DOI: https//doi.org/10.13037/gr.vol38n114.7310

Gestão & Regionalidade

Received: 26/08/2020 | Accepted: 31/08/2021

Economic Development of the State of Ceará: factorial and cluster analysis

Desenvolvimento Econômico do Estado do Ceará: análise fatorial e de cluster

Francisco Laercio Pereira Braga[1i]
Orcid: https://orcid.org/0000-0003-3145-2838

Kilmer Coelho Campos[2ii]
Orcid: http://orcid.org/0000-0001-7752-2542

Abstract

The objective of this project was to verify the economic development dynamics in the municipalities of the state of Ceará based on a set of 14 (fourteen) variables related to the year 2010. The methodological approach used was the factor analysis technique for the definition of factors, and cluster analysis for the formation of homogeneous groups. The empirical evidence pointed to an extraction of four factors that explain approximately 80% of the total variance of the development dynamics model of the municipalities. We observed that seven municipalities of the Metropolitan Region of Fortaleza (MRF) stood out in the economic dimension and that only the municipality of Fortaleza stood out in the human dimension. The results showed that the Northwest and South regions presented lower economic development dynamics than the others. Finally, we found the existence of imbalances in the development dynamics between the regions of Ceará.

Palavras-chave: multivariate analysis; ceará municipalities; economic development.

Resumo

O objetivo deste trabalho foi verificar a dinâmica de desenvolvimento econômico dos municípios do Estado do Ceará a partir de um conjunto de 14 (quatorze) variáveis referentes ao ano de 2010. Utilizou-se a técnica de análise fatorial para a definição dos fatores e análise de *clusters* para a formação dos grupos homogêneos. A evidência empírica apontou uma extração de quatro fatores que explicam aproximadamente 80% da variância total do modelo de dinâmica de desenvolvimento dos municípios. Notou-se que sete municípios da Região Metropolitana de Fortaleza (RMF) se destacaram na dimensão econômica e que somente o município de Fortaleza se destacou na dimensão humana. Os resultados mostraram que as regiões Noroeste e Sul Cearense apresentaram dinâmica de desenvolvimento econômico inferior às demais. Por fim, constatou-se a existência de desequilíbrios na dinâmica de desenvolvimento entre as regiões cearenses.

Keywords: análise de multivariada; municípios cearenses desenvolvimento econômico.

[1] Universidade Estadual do Ceará -Brasil. E-mail: laercio.braga@uece.br
[2] Universidade Federal do Ceará -Brasil. E-mail: kilmer@ufc.br

1 INTRODUÇÃO

The State of Ceará from the 1990s onwards has undergone political and institutional changes that have been reflected in the economic dimension, a moment in which the modernization of Ceará's institutions has been observed. In the social field, it was found that, in the same period, education and health are identified as important elements for growth and economic and social development in Ceará. Education, in particular, has remained on the agenda of public policies in Ceará, with the aim of improving its quality. That has made it present a comfortable situation when compared to other northeastern states, for instance, with a reduction in the illiteracy rate of 26.5% to 18.8% between 2000 and 2010 (SOUZA; TABOSA, 2016; MARIANO; ARRAES; BARBOSA, 2016).

In terms of territorial division, the State of Ceará is divided into seven geographic mesoregions with different economic, social, demographic, and physical characteristics. Concerning physical aspects, we can highlight mesoregions that undergo periods of intense drought, while others stand out due to a greater availability of water. In economic terms, it is noted that the Metropolitan Region of Fortaleza mesoregion accounts for 64% of Ceará's GDP, followed by the Northwestern mesoregion with a share of 10%. The Mid-Southern mesoregion, in turn, had the lowest participation in the State's GDP in 2010 (2.52%). Regarding the Municipal Human Development Index (2010) of the regions, it was 0.62 (on average), with emphasis on the Metropolitan Region mesoregion, which recorded the highest value with 0.67. According to Brasil (2019), the regions with the greatest participation in the state's agricultural units were the Metropolitan Region of Fortaleza, the Northwest and the Jaguaribe. The industrial units are concentrated in the Metropolitan Region (75%) and in the South (9.4%).

So, aiming to reinforce the debates around socioeconomic development strategies in the State of Ceará, which has become a challenge in view of a multiplicity of evidence, the following question arises: What factors can explain the dynamics of economic development in the State of Ceará? The hypothesis raised in this paper is that the economic variables, as opposed to the social variables, solidify as the variable group that mainly explains the development of Ceará's municipalities. This hypothesis is based on the fact that the economic dimension contains variables of the three major economic activities (agriculture, industry and services) that can make it so that municipalities that do not have industrial activity within their productive structure are influenced by trade activities and services.

That said, this project intended to verify the dynamics of economic development of the municipalities of the State of Ceará from a set of fourteen (14) economic and social variables referring to the year 2010. It intended, further, to establish a hierarchy between the municipalities by mesoregion that have greater or lesser development potential and to identify possible development clusters in the State. At the end, the paper will provide elements for decision making in the area of promotion and allocation of regional investments.

This article has four sections in addition to this introduction. The first brings a discussion of the debates on development and economic growth. The following section explains the methodological aspects used, with a presentation of the sources used for the database and the factor analysis and cluster analysis methods. The results are presented in the fourth section along with the hierarchy of Ceará's municipalities. And lastly, in section five, the final considerations are made.

2 THEORETICAL FRAMEWORK

2.1 The nuances of Regional Economic Development

Discussions about development date back a long time, are often complex and have undergone several changes over the years (SANTOS *et al.*, 2017). Adam Smith himself has said that the wealth of a nation would be the result of its productive work, the specialization of the workforce and the division of labor. For Joseph Schumpeter in "The Theory of Economic Development" (SCHUMPETER, 1982), development is not explained by economic growth, but by the process of spontaneous and discontinuous change, when innovation is presented as a consequence of new combinations.

Thus, within this theoretical scope, there is a need to debate the issue of regional development, which has undergone transformations over the decades of the twentieth century. The work arising from the studies of economists and geographers in the 1950s, after the World War II, allowed advances in the understanding of three concepts and strategies for regional development considered of fundamental importance for understanding existing regional inequalities. Therefore, more than fifty years later, regional economic studies encompass several approaches, theories and models to understand the trajectories of regional development (CAPELLO, 2009).

The first concept developed and brought to debate was the one developed by François Perroux in 1955 called "growth hub", which supports the idea that regional economic development takes place via growth and development hubs with strong connections with regions of influence (PERROUX, 1961). These hubs are thus characterized by a high rate of economic development and numerous cooperative connections (SZAJNOWSKA-WYSOCKA, 2009). However, according to Szajnowska-Wysocka (2009), areas strengthened or created as growth hubs end up dominating relatively weaker regions and become competitive at the expense of others. According to Capello (2009), the idea defended by Perroux is that the existence of development poles is concentrated due to the synergy and cumulative forces generated by stable and lasting local relationships, in which the physical space is conceived as diversified and relational.

The second theory refers to the "cumulative circular causation" (CCC) elaborated by Myrdal (1957) and exposes the instability and imbalance of the economic system. The third was the theory developed by Hirschman (1958) who used the theories of the former to detail the effects back and forth. In this line, the author considers that the key industries, or Perroux's driving industries, are stimuli to the potential growth of the product by induction in a given underdeveloped economy.

For Haddad (2009), in this theoretical context, development involves the well-being of society and, for this reason, the author takes into account the variable product per capita as an important element for measuring the productivity of a given economic region. However, Haddad (2009) points out two aspects inherent to the theme: i) first, he adds other relevant aspects for achieving economic development, they are: associative capacity, entrepreneurship and income distribution; ii) Haddad confirms, in this case, what was presented by Pelinski (2007), by stating that development is not something that occurs spontaneously, which therefore requires planning associated with governmental purposes, that is, it makes clear the need for State intervention in the regional economy.

Thus, economic development, especially when it comes to the regional issue, involves the consideration of a set of broad aspects, which in turn involve economic and social information, as they are what cause regional disparities. Economic variables, on the one hand, tend to reflect the level of economic growth

achieved by a region, but alone they are insufficient to say whether a region is developed or not, for this reason it is necessary to include other variables related to the other dimensions, among them, the social variables (EBERTHARDT; LIMA, 2012).

Šabić and Vujadinović (2017) reinforce that economic activity is spatially concentrated, as cities and even some developed economic regions use the effects of agglomerations to attract labor and capital to achieve more favorable economic conditions than other regions. The authors also emphasize that studies on European experiences have contributed to discrepancies between causes and consequences of regional inequalities. For this reason, regional development is a complex process, which consequently requires multidisciplinary and multidimensional approaches. In this aspect, there is an important contribution of the authors when they consider regional development as a process that requires adaptations and specific conditions, in which particular factors of each region must be considered in order to maintain its unique characteristics.

In this aspect, Szajnowska-Wysocka (2009) attests to the importance of the need to continuously observe and record the socioeconomic reality at a regional or local level in order to predict future development, especially in the European scenario of endogenous development.

Therefore, woven this brief theoretical background, we begin to present the empirical works regarding the issue of economic and regional development. For this reason, the authors state that economic development must lead people to obtain an improvement in their income, in addition to improving social conditions.

2.2 Discussion of empirical studies

Authors Rezende, Fernandes and Silva (2007) defined the development potentials regarding the municipalities in the southern region of the state of Minas Gerais (Brazil) in their study. And, to achieve their objective, the aforementioned authors used the statistical method of factor analysis to establish a hierarchy between the municipalities, as well as to identify the greater or lesser potential for development in the industrial, commercial and services aspects, in addition to social and agricultural development.

Silveira, Silva and Carvalho (2008) measured an index that showed the development reflected in the quality of life in the Northern region of Brazil. To achieve its purpose, the authors built a relative quality of life index called IRQV through the use of factor analysis. As main results, the study revealed heterogeneity in the quality of life of the states in the Northern region of Brazil.

In this scope, Lyrio, Barros and Menezes (2019) state that among the important dimensions for development, education is a decisive driving factor for its promotion. Also, according to Lins and Arbix (2011), educational concern plays a central role in increasing productivity, as highly educated people are more aware of their role within the productive structure and positively impact the economies of a region, making the most efficient. Also within the educational scope, Cerqueira *et al*. (2016) state that educational investment is one of the important instruments for controlling homicide rates.

In the study developed by Aves and Rodrigues in 2013, it is possible to verify elements related to sustainable development in the Southern mesoregion of the state of Ceará. To achieve this goal, the authors built the sustainable development index (IDS) for the municipalities and used the multivariate analysis technique. The results obtained showed, firstly, a strong presence of asymmetry in relation to the development trajectories of the analyzed municipalities in the Southern mesoregion of Ceará. The second aspect pointed out concerns the rethinking of the role of the State as an important actor to promote and coordinate

the development process that is aligned with specific local characteristics.

Cerqueira *et al.* (2016), in turn, aimed to map the educational conditions in schools and neighborhoods where the incidence of homicides was observed in 2014 in the state of Rio de Janeiro. It should be noted that this study is important because it works with two very important aspects that are considered to study the development of a region: education and violence. The results show that violent crime does not occur homogeneously in the municipalities, but it is concentrated in a few neighborhoods and locations within the municipalities.

Finally, it should be mentioned that the authors Lyrio, Barros and Menezes (2019) place their contributions to the debate on economic dynamics that economic development should not be focused solely on improving variables such as the growth of the Gross National Product (GNP), but it must take into account the health services. Having access to basic health services is an important element in development, as it allows for, for example, a reduction in infant mortality and an increase in society's life expectancy.

Thus, in the search for synthetic variables that can help to verify the dynamics of development of localities or regions, one is faced with obstacles given the impossibility of selecting elements that really represent the theme of development. For this reason, variables related to economic and social aspects (education and health) were chosen in this paper.

3 METHODOLOGY

3.1 Nature and source of data

The selected variables mostly refer to the year 2010, and many of them were taken from the 2010 Demographic Census. For others, due to lack of information regarding that year, variables that could be a proxy were used. Thus, the data used were taken from the Atlas of Social Development (UNDP/IPEA), Brazilian Institute of Geography and Statistics and Research (IBGE), Institute of Applied Economic Research (IPEA) and Annual Report on Social Information of the Ministry of Economy (RAIS/MTE).

In this way, fourteen (14) economic and social variables were collected from each municipality in the State of Ceará. It should be noted that the level of development of a given region has a multidimensional character, which explains the need to consider an expressive number of variables to characterize it in a comprehensive way. However, it was decided to use fourteen (14) variables in order to make the analysis shorter. Of these fourteen, seven variables were relativized to avoid major distortions in the cluster analysis (Chart 1).

Chart 1 - Variables and data sources for municipalities in the State of Ceará (2010)

Variablel Z (2010)		Source	Reference Studies
Gross Value Added of the Trade and Services sector at current prices (2009) per capita (VAB_CS pc)	Z_1	IBGE	Rezende, Fernandes e Silva (2007).
Gross Value Added of Industry at current prices (2009) per capita (VAB_Ind pc)	Z_2	IBGE	Rezende, Fernandes e Silva (2007).
Industrial units by formal link in the industry (Unid_Ind/Vinc.Ind.)	Z_3	RAIS	Rezende, Fernandes e Silva (2007).
Commerce and services units by formal links in commerce and service (Unid_CS/Vinc.CS)	Z_4	RAIS	Rezende, Fernandes e Silva (2007).
Agricultural units by formal link (Unid_Agrop/Vinc.Agrop.)	Z_5	RAIS	Rezende, Fernandes e Silva (2007).
Municipal Human Development Index (IDHM) (2010)	Z_6	PNUD / IPEA / FJP	Silveira; Silva e Carvalho (2008)
Percentage of young homicides (2012)- (Homic_Jov)	Z_7	IPEA	Cerqueira *et al*. (2016).
Life expectancy (EXP_vida)	Z_8	IPEA	Begnini e Almeida (2016).
Infant mortality (MORT_infantil)	Z_9	PNUD / IPEA / FJP	Rezende, Fernandes e Silva (2007); Alves e Rodrigues (2013); Begnini e Almeida (2016);
% of staff occupied with complete elementary school (OCUP_EFC)	Z_{10}	PNUD / IPEA / FJP	Begnini e Almeida (2016)

Chart 1 - Variables and data sources for municipalities in the State of Ceará (2010)

Variablel Z (2010)		Source	Reference Studies
% of employed persons with complete medium (OCUP_EMC)	Z_{11}	PNUD / IPEA / FJP	Lyrio, Barros e Menezes (2019); Lins e Arbix (2011).
% of employed persons with complete higher education (OCUP_EES)	Z_{12}	PNUD / IPEA / FJP	Lyrio, Barros e Menezes (2019); Lins e Arbix (2011).
Gross Domestic Product (GDP) per capita (GDP pc)	Z_{13}	IPEA	Alves e Rodrigues (2013); Silveira, Silva e Carvalho (2008);
Taxes, net of subsidies, on products, at current prices (2009) per capita (taxes pc)	Z_{14}	IBGE	Rezende, Fernandes e Silva (2007).

Source: Authors' elaboration (2020).

3.2 Factor analysis

Mingoti (2005) reinforces that the factor analysis method (FA) is a multivariate statistical technique used to represent complex relationships between sets of variables, in which the original variables have linear combinations of the common factors obtained and that explain, in turn, the total variance of each variable. According to Blbas, Mahmood and Omer (2017), FA analyzes the interaction of a large amount of information and, subsequently, explains this information in terms of its factors. Thus, the most correlated variables are aggregated within the same factor (dimension) that explain portions of the variations of the original variables (ALVES; RODRIGUES, 2013; REZENDE, FERNANDES; SILVA, 2007).

Using multivariate analysis, the exploratory factor analysis (f) technique and the principal component analysis (PCA) method were applied using the SPSS Statistics 20 software. PCA, in turn, provides as many factors as the number of original variables. However, few factors will be generated that explain, in some way, the total variance of the data (SILVEIRA; SILVA; CARVALHO, 2008).f

According to Hair *et al*. (2006), it is necessary to follow a few steps in order for the factor analysis of the observations and selected variables to be developed, namely: a) formulation of the study problem; b) construction of the correlation matrix; c) determination of the factor analysis method; d) determination of the number of factor rotations; e) calculation of factor loading or choice of substitutive variables; and, finally, f) determination of model adjustments. The factor analysis model can be expressed as follows (MINGOTI, 2005; BAI; HIRA; DESHPANDE, 2015; BLBAS; MAHMOOD; OMER, 2017):

$$z_1 = \alpha_{11}f_1 + \alpha_{12}f_2 + \alpha_{13}f_3 + \cdots + \alpha_{1m}f_m + \varepsilon_1$$
$$z_2 = \alpha_{21}f_1 + \alpha_{22}f_2 + \alpha_{23}f_3 + \cdots + \alpha_{2m}f_m + \varepsilon_2$$
$$z_3 = \alpha_{31}f_1 + \alpha_{32}f_2 + \alpha_{33}f_3 + \cdots + \alpha_{3m}f_m + \varepsilon_3$$
$$\ldots$$
$$z_m = \alpha_{m1}f_1 + \alpha_{m2}f_2 + \alpha_{m3}f_3 + \cdots + \alpha_{mm}f_m + \varepsilon_m \quad \text{(Equation 1).}$$

Equation 1 is a generalized model for the factor analysis method and expresses a linear combination between the observed variables (z_i) and between the m common factors (f). Thus, the basic model of factor analysis explains a correlation structure between the variables $z=z_1+z_2+z_3, \ldots,+z_m$, which are directly observed through a linear combination of variables that are not directly observed, called common factors, plus a residual component.

The factor rotation method is the next step in PCA, as it facilitates the interpretation of factors and thus seeks to minimize the number of variables with a high weight in a factor, as interpretation difficulties may occur due to the emergence of several different factors, which violates the assumption of orthogonality of the factors and makes it difficult to partition the original variables into "*m*" groups (CORRAR; PAULO; DIAS FILHO, 2012;

MINGOTI, 2005; CAMPOS; SILVA; CAMPOS, 2016).

According to Freitas, Paz and Nicola (2007), the orthogonality condition is one of the assumptions of the factor analysis model, in which each specific factor is orthogonal with all common values, and also assumes that they all have zero mean and the respective vectors in L-dimensional space module equal 1. The authors also emphasize that the diagonal of the matrix is called the commonality of the variable, that is, how much each variance is explained by common factors. Among the criteria used to find the orthogonal matrix, there is the Varimax criterion, which will be used in this paper.

Then, in FA, the factor scores for each category of factors obtained after the orthogonal rotation of the original factor structure must be estimated. So, for each sample element "*k*" ($k = 1, 2,..., n$), its score on the F_j factor (common factors) can be calculated according to Equation 2. According to Fávero *et al*. (2009), these factors are estimated from the linear combination of the original variables (z_i) considered in the model, where there is d_{jk} (the coefficients of the factor scores) (BAI; HIRA; DESHPANDE, 2015):

$$F_1 = d_{11}z_1 + d_{12}z_2 + d_{13}z_3 + \cdots + d_{1p}z_k$$
$$F_2 = d_{21}z_1 + d_{22}z_2 + d_{23}z_3 + \cdots + d_{2p}z_k$$
$$F_3 = d_{31}z_1 + d_{32}z_2 + d_{33}z_3 + \cdots + d_{3p}z_k$$
$$\cdots$$
$$F_{jk} = d_{j1}z_{1k} + d_{j2}z_{2k} + d_{j3}z_{3k} + \cdots + d_{jp}z_{pk} \quad \text{(Equation 2).}$$

For the formulation of the FA model, it is necessary to identify the suitability of the available data set (FÁVERO *et al.*, 2009; HAIR *et al.*, 2006). This verification took place through important steps (ALVES; RODRIGUES, 2013): i) number of correlations, for which, based on the significance level of the coefficients expressed in the correlation matrix, a substantial number of correlations greater than 0.30 is recommended; ii) sample adequacy measure using the Sample Suitability Measure (SSM), which varies between zero and one, and the cutoff criterion for excluding variables using SSM values lower than 0.5; iii) Kaiser-Meyer-Olkin (KMO) statistics and Bartlett's sphericity test.

The KMO is known as the sample adequacy index, or EFA's degree of adequacy to the dataset. The KMO result will be between zero and one, and the closer to zero indicates that the factor analysis is inappropriate. On the other hand, the closer to 1, the more appropriate the model will be. The KMO, according to Nugrahadi, Maipita and Situmeang (2020), tests whether the adequacy value is greater than 0.5, in which multicollinearity occurs in the research variables. Fávero *et al.* (2009) place the following critical points: values around 0.90: optimal adequacy of results; values around 0.80: very good adequacy; values around 0.70: good adequacy; values around 0.60: regular adequacy; and values below 0.50 characterize poor adequacy. Bartlett's sphericity test measures to what extent the correlation matrix is similar to an identity matrix. In this case, if the test indicates $p<0.05$, then the dataset is amenable to factorization (HAIR *et al.,* 2006).

3.3 Cluster analysis of homogeneous observations

Cluster Analysis (CA) is a multivariate analysis technique whose objective is to divide the elements of a population or a sample into groups with the same characteristics (CAMPOS; SILVA; CAMPOS, 2016; CORRAR; PAULO; DIAS FILHO, 2012). Thus, we chose to identify similar groups from the grouping of observations (municipalities) that had similarities, considering the values found for the factor scores for each municipality in

the State of Ceará in the formation of clusters. The criterion for grouping the municipalities was based on their proximity, indicated by a square Euclidean distance defined by Fávero *et al*. (2009) and Blbas, Mahmood and Omer (2017):

$$d(z_l, z_k) = \left[\sum_{i=1}^{p}(z_{il} - z_{jk})^2\right]^{1/2} \text{(Equation 3)},$$

on which:

Z_{il} = is the variable value for the observation *i*;
Z_{ik} = is the value of variable k for the observation *j*.

According to Faria *et al*. (2018), this calculation is a usual and recommended measure for the Ward method, which represents an agglomerative hierarchical technique, that is, it is assumed that municipalities are clustered sequentially according to their similarities. For the cluster analysis to occur, it was necessary to use the factor scores that were obtained by the EFA for each municipality and which, in turn, consider the standardization of data (FREITAS; PAZ; NICOLA, 2007; CAMPOS; SILVA; CAMPOS, 2016).

After performing these steps, it was possible to partition the observations with the choice of the clustering algorithm. According to Hair *et al.* (2006), the algorithm maximizes the differences between the clusters relatively within themselves. It is worth noting that the non-hierarchical method was used, in which the number of groups was 5 clusters, defined ex-ante. The choice was made, firstly, for obtaining results less susceptible to atypical observations in the data, the measure of distance used, and the inclusion of irrelevant variables for the case of selection of non-random seeds (HAIR *et al.*, 2006; CAMPOS; SILVA; CAMPOS, 2016).

After simulations were performed, the k-means method was used, which, according to Alves and Rodrigues (2013), allows the allocation of each of the existing objects in one of the pre-defined *k* groups, whose centroid (the clusters' central vector) is the closest to the vector of observed values for the respective object. The proof of significance will occur with the aid of the analysis of variance (ANOVA), which will inform, in the case, which variables allow the separation of these clusters, in such a way that the variables that most discriminate the groups are those that register the greatest variability between the groups and lower internal variability (higher F statistic value).

4 RESULTS AND DISCUSSION

According to Hair *et al*. (2006), the correlation matrix presented a substantial number of variables with values greater than 0.30 and statistically significant. Some of these variables showed high correlation (above 0.30) with almost all others and statistically significant at 5%, they are: industrial added value per capita (z_2), commerce and service units by formal commerce and service links (z_4), HDI (z_6), percentage of employed persons with complete primary education (z_{10}), complete secondary education (z_{11}), Gross Domestic Product per capita (z_{13}) and taxes per capita (z_{14}). The Anti-Image matrix indicates that all variables registered KMO above 0.50, with emphasis on those that reached values above 0.800, which are: z_4 (GDP per capita), z_6 (IDHM), and z_{14} (per capita taxes).

After checking the suitability of the variables, factor analysis was applied to the model, which allowed the extraction of four factors that could explain 78.84% of the total variance of the model (*eigenvalues* greater than unity). To check the consistency of the original data, the Kaiser-Meyer-Olkin (KMO) test presented an index equal to 0.807, a value considered very good and which confirms the indication of the factor analysis (Table 1). In the analysis of the second test, Bartlett's Test of Sphericity (BTS), it was found that

the correlation matrix is unlikely to be an identity, as it was statistically significant, and that the data set used was adequate for the technique used.

Table 1 – Eigenvalue values and percentage of total variance explained by the four factors

Components	Initial own values		
	Total	% of variance	% Cumulative
1	6.889	49.204	49.204
2	1.628	11.628	60.833
3	1.379	9.848	70.681
4	1.142	8.16	78.84
Kaiser-Meyer-Olkin measure of sampling adequacy *(KMO)*.			0.807
Bartlett's Test of Sphericity (BTS)	*Chi-square*		3339.401
df			91
Sig.			0

Source: Research results (2020).

Table 2 shows the factor loadings and commonalities for the considered variables. It is noted, first, that the value of commonality was below 0.50 only for the variable "Units of commerce and services by formal ties of commerce and services", while all the others expressed that more than half of the variance of the variable is reproduced by the common factors.

Factor 1 represents 49.2% of the total variance of the variables. By observing the variables, it is confirmed that the dynamics of the development of Ceará's municipalities involves variables of economic dimension, with high factor loadings and close to the unit, they are: Industrial value-added per capita, GDP per capita, taxes per capita, and value-added of trade and services. The fifth variable is the percentage of youth homicides, which was positively correlated with the dimension. For this aspect, violence has a strong relationship with education, but also with urban factors, that is, intense urbanization processes and the growth of cities. However, it should be noted that the sixth variable, Z_4, was weakly associated with the factor after rotation and has a negative sign, that is, an inverse relationship with factor 1. It should be noted that among these variables, GDP per capita is an important element to raise productivity and, consequently, impact the economic growth of the regions (HADDAD, 2009; EBERTHARDT; LIMA, 2012).

In turn, 11.62% of the total variability of the data is explained by Factor 2, which involves human aspects. Within this dimension, as theoretically expected, the nature of the variables has a positive relationship with the factor, they are: percentage with completed elementary, secondary and higher education, in addition to the HDI, which has a high correlation with the variables. These variables have the potential to influence the possibility of the trajectory of economic development in the regions, and for this reason, these aspects are prioritized by local and state public policies.

Factor 3 represents 9.84% of the total variability of the data, being composed of two variables related to the health dimension, which shows the importance of health policies in the development process of Ceará's population. For this factor, the two selected variables showed signs consistent with the theory, that is, the greater the development, the lower the infant mortality and the greater the life expectancy. Finally, with regards to Factor 4, which explains 8.16% of the data variability, the nature of the variables is related to the size of the industrial and agricultural production units, whose positive factor loadings are above 0.70.

Table 2 - Factor loadings after rotation and commonalities (2010)

Variables	Components				Commonalities
	1	2	3	4	
VAB_CS pc	.787	.406	.166	-.042	**.813**
VAB_Ind pc	.941	.063	.097	-.088	**.908**
GDP pc	.945	.176	.124	-.071	**.945**
Taxes pc	.937	.134	.134	-.075	**.919**
HOMIC_Jov	.716	.286	.302	-.099	**.695**
Unid_CS/Vinc.CS.	-.419	-.332	-.196	-.168	**.353**
IDHM	.472	.716	.279	-.188	**.848**
OCUP_EFC	.473	.778	.135	-.156	**.872**
OCUP_EMC	.395	.848	.126	-.098	**.901**
OCUP_ESC	-.061	.775	.113	.086	**.624**
EXP_vida	.228	.220	.935	-.062	**.978**
MORT_infantil	-.208	-.180	-.946	.062	**.974**
Unid_Ind/Vinc.Ind.	-.123	.065	-.230	.703	**.566**
Unid_agrop/vinc.Agro	-.010	-.136	.103	.783	**.642**

Source: Research results. Authors' elaboration (2020).

From what could be observed, in an overview of the municipalities of Ceará, concerning the economic dimension and its important variables, only 35.32% of the municipalities obtained values above the sample mean and the vast majority (119 municipalities) presented values below the average. For factor 2, human dimension, 91 municipalities (49.45%) are above the average, while 93 municipalities still present values below the average. Regarding factor 3, in general, 52.17% of the municipalities in Ceará recorded values above the average, compared to 88 municipalities below the average for the health dimension. And, for factor 4, 30.97% of the municipalities had values above the average, while 127 municipalities had values below the average in the dimension of productive units.

From Table 3, the scores of the five municipalities in Ceará are found in order of best and worst condition regarding the dimensions, which were also classified based on the mesoregions of the State. Thus, it is clear that among the five cities that are best placed in the economic dimension, all belong to the Metropolitan Region of Fortaleza (RMF) and are in Mesoregion 3. This result is explained by the spatial concentration of industrial production and service companies existing in the RMF, which represented 72% and 65% of industrial units and commerce and services existing in Ceará, respectively (BRASIL, 2019). This also reveals the existence of intra-regional inequalities, as at the same time that a large part of the municipalities was economically deficient, there are two macro-planning regions in the State of Ceará that are economically important, namely: Cariri/Mid-South (belonging to the Southern mesoregion) and Sobral/Ibiapaba (comprising the Northwestern mesoregion).

Regarding factor 2, it is noted that the city of Fortaleza occupies a prominent position in the ranking, along with the presence of three municipalities belonging to mesoregion 7 that presented a better situation for the human dimension. In this group, it is worth highlighting the investments that took place in the municipality of Sobral, with the implementation of city's literacy policy model between 1997 and 2000, which ended up serving as a reference for other municipalities in Ceará (LYRIO; BARROS; MENEZES, 2019). Overall, mesoregion 7 (South) is in the most comfortable situation in relation to factor 2, as only 40% of its municipalities are below average, which reveals a positive result of educational policies in the region.

Regarding the five worst placed, it is noted that two of them make up mesoregion 1, which reveals less potential for human

development. Such behavior highlights the need for more effective educational policies in the regions to increase the number of employed persons, especially with primary and secondary education. At this point, the results corroborate, in a way, with the conclusions obtained by the study by Santos *et al*. (2017), in which it is noted that the education variables are not spilling over into the development dynamics of the municipalities in question. This fact is worrisome, as the literature points to the human dimension as an important aspect for increasing productivity and, consequently, increasing the rate of economic growth (LINS; ARBIX, 2011).

In this aspect, the importance of the educational variable in Ceará is highlighted after the policies aimed at improving literacy indicators, especially those implemented in the 1980s and 1990s. Lyrio, Barros and Menezes (2019) reinforce the results obtained by placing the educational variable as promoting the growth and development of a given geographic region. For Lins and Arbix (2011), the educational concern should be a central focus of public policies to increase the economy's productivity. In this way, educational variables are important to leverage economic growth and reduce youth homicides in different municipalities (CERQUEIRA *et al*., 2016).

Table 3 - Ranking of the municipalities of Ceará according to their dimensions (2010)

Counties	Factor 1	Rank	Mesoregion	Counties	Factor 1	Rank	Mesoregion
The five municipalities above the average- Factor 1 (49,2%)				**The five municipalities above the average- Factor 3 (9,84%)**			
Eusébio	8.307	1	3	Sobral	2.268	1	1
Horizonte	5.861	2	3	Barbalha	2.142	2	7
Maracanaú	5.173	3	3	Crato	1.922	3	7
Fortaleza	2.113	4	3	Iguatu	1.816	4	6
Aquiraz	1.798	5	3	Caucaia	1.700	5	3
The five municipalities below the average- Factor 1				**The five municipalities below the average- Factor 3**			
Altaneira	-0.781	175	7	Groaíras	-1.959	175	1
Graça	-0.798	176	1	Barroquinha	-2.025	176	1
Lavras da Mangabeira	-0.813	177	6	Palhano	-2.048	177	5
Ererê	-0.816	178	5	Saboeiro	-2.064	178	4
Jati	-0.843	179	7	Jijoca de Jericoacoara	-2.122	179	1
The five municipalities above the average- Factor 2 (11,62%)				**The five municipalities above the average- Factor 4 (8,16%)**			
Fortaleza	4.902	1	3	Meruoca	5.562	1	1
Crato	3.716	2	7	Itatira	3.388	2	2
Juazeiro do Norte	2.499	3	7	Caririaçu	2.879	3	7
Penaforte	2.371	4	7	Morrinhos	2.698	4	1
Sobral	2.283	5	1	Reriutaba	2.579	5	1
The five municipalities below the average- Factor 2				**The five municipalities below the average- Factor 4**			
Ibaretama	-1.458	175	4	Antonina do Norte	-0.993	175	6
Alto Santo	-1.500	176	5	Granjeiro	-1.000	176	7
Tarrafas	-1.513	177	6	Groaíras	-1.018	177	1
Viçosa do Ceará	-1.523	178	1	Salitre	-1.060	178	7
Croatá	-1.639	179	1	Potiretama	-1.081	179	5

Source: Research results (2020).

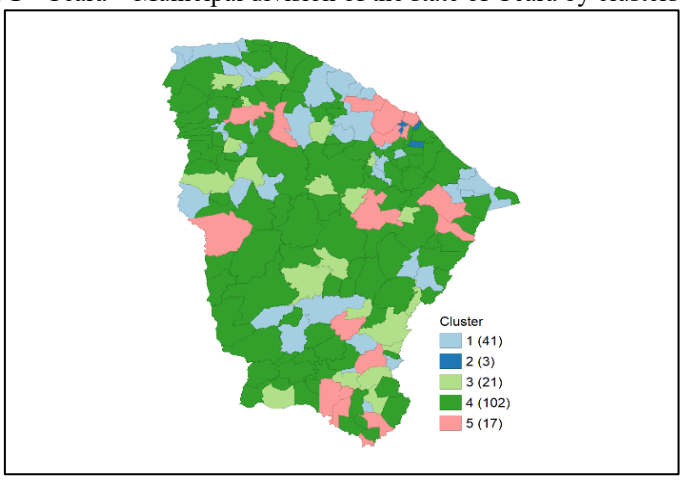
Regarding factor 3 (health dimension), there was the presence of two municipalities belonging to mesoregion 7 (South) among the best positioned and one (1) municipality in mesoregion 1. The two best placed were Sobral and Barbalha, with 2.26 and 2.14, that is, above the sample average, respectively. With regard to the five worst identified, three municipalities are in the Northwestern mesoregion (mesoregion 1). In this view, in relation to life expectancy and infant mortality, there is a need for greater attention to the health sector in towns in Ceará below average.

Lyrio, Barros and Menezes (2019) already reinforced that development must take into account the supply of health services, as the more access to these services, the greater the probability of having a reduction in infant mortality and raising the life expectancy of society.

In order to classify the municipalities into clusters, that is, group the factor scores according to the four dimensions (economic, human, health and productive units), the results of the analysis allowed for the distribution of the municipalities in Ceará into five distinct clusters (Figure 1). The largest amount was concentrated in cluster 4 (102 municipalities), followed by cluster 1, with forty-one (41) municipalities, cluster 3, with twenty-one (21) municipalities, and cluster 5, formed by seventeen (17) municipalities of Ceará. Finally, there is cluster 2, which grouped only three municipalities with similar characteristics. It should be noted that five clusters were chosen to obtain a better representation of the municipalities by grouping. It was found that the significance of the variables that most discriminate the groups and that register the greatest variability between groups and the lowest internal variability were all significant at 0% (highest value of the F statistic).

In view of the grouping results, cluster 4 shows that most of Ceará's municipalities have similar development dynamics. In this scope, it can be said that the municipalities that compose it had worse development conditions, because, when considering the economic dimension, human dimension and dimension of productive units by formal links, 75, 66 and 75 municipalities in Ceará presented values below the average of the sample, respectively. Only in the health dimension, factor 3, are 66.66% municipalities in this cluster above the sample mean. It is also worth mentioning that 68% of the municipalities that constitute group 4 belong to the Northwestern and Northern mesoregions, and have similar characteristics mainly in relation to the factors of industrial and agricultural production units (factor 4), economic dimension (factor 1) and human dimension (factor 2).

Figure 1 - Ceará – Municipal division of the state of Ceará by clusters (2010)



Source: Research results (2020).

In cluster 2, it is noted that, for the economic dimension, three municipalities are above the average, they are: Eusébio, Horizonte and Maracanaú. For the human dimension, the first two are below average, while for the health dimension, all three had negative values. With regard to the dimension of industrial and agricultural production units due to formal links created in these segments, only Eusébio is above the sample average.

In relation to cluster 3, in the economic dimension, only six (6) are above the average, which reveals a weak economic dynamic. Regarding the human dimension, fourteen municipalities in the cluster had results below average. For the health dimension, eleven (11) municipalities recorded values above the average and nine (9) below. Finally, in the dimension of productive units, all recorded values above the average, which revealed, for this factor, a positive dynamic.

Finally, in cluster 1, formed by forty-one (41) municipalities in Ceará, 59% of them had negative values below the average for the economic dimension, 100% below the average for the health dimension and 70% below the average for the dimension production units. Only for the human dimension was a different behavior noticed, as thirty-four municipalities (83%) obtained positive values.

It is noted, therefore, that the regional imbalances existing within the state of Ceará express the way that economic and industrial concentration occurred to the detriment of other regions of the state, reinforcing what Perroux (1961) affirmed about the formation of "islands" of economic growth surrounded by empty or stagnant economic spaces. It should be noted that public policies for the industrialization of Ceará boosted the formation of "islands", as well as the polarization and formation of gaps between inland municipalities and the Greater Metropolitan Region of Fortaleza. The policies adopted in Ceará prove the importance of their elaboration and implementation for the economic development of the municipalities, and this fact reveals the active role assumed by the State to generate and conduct institutional and legal conditions to achieve economic growth and development (PELINSKI, 2007; HADDAD, 2009).

These aspects corroborate the results of Alves and Rodrigues (2013), in which they affirm the strong presence of trajectory asymmetry in the municipalities of Ceará. Therefore, most of Ceará's municipalities need to strengthen institutional and public policy support to leverage their development, in which knowledge of the local characteristics of each municipality will allow the elaboration and implementation of public policies (SZAJNOWSKA-WYSOCKA, 2009; EBERTHARDT; LIMA, 2012; ŠABIĆ; VUJADINOVIĆ, 2017).

5 FINAL CONSIDERATIONS

This study sought to identify the dynamics of economic development in Ceará's municipalities in 2010 and establish a hierarchy based on factorial and cluster analysis, highlighting which municipalities in the State of Ceará have development potential. Thus, four common or specific factors related to developmental dynamics were obtained. These factors were called economic dimension, consisting of six initial variables; human dimension, formed by four explanatory variables; health dimension and dimension of industrial and agricultural production units, with two variables respectively.

It was found that the first factor 1 represented 49.2% of the variance of the variables and with a positive correlation for five of them, including the variable percentage of homicides, which contradicted what was theoretically expected, as a negative relationship was expected. Factor 2 accounts for 11.62% of the total variability of the data and encompasses the human aspect variables, and whose variables kept a positive relationship with the factor, proving the theory. Factor 3, in turn, was called the health dimension and accounted for 9.84% of the variance of the variables, consisting of the life expectancy and infant mortality variables, both with signs consistent with the theory. Finally, factor 4 (production units dimension) explained 8.16% of the total variability of the data, composed of industrial and agricultural units by formal employment.

The results of the ranking of the four factors made it evident that the dynamics of the development of Ceará's municipalities is somewhat unbalanced, especially when observing the behavior of municipalities in the interior of the State in relation to municipalities in the Metropolitan Region of Fortaleza. The study also identified the need for more effective care in the areas of education and health for the Mid-Southern, Northwestern and Northern mesoregions of Ceará, and stimuli for the productive units of the Mid-Southern and Northwestern mesoregions. In this way, compensatory measures can favor the less favored mesoregions of the State and reduce the existing regional and intraregional disparities.

These results certainly reveal important challenges to be overcome by the State and Federal Governments, if the objective is economic growth and development in the mesoregions of the State. First, there must be a concern with the still existing economic concentration of the Metropolitan Region of Fortaleza and, second, to alleviate intra-regional inequalities in the State of Ceará.

From the cluster analysis, the vast majority of municipalities (102) in Ceará were concentrated, according to their similarities, in cluster 4, while in cluster 2 there are only three (3) municipalities, in this case, Eusébio, Horizonte and Fortaleza. In this way, the use of municipal ranking helps to identify the potential for development for Ceará and to identify those regions that most need attention from the public authorities.

However, there is a need, firstly, for new studies that address a greater number of economic, social (educational and health) and demographic variables in the municipalities of Ceará, and secondly, to carry out a comparative analysis between other periods of time to verify the changes witnessed at the state level. Finally, at a later date, it is possible to elaborate an index based on this methodology that represents the dynamics of development for the municipalities of the State of Ceará and that allows for the comparison in regional terms.

REFERENCES

ALVES, C.L.B.; RODRIGUES, A.S. Elementos para pensar territorialidades e desenvolvimento sustentável na mesorregião sul cearense. **Revista Econômica do Nordeste,** Fortaleza, v.44, n. Especial, p. 251-274, 2013.

BAI, A.; HIRA, S.; DESHPANDE, P.S. An Application of Factor Analysis in the Evaluation of Country Economic Rank. **Procedia Computer Science**, v. 54, p. 311 – 317, 2015.

BEGNINI, S.; ALMEIDA, L.E.D.F. de. Grau de desenvolvimento dos municípios da mesorregião oeste catarinense: caracterização e classificação. **Revista Interações**, Campo Grande, v.17, n. 4, p. 547-560, 2016.

BLBAS, H.T.A.; MAHMOOD, S.H.; OMER, C.A. A Comparison results of fator analysis and cluster analysis to the migration of Young people from the Kurdistan Region to Europe. **ZANCO Journal of Pure and Applied Sciences,** v.29, n.4, p. 44-55, 2017**.**

BRASIL. Ministério da Economia. **Relatório Anual de Informações Sociais (RAIS)**. **2019**. Disponível em: http://www.rais.gov.br/sitio/sobre.jsf. Acesso em: 15 de ago. de 2019.

CAMPOS, K.C.; SILVA, F.D.V. da. CAMPOS, R.T. Perfil técnico e econômico da fruticultura irrigada na microrregião do Cariri, Ceará. **Revista Brasileira de Estudos Regionais e Urbanos**, v. 10, n. 1, p. 21- 43, 2016.

CAPELLO, R. Regional Growth and Local Development Theories: Conceptual Evolution over Fifty Years of Regional

Science. **Géographie, économie, Société**, v. 11, p. 9-21, 2009.

CERQUEIRA, D.; RANIERE, M.; GUEDES, E.; COSTA, J.S.; BATISTA, F. NICOLATO, P. Indicadores multidimensionais de educação e homicídios nos territórios focalizados pelo Pacto Nacional pela Redução de Homicídios. **Nota Técnica nº 18**. *IPEA*, Brasília, 2016.

CORRAR, L.J.; PAULO, E.; DIAS FILHO, J.M. **Análise multivariada para os cursos de Administração, Ciências Contábeis e Economia.** Fundação Instituto de Pesquisa Contábeis, Atuariais e Financeiras. São Paulo: Atlas, 2012.

EBERTHARDT, P.H. de; LIMA, J.F. de. Evolução e estágio do desenvolvimento econômico regional: o caso das regiões do Paraná. **Revista Desenvolvimento Regional em Debate**, ano 2, n. 1, 2012.

FÁVERO, L. P.; BELFIORE, P.; SILVA, F. L. da; CHAN, B. L. **Análise de dados: modelagem multivariada para tomada de decisões**. São Paulo: Campus, 2009.

FREITAS, C.A. de; PAZ, M.V.; NICOLA, D.S. Analisando a modernização da agropecuária gaúcha: uma aplicação de análise fatorial e *cluster*. **Revista Análise Econômica**, Porto Alegre, Ano 25, n. 47, p. 121-149, 2007.

HADDAD, P. Capitais intangíveis e desenvolvimento regional. **Revista de Economia. Curitiba**, v. 35, n. 3, p. 119-146, 2009.

HAIR JR., J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. **Multivariate Data Analysis**. 6th Ed. Pearson Prentice Hall, 2006**.**

HIRSCHMAN, Albert. O. **The Strategy of Economic Development**. Editora Yale University Press, 230p., 1958.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Censo Demográfico 2010**. Disponível em: http://www.censo2010.ibge.gov.br/sinopse / index.php?dados=8. Acesso em: 6 ago. 2019.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEA)/PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO (PNUD)/ FUNDAÇÃO JOÃO PINHEIRO (FJP). **Índice de Desenvolvimento Humano dos Municípios Brasileiros – IDH,** 2013. Disponível em: http://www.atlasbrasil.org.br/2013/. Acesso em: 20 ago. 2019.

LINS, L.M.; ARBIX, G. Educação, qualificação, produtividade e crescimento econômico: a harmonia colocada em questão. **Anais...I Circuito de debates acadêmicos**. IPEA, Code 2011. Disponível em: http://www.ipea.gov.br/code2011/chamada 2011/pdf/area3/area3-artigo5.pdf. Acesso em: 15 de ago. de 2019.

LYRIO, L.B.; BARROS, M.J.F.; MENEZES, A.M.F. Educação, um fator decisivo para o desenvolvimento: o modelo do município de Sobral/CE. *In*: ENANPUR, 18., 2019, Natal. **Anais** [...]. Natal, maio/2019.

MARIANO, F.Z.; ARRAES, R.A.; BARBOSA, R.B. Análise longitudinal para avaliação do ensino profissionalizante. *In*: ATALIBA, F.; BARRETO, F.D.; SOUSA, F.J. (Organizadores). **Economia do Ceará em Debate 2016**. 280p. IPECE, Fortaleza, 2016. Disponível em: http://www2.ipece.ce.gov.br/encontro/201 6/index.htm. Acesso em: 6 set. 2019.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada**: **uma abordagem Aplicada**.

Belo Horizonte: Editora: UFMG, 2005. 295p.

MYRDAL, G. **Economic Theory and Under-Developed Regions.** First edition. London: G. Duckworth, 1957.

NUGRAHADI, E. W.; MAIPITA, I; SITUMEANG, C. Dominant Socio-Economic Indicators on the Growth of Small-Scale Industrial Sector: Empirical Evidence with Principal Component Analysis. **Revista Espacios**, v. 41, n.1, p. 1-11, 2020.

PELINSKI, A. **Padrão de desenvolvimento econômico dos municípios do Paraná**: **disparidade, dispersão, e fatores exógenos**. Dissertação (Mestrado em Desenvolvimento Regional e Agronegócio) – Programa de Pós-Graduação em Desenvolvimento Regional e Agronegócio - Universidade Estadual do Oeste do Paraná (Unioeste), Toledo, 2007.

PERROUX, F. **L´economie du XXo siècle**. Paris: Presses Universitaires de France, 1961.

REZENDE, M.L.; FERNANDES, L.P.S.; SILVA, A.M.R. Utilização da análise fatorial para determinar o potencial de crescimento econômico em uma região do sudeste do Brasil. **Revista Economia e Desenvolvimento**. n 19, p. 92- 109, 2007.

ŠABIĆ, D; VUJADINOVIĆ, S. Regional Development and Regional Policy. **Geografski fakultet Univerziteta u Beogradu**, v.65, n. (1a), p. 463- 477, 2017.

SANTOS, N.B.; FARIA, A.M.M.; DALLEMOLE, D.; MANSO, J.R.P. Desenvolvimento e crescimento econômico das macrorregiões de Mato Grosso nos anos 2005 e 2013**. Revista Interações,** Campo Grande, v. 18, n. 3, p. 169-182, 2017.

SCHUMPETER, J. A. **Teoria do desenvolvimento econômico**: **uma investigação sobre lucros, capital, crédito, juro e o ciclo econômico**. São Paulo: Abril Cultural, 1982.

SILVEIRA, B.C.; SILVA, R.G. da.; CARVALHO, L.A. Índice relativo de qualidade de vida da região norte: uma aplicação da análise fatorial. **Revista Brasileira de Gestão e Desenvolvimento Regional**, Taubaté, v. 4, n. 4, p. 80-97, 2008.

SOUZA, H. G. de; TABOSA, F.J.S. Análise espacial do desempenho escolar da educação básica dos municípios do estado do Ceará. In: ATALIBA, F.; BARRETO, F.D.; SOUSA, F.J. (Organizadores). **Economia do Ceará em Debate 2016**. 280p. IPECE, Fortaleza, 2016. Disponível em: http://www2.ipece.ce.gov.br/encontro/2016/index.htm. Acesso em: 6 set. 2019.

SZAJNOWSKA-WYSOCKA, A. Theories of Regional and Local Development. **Abridged Review Bulletin of Geography APHY Socio–economic Series**, n. 12, p.75-90, 2009.

[i] Doutorando pelo Programa de Pós-Graduação em Economia Rural (PPGER), da Universidade Federal do Ceará.

[ii] Doutor em Economia Aplicada pela Universidade Federal de Viçosa.