Financing Agricultural Products: A Statistical Approach in sales data inside of a Brazilian State

C. P. Carvalho

Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Engenharia de Guarátinguetá. Avenida Doutor Ariberto Pereira da Cunha 33, Portal das Colinas, 12516410 - Guarátinguetá, São Paulo – Brazil.

Abstract—The work aims to present statistical data analysis on the sale of financing agricultural products in the regions of Vale do Paraíba, North Coast, and Alto do Tietê in the interior of São Paulo State, Brazil, obtained from agencies located in that regions, to analyze the best-selling products, and to understand the impacts that these regions have on the demand of agricultural producers. The data were collected from agents who participate in the National Rural Credit System through the National Agriculture Strengthening Program (PRONAF). The collected data were organized with the Excel® spreadsheet editor and analyzed using t-test and analysis of variance (ANOVA), to validate the null hypothesis which states that the sales volume was the same independent of the studied region. The results showed variation in sales numbers by region and they were discussed, and recommendations for future works were presented.

Keywords—Agribusiness, Agricultural financing, Agricultural products, Analysis of variance, T-test.

I. INTRODUCTION

Gross Domestic Product (GDP) of Brazilian agribusiness, calculated by the Center for Advanced Studies in Applied Economics (Cepea), esalq/USP, in partnership with the Confederation of Agriculture and Livestock of Brazil (CNA), continued to grow in May, completing five successive months of high in 2020. In the month, the expansion was 0.78%, leading to the growth of 4.62% in the period from January to May 2020 (Center for Advanced Studies in Applied Economics - Cepea, 2020). Agribusiness is the most important sector of the Brazilian economy, about 21.4% of Brazil's gross domestic product (GDP) (Instituto Brasileiro de Geografia e Estatística - IBGE, 2020). Due to the characteristics and diversity of climate and soil, which also contains highly fertile and still unexplored agricultural areas, Brazil presents a favorable scenario for the development and growth of agribusiness. When we analyze the increase in global demographics, therefore the demand for more food, we predict that Brazil should reach the level of the world leader in the supply of food and commodities linked to agribusiness, which will imply the solidification of its economy and catapulted its growth (Bacha, 2004). Rural credit in Brazil was systematized by Law No. 4,829/65 and regulated by Decree No. 58,380/66, being restricted to the specific field of financing rural activities. Rural credit finances the cost of normal expenses of production cycles, investment in goods or services, commercialization, and industrialization. Every year, banks have to allocate 30% of cash deposits, 60% of rural savings deposits, and 35% of ACL funding to apply to rural credit operations. The National Monetary Council (CMN) establishes sub directions for each rural segment according to the producer's profile. It aims to increase the productivity of cash resources in the sector and generate income in family farming. The Central Bank (BC) is the body responsible for managing the National Rural Credit System (SNCR), a set of financial institutions that provide rural financing. The SNCR aims at agricultural development in the country (Banco Central do Brasil - BCB, 2020). The National Program for Strengthening Agriculture (PRONAF) was created in 1996, to strengthen family farming, so that they can have more competitiveness with business agriculture, in addition to direct support to family producers, through credit, PRONAF has another line of action aimed at financing joint actions of municipalities and States, to eliminate bottlenecks that are slowing or preventing rural development in areas where the presence of family farmers predominates (Buainain, 1999). The National Program for Strengthening Family Agriculture (Pronaf) is one of the projects funded by the rural credit system. The initiative finances actions that generate income for family farmers and land-reform. Pronaf resources can be made available individually or collectively. The effective interest rates range from 2.5% per year (for the cultivation of rice, beans, cassava, tomato, onion, English potato, and wheat, among
others) to 5.5% per year (for the acquisition of animals intended for breeding and fattening and other crops and creations) (Banco Central do Brasil - BCB, 2020). Ranathunga, Wijemanna, Sathsara, and Gamage (2018), discussed in their work the status of the agriculture in Sri Lanka in a concise manner using the published official data and concluded that the authorities must attempt to achieve the self-sufficiency of each sub-sector to save the foreign exchange on continuous imports. The inclusiveness of the financial services of microfinance banks for inclusive Agricultural development in Anambra State, Nigeria, were analyzed, and data were elicited from three microfinance banks on the number of genders, location, and firms that had access to their financial services by using frequency, tables, mean, and standard deviation, coefficient of variation, t-test and bar chart and showed that there are significant imbalances in the financial services rendered by microfinance banks to gender, location, and firm (Odemero & Obianuju, 2016). Similar work was presented by Sroka, Dudek, Wojewodzic, and Król (2019), in which they described spatial variation in generational changes in farms located around large cities (metropolises) in Poland and to assess the factors affecting the scale of such changes, emphasizing the importance of the location of farms relative to large cities.

Another approach in the same research scope had questioned the relationship society's credit structures, describing that the growth of commodity-money relations, the need for cheap low-interest credit to dealers, lawlessness, the need for combining funds for bulk purchases and sales have increased credit cooperation development. As the result, it proposed that the analysis of legal support, the institutionalization of the system of public finance in Russia, has been carried out, as in a special role in crediting innovations in the field of agricultural production and modernization of peasant farms was played by the Moscow People's Bank (Galas & Seleznev, 2019). Durie (2018), had examined financing rural industrialization and employment creation practices and possibilities in Ethiopia. However, attempts were made to see at least the trends in agricultural commercialization, off-farm practices, the government's policy, the financial institution's practices, and above all how other countries approached rural industrialization and financing such industries. Using the descriptive statistics, the correlation analysis and the regression analysis analyses had conducted to understand how far the capital structure affects the profitability of agricultural holdings in the Czech Republic, considering a period of six years from 2008 to 2013 (Stekla & Grycova, 2016).

1.1 Statistics tools

Statistical methods are efficient direct approaches that provide objectivity and accuracy and provide tools that formalize and standardize procedures to obtain certain conclusions. The results of statistical analyses are the result of market intelligence services that involve the collection and detailed analysis of business data. There are some different methods and tests within the statistical analysis, the method used should be chosen according to the objective of the project. Variance analysis, also known as ANOVA, is an Inferential Statistics tool used to compare the distribution of three or more groups in independent samples and was the first tool used for our case study. This analysis is also a way of summarizing a linear regression model, by comparison, the sum of squares for each source of variation and, through the use of the F test, test the hypothesis that any source of variation in the model is equal to zero, the null hypothesis in this case. The conclusion of ANOVA is also based on Statistics F, which has Fischer-Snedecor F distribution with (k-1) and (n-k) degrees of freedom, where k is the number of groups and n is the number of observations. In case the F statistic is higher than the f tabled, it is concluded that the result is a much lower value than the calculated one since there is a very large difference between the values of the samples. In the present study, ANOVA was used to compare the sales per month of the agricultural company studied and analyze the influence of different locations on product sales (Garcia-Marques & Azevedo, 1995). ANOVA was also applied in a similar study addressing the development, dissemination, and assessment of a Food Safety System Management (FSSM) curriculum offered to college-aged, agribusiness students in Yerevan, Armenia (Pokharel, Marcy, Neilan & Cutter, 2017). Remenova and Jankelova (2019) in their work monitors the dependence between the decision-making style of agricultural managers and their personal and working parameters to identify the decision-making styles and explain how it can support the agricultural managers gain, and for this issue, a parametric test ANOVA was applied to assess potential differences in the score of decision-making. The parametric test ANOVA had also applied for the prediction of financial distress (default of payment or insolvency) of 250 agriculture business companies in the EU from which 62 companies defaulted in 2014 concerning lag of the user attributes (Klepac & Hampel, 2017). To make a more specific analysis, we chose to use the Student t-distribution comparing means two to two and for the study of the sale of a product for each region. Student's t-test is typically used when the test statistic follows a normal distribution, but the variance of the population is unknown. In this case,
the sample variance is used and, with this adjustment, the test statistic follows a t distribution of Student (Azevedo, 2016). Considering the t-test applied to agribusiness Dlamini and Huang (2020), conducted a study that was considering a gender-based comparative assessment of training needs for beef cattle farmers. Primary data were collected through personal interviews, guided by a reliability-tested questionnaire, from a sample of 397 farmers. The Borich Needs Assessment Model was adopted for data analysis and inferential statistics were employed to evaluate statistically significant differences between the gender groups. This research aims to study the sales of five agricultural products in four regions of the state of São Paulo, Brazil to identify the factors of greatest influence on the numbers of products sold and to facilitate the distribution of sellers in each region at the times of the year studied, reducing costs and increasing the company's profit. For this, it was based on Descriptive Statistics and used two statistical tools, ANOVA and t-test. In the sequence of this work, the Material and Methods that were used are described, detailing the area in São Paulo State, where this research was conducted for the researcher, the period of data acquisition, and how it was treated. Then, the results are presented and discussed. Finalizing the work, the conclusion is presented followed by the recommendations of the author to further researches.

II. MATERIAL AND METHODS

The study was conducted in four regions located inside the São State, Brazil. The four regions consolidated 39 cities, which represents a GDP of more than R$ 60billions and congregated 3 million people. The data was collected by the researcher during one year, by the financing agency companies spread into the four regions, taking into consideration the yearly volume of sales. Sales are related to the number of financing contracts for agricultural products by farmers established in the four regions. The financing contract accounted for were regarded as the financing for five different agricultural products: greenhouses, cattle, pickup truck, tractor, and a photovoltaic panel. The statistical investigation took place in several stages and involves specific and particular aspects reported for each agency. The first step consisted in the formulation of a survey to collect the data at agencies. At this point, it was necessary to consider whether or not the questions are appropriate and have a statistical nature, that is, involve or not variability in the data, classifying the data. The second step evolved was to carry on the data acquisition, until complete one year of information collection, when the appropriate plan for data acquisition had to be designed. Considering the next step, the researcher conducted the data analysis, starting with the choice of the most appropriate representation taking into account the nature of the data and the purposes in view, using Excel®. Finally, the interpretation of the results took account of the proposed question. Then, discussions regarded data was presented, focusing to emphasize what the statistical analyses were confirming or validating (Martins, 2010, p. 09). The analyses as described in the previous topics were made by the statistical ANOVA and t-test tools.

III. RESULTS AND DISCUSSION

Before the presentation of the results, a brief explanation is presented presenting the cities that compound each region. Then, the total sales unities per region were presented considering the main agricultural products requested by them, classified as greenhouses, pickup trucks, tractors, and photovoltaic panels. In the sequence, the total amount of sales per region were presented in Brazilian currency. The statistical tools ANOVA and t-test were applied to validate the null hypothesis considering the period comparing between two months the per studied region. For better classification of the data, the study was shared into four regions where each encompasses the following cities: Region 1: Caçapava, Lagoinha, Natividade da Serra, Pindamonhangaba, Redenção da Serra, São Luís do Paraítinga, Taubaté, Tremembé. Region 2: Arujá, Biritiba-Mirim, Guararema, Mogi das Cruzes, Santa Branca, Salesópolis, Santa Isabel, Suzano. Region 3: Jacareí, Jambéiro, Monteiro Lobato, Paraibuna, São José dos Campos, São Francisco Xavier, Santo Antônio do Pinhal, São Bento do Sapucaí. Region 4: Aparecida, Areias, Bananal, Cacheoira Paulista, Cruzeiro, Guaratinguetá, Lavrinhas, Lorena, Queluz, Piquete, Roseira, Silveiras, São José do Barreiro. Through the graphs generated in Excel®, a macro view of sales by product and region were illustrated in Fig. 1, 2, 3, 4, 5, 6, and 7 were obtained. In the sequence, statistical analysis was conducted, firstly per period, then per region, using ANOVA and t-test, always discussing the results.
From Fig. 1 it is clear that Region 2 leads the sales for greenhouses, followed by Region1.

In terms of cattle sales, Region 3 presented the biggest amount, followed by Region 1 and Region 4 with a very closed sales number.

Region 3 presented the highest volume of pickup truck sales, and Region 1 appeared in the second position.

In the consolidated numbers, Region 3 is the leader of sold products, followed by Region 1.
As can be observed from the Fig. 7, the leader is Region 1, in terms of total revenues per the period under analysis, since its products had more added value.

3.1 Analysis by period

From the collected data, ANOVA was used to compare sales per month. Table 1 presents the results that originated from the calculations performed.

**Table 1. ANOVA per period.**

| Source of variation | SS     | DF  | MS   | F      |
|---------------------|--------|-----|------|--------|
| Between groups      | 934.16 | 19  | 49,166.3157 | 3.7177 |
| Residual error      | 1058   | 80  | 13,225       |        |
| Total               | 1992.16| 99  |      |        |
| Table value (5%)    |        |     | 1.72          |        |

The calculated F value was 3.7177 and comparing with the table value for a significance level of 5%, it was possible to conclude that the null hypothesis should be rejected since the table value is lower. As the values of means are different, it is necessary to perform a more specific analysis to find out which means are different from each other. Therefore, the t-test analysis was conducted to compare the means two by two. Table 2 illustrates the results of this analysis.

**Table 2. t-test per period.**

| Medium variance       |  |  | t0  |  |
|-----------------------| | |    |  |
| January-February      | 3.18 | | 1.75 |  |
| January-March         | 4.24 | | 1.55 |  |
| January-April         | 2.71 | | 1.78 |  |

When comparing the results obtained with the table value (1.6886), also for a significance level of 5%, it is possible to notice in which comparisons the null hypothesis was rejected and in which it was accepted, based on the colors of the legend of Table 3.

**Table 3. Hypothesis identification – analysis per period.**

| Caption |  |
|---------|  |
| Reject $H_0$ |  |
| Accept $H_0$ |  |

The economic scenario is a factor that can explain this difference. There may be an economic crisis or an increase in the price of some component of the analyzed products (such as the agricultural film of the greenhouses), increasing the price of the product in question, and a possible drop in its sales. Geography is another point that can influence product sales. The climate changes in the analyzed period may change sales.

3.2 Analysis by region

The ANOVA was performed by region to analyze the influence of different locations on product sales. Tables 4, 5, 6 and 7 show the results of the calculations.

**Table 4. ANOVA – region 1.**

| Source of variation | SS     | DF  | MS   | F      |
|---------------------|--------|-----|------|--------|
| Between groups      | 470.24 | 4   | 117.56 | 5.8026 |
| Residual error      | 479.20 | 20  | 23.96 |        |
| Total               | 949.44 | 24  |      |        |
| Table value (5%)    |        |  | 2.87  |        |
Table 5. ANOVA – region 2.

| Source of variation | SS    | DF | MS    | F      |
|---------------------|-------|----|-------|--------|
| Between groups      | 214,80| 4  | 53,70 | 5,1536 |
| Residual error      | 208,40| 20 | 10,42 |        |
| Total               | 423,20| 24 |       |        |
| Table value (5%)    |       |    | 2,87  |        |

Table 6. ANOVA – region 3.

| Source of variation | SS    | DF | MS    | F      |
|---------------------|-------|----|-------|--------|
| Between groups      | 295,44| 4  | 73,86 | 5,1723 |
| Residual error      | 285,60| 20 | 14,28 |        |
| Total               | 581,04| 24 |       |        |
| Table value (5%)    |       |    | 2,87  |        |

Table 7. ANOVA – region 4.

| Source of variation | SS    | DF | MS    | F      |
|---------------------|-------|----|-------|--------|
| Between groups      | 187,84| 4  | 46,96 | 4,2459 |
| Residual error      | 221,20| 20 | 11,06 |        |
| Total               | 409,04| 24 |       |        |
| Table value (5%)    |       |    | 2,87  |        |

When comparing the table F value, to a significance level of 5%, with the calculated value, it was concluded that the null hypothesis was rejected, and a t-test analysis was necessary to obtain more detailed and accurate results. The colors of $t_0$ show whether the $H_0$ hypothesis was rejected or accepted as shown in Table 8. These conclusions come from the comparison of the calculated values with the tabulated value (1.64) to a significance level of 5%. For values above the tabled, the null hypothesis is rejected, for values below, it is accepted.

Table 8. Hypothesis identification – analysis per region.

|   | Reject $H_0$ | Accept $H_0$ |
|---|--------------|--------------|

Tables 9, 10, 11, and 12 present data from calculations performed separated by region:

Table 9. t-test for region 1.

| Region 1 | Average Variance | $t_0$ |
|----------|------------------|-------|
| Pick-up trucks and tractors | 1,30 | 2,50 |
| Pickup trucks and greenhouse | 1,80 | 2,78 |
| Pickup trucks and cattle | 3,25 | 2,46 |
| Pickup trucks and photovoltaic panels | 1,50 | 2,84 |
| Tractors and greenhouse | 0,80 | 0,35 |
| Tractors and cattle | 2,25 | 4,85 |
| Tractors and photovoltaic panels | 0,50 | 0,89 |
| Greenhouse and cattle | 2,75 | 4,58 |
| Greenhouse and photovoltaic panels | 1,00 | 0,32 |
| Cattle and photovoltaic panels | 2,45 | 5,05 |

Table 10. t-test for region 2.

| Region 2 | Average Variance | $t_0$ |
|----------|------------------|-------|
| Pick-up trucks and tractors | 1,95 | 1,13 |
| Pickup trucks and greenhouse | 1,95 | 2,26 |
| Pickup trucks and cattle | 1,85 | 1,39 |
| Pickup trucks and photovoltaic panels | 2,45 | 0,61 |
| Tractors and greenhouse | 0,70 | 3,02 |
| Tractors and cattle | 0,60 | 0,41 |
| Tractors and photovoltaic panels | 1,20 | 2,31 |
| Greenhouse and cattle | 0,60 | 3,67 |

Caption

Reject $H_0$
Accept $H_0$
In the table of region 1, it is noted that in 7 of the 10 comparisons, there is a rejection of the null hypothesis. In region 2, 5 are rejected. In region 3, 6 are rejected. And in region 4, 7 of the 10 null hypotheses are rejected. Also, it can be observed that the means differ for comparisons of different products in each region. Thus, it is concluded that the factors of each location influence the sale in the regions. Among them, there is the most accomplished agricultural activity in each location, the level of economic development of cities, climate, and relief, which are directly related to the use of photovoltaics and greenhouses, for example. After analyzing the results, it was concluded that sales show a large discrepancy between regions due to the difference between the goods and the needs of each region. The place that stood out most in sales was region 1 and the one with the lowest sales value was region 4, as shown in the total revenue in Fig. 7. There is also little variation in product sales over the months analyzed, which may indicate a slight growth in revenue in the region. Region 2 presented the highest sale of greenhouses and photovoltaic panels, representing greater sustainable development among the four localities, while region 1 showed higher cattle sales, pickup trucks, and tractors, indicating greater development of agriculture and livestock among the regions. Thus, it can be said that the region that presents the most market for the analyzed products is region 1. In the Paraíba Valley region, the largest sale was of tractors and cattle, due to the focus on agriculture, so the sale of products related exclusively to this area appears to be the best investment for the site and possibly generating a higher revenue than other products. In the Alto Tietê region, the most sold product in greenhouses, due to flower and vegetable producers, and the sale of photovoltaic panels, which is associated with largely Japanese producers, who have greater access to knowledge of technology for longer than producers in other regions. Pickup truck sales are relatively high in all regions, as it is a very useful tool in the daily life of any rural producer, regardless of specialization.
IV. CONCLUSION

The work had as a proposal to present statistical data analysis on the sale of financing agricultural products in the regions in the São Paulo State, using these tools in order to analyze the best-selling products, and validate the null hypothesis which states that the sales figures were the same among the regions involved in the research. As presented by the results and discussions made, the null hypothesis was rejected. This paper presented a first approach and a guide to help the understanding how the financing programs have impacted the sales of the agricultural products in the regions in the São Paulo State, and it can give a direction to explore more data in other regions in the same state or another. Limitations can be observed in this work as a short period collecting data, the number of regions, and the economic profile of each one. The t-test and ANOVA, could be improved by data analytics using algorithms.

V. RECOMMENDATIONS

Based on the results of the research, some recommendations were proposed: a. Amplification of the research for more regions in the São Paulo state. b. Comparison of how the PRONAF program impacts the sales in other Brazilian states. c. Conduct multi-criteria methods as Analytic Hierarchy Process (AHP), in other to understand the choice motivation from the farmers for agricultural products. d. Applied algorithm for data analytics to permit the research takes to account the whole parameters that cause influence on the product sales.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the University of the State of São Paulo, Faculty of Engineering of Guaratinguetá, for its assistance during the research period.

REFERENCES

[1] Azevedo, P. R. M. (2016). Introdução à estatística (3 ed.). https://repositorio.ufrn.br/pspi/html/123456789/21298
[2] Bacha, C. J. C. (2004). Economia e Política Agrícola no Brasil. São Paulo: Atlas, 2004.
[3] Banco Central do Brasil - BCB. (2020, August). Relatório Anual. https://www.bcb.gov.br/estabilidadefinanceira/credito_rural
[4] Buainain, A. M. (1999). Trajetória Recente da política Agrícola Brasileira [Tese de Doutorado], http://www.repositorio.unicamp.br/handle/REPOSIP/285712
[5] Center for Advanced Studies in Applied Economics – Cepea - esalq/USP. PIB. (2020, Augst 21).

https://www.cepea.esalq.usp.br/upload/kceditor/files/Cepea_PIB_CNA_MAIO_Agosto2020(1).pdf
[6] Dlamini, S. I., & Huang, W. -C. (2020). Towards Intensive Co-operated Agribusiness: A Gender-Based Comparative Borich Needs Assessment Model Analysis of Beef Cattle Farmers in Eswatini. Agriculture, 10(96), 1-17, https://doi.org/10.3390/agriculture10040096
[7] Durie, A. D. (2018). Financing Rural Industrialization and Employment Creation: The Case of Ethiopia. Independent Journal Of Management & Production (IJM&P), 9(4), 1317-1337, https://doi: 10.14807/ijmp.v9i4.803
[8] Galas, M.L., & Seleznev, P. S. (2019). Public Organizations of Russian Solving Socio-Economic Problems of the early of XX century. Bylye Gody, 52(9) https://doi: 10.13187/bg.2019.2.766
[9] García-Marques, T., & Azevedo, M. (1995). A inferência estatística múltipla e o problema da inflação do nível de alfa: a ANOVA como exemplo. Psicologia, 10(1/2), 195-220, https://doi.org/10.17575/psicpol.v10i1/2.655
[10] Instituto Brasileiro de Geografia e Estatística - IBGE. (2020, March). Agência de Notícias IBGE, https://agenciadenoticias.ibge.gov.br/agenciannoticias.pdf.ace sso 04/03/2020-march3rd2020
[11] Klepac, V., & Hampel, D. (2017). Predicting financial distress of agriculture companies in EU. Agricultural Economics – Czech, 63(8), 347–355, https://doi:10.17221/374/2015-AGRICECON
[12] Martins, E. (2010). Contabilidade de custos (10 ed.). São Paulo, SP: Atlas.
[13] Odemero, A. F., & Obianujunwa, E. P. (2016). Inclusive Financial Services of Microfinance Banks for Inclusive Agricultural Development in Anambra State, Nigeria. International Journal of Environment, Agriculture and Biotechnology (ISSN: 2456-1878). 1 (2), 149-154.
[14] Pokharel, S., Marcy, J. E., Neilan, A. M., & Cutter, C. N. (2017). Development, Dissemination, and Assessment of a Food Safety Systems Management Curriculum for Agribusiness Students in Armenia. Journal of Food Science Education, 16, 107-117, https://doi:10.1111/1541-4329.12122
[15] Ranathunga, L.N., Wijemanna, W. M. D. I. S., Sathsara, M. G. S., & Gamage, R.G.B.K. (2018). Agriculture in Sri Lanka: The Current Snapshot. International Journal of Environment, Agriculture and Biotechnology, 3(1), 118-125, http://dx.doi.org/10.22161/ijeab/3.1.15
[16] Remenova, K., & Jankelova, N. (2019). Decision-making style of agribusiness managers. Agricultural Economics – Czech, 65(7), 322–330, https://doi.org/10.17221/289/2018-AGRICECON
[17] Sroka, W., Dudek, M., Wojewodicz, T. & Król, K. (2019). Generational Changes in Agriculture: The Influence of Farm Characteristics and Socio-Economic Factors. Agriculture, 9, 264, https://doi.org/10.3390/agriculture9120264
[18] Stekla, J., & Grycova, M. (2016). Agricultural Economics – Czech, 62(9), 421–428, https://doi: 10.17221/231/2015-AGRICECON

ISSN: 2456-1878
https://dx.doi.org/10.22161/ijeab.55.1

1203