Final Submission
Innovation Youth

Cristina Esposito – Alya Gabsi

MMA TEAM

Alya Gabsi

Modeler – UX/UI Specialist

Cristina Esposito

Modeler – Data Architect

TABLE OF CONTENTS

01

Industry & Organization Overview

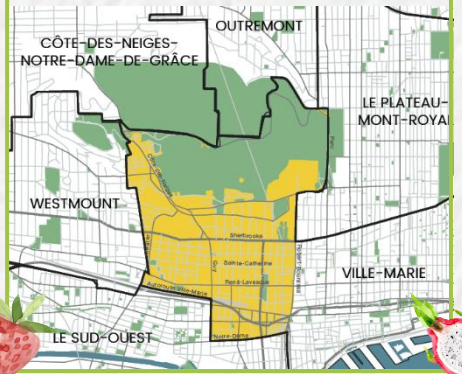
Food Assistance Landscape

The Industry: Food Assistance in Canada

The first food bank in Canada opened in 1981[5] in response to an economic recession. Since then, there has been a growing need for food assistance with the cost-of-living increasing, a key driver of food insecurity. There are many negative implications of food assistance, where the lack of access to good quality food leads to an increase in marginalization, a decrease in one's health, both physical and mental, and further stress to the healthcare system. According to a study done by PROOF, 1 in 8 Canadian households, equivalent to 4.4 million Canadians, were food insecure in 2018 (the last time there was a nationwide study on food insecurity)[1] . Quebec was the province with the lowest rate of food insecurity, coming in at 11.1% of the Quebec population. The need for food assistance became more apparent during the pandemic. According to Food Banks of Quebec, demand for assistance increased by 30%, much of which was due to people having lost their jobs[7] . Today, there are varying methods of food assistance besides food banks, such as community kitchens encouraging those in the community to cook healthy foods, and community solidarity markets that provide people with access to fresh produce at a cheaper price than supermarkets.

The Community: Peter McGill Neighbourhood

Peter McGill is in the heart of downtown Montreal with a population of over 35,000 residences. In 2019, a research study was done to understand the changing makeup of the neighbourhood and those who live there to provide them with the services that meet their needs. The neighbourhood is very diverse, where 63% of the population are 1st generation immigrants. The rise in immigrants has led to a rise in families and young children in the area. Of the 35,000 residents, 4,900 are seniors 65+ in age, and 7,800 are youth between the ages of 18-24. Over 42% of those in the community live below the poverty line, where the average household income is $8,000 less than the Ville-Marie average[9].



The Organization: Innovation Youth (IY)

Innovation Youth was created in 2006 as a community organization for young people that services the Peter McGill community in downtown Montreal with 4 areas of focus:

- Youth empowerment – focusing on education and academic success
- Children and family outreach and assistance for newly arrived immigrants
- Urban agriculture – focuses on growing food in partnership with 9 community garden sites
- Food security – food bank, community kitchen, and farmer's market

In terms of their food security initiative, the pandemic created an increased need for assistance within the community, with the cost of food rising and many having lost their jobs. To respond to this need, a subdivision was created in April 2020 called Innovation Assistance (IA). Their mission is to provide those in the community access to quality, healthy, fresh food. As of December 2021, they had over 300 members who received free food baskets on a monthly basis. Due to such high demand and wanting to reach out to as many people as possible, they changed their model into a solidarity farmer's market.

The Setting: IA's Solidarity Farmer's Market

The farmer's market was introduced to allow the organization to reach out to more members in the community who require food assistance. Originally once a month, the plan is to expand the market to twice a month during the summer months, and if successful, switch to weekly in the fall. The market has two sections for food – a members only section, and a section open to all (members and non-members). The members only section has food donated by Moisson Montreal and are free. The section open to all sells mostly fresh produce sourced from two suppliers (L'Autre Choix and Jardins Carya), as well as some produce grown in their own community gardens. The pricing model for members is 20 % of the retail price, 50 % of the retail price for non-members, and a pay-what-you-can model for those who simply can't afford the final cost of their bill.

For more information about the Food Assistance Industry, please refer to AS1. For the Peter McGill Community AS2. For Innovation Youth AS3. For the Solidarity Farmer's Market AS4

Overview of the Use Case

Objective: *Develop a forecast modelling framework that would allow Innovation Assistance to be able to predict how much produce to buy for the bi-weekly solidarity farmer's markets that could be used to make better ordering decisions to reduce left-overs , waste, and better manage costs.*

Operational Context

Pain Points: (ref. Strategy slide)
- Currently place orders based on intuition
- Missing a structured ordering process
- Tend to have leftovers that could not be resold (end up being donated or turned into food to resell)
- Faced with extremely limited fridge space
- Limited on the data tools available for use

Stakeholders:
- IA workers (ref. current ordering process slide)
- IA members and customers who attend the markets
- Suppliers of fresh produce (ref. AS4)

Deliverables and Success Criteria

- A forecast and optimization (ref. Modelling Process slide) modeling POC framework that utilizes IA's preferred tools that can be fully integrated and operational with real data
- Organized architecture structure to allow for the collection of data, along with a flow diagram displaying how the data relates to each other (ref. ERD slide)
- Documentation of a step-by-step process to facilitate seamless knowledge transfer and the operationalization of the POC (ref. Google Doc for documentation)
- Modified and improved user ordering process experience (ref. proposed ordering process slide)

Organizational Value (ref. Strategy Slide)

Quantitative Value Drivers:
- Reduce leftover and wasted food produce
- Better management on spending and limited budget
- Better data management and analysis

Qualitative Value Drivers:
- Improved bargaining power
- Happier customers
- Maintain/improve accessibility of fresh produce to those in need

Data Overview (ref. Data Sources slide)

Internal Data:
- Invoices in PDF format since August 2021
- Google sheet tracking the number of attendees at each market
- Real sales data in a Google sheet from the June 6th, June 20th, July 4th, and July 18th 2022 markets

Additional/External Data:
- Creation of synthetic sales data
- Seasonal produce guide for Quebec
- Weather data – temperature and if there was precipitation

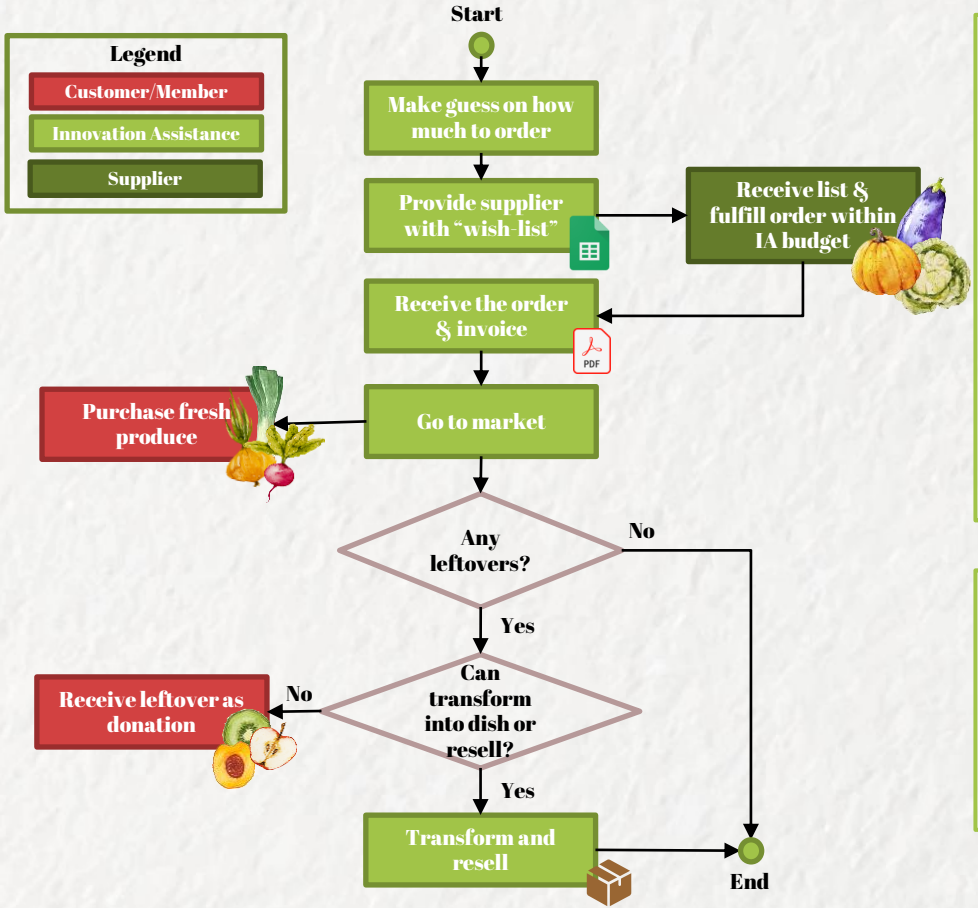
Architectural Tools Used in The Framework: (ref. Ideal proposed architecture slide)
- Google Sheets for data collection and dashboarding
- Google Colab for forecasting and optimization modelling
- Google docs for documentation
- Google Drive for storage of files

Note/Disclaimer*: Prior to deep-diving with IA on the state of the use case, it was written in the project proposal that this use case would be heavy modelling to forecast the demand of produce from their solidarity farmer's markets, in order to help them determine what they should be buying. However, since the deep-diving sessions, we had come to realize that the inexistence of demand data and data collection process required us to readjust the focus of our roles on elements related to UX/UI and Architecture. Therefore, certain components that would be seen in the modelling sections focusing on preprocessing, the models chosen, and the results will not be a focus, given the recommendation received by our coach from the SOW submission to focus more on building a working framework (architecture). To counter the lack of data, we created our own synthetic data, which will be discussed further.*

02

UX/UI

Current Ordering Process Flow



User Persona

innovation assistance

About Erin and Micah:
Role: Floor and Operations managers at IA
Characteristics: Looking to make the community a better place, non-technical know-how

Goals:
- Have an easy-to-use forecasting tool utilizing current Google suite to effectively plan for the farmer's markets

Frustrations:
- Lack data collection process
- Sometimes have a lot of leftovers
- Limited budget to run the market

Motivations:
Reduce leftovers from market

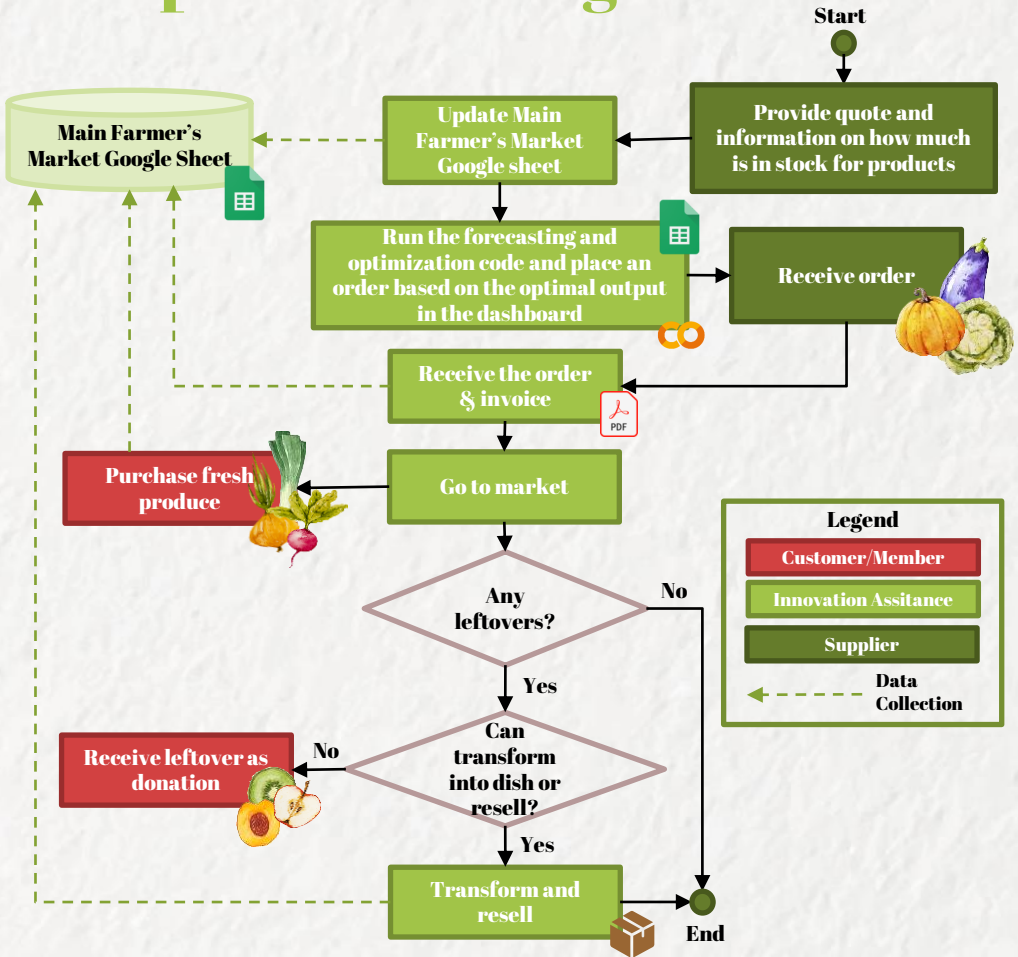
Improve budget management

Maintain/increase reach to the community

Key Points on the Experience:
- Make an educated guess on how much is needed → leads to potential over ordering (many leftovers and $ wasted) or under order (doesn't achieve their mission in fulfilling the needs of the community)
- Supplier is given purchasing power to make decision on what to order when certain products listed on the "wish-list" are not available → risks increasing cost and makes the ordering process more reactive
- No data collected on what was sold at the markets → lose visibility on demand that is crucial in making future purchasing decisions

Please refer to the "Ordering – IA x JM" Google Sheet to see the original wish list file

Proposed Ordering Process Flow



For our POC tool to work, we need to make modifications to the current purchasing process for it to be deemed usable in the future. The process should start with IA requesting a quote from the supplier that contains information on what products are available for sale, their quantities, and their cost. 1 master Google sheet is used, called the "Main Farmer's Market", and must be updated with the latest information from these quotes, along with invoices and what was sold at the markets. This file is then automatically loaded into a Google Colab file that only requires IA to click the "Run" button to generate the forecasted demand of products and the optimizer that will provide them with an optimized grocery list. With this optimized list, they can place their order with the supplier.

The benefit of including a step for a quote allows IA to have more certainty as to which products are available for sale, as well as what the expected cost is without putting the control in the hands of the supplier to sell them other products that may not be in-line with their budget. It essentially shifts their operating approach from one that is passive to one that is proactive. The forecast and optimizer tool will be able to generate a proposed grocery list within minutes, allowing IA to make more educated decisions with the data they have, which should help reduce spending and potential waste/leftovers.

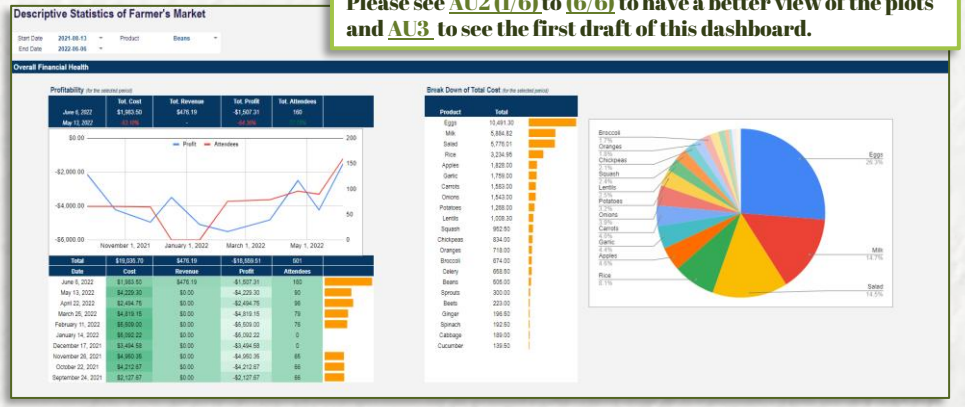
Key Points on the Experience:

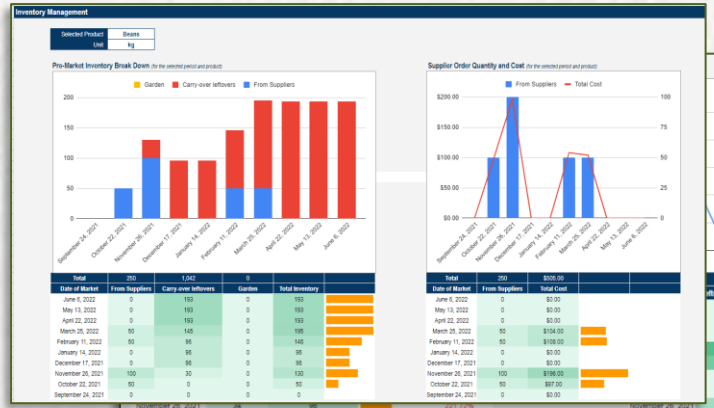
- Implemented forecasting and optimization models to provide a recommended grocery list that is based on previous demand patterns → reduces the potential of over or under ordering with more accurate ordering decisions
- Incorporated an easy-to-use dashboard interface using Google sheets → one-stop-shop location to get all information quickly that is important to IA to monitor spending and demand patterns
- IA is given purchasing power to make their own ordering decisions with the inclusion of a quote and in-stock information from the supplier → provides more control over their budget and changes the process to be more proactive by knowing what is in stock with the supplier before placing an order
- Inclusion of data collection at various stages of the process → allows for greater visibility on demand and any changes to customer behaviours that could arise

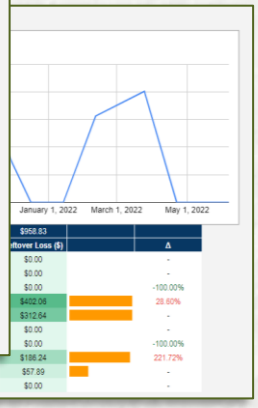
To see the step-by-step process of what needs to be done by IA in the forecasting process, please refer to AU1.

Descriptive Dashboard

Please see AU2 (1/6) to (6/6) to have a better view of the plots and AU3 to see the first draft of this dashboard.







Please refer to the "Dashboard" Google Sheet, tab "Statistics DB" to see the build

Purpose

Currently, Innovation Assistance does not have any way of understanding trends in their data. They simply rely on intuition to run their operations. This descriptive dashboard will allow IA to make better informed decisions to drive the success of their new solidarity farmer's market business model.

Characteristics

The dashboard is composed of two main sections governed by three filters; Start Date, End Date and Product. The first section called "Overall Financial Health" provides high-level statistics on the farmer's market without discriminating by product. This includes information such as Total Cost, Revenue, Profit and Attendees for each market, as well as an overall breakdown of Total Cost by produce. This section will help IA monitor the profitability of their market, as well as understand where their budget is spent.

The second section provides produce specific information, such as the amount and origin of pre-market inventory, supplier order quantity and cost, sales and post-market leftover loss (in respective units and dollars) for each produce. This will allow IA to identify any cost abnormalities and understand their demand for each produce. It will also help them monitor their leftovers, which is an important KPI for them as their goal is to minimize it (cf. Mixed-Integer Problem slide).

Updates

Updating this dashboard is relatively easy. In fact, it is a dynamic dashboard, meaning that the current graphs automatically update with new market data. To edit a chart or change the chart type, the UX/UI operator need only to (1) click on the graph (2) then on the three dots (3) and then on "edit chart". To add a new graph, the operator needs to import the necessary data using the IMPORTRANGE() function, if the data has not already been imported (cf. AU2. Dashboard - Tabs) and then can use the QUERY() function to query the data in the dashboard and create the new graphs.

Forecasting Dashboard

Please refer to the "Dashboard" Google Sheet, tab "Forecasting DB" to see the build

Purpose

This dashboard is the core of the project. It provides all the information needed for IA to prepare the inventory for their bi-monthly markets. It displays the output of the demand forecasting/optimizer integrated model that we have built. With this dashboard, IA will be able to know how much of each produce to order from their suppliers given garden production and previous market leftovers in order to minimize next period's market's leftovers. This tool will help IA better plan for their future and drive their new business model to success.

Characteristics

The dashboard shows the predictions of the ordering amount for next period (next market). It includes 4 sections :

Product:

- **Unit**: Unit clarification for each produce.
- **Product Name:** Produce name.

Total Available Supply

- **Total Suppliers Inventory**: Total quantity available to order from suppliers.
- **From Garden**: : Total quantity of produce grown in IY's garden.
- **From Previous Market**: Total quantity of leftovers from previous market.

Market Inventory Planning

- **Forecasted Demand**: Predicted demand from our forecasting models.
- **Quantity to Order from Suppliers**: Grocery list to be sent to suppliers.
- **Quantity from Garden/Previous Market**: Quantity to sell from garden and previous leftovers.
- **Forecasted Sales** Expected sales when taking into account demand and supply.
- **Forecasted Leftovers**: Expected leftovers when taking into account inventory and sales.

KPI

- **% Fulfilled Demand**: This represents how much of the demand was fulfilled.
- **Leftover reduction from previous market:** This represents the percentage decrease in leftovers.

Inventory Ordering Manager for Farmer's Market

Product		Total Available Supply			Market Inventory Planning					KPI	
Unit	Product Name	Total Suppliers Inventory	From Garden	From Previous Market	Forecasted Demand	Quantity to Order from Suppliers	Quantity from Garden/Previous Market	Forecasted Sales	Forecasted Leftovers	% Fulfilled Demand	Leftover reduction from previous market
lbs	Apples	46	0	563	64	0	64	64	499	100%	-11%
kg	Beans	80	0	193	0	0	0	0	193	-	0%
lbs	Beets	72	0	0	0	0	0	0	0	-	-
unit	Broccoli	26	0	0	0	0	0	0	0	-	-
lbs	Cabbage	29	0	0	0	0	0	0	0	-	-
lbs	Carrots	49	0	1462	28	0	28	28	1434	100%	-2%
unit	Celery	72	0	126	0	0	0	0	126	-	0%
kg	Chickpeas	66	0	358	0	0	0	0	358	-	0%
unit	Cucumber	39	3	0	86	39	3	42	0	49%	-
dozen	Eggs	70	0	1382	76	0	76	76	1306	100%	-5%
lbs	Garlic	47	0	479	14	0	14	14	465	100%	-3%
lbs	Ginger	100	0	66	12	0	12	12	54	100%	-18%
kg	Lentils	16	0	248	0	0	0	0	248	-	0%
litre	Milk	75	0	0	5	5	0	5	0	100%	-
lbs	Onions	82	0	1305	110	0	110	110	1195	100%	-8%
unit	Oranges	33	0	1314	239	0	239	239	1075	100%	-18%
lbs	Potatoes	77	0	1821	142	0	142	142	1679	100%	-8%
kg	Rice	95	0	1524	0	0	0	0	1524	-	0%
kg	Salad	28	0	0							
lbs	Squash	34	0	1258							

Please see AU3. Wireframes to see the first draft of this dashboard.

Updates

This dashboard has been fully integrated into the architecture as seen in the ERD slide. It is both connected to the python code and to IA's data. This means that it is automatically updated and tailored for each market. The introduction of a new product is also automatically handled by the dashboard. However, changing or adding columns can only be done in the code so python knowledge is required. See AA1. Optimizer Connectivity for more information on how the optimizer is structured.

03

Architecture

Ideal Proposed Architecture

Data Sources

The data sources used throughout the framework involve a combination of internal data on what products were sold and inventory, as well as external data from invoices, quotes, weather data, and a seasonal produce guide.

Input

These data sources are collected within the Main Farmer's Market Google sheet, where there are various tabs identified for each source.

Analysis

The Main Farmer's Market Google sheet is then used as an input within a Google Colab file. Python is used to preprocess the data, perform predictive demand modelling using 8 forecasting models, and an optimization model to provide a recommended optimized grocery list given the output from the forecast.

Output

The output from the Google Colab file is placed in a Dashboard Google Sheet that will showcase the optimized grocery list needed for the next market. An additional tab displays various exploratory graphs and tables pulled from the Main Farmer's Market Google sheet to showcase various business metrics.

Storage

The Main Farmer's Market Google sheet, Google Colab, and Dashboard Google sheet are stored within a folder on a Google Drive. Considering that the size and volume of the files are not large, this is a feasible storage solution that keeps all files in 1 location for easy access, convenience, and free of cost.

Why These Technologies?

For this POC solution to be as realistic as possible and to have any potential of future adoption, we had to use tools within the Google suite (Google Sheets, Google Docs, Google colab, and Google Drive) for various reasons. These are the only tools IY has access to/use within their day-to-day operations, as they cannot afford to spend money on subscriptions for other tools when most of their income comes from donations and their farmer's market. All income is supposed to be put back into serving the community. Luckily, these tools are free of cost for Google Drive storage up to 15GB. Considering that the volume and velocity of the data used within the solution is minimal, this makes it a viable solution. It is important to note that the level of technical knowledge of the users of this solution is also minimal, which makes it critical to use tools that they area already familiar with, will be fast and easy to implement, adopt, and maintain.

4Vs

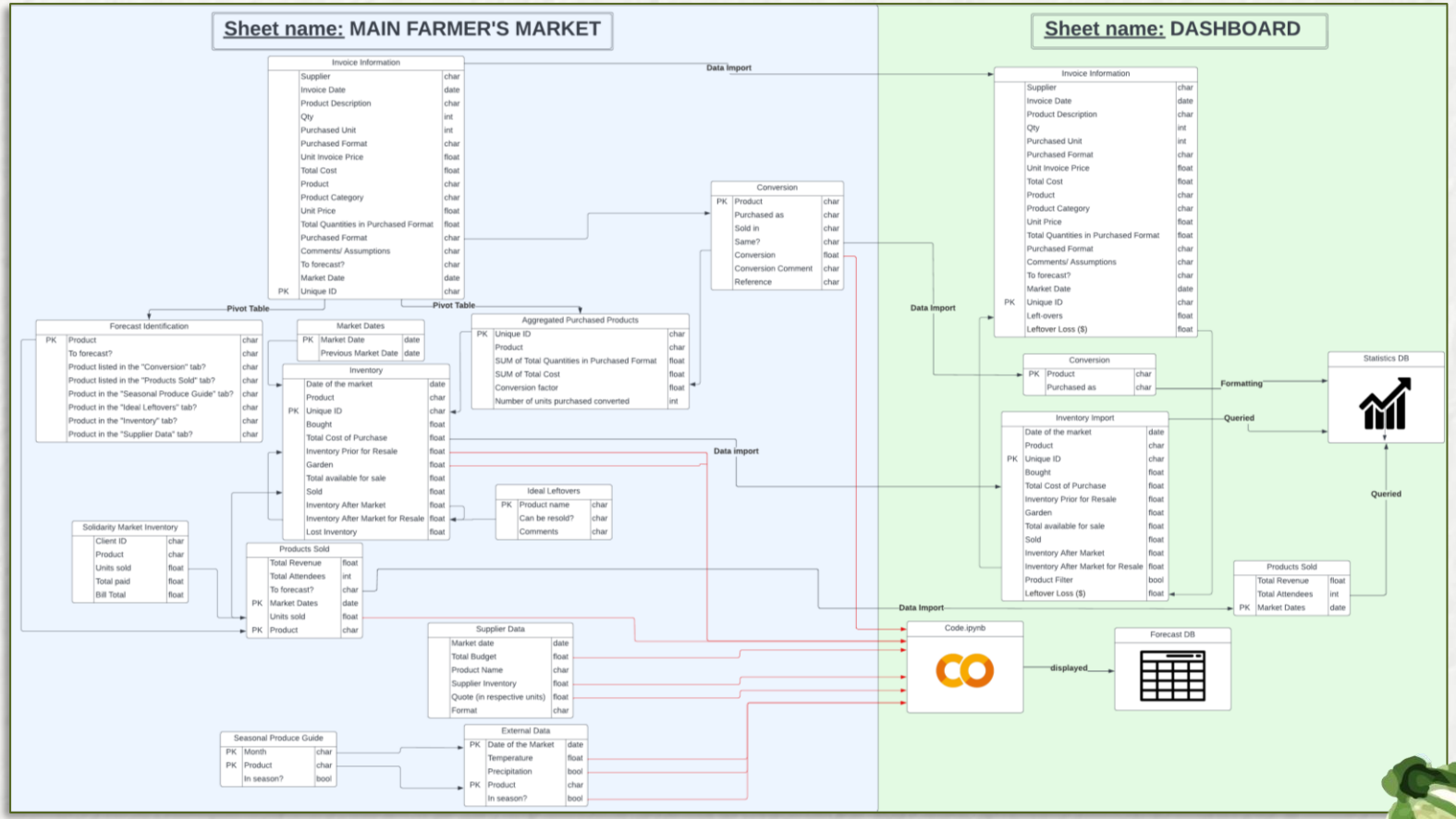
Volume
On average 18 unique products are purchased per market that requires updating within the Main Farmer's Market Google sheet.

Velocity
Data and output requires to be updated and generated twice a month (the frequency of the market).

Variety
Structured data is collected in the Main Farmer's Market Google sheet. Unstructured data from invoices mostly come in PDF file format.

Veracity
All data is collected and inputted within the Main Farmer's Market Google sheet manually, leaving for potential to make mistakes in logging the information

Entity Relationship Diagram (ERD)



To help visualize how all the tabs in the Main Farmer's Market and Dashboard Google sheets relate to one another, we have put together an ERD diagram for the tabs containing information (i.e., excluding any tabs that may be used for drop-downs). We have 2 Google sheets – in blue is the "Main Farmer's Market", and in green is the Dashboard. The Main Farmer's Market Google sheet contains information on invoices, what was sold at the markets, inventory, and external data. The Dashboard contains graphs and tables to showcase exploratory information (Stats DB), as well as the optimized grocery list (Forecast DB). Information that is used within the Google Colab file have arrows highlighted in red. For a full breakdown of each tab, please refer to AA2.

Please refer to the "Main Farmer's Market" Google Sheet to see the data collection build

04

Modelling

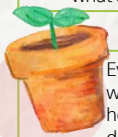
Data Sources

Internal Data

There were various datasets provided by IA at the beginning of the project (for the full list and our conclusion on relevance, please see AM1). The datasets used in the project were the following:

- Invoices – unstructured, comes in a PDF, from August 2021 until present, totaling 60 invoices with the following datapoints: Supplier name, invoice date, Product Description, Quantity Ordered, Unit Invoice Price, and Total Cost. For this to be used within our framework, we manually copied the information into the "Invoice Info" tab of the Main Farmer's Market Google sheet.
- Farmer's Market Tracking – structured, comes in a Google sheet with the number of attendees at each market. There were 3 dates that were missing the number of attendees. For these, we imputed them with a linear regression.
- Solidarity Market Inventory – structured, comes in a Google sheet, contains demand data requested by MMA to IA to collect how much of each product was sold at the markets, for June 6th, June 20th. The datapoints collected are the products sold by individual member. Even though we only need the total amount sold of each product at each market, we recommended they collect the data by individual so that after a year of collecting, we could expand the use case to perform market basket analysis to understand what are certain products that are more likely to be purchased together at the market.

External Data Overlay



Even though the focus of the project was no longer solely on modelling, in reality, external data would be a great source of identifying other factors that could influence demand, and demonstrate how this could be incorporated within our modelling framework. There are two sources of external data that we have included within our framework:

- Quebec Seasonal Produce Guide – unstructured, comes in a PDF, lists 49 different produce that is sold at Quebec farms, along with the months they are in season. This was provided to us by IA, as they had mentioned that they relied a lot on this guide to determine what and when to purchase. Therefore, we felt that it would be an important factor to consider and incorporate within our forecasting models, since what is in season greatly influences the demand for it. For this to be used within our framework, we manually copied the information into the "seasonal produce guide" tab using binary values to indicate if each product was in season (in season = 1, otherwise 0). There were an additional 15 products that IA had purchased in the past that were not included on this list. For these products, we added them to the "seasonal produce guide" tab, highlighting them in orange, and added a comment with our assumptions, whether taken from a website or just our own judgement.
- Weather data – structured, can get actual and forecasted weather from any weather network website. The historical weather data we used comes from the Government of Canada. The data points required would be the temperature in celcius of the day of the market and a binary value that would signal if there was precipitation. For actual market's in the past, we want to record the actual weather that occurred, and for the upcoming market, we want to record what it is being forecasted, along with the precipitation (defined as rain or snow) to be =1 if there was precipitation, and 0 otherwise. This data requires to be manually recorded within the "External Data" tab.

Synthetic Data Creation

Since the goal of this framework is to forecast the demand of individual products, we had to create synthetic sales data from previous markets. This is because for us to forecast demand, we need to know how much of a product was sold at each market. To generate this synthetic sales data, we used various sources of internal and external data available to simulate a random pattern of demand for each product. In total, we generated synthetic sales data for 20 of the most purchased products at IA. The following data points were used to generate the synthetic sales data:

- Date of the market from September 24th, 2021, to June 6th, 2022
- Number of attendees at each market – from the "Farmer's Market Tracking" Google sheet from IA
- % change in attendees – based on the difference in number of attendees between 2 consecutive market dates
- The product description
- Unique ID – concatenation between the market date and the product description, used as a VLOOKUP within the "Products Sold" tab to bring in the synthetic sales data values for each product and market date
- In season? – taken from the Quebec Seasonal Produce Guide
- Bought – how much of the product was purchased by IA from suppliers (in the unit in which it is purchased in by customers at the market), taken from the "Aggregated Purchased Products" tab
- Inventory prior to market
- Garden – how much of the product is coming from IY's own community gardens
- Total available for sale - how much of the product was purchased by customers at the market (in the unit in which it was purchased at the market)
- % sold – proportion of what was bought by customers over what was bought from suppliers by IY
- Inventory after the market for resale – based on online research in the "Ideal leftovers" tab to assume the shelf-life of each type of product

Prior to generating the synthetic data, we had spent some time scouring data sources on websites like Kaggle to see if what we needed for the framework already existed. We were not able to find data sources that fit what we needed, even if we had to apply additional data reformatting/cleaning.

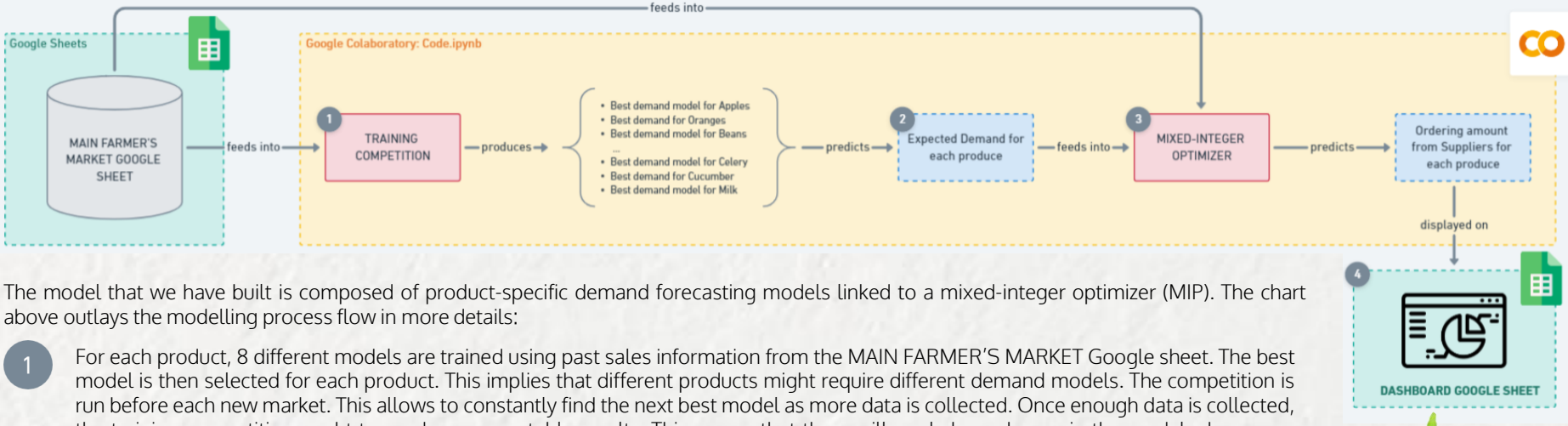
The process to generate the synthetic "Sold" value is as follows:

- Working from the most recent market date, we take the actual sales value recorded from June 6th (the latest real data we had at the time) as our base in projecting the values for the previous market
- For the previous market, we created a RANDBETWEEN() function that randomly generates a "Sold" value between:
  - (% change in attendees * % sold of the next market) + ((in season? * 5% increase) * total available for sale) – *10%* of what is available for sale)
  - (% change in attendees * % sold of the next market) + ((in season? * 5% increase) * total available for sale *+10%* of what is available for sale)
- Each product has its own *+/- %* increase or decrease to set an upper and lower bound. This value was set randomly.

The methodology behind this is to generate sales values based on various factors that could fluctuate demand, such as the number of attendees over time and whether the product was in season or not. For an example of what this looks like, please refer to AM2.

Please refer to the "Synthetic Data" Google Sheet & tab to see the build

Modelling Process



The model that we have built is composed of product-specific demand forecasting models linked to a mixed-integer optimizer (MIP). The chart above outlays the modelling process flow in more details:

1. For each product, 8 different models are trained using past sales information from the MAIN FARMER'S MARKET Google sheet. The best model is then selected for each product. This implies that different products might require different demand models. The competition is run before each new market. This allows to constantly find the next best model as more data is collected. Once enough data is collected, the training competition ought to produce more stable results. This means that there will rarely be a change in the model when more data is collected. For more details on this process please see the Models and Model Selection section in the Forecasting Models slide.

2. After selecting each model, they are used to predict the next period's demand for each product.

3. The predicted demand values along with other data points from the MAIN FARMER'S MARKET Google sheet are then fed into the MIXED-INTEGER OPTIMIZER. This optimizer then solves for the best ordering amount that minimizes leftovers given specific constraints. For more information on the MIP formulation please see the Mixed-Integer Problem slide.

4. The results of the optimizer are then displayed on the Forecasting dashboard.

Please refer to the "Code" Google Colab file to see the full modelling build

Forecasting Models

1. Data Imported and Preprocessing

The data sources used within the forecasting models pulls from 3 tabs within the Main Farmer's Market Google sheet: Products Sold (used to predict demand and help determine the amount to purchase), External Data (used within the supervised learning models), and Conversion (used to convert the forecasted value into the purchase unit of the supplier). The pandas package was used to convert these files into data frames and apply data preprocessing and formatting, such as filtering to have only the products selected to be forecasted, splitting the data points from the external data tab to be for training and prediction, and ensuring the data types for all values are in their respective formats.

When converting from univariate time series to supervised modelling, we had to reformat the data frame to incorporate the external data (temperature, precipitation, and in-season), as well as 3 columns to lag the sales (1 period, 2 periods, and 3 period lag). With this shift, we would be left with a few NA values, and therefore had to remove these prior to applying the supervised models.

3. Future Extension of the POC – Focus on Model Development and Accuracy

As a future extension of this POC, we recommend that once a minimum of 1 year's worth of real data is collected (i.e., minimum of 24 data points if the plan is to only stay with bi-weekly markets), to revisit the current models produced and focus on developing models with accurate results. Depending on the models developed, further data preprocessing may be required with the new data, such as normalizing/standardizing the data, applying feature selection/multicollinearity tests with the external data, and assess for trend and seasonality components for each product. If IA does decide to change the frequency of the markets to weekly, the forecasting model should then be able to forecast at minimum 2 periods in the future in order to account for the minimum 2 week lead time that they have with the main supplier.

```
{'Apples': 64,
 'Beans': 0,
 'Beets': 0,
 'Broccoli': 0,
 'Cabbage': 0,
 'Carrots': 28,
 'Celery': 0,
 'Chickpeas': 0,
 'Cucumber': 86,
 'Eggs': 76,
 'Garlic': 14,
 'Ginger': 12,
 'Lentils': 0,
 'Milk': 5,
 'Onions': 110,
 'Oranges': 239,
 'Potatoes': 142,
 'Rice': 0,
 'Salad': 0,
 'Squash': 0}
```

2. Models and Model Selection

The forecast models are used to forecast for 1 period ahead (i.e., the next market) on an individual product basis. Originally, we had wanted to do it by product grouping (i.e., leafy greens, fruits, legumes, etc.). However, we noticed that forecasting in an aggregated way would make the problem more challenging for IA in 2 ways:

- The optimized output would provide too much of a generalized recommendation and does not solve their problem in knowing what products need to be purchased
- Additional format standardization is required on the way the products are purchased from suppliers and sold at the markets. When aggregating by product group, this would mean that certain products in the group might be purchased and sold in different formats (i.e., pounds vs. units)

Since there are multiple products to forecast, each product having its own patterns with demand, having a "one size fits all" model would not be feasible. Therefore, our methodology was to create a for loop simulating a model competition framework for 8 different forecasting models, 6 of which use traditional univariate time series models (using the stats models package), and 2 supervised models (using the SkLearn package):

- Moving average
- Auto Arima
- Seasonal auto Arima
- Simple exponential smoothing
- Holt winter's exponential smoothing
- Naive forecasting
- Random forest
- Multiple linear regression

For more information on forecast modelling framework, along with the assumptions we took for variables used within the supervised models, please refer to AM3. The train/test split used to select the best model for each product is set to train on all records except the most recent record, which is used for training. We did this for 2 reasons:

- We only have 14 total synthetic data points
- Only forecasting for 1 period in the future. Therefore, we can use more data for training and capturing any seasonal patterns that may occur within the demand.

The measure of accuracy used within the model is the mean absolute percentage error (MAPE), because this was a measurement that was easiest to interpret by the IA team. Once the best model is selected, we take the already trained model and apply it to the whole data set to predict our final forecasted value for each product. These values, along with the name of the product, are stored in a data dictionary for easy use within the optimization model (see left image of the output – values are listed in the units in which they are purchased with the supplier. For example, the values for apples are in pounds. For the full list, please refer to the conversion table).

Please refer to the "Code" Google Colab file, parts 1 & 2 to see the data preprocessing and forecasting models

Mixed-Integer Problem

Please see AM4. Problem Justifications for a thorough walkthrough of our client's requirements & problem constraints.

The following problem is the formulation for the optimizer. We used the pulp package (one of the best free optimization packages on python) to implement this formulation.

Objective Function

Minimizing leftovers:

$minimize \ \sum_i (X_i + G_i + L_i - S_i)$

Decision Variables

X_i : Quantity of product i to order from suppliers
S_i : Expected Sales for product i
y_i : First binary variable to control for demand fulfillment constraint
δ_i : Second binary variable to control for demand fulfillment constraint

Data Required

I_i : Total suppliers' inventory for product i
D_i : Forecasted demand for product i
G_i : Total inventory produced in garden for product i
L_i : Total leftovers from previous market for product i
q_i : Supplier quote (in $ per respective unit) for product i
B: Total budget
M: Very big number

Please see AA1. Optimizer Connectivity for more information on the origin of the data

Constraints

Data type:
X_i is continuous and greater than 0
S_i is continuous and greater than 0
y_i is binary and can only take on the values 0 and 1
δ_i is binary and can only take on the values 0 and 1

Total cost should not exceed budget :

$$\sum_i q_i X_i \leq B$$

We cannot order more than what suppliers have in their inventory:

$X_i \leq I_i \ for \ all \ i$

Sales is the minimum between demand and total pre-market inventory:

$S_i \leq D_i \ for \ all \ i$
$S_i \leq X_i + G_i + L_i \ for \ all \ i$

Demand fulfillment constraint:

$X_i \geq (D_i - L_i - G_i) y_i + I_i (1 - y_i)$
$D_i + 0.001 - M(1 - \delta_i) \leq I_i + L_i + G_i \leq D_i + M\delta_i$
$1 - M(1 - \delta_i) \leq y_i \leq 1 + M(1 - \delta_i)$
$-M\delta_i \leq y_i \leq M\delta_i$

Please refer to the "Code" Google Colab file, part 3 to see optimization model

05

Strategy

Strategy

**Please refer to AS6 on the next steps/project timeline for moving the project forward*

Value Prop

Objectives	Pain Points	Solutions	Business Value
➢ Shifting business model to a bi-monthly farmer's market ➢ Maximizing profit ➢ Minimizing post-market leftovers ➢ Relying on own crop and reselling/transforming leftovers	➢ Scattered data files ➢ No previous records of sales data ➢ No methodical approach to ordering inventory ➢ Reliance on free data tools	➢ Setting up an easy-to-understand centralized database on Google Sheets ➢ Developing an integrated demand forecasting – MIP optimizer model to minimize leftovers that takes into account cost, leftovers and crop ➢ Build a descriptive dashboard to monitor operations for the long-term profitability of the market	➢ **Data Management Process Refinement**: Interconnected Google Sheets dashboard ➢ **Decision Automation**: Automated supplier grocery List ➢ **BI capabilities**: Comprehensive Descriptive Dashboard ➢ **KPI enhancements**: Minimizing leftovers and maximizing profit **NPV of $24,877.93 over 5 years**

**Please refer to AS5 for the NPV calculation breakdown*

Value Drivers

- **Reduce leftover and wasted food produce**: Our model is specifically built around minimizing post-market leftovers as much as possible given certain constraints. This allows IA to minimize food waste and maximize profit as less waste implies less unwanted expenses.
- **Better management on spending and limited budget**: Our model ensures that IA does not exceed their total budget, further maximizing their profit.
- **Better data management and analysis**: Through a centralized data architecture, IA will make better use of their data to infer insights on the health of their operations and profitability of their market.
- **Improved bargaining power**: By asking for quotes and inventory levels, IA will take the best decisions for their operations by not relying on their suppliers to make decisions for them.
- **Happier customers**: As our model solves for the best solution that minimizes the gap between supply and demand, IA will be able to fulfill more of the demand, fostering a more satisfied and happier clientele as they will have access to the fresh quality produce they need.
- **Maintain/improve accessibility of fresh produce to those in need**: Adopting our model will help IA achieve long-term profitability, thus maintaining/improving the accessibility of fresh produce to the community.

Recommendations for Improvement

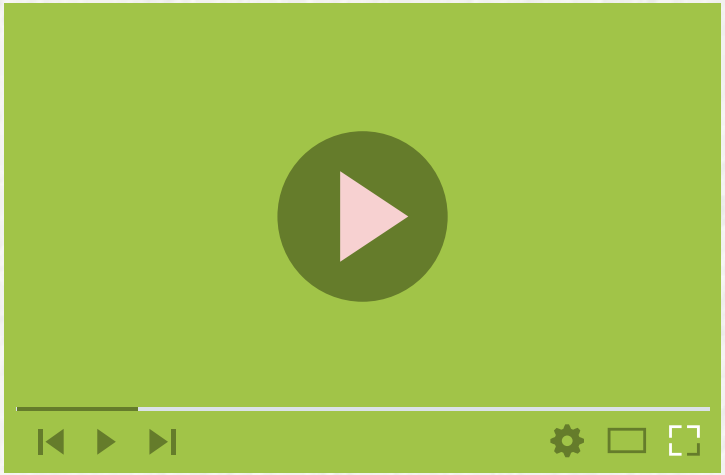
- **Adopting standardized process for collecting new data:** There is currently no standardized process for collecting new data, especially sales data. Through our project we propose a new way for IA to record data coming in, whether it is sales, invoice or even inventory data. IA will have to completely internalize our data collection process in order to use our tool to its best capacity.
- **Train and test models with real sales data:** After collecting at least one-year of data, IA will have to remove the fake data we generated from their database and re-train the model on real data only. The model competition will take care of selecting the best demand model for each produce. After training the model (demand-optimizer model), IA will have to test it out in their operations. If the model produces very low accuracy, other demand forecasting algorithms might have to be used in the model competition stage.
- **Updating dashboard with changing requirements:** If IA starts collecting other information on their market and wants to better analyze their operations and profitability, new descriptive plots should be added to the descriptive dashboard ("Statistics DB"). Similarly, if some plots that we have built do not serve them anymore, they should be removed and replaced with better ones.

06

Appendix

Video Demo
AS – Appendix for Strategy
AU – Appendix for UX/UI
AA – Appendix for Architecture
AM – Appendix for Modelling
Project Management Meeting Schedule with Innovation Youth
References
Google Drive folder with all supporting documents

Video Demo



To watch the video, please press the play button

Walkthrough Video

This video serves as a guide to understand how all parts of the project are inter-related. The video is divided in 3 sections:

1. Main Farmer's Market Sheet Overview (0:00): In this section, we will go over the description of each tab.
2. Code Overview (5:41): In this section, we will go over the different parts of the code; the demand forecasting model competition and the optimizer.
3. Dashboard Overview (10:12): In this section, we will go over the different tabs, descriptive graphs and how to read the optimizer output.

AS

Strategy Appendix

AS1. Food Assistance In Canada

Definition of Food Assistance

As per Santé Montreal, the definition of food assistance translated to English is as follows:
"Healthy eating consists of eating quality food and beverages, in adequate quantities to improve or maintain one's health, in all dignity, while obtaining pleasure, in a perspective of sustainable development.

- *Quality refers to food choices. It means eating less processed products more often, such as fruits and vegetables, dairy products and their alternatives, legumes and whole-grain cereal products. It also means eating foods high in fat, salt and sugar less often.*
- *Quantity refers to food intake, that is, eating food in adequate amounts to meet one's physiological needs.*
- *Dignity refers to the right to food and implies that food is obtained in a socially acceptable way (vs. charity).*
- *Pleasure refers to the fact that food is attractive, tasty, and eaten without guilt or obligation.*

The perspective of sustainable development is not a priority issue but faced with choices, the option that promotes sustainable development is preferred."[8]

Forms of Food Assistance


Food banks & baskets


Solidarity markets


Community kitchens & gardens

Key Stats on Food Assistance

+4000 food banks and agencies in Canada[3]

+1.3M total visits to a food bank in Canada in 2021 (only ~2300 food banks reporting), an increase of 20% since 2019[6]

33% of users were children, when they only represent less than 20% of the Canadian population[7]

Drivers of Food Assistance

Groceries - increased by 10%*

Gas – increased by 36.3%*

Inflation - consumer price index (CPI) increased by 7%*

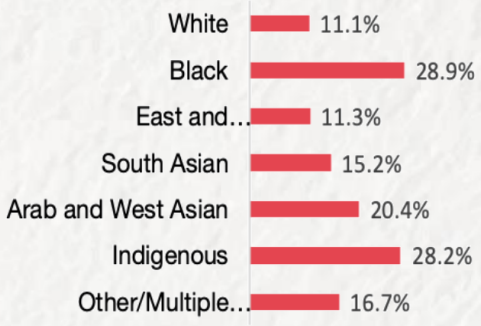
Housing/rent – housing increased by 20%**, rent increased by 4.5%*

Wages – increased by 3%***

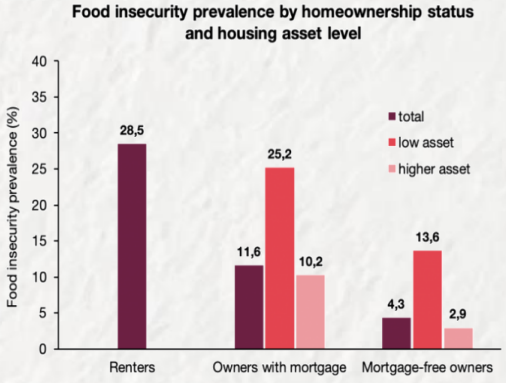
**increases as per Statistics Canada between April 2022 and the year prior*[10]
***increases as per CBC report from the Canadian Real Estate Association Canada between February 2022 and the year prior* [2]
****increases as per CBC report from Statistics Canada between February 2022 and the year prior* [4]

Back to Food Assistance landscape

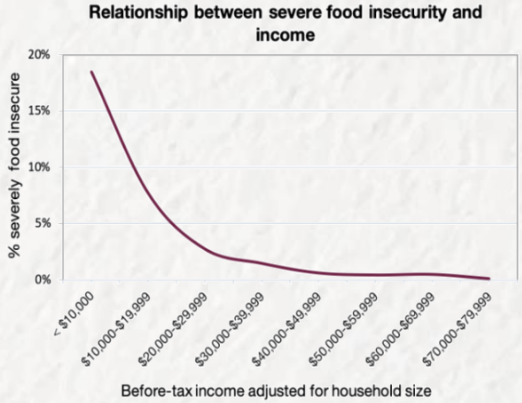
AS1. Food Assistance In Canada[1]



Food insecurity by race



Data Source: Statistics Canada, Survey of Household Spending (SHS) 2010. Low housing asset defined as home value ≤ $120, 000, representing the lowest decile of home value; higher housing asset defined as home value > $120, 000.



Data Source: Statistics Canada, Canadian Community Health Survey (CCHS), 2013-14

As previously mentioned, there was a study done in 2018 by PROOF, a research program with the University of Toronto, with the goal of understanding the food assistance landscape in Canada. In their study, they've found that those who are black or indigenous, renters, and have lower income, are the most at risk of requiring food assistance. Not having access to basic food needs, especially that of quality produce negatively impacts the quality of life of that individual, increasing the need to access health services, further straining the healthcare system, resulting in a vicious cycle. They recommend that for this cycle to be broken, Canada must implement policies that improve the financial circumstances of low-income households.

Back to Food Assistance landscape

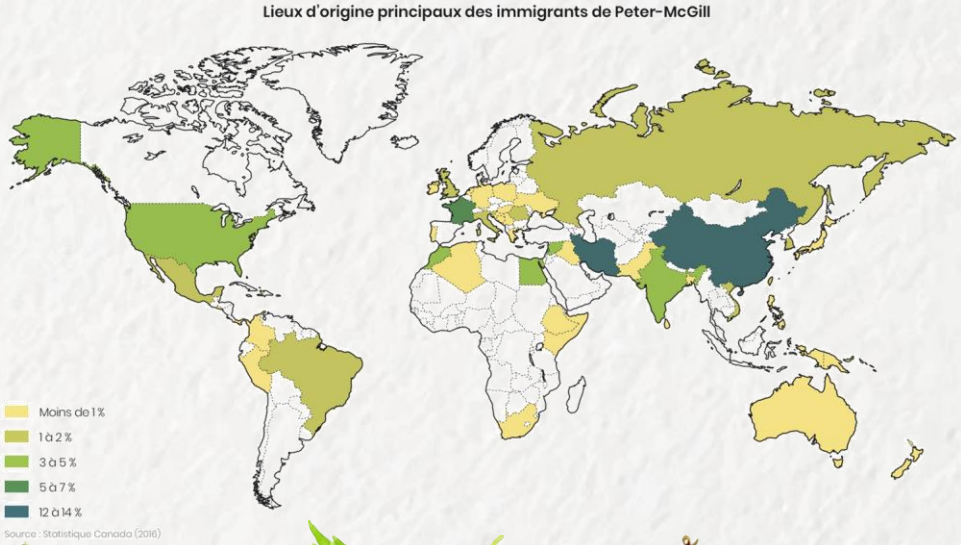
AS2. Peter McGill Community Demographics[9]

35,789 population

3,025 families

7,805 Youth between 18-24

Les principaux pays d'origine des immigrants sont la Chine, l'Iran, le Liban et la France.



Housing:

- Most housing is in 5+ story buildings
- Most housing is 1 bedroom apartments
- 1/10 households feel their apartment is too small
- 440 social housing units, representing 2.3% of households in the area

Youth and family issues

- Families in the area tend to have fewer children, approx. 60% having only 1 child, much of which is due to the high cost of living, few services, and limited access to large housing units
- Youth under 14 are few, mostly due to the families moving out of the area due to lack of public schools or requiring a larger living space
- 80% of families with children aged 0-4 are immigrants

Immigration:

- 22,260 first generation immigrants representing 63% of the population
- 13.8% of non permanent residents in Montreal live in this area alone, mostly due to the 2 universities and various colleges in the area
- Approx. 50% of residents' mother-tongue is not English or French
- Approx. 30% do not speak English or French at home

Back to Food Assistance landscape

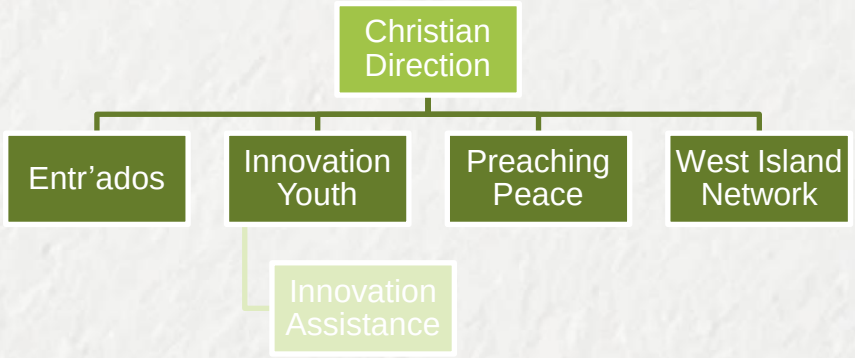
AS3. Innovation Youth

The focus of our project is done with members of Innovation Assistance, a branch of Innovation Youth. For simplicity, we have been referencing them throughout the project as Innovation Assistance (IA).

IY has a parent organization called Christian Direction, which is Christian faith-based social services organization with the mission of transforming urban neighbourhoods through various community development initiatives. Along with IY, Christian Direction has a total of 4 organizations throughout Montreal, each having their own mission:

- Entr-ados – located in Hochelaga-Maisonneuve, it's mission focuses on initiatives for youth and their families
- Innovation Youth – as previously mentioned, it's an organization focused on youth and their families in the Peter McGill community with 4 areas of focus. There are a total of 7 full time staff year round, with an additional 6 summer employees from Service Canada. Within a year, the organization sees a total of +80 volunteers, half of which are regulars
- Preaching Peace – group of pastors, leaders, and preachers that meet to discuss how to preach peace in the community
- West Island Network – group of churches and Christian organizations in the West Island whose goal is to provide social transformation work in the community

As previously mentioned, IY has a subdivision called Innovation Assistance, which focuses solely on the goal of providing those within the Peter McGill community access to fresh quality products. There are a total of 2 full time employees, Erin and Micah, whom we have had the pleasure of working with this summer.

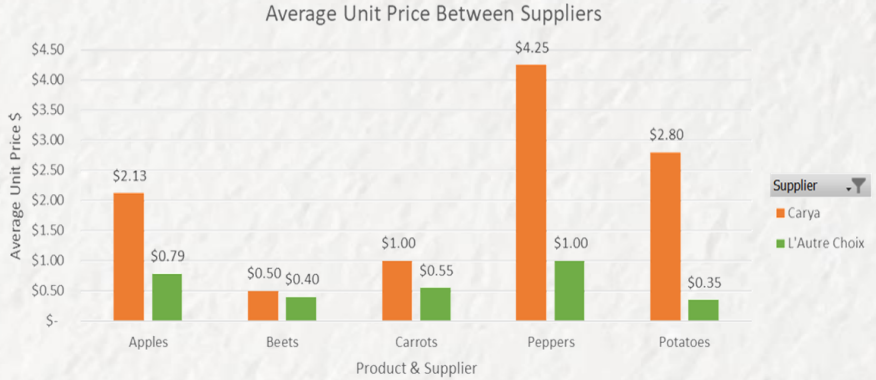


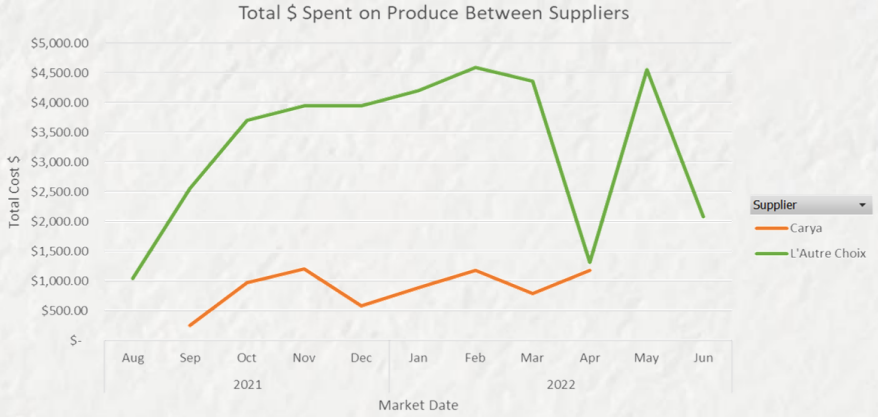
Back to Overview of the Use Case
Back to Food Assistance landscape

AS4. Solidarity Farmer's Market

L'Autre Choix (primary supplier) is an organic food store in Westmount, where the store owner has connections with farmers and sells to IA at reduced/wholesale prices. Lead times to purchase produce from them is anywhere between 1-2 weeks.

Jardins Carya has a network of farmers that produce micro greens, which typically have lead times of 3-6 weeks. They are more expensive and don't provide any special pricing.





Business recommendation/opportunity: *The McGill MBA program should incorporate a similar "community project" in order to find other sources of organic fresh produce supply for IA and negotiate lower prices. This would help the organization achieve its mission, whereby having lower costs means more funds available for more products, further increasing access to fresh produce in the Peter McGill community.*

Back to Strategy

AS5. Value Proposition Calculation

Please refer to the "Value Proposition" Google Sheet for the breakdown

	Year 1	Year 2	Year 3	Year 4	Year 5	Total	Assumptions	Links
BENEFITS								
Revenue	$ 24,000.00	$ 27,600.00	$ 31,740.00	$ 36,501.00	$ 41,976.15	$ 161,817.15	Assumes 15% growth rate per annum.	
Cost savings								
From Leftover Minimization	$ 5,979.36	$ 6,876.26	$ 7,907.70	$ 9,093.86	$ 10,457.94	$ 40,315.12	Assumes leftover loss of $498.28 per month (average leftover loss from real data) and a growth rate of 15% per annum.	
Cloud Storage	$39.99	$39.99	$39.99	$39.99	$39.99	$39.99	Assumes a cost saving of $39.99/year for 200TB of Google Storage	https://one.google.com/about/plans
Advanced Solver (e.g. Gurobi)	$ 4,000.00	$ 4,000.00	$ 4,000.00	$ 4,000.00	$ 4,000.00	$ 20,000.00	Assumes an average cost of $4,000 per annum for optimization solutions.	https://ampl.com/products/standard-price-list/
TOTAL BENEFITS	**$ 34,019.35**	**$ 38,516.25**	**$ 43,687.69**	**$ 49,634.85**	**$ 56,474.08**	**$ 222,332.22**		
COSTS								
Non-Recurring								
Staff Training	$ 561.00	$ -	$ -	$ -	$ -	$ 561.00	Assumes 7 tool operators: Six employees paid at $21/hour got access to our project's folder + one intern paid at $14.25/hour). Training should not be longer than 4 hours.	
Recurring								
Data recording	$ 2,052.00	$ 2,052.00	$ 2,052.00	$ 2,052.00	$ 2,052.00	$ 10,260.00	Assumes that an intern paid at $14.25/hour will record and insert data into our database at every market (6hours)	
Market Budget	$ 36,000.00	$ 36,000.00	$ 36,000.00	$ 36,000.00	$ 36,000.00	$ 180,000.00	Assumes a budget of $1,500 per market.	
TOTAL COST	**$ 38,613.00**	**$ 38,052.00**	**$ 38,052.00**	**$ 38,052.00**	**$ 38,052.00**	**$ 190,821.00**		
NET BENEFIT/PROFIT	**-$ 4,593.65**	**$ 464.25**	**$ 5,635.69**	**$ 11,582.85**	**$ 18,422.08**	**$ 31,511.22**		
ROI	-11.9%	1.2%	14.8%	30.4%	48.4%	16.5%		
Discount factor	0.95	0.91	0.86	0.82	0.78		Assumes a 5% discount rate per annum	
NPV	-$ 4,374.90	$ 421.09	$ 4,868.32	$ 9,529.24	$ 14,434.18	**$ 24,877.93**		

Back to Strategy

AS6. Next Steps

We are here

Proof of Concept (~ 2 months)

- Centralized Architecture (~1 month)
- Demand Forecasting Models (~1.5 weeks)
- Mixed-Integer Optimization (~1.5 weeks)
- Comprehensive and Automated Dashboard (~2 weeks)

User Acceptance Testing (+1 year)

- Collecting more Market Data (~at least 1 year's worth)*
- Model testing with real data (~2 months)
- Improving UX/UI interface to reflect new data (~2 weeks)

Project Roll-out (~ 1 month)

- Full integration of the tool in operations (~1 month)

*Please refer to AM5 on operationalizing the current models and data

AU

UX/UI Appendix

Back to Proposed Ordering Process

AU1. Step-by-Step Process for Forecasting Cycle

1. Choose which products to forecast for the next market by updating the "To forecast?" column in the "Invoice Info" tab in the master google sheet. Make sure that all the latest invoice information is added as well
2. In the "Forecast Identification" tab:
    - Double-check that each product listed within the "Forecast Identification" tab in the master google sheet has 1 choice in the "To forecast?" column. If any of the products have both "Yes" and "No" values, go back into the "Invoice Info" tab and make the correction.
    - If there are any products highlighted in red advising that they need to be added to a specific tab, please do so if the product is selected for the forecast. Otherwise, no need to add the information
        - Add products to the "Conversion" tab
        - Add products to the "Products Sold" tab
        - Add products to the "Seasonal Produce Guide" tab
        - Add products to the "Ideal-leftovers" tab
        - Add products to the "Inventory" tab
        - Add products to the "Supplier data" tab
3. In the "Aggregated Purchased Products", ensure formulas for columns E-G are copied for any additional data in the pivot table
4. In the "Products Sold" tab, ensure that all the latest market information was added

5. In the "External Data" tab,
    - Ensure a line is added for the market you are forecasting:
        - Add the market date
        - Add the forecasted temperature
        - Add the forecasted precipitation
        - Copy-paste the formula from column D onward
6. Ensure that the temperature and precipitation data of the previous market are updated with the actuals
7. In the "Market Data" tab, ensure that the latest market date and its previous date are in the table
8. In the "Inventory" tab, ensure that all the latest inventory information is up to date from the previous market
9. Received a quote from the supplier
    - Input quote information, along with the available budget for the next market in the "Supplier data" tab in the master google sheet
10. Run the forecast-optimization code in Google Collab
11. Observe the suggested grocery list in the Dashboard
12. Place an order with the supplier
13. Receive order and invoice
14. Add invoice information in the "Invoice Info" tab
15. Go to market
16. Update units sold in the "Product Sold" tab
17. Update units left-over in the "Inventory" tab
    - If any inventory is leftover that cannot be resold, donate it to members/transform it into food

Please refer to the "Documentation" Google Doc created for IA with all the information on using the proposed tool. All contents found in this PowerPoint are in the documentation

AU2. Dashboard - Tabs

Back to Descriptive Dashboard

Please refer to the "Dashboard" Google Sheet to see the build

General Category	TAB	CONTENT
Dashboard	Statistics DB	Descriptive Dashboard on Farmer's Market's health
	Forecasting DB	Grocery List for Farmer's Market
Data Import	Invoice Import	Data import from "MAIN FARMER'S MARKET" google sheet (Invoice Info tab)
	Inventory Import	Data import from "MAIN FARMER'S MARKET" google sheet (Inventory tab)
	Conversion Import	Data import from "MAIN FARMER'S MARKET" google sheet (Conversion tab)
	Products Sold Import	Data import from "MAIN FARMER'S MARKET" google sheet (Products sold tab)
	Drop-Down Formatting	Reference tab for filter formatting

Dashboard

The dashboard contains two main views; the Statistics DB and Forecasting DB views. The Statistics DB view provides IA with a descriptive understanding of their data. By using these plots, they will be able to perform exploratory data analysis and have a comprehensive view of their operations. The Forecasting DB view contains the grocery list that they will provide to their suppliers when ordering for their bi-monthly market. It also includes other information on their inventory such as amount of product produced at their garden or coming from leftovers of their previous market. Please view Forecasting Dashboard for more information on the content of the view.

Data Import

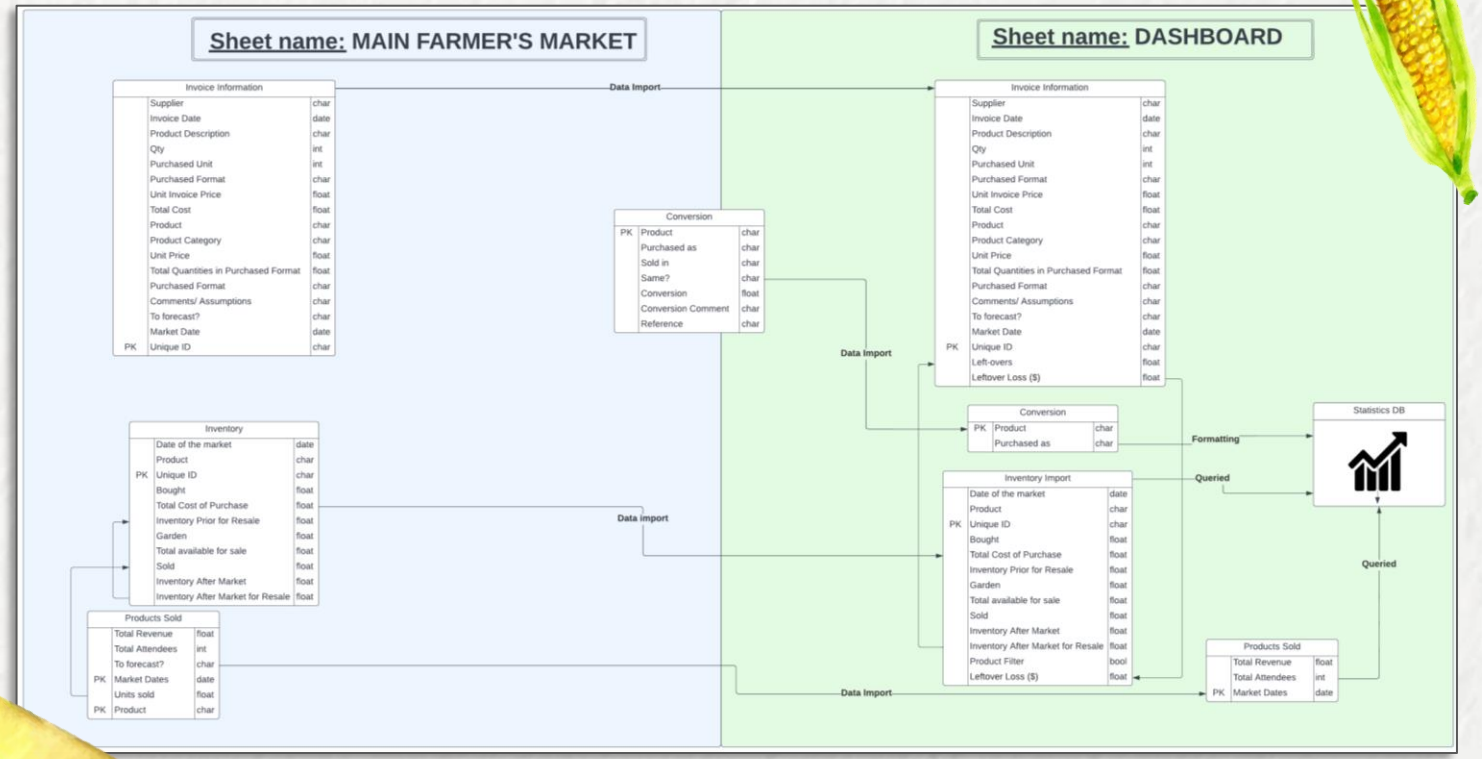
The data import tabs as seen in the Data Import Visualization slide, are only necessary to build the Statistics DB dashboard view. In fact, to build the descriptive plots, we had to query the relevant data from our database using the QUERY() function on Google Sheets. However, since the database is stored in another Google sheet (MAIN FARMER'S MARKET), we first had to important the relevant data to the DASHBOARD sheet using the IMPORTRANGE() function before using the QUERY() function.

By using the IMPORTRANGE() function, the dashboard becomes automated. In fact, changing or adding values in the relevant tabs of the MAIN FARMER'S MARKET sheet will automatically update the data imports, thus also updating the plots of the descriptive dashboard view.

AU2 Cont'd

Back to Descriptive Dashboard

AU2. Dashboard - Data Import Visualization



This part of the main ERD summarizes the different connections between the descriptive dashboard "Statistics DB" and the main database "MAIN FARMER'S MARKET". As you can see on the chart, four main tabs of the main database are imported to the DASHBOARD sheet. These are Invoice Information, Conversion, Inventory and Products Sold.

The next six slides give a general overview of the descriptive dashboard's plots including a description of which tab from the main database are associated to them as well as the query written to draw them.

AU2 Cont'd

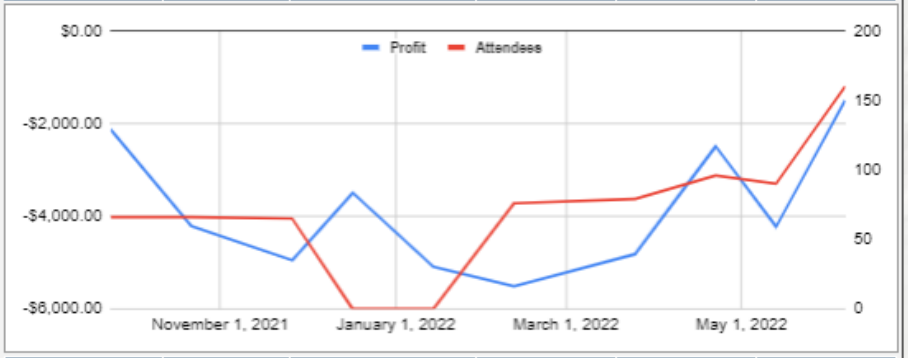
AU2. Statistics DB Plots (1/6)

Back to Descriptive Dashboard

Overall Financial Health Section

Profitability *(for the selected period)*

	Tot. Cost	Tot. Revenue	Tot. Profit	Tot. Attendees
June 6, 2022	$1,983.50	$476.19	-$1,507.31	160
May 13, 2022	-53.10%	-	-64.36%	77.78%



Total	$19,035.70	$476.19	-$18,559.51	501
Date	**Cost**	**Revenue**	**Profit**	**Attendees**
June 6, 2022	$1,983.50	$476.19	-$1,507.31	160
May 13, 2022	$4,229.30	$0.00	-$4,229.30	90
April 22, 2022	$2,494.75	$0.00	-$2,494.75	96
March 25, 2022	$4,819.15	$0.00	-$4,819.15	79
February 11, 2022	$5,509.00	$0.00	-$5,509.00	76
January 14, 2022	$5,092.22	$0.00	-$5,092.22	0
December 17, 2021	$3,494.58	$0.00	-$3,494.58	0
November 26, 2021	$4,950.35	$0.00	-$4,950.35	65
October 22, 2021	$4,212.67	$0.00	-$4,212.67	66
September 24, 2021	$2,127.67	$0.00	-$2,127.67	66

Description

This graph provides total cost, total revenue, total profit and total attendees data for IA's bi-monthly markets. It pulls data from the last 10 markets (5 months worth of markets – markets happen twice a month). The upper portion of the graphs provides data from the last market and compares these values to the second-to-last market's values. The line plot plots total profit and attendees over time. The table provides the same information for all last 10 markets ordered in descending order of dates (used to draw the plot).

Connectivity

The data is pulled from the Inventory and Products Sold import tabs.

Query

```
select A, sum(E) where A >= date """&TEXT(DATEVALUE(B3),"yyyy-mm-dd")&""" AND A <= date """&TEXT(DATEVALUE(B4),"yyyy-mm-dd")&"""
group by A order by A DESC limit 10 label A 'Date', sum(E) 'Cost'
```

where B3 and B4 are the Start Date and End Date filters, respectively.

The Revenue and Attendees values are pulled from the Products Sold table using a Vlookup of the dates. The Profit column was calculated by subtracting cost from revenue.

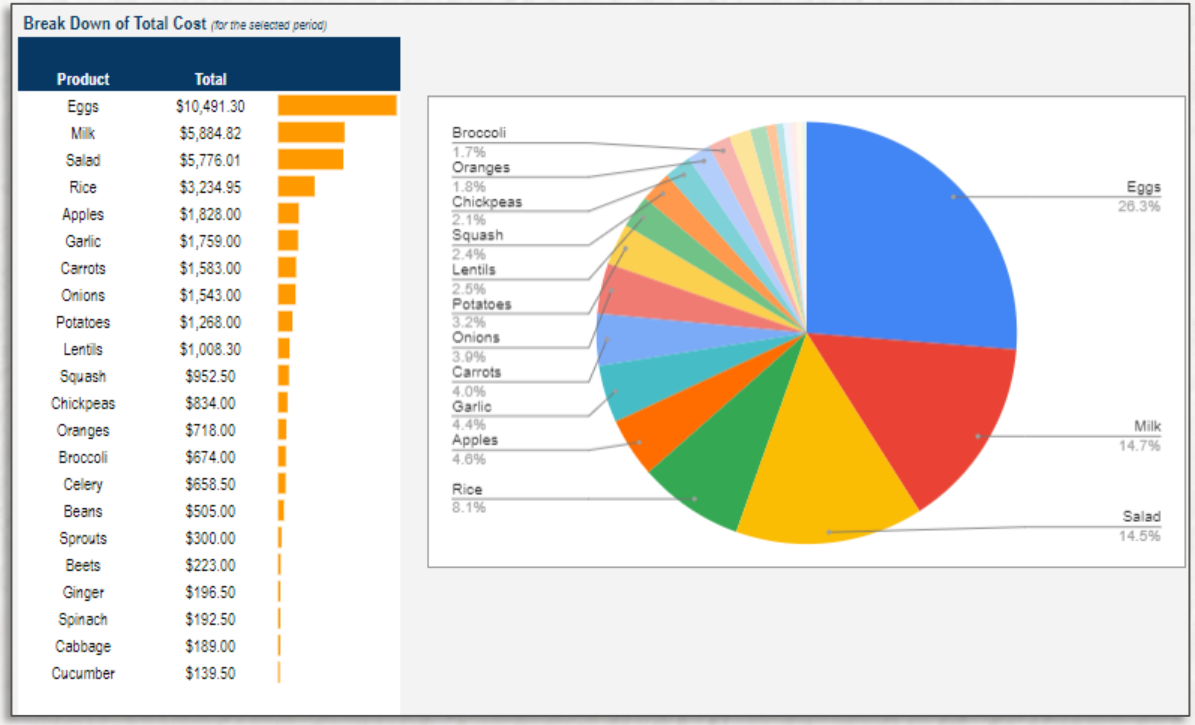
Please refer to the "Dashboard" Google Sheet, tab "Statistics DB" to see the build

AU2 Cont'd

Back to Descriptive Dashboard

AU2. Statistics DB Plots (2/6)

Overall Financial Health Section



Description

The dashboard gives a breakdown of Total Cost by produce. This allows IA to understand their spending habits overtime. The values are cumulated between the selected start and end date.

Connectivity

The data is pulled from the Inventory Import tab.

Query

```
select B, sum(E) where A >= date """&TEXT(DATEVALUE(B3),"yyyy-mm-dd")&""" AND A <= date """&TEXT(DATEVALUE(B4),"yyyy-mm-dd")&""" group by B order by sum(E) DESC limit 22 label B 'Product', sum(E) 'Total Cost'
```

where B3 and B4 are the Start Date and End Date filters respectively.

Please refer to the "Dashboard" Google Sheet, tab "Statistics DB" to see the build

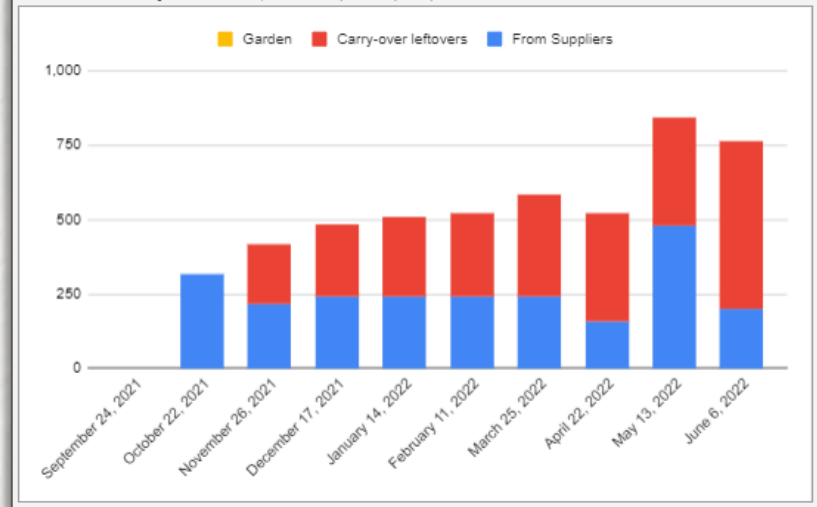
AU2 Cont'd

AU2. Statistics DB Plots (3/6)

Back to Descriptive Dashboard

Inventory Management

Pre-Market Inventory Break Down *(for the selected period and product)*



Total	2,329	2,631	0		
Date of Market	From Suppliers	Carry-over leftovers	Garden	Total Inventory	
June 6, 2022	200	563	0	763	
May 13, 2022	480	363	0	843	
April 22, 2022	157	363	0	520	
March 25, 2022	240	346	0	586	
February 11, 2022	240	281	0	521	
January 14, 2022	240	271	0	511	
December 17, 2021	240	243	0	483	
November 26, 2021	216	201	0	417	
October 22, 2021	316	0	0	316	
September 24, 2021	0	0	0	0	

Description

This graph showcases how much and where IA gets their inventory from. The From Suppliers tab shows how much IA ordered of each product from their suppliers. Carry-over leftovers shows how much inventory came from the previous market and finally Garden shows how much they produced at their community gardens.

Connectivity

The data is pulled from the Inventory Import tab.

Query

```
select A, sum(D), sum(F), sum(G) where A >= date """&TEXT(DATEVALUE(B3),"yyyy-mm-dd")&""" AND A <= date """&TEXT(DATEVALUE(B4),"yyyy-mm-dd")&""" and L = True group by A order by A DESC limit 10 label A 'Date of Market', sum(D) 'From Suppliers', sum(F) 'Carry-over leftovers', sum(G) 'Garden'
```

where B3 and B4 are the Start Date and End Date filters, respectively.

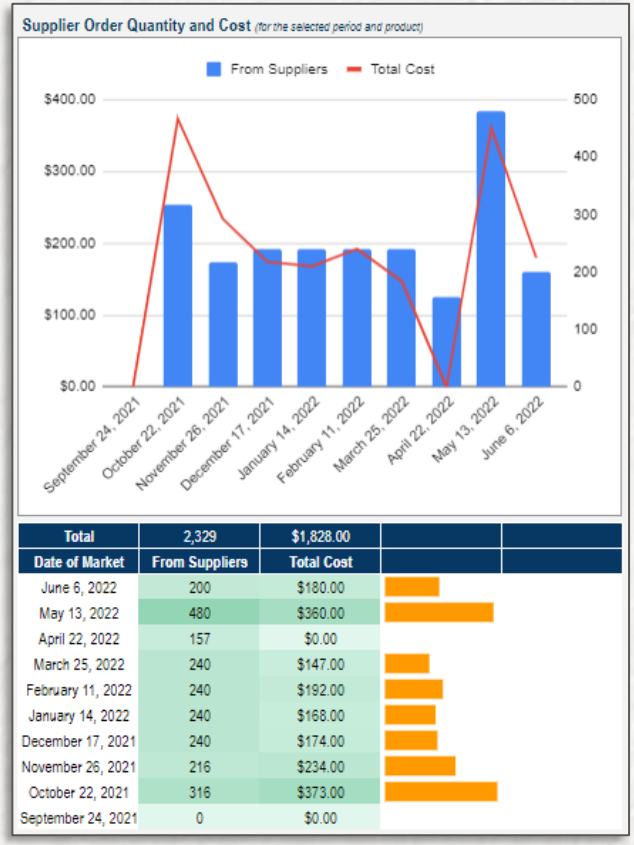
Please refer to the "Dashboard" Google Sheet, tab "Statistics DB" to see the build

AU2 Cont'd

Back to Descriptive Dashboard

AU2. Statistics DB Plots (4/6)

Inventory Management



Total	2,329	$1,828.00
Date of Market	From Suppliers	Total Cost
June 6, 2022	200	$180.00
May 13, 2022	480	$360.00
April 22, 2022	157	$0.00
March 25, 2022	240	$147.00
February 11, 2022	240	$192.00
January 14, 2022	240	$168.00
December 17, 2021	240	$174.00
November 26, 2021	216	$234.00
October 22, 2021	316	$373.00
September 24, 2021	0	$0.00

Description

This chart shows a side-by-side comparison of IA's ordering amounts and related costs. This graph would allow them to monitor any changes in quotes coming from inflation, supplier raises etc.

Connectivity

The data is pulled from the Inventory Import tab.

Query

```
select A, sum(D), sum(E) where A >= date """&TEXT(DATEVALUE(B3),"yyyy-mm-dd")&""" AND A <= date """&TEXT(DATEVALUE(B4),"yyyy-mm-dd")&""" and L = True group by A order by A DESC limit 10 label A 'Date of Market', sum(D) 'From Suppliers', sum(E) 'Total Cost'
```

where B3 and B4 are the Start Date and End Date filters, respectively.

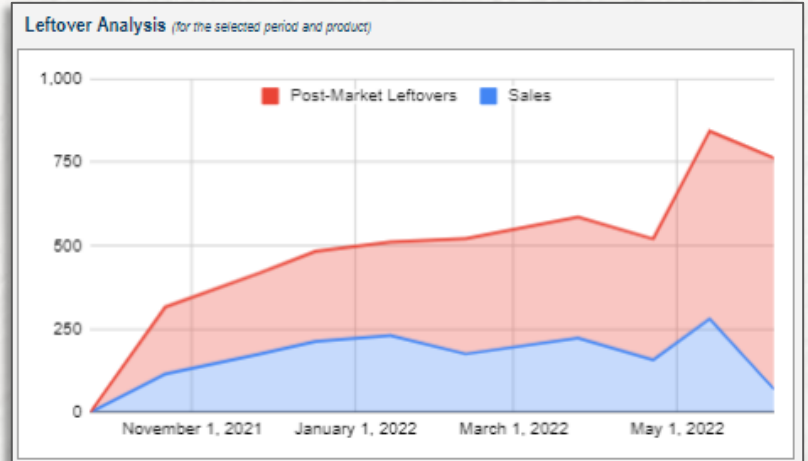
Please refer to the "Dashboard" Google Sheet, tab "Statistics DB" to see the build

AU2 Cont'd

AU2. Statistics DB Plots (5/6)

Back to Descriptive Dashboard

Inventory Management



Total	1,635	3,325		
Date of Market	Sales	Post-Market Leftovers		Δ
June 6, 2022	70	694		23.14%
May 13, 2022	280	563		55.05%
April 22, 2022	157	363		0.00%
March 25, 2022	222	363		5.11%
February 11, 2022	175	346		23.16%
January 14, 2022	230	281		3.69%
December 17, 2021	212	271		11.39%
November 26, 2021	174	243		20.70%
October 22, 2021	115	201		-
September 24, 2021	0	0		-

Description

This chart shows how much inventory was sold versus unsold (leftovers). This is a core graph for IA as their main goal is to minimize leftovers due to limited storing capacity.

Connectivity

The data is pulled from the Inventory Import tab.

Query

```
select A, sum(I), sum(K) where A >= date """&TEXT(DATEVALUE(B3),"yyyy-mm-dd")&""" AND A <= date """&TEXT(DATEVALUE(B4),"yyyy-mm-dd")&""" and L = True group by A order by A DESC limit 10 label A 'Date of Market', sum(I) 'Sales',sum(K) 'Post-Market Leftovers'
```

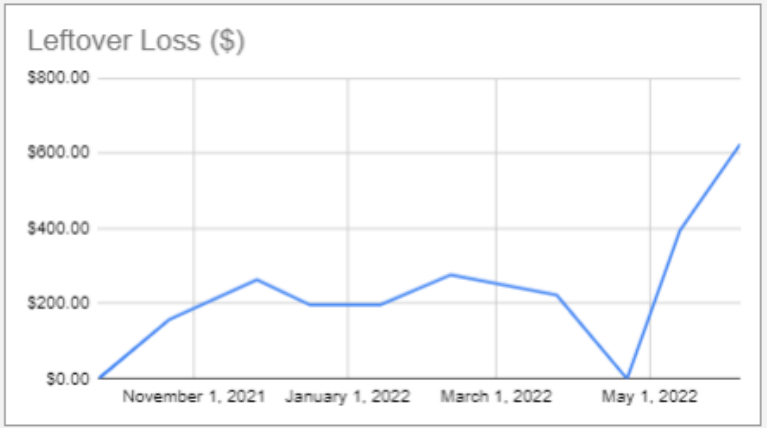
Please refer to the "Dashboard" Google Sheet, tab "Statistics DB" to see the build

AU2 Cont'd

AU2. Statistics DB Plots (6/6)

Back to Descriptive Dashboard

Inventory Management



Total	$2,330.25		
Date of Market	Leftover Loss ($)		Δ
June 6, 2022	$624.30		58.32%
May 13, 2022	$394.33		-
April 22, 2022	$0.00		-100.00%
March 25, 2022	$222.54		-19.52%
February 11, 2022	$276.53		40.75%
January 14, 2022	$196.47		0.12%
December 17, 2021	$196.23		-25.46%
November 26, 2021	$263.25		68.11%
October 22, 2021	$156.59		-
September 24, 2021	$0.00		-

Description

This graph showcases how much loss IA incurred from not selling all their inventory at their markets. Again, this is a very important chart as being a non-profit, IA needs to carefully invest their funds.

Connectivity

The data is pulled from the Inventory Import tab and from the Invoice Information import tab.

Query

```
select A, sum(M) where A >= date """&TEXT(DATEVALUE(B3),"yyyy-mm-dd")&"""  AND  A  <=  date  """&TEXT(DATEVALUE(B4),"yyyy-mm-dd")&""" and L = True group by A order by A DESC limit 10 label A 'Date of Market', sum(M) 'Leftover Loss ($)'
```

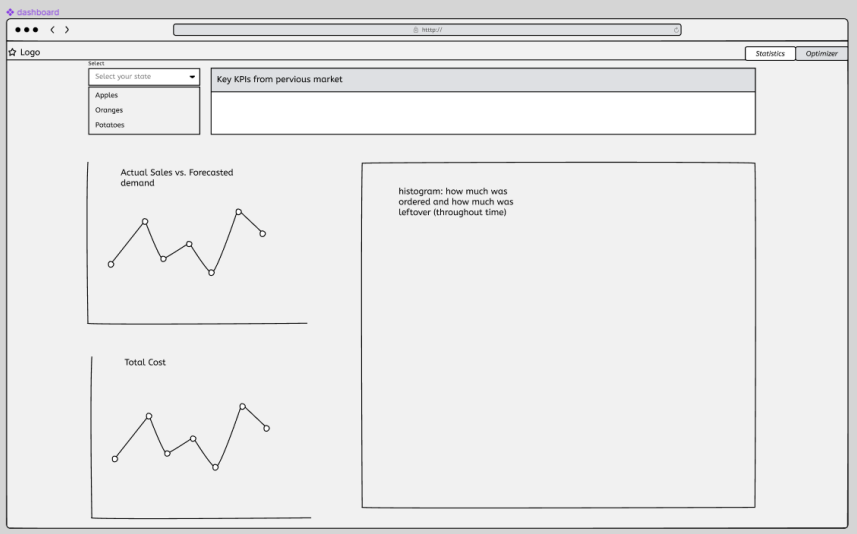
where B3 and B4 are the Start Date and End Date filters, respectively.

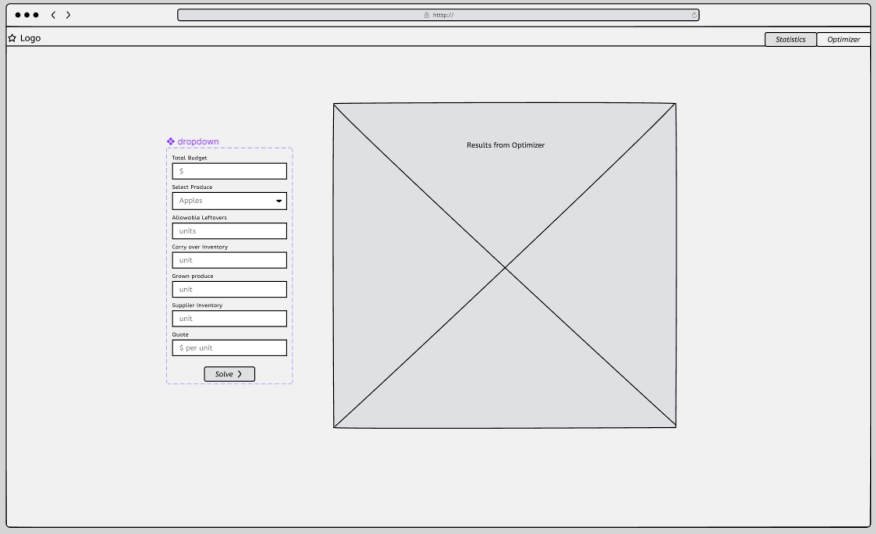
The M columns was created by using a Vlookup of unique ID (produce x market date pair)

Please refer to the "Dashboard" Google Sheet, tab "Statistics DB" to see the build

Back to Forecasting Dashboard
Back to Descriptive Dashboard

AU3. Wireframes





Background

Originally, we wanted to experiment with building a web app using Flask, JavaScript, CSS and Bootstrap. None of us had experience with any of those technologies so extended research had to be done. Due to time constraints, we then decided to simplify our UX/UI platform by opting for Tableau. However, we quickly realized two important caveats: (1) the tableau – google sheet integration was not properly developed as the tableau plots were not updating with new data, (2) with the free version of Tableau, it is not possible to save a local copy unless the dashboard is published online publicly. Due to privacy reason and customer requirements, we decided to build our dashboard on google sheets. Google sheets provides a user-friendly platform that can be fully integrated to our database and code which is ideal for our project.
Please see the Web dev research folder to see the preliminary work on Flask and HTML.

AA

Architecture Appendix

Back to Mixed-Integer Problem
Back to Forecasting Dashboard

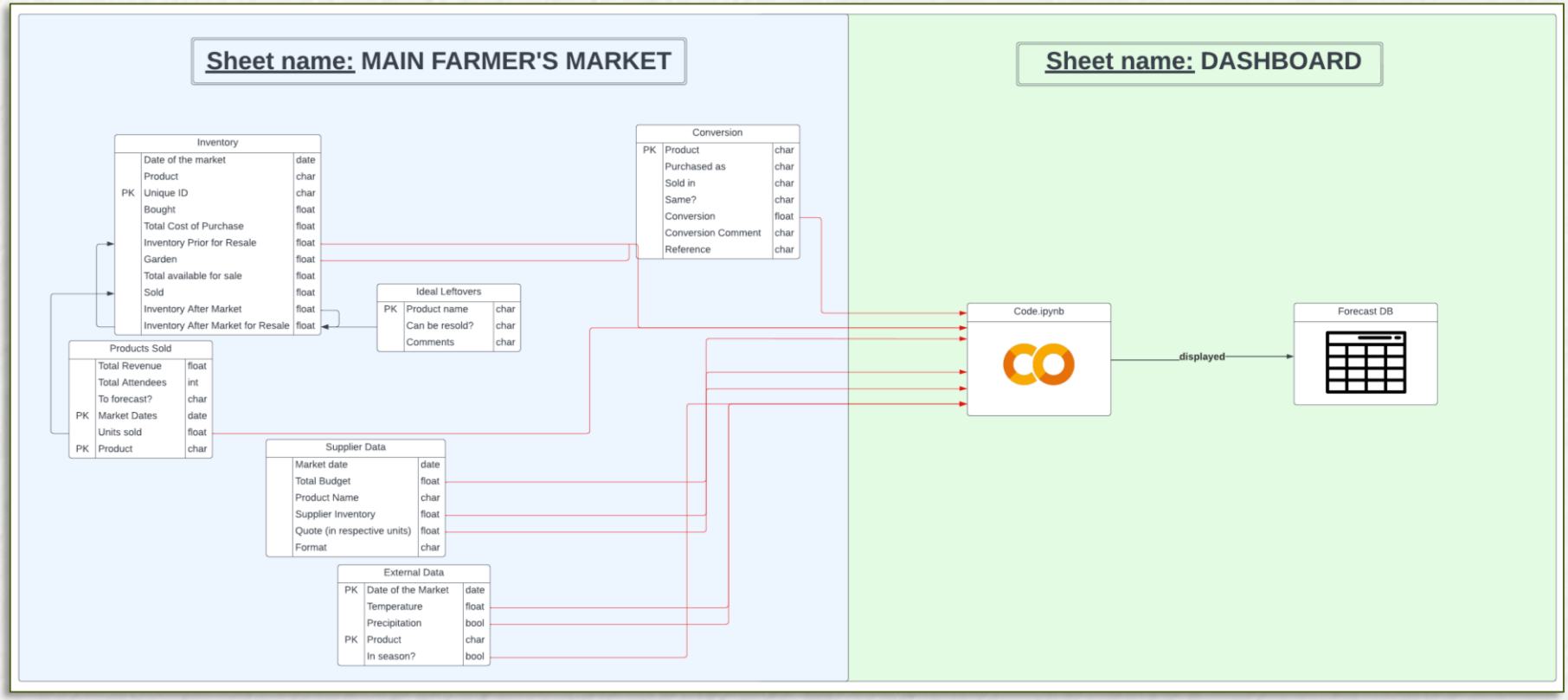
AA1. Optimizer Connectivity (1/2)

The optimizer requires some data that is directly pulled from the main database (MAIN FARMER's MARKET Google sheet). This direct connection allows the model to be flexible to new data changes. Please see AA2. Optimizer Connectivity (2/2) for a visual representation of the way the optimizer is connected to the database.

- I_i – Total suppliers' inventory for product I: This data is pulled from the Supplier Data tab in the MAIN FARMER'S MARKET Google sheet.
- D_i – Forecasted demand for product I: This data is directly calculated in the code file (Code.ipynb).
- G_i – Total inventory produced in garden for product I: This data is pulled from the Inventory tab in the MAIN FARMER'S MARKET Google sheet.
- L_i – Total leftovers from previous market for product I: This data is pulled from the Inventory tab in the MAIN FARMER'S MARKET Google sheet.
- q_i – Supplier quote (in $ per respective unit) for product I: This data is pulled from the Supplier Data tab in the MAIN FARMER'S MARKET Google sheet.
- B –Total budget: This data is pulled from the Supplier Data tab in the MAIN FARMER'S MARKET Google sheet.
- M – Very big number: This number is set in the code file (Code.ipynb).

AA1. Optimizer Connectivity (2/2)

Back to Forecasting Dashboard



Back to ERD

AA2. Main Farmer's Market Google Sheet - Summary Tab

When you open the Main Farmer's Market Google sheet (called "main farmer's market"), the first tab that opens is the "Summary", which lists all the tabs within the file, a link to that specific tab for easy navigation, and a brief description on what that tab contains:

	A	B	C
1	**General Category**	**TAB**	**CONTENT**
2	Invoices	Invoice Info	Contains information copied directly from invoices
3		Categories	Contains the grouped categories and dropdown list that is within the "Invoice Info" tab
4		Forecast Identification	Pivot table pulling from "invoice Info" that is used as a QA check to ensure that each product only has 1 value selected for forecasting
5		Conversion	Contains a table with the conversion factor between the units purchased and units sold at the market
6		Aggregated Purchased Products	Pivot table pulling from "Invoice Info" that automatically updates. It is aggregated by the unique product ID and sums the total quantity purchased and total cost
7	Products Sold	Products Sold	Contains the units sold of each product at the markets, as well as the total revenue made and the total number of attendees. This is manually recorded and pulls the list of products, along with it's forecast decision from the "Forecast Identification" tab
8		Seasonal Produce Guide	Contains information on when produce is in season at farm's in Quebec.
9		External Data	Contains external information such as the temperature on the day of the market, if there was precipitation, and pulls information from the "Seasonal Produce Guide" tab.
10	Inventory	Market Dates	Displays a list of the dates of each market, as well as the date of the previous market.
11		Ideal Leftovers	Used to help identify shelf-life of the produce and to set minimum quotas to have in stock for the next market
12		Inventory	Contains the list of total inventory left over that is still remaining for re-sale for the next market. It pulls purchased information from the "Aggregated Purchased Products" tab, the "Products Sold" tab, the "Market Dates" tab, and the "Ideal left-overs" tab.
13		Supplier Data	Used to record the supplier's Quote information regarding what they have available for sale.

Within the Google sheet, there are 3 key sections/categories that are important to keep updated, and that which are used within the tool:

- Invoice Information
- Products Sold
- Inventory

In the beginning stages of the project, we had originally separated each category as its own Google sheet. However, as we were advancing in the project, we noticed that it would be difficult to operate in this way, as there was information that needed to be imported within these tabs, creating a more complex file structure and redundant data being referenced in the files. Therefore, we decided it was best to combine all 3 categories into 1 master file as a "one-stop-shop" for data collection, making it much easier to maintain for the IA users. For the original files, please refer to the "Draft architecture files & synthetic data" folder on the Google Drive.

Please refer to the "Main Farmer's Market" Google Sheet to see this tab

AA2 Cont'd

Back to ERD
Back to Data Sources

AA2. Invoice Info Tab

This Google sheet contains information from invoices. Columns A to H has information that is copied directly from the invoices. Total Cost (Column H), is calculated as the quantity (Column D) x Unit Invoice Price (Column G). All this information must be manually entered (if the original invoice is only provided as a PDF) or can be copy-pasted from an excel invoice.

	A	B	C	D	E	F	G	H
1	Supplier	Invoice Date	Product Description	Qty	Purchased Unit	Purchased Format	Unit Invoice Price	Total Cost
2	L'Autre Choix	2021-10-24	Apples (cortland, spartan, and mcintosh 36lbs)	6	36	lbs	$28.00	$168.00
3	L'Autre Choix	2022-06-08	Apples (macintosh and cortland)	5	40	lbs	$36.00	$180.00
4	L'Autre Choix	2021-11-21	Apples (Macintosh, Spartan, Cortland 36lbs)	6	36	lbs	$39.00	$234.00

- Columns I to Q are additional information that is manually recorded.
- Column I is a general product description based on the Product Description in column C.
- Column J is a product category that is based on the value written in Column I. This is already predefined in a list within the tab called "Categories" and uses a VLOOKUP based on the value in column I. If there is a new product in column I, you must add it to the list in the "Categories" tab and assign it an appropriate category. For example, if you never purchased apples before, you need to record this and assign a group to it, such as "fruits".
- Column K is the unit price calculated as the division of the Unit Invoice Price (Column G) over the Purchased Unit (Column E).
- Column L is the total quantities within the purchased format and calculated as the Purchased Unit (Column E) multiplied by the Quantity (Column D).
- Column M references the Purchased Format in Column F for readability.
- Column N is for any comments or assumptions as required.
- Column O is a drop-down selection that allows you to determine whether you want to forecast this product. The data within the dropdown reference values within the "Categories" tab.
- Column P is the market date on which that purchased product was for.
- Column Q is a unique identifier that is to be used to link with other tabs within the Google sheet. It is the concatenation of the Market Date (Column P) and the Product description (Column I). Within the "Invoice Info" tab, the same Unique ID may exist in more than one row if the product was purchased more than once for the same market. This will not be an issue for linking between other tabs, as these values would be aggregated within a pivot table ("Aggregated Purchased Products") for referencing between tabs.

	I	J	K	L	M	N	O	P	Q
1	Product	Product Category	Unit Price	Total Quantities in Purchased Format	Purchased Format	comments/assumptions	To forecast?	Market Date	Unique ID
2	Apples	Fruits	$0.78	216	lbs		Yes	2021-10-22	2021-10-22-Apples
3	Apples	Fruits	$0.90	200	lbs	assuming 1 qty = 40lbs	Yes	2022-06-06	2022-06-06-Apples
4	Apples	Fruits	$1.08	216	lbs		Yes	2021-11-26	2021-11-26-Apples

For the values in Columns J and O, here is a sample of the options used within the drop-downs that come from the "Categories" tab:

	A	B	C	D	E
1	Product	Product Category			Yes
2	Apples	Fruits			No
3	Cabbage	Vegetables - Cabage			
4	Carrots	Vegetables - root			

Please refer to the "Main Farmer's Market" Google Sheet to see this tab

AA2 Cont'd

Back to ERD

AA2. Forecast Identification Tab

This tab contains a pivot table from the "Invoice Info" tab, in which the Product (column A) and the "To Forecast?" (column B) columns are used as rows to display the decision for each product on whether to forecast or not. This is used as a quality assurance (QA) check to ensure that all values of "To Forecast?" are the same for each product. If a product has both "Yes" and "No" written, you need to go back into the "Invoice Info" tab and make the correction. This pivot table will automatically update and requires no manual intervention. Columns D to I are additional columns used for QA to ensure that if there is a new product that is purchased and has "Yes" for the "To Forecast?", its information in these tabs must be added to ensure that the forecast can run without error. Information that is missing for new products uses a red highlight with conditional formatting, advising that the product is missing in the tab. These columns use various formulas that can simply be copy-pasted for new products and requires no additional modifications to.

- Column D uses a Vlookup with the "Conversion" tab
- Column E uses a Vlookup with the "Products Sold" tab
- Column F uses an Hlookup with the "Seasonal Produce Guide" tab
- Column G uses a Vlookup with the "Ideal leftovers" tab
- Column H uses a Vlookup with the " Inventory" tab
- Column I uses a Vlookup with the " Supplier data" tab

	A	B	D	E	F	G	H	I
1	*Product*	*To forecast?*	Product listed in the "Conversion" tab?	Product listed in the "Products Sold" tab?	Product in the "Seasonal Produce Guide" tab?	Product in the "Ideal left-overs" tab?	Product in the "Inventory" tab?	Product in the "Supplier data" tab?
2	Apples	Yes	Apples	Apples	Apples	Apples	Apples	Apples
3	Bananas	No	Need to add product to the tab	Bananas	Bananas	Bananas	Need to add product to the tab	Bananas
4	Beans	Yes	Beans	Beans	Beans	Beans	Beans	Beans
5	Beets	Yes	Beets	Beets	Beets	Beets	Beets	Beets
6	Broccoli	Yes	Broccoli	Broccoli	Broccoli	Broccoli	Broccoli	Broccoli
7	Brussels Sprouts	No	Need to add product to the tab	Brussels Sprouts	Brussels Sprouts	Brussels Sprouts	Need to add product to the tab	Brussels Sprouts
8	Cabbage	Yes	Cabbage	Cabbage	Cabbage	Cabbage	Cabbage	Cabbage
9	Canteloupe	No	Canteloupe	Canteloupe	Need to add product to the tab	Canteloupe	Need to add product to the tab	Canteloupe

Please refer to the "Main Farmer's Market" Google Sheet to see this tab

AA2 Cont'd

Back to Forecasting Models
Back to ERD

AA2. Conversion Tab

	A	B	C	D	E	F	G
1	Product	Purchased as	Sold in	Same?	Conversion	Conversion Comm	Reference
2	Apples	lbs	unit	different	0.3333333333	1 apple = X lbs	1 pound of apples = 3 apples, https://www.best-easy-apple-recipes.com/how-many-apples.html
3	Beans	kg	scoop (assume 1 cup)	different	0.18	1 scoop = X kg	1 cup of dried beans = 0.18 kg, https://www.howmany.wiki/vw/--1--cup--of--dried-beans--in--lb
4	Beets	lbs	unit	different	0.3333333333	1 beet = X lbs	1 pound of beets = 3 beets, https://www.howmuchisin.com/produce_converters/beet#:~:text=After%20surveying%20the%20vegetable%20selection,beets%20in%20a%20typical%20bunch.
5	Broccoli	unit	unit	same	1		
6	Cabbage	lbs	unit	different	2	1 cabbage = X lbs	1 cabbage = 2 pouds, https://cooking.stackexchange.com/questions/7239/cabbage-volume-to-weight-conversion
7	Canteloupe	unit	unit	same	1		
8	Carrots	lbs	unit	different	0.2	1 carrot = X lbs	1 pound of carrots = 5 carrots, https://www.thekitchn.com/heres-what-1-pound-of-carrots-looks-like-232314
9	Cauliflower	unit	unit	same	1		
10	Celery	unit	unit	same	1		
11	Chickpeas	kg	scoop (assume 1 cup)	different	0.211	1 scoop = X kg	1 scoop of chickpeas = 211 grams need to convert gram to kg (1g = 0.001 kg) https://freefoodtips.com/dried-chickpeas-volume-to-weight-conversion/
12	Corn	dozen	dozen	same	1		
13	Cucumber	unit	unit	same	1		
14	Eggs	dozen	dozen	same	1		
15	Garlic	lbs	unit	different	0.1	1 garlic = X lbs	1 pound of garlic = 10 garlic heads, https://garlicseed.ca/products/bulk-seed-garlic#:~:text=On%20average%20there%20are%2010,to%205%20cloves%20per%20bulb.
16	Ginger	lbs	unit	different	0.22	1 ginger = X lbs	1 ginger = 0.22 pounds (had to weigh it ourselves)
17	Lentils	kg	scoop (assume 1 cup)	different	0.3	1 scoop = X kg	1 serving of lentils = 1.5 cups, https://www.acouplecooks.com/lentil-to-water-ratio/#:~:text=How%20much%20is%201%20serving,of%20cooked%20lentils%20per%20serving. 1.5 cups of lentils = 0.3kg, https://coolconversion.com/cooking-volume-weight/1~cup~of~dry+lentils~to~gram
18	Milk	litre	bag	different	1.3	1 bag = 1.3 litre	1 bag of milk = 1.3L as per IA
19	Onions	lbs	unit	different	0.5	1 onion = X lbs	1 pound of onions = 2 onions, https://www.thekitchn.com/heres-what-1-pound-of-onions-looks-like-231826
20	Oranges	unit	unit	same	1		
21	Pear	unit	unit	same	1		
22	Peppers	lbs	unit	different	0.25	1 pepper = X lbs	1 pound of bell peppers = 4 peppers, https://www.howmuchisin.com/produce_converters/bell-pepper#:~:text=To%20answer%20How%20many%20bell,or%203%20large%20bell%20peppers.
23	Potatoes	lbs	unit	different	0.5	1 potatoe = X lbs	1 pound of potatoes = 2 potatoes, https://afoodloverskitchen.com/how-many-potatoes-are-in-a-pound/
24	Radish	kg	unit	different	0.0045	1 radish = Xkg	1 radish = 4.5g, converted as 0.0045kg, https://hannaone.com/Recipe/weightradish.html
25	Rice	kg	scoop (assume 1 cup)	different	0.19	1 scoop = X kg	1 cup of rice = 0.19kg, https://www.traditionaloven.com/conversions_of_measures/rice_amounts_converter.html
26	Salad	kg	kg	same	0.1333333333	1 salad = X kg	1 portion is 0.133 kg, come in 4kg bags and each bag has 30 portions
27	Spinach	unit	unit	same	1	1 spinach = X unit	
28	Sprouts	kg	kg	same	0.1333333333	1 sprouts = X kg	1 portion is 0.133 kg, come in 4kg bags and each bag has 30 portions
29	Squash	lbs	unit	different	2	1 squash = X lbs	1 squash = 2 pounds, https://www.howmuchisin.com/produce_converters/acorn-squash
30	Tomatoes	lbs	unit	different	0.3333333333	1 tomato = X pounds	1 pound of tomatoes = 3 tomatoes, https://tastylicious.com/how-many-tomatoes-in-a-pound/
31	Zucchini	lbs	unit	different	0.43	1 zucchini = Xlbs	1 zucchini = 196 grams, converted as 0.43lbs, https://recipeland.com/ingredients/zucchini-7900

This tab contains information on assumptions made for converting the purchased unit to the sold unit. For example, apples are purchased in pounds. However, when they are sold at the market, they are sold on a "per unit" (i.e., 1 apple) basis.

- Column A contains the product description (needs to be written the same way as the product descriptions in column I of the "Invoice Info" tab).
- Column B is the unit in which the product is purchased and uses a VLOOKUP to pull this information from column M in the "Invoice Info" tab.
- Column C is the unit the product is sold in (needs to match as a consistent unit across all markets).
- Column E contains the conversion value between the "Sold in" unit vs. the "Purchased as" unit.
- Column F clarifies what the value in Column F means. Wherever there is an "X", this means the value corresponds to what is in Column F.
- Column G is the reference used to make the conversion assumptions. It is important to note that if there is a product you want to forecast and it is not currently included in this tab, you must add it to this table, along with its appropriate conversion value.

Please refer to the "Main Farmer's Market" Google Sheet to see this tab

AA2 Cont'd

Back to Data Sources
Back to ERD

AA2. Aggregated Purchased Products Tab

This tab contains a pivot table using information from the "Invoice Info" tab, along with the "Conversion" tab. The pivot table uses the "Unique ID" (column A) and the "Product" description (column B) as rows, along with the "Total Quantities in Purchased Format" (column C) and "Total Cost" (column D) as aggregated sums. Column E references the value in column C, as making calculations using direct references from the pivot does not work. Column F uses a VLOOKUP to bring in the conversion factor value of the product from the "Conversion" tab. Lastly, Column G converts the value in Column E using the conversion value in Column F. It is ok to have values in this column as blank, so long as it is not a product to be forecasted. If it is a product to be forecasted, it is important to record the necessary information within the "Conversion" tab.

	A	B	C	D	E	F	G
1	*Unique ID*	*Product*	SUM of Total Quantities in Purchased Format	SUM of Total Cost		conversion factor	#units purchased converted
2	2021-08-13-Milk	Milk	208	$345.92	208	1.3	160
3	2021-08-27-Eggs	Eggs	225	$699.75	225	1	225
4	2021-09-24-Broccoli	Broccoli	90	$180.00	90	1	90
5	2021-09-24-Canteloupe	Canteloupe	40	$80.00	40	1	40
6	2021-09-24-Carrots	Carrots	288	$90.00	288	0.2	1440
7	2021-09-24-Celery	Celery	48	$52.00	48	1	48
8	2021-09-24-Corn	Corn	15	$69.00	15	1	15
9	2021-09-24-Cucumber	Cucumber	48	$72.00	48	1	48
10	2021-09-24-Eggplant	Eggplant	200	$80.00	200		

Please refer to the "Main Farmer's Market" Google Sheet to see this tab

AA2 Cont'd

Back to Data Sources
Back to ERD

AA2. Products Sold Tab

This tab displays how much of each type of product was sold at the markets, along with the revenue generated, and how many people attended the market.

- Row 2 contains the total revenue information that requires to be manually entered.
- Row 3 contains the total attendee information and requires to be manually entered.
- Column A contains the product description. These values are directly pulled from the "Forecast Identification" tab
- Column B contains the "To Forecast?" values directly from the "Forecast Identification" tab.
- Columns C to Q contains information on how much of each product was sold at the market. The values pull from the "Solidarity Market Inventory" file will automatically be populated by copying the formula for the previous market and just changing the name of the tab to the respective market's tab. **Important** - the "Solidarity Market Inventory" must be the same file structure each time. Therefore, in the second row, it is important to include the total sum of the products sold so that these values can get pulled correctly into the Products Sold tab.

For demonstration purposes, we created synthetic data within the sheet that comes from 3 sources:

- Market dates in green = real market data provided by IA
- Market dates in orange = uses synthetic data from the Synthetic Data Google sheet
- Market dates in yellow = uses a random function that pulls from the green and orange market dates

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1					Market												
2	Total Revenue		$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$476.19	$966.65
3	Total Attendees		35	76	42	0	66	66	65	0	0	76	79	96	90	160	0
4	Product	To forecast?	2021-07-09	2021-07-23	2021-08-13	2021-08-27	2021-09-24	2021-10-22	2021-11-26	2021-12-17	2022-01-14	2022-02-11	2022-03-25	2022-04-22	2022-05-13	2022-06-06	2022-06-20
5	Apples	Yes	586	187	279	58	0	344	523	637	690	525	667	470	840	209	192
6	Asparagus	No	0	0	0	0	0	0	0	0	0	0	0	0	0	0	68
7	Bananas	No	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	Beans	Yes	94	19	183	53	0	112	188	0	0	7	8	0	0	0	0
9	Beets	Yes	185	100	261	96	0	139	0	321	0	91	68	0	117	0	0
10	Broccoli	Yes	3	34	25	54	46	0	0	0	0	27	40	0	77.22	0	0
11	Brussels Sprouts	No	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	Cabbage	Yes	20	56	32	15	0	46	0	59	0	16	0	0	0	0	0

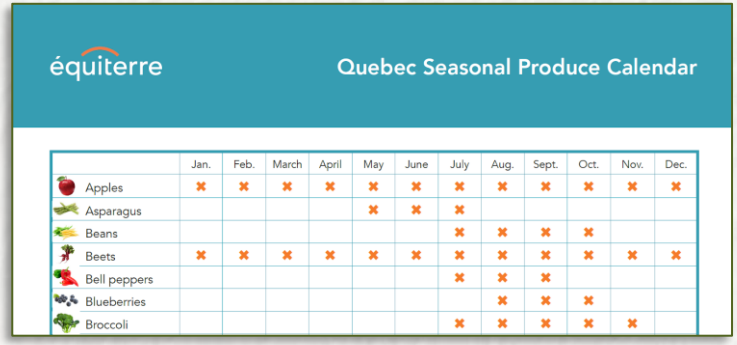
Please refer to the "Main Farmer's Market" Google Sheet to see this tab

AA2 Cont'd

Back to Data Sources
Back to ERD

AA2. Seasonal Produce Guide Tab

This tab contains information from équiterre on when products are in season in Quebec. This information was manually entered, along with additional products (highlighted in orange) that were added from online research. Where it is written "1" means that the product is in season during that month, otherwise, it is not in season. It is important to note that products that were purchased and that are not currently on the list must be added if they are to be forecasted. This is because some forecasting models incorporate this information, which would otherwise result in an error when running the code.



	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Apples	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Asparagus					✖	✖	✖					
Beans							✖	✖	✖	✖		
Beets	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Bell peppers							✖	✖	✖			
Blueberries								✖	✖	✖		
Broccoli							✖	✖	✖	✖	✖	

	A	B	C	D	E	F	G	H
1		Apples	Asparagus	Bananas	Beans	Beets	Bell Peppers	Blueberries
2	January	1	0	1	0	1	0	0
3	February	1	0	1	0	1	0	0
4	March	1	0	1	0	1	0	0
5	April	1	0	1	0	1	0	0
6	May	1	1	1	0	1	0	0
7	June	1	1	1	0	1	0	0
8	July	1	1	1	1	1	1	0
9	August	1	0	1	1	1	1	1
10	September	1	0	1	1	1	1	1
11	October	1	0	1	1	1	0	1
12	November	1	0	1	0	1	0	0
13	December	1	0	1	0	1	0	0

Please refer to the "Main Farmer's Market" Google Sheet to see this tab

AA2 Cont'd

Back to Data Sources
Back to ERD

AA2. External Data Tab

This tab contains various sources of external information that are used within the forecasting models. Columns A-C need to be manually entered, where column A contains the date of the market, column B the temperature in Celsius on the day of the market, and column C a binary value that indicates if there was precipitation (i.e., rain, snow). Columns B and C can utilize any weather forecast website to get this information. Columns D onward use an Hlookup formula, along with an "if" statement to pull in the seasonal produce guide information that would display whether each product is in season during the specified date of the market. The list of products in row 1 starting in column D are automatically updated and linked to the "Seasonal Produce Guide" tab.

D2 `=if(isblank(hlookup(D$1,'Seasonal Produce Guide'!$B$1:$BM$13,MONTH($A2)+1,FALSE)),0,hlookup(D$1,'Seasonal Produce Guide'!$B$1:$BM$13,MONTH($A2)+1,FALSE))`

	A	B	C	D	E	F	G	H	I	J	K	L
1	Date of the Mar	Temperature	Precipitation	Apples	Asparagus	Bananas	Beans	Beets	Bell Peppers	Blueberries	Broccoli	Brussels Sprouts
2	2021-07-09	18.5	0	1	1	1	1	1	1	0	1	0
3	2021-07-23	21.4	0	1	1	1	1	1	1	0	1	0

It is important to note that when you are planning on generating a forecast for your next market, **you must add the necessary information about that upcoming market to the file**. You would essentially have to add a new row, input the date of the market, what the weather and precipitation forecast is, and drag the formulas from column D onward to your new row. Once the market happens, it is important to update the weather and precipitation information to what was recorded on that day. It was designed this way to minimize any contact with the Google Colab code file, considering the users of the tool do not understand the code. The following sample shows how the values for the upcoming market are highlighted in yellow.

	A	B	C	D	E	F	G	H	I	J
1	Date of the Mar	Temperature	Precipitation	Apples	Asparagus	Bananas	Beans	Beets	Bell Peppers	Blueberries
2	2021-07-09	18.5	0	1	1	1	1	1	1	0
3	2021-07-23	21.4	0	1	1	1	1	1	1	0
4	2021-08-13	26.7	0	1	0	1	1	1	1	1
5	2021-08-27	21.3	0	1	0	1	1	1	1	1
6	2021-09-24	18.4	1	1	0	1	1	1	1	1
7	2021-10-22	8.9	1	1	0	1	1	1	0	1
8	2021-11-26	0.7	1	1	0	1	0	1	0	0
9	2021-12-17	4.7	0	1	0	1	0	1	0	0
10	2022-01-14	-14.8	0	1	0	1	0	1	0	0
11	2022-02-11	2.6	1	1	0	1	0	1	0	0
12	2022-03-25	5.4	1	1	0	1	0	1	0	0
13	2022-04-22	8.1	0	1	0	1	0	1	0	0
14	2022-05-13	22.5	0	1	1	1	0	1	0	0
15	2022-06-06	18.2	0	1	1	1	0	1	0	0
16	2022-06-20	25	0	1	1	1	0	1	0	0
17	2022-07-04	26	0	1	1	1	1	1	1	0

Please refer to the "Main Farmer's Market" Google Sheet to see this tab

AA2 Cont'd

Back to ERD

AA2. Market Dates Tab

This tab contains the list of market dates (column A), along with the date of the previous market (column B). Dates in column A must be manually entered for each market date. Dates in column B reference the previous date in column A. To add a new date, always insert from row 2 and copy the formula in B.

	A	B
1	Market Date	Previous market date
2	2022-06-20	2022-06-06
3	2022-06-06	2022-05-13
4	2022-05-13	2022-04-22
5	2022-04-22	2022-03-25
6	2022-03-25	2022-02-11
7	2022-02-11	2022-01-14
8	2022-01-14	2021-12-17
9	2021-12-17	2021-11-26
10	2021-11-26	2021-10-22
11	2021-10-22	2021-09-24
12	2021-09-24	2021-08-27
13	2021-08-27	2021-08-13
14	2021-08-13	2021-07-23
15	2021-07-23	2021-07-09
16	2021-07-09	

Please refer to the "Main Farmer's Market" Google Sheet to see this tab

AA2 Cont'd

Back to Data Sources
Back to ERD

AA2. Ideal left-overs tab

This tab contains information on the shelf-life of each product and whether it can be resold at the next market. Column A contains the list of products. Column B contains a "yes" or "no" value on whether the product can be resold at the market. This is an assumption made from the information in column C. Column C contains comments and web links to back up the assumptions written in column B. If there is a new product, it is important to add the information for it in this tab if it is being forecasted.

	A	B	C
1	**Product Name**	**Can be resold?**	**Comments**
2	Apples	Yes	longest is 4-6 weeks in the fridge, assuming they were still in their previous location where they had virtual unlimited fridge space https://www.healthline.com/nutrition/how-long-do-apples-last#:~:text=On%20the%20counter%3A%205%E2%80%937,8%20months%20in%20the%20freezer
3	Bags	Yes	
4	Bananas	No	can last 2-5 in the pantry, 5-7 days in the fridge, assume they can't be resold, https://www.stilltasty.com/fooditems/index/16451
5	Beans	Yes	dry beans, assume they can be resold
6	Beets	No	beets can last 2-3 weeks in the freezer, therefore can assume that no inventory was leftover https://www.myrecipes.com/extracrispy/how-to-store-beets#:~:text=Once%20you've%20chopped%20off,between%20two%20to%20three%20weeks.
7	Broccoli	No	broccoli lasts 3-5 days in the fridge, therefore can assume that no inventory was leftover https://www.allrecipes.com/article/how-to-store-broccoli/#:~:text=How%20Long%20Does%20Broccoli%20Last,12%20months%20in%20the%20freezer.
8	Brussels Sprouts	No	last 7-10 days in the fridge, therefore assume no inventory, https://www.doesitgobad.com/do-brussel-sprouts-go-bad/

Please refer to the "Main Farmer's Market" Google Sheet to see this tab

AA2 Cont'd

Back to ERD

AA2. Inventory Tab

This tab contains information regarding what is in stock/leftover inventory that can be resold at the next market. Columns A, B, and G require manual entry, and Columns C-F, H-L are automatically updated using various formulas.

- Column A contains the date of the market.
- Column B contains the product name.
- Column C is the unique ID that concatenates the values of the market date with the product name.
- Column D contains the total amount purchased of the product that Vlookup's the value from the "Aggregated Purchased Products" tab.
- Column E contains the total cost of the product that Vlookup's the value from the "Aggregated Purchased Products" tab.
- Column F contains the amount of inventory that was leftover from the previous market that is good for resale at the next market. This pulls the value from Column K based on the previous market date from the "Market Dates" tab.
- Column G contains how much of the product is available from the IY garden.
- Column H is the summation of columns D, F, and G to get the total amount of the product available for sale.
- Column I uses an index match function to pull information on how much was sold at the market from the "Products Sold" tab.
- Column J displays how much inventory there is leftover from the market by subtracting values from columns I to H.
- Column K uses a Vlookup and if statement to determine whether the value listed in column J can be resold at the next market.
- Column L calculates the difference between columns J and K in order to get the total lost inventory (i.e., products that could not be resold at the market and therefore are considered "wasted").

It is important to record this information after each market, as this will help determine how much of each product may need to be purchased for the next market. To do this, copy the list of products in column A in the "Products sold" tab, paste them into column B. Then input the market date in column A. For columns C to L, copy-paste a previous market's records and paste them. These columns contain formulas that will automatically update. Lastly, ensure that you adjust the "Garden" inventory accordingly.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Date of the Market	Product	Unique ID	Bought	Total Cost of Purchase	Inventory Prior For Resale	Garden	Total available for sale	Sold	Inventory After Market	Inventory After Market For Resale	Lost Inventory
2	2022-06-06	Apples	2022-06-06-Apples	200	$180.00	563	0	763	70	694	694	0
3	2022-05-13	Apples	2022-05-13-Apples	480	$360.00	363	0	843	280	563	563	0
4	2022-04-22	Apples	2022-04-22-Apples	157	$0.00	363	0	520	157	363	363	0
5	2022-03-25	Apples	2022-03-25-Apples	240	$147.00	346	0	586	222	363	363	0
6	2022-02-11	Apples	2022-02-11-Apples	240	$192.00	281	0	521	175	346	346	0
7	2022-01-14	Apples	2022-01-14-Apples	240	$168.00	271	0	511	230	281	281	0
8	2021-12-17	Apples	2021-12-17-Apples	240	$174.00	243	0	483	212	271	271	0
9	2021-11-26	Apples	2021-11-26-Apples	216	$234.00	201	0	417	174	243	243	0
10	2021-10-22	Apples	2021-10-22-Apples	316	$373.00	0	0	316	115	201	201	0
11	2021-09-24	Apples	2021-09-24-Apples	0	$0.00	0	0	0	0	0	0	0
12	2021-08-27	Apples	2021-08-27-Apples	19	$0.00	0	0	19	19	0	0	0
13	2021-08-13	Apples	2021-08-13-Apples	93	$0.00	0	0	93	93	0	0	0

Please refer to the "Main Farmer's Market" Google Sheet to see this tab

AA2 Cont'd

Back to ERD

AA2. Supplier Data Tab

This tab contains information that is manually entered from supplier quotes, along with the available budget for the market. Row 1 allows you to select a market date from a dropdown selection. Row 2 requires input on the budget constraint. Column A contains the list of products. If there is a new product, it needs to be added to the list. Column B contains information on how much is available for sale at the farmer's market. Column C displays the unit price of the product for sale. Column D contains the format in which the product is purchased.

	A	B	C	D
1	Market Date			
2	Total Budget	1500		
3	**Product Name**	**Supplier Inventory**	**Quote (in respective units)**	**Format**
4	Apples	79	1.03	lbs
5	Bags	11	3.75	roll
6	Bananas	42	0.49	lbs
7	Beans	87	2.02	kg
8	Beets	17	0.42	lbs
9	Broccoli	39	1.48	unit
10	Brussels Sprouts	30	2.36	lbs
11	Cabbage	95	0.44	lbs
12	Canteloupe	16	2.00	unit
13	Carrots	55	0.60	lbs

Please refer to the "Main Farmer's Market" Google Sheet to see this tab

AM

Modelling Appendix

Back to Data Sources

AM1. List of Data Sources Provided By IA

Name	Format	Content and Data Points	Utility
Invoices	PDF	60 invoices from suppliers (L'Autre Choix & Les Jardins Carya) on products they purchased from August 2021 until present. Includes the supplier name, invoice date, product description, quantity ordered, unit invoice price, and total cost.	Very relevant (quotes, quantity bought)
"Innovation Assistance Clients.xlsx"	Google Sheets	List of active members who receive assistance from Innovation Assistance. Includes names, addresses, phone numbers, food restrictions, delivery information etc.	Not relevant
"Pickup Tracking 2021.xlsx"	Google Sheets	File tracking pick-ups for their food basket pick-up service. Include number of portions expected to distribute in 1 week, number of portions picked up, and portions not picked up.	Not relevant for market
"Farmer's Market Tracking.xlsx"	Google Sheets	File tracking number of attendees, revenues, and expenses for their market. Also includes budget and allocated cost for their family pack distribution service.	Somewhat relevant (Budget information).
"Ordering – IY x JM.xlsx"	Google Sheets	"Wish-list" grocery list mentioned in the current process slide that contains the list of products (and format, quantity requested, and date needed) they requested from their supplier. Does not mean that all requests are fulfilled.	Somewhat relevant (may help fill in the gaps for what was demanded previously)
"Solidarity Market Inventory.xlsx"	Google Sheets	File recording items sold at the June 6th, June 20th, July 4th, and July 18th markets along with revenue (MMA requested IA to start collecting this data).	Very relevant (IA to continue recording)
"quebec_seasonal_product_calendar.pdf"	PDF	Table showcasing when each product is in season in Quebec.	Very relevant

Data Requests Made by MMA:

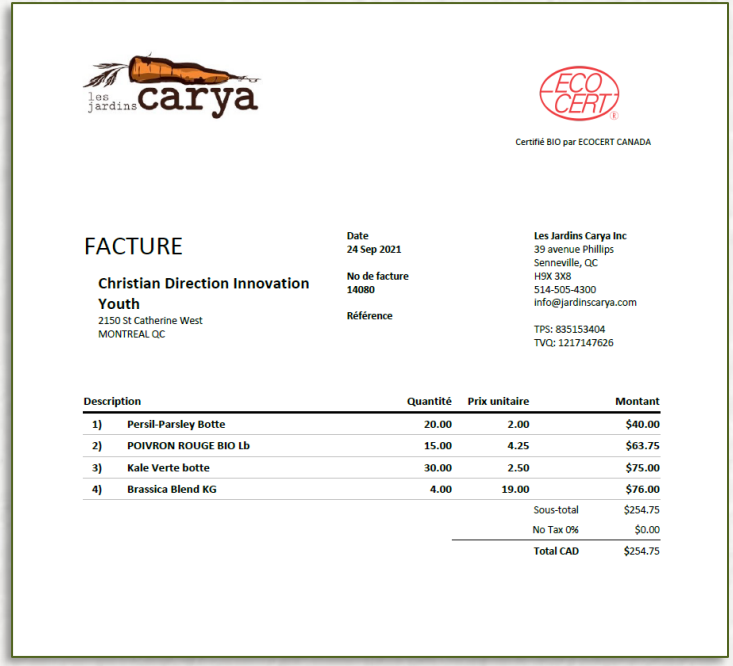
Invoices – Originally, we had invoices from December 2021 until present. To help us build the framework and generate synthetic data based on the invoices, we requested for all invoices they had on file on June 13th, 2022. Therefore, we had invoices from August 2021 until Present. However, some invoices from August were still missing.

Solidarity Market Inventory – Originally, this file did not exist until we started the project. IA was not collecting any data on how much was sold at their markets. However, since starting the project, we advised them that to predict demand, we needed to know how much was sold at the markets, and therefore requested they start collecting this data on June 6th, 2022.

AM1 Cont'd

Back to Data Sources
Back to List of Data Sources

AM1. Invoices



les jardins carya

Certifié BIO par ECOCERT CANADA

FACTURE

Christian Direction Innovation Youth
2150 St Catherine West
MONTREAL QC

Date
24 Sep 2021

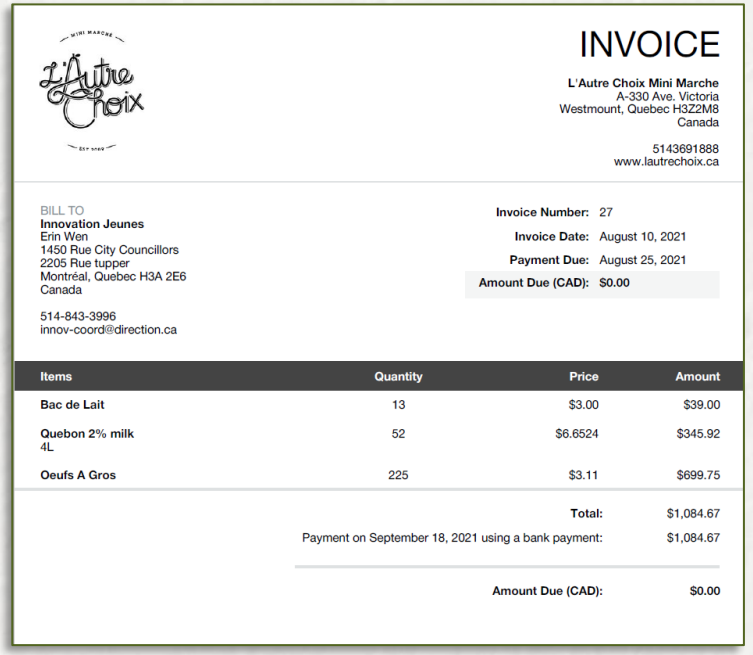
No de facture
14080

Référence

Les Jardins Carya Inc
39 avenue Phillips
Senneville, QC
H9X 3X8
514-505-4300
info@jardinscarya.com

TPS: 835153404
TVQ: 1217147626

Description		Quantité	Prix unitaire	Montant
1)	Persil-Parsley Botte	20.00	2.00	$40.00
2)	POIVRON ROUGE BIO Lb	15.00	4.25	$63.75
3)	Kale Verte botte	30.00	2.50	$75.00
4)	Brassica Blend KG	4.00	19.00	$76.00
			Sous-total	$254.75
			No Tax 0%	$0.00
			Total CAD	$254.75



L'Autre Choix

INVOICE

L'Autre Choix Mini Marche
A-330 Ave. Victoria
Westmount, Quebec H3Z2M8
Canada

5143691888
www.lautrechoix.ca

BILL TO
Innovation Jeunes
Erin Wen
1450 Rue City Councillors
2205 Rue tupper
Montréal, Quebec H3A 2E6
Canada

514-843-3996
innov-coord@direction.ca

Invoice Number: 27
Invoice Date: August 10, 2021
Payment Due: August 25, 2021
Amount Due (CAD): **$0.00**

Items	Quantity	Price	Amount
Bac de Lait	13	$3.00	$39.00
Quebon 2% milk 4L	52	$6.6524	$345.92
Oeufs A Gros	225	$3.11	$699.75
		Total:	$1,084.67
	Payment on September 18, 2021 using a bank payment:		$1,084.67
		Amount Due (CAD):	**$0.00**

Sample Invoices used within the “Invoice Info” tab.

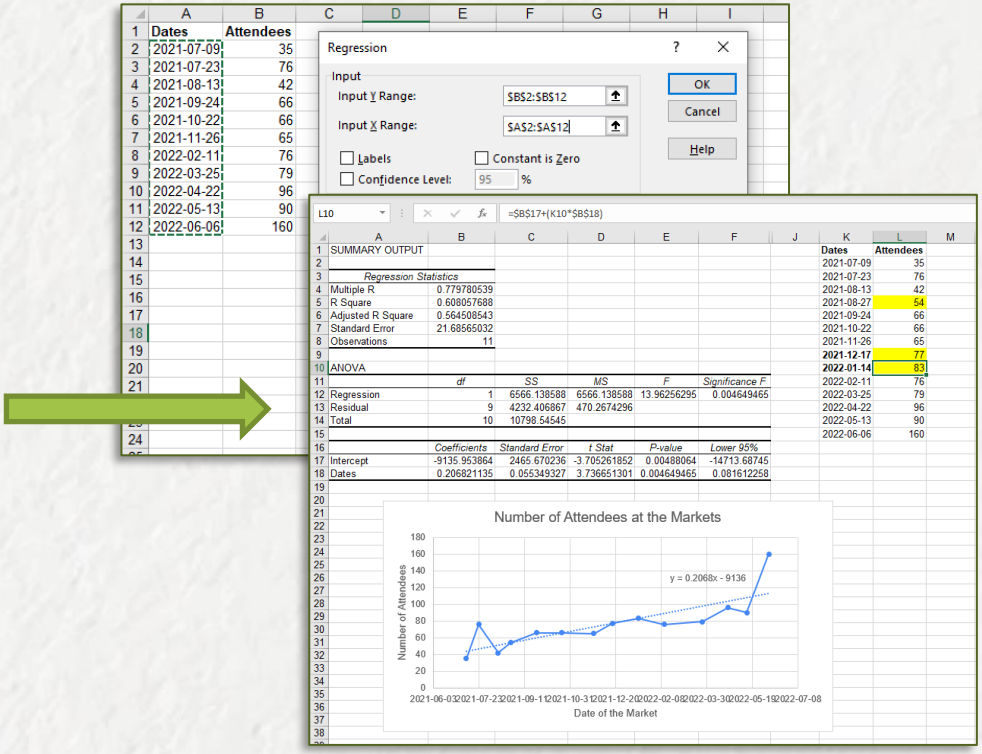
AM1 Cont’d

Back to Data Sources
Back to List of Data Sources

AM1. Farmer's Market Tracking

	A	B	C	D	E	F
1	**Dates**	**Attendees**	**Expenses**	**Revenues**	**Online giving?**	
2	July 9th	35	928.44	90.35		
3	July 23rd	76	869.5	310.55	110	
4	August 13th	42				
5	August 27th		878.8	274.6	25	
6	September 24t	66		356.3		
7	October 22nd	66	533.5	203.6		
8	November 26th	65		227.45	27	
9	Squarechip donations			154.42		
10	**2021 Total**			**1617.27**	**162**	
11	February 11th	76		277.2		
12	March 25	79		359.1		
13	April 22	96		308.95	115	
14	May 13	90		363.05		
15	**Total**			**2415.57**		
16						
17						
18						
19						
20	**Family pack spending**			**Alotted budget**		**Funding available (for April, May)**
21	October	1204.05			1200	
22	November	1199.04			1200	
23	December	1176.5			1200	
24	January				1200	
25	February	2233.16			1200	
26	March	1394.7			1200	
27	April				1200	
28	May				1200	
29	**Total**	**7207.45**		**Total**	**9600**	4808.12
30						
31						

Original Data Provided – missing attendees for August 27th, December 17th, 2021, and January 14th, 2022.



Applied a linear regression created in excel using the "data analysis" tool pack to fill in the attendees missing on August 27th, December 17th, 2021, and January 14th, 2022 (highlighted in yellow above).

Please refer to the "Farmer's Market Tracking" Google Sheet to see the file with our data imputation

AM1 Cont'd

Back to List of Data Sources
Back to Current Ordering Process

AM1. Ordering – IY x JM

	A	B	C	D
1	**Product**	**Format**	**Quantity**	**Date Needed**
2	Milk, 2%	4L	52	15-Jun-21
3	Eggs, medium	15 doz	15	15-Jun-21
4	Produce bags, plastic	10x15	6 rolls	15-Jun-21
5	Produce bags, plastic	10x20	8 rolls	15-Jun-21
6	Bags, paper	12x7x17	1200	15-Jun-21
7	Bags, paper	2lb	400	15-Jun-21
8	Bananas		120lbs	23-Jun-21
9	Oranges	72s or 88s	280lbs	23-Jun-21
10	Apples		200lbs	23-Jun-21
11	Garlic	heads	120lbs	23-Jun-21
12	Carrots	2 lbs	300lbs	23-Jun-21
13	Celery		60lbs	23-Jun-21
14	Bay Leaves		4lbs	23-Jun-21
15	Onions, yellow		400lbs	23-Jun-21
16	Potatoes, white or yellow		400lbs	23-Jun-21
17	Leeks*		20lbs	23-Jun-21
18	Tomatoes, large or cherry*	bulk	100lbs	23-Jun-21
19	Strawberries*	bulk preferred	50lbs	23-Jun-21
20	Zucchini, green or yellow*		80lbs	23-Jun-21

Sample of the "wish list" grocery list that is provided to the primary supplier. The supplier uses this excel sheet to see what they could provide.

Please refer to the "Ordering – IA x JM" Google Sheet to see the original file

AM1 Cont'd

Back to Products Sold
Back to Data Sources
Back to List of Data Sources

AM1. Solidarity Market Inventory

Client ID	Garlic	Green Onions	Onions	Coriander	Ginger	Potatoes	Sweet Potatoes	Carrots	Asparagus	Cucumber	Mushrooms	Spinach	Bok Choy	Watercress	Apples	Oranges	Strawberries	Value added items	Eggs	Milk	Bulk	Bill Total	Bill Total Paid
Grand total	135	81	215	38	64	193	159	139	68	86	28	97	47	30	192	247	53	70	75	57	30	1,002.10$	965.65$
Member (SoniS)	5	2	10	1	3	4	1	7	3	1	1	1		2	6	8	2		2	1		21.75$	20.20$
Non-Member	1																	1				2.60$	2.60$
Non-Member	2			1	1		2		1		1		4			4		2				15.45$	16.00$
Member (RahimN)		2	5	3			6	8	2	3	1	4		1	11	8	2	1	2	2	2	24.85$	24.85$
Member (NadlerM)		1																		1		0.85$	0.85$
Member (FilonenkoK)	2		10	1		6	3		1				1		8		1	3	2	2		16.45$	16.45$
Non-Member	2	1	3	2	3	3	3	2	1			2			2	2		1				22.00$	22.60$
Member (HananiaJ)	3	2		2			5			2	1				7	6	2	2	2			14.25$	14.30$
Member (ZhangQ)	6	4	10	2	3	12	10	12	3	3	1	5	10	2	20	8	2		2	2		42.35$	42.35$
Member (LimbasiyaP)	2	1	10	2	1	12		2		2						3	2		1	4		15.20$	15.20$
Member (PanchalS)		3	8	2	3	12		4		3		3		2	10	4	2		1	2		20.20$	20.20$
Member (SutariyaY)	5	2	8		1	10		8		2						4	2		2	2		16$	16$
Member (BhandariR)	1	1	8	2		3	1			3	1	3	1			8	2	2	2	1	1	17.85$	17.85$
Member (DivyangK)	3	4	6	2	3	12	5	6		3		2			10	4	2		1	2		24.80$	28.70$
Non-Member	2	1						2		2		1						1				8.00$	8.20$
Member (MarioD)		4	10	2		12	5		2	2	2	5	1	2	10	8	4	3	2	2		32.50$	10.30$
Non-Member	6	1					5	2	1		1	3		1	2	4	2	1				22.75$	22.75$
Non-Member	2	1		1		3				1						2						6.50$	6.50$
Non-Member		1			1		2			1	1		1			1						6.75$	10.00$
Non-Member			4	1								2										7.00$	7.80$
Member (ZhengJ)	6		6	2	3	10	10	6	3	3	1	3	2		5	8	2	1	2	1		27.30$	27.30$
Member (Sajida)	6	4	6	2	3	12	10	7	3	3	1	6	2		10	8	2	3	2	2	5	37.80$	0.00$
Non-Member			6	1		6		6		2	1	1			4	8						24.00$	24.85$
Member (TonyS)				1			1		1									1	1			3.20$	3.20$
Member (HaghjooN)			4	1		1					1				2	3	2	1				7.35$	7.35$
Member (AbbasY)			4	1		1					1				2	5	2	1	2			9.55$	9.55$
Member (SalarbashisharM)			4			1					1				2	4	2	1	2	1		9.80$	9.80$
Member (MemarianH)			3	1						1	1					6	1	1	2	2		10.00$	11.20$
Non-Member									2	1	1	1						2				12.00$	12.55$
Member (FamiF)		1	6		1	3				1	1					2	1	2	1	1		9.05$	9.05$
Member (TremblayG)																			2			0.85$	0.15$
Member (GhanimeA)	6	4	10			4		6	3	3	1			2	7	5	2	1				19.35$	15.00$
Non-Member		1		1					1	1	1		1				1	2				11.35$	11.35$
Member (?)		2		1		6	4			2		2				3			2	2		9.10$	9.55$
Non-Member																		2				4.00$	4.00$
Member (BerthiaumeC)	6	4		1	1		1	1	3	3	1	4	1	2	4	5	2	1				0.00$	0.00$
Non-Member					1		4	2	1	1	1		1	1		2	2	1	1			14.35$	14.35$
Member (TanoL)	4	4	20	2	3	12	8	6	3	3	1	6	1	2		8	2	1	2	1	4	31.80$	31.80$
Non-Member					1		2			1												5.50$	8.50$

Sample of the data collected on what was sold at the solidarity markets. Although we would only need the total sold for each product, we recommended that if it was not too much trouble for IA, to collect what was purchased by everyone. This is because as IA collects data over the year, we could expand the project to apply market basket analysis on what products are commonly purchased together, as well as to create a prediction model on what members of IA routinely purchase at the market, to potentially help IA make sure that frequently purchased products are available for sale at the market. The aggregations of how much was sold for each product is inputted within the "products sold" tab of our excel sheet.

Please refer to the "Solidarity Market Inventory" Google Sheet to see the original file

AM1 Cont'd

AM1. Quebec Seasonal Product Calendar

Back to Data Sources

Back to List of Data Sources

équiterre

Quebec Seasonal Produce Calendar

	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Apples	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Asparagus					✖	✖	✖					
Beans							✖	✖	✖	✖		
Beets	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Bell peppers							✖	✖	✖			
Blueberries								✖	✖	✖		
Broccoli							✖	✖	✖	✖	✖	
Cabbage	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Cantaloupe								✖	✖			
Carrots	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Cauliflower								✖	✖	✖	✖	
Céleriac	✖	✖	✖	✖	✖				✖	✖	✖	✖
Célery							✖	✖	✖	✖	✖	
Corn							✖	✖	✖	✖		
Cranberries									✖	✖	✖	✖
Cucumber			✖	✖	✖	✖	✖	✖	✖	✖	✖	
Eggplant							✖	✖	✖	✖	✖	
Endive	✖	✖	✖								✖	✖
Fennel						✖	✖	✖	✖			
Fresh herbs	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Garlic	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Grapes								✖	✖	✖		
Green onions							✖	✖	✖	✖		
Ground cherries								✖	✖			
Jerusalem artichoke	✖	✖									✖	✖
kale							✖	✖	✖	✖	✖	
Leeks	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Lettuce					✖	✖	✖	✖	✖	✖		
Mushrooms	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Onions	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖

* This list is not exhaustive. Availability varies by region and climate factors.

Please refer to the "Quebec Seasonal Produce Calendar" for the original copy

Back to Data Sources

AM2. Sample Synthetic Sales Data

Synthetic Dataset

Please refer to the "Synthetic Data" Google Sheet & tab to see the build

	A	B	C	F	G	H	I	J	K	L	M	N	O	P
1	Date of the Market	Number of attendees	% change in attendees	Product	uniqe id	In season?	Bought (converted to sold unit)	Inventory prior to market	Garden	Total available for sale	Sold	% sold	Inventory after market	Inventory After market for resale
2	2022-06-06	160	78%	Apples	2022-06-06-Apples	1	600	1162	0	1762	209	35%	1553	1553
3	2022-05-13	90	-6%	Apples	2022-05-13-Apples	1	1440	532	0	1972	810	41%	1162	1162
4	2022-04-22	96	22%	Apples	2022-04-22-Apples	1	0	1051	0	1051	519	49%	532	532
5	2022-03-25	79	4%	Apples	2022-03-25-Apples	1	720	1019	0	1739	688	40%	1051	1051
6	2022-02-11	76	-7%	Apples	2022-02-11-Apples	1	720	860	0	1580	561	36%	1019	1019
7	2022-01-14	82	8%	Apples	2022-01-14-Apples	1	720	784	0	1504	644	43%	860	860
8	2021-12-17	76	17%	Apples	2021-12-17-Apples	1	720	690	0	1410	626	44%	784	784
9	2021-11-26	65	-2%	Apples	2021-11-26-Apples	1	648	552	0	1200	510	43%	690	690
10	2021-10-22	66	0%	Apples	2021-10-22-Apples	1	948	0	0	948	396	42%	552	552
11	2021-09-24	66	25%	Apples	2021-09-24-Apples	1	0		0	0	0	#DIV/0!	0	0

Products Sold Tab

	A	B	C	D	E	F	G	H		J	K	L	M	N	O	P	Q
1			Market														
2	Total Revenue		$966.65	$476.19	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
3	Total Attendees		0	160	90	96	79	76	0	0	65	66	66	0	42	76	35
4	Product	To forecast?	2022-06-20	2022-06-06	2022-05-13	2022-04-22	2022-03-25	2022-02-11	2022-01-14	2021-12-17	2021-11-26	2021-10-22	2021-09-24	2021-08-27	2021-08-13	2021-07-23	2021-07-09
5	Apples	Yes	192	209	840	470	667	525	690	637	523	344	0	58	279	187	586
6	Bananas	No	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	Beans	Yes	0	0	0	0	8	7	0	0	188	112	0	53	183	19	94
8	Beets	Yes	0	0	117	0	68	91	0	321	0	139	0	96	261	100	185
9	Broccoli	Yes	0	0	77.22	0	40	27	0	0	0	0	46	54	25	34	3
10	Brussels Sprouts	No	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	Cabbage	Yes	0	0	0	0	0	16	0	59	0	46	0	15	32	56	20

For the values of products sold for the market dates highlighted in yellow, we used the RANDBETWEEN() function, with the lower and upper limits being the minimum and maximum value from the September 24th and June 20th markets.

Note: the "sold" value in the synthetic dataset will not match what is listed in the products sold tab. This is because the values change constantly in the synthetic dataset due to the use of the RANDBETWEEN() function. The values were copy-pasted as values into the products sold tab.

Back to Forecasting Models

AM3. Forecasting Framework and Assumptions

Assumptions

As previously learnt in Professor Cohen's Revenue management class, a typical demand function is the Quantity on the x-axis, with the price on the y-axis. Considering that prices being charged at the market are low and there being a pay-what-you-can system, we have assumed that price would not be a significant factor in determining demand within our forecasting framework. However, to identify other factors that may impact demand, we've identified weather, precipitation, and a seasonal produce guide as potential factors that could impact model performance in our supervised models. We can assume that when the weather is not nice/has precipitation in the forecast, this may potentially impact the total number of attendees to the market. The seasonal guide is also a good source in determining demand because all the produce that is purchased comes from Quebec farms and therefore would help signal in the model that demand for such a product may decrease if it is no longer in season. Of course, there could be potential to have multicollinearity between the seasonal guide, weather, and precipitation. However, once more real data has been gathered, a multicollinearity analysis should be done to not only identify which variables should be kept within the models, but to also see how effective these sources of data really are at predicting demand.

Framework

1. For loop used to go through each product we want to forecast and allow for the 8 models to always be updated for each forecasting cycle.

2. Each model's predicted results on the test dataset are stored in 1 data frame.

3.

0	1	2	3	4	5	6	7
Moving Average	Auto Arima	Seasonal Auto Arima	Simple Exponential Smoothing	Holt's Winter's Exponential Smoothing	Naïve	Random Forest	Multiple Linear Regression

From the results data frame, we calculate the MAPE for each model and store these results within a list.

4. The list index with the smallest MAPE is then used within an if-else statement to generate the prediction of the next period using the already trained model with the entire dataset.

5. These predicted values are converted into their appropriate purchased format and stored in a dictionary for input within the optimization model.

References used to build the forecasting models:
Moving average and exponential smoothing models: https://machinelearningmastery.com/time-series-forecasting-methods-in-python-cheat-sheet/
Arima models: https://www.machinelearningplus.com/time-series/Arima-model-time-series-forecasting-python/
Converting time series to supervised learning (for random forest and multiple linear regression): https://machinelearningmastery.com/convert-time-series-supervised-learning-problem-python/

Back to Mixed-Integer Problem

AM4. Problem Justifications

Client's Requirements

Our client had a few important requirements that had to be embedded in our MIP.

- **Budget**: IA has a strict budget to follow. They have a set budget of around $1,500 per market however, looking at their data, IA spends more than this amount on average. The optimizer will ensure that IA does not surpass their budget.
- **Leftovers**: IA needs to minimize their post-market leftovers. They do not have the capacity to store all their leftovers as they only have two single glass door fridges. They either throw away leftovers, transform them into prepared meal to be sold at the next market or store the rest inside or outside their fridges depending on the shelf-life of their products.

At first, the optimizer was minimizing budget and had leftover minimization as a constraint, however, it made more sense to have it the opposite way. In fact, having leftover minimization as a constraint implied that IA needed to set up allowable leftover amounts that the optimizer won't exceed which is an additional step for them to use the tool. Additionally, if the budget was too low to purchase enough inventory to have leftovers within the allowable leftover range, the optimizer broke. Therefore, having leftover minimization as the objective function and budget minimization within the constraints was the best option for the tool.

Problem Constraints

Other constraints were also important to include in our MIP to reflect the reality of IA's market operations.

- **Inventory limitations**: IA can only order as much as their suppliers are willing to sell which is why we had to add this constraint in our MIP.
- **Sales determination**: Sales can only be as much as the minimum between demand and total pre-market inventory. Either demand exceeds supply, in which case sales would be equal to supply or supply exceeds demand in which case sales would be equal to demand.
- **Demand fulfillment:** IA needs to order as much as their demand. And so, if the sum of their supplier's total inventory, and inventory from leftovers and garden is enough to cover for the demand then IY only needs to order what is needed to exactly match the demand. On the other hand, if it is not enough then IA needs to order all their supplier's total inventory.

To see the different iterations of problem formulation please see the Optimization Formulation Folder.

Back to Next Steps

AM5. Operationalizing the models using real data

As mentioned previously in the Products Sold tab, most of the data used throughout the POC is synthetic. We recommend that IA collect information on **1 year's worth of markets** (assuming a minimum of 24 data points if IA decides to maintain only a bi-weekly market), to have enough data that captures all seasons.

Once this data is collected, for the framework to be operationalized, IA must delete the data in the following tabs of the Main Farmer's Market Google sheet:

- **Products Sold tab**: delete all columns with dates in row 4 highlighted in orange and yellow

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1					Market													
2	Total Revenue		$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$476.19	$966.65	$975.48
3	Total Attendees															160	0	0
4	Product	To forecast?	2021-07-09	2021-07-23	2021-08-13	2021-08-27	2021-09-24	2021-10-22	2021-11-26	2021-12-17	2022-01-14	2022-02-11	2022-03-25	2022-04-22	2022-05-13	2022-06-06	2022-06-20	2022-07-04
5	Apples	Yes	586	187	279	58	0	344	523	637	690	525	667	470	840	209	192	267
6	Asparagus	No	0	0	0	0	0	0	0	0	0	0	0	0	0	0	68	57
7	Bananas	No	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	Beans	Yes	94	19	183	53	0	112	188	0	0	7	8	0	0	0	0	0
9	Beets	Yes	185	100	261	96	0	139	0	321	0	91	68	0	117	0	0	0
10	Broccoli	Yes	3	34	25	54	46	0	0	0	0	27	40	0	77.22	0	0	0
11	Brussels Sprouts	No	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	Cabbage	Yes	20	56	32	15	0	46	0	59	0	16	0	0	0	0	0	0
13	Canteloupe	No	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

- **Inventory tab**: delete all rows where the "Date of the Market" column has dates corresponding to the orange and yellow dates in the above screenshot.
  - For column K "Inventory After Market for Resale", double check that the values listed here are correct in terms of inventory leftover. As previously explained, this column uses an if and Vlookup statement to determine whether the amount left over listed in column J "Inventory After Market" can be resold based on the "Ideal leftovers" tab. Here, we had assumed that if the product can be resold based on the "Idea leftovers" tab, then it will take the full value listed in column J "Inventory After Market", otherwise, the inventory leftover would be 0. It may be that if the product can be resold, not all of the inventory values listed in column J will end up being resold (i.e., may decide to still donate some of the food or convert it to food anyways). If this is the case, you can manually override the formula in column K to adjust the final value available for resale at the next market.

Project Management – Meeting Schedule with Innovation Youth

Date	Duration of Meeting	Attendees	Topics of Discussion
June 2nd, 2022	1 hour (9-10 am)	Christa Smith (IY) Micah Angell (IA) Erin Wen (IA) Cristina Esposito (MMA) Alya Gabsi (MMA)	Kick-off meeting to deep-dive into the use case and understand pain points, expectations, and available data sources. For the notes from this meeting, please refer to the Google Doc "MMA-IA Meeting 2022-06-02".
June 6th, 2022	30 minutes (1-1:30pm)	Christa Smith (IY) Micah Angell (IA) Erin Wen (IA) Cristina Esposito (MMA)	Visited the IA farmer's market to observe how it is set up. Requested that IA count how much was sold at the market for us to use within the framework.
June 9th, 2022	30 minutes (11-11:30 am)	Micah Angell (IA) Erin Wen (IA) Cristina Esposito (MMA) Alya Gabsi (MMA)	Continued deep-diving the use case and asked questions about the data sources provided. For the list of comments and questions prepared, please refer to the Google Doc "MMA-IA Meeting 2022-06-09".
June 23rd, 2022	15 minutes (11-11:15 am)	Micah Angell (IA) Erin Wen (IA) Cristina Esposito (MMA) Alya Gabsi (MMA)	Quick touch-base meeting to get feedback on how they would like us to forecast (aggregated categorical level or individual product level), as well as suggesting changes to their current ordering process for the POC tool to work.
July 19th, 2022	30 minutes (11:30-12 pm)	Micah Angell (IA) Erin Wen (IA) Cristina Esposito (MMA) Alya Gabsi (MMA)	Preliminary walk-through of the tool we built and to gain feedback on what they like/don't like, things they would like to see added, and additional user testing. Gave access to the full tool to allow IA to play around with it and get adjusted.
July 25th, 2022	1 hour (3-4 pm)	Christa Smith (IY) Micah Angell (IA) Erin Wen (IA) Cristina Esposito (MMA) Alya Gabsi (MMA)	Final deep-dive walk-through and demo of the tool with the changes made to the tool from user feedback. Handed over all files for the project.

Please refer to the "Project Management" folder for all related project management documents

References

1. Administrator, Proof. "Household Food Insecurity in Canada." PROOF, 19 July 2022, proof.utoronto.ca/food-insecurity.
2. *"Average Canadian House Price Hits $816,720 — up 20% in Past Year." CBC, 15 Mar. 2022, www.cbc.ca/news/business/crea-housing-february-1.6385274.*
3. CanadaHelps. "Hunger in Canada –A Sobering Reality." CanadaHelps - Donate to Any Charity in Canada, 25 Aug. 2015, www.canadahelps.org/en/giving-life/connecting-with-charities/hunger-in-canada-a-sobering-reality.
4. *Evans, Peter. "The Cost of Living May Be Going up, but Wages Are Rising, Too." CBC, 18 Mar. 2022, www.cbc.ca/news/business/inflation-rising-wages-1.6384530.*
5. "Food Insecurity and the 'Temporary' Relief of Food Banks." Active History, 17 Aug. 2015, activehistory.ca/2015/08/food-insecurity-and-the-temporary-relief-of-food-banks.
6. *"HungerCount 2021." Food Banks Canada, 2021, hungercount.foodbankscanada.ca.*
7. Les Banques alimentaires du Québec. "About Us." Les Banques Alimentaires Du Québec, 14 Mar. 2022, banquesalimentaires.org/en/about-us.
8. Santé Publique. "SAINE ALIMENTATION." Santé Montréal, 24 July 2022, santemontreal.qc.ca/en/professionnels/drsp/sujets-de-a-a-z/saine-alimentation/information-generale.
9. Table de quartier Peter-McGill, Peter-McGill Community Council. "Food Security :" Peter-McGill Community Council, 3 Nov. 2021, petermcgill.org/en/food-security.
10. *"The Daily — Consumer Price Index, April 2022." Statistics Canada, 18 May 2022, www150.statcan.gc.ca/n1/daily-quotidien/220518/dq220518a-eng.htm.*