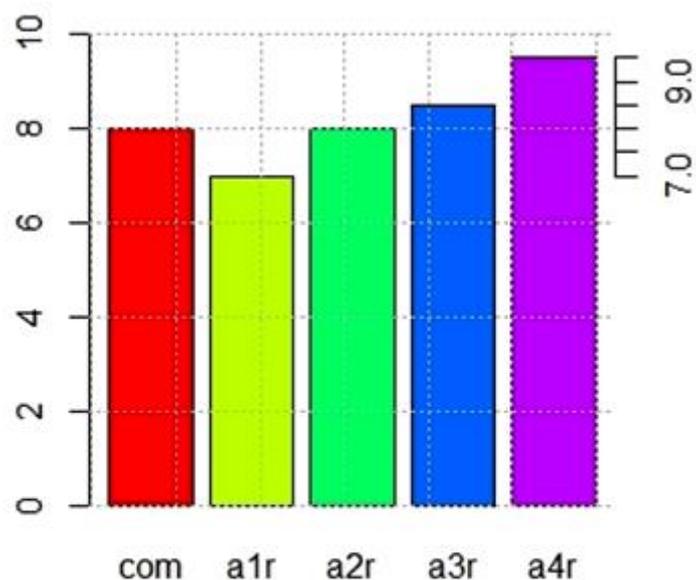


Figure 6
User preferences (vertical axis) and considered alternatives. Source: Author.

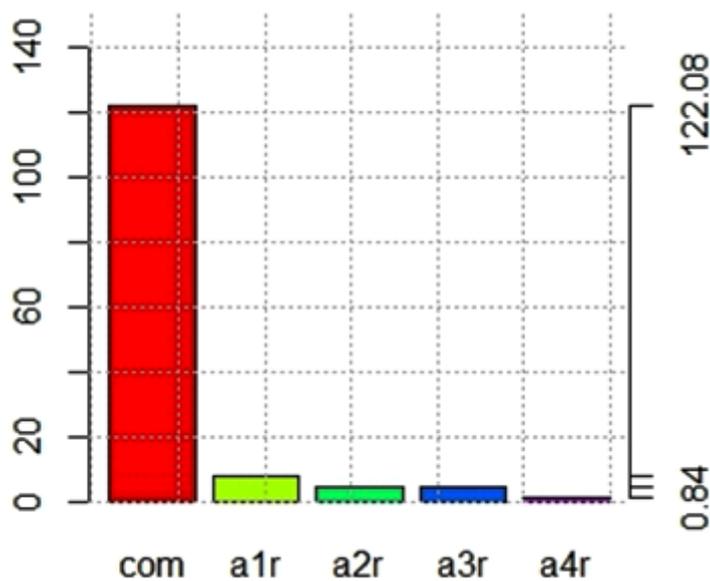


Figure 7
Cost/analysis (\$) (vertical axis) and alternatives. Source: Author

Comparing Figures 6 and Figure 7, which represent the costs of each alternative, we can infer that the increase in interface level resulted in an additional cost reduction due to the fact that the license payment is not necessary, which is represented by the difference of the red bar (commercial alternative) and the green bar representing Alternatives1-R.

This fact shows the impact of the evolution in interface level on user's preference and on the costs related to this application.

7. CONCLUDING REMARKS, LIMITATIONS E RECOMMENDATIONS FOR FUTURE WORK

This paper showed, in an interdisciplinary context, the use of tools for decision analysis in IT management. Through the use of the tools discussed in (Howard, 1988; Howard, 2004), it was determined the best alternative, from a set of five, to conduct a means comparison test with the use of statistical software in an enterprise in corn breeding and seed production.

A reduction of the cost related to analysis process was obtained. The paper also demonstrated how an R function can serve as an efficient interface and the influence of interface level on user's preference regarding different alternatives.

The possibility using free and open platforms in enterprises was shown. In addition, it was demonstrated how R programming skills impacted the decision regarding the software to be chosen for the statistical analysis system of the enterprise.

As limitations of the work we can cite the solution of just one specific application, to adjust the work to daily activities of the enterprise in which the work was done. More robust solutions considering all the applications used by the enterprise would be possible. Such solutions could use a package with functions for all the applications used in the enterprise to be loaded on R console in a perspective similar to the one proposed in (Wickham, H. 2014). Another limit of the work was that the solution was conducted in only enterprise in the related sector. If the possibility of solving the same problem in other enterprises exists, a more generalist solution could be built considering the whole sector.

We recommend for future works an broader integration of tools for generating the statistical reports by the selected alternative that comprises the use of tools as Rstudio, available at <http://www.rstudio.com> and the use of packages Knitr, which uses markdown structures to generate reports, and Shiny for dynamic applications building in R, what would make the scores of users preferences for the alternative greater that the ones observed in this paper, given the fact that such tools would produce an even more user friendly environment for the user of the application. The packages discussed are available on the R website.

ACKNOWLEDGMENT TO CRIAGENE SK NORMAL AND SPECIALTY CORNS RESEARCH AND DEVELOPMENT. THEY SPONSORED THE WORK AS PART OF THEIR COMMITMENT TO DATA SCIENCE DIFFUSION ON AGRIBUSINESS

REFERENCES

- Ariely, D. (2013). *The Irrational Bundle: Predictably Irrational, The Upside of Irrationality and The Honest Truth About Dishonesty*. New York. Harper.
- Bell, E. & Raiffa H. (1990). *Decision Making: Descriptive, Normative and Prescriptive Interactions*. London, Cambridge Press.

- Cormen, T. H. & Leiserson, C. E. & Rivest, R. L. & Stein (2009). *C. Introduction to Algorithms*. Cambridge MA, MIT Press.
- Cox, D. R. & Reid, N (2000). *The Theory of the Design of Experiments*. Boca Raton, Chapman and Hall.
- D. Knuth (1997). *The Art of Computer Programming: Fundamental Algorithms*. Massachusetts, Addison Wesley Longman.
- Gardener, M. (2012). *Beginning R: The Statistical Programming Language*. New York. John Wiley and Sons, Inc.
- Givens, H. G. & Hoeting, J.A. (2013). *Computational Statistics*, New York. John Wiley and Sons, Inc.
- Hinkelmann, K. & Kempthorne (2005). *O. Design and Analysis of Experiments*. New York, JohnWiley and Sons.
- Hogarth, R. (1991). *Judgement and Choice*. New York, John Wiley and Sons.
- Howard, R. (1988). *Decision Analysis practice and promise*. Management Science, 6, 679-675.
- Howard, R. (2004). *Speaking of decisions: Precise decision language*. Decision Analysis, 1, 71-78.
- Mapa (2014), *Registro Nacional de Cultivares*. Retrieved from: <http://www.agricultura.gov.br/vegetal/registros-autorizacoes/registro/registro-nacional-cultivares>.
- Mason, R. L. & Gunst, R. F. & Hess, J. L (2003). *Statistical Design and Analysis of Experiments*. New York, John Wiley and Sons.
- Micheaux, P. L. & Drouilhet, R. & Liquet, B. (2013). *The R Software: Fundamentals of Programming and Statistical Analysis*. New York. Springer.
- Montgomery, D, (2013). *Design and Analysis of Experiments*. New York. John Wiley and Sons, Inc.
- Muller, J.M (2006). *Elementary Functions Algorithms and Implementations*. Boston, Birkhäuser.
- Munchen, R.A. (2012). *The Popularity of Statistical Software*. Retrieved from: <http://r4stats.com/articles/popularity/>.
- Plessner, H. & Bletsch, C. & Bletsch, T. (2007). *Intuition in Judgement and Decision Making*. London, Lawrence Erlbaum.
- Silberberg, E. & Suen, W. (2001). *The Structure of Economics: A Mathematical Analysis*. New York, Mc Graw Hill.
- Stackelberg, H. (2014). *Foundations of a Pure Cost Theory*. Berlin. Springer.
- Teetor, P. (2011). *R Cookbook: Proven Recipes for Data Analysis, Statistics and Graphics*. Cambridge, O'Reilly.
- Wickham, H. (2014). *R packages*. Retrieved from: <http://r-pkgs.had.co.nz/demo.html>.