

# Representation of Learning Strategies through Visual Techniques of Multivariate Analysis in University Students

María Arely López Garrido<sup>1</sup>, Erika Yunuen Morales Mateos<sup>1</sup>,  
José Alberto Hernández Aguilar<sup>2</sup>, Carlos Alberto Ochoa Ortíz<sup>3</sup>,  
Carolina González Constantino<sup>1</sup>

<sup>1</sup> Universidad Juárez Autónoma de Tabasco Cunduacán, Tabasco,  
México

<sup>2</sup> Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos,  
México

<sup>3</sup> Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, Chihuahua,  
México

{arely.lopez, erika.morales, carolina.gonzalez}@ujat.mx, jose\_hernandez@uaem.mx,  
alberto.ochoa@uacj.mx

**Abstract.** This research aims to identify the learning strategies employed by university students; and represent them through graphic techniques of multivariate analysis. The sample consisted of 30 students from the Multidisciplinary Academic Division of the Rivers; of the degrees in Administration and Administrative Computing of the generational cohorts 2010, 2011 and 2012. The instrument used is the Inventory of Strategies of Learning and Motivational Orientation (EDAOM), this instrument consists of a self-assessment that the student makes on the learning strategies that employs. For the implementation of the visual techniques was used the Language R, with these techniques were represented individually to the students with their respective learning strategies. The results show the interaction of the students with the learning strategies and the academic performance, visualization that the students of the Degree in Administrative Computing obtain the lowest values in the strategies of performance and academic performance.

**Keywords:** EDAOM, learning strategies, data analysis, visualization, graphs.

## 1 Introduction

Higher education institutions have a large amount of data that has been generated from academic and administrative activities, in many cases this information has not been explored, analyzed and visualized. In this paper we used the data obtained from the

application of the Inventory of Strategies of Learning and Motivational Orientation (EDAOM), to new students; with the purpose of identifying the learning strategies employed by students of the Bachelor's in Management and Administrative Computing and represent them through graphic techniques of the multivariate analysis.

For the analysis of a set of data with many variables it is convenient to use multivariate visualization techniques, which allows in this case in an individual and group way the interaction of the learning strategies with academic performance; In this sense, several studies have been found in the literature in which they employ multivariate data analysis techniques such as GGE Biplots, generally applied to agriculture with the objective of evaluating the variability in the yield of corn grain and determining the response patterns between hybrids and environments, with the analysis of the biplot they identified the hybrids of superior performance [1]; another study consisted in determining the magnitude of the genotype x environment interaction in native corn populations of Chiapas in three contrasting environments, obtaining significant differences between environments, populations and interaction environments by populations [2], other studies of this type can be consulted in [3,4].

### **1.1 Multivariate Analysis**

The main objective of data analysis is to summarize and interpret data sets, detailing the contents and showing the important characteristics. In this sense the visualization of the multivariate data plays an important role in the fulfillment of the objectives [5].

Some studies in the educational area use graphical visualization techniques such as radar charts to make comparisons that are easy to interpret [6, 7]; In other studies, on university academic performance and the aspects that affect it, data analysis techniques have been applied [8, 9] or techniques of evaluation of results obtained [10].

Multivariate analysis is the part of statistics and data analysis that studies, analyzes, represents and interprets data that results from observing more than one variable of an object [11].

### **1.2 Visual Techniques of Multivariate Analysis**

The problems of multivariate space are difficult to understand, due to the large amount of data that must be represented in a very limited space; the problem is to consider a large number of variables and their relationships simultaneously. Traditional methods only produce a partial view of information. Visualization methods are the most convenient way to present data and achieve a graphic illustration of all the information in which it shows the essential facts, make comparisons and identify trends [5].

Advances in the development of automated information processing tools have simultaneously provided the creation and use of multivariate analysis techniques [12], increasingly sophisticated, dynamic and interactive graphs to visualize data or model results statistics [13].

Through the visual representation of data, the relationships described by graphs are easily understood and are easy to remember, in addition to a graph are displayed large amounts of data [14].

The graphs used for the visual representation were: of stars and GGE Biplots representing each of the individuals of a population or sample.

Stars's graphs represent a method for representing multivariate data in which each star represents an element of the sample or population, and each axis represents a variable, its length is proportional to the value of the variable. The ends of the rays are connected with segments of straight lines to form a star [15]. Star charts allow the identification of dominant variables in an element of the population; visualizes if there are similarities between the elements of the population and therefore if there are groupings, also the atypical values can be identified if they exist [16].

The GGE Biplots graphics employ the combined effects of genotypes (G) and genotype-environment interaction (GE), which facilitate the visual identification of genotypes and evaluation environments. These graphs are constructed using the first two main components, derived from the decomposition of the combined effects of G + GE. The first component, when highly correlated with the main effect of the genotype, represents the proportion of the yield that is due only to the characteristics of the genotype. The second main component represents the part of the yield due to the genotype-environment interaction [17].

The importance of identifying the learning strategies used by the students consists of knowing the causes of low or high performance [18], that is, improving learning outcomes can be achieved by incorporating assessment procedures to identify risks and determining strategies should be encouraged in students [19].

## **2 Method and Tools**

The scope of this research is descriptive because they will specify properties and characteristics describing trends, the design is non-experimental transactional [20].

The population is made up of the new students belonging to the cohorts 2010, 2011 and 2012 of the Bachelor's Degree in Administration and Administrative Computer Science, the sample is of 30 students who answered in their entirety the inventory of Strategies of Learning and Motivational Orientation (EDAOM).

### **2.1 Inventory of Learning Strategies and Motivational Orientation (EDAOM)**

The instrument used to know student learning strategies is the Learning Styles and Motivational Orientation questionnaire (EDAOM), the result of this questionnaire is a self-assessment that the student makes about their learning strategies and motivational orientation to the study [19]. The self-report section measures self-assessments of students on: a) frequency, b) ease or difficulty, and c) the results of using a wide variety of learning strategies.

The EDAOM is composed of four scales and 13 subscales; the Self-regulation, Metacognitive and Metamotivational scale consists of three components: those of the person, those of the learning task and learning materials; Table 1 shows the structure of the EDAOM, the first column specifies the scales and the second the subscales that make it up.

**Table 1.** EDAOM scales and subscales.

Scales	Subscales
Acquisition of information	Selective Generative
Recover of information	Subjected tasks During exams
Processing information	Convergent Divergent
Self-regulation, Metacognitive Metamotivational	Person: Efficiency perceived Internal Contingency Perceived autonomy Orientation to external approval Learning Task: Orientation to the task itself Orientation to the achievement of goal
	Materials

In this work the learning strategies were analyzed: selective, generative, task recovery and during exams and convergent and divergent processing.

## 2.2 Software Used for Multivariate Analysis

A very important element to implement the techniques of multivariate analysis are the tools, in this case we used the software SPSS and language R.

IBM SPSS Statistics 22 is a software for statistical analysis that contains the primary functions to perform a complete analysis process. It is easy to use and contains an extensive category of procedures and techniques that help direct research and make the best decisions. Its main characteristics are to have a wide range of statistical procedures to perform accurate analysis, includes techniques to prepare data for analysis quickly and easily, can create very efficient charts for reporting and supports all types of data [21].

R is a free software for statistical computing and graphics. Compile and run and compile on a variety of UNIX, Windows and Mac OS platforms. It is widely used for the development of statistical software and data analysis, provides a wide variety of statistical models: linear and nonlinear models, classical statistical tests, time series analysis, classification, grouping among others), and graphic techniques. R is an environment in which statistical techniques are implemented and which is extensible through packets, there are about eight packets that are supplied with the R distribution [22].

## 2.3 Descriptive Analysis

A descriptive analysis of the learning strategies was carried out: selective, generative, retrieval of information before tasks and during exams, processing of convergent and

divergent information, the results obtained are shown in Table 2, in this it is observed that the value minimum for the majority is 60 or more with the exception of the acquisition of selective information that obtained 54 and the processing of divergent information 51; in the maximum value values equal or greater than 90 were obtained; in terms of the average the strategy of acquiring selective information and retrieval of information before tasks obtained a value greater than 70 and the other strategies values greater than 80. The minimum average bachelor is 7.30, the highest average is 9.95 and the mean of the averages is 8.70.

Academic performance is classified as low, regular and high; was derived from the indicators: approval rating in ordinary, average and promotion index in the first year [23]. Table 3 shows that 36.7% of the sample had low academic performance; 46.7% high and 16.7% regular.

**Table 2.** Descriptive Statistics.

<b>Subscale</b>	<b>Min</b>	<b>Max</b>	<b>Half</b>	<b>Standard deviation</b>
Selective	54	94	73.03	8.818
Generative	69	100	82.90	10.018
Before Task	60	90	72.97	7.025
During Exams	60	100	85.77	9.957
Convergent	63	97	81.60	10.074
Divergent	51	100	81.27	10.041
Average High School	7.30	9.95	8.7037	.75926

**Table 3.** Frequencies of Academic Performance.

<b>Performance</b>	<b>Frequency</b>	<b>Percentage</b>
<b>Low</b>	<b>11</b>	<b>36.7</b>
<b>Regular</b>	<b>5</b>	<b>16.7</b>
<b>High</b>	<b>14</b>	<b>46.7</b>

### 3 Results

Pearson's correlation was made between learning strategies and academic performance; the correlation measures the linear association between learning strategies, the value of the association may be from -1 to 1, the closer the value is to 1 the greater the association between variables [24]. Figure 1 shows the correlation between learning strategies and baccalaureate average and academic achievement in which there is a strong association between the average baccalaureate and academic achievement; learning strategies are correlated with values greater than .5 and .8.

The correlations between the subscales of Acquisition, Recovery and Processing of EDAOM Information, it is observed that they are related to each other, so it is

determined that they are not mutually exclusive, that is, a student can have a good command of the different learning strategies [25, 7], in the same way a strong association of the high school average with the academic performance obtained in the first year of the bachelor's degree is observed. The subscales that present a higher relation considering the selection criteria equal to or greater than .70, according to Figure 1, identify the selective subscale with recovery during the exams and convergent processing; and the generative with during the exams and divergent.

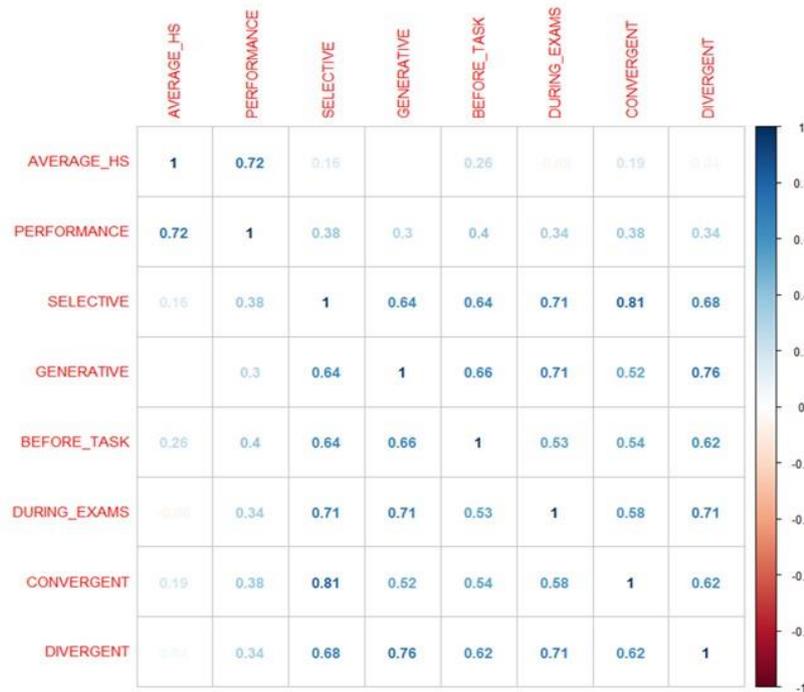


Fig. 1. Correlation of EDAOM learning strategies and academic performance, Source: (Self-realization, 2017).

### 3.1 Data Analysis with Star Graphics

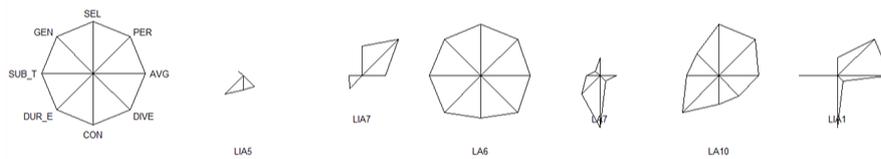
The chart of Stars with the language R was made, in Table 4 the variables of each axis with their respective meaning are shown.

Table 4. Definition of the variables (axes or rays of each star).

Variable	Significado
SEL	Selective
GEN	Generative
SUB_T	Subjected_Task
DUR_E	During Exams
CON	Convergent

DIVE	Divergent
AVG	Average HS
PER	Performance

Graph 2 shows 6 stars where each one represents a student and the values obtained in each of the variables: Selective, Generative, Homework Recovery, during exams, convergent, divergent, baccalaureate average and academic performance, to identify the degree to which they belong were assigned a label composed by the initials of the degree and a consecutive number. This graph allows us to observe that students LIA7, LA6, LA10 and LIA1 have a high academic performance, the student LIA5 has very low results in almost all variables the highest value is shown in the variable of information retrieval during exams, the LIA7 student has a good academic performance, an average high school diploma and a high value in selective strategies. Students LA6 and LA10 observe a similar pattern obtaining high values in the averages and learning strategies.

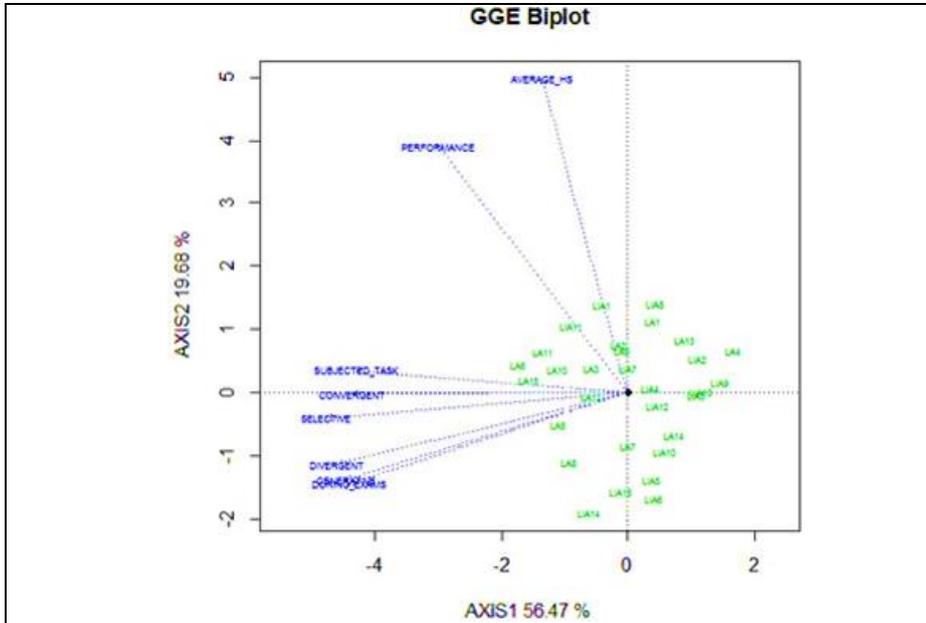


**Fig. 2.** Graph of stars of the learning strategies and the academic performance of the students of the Bachelor's degrees in Administration and Administrative Informatics, Source: (Self-realization, 2017).

### 3.2 Data Analysis with Graphs GGe Biplots

Another visualization technique applied to the data was the GGE Biplot graph for each case of the sample, in relation to the learning strategies, average of bachiller and academic performance. GGE Biplot is effective for the analysis of mega-environments, for example (which genotype (student), -got-where (in which locality (learning strategies), the evaluation of genotypes (from the point of view of stability and performance) and the evaluation of environments (discriminant power), there are several studies that have shown that this graph describes the IGA of populations in an effective way[26], due to this characteristic its use is adequate for the representation of student populations and their interaction with the strategies of learning used with academic performance obtained.

In Figure 3, the genotypes (G), which are the students, are identified where they identify by race and a unique additional number, also identifies genotype-environment interaction (GE), such as academic performance and strategies of learning: selective, generative, task recovery and during exams, convergent and divergent processing. The graph shows the first two main components, derived from the decomposition of the combined effects of G + GE [17].



**Fig. 3.** Graph GGE Biplot of the learning strategies and the academic performance of the students of the degrees of Administration and Informatics Administrative, Source: (Self-realization, 2017).

The first main component explains 56.47% and the second main component 19.68%, obtaining a total 76.15% variability of the two axes and The GGE Biplots graphs are interpreted in terms of distances between elements and orientation of the axes, this allows to know roughly, from the graph, the values of all variables for each genotype [17]. For the description of the results, genotypes with more representative values (extremes), have been considered, and other cases with characteristics close to them can be observed in the graph. The academic performance environment takes high values for the majority of the genotypes and lower for the genotypes LIA14, LIA15, LIA6, LIA5, LIA10, LA14 and LA4 observing that the students of the Degree in Computer science have a lower academic performance. As for the learning strategies, it is observed that the genotypes LIA2, LIA4, LIA9 and LIA13 have the lowest values. The genotypes with low values in both academic performance and strategies are: LIA14, LIA5, LIA15, LIA6, LIA10 and LA14.d being greater than 50% [17], are considered valid interpretations presented.

#### 4 Conclusions

The graphical techniques of multivariate analysis provide numerous advantages over any other type of representation, being more attractive, provoking interest and facilitating the transmission of information in an easy and fast way.

The visualization in the analysis of data allows showing information easy to assimilate and to understand. Also in these representations of information like the graphs of stars and the GGE Biplot one can identify groupings, atypical values and tendencies.

In this work the objective was to use visual techniques of the multivariate analysis to represent the learning strategies used by students, it is proposed the continuation of this work by implementing other visual graphs that describe characteristics of students and are easy to understand for the academic administration or tutors to help them in decision making.

## References

1. Ibañez, M., Cavanagh, M., Bonamico, N., Renzo, M.: Análisis gráfico mediante biplot del comportamiento de híbridos de maíz. *Revista de Investigaciones Agropecuarias*, 35(3):83–93, 86435306 (2006)
2. Martínez-Sánchez, J., Espinosa, N., Villegas, Y.: Interacción genotipo-ambiente en poblaciones de maíz nativo de Chiapas. *Revista Mexicana de Agroecosistemas*, 3(1):38–48 (2016)
3. Ledesma-Ramírez, L., Solís-Moya, E., Suaste-Franco, M.P., Rodríguez-Caracheo, J.F., de la Cruz-Gonzalez, M.L.: Análisis GGE Biplot del rendimiento de trigo (*triticum* spp.) con riego normal y restringido en el bajío. *Agrociencia*, 46(2):119–131 (2012)
4. Lozano-Ramírez, A., Santacruz-Varela, A., San-Vicente-García, F., Crossa, J., Burgueño, J., Molina-Galán, J.: Modelación de la interacción genotipo x ambiente en rendimiento de híbridos de maíz blanco en ambientes múltiples. *Revista Fitotecnia Mexicana*, 38(4):337–347 (2015)
5. Alvarado, J.: Visualización multidimensional, nuevas técnicas de exploración de datos y representación de información. Ediciones Universidad Cooperativa de Colombia (2010)
6. Morales, E.Y., Hernández, J.A., Ochoa, C.A., Lopez, M.A.: Comparison Represented in the Form of Radar of University Student Engagement in Degrees in Technologies. *Research in Computing Science*, 122 (2016)
7. López, M.A., Hernández, J.A., Ochoa, C.A., Morales, E.Y., González, C.: Comparative Study of Learning Strategies of Bachelor Students in Nursing. *Research in Computing Science*, 122 (2016)
8. Quintana, M., Trinidad, J.C., Morales, S.J., Landassuri, V.M.: Análisis Comparativo de Algoritmos de Minería de Datos para Predecir la Deserción Escolar. *Research in Computing Science* 67, Special issue Advances in Computer Science, pp. 51–61(2013)
9. Gudino-Penalosa, F., González-Mendoza, M., Mora-Vargas, J.: Uso de técnicas de agrupamiento en la clasificación de estilos de aprendizaje. *Research in Computing Science*, 95, pp. 135–146 (2015)
10. Cruz, C., Reyes, L. A., Hernandez, G., Sanchez-Cervantes, J.L.: Diseño de una ontología para el proceso de evaluación de las asignaturas técnico científicas del Instituto Tecnológico de Orizaba. *Research in Computing Science*, 128 (2016)
11. Balzarini, M., Bruno, C., Córdoba, M., Teich, I.: Herramientas en el Análisis Estadístico Multivariado. Escuela Virtual Internacional CAVILA. Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba. Córdoba, Argentina (2015)
12. Pérez, M.: Minería de Datos a través de ejemplos. AlfaOmega (2014)

13. Ledesma, R., Molina, J., Forrest, W., Valero-Mora, P.: La visualización múltiple en el análisis de datos: una aplicación en Vista para el análisis de componentes principales. *Psicothema*, 19(3), pp. 497–505 (2007)
14. Sáez, M.: Diseño e implementación de una aplicación en Processing para la representación de datos multidimensionales (2015)
15. Schiattino, I., Silva, C.: Representación gráfica de información multivariante. Aplicación al sistema de salud de Chile (2010)
16. Cabas, C.: Análisis de datos multivariantes (2017)
17. Guisande, C., Vaamonde, A.: Gráficos estadísticos y mapas con R. España. Ediciones Díaz de Santos (2013)
18. Beltrán, J.A.: Estrategias de aprendizaje. *Revista de Educación*. Madrid, 332, pp. 55–73 (2003)
19. Castañeda, S.: Educación, aprendizaje y cognición. Teoría en la práctica (2004)
20. Hernández, R., Fernández, C., Baptista, M.: Metodología de la investigación. McGraw-Hill Interamericana (2010)
21. IBM SPSS Statics: SPSS Statics Base. Recuperado de: <http://www-03.ibm.com/software/products/es/spss-stats-base> (2015)
22. R-Project: R-Project. <https://www.r-project.org/about.html> (2017)
23. Chain, R., Cruz, N., Martínez, M., Jacomé, N.: Examen de selección y probabilidad de éxito escolar en estudios superiores. Estudio en una universidad pública estatal mexicana. *Revista Electrónica de Investigación Educativa*, 5(1) (2003)
24. MacRae, S.: Modelos y métodos para las ciencias del comportamiento (1995)
25. López, M.A.: Predicción del rendimiento académico de estudiantes universitarios aplicado minería de datos (2017)
26. Frutos, M.: Interacción genotipo – ambiente: GGE Biplot y modelos AMMI, <https://gredos.usal.es/jspui/handle/10366/121900> (2011)