NORTHWEST NAZARENE UNIVERSITY

Taggart Farms Database System

THESIS
Submitted to the Department of Mathematics and Computer Science
in partial fulfillment of the requirements
For the degree of
BACHELOR OF ARTS

Amiee Petersen
2023

THESIS
Submitted to the Department of Mathematics and Computer Science
in partial fulfillment of the requirements
For the degree of
BACHELOR OF ARTS

Amiee Petersen
2023

Taggart Farms Database System

Author: Amiee Petersen
Amiee Petersen

Approved: Kevin S McCarty
Kevin McCarty, Ph.D., Associate Professor of Computer Science, Department of Mathematics and Computer Science, Faculty Advisor

Approved: Jennifer King
Jennifer King, Assistant Professor and Assistant Director of the Center for Academic Success and Advising, Second Reader

Approved: Barry Myers
Barry Myers, Ph.D., Chair, Department of Mathematics and Computer Science

ABSTRACT

Taggart Farms Database System Creation.

PETERSEN, AMIEE (Department of Mathematics and Computer Science), MCCARTY, DR KEVIN (Department of Mathematics and Computer Science).

The goal of this project was to update and streamline data capturing for a five-generation family farm named Taggart Farms. The current state of the farm's record keeping and reporting system was being done through unorganized excel files and paper hard copies kept in multiple locations. There is now a centralized database that holds all this information. Over the last few months, every stack of paper has been analyzed and turned into electronic records of the information. With the creation of the database the business will be able to easily maintain and update information and be able to draw conclusions about certain business dealings that can help inform decisions.

ACKNOWLEDGEMENTS

I would like to formally thank the members of Taggart Farms Inc for the funding, support, and access to proprietary information regarding the historical data that is vital to this project's success. I would also like to thank my advisor for assisting me with scheduling reminders, and questions that have come up during the process of this project.

TABLE OF CONTENTS

TABLE OF FIGURES

BACKGROUND

This project was intended to simplify and replace the record keeping process at Taggart Farms Inc. The former process was found to be outdated and unorganized, making it hard to utilize the data the company had. Through several meetings with stakeholders, we were able to determine the goals that the company wanted to achieve through the creation of a database system for the farm. Using the database, the farm wanted to be able to track contracts, create visualizations, and keep track of assets owned by the farm.

In order to understand the complexity of Taggart Farms it is important to explain a few basic things about the company. Taggart Farms has been a family-owned company for over five generations in southern Idaho. Primarily, Taggart Farms produces seed crops, as opposed to fresh vegetables that go straight to market. Since Taggart Farms produces seed, they have created relationships with large seed companies that contract the farm to grow a certain vegetable for them. The contracts signed for each crop are how the farm earns most of their income. The terms of the contracts get more complicated than that but for the purpose of this project it is not necessary to go into.

After meeting with the farm owners and learning more about the farm as a business, an understanding was decided on what information was necessary to have in the database and what kind of information they wanted to learn from their data. This first meeting was very important and provided guidelines and structure for the planning process of the database.

Following the meeting with Taggart Farms was the process of going through the old system and locating all the information that was necessary for the database. The old system currently is functioning; however, it was made clear that it would greatly benefit the business to centralize the farms data to one platform and use the data they have collected over the years to be able to inform and change the way that they make business decisions. All the historical data for the farm was located on paper, or in Microsoft Excel. Each excel spreadsheet had a different format from previous years, broken formulas, and missing and incomplete data.

INTRODUCTION

Farms have complicated business models and require a lot of hard labor and attention to detail. Failure to adequately perform these tasks can lead to having to sell the properties. An important task in the beginning stages of the project was to determine a period of time to cover. Considering that Taggart Farms is over 100 years old, there were going to be vast changes in how the business was operated then versus how the business is currently running. Meeting with the stakeholders cleared up any of these questions. It was decided that the past six years of data would be enough for the scope of this project, and for it to be applicable data for what they are currently growing.

After determining the years of data that were necessary, all access to farm data was given, information was cleared up so that the database only carried important contract information pertinent to the farm, as opposed to the seed company contracting the farm. After performing some basic organization, the process of pulling all important data started. To easily break down the scope of the project, the following sections are explinations of a data management process called ETL. ETL stands for extract, transform, load.

EXTRACT

To centralize all the information that was gathered from the hard copies and excel sheets, a master excel sheet was created. The new excel system was formatted specifically for the purpose of importing the new files into the database at the end of the collection process. This was not as simple a process as initially thought to be. Since the records were in multiple spreadsheets, it sometimes required information from multiple sheets to create a single updated table for the new system. In addition to information being in different places, there was also data that had been copied and pasted from a previous year into a following years file, meaning that we had to go back and either do the math to fix the problem or contact the farm owners to get the correct data. The process of data extraction over the last six years, along with the simultaneous data validation, the extraction phase took about two full weeks.

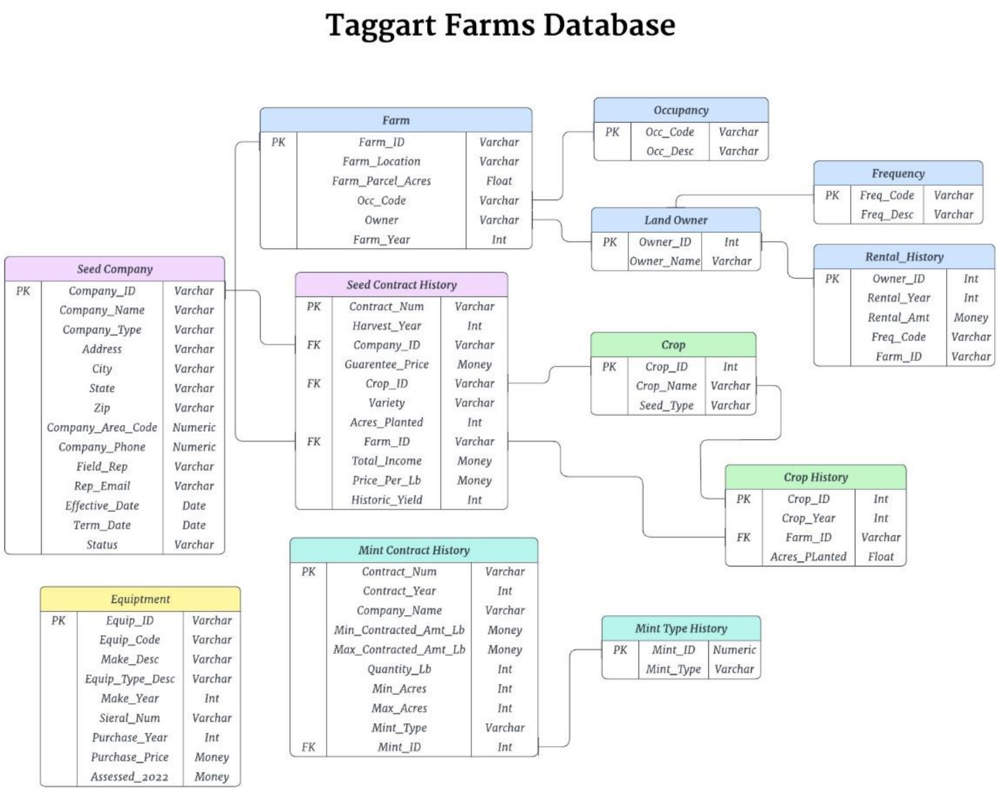
TRANSFORM

Taggart Farms has relationships with nine different companies, which means nine different formats of contracts, all written in different ways and even in different metrics. To help make the process of migrating the data that was in written form easier, a new meeting with Justin Taggart was scheduled. Since Justin would be the primary stakeholder in future farm business dealings and would also be using the database system the most. Contracts signed between the farm and companies we work for can be hard to read and sometimes contain information that is only beneficial for the companies we work for. To simplify the contracts table, the only information that was pulled from each contract was what was defined as being important to the stakeholders. This has proven to be an important decision to make since the farm needs different information from the contracts than the seed companies need from them. The information that was decided to be pulled was the following:

- Contract numbers
- Harvest year
- Company ID
- Guaranteed price
- Crop Name
- Variety of the crop
- Amount of acres planted
- The farm it was planted on
- The amount of income generated from specific contract
- The price per pound of the seed
- The historic yield of the variety provided by the company.

As tables were being created, each one was built into an entity relationship diagram, which is a tool used to show the relationships between each table in the database. Having this diagram helps make connections between each table and visualizes the database. Figure 1 is the completed entity relationship diagram. In the database, there are currently twelve different tables. There are plans to increase the size of the database, but for the scope of this project it was determined that this would be enough data to start with.

Figure 1: Entity relationship diagram



LOAD

In order to properly load the data into the database the first step would be to create the database in SQL Server Management Studio. To avoid as many bugs as possible during import, each table was hard coded into the database with only the column names entered. After that was completed, the import into the database took place. Each table's data was saved and imported using the import wizard, in the wizard the parameters for the table constraints were set. This included the variable types, sizes, and primary keys. Importing this way seemed to make the process easier for my system to process the import. It also became easy to see if there was an issue in the imports and if any changes needed to be made. In that case, all that needed to be done was to write a query that deleted the table that was just created, make changes, and created the updated table. After each upload there would be a check process to make sure each table was successfully loaded. A refresh would happen in the management server to see the creation of a table, additionally a query was done to select all the contents of the table to verify that it looked like the master table created in excel.

VISUALIZATION

Visualization was probably one of the most important aspects of this project to accomplish. Not only is it important for Taggart Farms to have the ability to see what is going on throughout the company. It is also important because it shows how helpful this new system can be to people who do not know agriculture very well. In addition, it also shows very well how databases can be used as tools we create to visualize a business and its operations. The ability to identify aspects of a business that do not normally get seen can make a huge difference in the overall success of any company.

When deciding on which of the different platforms for visualization we would want to use, it was important for the stakeholders to use systems that other companies are familiar with and standard practice. To find this information job sites were searched to see what platforms were most popular in the field of data science and analytics right now. The main two that were found in most descriptions were Power BI and Tableau. Since the database system was created using Microsoft’s SQL Server Management Studio, we figured it would be best to go with Power BI since it is also a Microsoft product. To gain experience with Power BI, video tutorials proved to be very helpful with the process of getting familiar with the software. The next three visualizations are some examples of what my database can tell Taggart Farms.

Figure 2: Income by Company

The income by company table was created by using the data gathered from contract information and showing us which company, we work with brings us the most income.

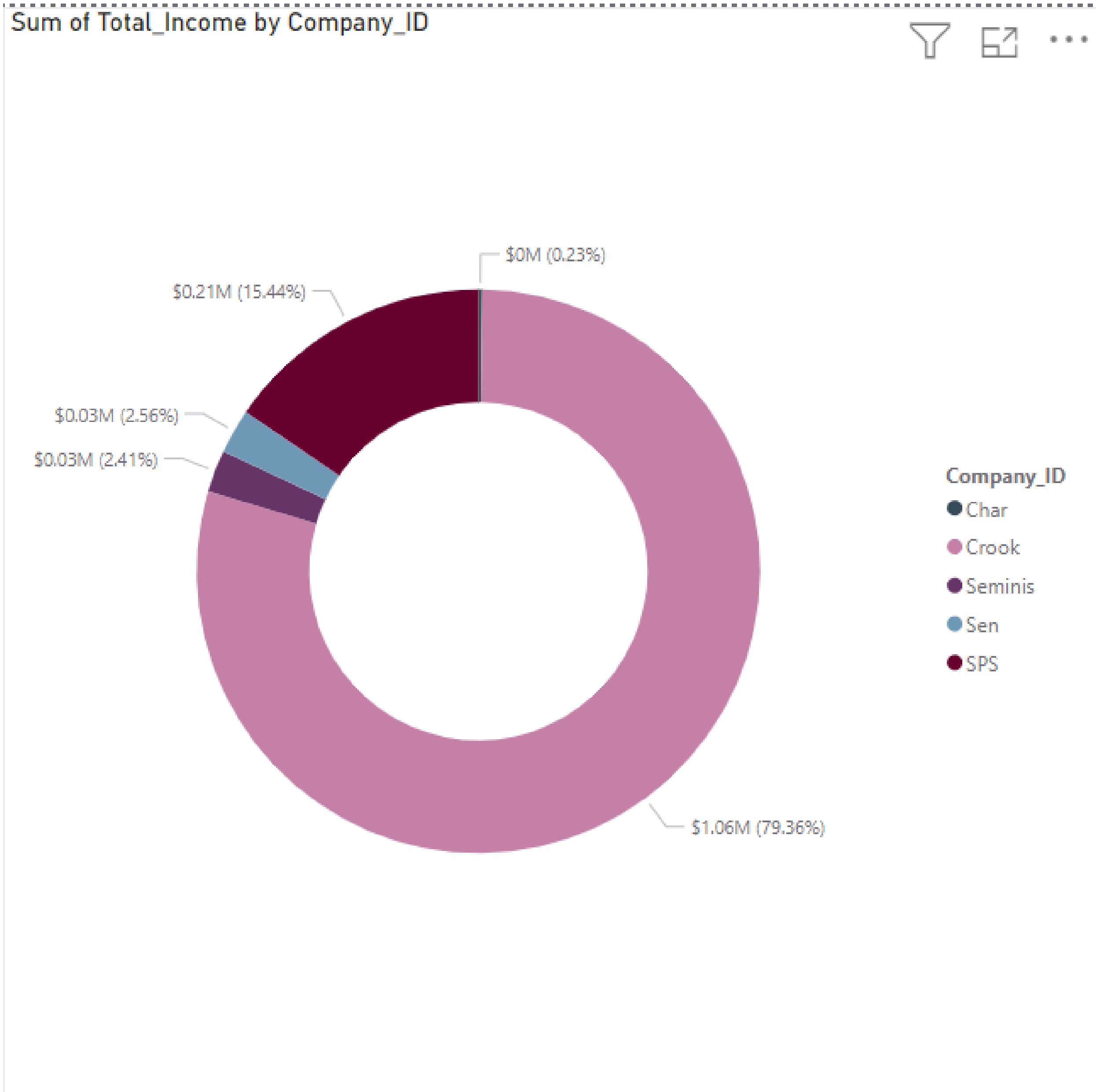


Figure 3: Income by Farm

Figure three shows which farm on the properties has historically brought in the most money. This visual was important for the stakeholders to see because some of the acreage they farm is rented. Meaning that on top of normal farming costs, there is the added costs of rental price. This data was used for the 2023 farm planning season where two farms are no longer rented.

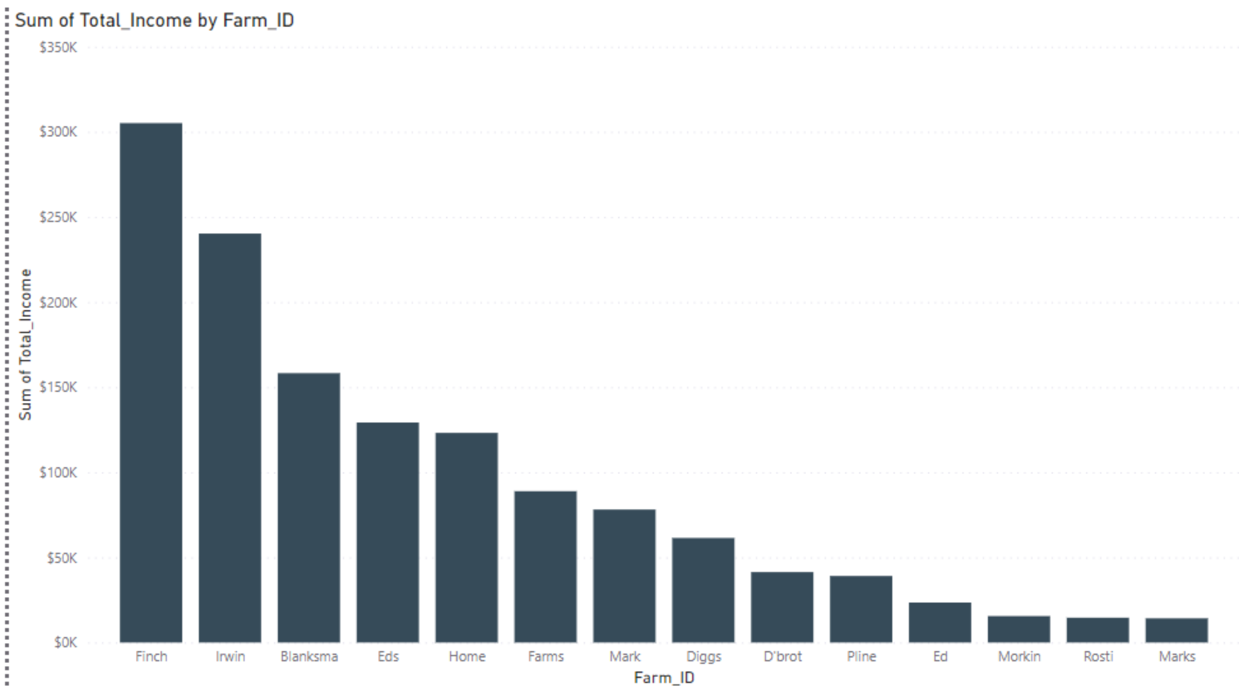
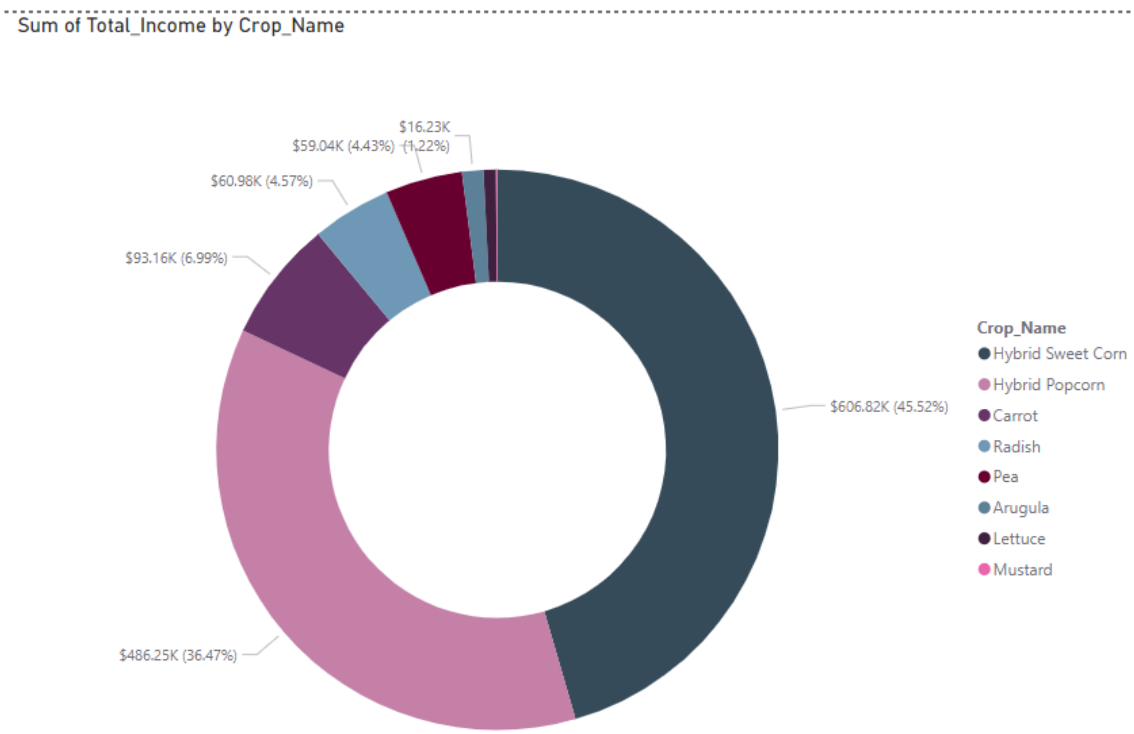


Figure 4: Income by Crop

This image shows the relationship between the amount of money made and which crop made that amount of money. The circle chart shows well which crops have made the most amount of money for the farm. A visual like this would be necessary in a situation where you are comparing what crops grow season to season.



TRIAL AND ERROR

One of the most challenging aspects of this project was the data validation. In addition to working with multiple media types, the main contributor to Taggart Farm's record keeping is an eighty-six-year-old man, with the introduction of human error some files can have missing or incomplete data making documents confusing to read. Starting with the paper copies of the contracted crops was to ensure that it would all match up with what the digital records had. When a discrepancy was located, the paper trail would need to be followed in order to fix it, and sometimes even do some math to double check that the information was correct. The other challenging portion of the project was keeping in mind that there was no access to full financial records, having this information in mind ended up skewing my visuals a little bit, because the farm does have other sources of income.

FUTURE WORK

There are a few key upgrades that are in the works for the Taggart Farms Database. The first thing that the company is going to want to do is create a user interface. Towards the end of the year, the farm starts receiving data from the companies that they grow crops for. A user interface would provide the members of Taggart Farms to edit tables and add data, providing us with a more real-time view of our data. Additionally, Taggart Farms is interested in connecting the database to their QuickBooks system, being able to quickly connect seed production with payment information will help them keep track of which companies owe them money and provide quick views on which crops make the most money.

Finally, there is an expansion planned for the database and a transition to move towards predictive analytics. Gathering data about market pricing, chemical costs, and other things like gas and diesel prices can help provide us with some good estimates to analyze the cost analysis and future estimated income.

The creation of the database has thus far proven to be worth the time investment. As the new farming season approaches, newly gathered data is being used to draw models for the seed companies that we grow for. Showing the field reps where their company has performed for Taggart Farms versus all the other companies that contract the farm, has given Taggart Farms leverage when it comes to negotiating new contracts and prices.

REFERENCES

Haan, K. (2023, April 22). *The Best Data Visualization Tools of 2023*. Forbes. Retrieved April 27, 2023, from https://www.forbes.com/advisor/business/software/best-data-visualization-tools/

YouTube. (2020). *Kevin*. *YouTube*. Retrieved April 27, 2023, from https://www.youtube.com/watch?v=TmhQCQr_DCA.