STATISTICAL MODELING OF SUSTAINABILITY IN AGRICULTURE IN THE STATE OF PARÁ IN 2022: AN ANALYSIS USING PEARSON'S CORRELATION.

.Eduardo Da Silva Pinheiro[1], Amanda Milene De Sousa Monteiro[2]

Federal University of Pará

Belém, Brasil

eduardosilva3256@gmail.com

ABSTRACT

In recent years, agriculture has experienced significant growth, aligned with the Sustainable Development Goals (SDGs), aiming for clean and sustainable production amidst climate change, increasing urbanization, and resource scarcity, directly affecting prices and production. Therefore, this study aims to investigate the relationship between variables associated with agricultural production in the region, using Pearson's correlation coefficient as a precise quantitative measure of these relationships. To achieve this goal, Pearson's correlation coefficient will be employed to measure the direction and degree of the linear relationship between two quantitative variables. Data will be collected on agricultural production in the Pará region in 2022, as well as information related to climatic conditions and biological factors influencing this production. The results obtained will reveal the correlations between variables associated with agricultural production, and it is expected to identify patterns and significant relationships between climate, biological factors, and agricultural production in the region. Therefore, the use of Pearson's correlation coefficient will provide an understanding of what influences agricultural production aiming to promote more sustainable agriculture.

Keywords: Statistical modeling. Multivariate analysis. Influencing factors.

1 INTRODUCTION

Agriculture has grown significantly in recent years, during which, in parallel with the Sustainable Development Goals (SDGs), it has been striving for clean and sustainable production, taking into account climate change, increased urbanization, and resource scarcity (Massruhá, 2020). Moreover, agricultural activity faces various interferences, directly affecting prices and production; these risk factors include climate change and biological aspects. Therefore, current tools can be useful in minimizing this impact by analyzing these risks (Raucci, 2020). The region in question plays a crucial role in agricultural production, which is closely linked to rainfall patterns. As observed by Carmello and Sant'Anna Neto (2015), despite investments aimed at mitigating the effects of rainfall variability, agricultural production remains heavily dependent on weather conditions. Indeed, both weather and climate account for 60% to 70% of the challenges farmers face in their activities. Thus, Pearson correlation can be used, which is a statistical method employed to measure the direction and degree of the linear relationship between two quantitative variables. The linear relationship between two variables can be defined as the similarity between the numerical distribution of their paired data (Vaz, 2020). In this context, the focus of the research is directed towards the correlation between variables associated with agricultural production, using the Pearson coefficient as a precise quantitative measure of these relationships, to promote a deeper understanding of the dynamics driving agriculture in Pará in 2022.

2 MATERIAL & METHODS

2.1. Study Area Characterization

The study area comprised the main cities that had the highest production in 2022, namely, Cametá, Altamira, and Conceição do Araguaia. These locations were selected due to their different geographic regions within Pará and agricultural production patterns in the state. The crops studied for analysis were soybeans, açaí, cocoa, and cassava. These were chosen for being key crops in Pará's agricultural economy. The data used in this research were collected from national institutes to analyze the relationships between production quantity, monetary value of crops in the studied year, and to calculate the climate's relationship and its impact on the manufacturing of these crops.

2.2. Data Retrieval

For data retrieval in the year 2022, a screening of data from national institutes such as the Brazilian Institute of Geography and Statistics (IBGE), the State Secretariat of Environment and Sustainability (SEMAS), the National Institute of Meteorology (INMET), and the 2022 Climate Yearbook of Pará was used. Integrating these data sources allowed for correlation analysis and assessment of climate influence on the performance of the studied crops in the state.

2.3. Correlation Analysis Using Rstudio

Correlation analysis was conducted using the statistical programming environment Rstudio, employing the R programming language, which was essential for evaluating the relationships between production quantity and value of crops. Therefore,

correlation with state climatic conditions was performed through Pearson correlation calculation and the 'cor()' function used to calculate correlations between variables along with data, typically utilizing Pearson correlation as the default method, which measures the strength and direction of the linear relationship between two continuous variables. Additionally, graphic packages such as 'ggplot2' were used for creating scatter plots and data visualizations, with this specific function defining the initial aesthetics of the plot.

2.3.1. Correlation between Quantity and Value of Production

The 'cor()' function in R will be used to calculate the correlation between the quantity and value of production of crops (Soybeans, Açaí, Cocoa, and Cassava), employing the Pearson correlation coefficient as the default method. This choice will evaluate the strength and direction of the linear relationship between two continuous variables: produced quantity and production value. The relevant variables, in this case, the quantity and value of production of each crop, are specified in the function's argument. Furthermore, the code 'cor(data[, c("Quantity_Soy", "Value_Soy")], method = "pearson")' will be used, generating a number ranging from -1 to 1, where a value close to 1 indicates a strong positive correlation, while a value close to -1 indicates a strong negative correlation, and a value of 0 indicates no linear correlation.

2.3.2. Correlation of Climatic Conditions

To evaluate the correlation between climatic conditions, such as maximum and minimum temperatures, and their possible influences on agricultural production, the same principle of the 'cor()' function is applied. The relevant climatic variables, such as maximum and minimum temperatures, are specified in the function's argument using the code: 'cor(data[, c("Max_Temperature", "Min_Temperature")], method = "pearson")'. Similar to the correlation between quantity and value of production, the 'cor()' function is used to calculate the Pearson correlation coefficient between the selected climatic variables, providing a measure of the linear relationship between maximum and minimum temperatures.

3 RESULTS & DISCUSSION

3.1. Characterization of the study area

The image highlights the main cities studied, distributed in different regions of the state of Pará. It was correlated to the quantities of crops and detailed information about agricultural production in the main municipalities: Altamira, Cametá and Conceição do Araguaia.

Figure 1 - Map of the location of cities and their main production

Source: author, 2024.

Data on agricultural production and municipal statistics were collected, focusing on specific climatic information such as maximum and minimum temperatures. Sources such as INMET and regional climate reports, including the 2022 Pará Climate Yearbook, were explored to ensure accurate data. In 2022, the state produced between 2,574,954 and 145,994 tons, with monetary values ranging between R$7,483,979.00 and R$1,894,303.00 million Brazilian reais, according to IBGE reports. Climatic conditions were obtained directly from researched cities in Pará, such as Cametá, Altamira, and Conceição do Araguaia, recognizing significant variations between these locations.

Table 1: Quantity and values of production in the state of Pará in 2022

Cultura	Quantity (tons)	Production Value (R$)
Soya	2.574.954	R$ 7.483.979,00
Açaí	1.595.455	R$ 5.927.144,00
Cassava	4.157.308	R$ 3.174.826,00
Cocoa	145.994	R$ 1.896.303,00

Source: author, 2024.

Table 1 highlights agricultural production in Pará in 2022, showcasing the significant contribution of crops. Soya leads with 2.57 million tons, generating R$7.48 million. Açaí, with 1.59 million tons, stands out with R$5.92 million. Cassava leads in quantity with 4.16 million tons, generating R$3.17 million. Cocoa, although smaller in quantity, contributes R$1.89 million. These data highlight the diversity and economic relevance of crops in the state.

3.2. Data Search

The results of multivariate analysis provide a comprehensive view of correlation patterns in the quantities of various crops, offering significant insights into agricultural production in Pará municipalities. In 2022, according to IBGE data, several crops stood out both in value and quantity of production. Among the main crops are soya, açaí, cassava, and cocoa. This multivariate approach enhances understanding of the interrelationships between different crops, contributing to a more comprehensive and informed analysis of the agricultural landscape in the state. Additionally, for climatic correlation, records were extracted from the 2022 Pará Climate Yearbook, prepared by the Secretariat of Environment (SEMAS). These climatic information are based on records from the National Institute of Meteorology (INMET), reinforcing the solidity of the climatic data used in this analysis.

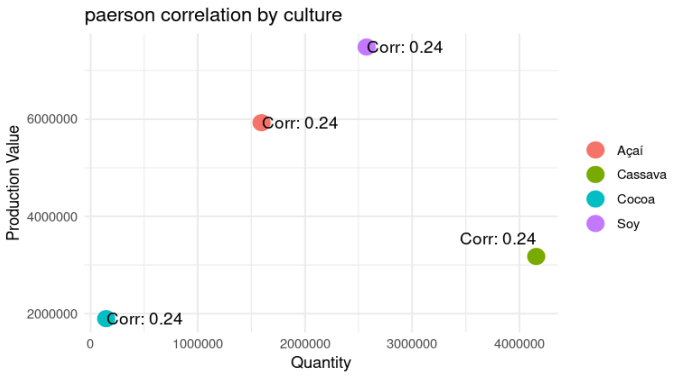
3.3. Correlation analysis using Rstudio language

This analysis allowed a detailed investigation of the relationships between quantity, production value and climatic conditions, providing a more comprehensive understanding of the factors that influenced agricultural production in Pará in 2022. The code “Color” was used to analyze, as shown in Figure 2 .

3.3.1. Correlation between Quantity and Production Value

Pearson's correlation analysis between the quantity and value of agricultural production in the state of Pará in 2022 revealed an overall correlation of 0.24, as shown in figure 3. This correlation coefficient quantifies the strength and direction of the linear relationship between quantities produced and the corresponding monetary values, providing a significant statistical measure of the association between these variables during the analyzed period.

Figure 3 - Pearson correlation between harvest quantities and values



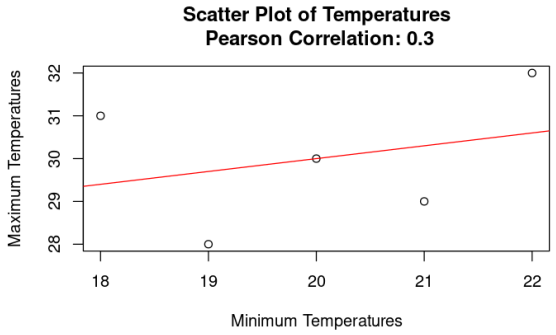
Source: author, 2024.

In 2022, the correlation of 0.24 between the quantity of production and the monetary value in Pará reveals a moderate relationship. This indicates that, in general, the increase in production is not strongly linked to a proportional increase in value. Figure 3, with colored balls representing each crop, highlights açaí (red), cocoa (green), cassava (blue) and soybeans (lilac). These colors symbolize the individual correlations with the overall relationship of 0.24. For example, açaí (red ball) shows a specific correlation of 0.24. This detailed approach offers specific insights into how each crop relates to quantity and value, providing a more refined view of agricultural economic dynamics in Pará in 2022.

3.3.2. Correlation of Climatic Conditions

Exploring the intricate climate records of the state of Pará, we delve into data from the 2022 Climate Yearbook, meticulously prepared by the dedicated team at the Secretariat of the Environment (SEMAS). This precious information is extracted from the detailed records of the National Institute of Meteorology (INMET), providing an in-depth look at the maximum and minimum temperatures of the main agricultural municipalities in the state.

Figure 3: Pearson correlation of maximum and minimum air temperatures between municipalities.



Source: author, 2024.

The correlation made through RStudio between the maximum and minimum temperatures in the main agricultural municipalities of Pará proved to be remarkable, reaching a value of 0.3 This strong negative correlation, obtained from data from the Brazilian Institute of Geography and Statistics (IBGE), Agência Pará and the 2022 Climate Yearbook of Pará from the Secretariat of the Environment (SEMAS) with support from the National Institute of Meteorology (INMET), indicates a significant inverse relationship between maximum and minimum temperatures.

4 CONCLUSION

In this study, we explored the correlation between agricultural crops and climatic variables in the state of Pará in 2022, using Pearson correlation. The results revealed a moderate correlation of 0.24 between the quantity of production and the generated monetary value, suggesting a significant yet nonlinear relationship between these variables. Additionally, we observed a weak positive correlation of 0.3 between maximum and minimum temperatures, indicating a moderate association between these climatic variables. These findings highlight the complexity of interactions within the Pará agricultural sector and underscore the importance of considering multiple factors in the pursuit of sustainability. For the future, these results can guide more effective agricultural policies and practices, promoting sustainable development and resilience in the Pará agricultural sector.

REFERENCES

1 Carmello, V., & Sant'Anna Neto, JL (2015). Variabilidade das chuvas na vertente paranaense da bacia do rio Paranapanema - 1999-2000 a 2009-2010. Raega - O Espaço Geográfico em Análise, 33, 225-247.

2 MASSRUHÁ, Silvia Maria Fonseca Silveira et al. A transformação digital no campo rumo à agricultura sustentável e inteligente. 2020.

3 RAUCCI, Gian Lucca; CAPITANI, Daniel Henrique Dário; SILVEIRA, Rodrigo Lanna Franco da. Derivativos climáticos na agricultura: uma revisão da literatura. Revista de Política Agrícola, v. 3, pág. 83, 2020.

4 Vaz, Émerson Feix e Farret, Félix Alberto. (sem dados). Correlações de Pearson entre o Consumo de Energia Elétrica e os Índices de Desenvolvimento Humano e Econômico.

5 INMET. Instituto Nacional de Meteorologia. Disponível em: https://portal.inmet.gov.br/. Acesso em: Janeiro de 2024.

6 IBGE. Instituto Brasileiro de Geografia e Estatística. Disponível em: https://www.ibge.gov.br/ . Acesso em: novembro de 2023

7 SEMAS. Secretaria de Meio Ambiente do Pará. Disponível em: https://www.semas.pa.gov.br/ . Acesso em: novembro de 2023

8 Anuário Climático do Pará 2022. Disponível em: https://www.semas.pa.gov.br/. Acesso em: Dezembro de 2023

ACKNOWLEDGEMENTS

I would like to thank Amanda Monteiro, bioprocess engineering student at the Federal University of Pará