Webology (ISSN: 1735-188X)
Volume 19, Number 3, 2022

Evaluation Of Mechanism For Decision Making In Smes To Achieve Competitiveness

Hugo G. Hernández Palma[1] , William A. Niebles Núñez[2] , Carlos Pacheco Ruiz[2]

[1] Business Administration Program, Universidad del Atlántico, Colombia.

[2] Faculty of Economic and Administrative Sciences, Universidad de Sucre, Colombia.

ABSTRACT

The benefits that mathematical sciences from statistics can generate for companies are still unknown. However, thanks to the technological advances that have taken place in recent decades, we can observe a rapprochement of business environments to this type of instrument. To analyze the most outstanding aspects that statistics offers for the management oof information at the enterprise level, the data collected after consulting research in Spanish and English are presented below. The results presented are a guide for entrepreneurs, students, and those who are considering the option of starting a business idea, to choose one of the above options as a support to achieve better results.

Keywords: Applied statistics, business, entrepreneurship, engineering, mathematics.

1 INTRODUCTION

Statistics are characterized as a mathematical science connected with the assortment, examination, translation, and presentation of mathematical information masses to make pertinent inferences [1]. Statistics are a type of numerical examination that utilizations evaluated models, portrayals and summaries for a given arrangement of trial information or genuine investigations [1].

Statistical methods are useful in many types of scientific research; they are particularly useful in situations where there is experimental uncertainty [2], with the main objective of supporting decision-making [3]. Real-world statistics are not only about calculating an average and a standard deviation, but they are also not always an exact and very precise science; statistics involve collecting data and filtering large amounts of information to a reasonable and accurate conclusion [4]; so, the researcher, once he has an important amount of data in his hands, will be obliged to decide what to do to find what they indicate [5].

Exploratory data analysis is an analysis performed by a researcher without any preconceived ideas to discover what the data tell them about the phenomenon studied[6]; it is concerned with developing an understanding of the data, considering to explore the nature of the distributions of the variables included, and the relationships between the variables[7]; on the other hand, the optimization process includes studying the response based on the combinations, estimating the coefficients, adjusting the experimental data, predicting the response and verifying the adequacy of the adjusted model [8]. Rencher and Christensen [9] highlight four types of data:

- One individual sampling with a great number of variables measured in each sample unit.
- One single sample with two sets of measurement variables, in each unit.
- Two samples with different variables measured in each unit.
- More than three samples with quite a few of measurement variables in each unit.

2 METHODOLOGY

The design of the research is documentary in nature, since it allows for a systematic process of collecting, organizing, analyzing, and interpreting information from different sources to address a specific topic [5]. In this case, this document is framed within this approach, since a bibliographic review of the most relevant references for the subject under study will provide a compendium of the tools applied at the business level from the statistical models.

3 USES OF STATISTICS AT THE ENTERPRISE LEVEL

One of the main uses of statistics is the description, inference, testing and prediction of information from visual inspection and numerical indicators; graphs are useful for searching for patterns or trends in a large amount of data: at the same time, simple arithmetic can be used to calculate unique numbers that summarize the data, which support analyses derived from visual inspection; from these results the statistician can make inferences or reach conclusions in the studies carried out [10].

In an increasingly complex world, researchers rarely use the entire population or case universe to conduct a study; instead, they use a subset of the population called a sample. While each research project is unique, four main forms of statistical use can be highlighted, represented in figure 1 [10].

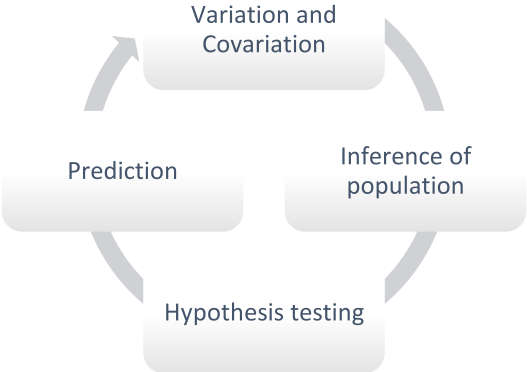

Figure 1. Use of statistics

The detailed components are defined as follows:

- **Description of the variation and covariation**

Descriptive statistics are tools utilized to sum up and tabulate data around the fluctuation in the distribution of values in cases with one or more variables [11]. Mendenhall et al. [12] define descriptive statistics as the procedure for summarizing and describing the important characteristics of a set of measurements; it consists of learning qualitative characteristics from data that were not previously known [11]; that is, summarizing large amounts of information in such a way that it is possible to communicate quickly, accurately, and honestly about the main patterns that are present in the data [10].

- **Inference of the population**

Another main task of statistics is to make inferences about fluctuations and covariations in a population, based on the information available from a sample drawn from that group [10]. A number that summarizes the variation or covariation in a sample is known as a statistic, on the other hand, if a number is calculated using the population, it is described as a parameter; likewise, this form of use is composed of two types of activities, the estimation of value from some statistics, and activities to assess the level of security of said estimation [10]. The inferential statistic then consists of the technique that allows us to study the samples and generalize about the selected population [13].

- **The Hypothesis Test**

Hypothesis testing begins with a speculation on how the world works, from which one logically deduces what should be observed in a particular sample to determine whether the theory is correct, and what might be observed if it is not correct. Hypothesis testing uses descriptive and inferential statistics to assess whether a set of observations is consistent with a theory; considering that a statistic is a characteristic of a sample, and that the parameter is the characteristic derived in the population, it is possible to establish that the hypothesis test uses sample statistics to arrive at better guesses or estimates of population parameters [10].

- **Prediction - Prediction**

In addition to describing, inferring, and testing hypotheses, statistical techniques are used to make predictions; in this sense, a model is created that describes the knowledge of how one or more causes produce a result, and then used to describe the patterns we see in the sample [10]. In the models for statistical prediction, tools of description, inference and hypothesis research are applied; the difference lies in the fact that the objective of analysis is the creation of a formal and mathematical description of the processes, in addition to establishing that said model can help explain current observations and predict new observations [10].

Predictive analysis is widely used in the treatment of Big Data, understanding this as the information captured from social interactions, in such a way that it seeks to discover patterns and determine relationships in the data. Predictive analysis techniques are divided into two groups; some techniques, such as moving averages, attempt to discover historical patterns in one or more resulting variables to extrapolate them into the future. Others, such as linear regression, aim to capture the interdependencies between the outcome variables and the explanatory variables and exploit them to make predictions. In this sense, the techniques available for this purpose could be marked as well into two groups: regression techniques and automatic learning techniques [14]. In general, the techniques for predictive analysis are related to the following characteristics:

- Heterogeneity: a common characteristic of Big Data originating from the source of the information, from various sources 14].
- Noise accumulation: some variables with significant explanatory power may be overlooked because of noise accumulation [14].
- Illegitimate correlation: refers to uncorrelated variables that are falsely correlated due to the large size of the data set [14].
- Incidental endogeneity: Refers to a genuine relationship between the variables and the error term [14].

To deepen the analysis of data, three stages are considered, the data entry, analysis and output results [15], represented in figure 2, which will be expanded upon below.

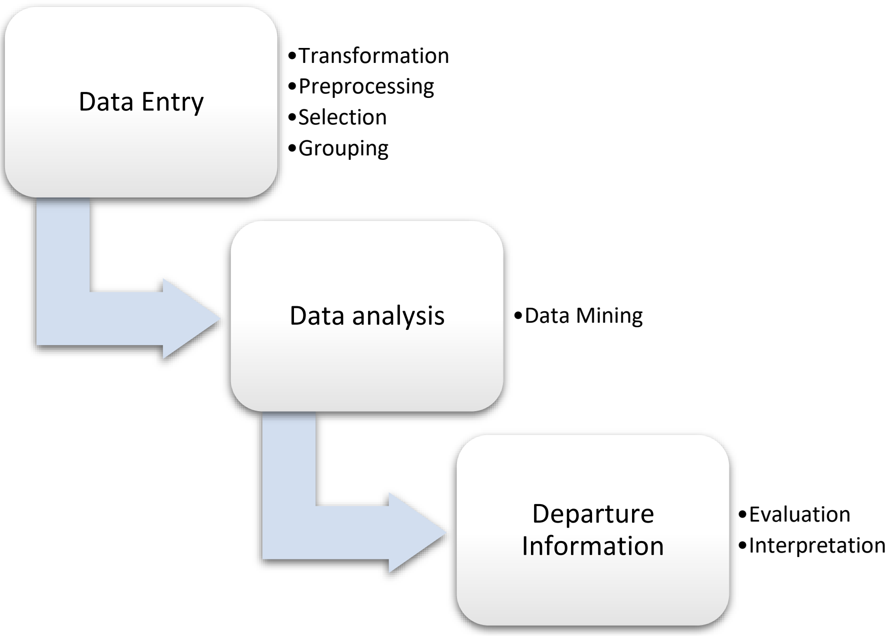

Figure 2. Knowledge discovery process in databases

- **Input of data**

According to what is seen in the figure, the operators of collection, selection, pre-processing, and transformation make up the entry stage, where the selection operator is generally the one who defines what type of data are necessary for the analysis, and in this way selects the relevant information from the data or databases collected. On the other hand, the pre-processing operator processes the input data with the aim of detecting, cleaning, and filtering unnecessary, inconsistent, and incomplete data into useful data. The transform operator then transforms the heterogeneous data into a single format compatible with the data mining technique, including dimensional reduction, sampling, coding, and transformation [15].

- **Analysis of data**

Data mining is used to perform information processing [15], analytical techniques must be carefully chosen to meet analysis objectives and data properties [16]; some of the most used algorithms are rules of association, automatic learning, metaheuristic algorithms and distributed computing. Grouping techniques are generally used to understand new input data, separating a set of unlabeled input data into different groups as defined by k-means. On the other hand, the classification techniques are based on a set of input data tags to build a set of classifiers [17], which will later be used to categorize untagged characters, such as the naïve Bayesian classification, decision tree-based algorithm, and support vector machine (SVM). Finally, the association rules seek to find all the concurrency relationships between the input data [15].

- **Departure Information**

Evaluation and interpretation are two results of the output. Evaluation regularly assumes the part of estimating results. It can likewise be one of the administrators of the Mining Algorithm information, for example, the number of square mistakes that was utilized. It is expected that the summary data will be provided in a concise and simple manner, as far as possible averaged over a graphical interface [15]. Myatt and Johnson [17] summarized the main components of a data analysis, shown in Figure 3.

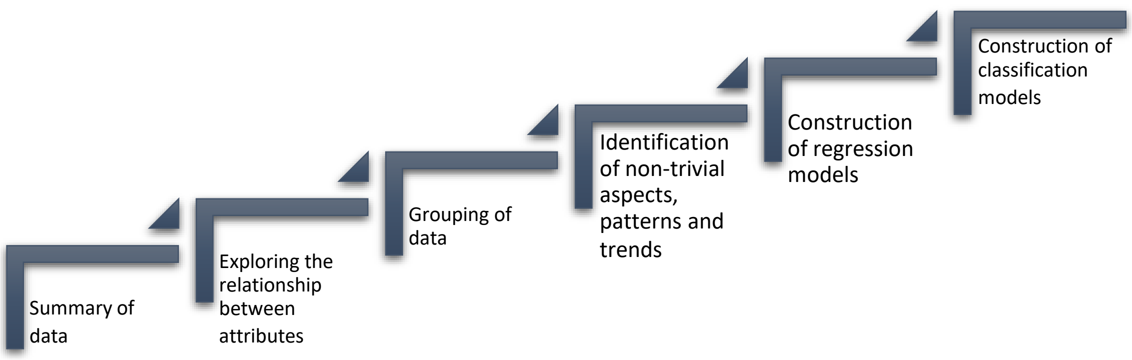

Figure 3. Components of a data analysis

In terms of data-intensive analysis, there are classic approaches used to assess group differences, either univaried, i.e., parameter by parameter, or using multivariate techniques [18]. Multivariate statistics refer to a variety of statistical methods that have been developed to handle situations involving multiple variables or measures [19]. In the case of multivariate data analysis, statistics help to study characteristic aspects of information [20]:

- Interdependence analysis: The objective of this technique is to transform the variables observed into a new set of uncorrelated and ordered variables in decreasing order of importance, reducing the dimensionality of the problem, and finding new variables that facilitate the understanding of the data.

- Factorial analysis: considers statistical models that take the error into account.

- Dependency analysis: implies the existence of dependent variables and descriptive variables. Multiple regression can explain the dependence of one variable on another, while multivariate regression analyses the dependence between several variables.

- Classification: uses cluster analysis to determine whether a data group forms groups, while discriminant analysis differentiates between different data and the group.

4 CONCLUSION

After carrying out this research, the final reflections point to the fact that the statistical tools are invaluable in facilitating decision-making. In a global world where economic changes are pushing organizations to permanently transform themselves, having instruments of this type adds value to business strategies.

Thus, in general terms, it can be said that mathematical tools are currently a highly appreciated tool for organizations, because due to the demanding dynamics of global markets, it is essential to have the possibility of complex and structured approaches, especially when defining aspects such as products, services and investments, among others.

As markets are a particular universe, it can be inferred that the reflections and the level of knowledge we want to have about them evolve as individuals, products, goods and services change. Thus, we can observe that there are engineering tools that allow us to perform such simple analyses that approach a single variable and require a low and attainable investment, like other studies that integrate several variables, broadening the scope of the analyses, but also require a larger investment.

REFERENCES

1. Trajkovski, V.: How to select appropriate statistical test in scientific articles, Journal of Special Education and Rehabilitation, vol. 17, nº 3-4, pp. 5-28, (2016).
2. Chatfield, C.: Statistics for technology. A course in applied statistics, Chapman & Hall/CRC, (2018).
3. Sachs, L.: Applied Statistics: A Handbook of Techniques, Springer Science & Business Media, (2012).
4. Jarman, K.: The Art of Data Analysis. How to Answer Almost Any Question Using Basic Statistics, Wiley & Sons, (2013).
5. Lee, E., Wenyu J.: Statistical Methods for Survival Data Analysis, Wiley, (2013).
6. Martinez, W., Martinez, A., J. Solka: Exploratory Data Analysis with MATLAB, Chapman & Hall/CRC, (2017).
7. O'Keefe, C.: Confidentialising Exploratory Data Analysis Output in Remote Analysis, Journal of Official Statistics, vol. 28, 4, p. 591–613, (2012).
8. A. E. A. y. I. R. F. Baradeswaran: A Statistical Analysis of Optimization of Wear Behaviour of Al-Al2O3 Composites Using Taguchi Technique. Procedia Engineering, vol. 64, 973-982, (2013).
9. Rencher, A.W. Christensen: Methods of Multivariate Analysis, Wiley, (2012).
10. Hanneman, R., Kposowa, A., Riddle, M.: Basic Statistics for Social Research, United States of America: Jossey-Bass, (2013).
11. Ivezic, Z., Connolly, A., Vander Plas, J., Gray, A.: Statistics, Data Mining, and Machine Learning in Astronomy, Princenton University Press, (2014).
12. Mendenhall, W., Beaver, R., Beaver, B.: Introduction to Probability and Statistics, Cengage Learning, (2012).
13. Gravetter, F., Wallnau, L.: Statistics for The Behavioral Sciences, Cengage Learning, (2016).
14. A. y., M. Gandomi, «Beyond the hype: Big data concepts, methods, and analytics, International Journal of Information Management, vol. 35, nº 2, pp. 137-144, (2015).
15. C. W. Tsai, C. F. Lai, H. C. Chao y A. V. Vasilakos: Big data analytics: a survey, Journal of Big Data, vol. 2, nº 21, pp. 1-32, (2015).
16. S. G. Heeringa, B. W. West, P. A. Berglund: Applied Survey Data Analysis, Chapman & Hall /CRC, (2017.
17. G. J. Myatt y W. P. Johnson: Making Sense of Data I. A practical Guide to Exploratory Data Analysis and Data Mining, Wiley, (2014).

18. J. Bartel, J. Krumsiek y F. J. Theis: Statistical methods for the analysis of high-throughput metabolomics data., Computational and structural biotechnology journal, vol. 4, nº 5, pp. 1-9, (2013).
19. G. Marcoulides y S. Hershberger, Multivariate Statistical Methods. A first course, Routledge, (2014).
20. C. Cahtfield y A. Collins, Introduction to Multivariate Analysis, Chapman & Hall /CRC, (2000).
21. Parody A. et al. Application of a Central Design Composed of Surface of Response for the Determination of the Flatness in the Steel Sheets of a Colombian Steel. In: Tan Y., Shi Y., Tang Q. (eds) Data Mining and Big Data. DMBD 2018. Lecture Notes in Computer Science, vol 10943. Springer, Cham (2018).
22. Parody A., Viloria A., Lis J.P., Malagón L.E., Calí E.G., Hernández Palma H.: Application of an Experimental Design of D-Optimum Mixing Based on Restrictions for the Optimization of the Pre-painted Steel Line of a Steel Producer and Marketing Company. In: Tan Y., Shi Y., Tang Q. (eds) Data Mining and Big Data. DMBD 2018. Lecture Notes in Computer Science, vol 10943. Springer, Cham (2018).
23. Ortíz-Barrios M., Neira-Rodado D., Jiménez-Delgado G., Hernández-Palma H.: Using FAHP-VIKOR for Operation Selection in the Flexible Job-Shop Scheduling Problem: A Case Study in Textile Industry. In: Tan Y., Shi Y., Tang Q. (eds) Advances in Swarm Intelligence. ICSI 2018. Lecture Notes in Computer Science, vol 10942. Springer, Cham (2018).