MATHEUS FERNANDES BARBOSA

DATA-DRIVEN DECISION-MAKING IN THE BRAZILIAN FLOWER INDUSTRY

São Paulo
2021

MATHEUS FERNANDES BARBOSA

DATA-DRIVEN DECISION-MAKING IN THE BRAZILIAN FLOWER INDUSTRY

Thesis presented to Polytechnic School of the University of São Paulo to obtain the Bachelor's Degree in Mechatronics Engineering.

São Paulo
2021

MATHEUS FERNANDES BARBOSA

DATA-DRIVEN DECISION-MAKING IN THE BRAZILIAN FLOWER INDUSTRY

Thesis presented to Polytechnic School of the University of São Paulo to obtain the Bachelor's Degree in Mechatronics Engineering.

Academic Supervisor:

Marcos de Sales Guerra Tsuzuki

São Paulo
2021

Cataloging-in-Publication

Fernandes Barbosa, Matheus
Data-Driven Decision Making in the Brazilian Flower Industry / M. Fernandes Barbosa -- Sao Paulo, 2021.
87 p.

Trabalho de Formatura - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos.

1.Flower Industry 2.Digitalization 3.Price Forecast 4.Software Development I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos II.t.

ACKNOWLEDGMENTS

This work closes the chapter in my life of the period that I studied in Sao Paulo. It was six long years with intensity, emotion, and learning. Although the pandemic affected it somehow for the final years, I prefer to focus on the good side of things. That being said, let's focus where matters:

First of all, I need to say thank you to my family, my father Luiz Barbosa and my mother Ilma Fernandes. I know all the effort that you both put into my development and I do am grateful for this. Your support is really important for any achievement that I reach.

I would like also to say thank you to my girlfriend, Lyna Almeida. Thank you for being together and always encouraging me with your words. You make me better.

Thank you also to Professor Marcos Tsuzuki, that accepted me for this thesis on a topic of research that I felt I should focus on, always motivated me weekly, and made important comments for my paper.

Finally, I need to say thanks to my sister, Bruna Barbosa, and also to all my colleagues from the Polytechnic School that made my experience in Sao Paulo amazing. I know that it is the end of a chapter, but the memories I will keep with me for a long time.

"If you're not making someone else's life better, then you're wasting your time."

-- Will Smith

ABSTRACT

This project proposes to develop a platform including an algorithm for price forecast to support the most critical decisions in the daily routines of the Brazilian flower growers'. As a secondary objective, there is the development of a mobile application to structure the data collection between the final consumer and the grower, in a way to connect the last to the first part of the value chain. The work was defined in two main phases, the first was the qualitative research focused on literature review, event participation, and interviewing with experts to define the problem, understand the main decisions that must be made, synthesize and ideate on it to define the project requirements. The second phase was based on the input collected before defining which initiative should be prioritized for the development and, later, testing the prototypes with growers and consumers. The results for the price forecast shows that for data of 2019 (using 2017 and 2018 to train) the Random Forecast model achieved the best result, 5% better than the baseline, but for 2020 and 2021 the Naive Forecast presented the best marks, concluding that a daily price forecast did not present outstanding results using the data from only one grower, given the other irregularities of the market and the comparison with the baseline. On the other hand, the prototype for the web system can create value for the growers, since the majority of the survey's answers agreed that the system development was better than the one currently in place. For the secondary objective of the mobile application, it was possible to confirm with the survey that users are willing to use new applications and that there is space to implement such a tool to gather more data and, in the future, develop new solutions and products for this market.

Keywords – Flower Industry, Digitalization, Price Forecast, Software Development

RESUMO

Este projeto se propõe a desenvolver uma plataforma com um algoritmo de previsão de preços para apoiar as decisões mais críticas no dia a dia dos produtores de flores do Brasil. Como objetivo secundário, existe o desenvolvimento de um aplicativo para celular para estruturar a coleta de dados entre o consumidor final e o produtor, de forma a conectar a última à primeira parte da cadeia de valor. O trabalho foi definido em duas fases principais, a primeira composta por uma pesquisa qualitativa com a revisão bibliográfica, participação em eventos e entrevistas com especialistas para definir o problema, entender as principais decisões que devem ser tomadas, sintetizar e definir os requisitos de projeto. A segunda fase baseou-se nos insumos coletados anteriormente para definir as iniciativas priorizadas no desenvolvimento e, na sequência, foram realizados testes de protótipos com produtores e consumidores. Os resultados mostram que para os dados de 2019 (usando 2017 e 2018 para treino) o modelo utilizando Random Forest alcançou o melhor resultado, 5% melhor que o baseline, mas para 2020 e 2021 o Naive Forecast apresentou os melhores resultados, concluindo que a previsão diária utilizando dados apenas de um produtor não apresenta resultados relevantes, dadas as irregularidades do mercado e a comparação com o baseline. Por outro lado, o protótipo do sistema web pode criar valor para os produtores, uma vez que a maioria dos respondentes da pesquisa concordaram que o protótipo é melhor do que o que o atual sistema. Para o objetivo secundário do aplicativo móvel, foi possível constatar que os usuários estão dispostos a utilizar novos aplicativos e que existe espaço para implementar tal ferramenta para recolher mais dados e, futuramente, desenvolver novas soluções e produtos para este mercado.

Palavras-Chave – Indústria de Flores, Digitalização, Previsão de Preços, Desenvolvimento de Software

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

"Thriving as a mainstream company today means being data-driven, but cultural challenges — not technological ones — represent the biggest impediment around data initiatives" [1]. The amount of data generated by different businesses continues to grow, mainly pushed by technology companies, but even the more traditional ones can collect benefits from a more data-driven approach.

1.1 Overview

Today, the advance of technology has enabled different companies to manage their data assets without the need for a heavy cost structure, and this "increased accessibility of data has enabled a different way of making decisions that involve more empirical evidence rather than personal experience, intuition, or belief" [2]. In this sense, the decision-making process may change, but only if the digital transformation is carried out with a cultural change of the organization.

Many enterprise information technologies (IT) are being used to capture a large amount of data in their regular activities. "Increasingly, these systems are imbued with analytical capabilities, and these capabilities are further extended by Business Intelligence (BI) systems that enable a broader array of data analytic tools" [3] and this can be applied to operational data. These digital assets are valuable and "Many organizations are hungry to use data to grow and improve performance - and multiple players see market opportunities in this explosion of demand" [4]. Usually, there are typically many steps between raw data and actual usage, and there are openings to add values at various points along the way with advanced analytics.

This work proposes to look at how to leverage the power of data in the Brazilian flower industry, using the most advanced tools available and benchmarking with other industries and countries that recently passed through this digital transformation. "The true power of data is unlocked when organizations move their thinking beyond score-keeping and toward a forward-looking mindset" [5].

1.2 Motivation

Despite its size, many of the managerial decisions in this sector still rely on a leader's "gut instinct" and a few financial metrics to support its decision-making. Which variety to plant, which price to trade, when to sell the product are some questions that need to be answered routinely. "Just like yesterday's newspaper, there is little demand for the last week's fresh flowers" [6].

In 2017, Royal FloraHolland, a Dutch organization leader in the commercialization of flowers, started a process to move all direct transactions to digital to increase the efficiency of both supply and demand and comply with the financial laws in Europe. Their work is still in progress with the project called Floriday, which was a large step to increase the inflow of data from its operations and create the building blocks for their data platforms to support decision-making in the flower industry.

On the other hand, the biggest flower cooperative in Brazil, since its foundation, has been known as an implementer, not a developer, of information technology. However, in the last couple of years, they began to gain international recognition as an important IT developer in this industry, especially in proposing additional tools for the original model [7]. This work aims to design and implement a personalized and local tool for the sector and improve the visibility of Brazilian flower research in the global market.

About 97% of the production of flowers in Brazil are sold in the internal market [8]. Once the country's production is not integrated with the commercial flow of the international market, most of local growers and buyers have access only to national support. In this sense, internal research and development have the space to create its solutions for this market, based on the knowledge of other countries and industries.

1.3 Objective

"One of the main characteristics of buyers of flowers and ornamental plants in Brazil is to concentrate their demands strongly on a few specific dates throughout the year" [9], so advanced planning is critically important. Once the decision-making in the Brazilian flower industry is heavily based on the intuition of a few local experts, it can become a problem for the perpetuity and the standardization of strategic decisions. This thesis proposes to leverage the power of data to support decision-making in this industry looking from the growers' perspective.

The primary objective is to develop a platform including an algorithm to support the most critical decisions in the growers' daily routines. "Organizations can benefit by focusing their data initiatives on clearly identified high-impact business problems or use cases" [1], so this work proposes to start looking for organizations to get its initiatives and identify the most promising opportunities to pursue. The web platform will be restricted to growers associated with the cooperative Veiling Holambra, due to the data available for research (that was obtained with one grower), but it was developed using a modular structure that can be adapted to other data structures.

It is expected that the platform can process large amounts of data and the system is protected with high standards of data protection. As a secondary objective, there is the development of a mobile system to structure the data collection between the final consumer and the grower, in a way to connect the last to the first part of the flower value chain. It is not planned, however, to build a model to analyze the data that will be collected by this new system.

2 STATE OF THE ART

The floriculture supply chain is large. To get to the final customer, there are three main players in the production chain: the growers, the wholesalers, and the retailers. Figure 1 illustrates the entire production flow of the flower industry.

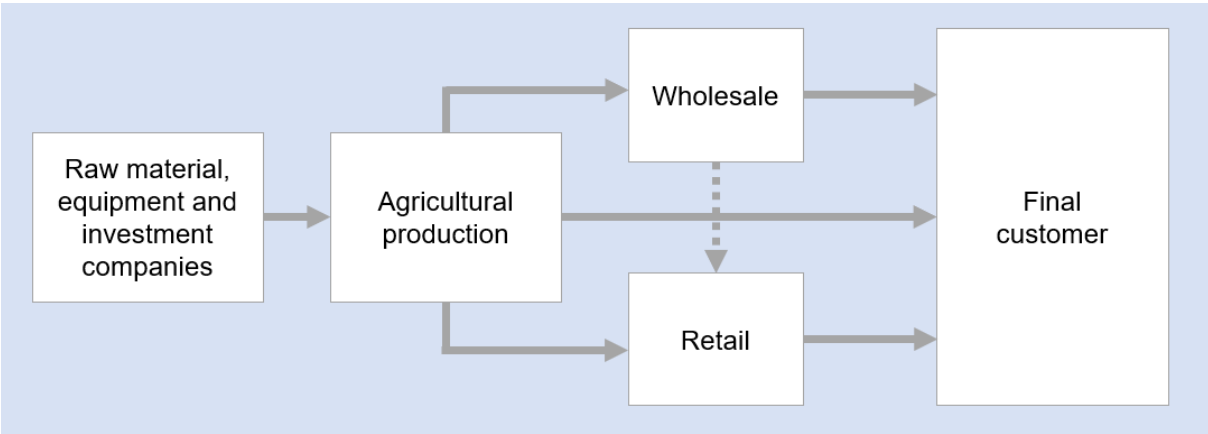

Figure 1: Production chain of flowers and ornamental plants in Brazil. Adapted from [8].

Once the flower is ready to be sold, there are many ways for the grower to do it, depending on their structure and the relation with the market. Figure 2 displays the main formats of sales offered today by growers and for the second commercialization, by different wholesalers and retailers.

When humans make decisions, the process is often muddy, biased, or limited by the inability to process information overload. Data and analytics can change all that by bringing in more data points from new sources, breaking down information asymmetries, and adding automated algorithms to make the process instantaneous [4]. This research presents a broad view which are the main strategic decisions that are done by the growers' side. The conclusions came through both an intellectual and qualitative approach.

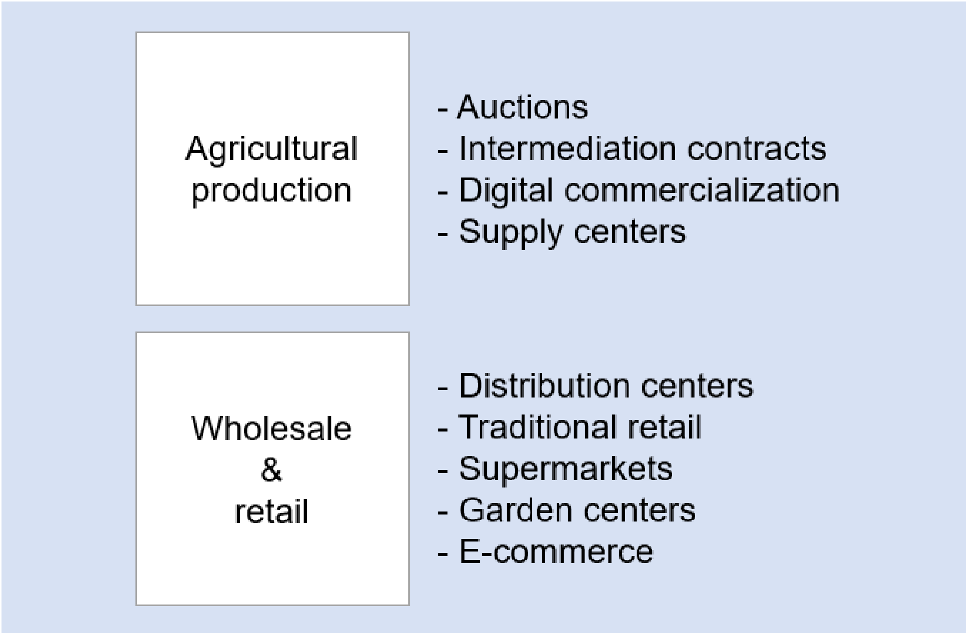

Figure 2: Commercialization flow in the internal market. Adapted from [10].

2.1 Product Chain

There are different paths in which the flower can go through the farm to the final consumer. Different players act in this process, according to the product, the format of the sale, and the technology involved. The reason for an immersion in the process is to evaluate in which situation a data-driven support mechanism can create more value for the grower and understand the whole scope of the process.

2.1.1 Raw material, equipment, and investment companies

In 2014, the raw material, equipment, and investment companies, also known as "before the farms" were responsible for 13% of the total financial movements of the flower industry, equivalent to R$ 1.3 billion (US$0.2 billion) [8]. This can be divided into two parts: companies that provide raw materials (and other operational expenditures - OPEX) and companies focused on equipment and investment (defined as capital expenditures - CAPEX). Table 1 presents information about each of the subsectors.

2.1.2 Agricultural Production

Agricultural production can be divided into three main types of groups, as defined in Table 2. Together, they made R$ 2.1 billion (US$ 0.4 billion) in revenue in 2014 [8].

Type	Sector	Revenues
OPEX 66.3%	Seedlings, seeds and bulbs	19.3%
	Substrates	13.3%
	Vases	11.8%
	Packaging	7.7%
	Fertilizers	6.4%
	Energy supply	2.5%
	Pesticides	2.4%
	Heating fuel	1.9%
	Water supply	0.4%
	Individual protection equipment	0.3%
	Biological control	0.3%
	Pruning and harvesting tools	0.3%
CAPEX 33.7%	Greenhouse	11.4%
	Plastics and agro-textiles	9.1%
	Irrigation and fertigation structures	6.3%
	Tables, railings	4.0%
	Air conditioning structure	1.2%
	Filling machines	1.0%
	Lighting equipment	0.5%
	Trays and boxes	0.1%

Table 1: "Before the farms" revenue structure. Adapted from [8].

Type	Revenues
Growers in cooperatives	35%
Growers in supply and distribution centers	25%
Independent growers	40%

Table 2: "At the farm" revenue structure. Extracted from [11].

Growers in cooperatives are those that use an organizational structure to commercialize their products. This organization does not have the product but serves as the interface between growers, wholesalers, retailers, and final consumers. According to the managers of the biggest flower cooperative in Brazil, the company's main role is to facilitate the interaction between the producers and buyers, integrating the production with the market [12].

Growers in supply and distribution centers sell their products in specialized centers, or not, in flowers to the final consumer. Finally, independent growers are those that sell their products directly to the final consumer, or other growers.

2.1.3 Wholesalers

The wholesalers - also known as cash and carry - represent the main intermediaries in the production flow of the flower industry. They are divided into three main categories, as presented in Table 3. They are in the middle of the growers and the final retailers, described in more detail in the next section. The wholesalers presented are not the only ones, but concentrate on the main financial movements of the sector. The difference between line and central wholesalers is the operation format: line wholesalers sell their production for different clients in a predefined path, while the central ones usually have a main hub where clients go to get its products. This segment movements R$ 2.1 billion (US$ 0.4 billion) in 2014 [8].

Type	Revenues
Central wholesalers	3.6%
Line wholesalers	70.0%
Garden centers	26.4%

Table 3: "After the farm: wholesale" revenue structure. Adapted from [8].

For some regions outside the Sao Paulo center of commercialization, the routes developed by the wholesalers had a dual function, while they have sold flowers to retailers, in return for their property, they bought flowers from the regional producers, thus, besides reducing the buying and selling costs, they enlarged the line of commercialization products [13].

2.1.4 Retailers

Finally, the retailers are the last point before the product passes to the final consumer. In 2014, the financial transactions in this part of the process were around R$ 4.4 billion (US$ 0.8 billion) [14]. The way it was distributed is shown in Table 4.

Type	Revenues
Floricultures	22.6%
Supermarkets	8.8%
Decorators	53.7%
Landscapers	14.9%

Table 4: "After the farm: retail" revenue structure. Adapted from [8].

The supermarkets have been gaining market share since the beginning of the millenium, where the participation of this segment is increasing, from an inexpressive participation in 2003 to almost 10% in 2014. This behavior of the consumer is explained by the increase in the average income of the population and, mainly, by the more accessible prices practiced by the supermarkets.

It is important to define the decorators and landscapers, once both segments represent over sixty percent of the total retail sales of the sector. Moreover, [8] used the assumption to build Table 4 that 40% of the revenue from the flower shops are for the activities of decorators, so the 53.7% of participation of decorators, do not include part of the revenue of the flower shops, due to the 'service as a product' part of its context. These companies do not only sell flowers, there is a differential and value-added by the service they provide to their clients.

> "Landscape: it is responsible for (i) the development of the landscape project, (ii) the purchase of flowers and ornamental plants, (iii) execution, and (iv) maintenance of the project within a scheduled period. Its service is widely used in commercial sectors and condominiums, as well as in homes and businesses. Given the fact that the project does not have a daily effect, there is constant maintenance of flowers and ornamental plants, with replacement whenever necessary."[8]
>
> "Decorator: characterized by the signature of unique, fast projects that do not require constant replacement. Generally, the decorator buys a large number of flowers and ornamental plants, so that he uses the marketing centers and marketing cooperatives as the main purchasing channel whenever they are close by. When distant, the decorator buys via order and the wholesale distributor makes the delivery. In more distant regions, such as the state of Amazonas, for example, orders are transported by air, which increases the value of the final product." [8]

2.1.5 Final Consumer

The final consumer, the one that buys the product in a flower shop, supermarket, or even online, is the one that makes the final decision. The average annual Brazilian per capita consumption of flowers and ornamental plants is estimated at R$ 45 (US$8) in 2020, considering the R$ 9.5 billion (US$ 1.7 billion) that this sector movement at the level of the final consumer [11]. On average, women spend R$ 100 (US$18) to R$ 200

(US$36) for one year on flowers and ornamental plant products, and men, around half, R$ 50 (US$9) to R$ 100 (US$18) [15].

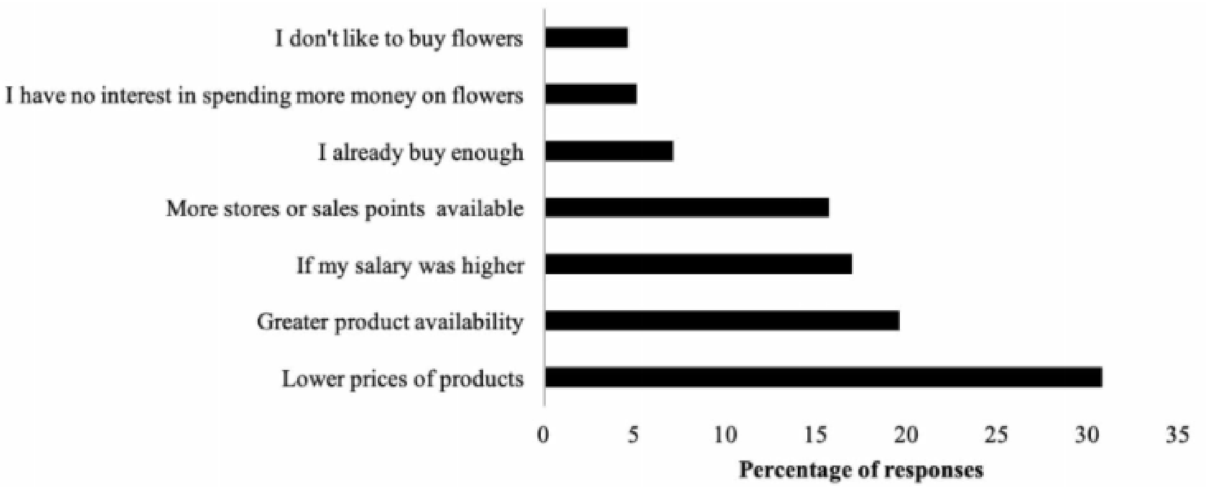

Figure 3: Situations that may encourage greater consumption of flower and ornamental plants according to consumers' view in Brazil (multiple responses). Extracted from [15].

According to research conducted during 2017 and 2018, what could drive more consumption of these products is if they had lower prices, as presented in Figure 3. This is a result of a cultural statement in which, unfortunately, flowers and ornamental plants are considered expensive products and not essential for daily life. On the other hand, lower prices may reflect limitations for the producers. Greater availability of products was also appointed as the second main issue that could stimulate responders to raise their purchases. This is an indication that flower shop owners need to diversify more and offer different options to retain a greater number of customers. However, for retails, a great diversity of products sometimes represent a problem since they may not be sold and represent lost money [15].

If Brazilian consumption is compared with other countries, the country is considered a market under development. According to [9], Brazil still has a low consumption per capita, if compared to developed markets, although this can be considered a signal of the potential of expansion to reach the consumption level of more saturated markets. Table 5 presents the main differences in the consumption of emerging markets, growth markets, and saturated markets.

2.2 Commercialization Flow

The commercialization flow depends heavily on the structure that the growers have and the professionalization of their market. Figure 2 presented an overview of the different

Groups of countries by development stages of their markets	Main consumption characteristics
Emerging markets	Low per capita consumption; Low percentage of buyers; Tradicional assortment; Special occasions of consumption (e.g. Mother Day, Valentine's Day, Women's Day, Weddings, Funerals)
Growth markets	Strong growth of per capita consumption; Growth percentage of buyers; Consumer wants to have more choice than the traditional; More gift occasions are developing (e.g. birthdays, Easter, Christmas, visits, Friday Bouquet).
Saturated markets	Minimal growth in consumption or even stagnation or decrease; Flowers for everybody, every day; Much interest in innovation of assortment; Trends in flowers and plants are important (interior decoration and personal style).

Table 5: Main consumption characteristics of flowers and ornamental plants, in different countries in the world-wide market. Extracted from [9].

types of sales that may exist, for the point of view of the growers which productions are sold to other companies, and for the point of view of these other companies, that mostly have as the buyer the final consumer. This section is intended to make a broader description of these transactions.

2.2.1 Agricultural Production

Part of the Brazilian cooperatives uses the Dutch Auction System. It is implemented using fast-paced auction clocks displayed on a digital board. Aside from the current asking price, each clock also contains information about the setup of the current auction. As the clock ticks down counterclockwise, each bidder can stop the clock by pressing a button indicating that she is willing to accept the price corresponding to the current clock position. The first bidder who makes a bid wins [2].

For the buyers, there are also different ways to participate in the auction. One new feature used is an advanced bid, a bid done before the auction occurs. The growers insert sales proposals for the next few days, allowing the customer to view the available products and manage their purchases, guaranteeing the amount paid at the time of purchase, regardless of the price they could reach at the auction [16]. Figure 4 represents the characterization of the buyers from a flower auction in the two main scenarios.

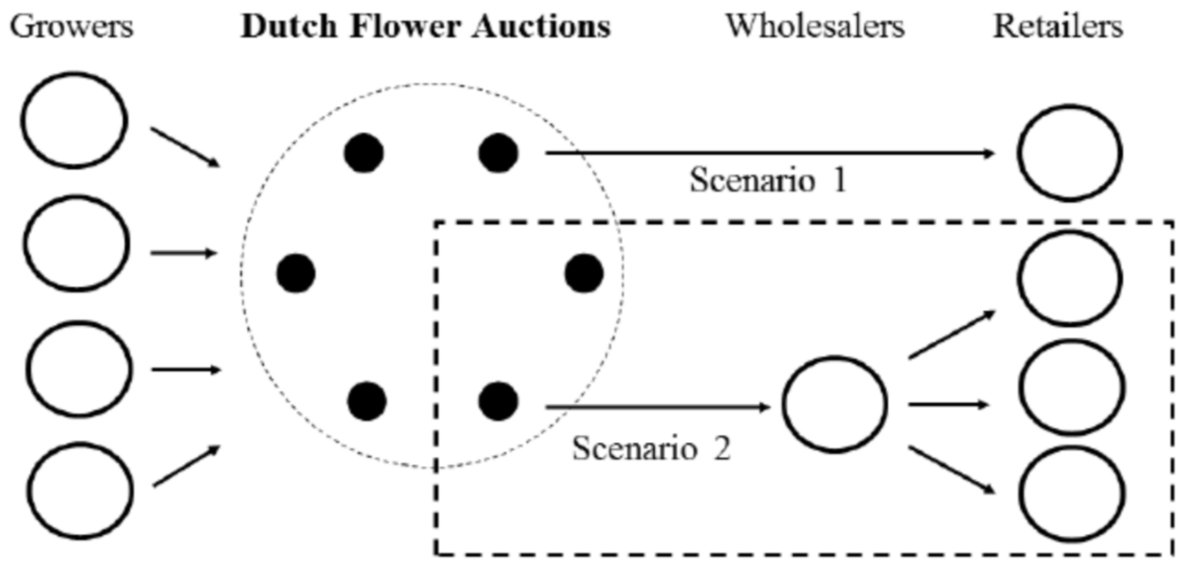

Figure 4: Overview of the two buying scenarios. In the first scenario, retailers buy for themselves. In the second scenario, wholesalers buy for their retail clients. Extracted from [17].

Intermediation contracts (also known as direct sales) are a system in which producers and distributors close short, medium, or long-term contracts, formal or informal. Prices, product features, and delivery times are agreed upon when closing contracts [10]. This type of sale creates an opportunity to make sales even before the products are harvested from the sites, thereby reducing the risk of loss of sales and speculation about the excess stock available.

Via web services, registered buyers can acquire directly from the growers' different types of products. The growers register their products in advance (even before harvesting the products), providing information about quantity, quality, and product. The buyers have access to this information and can negotiate with the specific grower [12]. The cooperatives work like a marketplace where growers can display their products and registered buyers are allowed to make deals.

The supply centers are the oldest and most traditional marketing system, in which producers stand side by side, offering their products to customers [10]. This system is mostly used by independent growers, once those that sell its products in cooperatives sign exclusivity contracts (so they are not allowed to perform sales in the channels available in the cooperatives).

2.2.2 Wholesalers and Retailers

The distribution centers aim to meet the needs of wholesale sectoral marketing in states or municipalities where there are no supply centers or where they do not have this type of wholesale in their functional structures [10]. They differ from the supply centers because of the private structure.

Traditional retail is mainly represented by florists, street markets, and street commerce, which, in most cases, operate only seasonally on special dates, notably Mother's Day, Valentine's Day, and All Souls' Day [10]. On the other hand, holidays near the weekend, which inhibit the holding of wedding parties, and holiday months, when there are fewer parties, are periods with a decrease in demands for the decorators and other resellers [18].

The supermarket and garden centers are two different structures but focused on the commercialization of flowers together with other products. In the productive chain of flowers and ornamental plants, self-service has been gaining space and assuming more and more importance [8].

E-commerce is getting more attention in the commercialization flow. The technological aspects related to the operationalization and ease of navigation on the purchase sites also can be identified as decisive aspects for the growth in sectoral importance [10]. With the pandemic, that affected Brazil from 2020, the use of e-commerce and social media as the most frequent action to restart the sales during the closed periods[19].

3 STATE OF THE MARKET

This work proposes to leverage the power of the data for the flower industry, so a deep dive in the market is critical. This section proposes to present an overview of the main products, tendencies, and issues of the industry.

3.1 Products

"The product range in flowers and plants is huge, and this diversity in assortment requires strong operational management in floriculture wholesale, and trade business" [20]. This variation is not only related to biological characteristics, it also can be modified by supply and demand periods along the year. For example, on the one hand, the production can vary from summer to winter, climatic conditions, and on the other hand, the product demand on Valentine's day is completely different from the one of the All Souls' day. Figure 5 presents research conduction in Brazil from 2014-2016 of how different products are consumed in different holidays, based on the ceremony.

The products can be divided into two groups: cut flowers and potted plants. "Cut flowers, the largest product group, can be split into mono-bouquets or mixed bunches. In potted plants, there are numerous types, colors, and sizes. The main distinction for plants is the difference between flowering plants and leafy plants" [20].

In Brazil, the main species cultivated for cut flowers are roses (30%), chrysanthemums (15%), lisianthus (12%), lily (7%), and gerbera (6%). The potted plants have a wide mix of varieties, where the six main species are: orchids (14%), lily (7.5%), chrysanthemum (7%), kalanchoe (6.4%), violet (6%) and bromeliad (6%) [9].

Date	Roses	Orchids	Mixed bouquets	Poted flowers	Baskets and flower arrangements	Chrysanthe-mums	Others
European Valentine's Day Feb 14th	51	10	24	5	5	0	5
International Women's Day Mar 8th	61	7	9	11	9	0	3
Easter April	11	13	0	0	47	8	21
Monthers Day 2nd Sunday of May	38	22	5	18	6	0	11
Valentine's Day Jun 12th	64	13	6	3	6	0	8
Father's Day 2nd Sunday of Aug	0	19	0	0	28	15	38
Beginning of Spring Sep 23th	34	19	12	8	0	12	15
Secretary's Day Sep 30th	45	19	10	13	0	10	3
All Souls' Day Nov 2nd	21	1	0	14	3	53	8
Christmans and New Years's Eve Dec 25th and 31th	17	6	0	19	11	0	47

0% 10% 20% 30% 40% 50%

Figure 5: Main flowers offered as gifts on holidays in Brazil, average 2014-2016 (%). Data extracted from [9].

3.2 Digital Transformation

Digital transformation is one of the components of productivity growth, where it enables an organization to put its resources to the most efficient use, reduce losses, and optimize resource management. In the past, being digital was a differentiator, but today it became a necessity for incumbents not to lose their market to the emerging new digital companies.

The four main domains in digital transformation include technology, data, process, and organizational change. The technology brings the building blocks for transformation, including the Internet of Things (IoT), blockchain, data lakes, artificial intelligence, and others. The data can be perceived as the flow of information throughout the company and is closely related to the tools presented before. Transformation requires an end-to-end mindset, so the process is important for the transformation not to be reduced to a series of incremental improvements. Finally, the organization means embracing the human side of the change, so the new process and technologies can be incorporated into the daily routine inside the company [21].

This thesis will focus on the technology and data aspects of a transformation. This section is divided into a discussion of how the floriculture sector is in the aspect of digital transformation in the process, focusing on the Brazilian hub, with benchmarking across the leading companies in the world.

3.2.1 Internet of Things

“Supply chains are increasingly being virtualized in response to globalization and emerging market challenges” [22]. In this context, IoT technologies allow the chain to gather data from physical elements and, consequently, analyze, control, and interact with devices, equipment, and people [22]. The IoT systems are usually combined with robotics, automation, and analytics to create devices for capturing new data, decision support software, and big data analytics. Also, it can be combined with on-farm machinery, automation, drone manufacturers, and growing equipment.

The IoT-based systems in the supply chain “build on traceability systems that provide information to track the location of certain objects (e.g., products, box, pallets, truck) and trace its history” [22]. But in the context of flowers, the IoT technology can go far beyond the trackability function. There are many IoT tools for monitoring ambient conditions (e.g., temperature and humidity) and other quality parameters, so Internet-connected actuators can equip objects with remote operating such as coolers, irrigation, and other functions [22].

Although IoT technologies are widely spread across industries, increasing the amount of data generated by the process can also lead to friction. One example is using IoT sensors to control the temperature in the transportation of flowers, a process that is highly related to flower durability. Temperature sensors combined with warning systems could create awareness of problems in the process, but problems in the relationship between the growers and transporters are due to a more transparent ecosystem.

3.2.2 Blockchain

“Organizations that invest in emerging supply chain transparency technologies, such as blockchain or smart contracts, often realize more efficient visibility of partners’ activities and minimize supply chain risks” [23]. Blockchain technology has the potential to transform the collaboration and share of knowledge between supply chain partners, and this is also applied in the floriculture sector.

According to Wageningen University, in horticulture, various platforms enable companies to share data. The blockchain technology can link all participants in a network in which they could reliably share and retrieve information. The Public-Private Partnership Project (PPP) "Blockchain: Automated Compliance" is investigating the conditions under which blockchain technology has added value for breeders, growers, and traders in floriculture to increase transparency about the use of improved crop protection products at the batch level throughout the chain [24].

The data ecosystem in the sector still consists of many linear data flows, once that each part within the production chain has its structure. However, this is "a fragile digital ecosystem when a party decides not to pass on that data and information to the next player in the chain" [24]. There is also often a lack of feedback from data in the chain, as presented in Figure 6, so an individual consuming at the end of the chain has no, or limited, information about the origin of its acquisition.

Figure 6: Fragility of the data chain in the supply chain. Extracted from [24].

Although the technology exists (blockchain), its implementation must have a goal: increase the value added for the final consumer or reduce the costs for the members of the supply chain. This case is applied in the floriculture of developed countries where "it costs to breeders, growers, traders, and other parties a lot of time and money to demonstrate that their products meet the legal or nonstatutory requirements" [24], so an implementation of an integrative system could be one solution to reduce costs for certifications because of a less amount of inspection and audits. "Blockchain can reduce the burden of proof and costs for growers by sharing and maintaining data and information in a standardized and reliable way for buyers and certifiers" [24].

Some of the challenges for blockchain technology to move from early stages applications to large-scale production are divided between the technological and implementation challenges.

For the technological side, there is limited transaction speed, once it can only execute a limited number of transactions per second and industry and business often require high levels of processing power; latency, since the frequency with which blocks become available and can be added to a blockchain varies; lack of uniformity, once the Key Data Elements (KDEs) used to store data are not harmonized across industries and users; and the energy costs, since computers require a large amount of energy to mine blocks for the blockchain.

On the other hand, the implementation challenges include transparency issues. given that competing development of private blockchains inhibits transparency; access, once there are potentially limited access to blockchain technology in developing and rural communities; costs, because of the high upfront cost for the technological infrastructure; and the behavior change, that is, the willingness of farms to adopt blockchain practices.

3.2.3 Artificial Intelligence

Artificial Intelligence (AI) algorithms are becoming popular within the industry and agriculture sector. The algorithms can manage complex data and create information and insights to help companies to make the right decisions. From the daily supply to the cooperative, to the type of sale (clock presales, auction clock, direct trade), the growers need to decide when to offer their products and how.

When deciding where to apply AI, the first step is to identify where technologies should be incorporated. The key decision points where there is a constant need for individual involvement are often the best place to install smart tools and to maximize the investment, but to work properly, a high amount of data is needed.

From price predictions to which plant to cultivate, different models that apply AI and machine learning can leverage the power of data for the sector, but a digital transformation must be held with a change of culture. To give an example, the recommendation from the algorithm should be used to identify leads and maximize the revenue for the grower. Choosing the right customers with the highest propensity to buy, with the right opportunities for each segment at the right time can support campaigns and the commercial department to increase conversion.

Usually, the AI-based approach is used when high availability of historical data can be used to train the machine learning model, like historical sell-out data, sales transactions, and data related to temperature, rainfall, wind speed, and humidity.

3.3 Cultural and Technological Status

This chapter focus is to get an overview of what is the current technological level of adoption for the Brazilian floriculture sector, and how it is incorporating the new technologies available in the market. The first section is about the growers, followed by a section about the cooperatives and finally, the perspective of the wholesalers and retailers. "While technology is fast maturing, innovation seems to be delayed by organizational barriers" [22], and many are the complaints and context for these barriers as presented next.

3.3.1 Growers

The floriculture sector involves many small players, as presented in the previous section of the industry overview. The average size of a company in this sector in Brazil is 1.88 hectares, really low if compared to other sectors (e.g., in the soy segment, one of the main food commodities exported by the country, the average property has 145.72 hectares of production [25]).

The mix of small companies create some constraints for the change in processes (e.g., adoption of new technologies), but a "the combination of 'the sermon' (communication, persuasion, and information), 'the carrot' (bonus or other benefits for the first growers who switched) and 'the whip' (fine or higher costs for growers who did not enter digitally)" [24] can help to boost the implementation. An example of this happened in the early 1990s when a first step was taken in floriculture with the automation of auctions (where the growers had to purchase computers and change the way they transmit information for the cooperatives).

In this phase of the project, different growers from the main Brazilian cooperatives were selected to have different perspectives for this research. From the qualitative interviews, it was possible to define the main strategic decisions to be made in the operational process and what are the barriers and issues for them related to the digital transformation.

The decision to expand the production is heavily influenced by the growers' expertise in the field, aligned with client feedback about current and new products. The current product's performance (e.g., demand in peak dates) is known by the growers, according to the conversations. The cooperative website has a page to display the sales data, although currently, only the past three months of sales are presented. If the grower does not save this data recurrently, he has to pay a fee and request the data from the IT department.

When the focus is on testing new products, the first step is a productivity test inside the farm, to see if the product adapts itself to the climate conditions and the number of flowers collected per year is expected and performed by other growers. In parallel, the market acceptance of the product is also evaluated in the event of exposition (usually provided by the cooperatives). As it is an important decision to be made, there is a high effort to perform these tests before the final decision of which product to expand the production, however, it is still highly influenced by the human knowledge of the field, and less by data or other technology tools.

For some growers, the cooperative offers two ways to sell the products, as presented before. According to the interviews, direct sales (intermediation) have a higher return per flower compared to those that are sold in auctions. However, the sales data show that this is not always the case, so for the peak dates, an important decision is if the product should be sold in advance to the clients, and if so, which price to perform. To answer this, historical data questions are used, as well as the knowledge of the production for the period, but no developed tool helps in this process.

Moreover, it was possible to see that there is a client relationship to offer special discounts and encourage recurrency, but not in a structured way (e.g., all contracts of recurrent acquisitions are verbal for the growers interviewed). Based on the sales data, a data mining approach could be developed to cluster the clients and prioritize the sales channels. The goal of this technique could be to give more attention to buyers with a good history of purchase, as well as transform auction sales into direct sales in the long term.

3.3.2 Cooperatives

The main goal of the interviews with cooperatives was to define their level of technology implemented in the process and their objectives for the short-term. To do so, one executive director from Cooperflora, the biggest Brazilian cut flower cooperative, was invited to a conversation. Information about Veiling Holambra was obtained through desk research and participation in events. Royal FloraHolland was reached by one interview with a product owner of the data department.

3.3.2.1 Brazil

The aim of the cooperative is commercialization, and to achieve this objective, Cooperflora has different departments to be the marketplace between growers and buyers. However, different from an online marketplace, the cooperative, as like others, works as the main center to get the products and take care of the process of the sale.

Logistics is an important part of the operation. For regions near the cooperative (e.g., the city of Sao Paulo), they deliver the product store by store and not in the distribution center of the client (e.g., some chains of grocery stores adopt this service). The reason behind this action is the perishability of the product. Even though it is possible for the buyer to get the product from the growers' farm, this is not common, once that buyers usually have a diversified portfolio of acquired products and it would not be feasible to get the products farm by farm.

In the first quarter of 2021, around 65% of the sales of Cooperflora were performed using a digital marketplace. The tendency is that this number continues to grow, but due to cultural conditions the expectation is to reach 90% as the maximum penetration rate for this channel - some customers still prefer to have the person contact inside the company to perform their acquisitions.

There are up to 13 pricing modifications during the week for the same product (in 2020 modifying the price more than twice per week was not common). This is culturally disruptive from the buyer's point of view, but the tendency is that the price becomes increasingly volatile. There is the use of algorithms to define the price for each product, for each grower, each day. The commercial department defines the index, but this algoritmus, based on the historical data, performs the fine adjustment.

By the number of uncompleted orders, the cooperative has data about which products there is a pent-up demand. They share this information with all growers, but the decision to cultivate or not, and the size of the plantation is up to the decision of the individual grower. The sharing of data between the growers is transparent by cooperation. One grower can see his average price, as well as the ones from his peers. The volume sold is also available, however, the buyer does not have access to this information.

About Veiling Holambra, currently the cooperative is developing a new platform to integrate the entire chain of flowers, they are willing to bring new tools to facilitate the process for their clients using a digital platform. Their vision is to create a digital ecosystem, connecting the product to the market: growers, cooperatives, clients, and sales points. The plan is to have a marketplace until the end of 2021, where the growers can present their products and the wholesalers can use the integrated system to offer the products to their customers - flower shops, decorators, among others. Their projects also include the development of algorithms to, for example, purchase suggestions for their clients, based on historical data.

3.3.2.2 Netherlands

Royal FloraHolland was reached by one interview and more information was obtained by participation in their events. The cooperative is changing from an auction-focused platform to a digital marketplace, with three main building blocks: ordering; payment, and delivery. Ordering is related to Floriday, their digital B2B platform that provides growers and buyers the opportunity to choose the way of doing business that best suits them. Payments are related to the international market, connected by the cooperative with different currencies. The delivery services are logistics solutions and transform all orders from the greenhouse to the buyer to deliver more often, faster and fresher. The company started a joint venture with four carriers to improve the delivery service inside the Netherlands.

There are only a few software companies focused on the Brazilian flower market, due to its smaller size compared to other sectors. In this sense, the Dutch market was selected and an interview was conducted with the Product Owner of Insights, an additional service offered by Royal FloraHolland to their growers. The conversation started with the long-term strategy for the company and how they are passing through the digital transformation in the Netherlands.

Traditionally, FloraHolland did auction clocks, but in the last few years, they are creating a new digital platform, Floriday, because they understand that the original way of selling can't be held forever. The main goal of the cooperative is the creation of an ecosystem to connect growers from all around the world and get more efficient logistics.

Floriday Insights is part of this transformation into a digital tool that allows growers to perform different analyses of their sales trends and historical information of the company. It is a service available for growers from FloraHolland and it is already being integrated with other companies from Germany and Belgium. Although there is this need for the digitalization of the sector, the cooperative is facing different issues with its clients, where the adaptation is a slow process due to different stakeholders with different capabilities that in the end have the same power of decision in assemblage, the most common way of making decisions in a cooperative.

One goal for moving to digital is to perform an improvement in the carbon print of the sector. With a 24/7 marketplace, the goal is to work together with the transportation companies to work more efficiently. They started the cooperation with the three largest transport companies to do it together and make it more efficient, because they think that they lose a lot of money by not being very efficient in a logistical way, between the different locations of growers and buyers. If it is done in a large body, it could be much more efficient.

FloraHolland does have price predictors, but they are not allowed to use most of them because of regulations of the fair market. If FloraHolland predicts the price, it will be because they have too much data, too many growers next to them. Therefore they can predict it pretty well, but if they do that, it will be a self-fulfilling prophecy. Therefore, that is why they are not allowed to predict prices in general, for their growers.

To solve the issue of a fair market, some services have different options. Consequently, they have a price prediction with five different options and settings, that uses more actual data using the grower's data, uses wide data, and then the user has to choose what to use. And then it is allowed because depending on the choice the grower creates, he gets a different prediction.

FloraHolland also wants to make smarter connections, although it may be very risky to do that in a fairway. For example, if a certain flower company is always buying red roses from two different products, and FloraHolland suggests that they can also buy another product that is not allowed because other growers may also have good products and FloraHolland is an independent marketplace. FloraHolland has a lot of services, but they always have to be cautious in what they do and automate.

3.3.3 Wholesalers and retailers

To achieve the overall players of this segment, qualitative interviews were developed with the procurement sector of supermarkets, with the main goal of understanding their position in respect to the flower segment. From these interviews, it was possible to infer a willingness to buy directly from the cooperative to not pay taxes to the intermediaries, once they do weekly acquisitions. The procurement department is interested in the weekly price (daily changes are not essential).

For the supermarkets, the product is used as a gift, so you need to make the client want the product. Usually, there's no markdown approach for old products, once it's a premium network of supermarkets and it would not fit its brand. In the supermarket context, there is no need for a large variety of products. There was an occasion where they reduced the number of them and the sales increased. Variety does not mean better sales.

4 METHODOLOGY

The Brazilian flower industry is a complex environment, due to its fresh products and the size of the value chain, with many stakeholders involved and different parameters of quality available.

The work was defined in two main phases, the first was the qualitative research focused on literature review, event participation, and interviewing with experts to define the problem, understand the main decisions that must be made, synthesize and ideate on it to define the project requirements. The second phase was based on the input collected before defining which initiative should be prioritized for the development and, later, testing the prototypes with growers and consumers.

4.1 Qualitative Research

To create the baseline and support the research for the next topics, the first phase of the literature review had the following goals:

- Understand the Brazilian Flower Industry;
- Get the tendencies happening abroad;
- Deep dive to the stakeholders of the value chain;
- Define the strategic decisions to be made by growers.

Participation in events in the field is intended to gather data from people that were not reachable by interviews. Two online events were selected due to their importance in the global and regional context: Cooperative Day, held by Royal FloraHolland on 15th December 2020, and Veiling Market, held by Veiling Holambra from 14th to 16th April 2021. From the first event, the objective was to have an immersion in the global context of the flower sector and what are the goals for the next years, and, for the Veiling Market, the intention was to look at the same aspects, but this time focusing on the Brazilian market.

Initially, interviews were held with business employees of the cooperative of growers, to understand how the process is structured and get information that is not described in the literature. A semistructured interview, with a question script, was applied as the data collection instrument using the in-depth interview approach, where "interviews are often used to provide context to other data, offering a more complete picture of what happened in the program and why" [26].

To obtain valuable results from the interviews, a few goals were defined to structure the conversation with each stakeholder. Below is presented the main objectives for each group and the target interviewed.

The target growers for this research are the Brazilian ones, from different cooperatives. It is aimed to interview the operational managers and employees responsible for the pricing and client relationship. From the interviews, the goals were to:

- Understand what are the strategic decisions in the flower industry;
- Find where there is and where there is no data provided to them;
- Understand their long-term objectives;
- Level of information technology adopted;
- Type of sale with higher return.

The cooperatives are the fundamental base for the movement of products to the final consumer. Therefore, for this research, both main cooperatives from Brazil, Veiling Holambra, and Cooperflora, were contacted. Moreover, the main global flower cooperative, Royal FloraHolland from the Netherlands, was selected due to its pioneer in the field. The conversations were willing to:

- Understand in which level the growers and buyers today use data for their strategic decisions and what is the role of the cooperative;
- Get a view of the goals for the short term;
- See what is common to be outsourced and what they build internally.

There are not many software companies focused on the Brazilian flower market. Due to this lack of expertise outside the cooperatives, the project selected one Dutch company that is a spin-off of the main cooperative in the field: Floriday Insights. The service is an additional feature that offers analytical tools for the growers. The conversation had the following objectives:

- What are the main challenges in developing tools for this sector;
- Good practices learned with the time;
- The main goals for the next few years.

The last segment before the customers are the buyers. They are divided into wholesalers and retailers. For the wholesalers, the decision was to select the logistics and procurement departments of these companies to perform a qualitative interview. The goals for the qualitative interviews were:

- Get the importance of the flowers in their segment;
- Main problems occurred in the past and learnings;
- See if their suppliers are many or focused on just one farm/cooperative.

4.2 Project Requirements

The main purpose of this project is to enable a “gut instinct” decision-maker with tools to support his strategic decisions using a data-driven approach. To achieve this goal, the work starts by defining a persona with many interviews addressing the main stakeholder of this project: the grower. Once the persona was defined, the project created a platform with different algorithms to support the decision-making that makes the most impact in this environment.

Two stakeholders were defined: the manager of the farm that is responsible to decide the price to request from the clients is the persona for the primary objective, and the final consumer that sporadically buys flowers and plants is the persona for the secondary objectives. For these two personas, the system requirements were defined in Table 6, where the primary requirements are related to the grower and the secondary requirements, related with to final consumer.

Objetive	#	Requirements
Primary	SR0	Functional - The system should be able to generate a report for a given sales period
	SR1	Functional - The system should have new transaction data every day
	SR2	Functional - The algorithm should offer suggestions personalized for each user
	SR3	Non-functional (product) - The algorithm developed should be able to overperform the baseline for the price prediction
Secondary	SR4	Functional - The application should be able to collect the user data when the consent is given (e.g.: location, photo, comments)
	SR5	Non-functional (product) - The application should allow users to access the main functions without being logged in
Both	SR6	Non-functional (external) - The system should fulfill all the Brazilian LGPD (General Data Protection Law) requirements
	SR7	Non-functional (external) - The system should protect its user data

Table 6: System requirements.

4.2.1 Evaluation Metrics

For most of the requirements, the fulfillment is dichotomous: or it is achieved, or it is not. This is the case for SR0, SR1, SR2, SR4, SR5, SR6, and SR7. From the technical point of view, these requirements can be achieved if all functionalities are created and work smoothly. This can be checked by performing tests in the system or referring to external requirements (e.g., the LGPD case).

For the requirement SR3, on the other hand, it is not a true or false situation. To help people make the right decision with uncertainty, a forecast model can be used. Before creating different models and comparing the key metrics, it is necessary to define the baseline, that is, the principal number that the model should overperform.

The baseline should be how the process works "as is", so adapting to the flower case, to define the price for tomorrow's sales, the price of today is the preferred metric, based on the interviews with growers. Therefore, any model developed should be able to overperform this baseline. To compare the results, 3 metrics are going to be used: the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), and the Mean

Absolute Percentage Error (MAPE). The RMSE is traditionally used for measuring the differences between the estimated values and actual values, in the MAE, it is possible to see the error in the same scale of the data, and MAPE indicates the average percentage difference between the predicted value and the original value.

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right|$$

Where $\hat{y}_i$ is the predicted value and y_i is the actual value for the time i. And n is the total of days used to compare the models.

4.3 Product Definitions

Once defined the options and the ideas of the projects to be implemented, the prioritization was by the activities that would create more value for the users.

The algorithm for the price predictor was developed using the Python programming language. To test its results, two approaches were performed: comparing the results with a predefined baseline; structuring an online dashboard to make again use of interviews, this time validating the insights from the tool.

The online dashboard was developed using React and Javascript. The database was structured using MySql and the data flow was performed by APIs created using PHP. The mobile application was developed using React Native and Javascript. The data flow used the same structure that the one used for the online dashboard.

5 DEVELOPMENT

This chapter focuses on describing how the methodology was implemented for this project. After phase one of the research, different use cases were created and databases were obtained for this study. After it, the data relation was established and the architecture of the system was created. Then it is presented the structure of the price prediction algorithm and the development of the web and mobile application for the platform. Finally, an overview of how the system was integrated is presented in the last section.

5.1 Use Cases

Before any part of the development, the use cases of the system were created. Using qualitative interviews and to reach the primary and secondary objectives of the project, two applications were defined with different purposes, one for computers and the other for mobile devices. Figure 7 presents the main actions of both interfaces. The web application is intended for the internal growers of a cooperative, and the mobile is designed for the final consumer.

The use cases for the web have two actors: the grower and the webmaster. The grower is defined as someone with the production of flowers and plants and affiliates with a cooperative, and the webmaster is an employee of the company that authorizes the registration of the users and manages them. The grower can use the system to check the past sales and analyze the market tendency, the main use case for the system.

The mobile application, in contrast, is designed to be an open application with access to a registered or anonymous user. There are three actors in the system: the registered user, the anonymous user, and a webmaster. An anonymous user is someone that downloads the application, decides to not sign up but has access to some features of the system. The registered user has access to all functions of the system and there is, again, the use of an actor as the webmaster to stay in the background of the application. The platform is designed to capture data from users that are interested in this segment: flowers and plants. It is a place to center the communication between the grower and the final clients, as well as to promote new brands and products.

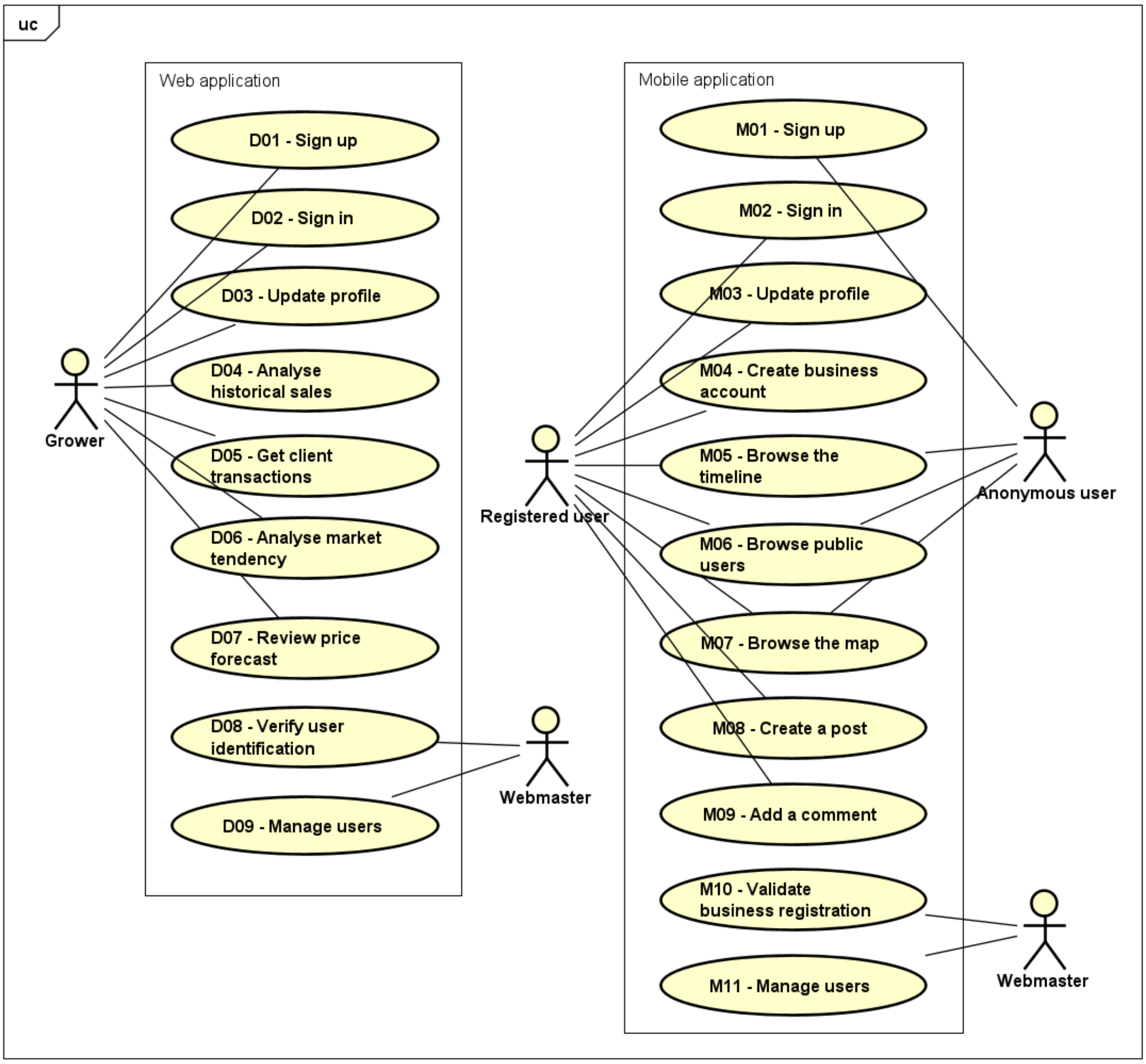

Figure 7: Use cases for the web and mobile platform.

5.2 System Architecture

This project was divided into three parts: the front-end to interact with the user, a back-end with a relational database for support, and an analytical part created with different algorithms and data requests. This section will describe how the system works and which integrations were used. Figure 8 presents an overview of the platform. Both the mobile (React Native) and web (React) systems were created using the same architecture, although the mobile application does not have the use of any Python algorithms.

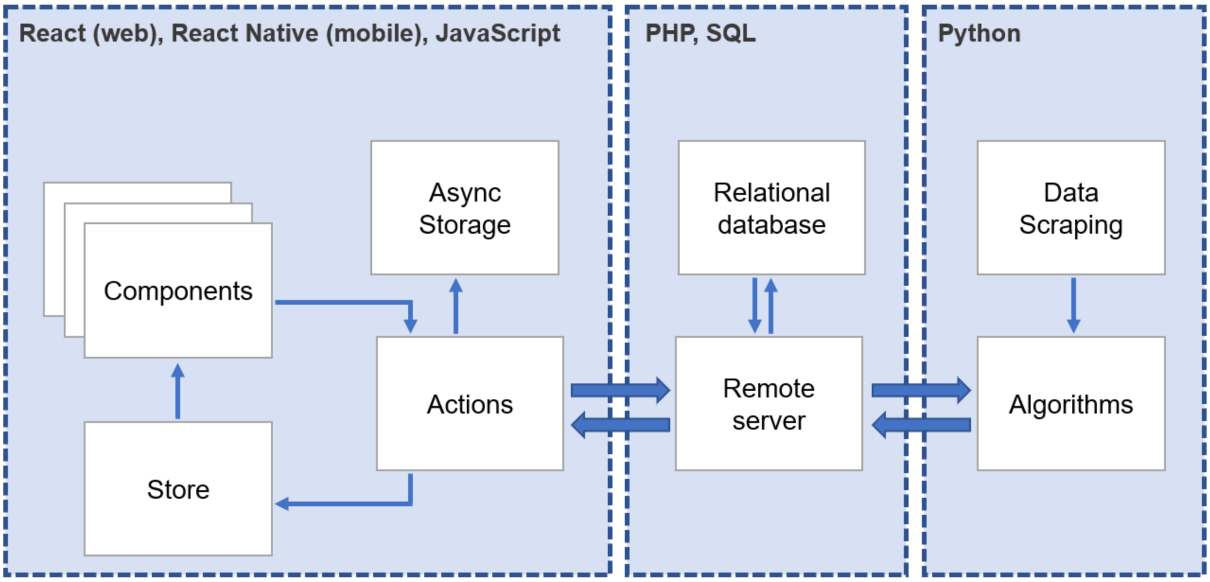

Figure 8: System architecture for the mobile and web application.

The first part created for this project was the data scraping function (explained in detail by Figure 9), using Python and some predefined functions. The algorithms created will be presented in detail in the next sections, but they use data from the data scraping, along with data from the remote server.

Once the structure of the data was designed, the next part of the development focused on the front-end to the final user. The use case, described by Figure 7, guided the development of two different applications, a mobile (React Native) and a web (React) one. Both share the same architecture, defined by actions, components, stores, and storage. The state of the application is kept in a store and every change to the state re-renders the necessary parts. Components dispatch actions that do something (e.g., send a message to the remote server) and then update the store.

Finally, once the data was already available, and a framework was designed to interact with the users, a remote server was created. Using PHP and MySQLi functions, some APIs (Application Programming Interface) were programmed to adapt for different uses of the system. The APIs are intended to connect the functions designed to the platform available to the user. More details of the APIs were given in the next section.

5.3 Web Application for Growers

This section includes all the development to create the web application for the growers.

5.3.1 Data Collect

The main goal of this work is to leverage the power of data but to do so, one important building block is the origin of the information used in the platform. This section will illustrate the flow of data that allows the algorithms to work.

The web application uses the growers' sales data. This data is collected from the cooperative system, using a daily POST request with Python. Figure 9 describe this process of data flow. The cooperative system authorizes access to the data with only authenticated users, so the first step is to send a GET request with the grower username and password. If both are correct, the system returns a session token and redirects to the main page. Then, to request the sales data, the process is to send a POST request with the grower identification and the date of the request, together with the session token. Unfortunately, it was not found a direct connection with the cooperative API, so the response is an HTML page with the transactions data in a table that is handled with Python in the Jupyter notebook to save the data in a structured and tabular way.

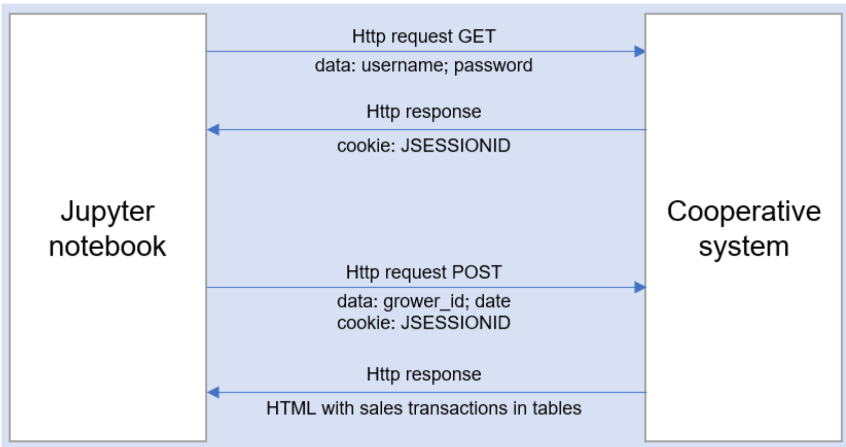

Figure 9: Data flow with the cooperative for the web application.

5.3.2 Data Modeling

This section presents all database structures for the web application. Using a relational database, different tables were designed and implemented to manage the connections between the data structures.

The web application uses a data structure that contains 8 tables, as presented in Figure 10. There are 4 fact tables, one for the relation between users and suppliers (e.g., one user can be connected to one or more suppliers) and the other for the sales transactions. There are 5 dimension tables, containing data from the tokens, users, suppliers, buyers, and products.

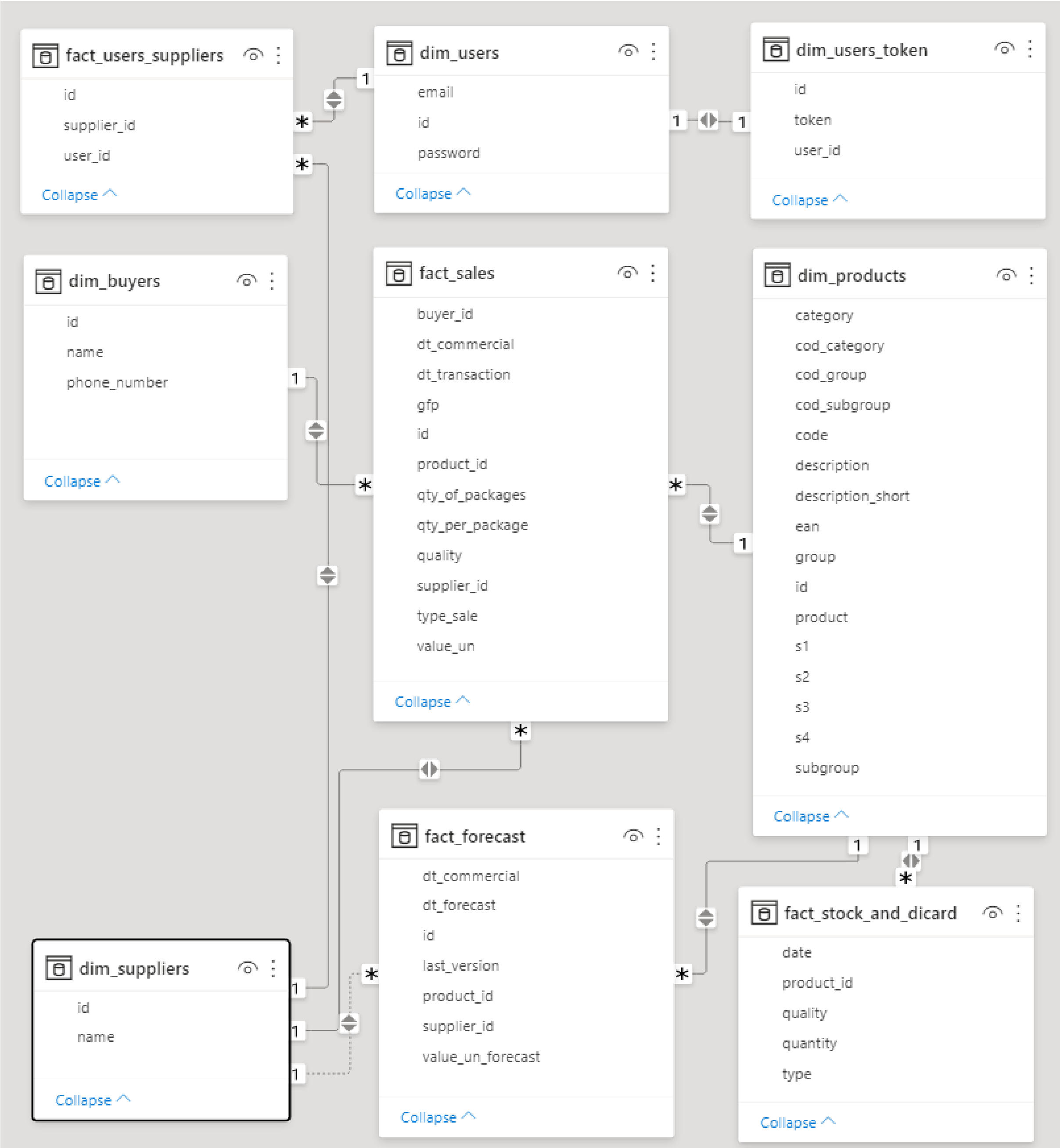

Figure 10: Relational database for the web application.

5.3.3 API Reference

An API can be defined as a software intermediary that allows two applications to talk to each other. For this project, numerous APIs were created to fulfill the use cases needed, as presented in this section.

Params flagged with a * are required. All params are shared with JSON protocols.

- POST /authentication

  Authentication required to use the application. If the params match to those in the database, return a unique token for the user and the user data. Params needed:

  - *authentication[email]: text;
  - *authentication[password]: text.

- GET /stock_and_disposal (Auth required)

  Primary endpoints for retrieving stock and disposal of the products. Params needed:

  - *stock_and_disposal[token]: information automatic added by the application to identify the user;
  - *stock_and_disposal[visualization]: identification of how to group the data. Allowed values: w, m, y;
  - *stock_and_disposal[hier]: Select the group type of the data. Allowed values: 4 (species) or 5 (specie and specific size);
  - stock_and_disposal[product_id]: identification of hier 5;
  - stock_and_disposal[cod_specie]: identification of hier 4.

- GET /products

  Primary endpoints for retrieving products. Params needed:

  - *products[token];
  - *products[query_type]: identification of the page retrieving the products. Allowed values: get-products-supplier, get-categories, get-groups, get-subgroups, get-species and get-size;
  - products[cod_category]: identification of the category;
  - products[cod_group]: identification of the group;
  - products[cod_subgroup]: identification of the subgroup;
  - products[cod_specie]: identification of the specie;

- POST /sales

  The entry point for the sales from the cooperative system. Obtained by data scraping and sent using Python algorithms.

  - *sales[sales]: a list of objects containing all transactions and details in the cooperative system.

- GET /sales (Auth required)

  Primary endpoints for retrieving sales. Params needed:

  - *sales[dt_initial] initial date to filter the sales;
  - *sales[dt_final] end date to filter the sales;
  - *sales[supplier_id] grower id;
  - *sales[sale_type] type of sale (Klok or intermediation). Allowed values: K, I;
  - *sales[cod_product];
  - *sales[product_id];

- POST /prediction

  Entry points for the price prediction created by the Python algorithm. Params needed:

  - *prediction[dt_prediction]: the prediction date for the unitary value;
  - *prediction[dt_commercial]: day that the prediction was made;
  - *prediction[value_un_prediction]: the value predicted, in Brazilian reais, of the product;
  - *prediction[supplier_id];
  - *prediction[cod_product];

- GET /prediction (Auth required)

  Primary endpoints for retrieving the predictions. Params needed:

  - *prediction[token];
  - *prediction[dt_commercial];
  - *prediction[supplier_id];
  - *prediction[cod_product];

5.3.4 Price Prediction Algorithm

The price predictor was a use case requested by some growers interviewed in the first part of the research. There are in the commercial routine many deals agreed upon daily, but there is not much to help in the decision-making. Today's price is the best guess for tomorrow's price, but could other factors, like the day of the week or period of the month contribute to making better deals and increasing the revenue from sales?

Any new model developed must be comparable to a baseline. This is necessary because a meaningful reference point, usually a simple logic that is behind the "gut-instinct" of the experts, needs to be outperformed. If with the available data it is not possible to predict better than this, other data points that should be used, or the problem may need to be reframed. As pointed in the section of the project requirements, the metrics to evaluate the models are going to be the RMSE, MAE, and MAPE.

5.3.4.1 Selection of the data

To create the model, it was used the transactional data from the sales that one grower had. The dataset was composed of the transactions from 2017 to 2021. Each transaction is characterized by a date, a buyer, product, the number of products sold, and the value per unit. To use the dataset, it was applied preprocessing to transform the data string into variables that could be used in the model, like weekday, month, and day of the year.

Figure 10 presented the structure of the data available in the web application. For the price predictor algorithm, data from table "fact_sales" and "fact_stock_and_discard" were used. The fact sales table presents transactional data for each sale, as presented in Table 7. The fack stock and discard, on the other hand, presents the daily stock and discard of the products of the cooperative, including all growers. The structure of the data is available at Table 8.

To select which product from the grower would be used in the model, it was analyzed the volume sold in auctions for each product compared to the total volume sold by the company, so there would be enough data points for the analysis. It was selected one product for the analysis: Rose Freedom, which corresponded to 64% percentage of the volume sold by the grower. Figure 11 presents the daily price for Rose Freedom. Notice that there is a correlation between the price with another product, Rose Revival, which is expected given their similarities.

With an exploratory data analysis, it is possible to notice, with Figure 12, the yearly

Variables	Description
Date	Day of the transaction
Buyer Id	Reference of the consumer
Supplier Id	Reference of the grower
Product Id	Reference of the product
Quality	Product quality classification
Invoice Id	reference to the transaction id
Quantity of packages	Total of packages sold in the transaction
Quantity per package	Quantity of products stored in a package
Unit price	Sale price
Type of sale	Auction sale or direct sale

Table 7: Structure of the data table fact sales.

Variables	Description
Date	Day of the reference
Product Id	Reference of the product
Quality	Product quality classification
Quantity	Total of units available
Type	Classification of stock or discard

Table 8: Structure of the data table fact stock and discard.

season of the variation of prices. Connecting to the events presented in Figure 5, event A represents the peak of sales for International Women's day, event B represents Mothers Day, and event C represents Valentine's Day. It is possible to notice that from 2017 to 2019, the price of the rose followed the same trend, but after the pandemic of 2020, the peaks and valleys are very different now, which can be a difficulty for the performance of a price predictor algorithm.

5.3.4.2 Baseline Definition

For this specific prediction problem, after the interviews with growers, it was clear that the best baseline for tomorrow's price is the one performed today. When defining the price for advanced transactions, the commercial department of the company uses the price performed today as a baseline. All models developed will be compared and try to improve the evaluation metrics with this approach. If the model does not overperform the baseline, it should be discontinued. If it performs better than the baseline, it could be used, although the commercial department would need to see the value of this algorithm, that is, the improvement should justify its implementation cost. This baseline definition

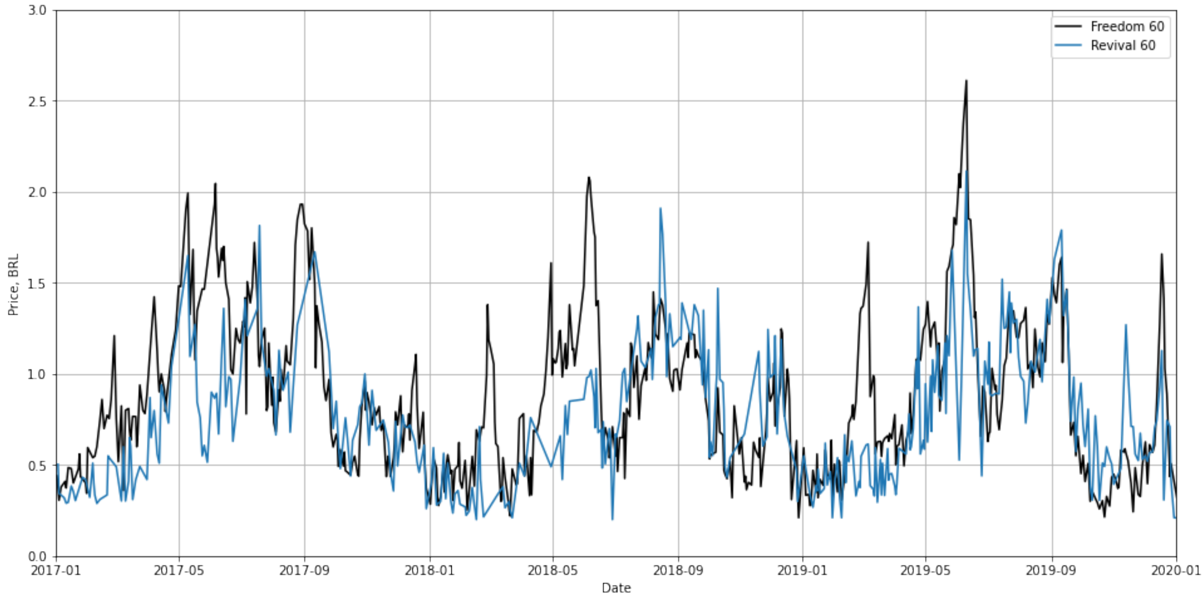

Figure 11: Auction daily average price for the Rose Freedom and the Rose Revival for the period of Jan/17 to Dec/19.

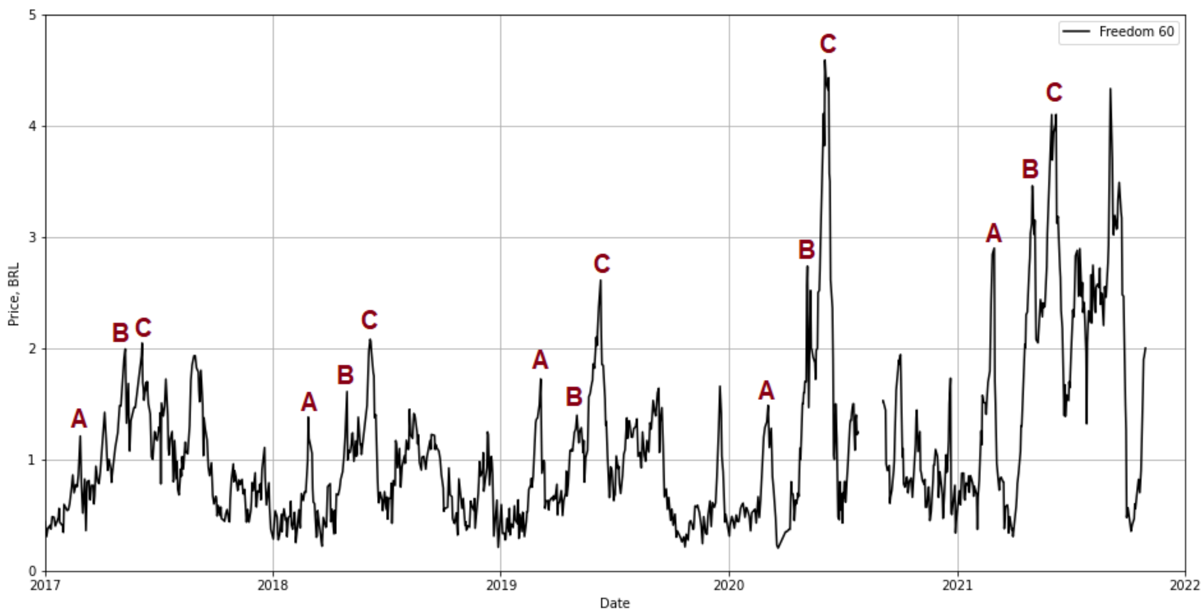

Figure 12: Auction daily average price for the Rose Freedom period of Jan/17 to Oct/21. Events A, B, and C mean, respectively, International Women's Day, Mothers' Day, and Valentine's Day.

is also cited in the literature as the Persistence Model Forecast or the Naive Forecasting.

$$\hat{Y}_{t+1} = Y_t$$

5.3.4.3 Model ARIMA

The Autoregressive Integrated Moving Average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model, commonly used in statistics for time series data. According to [27], “An ARIMA model is labeled as an ARIMA model (p, d, q), wherein:

- p is the number of autoregressive terms;
- d is the number of differences;
- and q is the number of moving averages.

Given the seasonal component of the data, it was used the Seasonal Autoregressive Integrated Moving Average (SARIMA) to create the forest, where the automatic method from the Python library auto-ARIMA was used to find the best parameters for the model, that is, the one with the lower error rate. For the SARIMA, more the only the seasonal components, it is requested to define the not-seasonal (P, D, Q) for the model. The auto-ARIMA finds the optimal that are suitable for the dataset to provide good predictions, through gradual execution of hyperparameters and comparative through the AIC (Akaike Information Criterion) metric.

After the execution of the auto-ARIMA, the best model for the dataset was the following: ARIMA(0,1,1)(5,1,0) where (0,1,1) are the seasonal components and (5,1,0) the non-seasonal components.

5.3.4.4 Model Random Forest

A random forest is an estimator that uses a defined number of classifying trees, which in turn are sets of splits or decisions on how to separate the data. After going through the splits and getting to the sheet, the data label is determined and then uses averaging to improve the predictive accuracy. “A random forest is random in two ways: (i) each tree is based on a random subset of observations, and (ii) each split within each tree is created based on a random subset of candidate variables. Trees are quite unstable, so that this randomness creates differences in individual trees' predictions", as defined by [28].

In this work, the model was tested using different features available in the dataset, and for the final model, it was selected the mix of features that performed the smaller RMSE error was. The following features were used: month, week, year, year-day, last-day price, weekday, holiday (true or false for Women's Day, Mother's Day, and Valentine's Day), and volume sold.

The algorithms was implemented using Python with the library Scikit-Learn. The number of trees in the forest was defined as 100.

5.3.4.5 Model LSTM

The Long Short-Term Memory (LSTM) is a specialization of a Recurrent Neural Network (RNN) that explicitly adds the manipulation of the order between observations when learning the function of mapping the input of the observations to the output.

For the construction and testing of the LSTM-based model, the Keras API of the TensorFlow Python package was used, which already has the implementation of the model natively. After some tests with the model, the quantity of neurons was defined as 50 for the data available.

5.3.4.6 Model Prophet

According to its library, Prophet is a procedure for forecasting time series data based on an additive model where nonlinear trends are fit to yearly, weekly, and daily seasonality, plus holiday effects. It was implemented using the Python language.

Prophet uses "a decomposable time series model with three main model components: trend, seasonality, and holidays. They are combined in the following equation" [29]:

$$y(t) = g(t) + s(t) + h(t) + e_t$$

$g(t)$ is "the trend function which models non-periodic changes in the value of the time series, $s(t)$ represents periodic changes (e.g., weekly and yearly seasonality), and $h(t)$ represents the effects of holidays which occur on potentially irregular schedules over one or more days." [29].

By default, the Prophet algorithm needs only the data (t) and the Y variable (in this

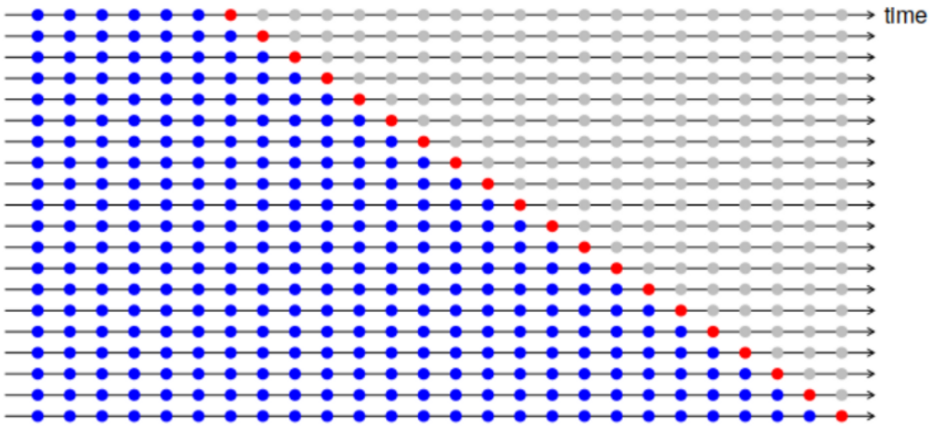

Figure 13: Illustration of the Walk Forward methodology. Extracted from [30]

project, the average price), so the implementation is quite easy. Another variable used in the algorithm can be the definition of important holidays as another variable for the models. It was defined the Mother's Day, Valentine's Day, and Women's Day.

5.3.4.7 Walk Forward Validation

The validation follows the Walk Forward Validation method, commonly used for time series forecasts. In essence, it divides the dataset between train and test data, and "each time it feeds the historical data to the new data to enrich it to improve the new model and forecasting" [27]. Figure 13 illustrates this concept. With multiple splits across different periods with the training data expanding each time.

5.3.5 User Interface

The objective of this section is to display some images of the MVP (the minimum viable product) developed for this work following the main use cases presented before.

Figure 14 presents an overview of the main dashboard of the web application, where the user can access the company results in terms of sales. This screen and all following dashboards are in an internal part of the website, that is, only authenticated users have access. Therefore, the precondition is to have a username and password in the system and the company should have been admitted beforehand for the data integration.

The objective of the screen presented in Figure 14 is to allow online reports for a predefined period. There are some fields at the top of the screen where the user can select the type of sale to get more details (auction sales, or intermediation), the period

Figure 14: MVP Web - sales dashboard: filters, data summary and visualization format.

of review (begin and end date), and, optionally, select a specific product and dimension. With each action from the user, the system sends a request to the database that contains the classified sales data and, after the response, the data is presented in the chart. Then, there are four big blocks presenting the main characteristics of the transactions for the requested period: revenue, volume, average unit price, and the number of buyers.

In the bottom part of Figure 14 there are four big buttons, they are intended to change the visualization format of the charts on a timeline basis: daily, weekly, monthly, or yearly. Figure 15 presents the charts using the monthly format for the product Rose Freedom of 60 centimeters. The first chart is the revenue from the product, followed by two charts, one presenting the volume sold and the second with the average price for the period.

Finally, the end of the main dashboard is presented with Figure 16. The bottom of the page is composed of different tables presenting the data. The first table presents the sales accumulated by-product, the second table presents the sales accumulated by-product (with the division of size), the third one presets the buyers, and the last one the daily total revenue.

The functionalities presented in the main dashboard are intended to serve as an online report for the grower to have easy access to the sales information without the need of downloading information from the cooperative system and creating its own data visualization.

The second important dashboard is presented by Figure 17: the daily sales dashboard. It uses the same structure from the previous dashboard, with the filters in the top, main numbers in the middle, followed by the charts, and then tables presenting the complete

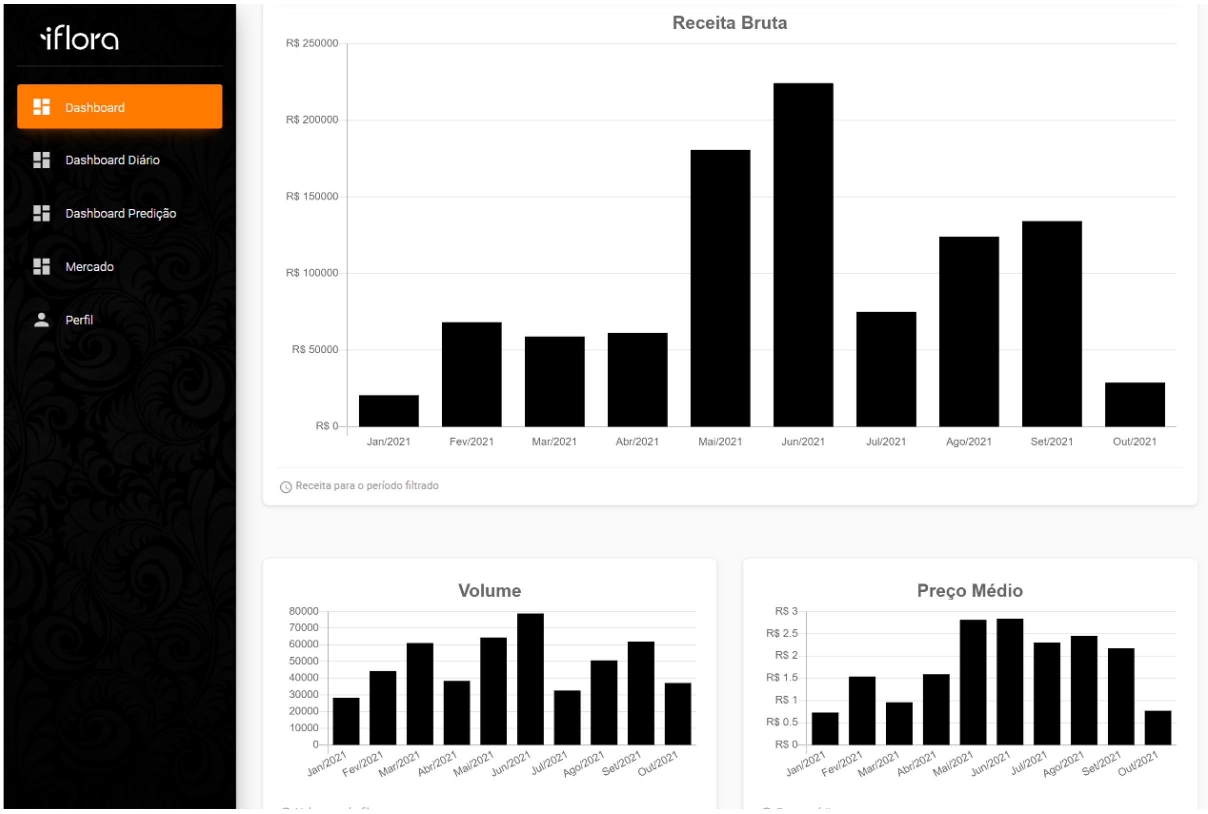

Figure 15: MVP Web - sales dashboard: charts for the period of sales from Jan/2021 to Out/2021 for the product Rose Freedom of 60 centimeters.

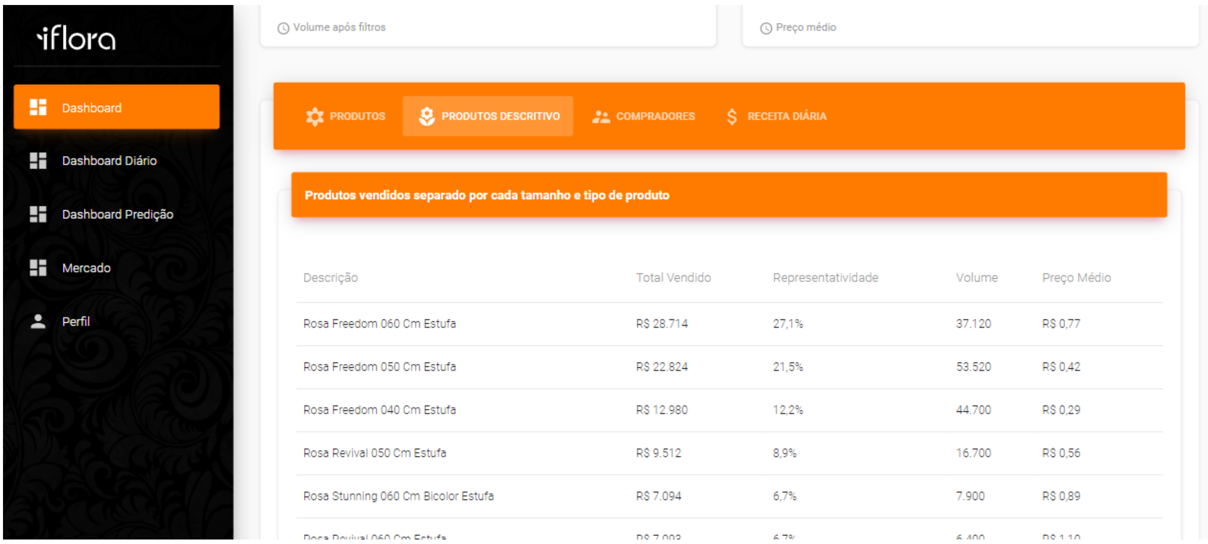

Figure 16: MVP Web - sales dashboard: tables for the period of sales of Oct/2021 for all the products

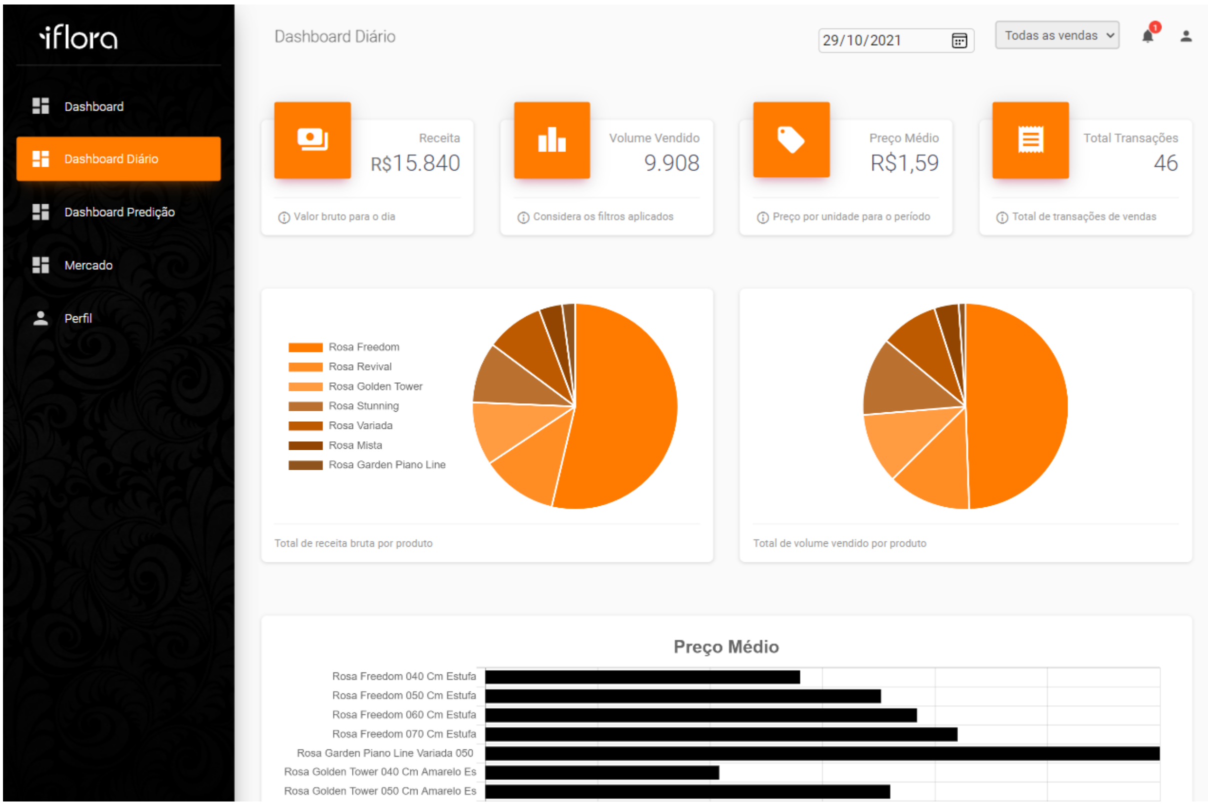

Figure 17: MVP Web - daily sales dashboard.

information.

The next dashboard, exhibited by Figure 18, presents the prediction dashboard. This page intends to present the output from the model, using the APIs developed. Basically, at the top of the page, the user can filter the product and the system will display the predictions for the next day. At the bottom of the page, there is the average price of the last 8 days of transactions for each one of the product sizes, grouped by type of sale (auction sale as K and direct sale as I).

Finally, the last dashboard, presented by Figure 19, is the market dashboard. It presents the stock and discard data from the entire cooperative and can be used for strategic decisions. The grower can explore different products and check its availability and annual performance.

5.4 Mobile Application for Final Consumer

This section includes all the development to create the mobile application for the final user.

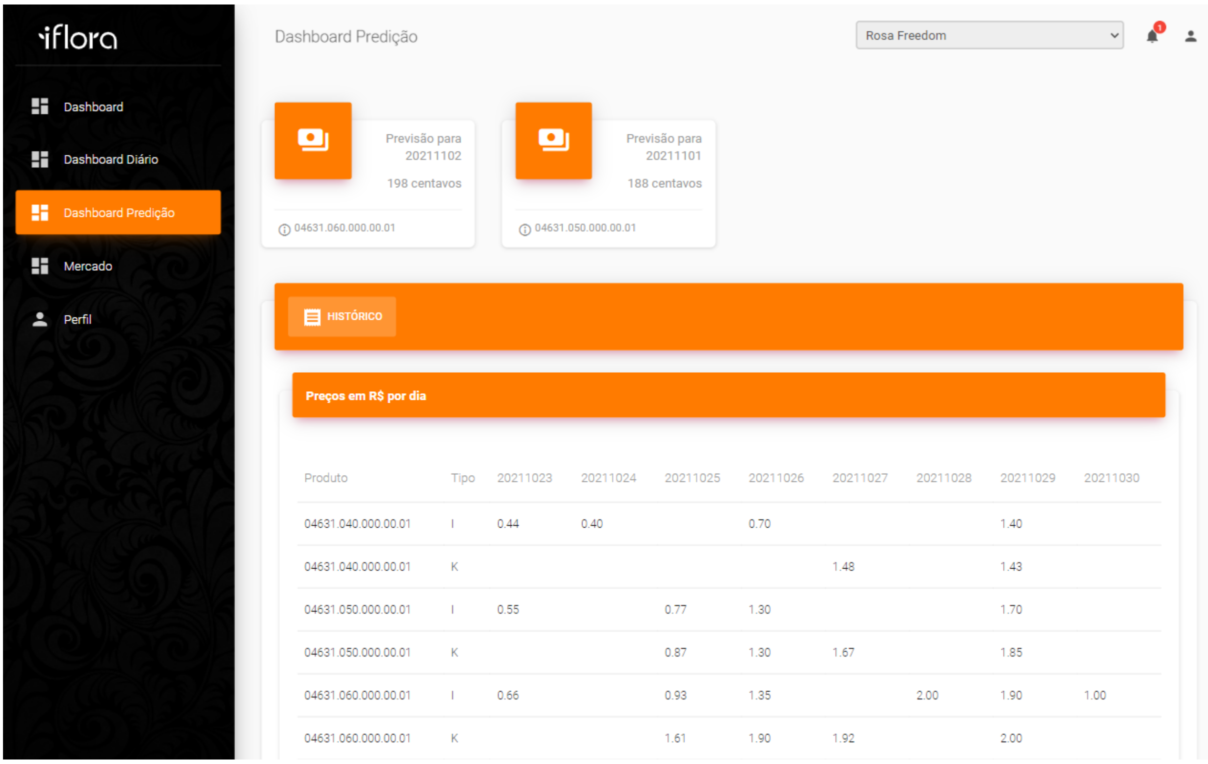

Figure 18: MVP Web - prediction dashboard.

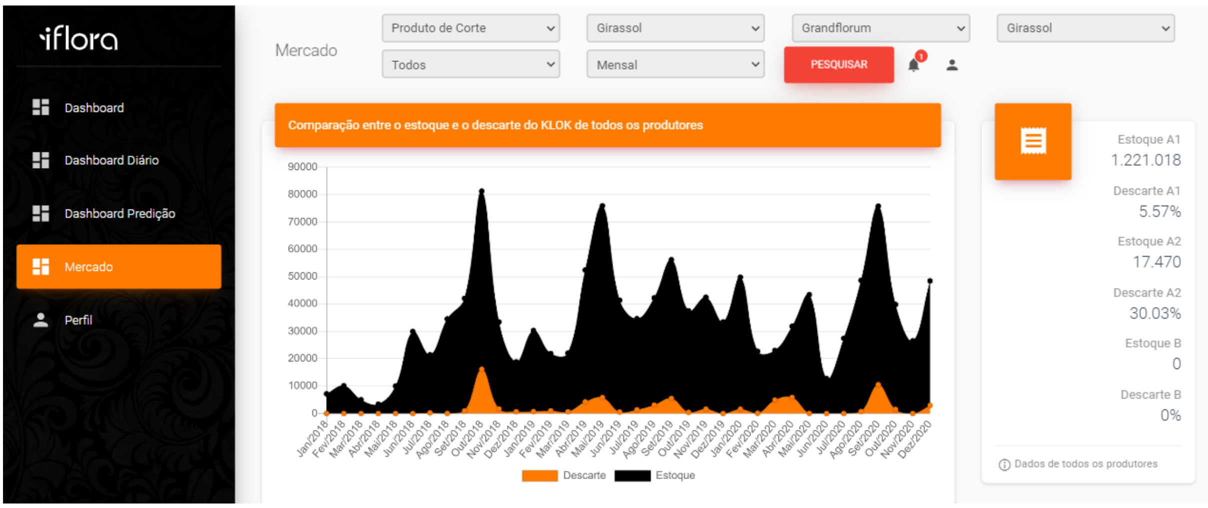

Figure 19: MVP Web - market dashboard, filtered by the product Sunflower.

5.4.1 Data Modeling

With a relational database, different tables were created and implemented to manage the connections between the data structures of the mobile application.

The mobile structure is presented by Figure 20. There are 6 tables, of which 3 are facts and 3 are dimension tables. The data is obtained using the application, where currently, only one information is added manually (latitude and longitude for the store address).

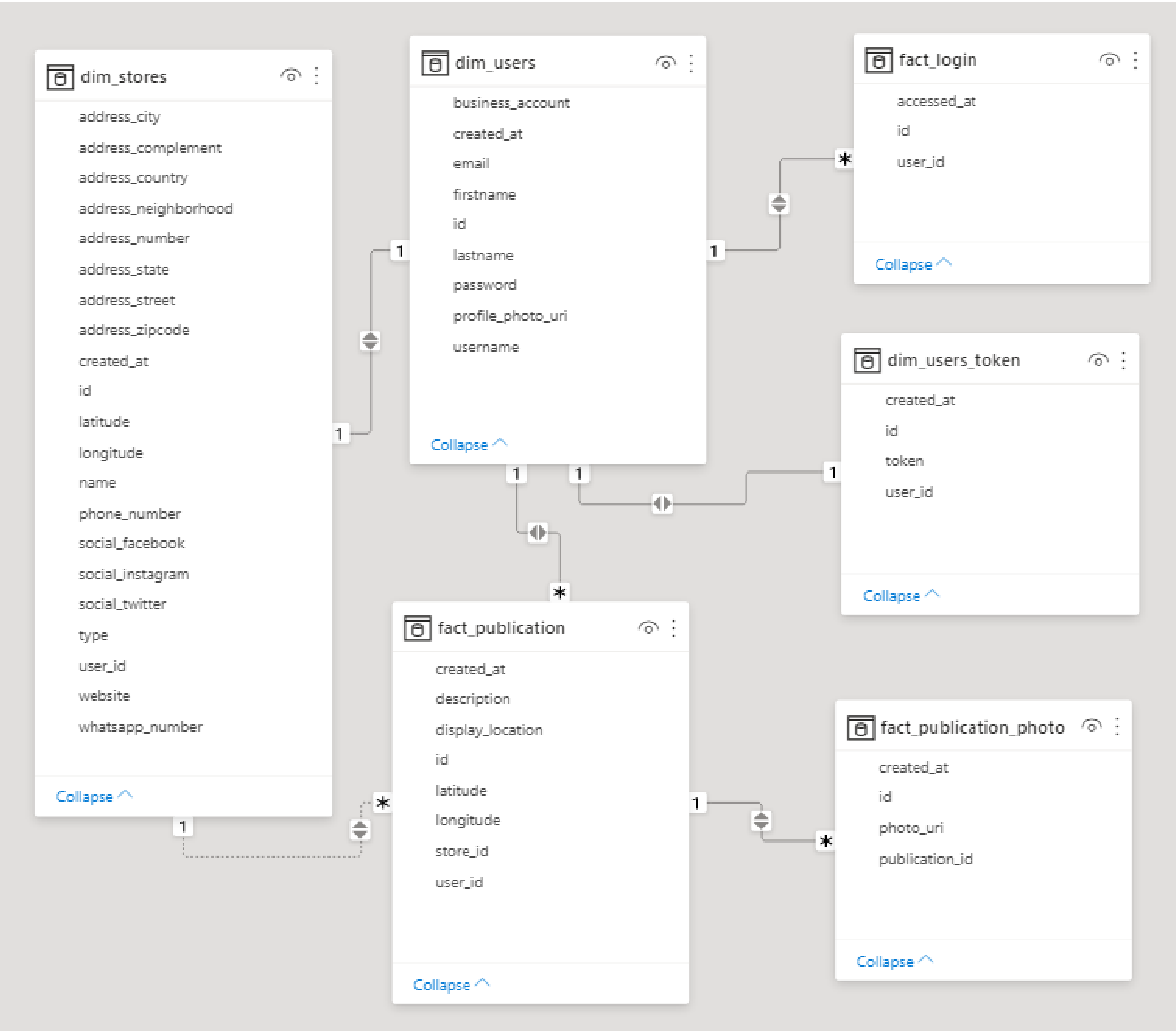

Figure 20: Relational database for the mobile application.

5.4.2 API Reference

Params flagged with a * are required. All params are shared with Json protocols, except the images.

- POST /registration

Primary endpoints for creating new users. POST params should be identified for a single user. If all required params, the server response will be that the user was created. Params needed:

  - *registration[email]: allowed values include e-mails that not are subscribed (if the e-mail is already in the database, an error is returned);
  - *registration[username]: allowed values include usernames that are not already used (if found in the database, an error is returned);
  - *registration[firstname]: text;
  - *registration[lastname]: text;
  - *registration[password]: text that are saved as a new password hash using a strong one-way hashing algorithm.

- POST /authentication

  Authentication required to use the application. If the params match to those in the database, return a unique token for the user and the user data. Params needed:

  - *authentication[email]: text;
  - *authentication[password]: text.

- POST /change_password

  This function is used to change the user password. First, the user can send his e-mail, if found in the database, the API can send a four-digit code for the e-mail, and the user has to confirm this number in the application. Then, if the pin is correct, the user can change his password.

  - *change_password[email]: text;
  - *change_password[pin_code]: four digits int;
  - *change_password[new_password]: four digits int;

- POST /registration_store (Auth required)

  Primary endpoint for creating a new store (business account), associated with an user. If all required params, the server response will be that the store was created. Params needed:

  - *registration_store[token]: information automatic added by the application, to identify the user;

  - *registration_store[type]: text;
  - *registration_store[name]: text;
  - *registration_store[whatsapp_number]: text;
  - *registration_store[website]: text;
  - *registration_store[address_street]: text;
  - *registration_store[address_number]: text;
  - registration_store[address_complement]: text;
  - *registration_store[address_zipcode]: text;
  - *registration_store[address_neighborhood]: text;
  - *registration_store[address_city]: text;
  - *registration_store[address_state]: text;
  - *registration_store[address_country]: text;
  - *registration_store[social_instagram]: text;
  - *registration_store[social_twitter]: text;
  - *registration_store[social_facebook]: text;

- POST /edit_profile (Auth required)

  Allow the user to edit his profile information. Params needed:

  - *edit_profile[token];
  - *edit_profile[query_type]: identification of a POST query. Allowed value: edit_photo, edit_business_account, edit_profile;
  - edit_profile[file]: image;
  - edit_profile[firstname]: text;
  - edit_profile[lastname]: text;

- POST /publication (Auth required)

  Primary endpoints for creating new publications (or modifying existing ones). Params needed:

  - *publication[token];
  - *publication[query_type]: identification of a POST query. Allowed values: send_pu- blication, edit_publication

  - *publication[description]: text;
  - *publication[display_location]: 1 if the user shares the photo location, otherwise 0;
  - publication[latitude]: numeric information obtained by the map;
  - publication[longitude]: numeric information obtained by the map;
  - publication[store_id]: optional information if the user select the store that he acquired the product displayed.

- POST /publication_photo (Auth required)

  Primary endpoints for adding photos to an publication. Params needed:

  - *publication_photo[token];
  - *publication_photo[publication_id]: the id of the publication (used to display all details of a publication).
  - *publication_photo[file]: image;

- GET /publication

  Primary endpoints for retrieving publications. Params needed:

  - *publication[token];
  - *publication[query_type]: identification of the screen retrieving the publications. Allowed values: get_publications_map, get_publications and get_publication_details
  - publication[latitude_top]: maximum latitude to filter the publications;
  - publication[latitude_bottom]: minimum latitude to filter the publications;
  - publication[longitude_top]: maximum longitude to filter the publications;
  - publication[longitude_bottom]: minimum longitude to filter the publications;
  - publication[limit_begin]: used to get slices of the response;
  - publication[limit_end]: used to get slices of the response;
  - publication[specific_user]: if equal to 1, will return only the publications for the user id requested;
  - publication[requested_user_id]: the user owner of the publications;
  - publication[publication_id];

- DELETE /publication (Auth required)

  Delete a publication. Authenticated users must own the publication.

  - *publication[token];
  - *publication[publication_id];

- GET /stores

  Primary endpoints for retrieving stores. Params needed:

  - *stores[token];
  - *stores[query_type]: identification of the GET query. Allowed value: get_stores_map, get_stores_search, get_store_details;
  - stores[latitude_top];
  - stores[latitude_bottom];
  - stores[longitude_top];
  - stores[longitude_bottom];
  - stores[store_id];
  - stores[text_researched] text;

- GET /users

  Primary endpoints for retrieving user details and username availability. Params needed:

  - *users[token];
  - *users[query_type]: identification of the GET query. Allowed value: get_user_details, check_username_availability;
  - *users[requested_user] id of the user;
  - *users[requested_username] username requested (used in the registration);

5.4.3 User Interface

To validate the project requirements, the mobile application was created from the beginning, in a configuration that would allow access for registered or anonymous users. This configuration allows the user to quickly have an overview of the main functions of the application, diminishing the churn for the users that do not want to create a new account. Figure 21 presents the initial screen of the application. It is possible to notice that users can access the application with their credentials or take a look at the app (without all functions available) without the need of logging in. Still, in Figure 21, the second screen presents the register screen and the last one presents the timeline screen. The timeline screen presents the publications from all users, ordered by chronological order. The objective of the timeline screen is to allow users to interact with each other and, ideally, there would be many publications about plants and flowers from flower shops to final consumers.

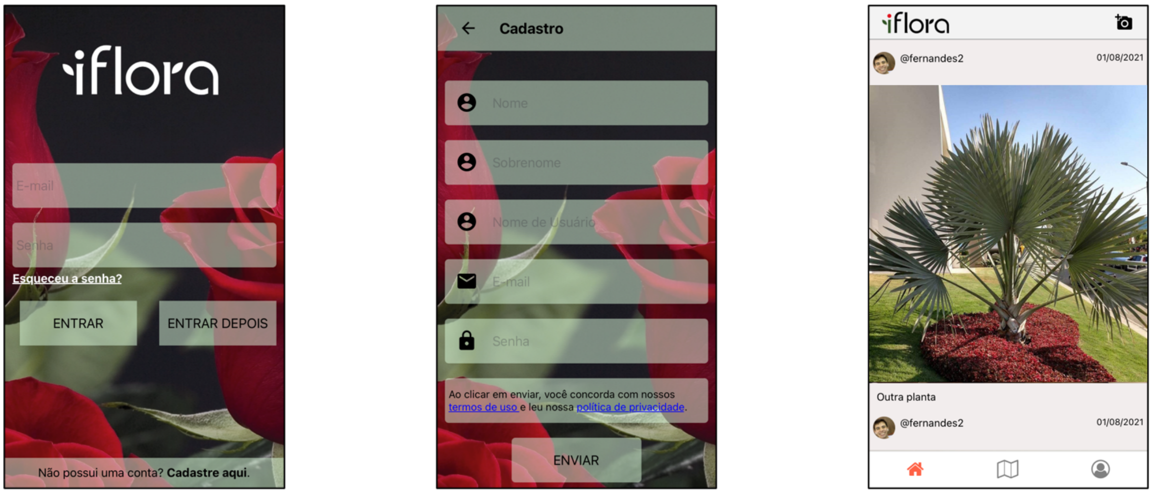

Figure 21: MVP Mobile - screens from the left to the right: initial screen, register screen and timeline screen.

Following the user journey, Figure 22 presents the different functions of the application. The first screen allows the user to create a new publication. First, it is possible to take a photo with the application or get one from the gallery. Then, the user can add a description of the image and choose the store from which the product was acquired. Finally, the user can add the location where the photo was taken, using the GPS of the smartphone. The second screen presents the map, where all publications with locations are presented, as well as some stores registered in the app. Each pin in the map represents a publication that can be clicked to get more details, as presented in the last screen, where a pin of a store was clicked and more detailed information is available.

Then, Figure 23 presents the profile screen. It contains all user publications and information. The settings screen, the second one in the figure, allows the user to log out from the application, modify his profile, or create a business account. The business account, presented by the last two screens of the figure, is to transform the user into a business profile, where the company would be able to access the app and publish on behalf of that account. The idea was to test if users (from flower shops to growers) would be

Figure 22: MVP Mobile - screens from the left to the right: add publication screen, map screen, store screen.

interested in publishing for their companies.

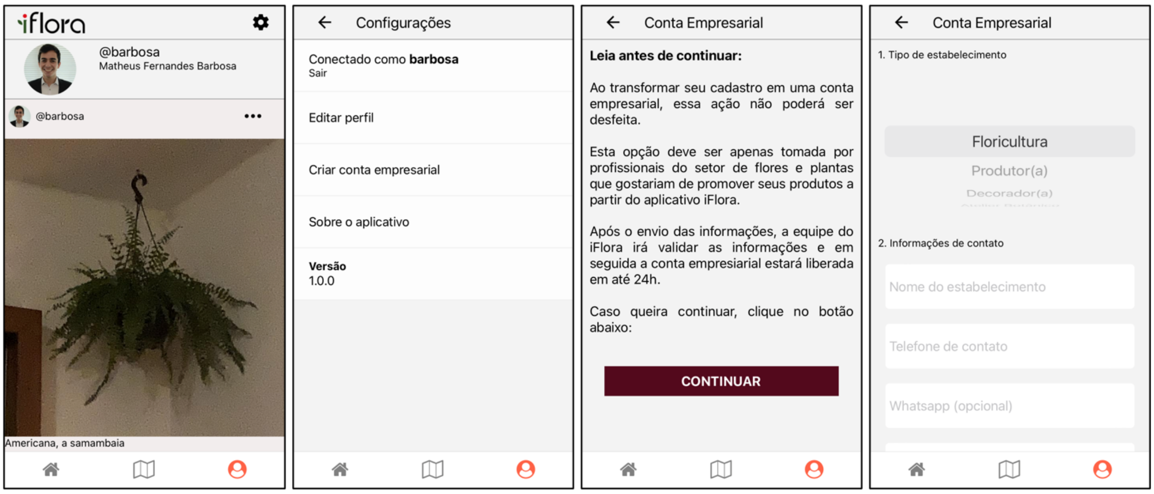

Figure 23: MVP Mobile - screens from the left to the right: profile screen, settings screen, and business account screen.

A mobile application was created to test the hypothesis: users would be interested in using an app dedicated to the niche of flowers and plants if they could interact with growers and flower shops. On the other hand, the companies could use the application to promote their products and create a new channel of interaction with the final user.

6 RESULTS

To evaluate the performance of this project, a qualitative approach was used to obtain feedback from the users on the experience in the website and mobile application, as well as about the intention to keep using the service or product. And, for the price predictor, a quantitative analysis was performed comparing the different models developed.

6.1 Web Application Reviews

The section of the web application is divided into two parts: first, the output from the survey that was applied to growers about the usability of the system. The second part is a review of the price predictor algorithms.

6.1.1 Questionnaire with users

To review the implementation of the web dashboard, the system was deployed online with the name iFlora and it was tested with some people from the company that provided the data. In total, 3 people were interviewed and replied to the survey.

The first question was intended to test the system's usability. According to the survey, 100% of the responses agreed that the system was easy to navigate and they were able to find what they needed without problems. Figure 24 presents the results of the question.

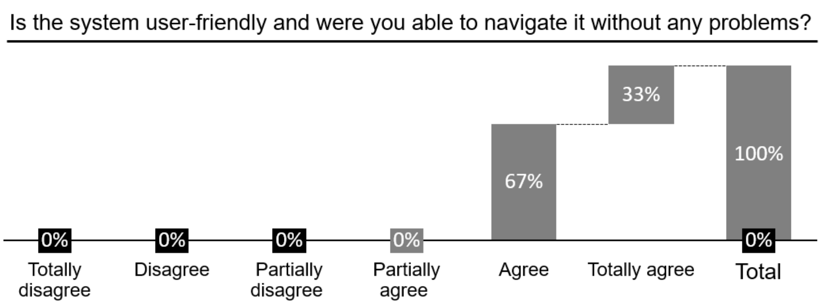

Figure 24: Web survey - responses about the usability of the application (3 answers).

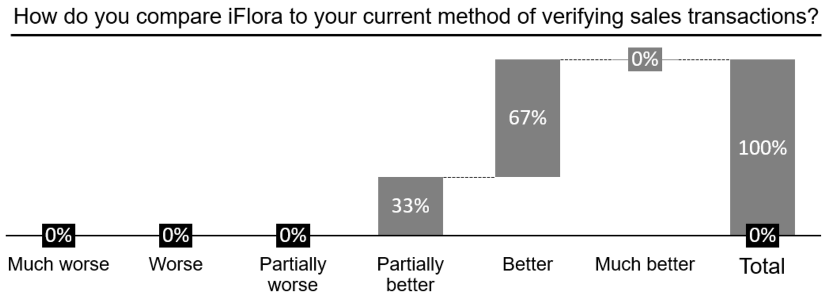

Figure 25: Web survey - responses for comparison with current system (3 answers).

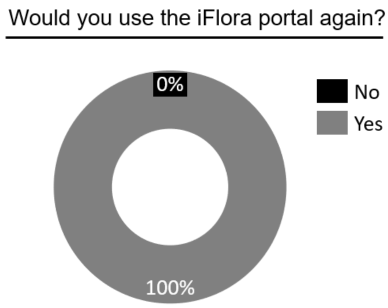

Figure 26: Web survey - responses for the willingness to use the application again (3 answers).

Then, the next question asked the user to compare the web system with his current system for checking the transactional data. Figure 25 shows the results for this question. The next question inquired if the user would use the application again. Figure 26 presents the results for this question.

The following questions require that the respondent rate on a scale from 1 to 5 some features that were already implemented in the system and some that are on the pipeline of production. On the scale 1 means that the function “is not important” and 5 means that it is “very important”. The intention of these questions was to verify the importance of each component. Figure 27 presents the affirmations and responses for each affirmative.

6.1.2 Price predictor analysis

The accuracy of the price predictor was evaluated by comparing the models and the real prices for the period, using the Walk Forward validation, as discussed in the development chapter.

For the purposes of this project, the first model was developed on data from 2017 and was then tested on data from March 2018. As such, the outputs of the first model for 2018 are out-of-time for previous years (i.e. 2017), and out-of-sample for transactions

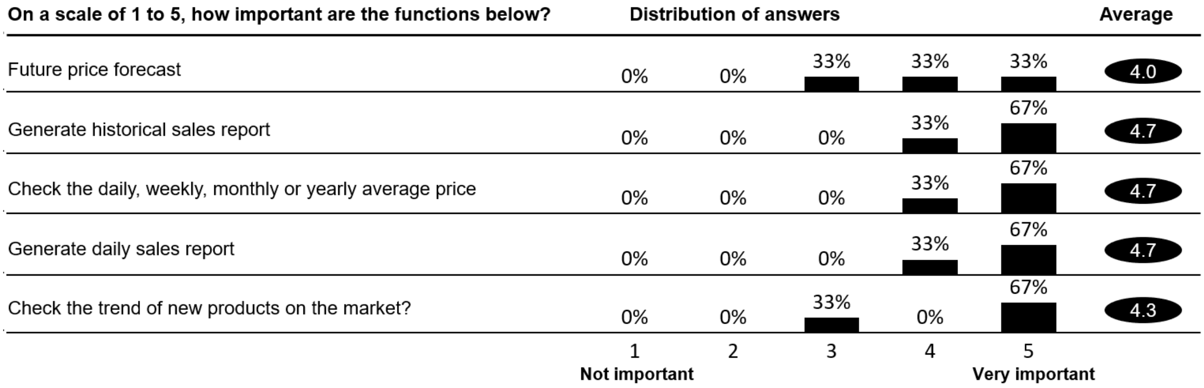

Figure 27: Web survey - responses for features prioritization (3 answers).

Test Period	Naive	Random Forest	Prophet	SARIMA	LSTM
19/Feb/18 to 07/Mar/18	0.2009	0.2163	0.5907	0.1964	0.2044
18/May/18 to 11/Jun/18	0.2607	0.2669	0.6709	0.2505	0.2728
08/Oct/18 to 26/Oct/18	0.2276	0.1926	0.3914	0.2707	0.2167
15/Feb/19 to 06/Mar/19	0.1898	0.1734	0.4788	0.2054	0.2202
27/May/19 to 11/Jun/19	0.2196	0.2843	0.3232	0.2368	0.2210
09/Oct/19 to 30/Oct/19	0.1013	0.1258	0.2619	0.1148	0.0903
12/Feb/20 to 06/Mar/20	0.2483	0.2783	0.2891	0.2634	0.2562
26/May/20 to 11/Jun/20	0.5356	0.8222	1.0432	0.4967	0.5421
13/Oct/20 to 30/Oct/20	0.2161	0.2202	0.2492	0.2545	0.2214
17/Feb/21 to 05/Mar/21	0.4699	0.5606	0.5845	0.5421	0.4885
25/May/21 to 11/Jun/21	0.4342	0.4102	0.4008	0.5238	0.4371
12/Oct/21 to 29/Oct/21	0.2650	0.3219	0.2762	0.3353	0.2777

Table 9: RMSE between the prediction of the models and the test real value. Data for rose Freedom of 60 centimeters.

that happened after 2017 (i.e. in 2018). Then, the model was reestimated using data from 2017 to 2018, and tested in October 2018, and so on.

It was selected 3 periods with 10 days per year for the analysis. The first period consists of 10 auction sales before International Women's Day (8th March), which is by default characterized by an increase in the average price some days before the holiday, followed by a decrease in the prices. The second period defined was with the days before Valentine's day and the third period defined was a month without holidays affecting the average price, the month of October. Table 9, 10 and 11 presents the different error rates for the models developed.

A deep dive in some of the cases was conducted by plotting the forecast for the auction price of each model, as presented by Figure 28. The figure illustrates the real average

Test Period	Naive	Random Forest	Prophet	SARIMA	LSTM
19/Feb/18 to 07/Mar/18	0.16	0.17	0.53	0.14	0.15
18/May/18 to 11/Jun/18	0.20	0.23	0.57	0.19	0.21
08/Oct/18 to 26/Oct/18	0.17	0.16	0.35	0.20	0.16
15/Feb/19 to 06/Mar/19	0.14	0.11	0.38	0.16	0.18
27/May/19 to 11/Jun/19	0.17	0.21	0.27	0.18	0.18
09/Oct/19 to 30/Oct/19	0.08	0.12	0.24	0.10	0.07
12/Feb/20 to 06/Mar/20	0.19	0.20	0.25	0.22	0.21
26/May/20 to 11/Jun/20	0.40	0.67	0.93	0.37	0.39
13/Oct/20 to 30/Oct/20	0.16	0.16	0.19	0.18	0.17
17/Feb/21 to 05/Mar/21	0.34	0.48	0.48	0.40	0.37
25/May/21 to 11/Jun/21	0.32	0.33	0.32	0.40	0.32
12/Oct/21 to 29/Oct/21	0.18	0.22	0.23	0.27	0.19

Table 10: MAE (R$) between the prediction of the models and the test real value. Data for rose Freedom of 60 centimeters.

Test Period	Naive	Random Forest	Prophet	SARIMA	LSTM
19/Feb/18 to 07/Mar/18	16.99	18.60	53.84	15.14	16.31
18/May/18 to 11/Jun/18	13.25	14.24	32.11	12.06	13.41
08/Oct/18 to 26/Oct/18	27.28	28.64	74.67	33.16	27.19
15/Feb/19 to 06/Mar/19	12.27	9.38	29.42	13.48	15.77
27/May/19 to 11/Jun/19	8.23	9.72	12.5	8.14	8.47
09/Oct/19 to 30/Oct/19	26.75	40.27	84.94	31.96	22.63
12/Feb/20 to 06/Mar/20	22.87	21.2	29.53	26.37	23.62
26/May/20 to 11/Jun/20	10.27	17.62	23.66	9.95	10.21
13/Oct/20 to 30/Oct/20	15.81	15.02	17.67	17.44	16.28
17/Feb/21 to 05/Mar/21	20.82	26.97	26.69	23.43	22.44
25/May/21 to 11/Jun/21	9.26	9.62	9.4	11.6	9.15
12/Oct/21 to 29/Oct/21	16.73	15.38	17.65	25.57	16.95

Table 11: MAPE (%) between the prediction of the models and the test real value. Data for rose Freedom of 60 centimeters.

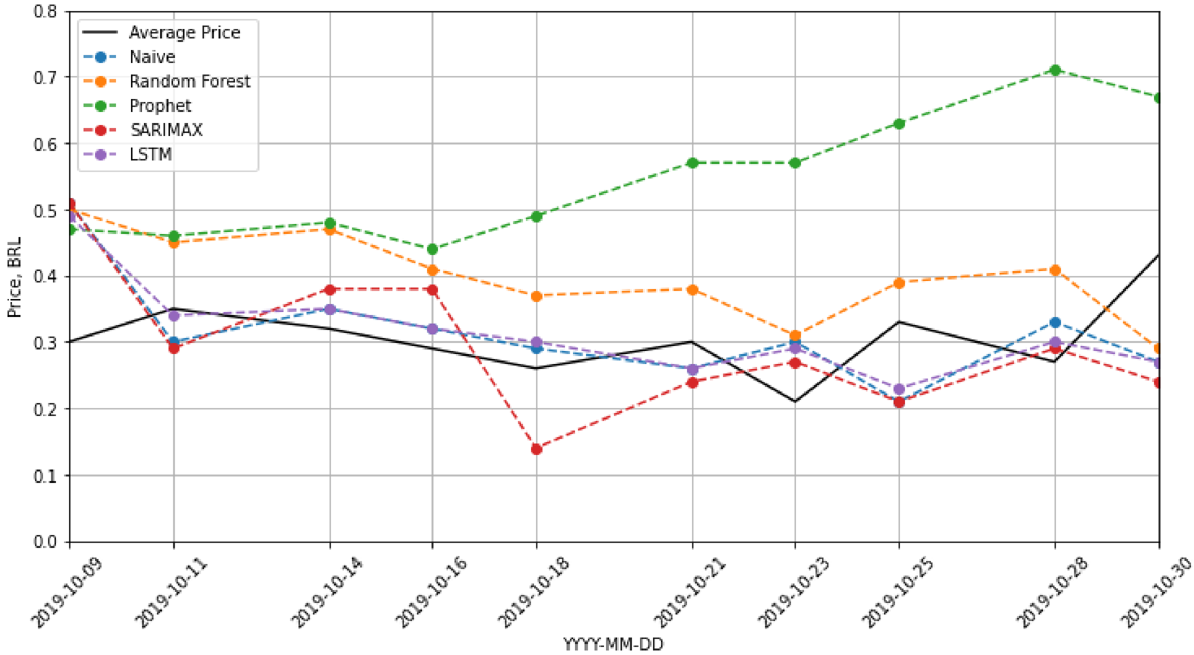

Figure 28: Price variation for the Rose Freedom 60 centimeters for the period of 09 to 30 of October 2019 including different model forests.

price of the auction and the price forecast by each model, given the day of the auction. For this specific period (month of October), the LSTM model presented the lowest RMSE value (equal to 0.0903).

Figure 29 illustrates the price for the days preceding Women's day in 2020. For this period, the model of the Naive presented the lower RMSE (0.2483). Figure 30 shows the price for the days before Valentine's Day in the 2021 holiday in which the model of the Prophet presented the lowest RMSE (0.4008).

6.2 Mobile Application Reviews

To review the results from the mobile application, the software was deployed in the Google Play Store and a questionnaire was applied to some users after they used the application. Although the application was ready also for publication in the Apple Store, once the React Native language works for both systems, the decision to restrict the application to one store was to have a deep test before going into production. In total, 16 people replied to the survey.

The first question was to test the system's usability. According to the survey, 100% of the responses agreed that the system is user-friendly and they were able to navigate on it without problems. Figure 31 presents the results of the question.

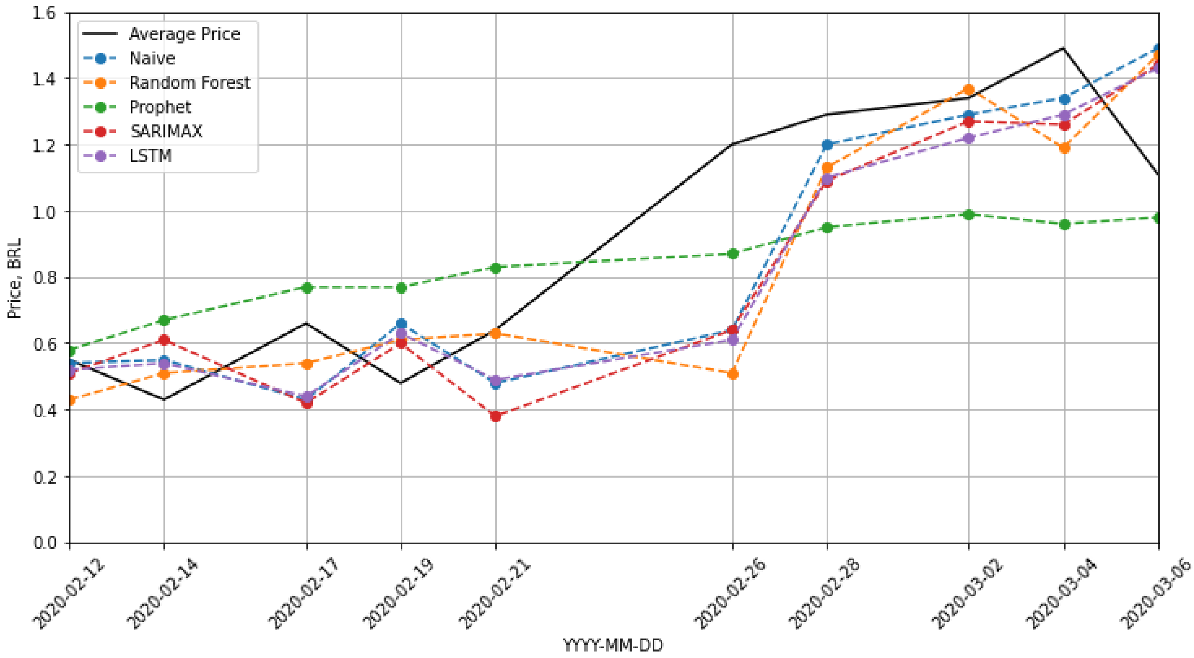

Figure 29: Price variation for the Rose Freedom 60 centimeters for the period of 02 February to 06 March 2019 including different model forests.

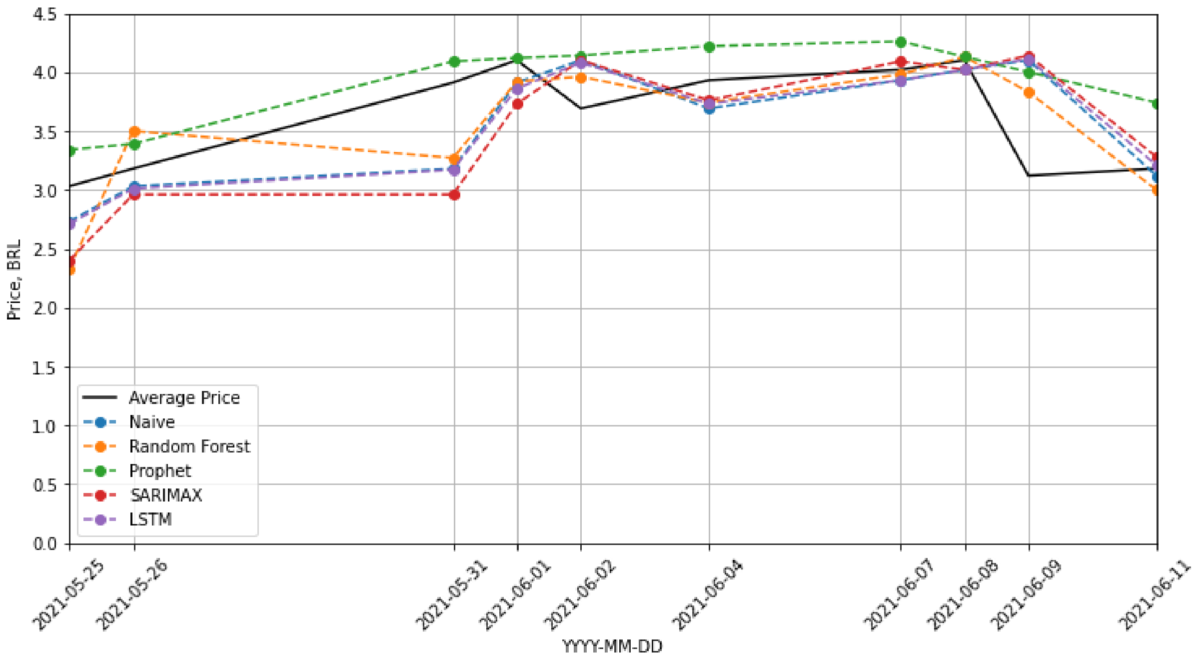

Figure 30: Price variation for the Rose Freedom 60 centimeters for the period of 25 May to 11 June 2021 including different model forests

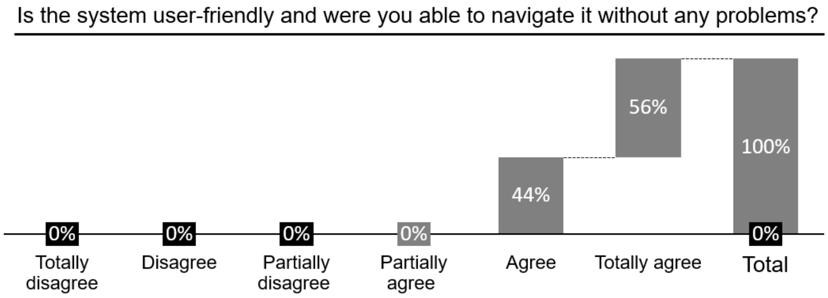

Figure 31: Mobile survey - responses about the usability of the application (16 answers).

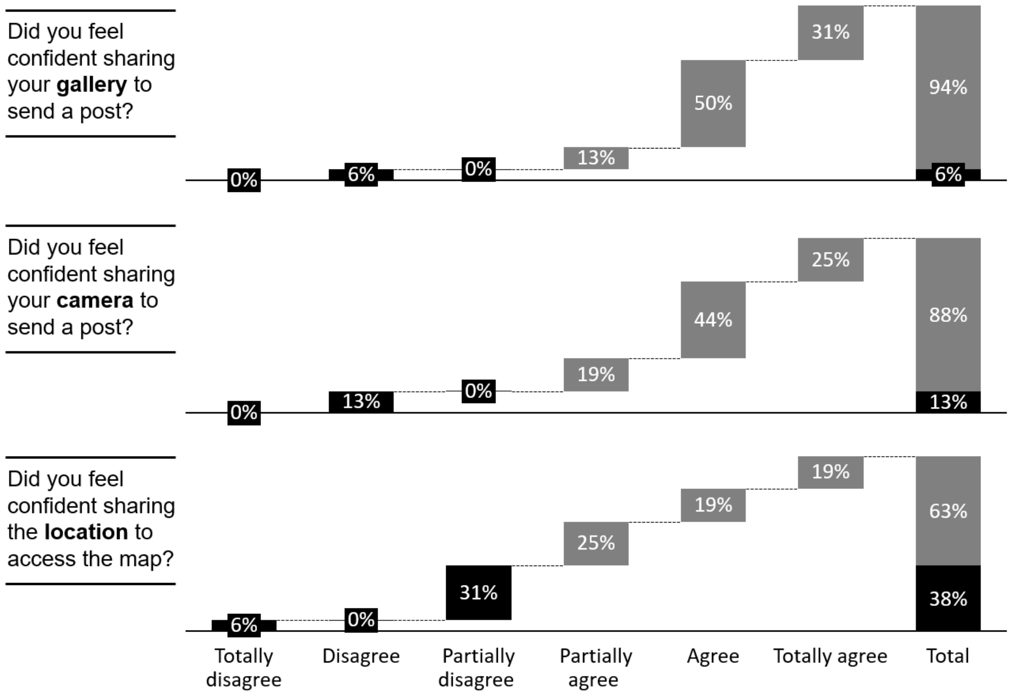

Figure 32: Mobile survey - responses for confidence in sharing gallery, camera and location (16 answers).

The following three questions were intended to test the feeling of safety using the application. First, it was asked if the user was confident when asked to share the gallery to send a publication. Then, it was asked the same question, but focusing on the camera permissions. The last one in this part was about if the user was confident sharing his mobile location to see the publications on the map. Figure 32 presents the results for these questions.

The next question inquired if the user would use the application again. Figure 33 presents the results for this question.

The following questions required that the respondent rate on a scale from 1 to 5 some features that were already implemented in the application and some that were on the pipeline for production, where the scale 1 means that the function " is not important"

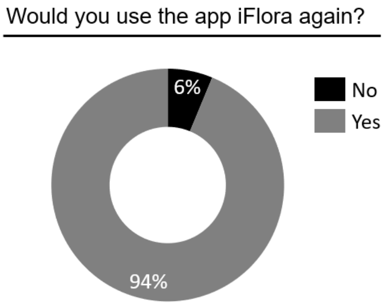

Figure 33: Mobile survey - responses for willingness to use the application again (16 answers).

and 5 means that it is “very important”. The intention for these questions was to verify the importance of each component. Figure 34 presents the affirmation and responses for each affirmative.

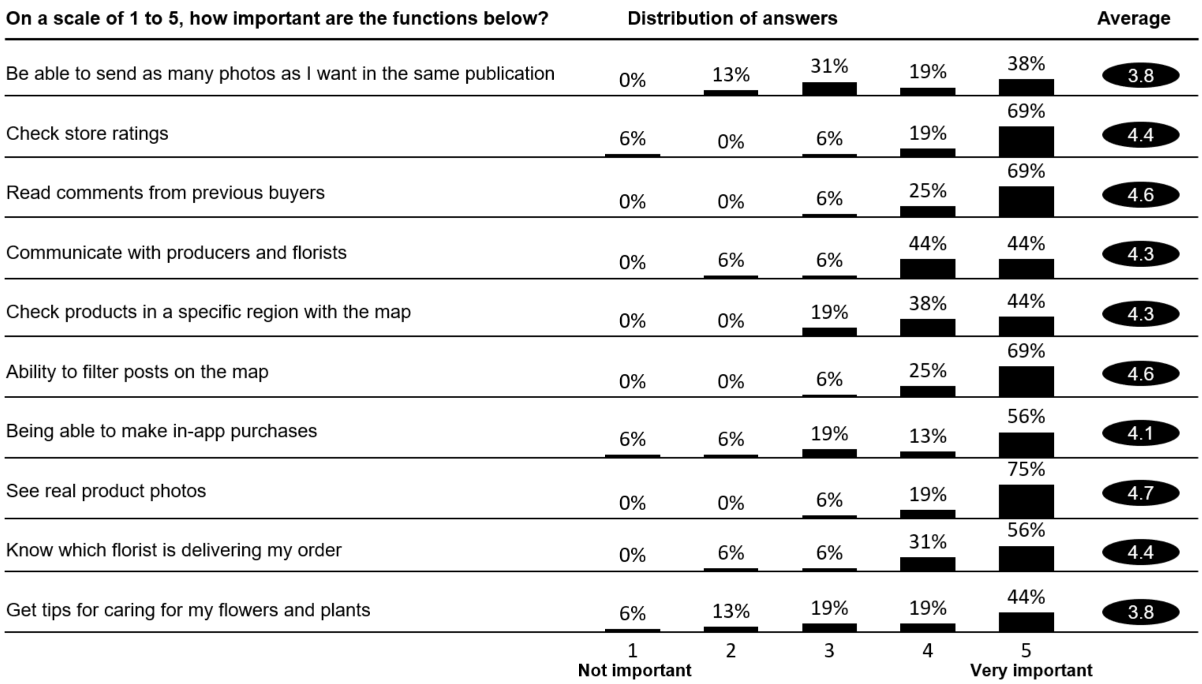

Figure 34: Mobile survey - responses for features prioritization (16 answers).

7 DISCUSSION

Before going deeper into the results of the project, it is important to remember what was the goal behind each one of the platforms. The website had the objective to test the concept: an online dashboard could add value for growers in their decision-making routine and, to do so, some visualizations were created together with an algorithm to help the price definition. The mobile application, on the other hand, has the goal of connecting growers and consumers in one single platform, to increase the engagement and interaction between these two sides of the value chain.

Given that context, the prioritization of the development phase was to achieve the objectives: focusing on the quantitative part of the growers' platform (including the price predictor and the sales report) and on the qualitative aspect of the mobile application, including creating a platform to start getting data from this industry.

7.1 Web application

To test the prototype, it was not possible to simply send the survey to some users to test the system. An introduction was needed and an explanation of each one of the pages was requested. This was needed once it is a new platform and there is a learning curve for the new users. If implemented for a broader set of consumers, some tutorial videos should be developed explaining each one of the functions.

With the user interviews, it is possible to see that the system was easy to use and the users could navigate in the pages without problems, as presented by Figure 24 with 100% of the users that agreed with this point. Although the price predictor was not the most useful tool for the system, the other applications, like the sales report and market trend, were very insightful, according to the interviews. Figure 27 presented that the sales report, historical and daily price, were the most important functions of the application since all the replies voted these features with grade 4 or 5.

Test Period	Metric	Naive	Random Forest	Prophet	SARIMA	LSTM
Nov/20 to Oct/21	RMSE	0.3225	0.3532	0.5303	0.3522	0.3300
	MAE	0.23	0.24	0.42	0.26	0.24
	MAPE	17.93	17.87	34.61	20.93	18.29
Jan/19 to Dec/19	RMSE	0.1983	0.1893	0.2616	0.2178	0.199
	MAE	0.15	0.14	0.20	0.17	0.15
	MAPE	21.62	21.12	29.76	23.93	21.33

Table 12: Metrics for a 12-months time period for the rose Freedom.

The last question was an open question requiring that the user presents some new features that they would like to see in the system. Some interesting replies included adding functions that already exists in the cooperative system, like the market price (currently the cooperative shows it in a tabular way). Also, some recommendations were related to given an introductory explanation about the tools for each page and to add a search button when the select presents many options (e.g., when it displays all the products available).

7.2 Price predictor analysis

Although in the short-term, as presented by Table 9, 10 and 11, the models had different results of performance, in the long-term the results are close to the same, as presented by Table 12. For 2021, a year with high variation of the price, the Naive Forecast achieved the best RMSE metric (RMSE equals 0.3225), but for the year 2019, the Random Forest achieved the best result (RMSE equals 0.1893), 5% better than the baseline defined by the Naive Forecast. All models, not including the Prophet model, achieve similar MAE results, which demonstrate a type of randomness of the results, given that the average price distortion was close to the same.

With these results, it is possible to conclude that, still, the Naive Forecast is the process that had the best performance among the models. Although the data follows some tendencies in the long term, as presented by Figure 12, this is true only in the monthly scenario and for some specific holidays. The pandemic situation distorted the historical data and, because of that, there is the need for some months so the market can return to normality.

Comparing the results from the different predictors, Figure 28 shows that the Random Forest is the only model that does not act only following the last day price but anticipates some price increases, like the data point from 23/Oct to 25/Oct. This can be associated with some variables that are used in this model, like the total volume sold by the grower. However, this better result, if compared with the other models, is not enough to fulfill the computation cost to implement this system, once the 5% better estimation is not enough to cover the investment needed, according to the growers.

7.3 Mobile application

To receive answers for the usability, functions, and features of the mobile application, the app was shared in Whatsapp groups, paid Instagram ads, and posted on educational websites (Politécnicos and Polishare). From all those channels, the Whatsapp application was the best way to get users, since the friction of the test: the user needed to download the application and (optionally) register an account before being allowed to answer the survey.

From the 16 replies obtained, it is possible to see that people enjoyed the application, as presented by Figure 33 with 94% of the respondents saying that they would use the app again and 100% agreed that the usability was good, as presented in Figure 31.

When defining what kind of information the user is willing to share without any friction, it was possible to see by Figure 32 that 94% were confident sharing the gallery, 88% sharing the camera, and only 63% sharing the location. These numbers illustrate that an application should request these types of data only in the situation of really need since the unwanted requests can represent more users stopping using the application.

Finally, with the scale question, represented by Figure 34, it is possible to define what should be prioritized and what should not in the roadmap of new functions for the app: technological functions like sending many photos or a dedicated blog for tips about taking care of flowers and plants are the less essential functions according to the survey. On the other hand, the need for the users to see only real photos, be able to apply filters in the map, and read the comments from previous buyers are the most important functions that should be prioritized, with average grade of 4.7, 4.6 and 4.6, respectivelly.

The last question was an open question requiring that the user comments about what he misses in the application. 2 out of the 16 responses mentioned that the purpose of the application was not clear, which indicates that some information content should be added in the initial screens, as well as in the app description. Other requests include an option to comment on the other's user publications, add photos for a pre-defined period, and bookmark posts. All these comments are useful for the prioritization and pipeline for future functions.

8 CONCLUSION

The web system developed had the objective to be more intuitive and give more insights for the growers than the solution that there is today in place with the cooperative system. With the results of the survey for the web system, conclude that it can overperform the current methodology of checking the sales transaction, with the highlight for the pages presenting the historical and daily sales and the functions of checking the daily, weekly, and monthly average price. With these results, the project requirements SR0 and SR1 were achieved.

For the algorithm of the price forecast, the approach used was to define a baseline (according to the problem definitions), where the selected one was the Naive Forecasting, and then build other models to improve the metrics. For the situation of normality (before the covid pandemic), the model Random Forest presented better results if compared with the baseline (however only 5% better in the RMSE metric), but for 2020-21, the Naive presented the best result comparing the RMSE among the models. It is possible to see that all the models (except the Prophet) had a difference of no more than 3 cents when the metric MAE is used for comparison. Given these results and discussion made, although the project requirement SR2 was technically built, the project requirement SR3 is only reached in the situation of normality, that is, with seasonal data, like the year of 2019 and before. Hopefully, from 2022 on, this situation will be reached again and the models can obtain better results.

The results of the survey for the mobile application, conclude that people are willing to use new systems and test new functionalities since the majority of the responses would use the app again. The SR5 was achieved in the system development since the users do not need to create a user to access the application, and the SR4 was reached with the consent of the users. With the survey, it was possible to see that, although the majority of the users are confident sharing the gallery, camera, and location, this can be friction for some of them, so this data should be asked only in essential requests and the return for a declined from the user should be elaborated.

The SR6 and SR7 are external requirements that were used as the base for the building of the platforms. The LGDP requirements are respected given the policies created for both applications. Overall, the systems fulfilled the majority of the project requirements, and the considerations about the development are going to e discussed in the next chapter.

9 FINAL CONSIDERATIONS

From the first day of the project, the intention was to develop a system take could achieve scale and be implemented in production. For this reason, the project started with desk research and interviews then passed with a development phase and ended with a qualitative survey. With that in mind, the next steps include using the feedback from the users to plan the pipeline to reach production.

The next step for the web application is to reach other growers to connect their data in the system and create new features given the feedback from the survey. For the mobile application, the points obtained from the survey should be used to prioritize the next developments and move the application from the test environment to production. Also, the launch of the product for the IOS devices should be prioritized.

For the price forecast, more interviews are needed to understand if the forecast weekly could be useful, and if so, the models should be adapted for this scenario. Another opportunity is to create a partnership with the cooperative or a group of growers, to have more transaction data for the analysis.

REFERENCES

1 BEAN, R. *Why Is It So Hard to Become a Data-Driven Company?* 2021. Harvard Business Review. Available at: ⟨hbr.org/2021/02/why-is-it-so-hard-to-become-a-data-driven-company⟩. Accessed on: 30/08/2021.

2 LU, Y. *Data-Driven Decision Making in Auction Markets*. Thesis (Doctor of Philosophy) — Erasmus Research Institute of Management, 2014. EPS-2014-314-LIS.

3 BRYNJOLFSSON, E.; HITT, L. M.; KIM, H. H. Strength in numbers: How does data-driven decisionmaking affect firm performance? *International Conference on Information Systems*, v. 1, p. 541–558, 2011.

4 HENKE, N. et al. *The Age of Analytics: Competing in a Data-Driven World*. 2016. McKinsey Global Institute. Available at: ⟨www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world⟩. Accessed on: 10/07/2021.

5 PEETERS, M. P.-S. P.; SCHNITZLER, C. *A data-driven approach to improving net working capital*. 2021. McKinsey and Company. Available at: ⟨www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/a-data-driven-approach-to-improving-net-working-capital⟩. Accessed on: 10/07/2021.

6 GJOLBERG, O.; STEEN, M. Forecasting prices at the dutch flower auctions: Calendar patterns and other time regularities. 1997. Available at: ⟨orgprints.org/id/eprint/6192/1/D9710.PDF⟩. Accessed on: 24/07/2021.

7 POZZEBON, M.; HECK, E. van. Local adaptations of generic application systems: The case of veiling holambra in brazil. *Journal of Information Technology*, v. 21, p. 73–85, 2006.

8 NEVES, M. F. et al. *Mapeamento e Quantificação da Cadeia de Flores e Plantas Ornamentais do Brasil*. 2015. Ministério da Agricultura, Pecuária e Abastecimento.

9 JUNQUEIRA, A. H.; PEETZ, M. Brazilian consumption of flowers and ornamental plants: habits, practices and trends. *Ornamental Horticulture*, Brazilian Society of Floriculture and Ornamental Plants, v. 23, p. 178–184, 2017. ISSN 2447-536X.

10 JUNQUEIRA, A. H.; PEETZ, M. D. S. Mercado interno para os produtos da floricultura brasileira: características, tendências e importância socioeconômica recente. *Revista Brasileira de Horticultura Ornamental*, v. 14, p. 37–52, 2008.

11 SCHOENMAKER, K. *O Mercado de Flores no Brasil*. 2021. Instituto Brasileiro de Floricultura. Available at: ⟨www.ibraflor.com.br/numeros-setor⟩. Accessed on: 25/08/2021.

12 VIZZOTO, F. *E-commerce in the Brazilian Flower Sector: Characterization, analysis and trends*. Dissertation (Master) — Wageningen University, 2011.

13 ANACLETO, A. Flowers supply chain in paraná coast - brazil: an application of miles and snow taxonomy. *Revista Capital Científico - Eletrônica*, v. 14, 2016.

14 GASQUES, J. G. et al. *Projeções do Agronegócio Brasil 2019/20 a 2029/30*. 2020. Ministério da Agricultura, Pecuária e Abastecimento.

15 PAIVA, P. D. de O. et al. Flower and ornamental plant consumers profile and behavior. *Ornamental Horticulture*, Brazilian Society of Floriculture and Ornamental Plants, v. 26, p. 333–345, 2020. ISSN 2447-536X.

16 OSORIO, R. F. *Navegar é preciso, digitalizar é essencial. - A nova Revolução Comercial*. 2021. Veiling Holambra. Available at: ⟨www.veiling.com.br/noticias/navegar-e-preciso-digitalizar-e-essencial-a-nova-revolucao-comercial⟩. Accessed on: 20/07/2021.

17 LU, Y. et al. Dynamic decision making in sequential business-to-business auctions: A structural econometric approach. *Management Science*, v. 65, p. 3853–3876, 2019.

18 HUMMEL, M.; MIGUEL, L. A. P. Gerando valor na cadeia de flores de corte no mercado brasileiro. *Revista Práticas em Contabilidade e Gestão*, v. 5, 2017.

19 ANACLETO, A. et al. Between flowers and fears: the new coronavirus pandemic (covid-19) and the flower retail trade. *Ornamental Horticulture*, Brazilian Society of Floriculture and Ornamental Plants, v. 27, p. 26–32, 11 2020. ISSN 2447-536X.

20 HOREN, L. van; BAC, A. *Kleur bekennen in de sierteelthandel*. 2021. Rabobank. Available at: ⟨research.rabobank.com/far/en/sectors/fresh-produce/kleur-bekennen-in-de-sierteelthandel.html⟩. Accessed on: 26/07/2021.

21 DAVENPORT, T. H.; REDMAN, T. C. *Digital Transformation Comes Down to Talent in 4 Key Areas*. 2020. Harvard Business Review. Available at: ⟨hbr.org/2020/05/digital-transformation-comes-down-to-talent-in-4-key-areas⟩. Accessed on: 26/08/2021.

22 SALVINI, G. et al. Enhancing digital transformation towards virtual supply chains: a simulation game for dutch floriculture. *Production Planning & Control*, 2020.

23 MONTECCHI, M.; PLANGGER, K.; WEST, D. C. Supply chain transparency: A bibliometric review and research agenda. *International Journal of Production Economics*, Elsevier, v. 238, p. 108152, 2021. ISSN 0925-5273.

24 WASSENAER, L. van et al. *Digitale transformatie in de sierteelt*. 2020. Wageningen University. Available at: ⟨research.wur.nl/en/publications/digitale-transformatie-in-de-sierteelt-van-registratiedruk-naar-s⟩. Accessed on: 20/08/2021.

25 RIBEIRO, V. S.; FINCO, M. V. A. Biodiesel e agricultura familiar no tocantins: Uma analise a partir da teoria dos conjuntos fuzzy. *Revista Esmat*, v. 6, p. 167, 2015.

26 BOYCE, C.; NEALE, P. *Conducting in-depth interviews: a guide for designing and conducting in-depth interviews for evaluation input*. 2006. Pathfinder International.

27 FATTAH, J. et al. Forecasting of demand using arima model. *International Journal of Engineering Business Management*, v. 10, p. 184797901880867, 1 2018. ISSN 1847-9790.

28 GRöMPING, U. Variable importance assessment in regression: Linear regression versus random forest. *The American Statistician*, v. 63, n. 4, p. 308–319, 2009.

29 TAYLOR, S. J.; LETHAM, B. Forecasting at scale. *PeerJ Preprints*, 2017.

30 HYNDMAN, R. J.; ATHANASOPOULOS, G. *Forecasting: principles and practice, 2nd edition.* [S.l.]: OTexts: Melbourne, Australia, 2018.

31 HUMMEL, M.; SILVA, A. A. da. Modelo de negócios em plataforma digital para comercialização de flores no brasil. *Navus: Revista de Gestão e Tecnologia*, v. 10, p. 1–17, 2020. ISSN 2237-4558.

32 RIJSWICK, C. van. *World Floriculture Map 2016: Equator Countries Gathering Speed.* 2016. Rabobank. Available at: ⟨research.rabobank.com/far/en/sectors/regional-food-agri/world_floriculture_map_2016.html⟩. Accessed on: 20/08/2021.

33 RIKKEN, M. *Cut Flowers and Foliage.* 2016. CBI Trade Statistics - Ministry of Foreign Affairs. Available at: ⟨www.cbi.eu/market-information/cut-flowers-foliage/trade-statistics⟩. Accessed on: 14/08/2021.

34 CARDOSO, B. F. et al. Trade dynamics in the italian floriculture sector within eu borders: A gravity model analysis. *Agris On-line Papers in Economics and Informatics*, Faculty of Economics and Management, v. 9, p. 23–32, 2017.

35 HOREN, L. van; BAC, A. *Volop kansen voor de Nederlandse sierteelt.* 2017. Rabobank. Available at: ⟨www.rabobank.nl/images/pdf_rabobank_volop_kansen_nederlandse_sierteelt_sep2017_29924822.pdf⟩. Accessed on: 28/07/2021.

36 FREDENBURGH, J. *The 4,000 mile flower delivery.* 2018. Made On Earth BBC. Available at: ⟨www.bbc.com/future/bespoke/made-on-earth/the-new-roots-of-the-flower-trade/⟩. Accessed on: 14/07/2021.

APPENDIX A – MARKET SIZE

Although ornamental plants are included in the agriculture sector, these products cannot be considered commodities, due to its wide variety "involving cut flowers, foliage, bulbs, flower pots, green plant pots, lining, and plants for landscaping, representing a world market with a turnover of around US$ 90 billion/year" [31], from which it is estimated that US$ 55 billion is related to floriculture and US$ 35 billion to the global nursery production [32].

"The commercial floriculture understood as the professional and business activity of production, commerce, and distribution of flowers and plants cultivated for ornamental purposes, represents one of the most promising segments of contemporary Brazilian agribusiness" [9], a sector that generated R$ 9.5 billion (US$ 1.7 billion) in value for the GDP of the country in 2020. Considering the analysis that an Institut reported in 2015 [8], around 46% of its GDP is represented by the growers' sales (direct or by cooperatives), so an amount of R$ 4.4 billion (US$ 0.8 billion) can be estimated as the turnover for this part of the value chain in 2020.

The Netherlands is the leader in the exportation of these products around the world. "The wholesale business brings cut flowers and plants with the right specifications from growers to florist shops, garden centers, and supermarkets" [20], with a turnover that reached € 7 billion in the level of trading companies in 2020.

In Europe, the data provided by CBI in 2013 presented, as the biggest markets in terms of consumption value of cut flowers, the following countries: Germany (€4.3 billion), France (€3.1 billion), the UK (€2.9 billion), and Italy (€2.7 billion) [33].

The market size of the floriculture sector is increasing globally, even if at different growth rates. In this section, it is presented an overview of the size of the sector, firstly with a global perspective, and then locally, with a focus on the Brazilian sectors, the objective of this research. "Floriculture consumption is strongly related to income levels, thus clarifying why markets with high purchasing power also have high consumption levels" [34] and why economics in developing countries have a strong growth rate compared to developed countries. According to research developed by Rabobank in 2017, it is expected a 2% growth per year in cut flower and potted plant expenditures in Europe and North America, with Asia growing 6-8% annually [35].

A.1 Global

Comparing the different regions around the globe, Asia presents the most optimistic growth expectation. From 2017 to 2027, the consumption value is due to grow by 80%, with the assumptions of continuing economic growth and a strong relation between purchasing power and floriculture expenditures [35]. Figure 35 presents the market value for each region and the expectation of growth for the next few years. The European region and the United States are the biggest buyers, but the biggest growers and exporters are the Netherlands, Ecuador, Colombia, Kenya, and Ethiopia [36].

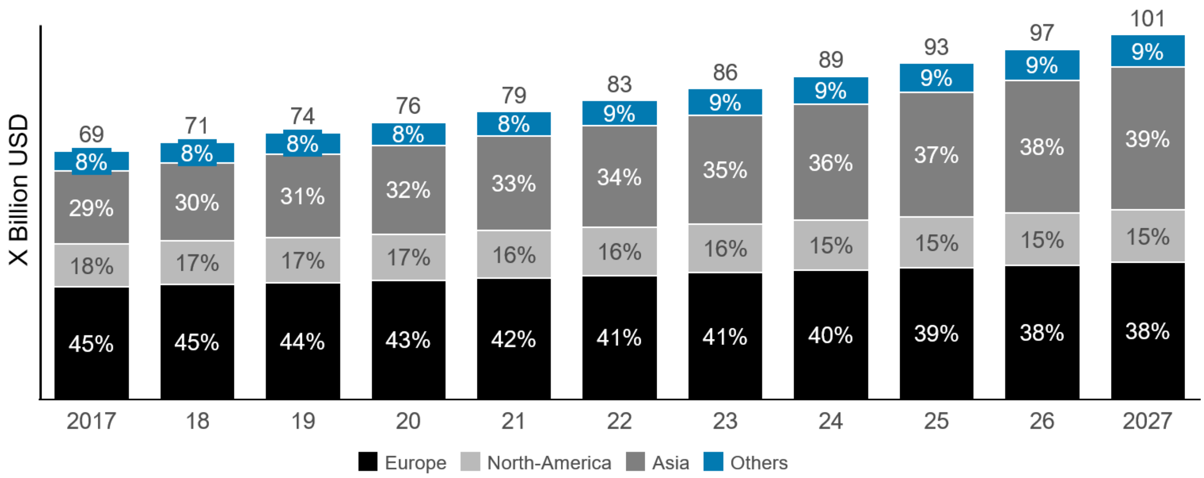

Figure 35: Expected development of consumption value of flowers and potted plants 2017-2027. Adapted from [35].

The cut rose is planted all around the world, as presented in Figure 36. Although the Netherlands is commonly known as the Auction of the world, their production has, in the last few years, followed a downward trend. The countries on the equator gained, in the last years, increasingly space as important growers as the technology allowed transportation for a reasonable price. "Cut flowers have to be transported quickly using a "cold chain" – a series of refrigerated facilities on farms, lorries, planes, and boats – which put the flowers into a dormant state, so they stay fresh" [36].

Sylvie Mamias, secretary-general of Union Fleurs, the international flower trade asso-

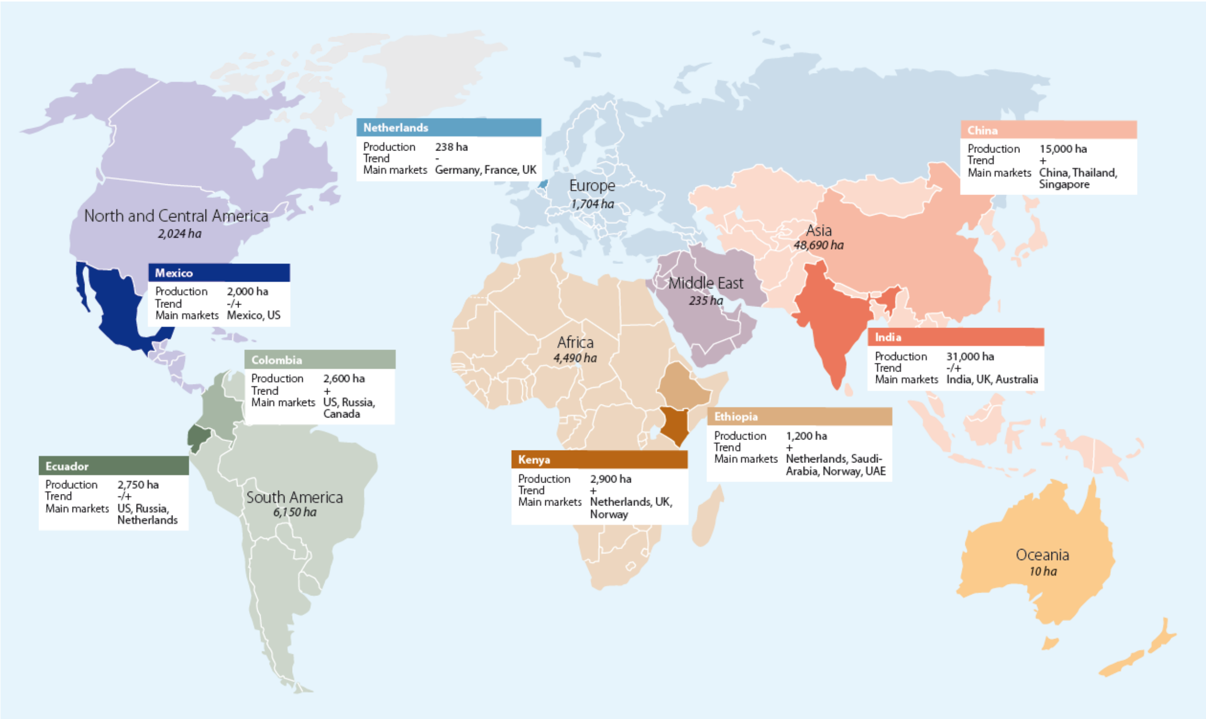

Figure 36: Comparison of cut rose-producing regions and nations. Extracted from [32].

ciation, says that "one of the reasons behind the increase in flower exports in Africa dates from the 1970s when an oil crisis increased the cost of heating greenhouses in northern countries". In this sense, the production moved to the south where flowers could be grown with little energy input all year round. For Europe, this meant seeing more flowers imported from Israel and Morocco, and later East Africa, while US buyers developed trade with Latin America [36].

There are some characteristics in common with the new producers, "areas of high altitude with cool nights, which many flowers benefit from, proximity to the equator for maximum hours of sunlight, and cheaper labor". Moreover, the change from the north to the south meant an "end to seasonal production and the beginning of a 365-day-a-year international competitive trade". [36].

A.2 Brazil

The Brazilian expenditures for the floriculture sector are increasing year after year. Considering the period of 2012 to 2020, the annual growth had a 9% rate. Figure 37 presented the date provided by IBRAFLOR (Instituto Brasileiro de Floricultura - Brazilian Floriculture Institute).

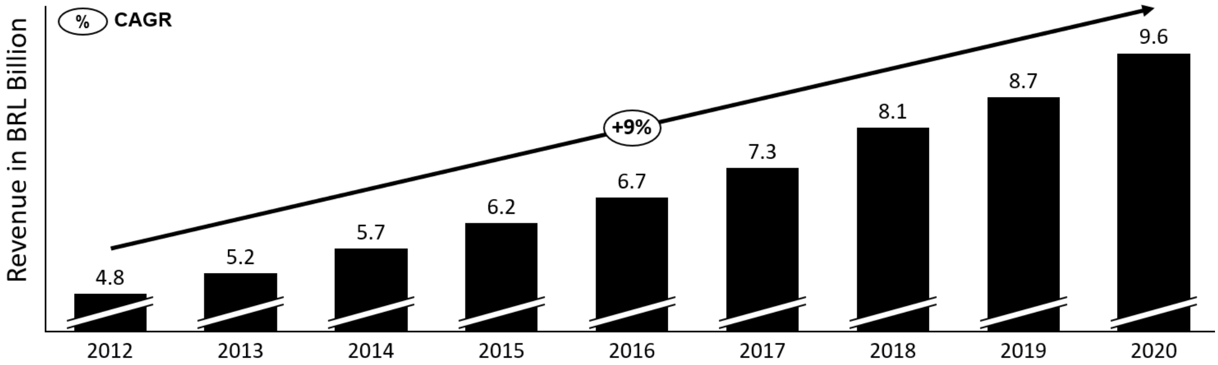

Figure 37: Revenue of the floriculture sector at consumer level in Brazil. Data extracted from [11].

Despite taking place, direct trade between the grower and the final consumer, or even from wholesale to the consumer, retail is still the main channel for the commercialization of flowers and ornamental plants to the end-user. Among the main retail channels are florists, supermarkets, decorators, and landscapers, the latter two of whom are linked to the service sector [8]. Figure 38 presents the participation of the different stakeholders in the value aggregated to the final user. The characterization of the different players in the value chain will be detailed in the next sections.

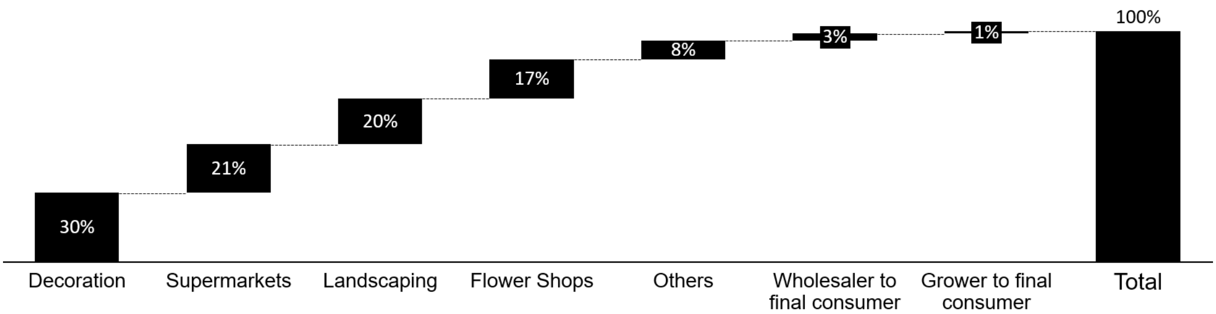

Figure 38: Revenue distribution of the floriculture sector per segment in 2020. Data extracted from [11].

Although in 2015, the GDP estimated of the floriculture sector represented only 0.6% from the total agricultural GDP, the employment of the sector represented around 7% (51 thousand out of 780 thousand registered workers in rural areas) of the total formal jobs of the field. “These figures show that the sector is labor-intensive and contributes to the retention of the population in rural areas”, as presented in the research developed in [8]. It is a technical work that is difficult to automatize, compared to other sectors, what is the main reason for this level of labor-intensive. Table 13 presents how the floriculture sector employs people in different activities.

Employment type	# of employees
Agricultural production	81,000
Wholesaler employees	9,000
Retail employees	112,000
External workers	7,000
Total	**209,000**

Table 13: Brazilian employment structure in the floriculture sector in 2020. Extracted from [11].

According to the data provided by Ibraflor [11], the country has a total of 8.300 growers, with 15.600 hectares as the area of these products. Each farm has an average of 1.88 hectares (18.800 square meters) with an average number of 8 employees per hectare. Figure 39 presents how the space is divided in the different technologies adopted in Brazil and the area used by each product.

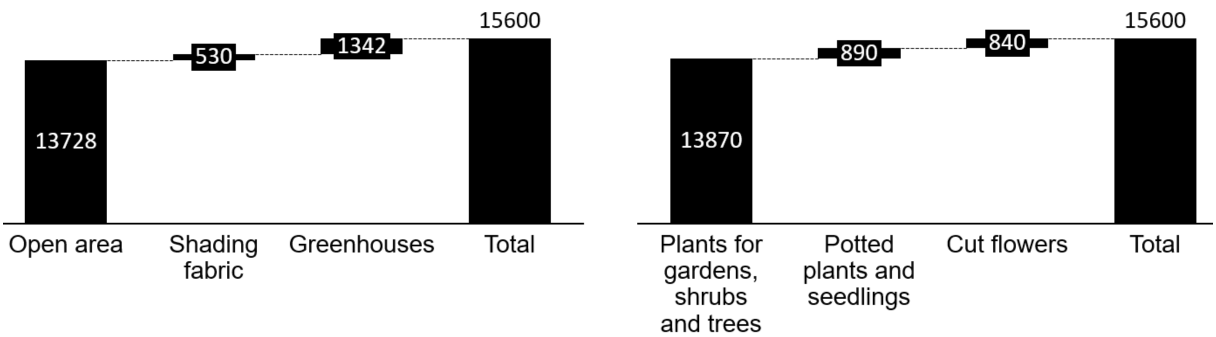

Figure 39: Segmentation of type of infrastructure and product for the different growers in the Brazilian floriculture sector. Data from [11].

Figure 40 presents the main area of production in Brazil, where the states in the figure represent 77% of the total flower cultivated area in the country, the remaining is divided into the other states.

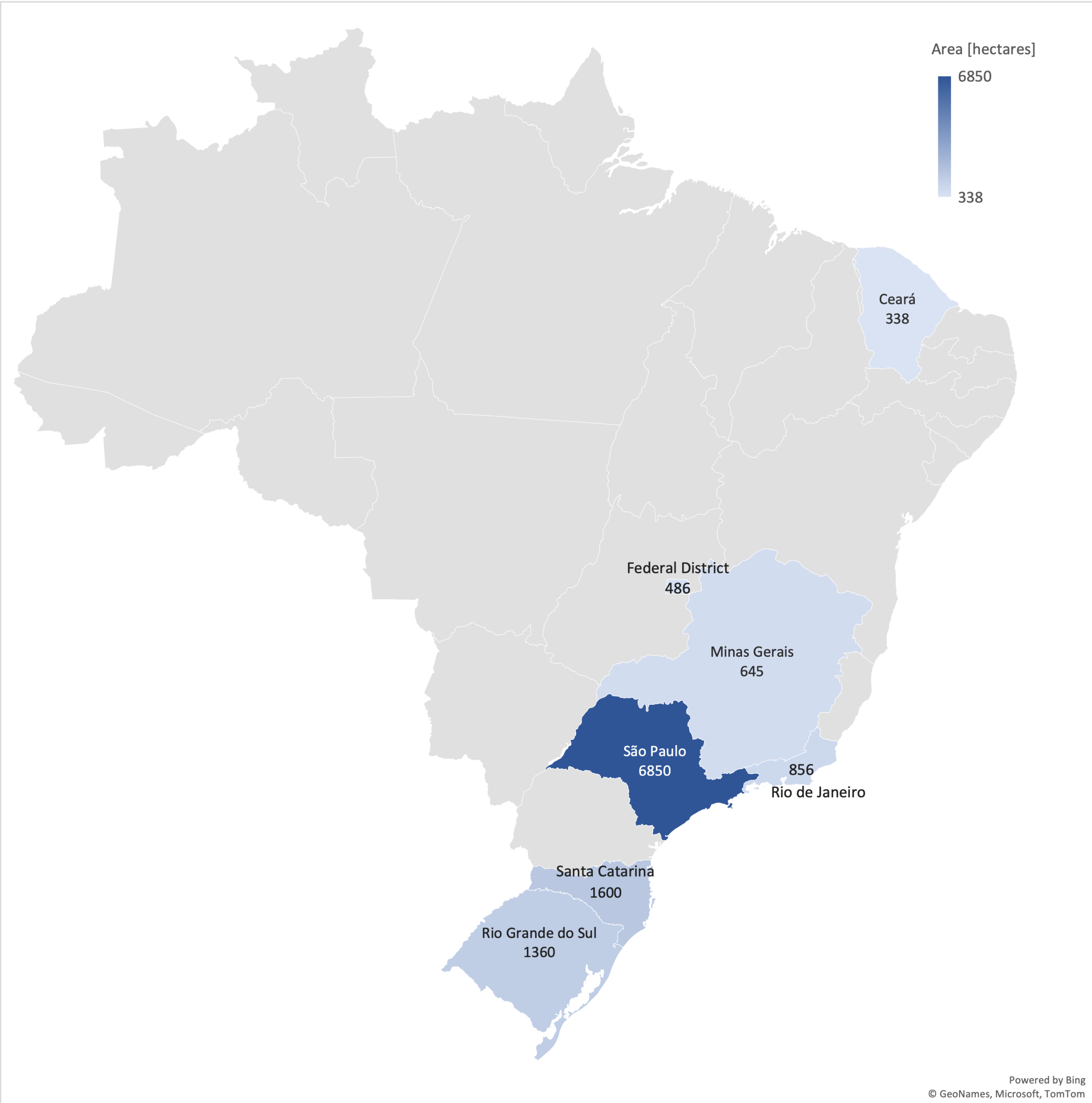

Figure 40: Cultivated area per region of the floriculture sector in the main regions of Brazil in 2015. Data from [8].