IOSR Journal of Business and Management (IOSR-JBM)
e-ISSN:2278-487X, p-ISSN: 2319-7668. Volume 26, Issue 3. Ser. 2 (March. 2024), PP 56-71
www.iosrjournals.org

The Latent Characteristics of Family Farming in Pará: A Principal Component Factor Analysis

Hermógenes José Sá De Oliveira[1]*; Anderson Rodolfo De Lima [2];
Tiago Santos Barreto Thomaz[3]

[1] *Peabiru Institute. Phd Candidate in The Postgraduate Program in Economics at The Federal University of Pará.*
[2] *Federal Institute of Mato Grosso. Professor of Administration in basic technical and technological education.*
[3]*Federal Institute of Education, Science and Technology of Roraima. Teacher Of Basic, Technical and Technological Education in The Area of Administration. Master In Education Sciences/Educational Administration.*

Abstract:
The research aimed to identify latent phenomena in family agriculture in the state of Pará. The multivariate technique of principal component analysis was used, based on 16 indicators derived from variables from the 2017 Agricultural Census. The main results of the research identified three factors that, together, explain about 67% of the total data variation. Factor 1 is related to a more commercial approach to family agriculture, highlighting variables such as income, expenses, and production value. In terms of geographical distribution, although this factor is spread throughout Pará, its presence is most pronounced in the municipalities of the southern and southeastern agricultural frontiers of the state. Factor 2 is associated with soil management practices, correction, fertilization, and irrigation. Spatially, municipalities with high scores in this factor are concentrated in the oldest colonization regions of the Amazon, especially in the Metropolitan Region of Belém, Castanhal microregion, and Northeast Pará, in the Bragantina and Salgado microregions. Factor 3 is an extension of the first and is observed in municipalities with higher proportions of agricultural areas destined for family agriculture. In these locations, family farming establishments predominantly direct their production towards commercialization, with agricultural activities being the main source of income. The geographical distribution of this factor forms corridors in regions such as Marajó, Northeast Pará, and municipalities influenced by military dictatorship colonization projects, notably those related to the Transamazon Highway (BR-230).
Keywords: *Principal Component Analysis (PCA); latent phenomena, family agriculture in the State of Pará.*

Date of Submission: 27-02-2024 Date of Acceptance: 07-03-2024

I. Introduction

According to the 2017 Agricultural Census (IBGE, 2019), out of over 5 million agricultural establishments (AE) identified throughout Brazil, 77% are classified as family farming (FF), with its highest proportion found in the Northern region, where 83% of AE are classified as FF, and the lowest in the Midwest region, with 64%. In Pará, the state analyzed in this research, family farming accounts for 85% of the 281,699 agricultural establishments surveyed, which equals 239,737 units.

The state of Pará, located in the northern region of Brazil, has a population of 8.811.659[1] from people distributed across 144 municipalities within an area of 1,248,042 square kilometers, rendering it the second largest state by territorial extension, encompassing 16.66% of the national territory, and occupying 26% of the Brazilian Amazon, the largest tropical rainforest in the world (FAPESPA, 2022a).

According to FAPESPA (2022b), in 2020, the GDP of Pará was approximately R$ 215 billion, which corresponds to 2.8% of the national economy, totaling R$ 7.6 trillion, and ranks the Pará economy in the 10th position among the federal units. However, concerning the Northern region, Pará is the main economy, representing 45.2% of the region's GDP.

The social group representing family agriculture is characterized by work carried out with labor relations based on kinship and production destined for both family consumption and commercialization. It is composed of small landowners living in small communities or rural settlements (Schneider & Cassol, 2017). Other characteristics include relative autonomy compared to the surrounding society and non-business orientations and rationalities of economic orientation (Grisa & Sabourin, 2019).

In terms of political-institutional aspects, Law No. 11,326 of July 2006 establishes that family agriculture, unlike non-family agriculture, has limitations on productive area of four modules, must manage

[1] IBGE (2021) - Population Estimates

property and production on family bases, must have its main income related to agricultural activities, and use the family nucleus as the main source of labor (Cruz et al., 2020).

However, the current recognition of family agriculture institutionally and politically has been a long social construction. Only in terms of public policies directed towards family agriculture, there have been over twenty years in Brazil, with a representative milestone being the creation of PRONAF - National Program for Strengthening Family Agriculture (Sabourin, 2017).

The implementation of policies to stimulate and support family agriculture, such as in Pronaf, was justified by its socioeconomic importance as a fundamental activity for the production of basic and diverse foods for the national market and for exports, as well as for maintaining people in rural areas with generation of work and income (Cruz et al., 2020).

In this context, the characterization, in terms of meaning and characteristics, of the social group called family agriculture has advanced considerably in recent years, especially in the recognition of its economic diversity and social heterogeneity (Schneider & Cassol, 2017). In the past, family farmers were classified, among other definitions, as settlers, peasants, small producers, small landowners, and small farmers (Cruz et al., 2020).

In parallel with the evolution of the definition of family agriculture, the concept of rural development is still under construction (Stege & Parré, 2013); however, there is some assurance affirming that this process is not limited to the economic growth of the socio-spatial framework of reference but, analytically speaking, it involves a dynamic transformation of reality that has gradually incorporated multiple dimensions: economic, social, political-institutional, and environmental (Schneider & Tartaruga, 2004; Schneider, 2010; Schneider & Escher, 2011).

These multiple dimensions reinforce the territorial approach as a way to explain the relevance of the socio-spatial context as an important factor for the development process since the proximity of actors and the density of social relations potentiate collective actions of cooperation (Schneider & Tartaruga, 2004).

In this sense, the analysis of rural development involves considering it as a complex phenomenon with multiple dimensions involving processes of social, technological, distributive, and economic transformations (Melo & Parré, 2007).

This complexity is not alien to the Brazilian rural context; on the contrary, in recent decades, rural Brazil has undergone a significant process of "heterogenization," with a slow reduction of essential asymmetries in relation to what would traditionally be urban Brazil (Favareto, 2010). However, there is an indisputable aspect in the Brazilian rural scenario: the relevant role of family-based agriculture (Schneider, 2010).

The concept of development, being difficult to define, is more easily achieved through simplifications. In this sense, to find approximate measures of the degree of development, schemes of decomposition of its most relevant aspects such as economic, and social, among others, are usually used (Kageyama, 2004).

With this objective, Rebello and colleagues (2011) and Lobão and Staduto (2020) seek to analyze and measure the process of agricultural modernization in the Northeast of Pará, the former, and in the Amazon as a whole, the latter. Both works use factorial analysis to find latent variables in the dynamics of agricultural transformation.

Rebello et al. (2011) worked with four factors that capture the joint variance of 17 variables. The "Intensity of capital use" factor is related to the use of tractors, investments, and total expenses. The "Use of modern technology" factor indicates practices of fertilization, use of agrochemicals, irrigation, and access to electricity. The "Productivity of factors - land, capital, and labor" factor strongly and positively correlates with the value of financing and production. The "Coverage of technical assistance and financing" factor correlates positively and strongly with indicators of access to technical assistance services and rural extension (ATER) and financing.

Lobão and Staduto (2020) capture 5 factors that highlight the importance of agricultural expenses, land use, and intensive labor. The analysis demonstrated heterogeneity in the process of agricultural modernization between the western and eastern sides of the Amazon, with municipalities on the latter side, where the agricultural frontier expands, showing better results.

Thus, recognizing that the analysis of development is not a simple task and that family-based agriculture is a relevant factor in this context, the present study aims to identify and describe factors that characterize family agriculture in the municipalities of Pará, based on the results of the 2017 Agricultural Census and using the multivariate technique of principal component analysis (PCA) together with a geospatial analysis of the results.

The text is divided into four sections. In addition to this brief introduction, the data collection process and the construction of factors or principal components are presented in the Materials and Methods section. Then, the Results section presents in detail and analytically the main findings of the research. A summary of the work is presented in the final remarks.

II. Materials And Methods

In its general purposes, this research was structured along two axes, each with distinct delineations. An exploratory part, which, through bibliographic research on family farming and rural development, sought greater familiarity with the relevant factors and determinants for the studied phenomenon. And a descriptive part that, utilizing techniques for collecting secondary data and multivariate data analysis, sought to capture factors describing the dynamics of family farming in the State of Pará, based on the 2017 Agricultural Census.

To establish the initial theoretical foundation, bibliographic research was conducted on rural development and family farming to deepen the knowledge about these themes. This stage aimed, as an expected outcome, to support the interpretation of the relationship between variables and any "hidden" phenomena in the factors captured by Principal Component Analysis (PCA).

Very useful when there is relatively high correlation between the researched variables, PCA, as one of the multivariate factor analysis techniques, seeks to analyze the joint behavior of interdependent original variables by discovering a smaller quantity of resulting variables, called factors, which function as a kind of grouping of the correlations of the original variables (Fávero and Belfiore, 2017).

The multivariate data analysis for factor extraction sought, in an exploratory manner, to prospect and capture latent variables (factors) indicating non-explicit territorial phenomena of family farming reality that, nevertheless, become evident through joint observation, at the municipal level, of the various variables available in the 2017 Agricultural Census.

Factor analysis, as a multivariate technique, seeks to replace the difficult task of analyzing a large number of variables by identifying a smaller number of factors that capture the behavior of these original variables jointly (Fávero and Belfiore, 2017). This generic term is an "umbrella" for a set of techniques and procedures that enable data reduction and summarization.

These data summarizing the relationships between the data are called factors, which are nothing more than new extracted variables representing underlying dimensions that explain the correlations among an independent set of original variables (Malhotra, 2019; Fávero and Belfiore, 2017).

The proposed analysis began with the collection of secondary data from the 2017 Agricultural Census, via the Brazilian Institute of Geography and Statistics (IBGE) Automatic Recovery System (SIDRA), focusing on the municipality as the unit of analysis. The same variables were surveyed for each of the 144 municipalities in Pará, from the following tables: Characteristics of Agricultural Establishments, Characteristics of Producers, Financial Movement, Employed Personnel, and Brazilian Confederation of Agriculture and Livestock (CNA).

To enable a socio-productive comparison of the municipalities among themselves, it was decided to use relative values rather than the absolute values of municipal variables. In this sense, the first part of data processing was to generate new variables with percentage values, which basically involved dividing the absolute value of each variable by the total value of each municipality, whether related to the universe of agricultural establishments (EA) of family farming (AF) or of the entire municipal agriculture (family or not). A first database with 40 variables and 144 observations was processed to be in an appropriate format for factor analysis.

As the analytical process is based on a correlation matrix, the analysis begins with the construction of this matrix containing Pearson's linear correlation values, denominated the ρ matrix.

The ρ matrix presents correlation coefficients between all pairs of variables, being symmetric with respect to the main diagonal, whose values are equal to 1. Since these correlation values can vary on a scale between -1 and 1, pairs of variables with correlation values closer to 1, in modulus, indicate a strong linear relationship between such variables and should indicate a correlation with the same extracted factor.

On the other hand, correlation values close to 0 indicate weak or nonexistent correlation between variables and the need for extraction of different factors to capture the variance or, ultimately, that factor analysis is inappropriate (Malhotra, 2019; Fávero and Belfiore, 2017). In this sense, for an adequate extraction of factors, the ρ matrix must present statistically significant and relatively high values.

The next step of the analysis was then to test the degrees of correlation between the 40 initial variables, which represented the first indication for cutting any variables not suitable for the analysis. A Pearson correlation ρ matrix was elaborated for this initial dataset, highlighting the variables with statistically zero correlation. Figure 1 illustrates the result of this initial ρ matrix. The white squares represent correlations equal to zero.

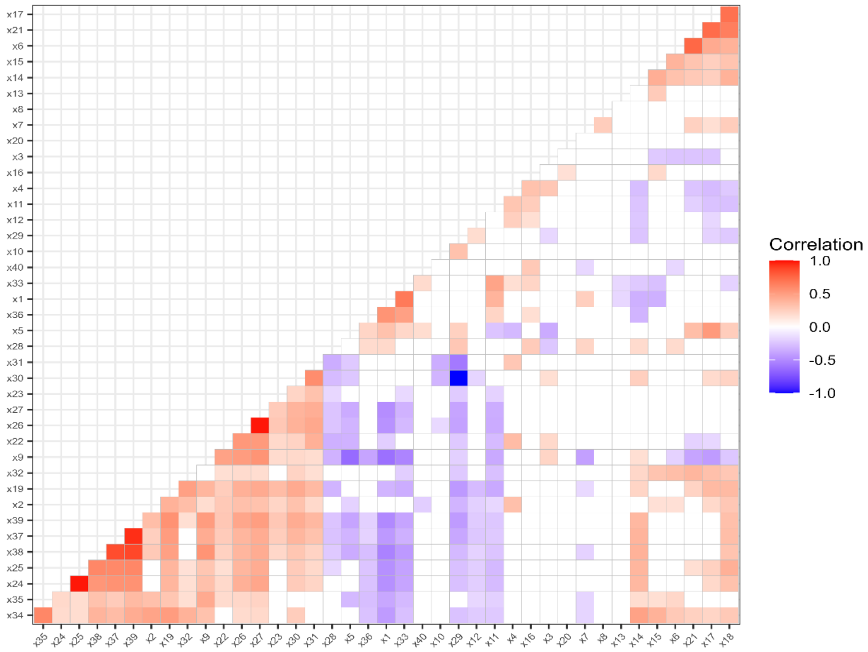

Figure 1. Pearson Correlation Coefficient Matrix
Source: Original research data

After an initial assessment of correlation coefficient values and some preliminary experiments of factor loadings and communalities resulting from PCA, 16 variables were selected to compose the definitive database.

Table 1. Variables used and description of variables

Variable	Variable description
Pesticides	Percentage of agricultural establishments (EA) of family farming (AF) using pesticides.
hort_flor	Percentage of EA of the AF of the economic activity group Horticulture and floriculture.
fam_farm	Percentage of FA AE in relation to the total number of agricultural establishments in the municipality.
temp_crops	Percentage of EA of the AF of the economic activity group production of temporary crops.
livestk	Percentage of AE of the AF of the economic activity group Livestock and breeding of other animals.
ff_area	Percentage of the AF area in relation to the total area of agricultural establishments in the municipality.
soil_corr	Percentage of agricultural establishments (EA) of family farming (AF) with soil correction.
fertil	Percentage of agricultural establishments (EA) in family farming (AF) with fertilization.
irrigation	Percentage of agricultural establishments (EA) of family farming (AF) with irrigation.
expenses	Average value of expenses of AF EAs that incurred expenses in the municipality.
renda_princ	Percentage of agricultural establishments (EA) of family farming (AF) in which the main income is obtained from activities carried out in the establishment.
comm_purple	Percentage of agricultural establishments (EA) of family farming (AF) whose main purpose of production is commercialization.
vehicles	Percentage of agricultural establishments (EA) of family farming (AF) with vehicles.
prod_value	Average production value of AF AEs with production in the municipality.
revenue	Average revenue value of AF EAs that obtained revenue in the municipality.
financing	Percentage of agricultural establishments (EA) from family farming (AF) that received financing.

A new matrix ρ was generated and exhibited improved correlations among the variables, as illustrated by Figure 2:

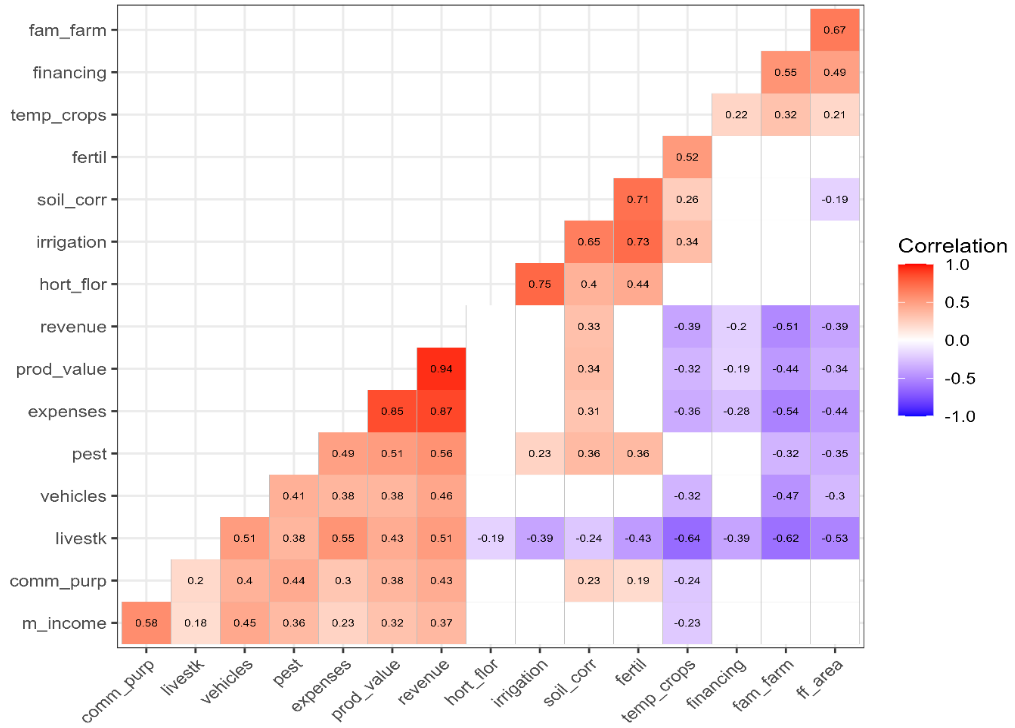

Figure 2. Pearson Correlation Coefficient Matrix
Source: Original research data

The next step involved assessing the overall adequacy of the factor analysis. The study utilized Bartlett's sphericity test (Fávero & Belfiore, 2017):

$$\chi^2_{Bartlett} = -\left[(n-1) - \left(\frac{2k+5}{6}\right)\right] \cdot ln\ |D| \qquad (1)$$

Where n represents the number of observations and k the number of variables, and D denotes the determinant of the matrix ρ. The degrees of freedom are calculated by $\frac{k \cdot (k-1)}{2}$.

The test compares the matrix ρ with an identity matrix I of the same dimension, checking whether the values outside the main diagonal are statistically different from 0, for a certain number of degrees of freedom and a specified level of significance. For the research matrix ρ, with 120 degrees of freedom, the p-value was 1.040742e-278. As the null hypothesis ($H_{0:\, \rho = I}$)

The analysis was rejected, and the procedure proceeded to extract factors based on eigenvalue computation.

From the stats package of the R language, version 4.1.3, the prcomp() algorithm, as described by Venables and Ripley (2002), Becker et al. (1988), and Mardia et al. (1979), was utilized for principal component extraction. Prior to this extraction, standardization was performed using the z-score procedure on the database. Sixteen principal components and their corresponding eigenvalues were obtained. Each eigenvalue is associated with a factor or principal component and represents the amount of total variance attributed to its corresponding factor (Malhotra, 2019). In this study, the eigenvalue of PC1 captures the variance of 5.4 variables, which corresponds to approximately 33% of the total shared variance among the variables. Only the first four factors capture more than 70% of the joint behavior of the database. Table 2 presents the 16 extracted factors in the form of principal components (PC), their respective eigenvalues, the shared variances captured, and the cumulative variance

Table 2. Principal components, eigenvalues, percentage of shared variance (%) and cumulative variance explained (%)

Principal Component (PC)	Autovalor	Shared variance	Cumulative variance
PC1	5,420348082	0,33877	0,33877
PC2	3,476734404	0,2173	0,55607
PC3	1,761040132	0,11007	0,66613
PC4	1,086418251	0,0679	0,73403
PC5	0,990610473	0,06191	0,79595
PC6	0,655329939	0,04096	0,83691
PC7	0,539243708	0,0337	0,87061
PC8	0,450280105	0,02814	0,89875
PC9	0,409012175	0,02556	0,92431
PC10	0,331957478	0,02075	0,94506
PC11	0,243371949	0,01521	0,96027
PC12	0,195900578	0,01224	0,97252
PC13	0,149500268	0,00934	0,98186
PC14	0,135950463	0,0085	0,99036
PC15	0,106621701	0,00666	0,99702
PC16	0,047680295	0,00298	1

The first part of the analysis was concluded with the calculation of factor loadings, which are Pearson correlations between the original variables and the factors, allowing for the interpretation of the importance of each variable in the construction of each factor, i.e., variables that exhibited high loadings with respect to it (Malhotra, 2019). Figure 3 illustrates the relationship of the variables with each principal component.

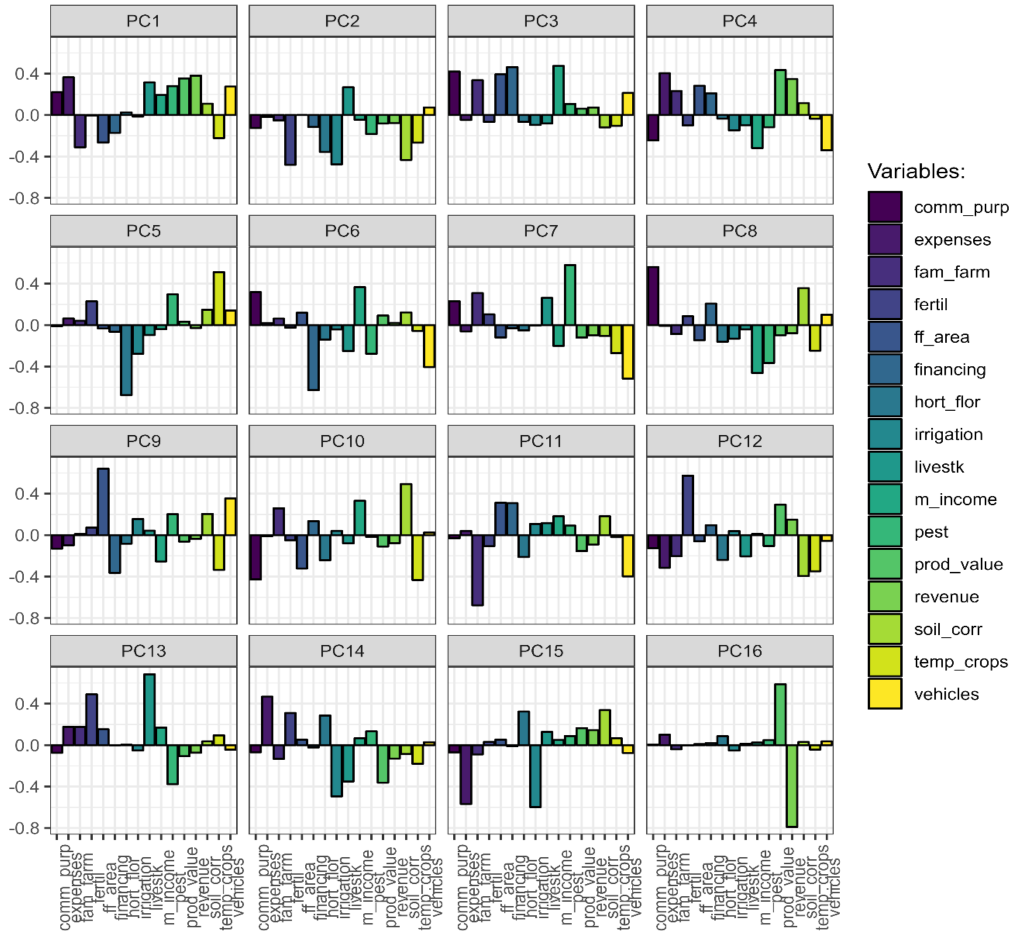

Figure 3. Relationship between original variables and principal components
Source: Original research data

The analysis proceeded with the determination of the number of factors used in the study and the calculation of factor scores. Since the number of principal components equals the number of existing variables (Malhotra, 2019), there are several procedures for determining the quantity of factors selected for analysis. This study employed as criteria for selecting principal components the latent root (approach based on eigenvalues greater than 1) and the cumulative percentage of extracted variance, relying on the Scree Plot.

Although according to the latent root criterion, the first four principal components would be chosen, through the analysis of cumulative variance increment, only the first three were selected, as PC4 adds only a little over 6% to the total variance already accumulated by the first three factors. Figure 4 illustrates this argument

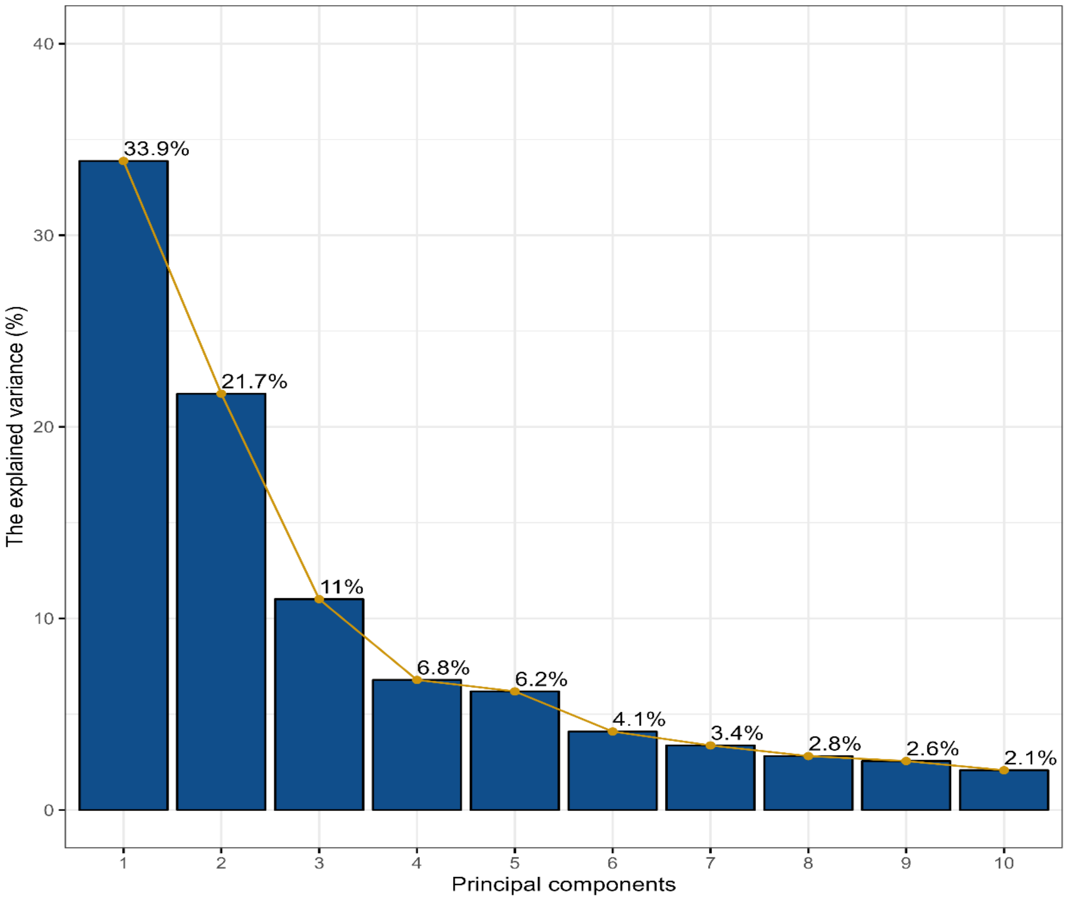

Figure 4. Scree Plot of the first 10 principal components

With the selection of the three main components, the variable with the lowest communalities, percentage of shared variance with all selected factors, was "hort_flor", with 45%; however, since this variance was captured almost exclusively in the second principal component, it was retained. The remaining variables have communalities greater than 50%, with 11 out of 16 variables presenting communalities greater than 60%. Table 3 displays the factor loadings of each variable and their respective communalities.

Table 3. Factor loadings of variables with each principal component and their respective communalities.

	Factor Loadings			
Variable	PC1	PC2	PC3	Communalities
revenue	0,886	-0,143	0,098	0,815
fertil	-0,012	-0,895	-0,089	0,809
irrigation	-0,033	-0,888	-0,127	0,806
livestk	0,733	0,499	-0,108	0,799
soil_corr	0,254	-0,807	-0,157	0,741
fam_farm	-0,724	-0,096	0,446	0,732
expenses	0,847	-0,032	-0,064	0,723
prod_value	0,824	-0,153	0,081	0,709
ff_area	-0,619	0,000	0,521	0,655
comm_purple	0,512	-0,232	0,556	0,625
renda_princ	0,450	-0,086	0,629	0,605
financing	-0,400	-0,215	0,615	0,584

	Factor Loadings			
Variable	PC1	PC2	PC3	Communalities
Pesticides	0,651	-0,343	0,143	0,562
temp_crops	-0,517	-0,497	-0,137	0,533
vehicles	0,639	0,138	0,283	0,507
hort_flor	0,061	-0,664	-0,088	0,453

The next step involved calculating the factorial scores, which relate a specific factor to the original variables, akin to the parameters of a linear model, and are computed by dividing the eigenvectors by the square root of the eigenvalues associated with each component (Fávero & Belfiore, 2017). Table 4 presents the score of each variable in each principal component, used for the formation of the municipalities' scores in each of the three principal components.

Table 4. Scores factorials of the variables in each main component

	Scores factorials		
Variable	PC1	PC2	PC3
Pesticides	0,12009	-0,09865	0,08099
hort_flor	0,01132	-0,19105	-0,04971
fam_farm	-0,13356	-0,02750	0,25346
temp_crops	-0,09532	-0,14297	-0,07796
livestk	0,13530	0,14359	-0,06134
ff_area	-0,11420	0,00011	0,29608
soil_corr	0,04692	-0,23218	-0,08929
fertil	-0,00216	-0,25740	-0,05055
irrigation	-0,00614	-0,25543	-0,07216
expenses	0,15635	-0,00928	-0,03620
renda_princ	0,08300	-0,02464	0,35703
comm_purple	0,09440	-0,06669	0,31589
vehicles	0,11783	0,03970	0,16054
prod_value	0,15203	-0,04388	0,04623
revenue	0,16347	-0,04115	0,05538
financing	-0,07388	-0,06195	0,34900

With the defined scores, the values of each principal component, or factor, were calculated for each municipality. These values were then plotted on the map of the state of Pará for a geospatial analysis of the result.

III. Results And Discussion

The three factors emerging from the analysis are independent of each other and thus should be interpreted separately, albeit allowing for a complementary evaluation among them.

PC1 exhibited a strong positive correlation with the variables in the financial transaction table and with the variable "livestock," while showing a negative correlation with the variables "fam_farm" and "ff_area," which represent the percentage presence of family agriculture in the municipalities. Conversely, PC2 demonstrated positive correlations with activities related to soil management, represented by the variables "irrigation," "fertilization," "soil_color," and "hort_flor," and negatively with the variable "livestock." Finally, PC3 captures the variance of the variables "principal_income," "financing," "commercial_purpose," and "ff_area."

Figure 5 depicts the relationship of the variables with the first two principal components.

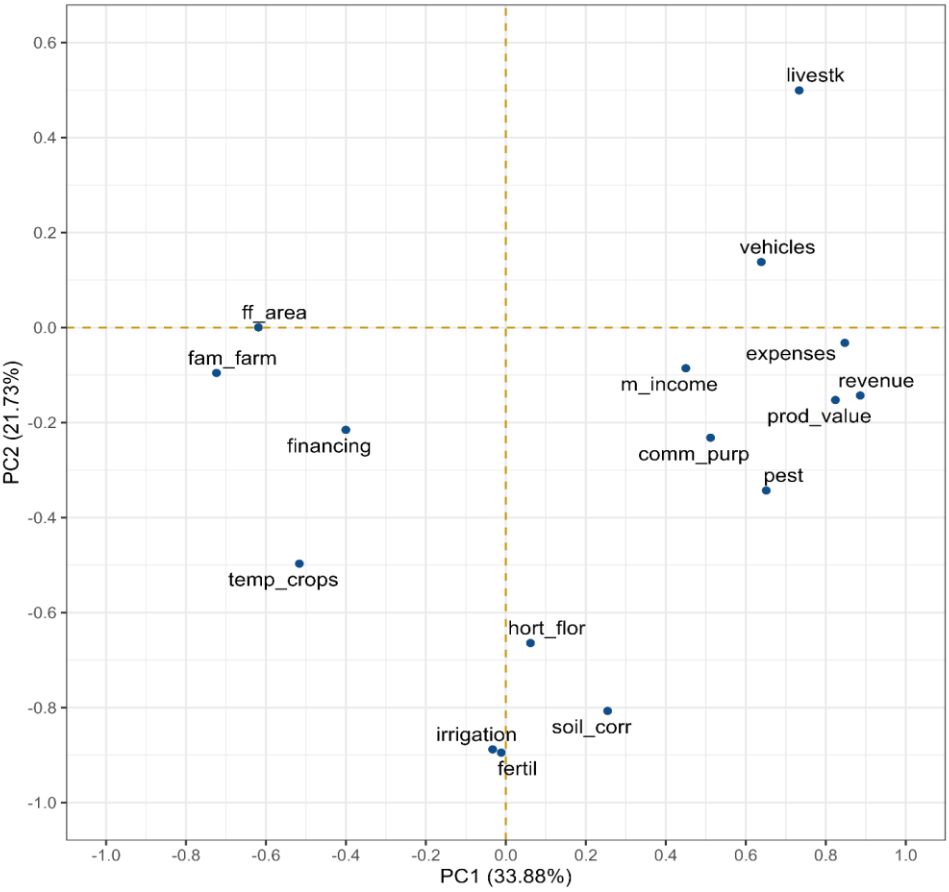

Figure 5. Factor loadings of variables on the first two principal components
Subsequently, each factor is presented in greater depth.

Factor 1 - Agricultural Financial Dynamism

Factor 1 is closely related to municipalities with a more commercially oriented family agriculture, with emphasis on variables derived from the financial turnover table, such as revenues, expenses, and production value. Table 5 presents the top ten municipalities with the highest scores in this factor, as well as the top five in each quartile of 36 municipalities. In quartile 1, the group of municipalities with the lowest scores in the factor, the five municipalities with the lowest scores in the overall ranking are also presented.

Table 5. Factor 1 of the top 10 municipalities, the top 5 in each quartile, and the bottom 5 in the overall score.

				Position	
cod_municipio	County	Fator1	Quartile	In the quartile	General
1502939	Dom Eliseu (PA)	4,00162	Q4	1	1
1502772	Curionópolis (PA)	2,28380	Q4	2	2
1505031	Novo Progresso (PA)	2,26518	Q4	3	3
1506161	Rio Maria (PA)	2,18887	Q4	4	4
1507458	São Geraldo do Araguaia (PA)	2,03339	Q4	5	5
1500347	Água Azul do Norte (PA)	1,86595	Q4	6	6
1508084	Tucumã (PA)	1,85689	Q4	7	7
1502152	Canaã dos Carajás (PA)	1,80099	Q4	8	8
1505437	Ourilândia do Norte (PA)	1,68163	Q4	9	9
1505635	Piçarra (PA)	1,62536	Q4	10	10
1500131	Abel Figueiredo (PA)	0,78838	Q3	1	37
1507508	São João do Araguaia (PA)	0,72643	Q3	2	38
1506500	Santa Izabel do Pará (PA)	0,67225	Q3	3	39
1505502	Paragominas (PA)	0,64025	Q3	4	40
1507953	Tailândia (PA)	0,59656	Q3	5	41
1507961	Terra Alta (PA)	-0,24081	Q2	1	73
1500503	Almeirim (PA)	-0,25035	Q2	2	74
1504703	Moju (PA)	-0,27484	Q2	3	75
1507102	São Caetano de Odivelas (PA)	-0,27874	Q2	4	76

				Position	
cod_municipio	County	Fator1	Quartile	In the quartile	General
1504950	Nova Esperança do Piriá (PA)	-0,27959	Q2	5	77
1503309	Igarapé-Miri (PA)	-0,77429	Q1	1	109
1504406	Marapanim (PA)	-0,77927	Q1	2	110
1500206	Acará (PA)	-0,83413	Q1	3	111
1501501	Benevides (PA)	-0,85625	Q1	4	112
1505304	Oriximiná (PA)	-0,86104	Q1	5	113
1505205	Oeiras do Pará (PA)	-1,45548	Q1	32	140
1506203	Salinópolis (PA)	-1,50067	Q1	33	141
1504505	Melgaço (PA)	-1,57672	Q1	34	142
1502608	Colares (PA)	-1,74584	Q1	35	143
1506104	Primavera (PA)	-1,96958	Q1	36	144

Source: Original survey results

"Municipalities with higher scores in this component exhibit higher average revenue per total number of family farming establishments, as well as higher average expenditures in their establishments, indicating significant financial dynamism. This is reinforced by two other important variables in the composition of this factor: family farming establishments with higher average production values and production primarily destined for commercialization. Extrapolating this factor to the state level is consistent with the study on the modernization of agriculture in municipalities in the Northeast of Pará, conducted by Rebello et al. (2011). The authors identify Factor 1 (Capital Intensity) strongly correlated with variables related to investments, expenses, and tractor usage.

In terms of activity, livestock farming stands out as the primary activity in these municipalities. Greater technological intensity may be indicated by the prominent use of pesticides and a higher presence of vehicles in family farming establishments in the higher scoring group of this component. Conversely, municipalities in the upper quartiles in this factor have a lower percentage of family farming establishments compared to the total number of agricultural establishments and a smaller area dedicated to family farming compared to the total agricultural area. In terms of activities, temporary crops increase their presence in municipalities in the lower quartiles, i.e., with the lowest scores in this component. Table 6 presents the means of the values and proportions, by quartile of 36 municipalities, of the variables with the highest factor loadings of this principal component, considering an increasing score. There is a clear positive correlation between the 1st principal component and the variables 'revenue', 'expenses', 'prod_value', 'livestock', 'pesticides', and 'vehicles'; and a negative correlation with the variables 'temp_crops', 'area_ff', and 'fam_farm'."

Table 6. Means of variables with highest factor loadings on PC1 by municipality quartile.

Variable		Quartile 1	Quartile 2	Quartile 3	Quartile 4
revenue	R$	14.004,44	20.983,90	27.630,60	44.381,67
expenses	R$	2.951,96	5.319,05	9.870,83	19.351,04
prod_value	R$	14.289,67	20.529,92	25.005,65	40.654,76
livestk	(%)	0,07	0,19	0,47	0,71
Pesticides	(%)	0,05	0,11	0,20	0,37
vehicles	(%)	0,07	0,13	0,26	0,43
comm_purple	(%)	0,58	0,65	0,72	0,88
temp_crops	(%)	0,52	0,50	0,32	0,15
ff_area	(%)	0,65	0,48	0,36	0,26
fam_farm	(%)	0,91	0,87	0,81	0,74
Source: Original research findings					

In terms of territorial distribution, as depicted in Figure 6, although there is considerable diffusion of this factor in the state of Pará, it is observed that municipalities with higher scores are located in the southern and southeastern edges of the state, an agricultural frontier region linked to agricultural activities and also known as the "Arc of Deforestation" in the Amazon, due to the predatory cycle of deforestation and degradation of natural resources, accompanied by violence in rural areas, with subsequent cattle ranching as a subsequent activity.

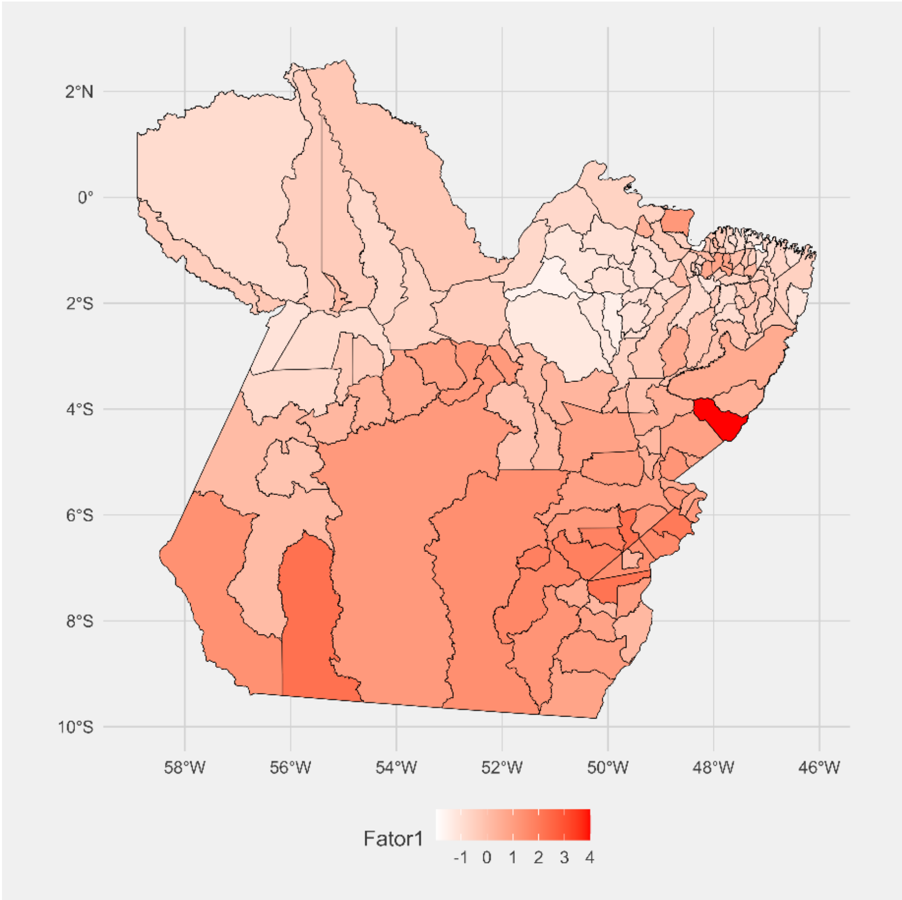

Figure 6. Map of Pará with the classification of municipalities according to Factor 1

Factor 2 - Soil and Crop Management

The main component 2 is related to the set of variables associated with soil management practices for agricultural production, such as correction and fertilization activities, with the presence of irrigation. This factor is composed of the activity groups Temporary Crop Production and Horticulture and Floriculture, positively correlated, and livestock, negatively correlated.

Municipalities with higher scores in this component have higher percentages of family farming establishments with fertilization, irrigation, and soil correction practices. Additionally, 7% of family farming establishments in the upper quartile municipalities of this component engage in horticulture and floriculture as their main activity, while only 1% engage in this activity in the other quartiles. Meanwhile, temporary crop production is practiced by 55% of establishments in the 4th quartile and by 52% in the 3rd quartile, while in the lower quartiles, it is practiced by only 28% of establishments in the 2nd quartile and by 14% in the 1st quartile.

Table 7 presents the top ten municipalities with the highest scores in this factor, the five best in each quartile of 36 municipalities, and the five with the lowest scores in the overall ranking.

Table 7: The 10 municipalities with the highest scores on this factor, the top 5 in each quartile out of 36 municipalities, and the 5 with the lowest scores in the ranking.

				Position	
cod_municipio	County	Fator1	Quartile	In the quartile	General
1507003	Santo Antônio do Tauá (PA)	3,34051	Q4	1	1
1506500	Santa Izabel do Pará (PA)	2,97396	Q4	2	2
1502905	Curuca (PA)	2,87623	Q4	3	3
1502400	Castanhal (PA)	2,87004	Q4	4	4
1507409	São Francisco do Pará (PA)	2,85666	Q4	5	5
1502939	Dom Eliseu (PA)	2,47517	Q4	6	6
1503200	Igarapé-Açu (PA)	2,29560	Q4	7	7
1507466	São João da Ponta (PA)	2,01255	Q4	8	8
1506609	Santa Maria do Pará (PA)	1,94753	Q4	9	9
1505007	Nova Timboteua (PA)	1,92683	Q4	10	10
1506351	Santa Bárbara do Pará (PA)	0,53811	Q3	1	37
1504950	Nova Esperança do Piriá (PA)	0,50258	Q3	2	38
1503044	Araguaia Forest (PA)	0,45753	Q3	3	39
1502756	Concordia do Pará (PA)	0,42006	Q3	4	40
1501709	Bragança (PA)	0,41628	Q3	5	41
1505494	Palestine of Pará (PA)	-0,34511	Q2	1	73
1504208	Marabá (PA)	-0,35457	Q2	2	74
1506559	Santa Luzia do Pará (PA)	-0,38040	Q2	3	75
1502103	Cametá (PA)	-0,38341	Q2	4	76
1500602	Altamira (PA)	-0,38431	Q2	5	77
1501402	Belém (PA)	-0,69496	Q1	1	109
1506161	Rio Maria (PA)	-0,69948	Q1	2	110
1503705	Itupiranga (PA)	-0,70616	Q1	3	111
1501956	Piriá Waterfall (PA)	-0,71054	Q1	4	112
1502004	Arari Waterfall (PA)	-0,71305	Q1	5	113
1502509	Chaves (PA)	-1,14900	Q1	32	140
1501253	Bannach (PA)	-1,20366	Q1	33	141
1507755	Sapucaia (PA)	-1,42857	Q1	34	142
1507904	Soure (PA)	-1,49570	Q1	35	143
1506401	Santa Cruz do Arari (PA)	-1,69495	Q1	36	144

Source: Original survey results

These results also find equivalence with Rebello et al. (2011), in Factor 2, denominated "Use of modern technology," which showed a strong positive correlation with the use of fertilizers, irrigation, and soil correction practices. Table 8 presents the means of the proportions, by quartile of 36 municipalities, of the variables with the highest factorial loadings of this principal component, considering an increasing score. A positive correlation is observed between the 2nd principal component and the variables "fertilization," "irrigation," "soil_correction," "horticulture_floriculture," and "washing_time"; and a negative correlation with the variable "livestock."

Table 8. Means of the variables with the highest factor loadings of PC2 by quartile of municipalities

Variable		Quartile 1	Quartile 2	Quartile 3	Quartile 4
fertil	(%)	0,04	0,08	0,23	0,53
irrigation	(%)	0,01	0,02	0,06	0,21
soil_corr	(%)	0,01	0,02	0,05	0,12
hort_flor	(%)	0,01	0,01	0,01	0,07
temp_crops	(%)	0,14	0,28	0,52	0,55
livestk	(%)	0,59	0,43	0,26	0,15

Source: Original research findings

Territorially, municipalities with higher scores in this component are concentrated in the region of the oldest colonization in the Amazon, especially in the metropolitan mesoregions of Belém, Castanhal microregion, and Northeastern Pará, Bragantina and Salgado microregions (Rabelo et al., 2011). Among the top 15, there is only the municipality of Dom Eliseu outside this concentration region, in Southeastern Pará, ranking 6th. It is important to note that this municipality ranked first in Factor 1. Figure 7 illustrates this spatial distribution.

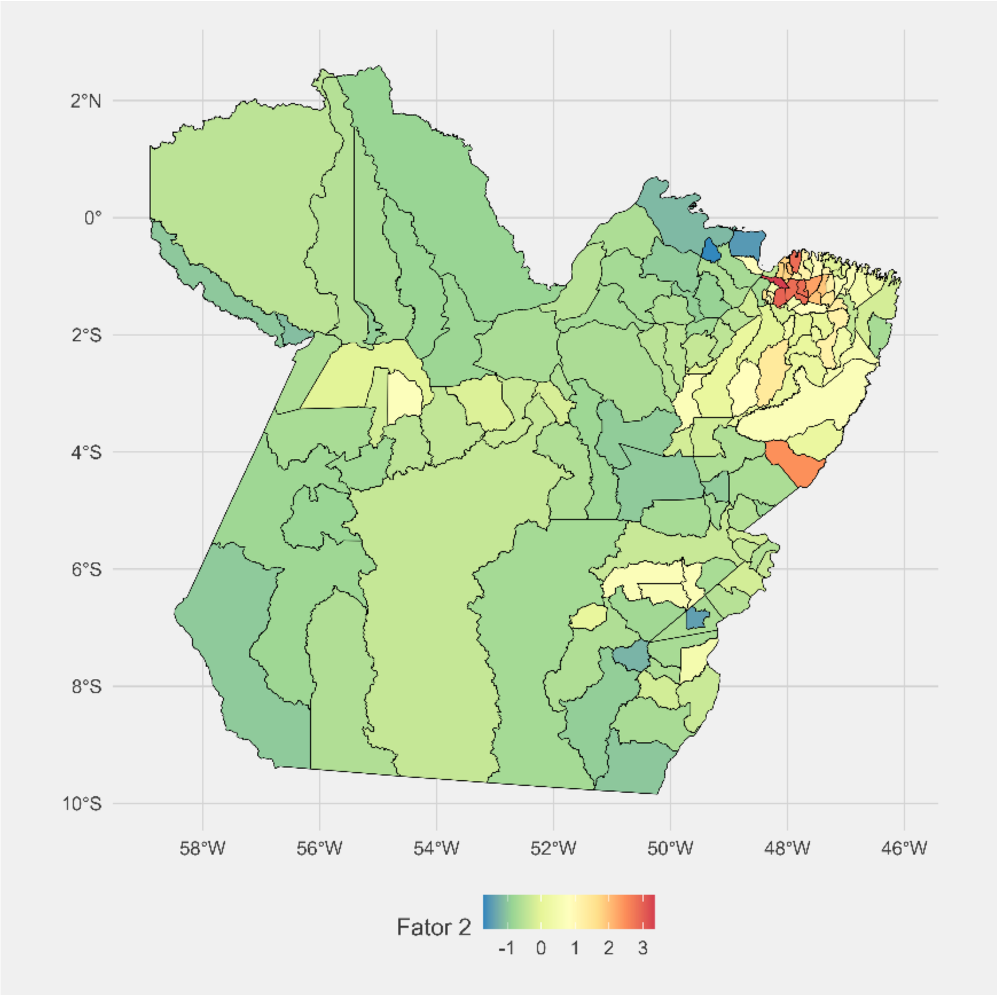

Figure 7. Map of Pará with classification of municipalities according to Factor 2

Factor 3 – Financing of Family Agriculture

Principal component 3 is a sort of subset of principal component 1 for municipalities with higher percentages of agricultural areas dedicated to family farming. These are municipalities where agricultural establishments of family farming predominantly allocate their production for commercialization, and where the primary income is derived from agricultural activities carried out on these establishments. Table 9 presents the top ten municipalities with the highest scores in this factor, the top five in each quartile of 36 municipalities, and the five with the lowest scores in the overall ranking.

Table 9. Factor 3 of the top 10 municipalities, the top 5 in each quartile, and the bottom 5 in the overall score.

				Position	
cod_municipio	County	Fator1	Quartile	In the quartile	General
1503101	Gurupa (PA)	2,395743	Q4	1	1
1502954	Eldorado do Carajás (PA)	1,765290	Q4	2	2
1501303	Barcarena (PA)	1,696750	Q4	3	3
1500305	Afuá (PA)	1,674036	Q4	4	4
1505650	Plates (PA)	1,654037	Q4	5	5
1501105	Bagre (PA)	1,613537	Q4	6	6

				Position	
cod_municipio	County	Fator1	Quartile	In the quartile	General
1508084	Tucumã (PA)	1,563738	Q4	7	7
1502806	Curralinho (PA)	1,553980	Q4	8	8
1503705	Itupiranga (PA)	1,537581	Q4	9	9
1507201	São Domingos do Capim (PA)	1,492589	Q4	10	10
1502301	Captain Poço (PA)	0,646684	Q3	1	37
1502756	Concordia do Pará (PA)	0,641882	Q3	2	38
1507508	São João do Araguaia (PA)	0,641723	Q3	3	39
1504208	Marabá (PA)	0,640814	Q3	4	40
1507607	São Miguel do Guamá (PA)	0,623560	Q3	5	41
1500800	Ananindeua (PA)	0,031691	Q2	1	73
1506708	Santana do Araguaia (PA)	0,019165	Q2	2	74
1501709	Bragança (PA)	0,012894	Q2	3	75
1504422	Marituba (PA)	-0,004928	Q2	4	76
1506807	Santarém (PA)	-0,067939	Q2	5	77
1500404	Alenquer (PA)	-0,639142	Q1	1	109
1506203	Salinópolis (PA)	-0,664077	Q1	2	110
1501204	Baião (PA)	-0,701243	Q1	3	111
1506302	Salvaterra (PA)	-0,706036	Q1	4	112
1507003	Santo Antônio do Tauá (PA)	-0,716901	Q1	5	113
1504109	Magalhães Barata (PA)	-1,979808	Q1	32	140
1508126	Ulianopolis (PA)	-1,992532	Q1	33	141
1506401	Santa Cruz do Arari (PA)	-2,868233	Q1	34	142
1507904	Soure (PA)	-3,465559	Q1	35	143
1507755	Sapucaia (PA)	-3,618772	Q1	36	144

Source: Original survey results

An important variable within the composition of this factor is access to financing, which demonstrates a certain convergence with Factor 3 in Rebello et al.'s (2011) analysis, as financing is a significant variable in this study. Factor 3 may indicate a certain political-institutional dimension of social organization among the producing families in these municipalities, as access to financing for family-based agricultural activities is typically associated with higher levels of education and income, and is recognized as deficient in rural areas of the Amazon as a whole (Cruz et al., 2020).

It is interesting to observe the relationship between the availability of land, surplus production, and access to financing on one hand, and the possibility of predominantly agricultural income on the other. Municipalities with higher scores on this component exhibit higher percentages of agribusiness establishments engaged in family farming.

Table 10 presents the means of the proportions, by quartile of 36 municipalities, of the variables with the highest factorial loads for this principal component, considering an increasing score.

Table 10. Means of variables with the highest factor loadings of PC3 by quartile of municipalities.

Variable		Quartile 1	Quartile 2	Quartile 3	Quartile 4
renda_princ	(%)	0,42	0,56	0,63	0,72
financing	(%)	0,68	0,79	0,82	0,87
comm_purple	(%)	0,52	0,71	0,76	0,85
ff_area	(%)	0,31	0,41	0,45	0,59
Source: Original survey results					

Analyzing the spatial distribution of the factor, one can observe certain corridors of municipalities: Marajó (Gurupá, Afuá, Bagre, and Curralinho), Northeastern Pará, and municipalities influenced by the colonization projects of the military dictatorship, especially by the Trans-Amazonian Highway (BR-230), such as Uruará, Rurópolis, Itupiranga, Senador José Porfírio, Tucumã, and Eldorado do Carajás. This configuration reinforces the possibility of a higher degree of political and institutional organization indicated by this factor, as the historical formation of these municipalities encourages the construction of a peasant identity. Figure 8 illustrates this spatial distribution.

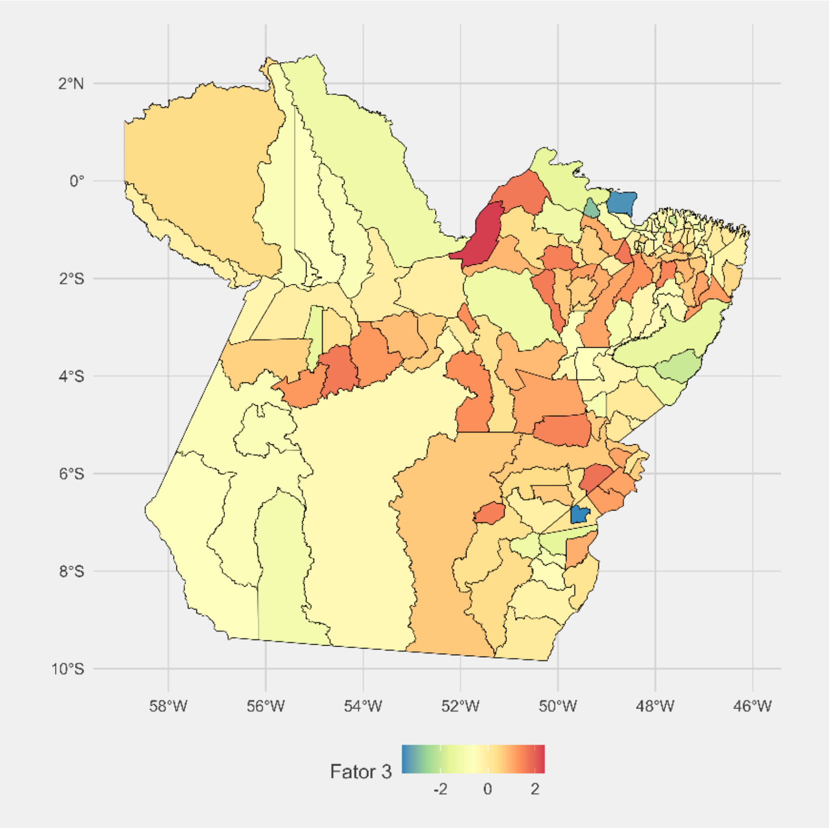

Figure 8 depicts a map of the state of Pará, illustrating the classification of municipalities according to Factor 3.

IV. Final Considerations

Family farming is a key component of any rural development process, and in Pará, its importance is growing due to its strong presence in all rural areas of the state. It was from this perception that this research sought to capture latent variables that reflected "hidden" phenomena of family-based agriculture dynamics.

The multivariate technique of principal component factor analysis was used, based on 16 indicators derived from variables available in the 2017 Agricultural Census. The research successfully captured factors that demonstrate social relations in the Pará territory and are somewhat supported by previous research. Three factors were captured that, in a shared manner, represent approximately 67% of the total variance of the database.

Factor 1, labeled Agribusiness Financial Dynamism, is related to a more commercial character of family agriculture, with emphasis on variables such as revenue, expenses, and production value. In terms of territorial distribution, although there is considerable diffusion of this factor in the state of Pará, it is noticeable that municipalities with higher scores are located in the southern and southeastern borders of the state, a region of agricultural frontier.

Factor 2, labeled Soil and Crop Management, is related to the set of soil management practices, soil correction, fertilization, and irrigation. Territorially, municipalities with higher scores in this component are concentrated in the region of the oldest colonization of the Amazon, especially in the Metropolitan region of Belém, Castanhal microregion, and Northeastern Pará, Bragantina, and Salgado microregions.

Finally, Factor 3, labeled Family Agriculture Financing, is a kind of subset of Factor 1 for municipalities with higher percentages of agricultural areas dedicated to family farming. These are municipalities where family farming establishments primarily allocate their production for commercialization and where the main income comes from agricultural activities. The factor is spatially distributed in some corridors of municipalities: Marajó, Northeastern Pará, and municipalities influenced by the colonization projects of the military dictatorship, especially by the Trans-Amazon Highway (BR-230).

References

[1]. Becker, R. A.; Chambers, J. M.; Wilks, A. R. 1988. The New S Language. Wadsworth & Brooks/Cole.

[2]. Cruz, N. B.; Jesus, J.G.; Bacha, C.J.C.; Costa, E. M. 2020. Family Farming Access To Credit And Technical Assistance In Brazil. Journal Of Rural Economics And Sociology, 59(3), 1–20.Https://Doi.Org/10.1590/1806-9479.2021.226850

[3]. Favareto, A. 2010. The Territorial Approach To Rural Development-Institutional Change Or "Innovation By Addition"?. Advanced Studies, V. 24, No. 68, P. 299–319.

[4]. Fávero, L. P.; Belfiore. 2017. P. Data Analysis Manual: Statistics And Multivariate Modeling With Excel®, Spss® And Stata®. Elsevier Brazil.
[5]. Amazon Foundation To Support Studies And Research [Fapespa]. 2022a. Pará In Numbers - 2022. 4th Edition. Available At: <Https://Www.Fapespa.Pa.Gov.Br/Node/81>. Accessed On: April 4, 2023.
[6]. Amazon Foundation To Support Studies And Research [Fapespa]. 2022b. Pará In The National Context 2022: Gross Domestic Product Table (Current Values), According To Brazil, Major Regions And Federation Units - 2016-2020. Available At: <Https://Www.Fapespa.Pa.Gov.Br/Sistemas/Pcn2022/Tabelas/11-Pib/1-Produto-Interno-Bruto-Valores-Corries-2016-2020.Htm>. Accessed On: April 4, 2023.
[7]. Grisa, C.; Sabourin, E. 2019. Family Agriculture: From Concepts To Public Policies In Latin America And The Caribbean. 2030-Food, Agriculture And Rural Development In Latin America And The Caribbean, 15, 3–19. Http://Www.Fao.Org/Publications/Es
[8]. Brazilian Institute Of Geography And Statistics [Ibge]. 2019. Agricultural Census 2017. Available At: <Https://Sidra.Ibge.Gov.Br/Pesquisa/Censo-Agropecuario/Censo-Agropecuario-2017/Resultados-Definitivos>. Accessed In: Sea. 2023.
[9]. Kageyama, A. 2004. Rural Development: Concept And Measurement. Cadernos De Ciência & Tecnologia, 21(3), 379–408.
[10]. Lobão, M. S. P.; Staduto, J. A. R. 2020. Agricultural Modernization In The Brazilian Amazon. Journal Of Rural Economics And Sociology, 58(2), E188276. Https://Doi.Org/10.1590/1806-9479.2020.182276
[11]. Malhotra, N., K. 2019. Marketing Research: An Applied Orientation. Pearson.
[12]. Mardia, K. V.; Kent, J. T.; Bibby, J. M. 1979. Multivariate Analysis, London: Academic Press.
[13]. Melo, C. O.; Parré, J. L. 2007. Rural Development Index Of Municipalities In Paraná: Determinants And Hierarchy. Journal Of Rural Economics And Sociology, V. 45, No. 2, P. 329–365.
[14]. Rebello, F. K.; Santos, M. A. S.; Homma, A. K. O. 2011. Modernization Of Agriculture In The Municipalities Of Northeast Pará: Determinants And Hierarchization In The Year 2006. Revista De Economia E Agribusiness/Brazilian Review Of Economics And Agribusiness, 9, 209–232. Https://Doi.Org/10.22004/Ag.Econ.121296
[15]. Sabourin, E. 2017. Origins, Evolution And Institutionalization Of Family Farming Policy In Brazil. In G. C. Delgado; S. M. P. P. Bergamasco (Eds.). Brazilian Family Farming: Challenges And Future Perspectives (Pp. 84–109). Brasília, Df, Brazil.
[16]. Schneider, S. 2010. Situating Rural Development In Brazil: The Context And Issues Under Debate. Brazilian Journal Of Political Economy, V. 30, No. 3, P. 511–531.
[17]. Schneider, S.; Cassol, A. 2017. Diversity And Heterogeneity Of Family Farming In Brazil And Implications For Public Policies. In G. C. Delgado; S. M. P. P. Bergamasco (Eds.). Brazilian Family Farming: Challenges And Future Perspectives (Pp. 84–109). Brasília, Df, Brazil.
[18]. Schneider, S.; Escher, F. 2011. Karl Polanyi's Contribution To The Sociology Of Rural Development. Sociologies, V. 13, No. 27, P. 180–219.
[19]. Schneider, S.; Tartaruga, I. G. P. 2004. Territory And Territorial Approach: From Cognitive References To Contributions Applied To The Analysis Of Rural Social Processes. Roots, 23(02), 99–116. Http://Www.Ufcg.Edu.Br/~Raizes/Artigos/Artigo_4.Pdf
[20]. Stege, A. L.; Parré, J. L. 2013. Factors That Determine Rural Development In Microregions Of Brazil. Confins, 19. Https://Doi.Org/10.4000/Confins.8640
[21]. Venables, W. N; Ripley, B. D. 2002. Modern Applied Statistics With S, Springer-Verlag.