Analysing College Campus Dining Patterns

Dheeraj Reddy Kukkala
Department of Computer Science
Golisano College of Computing and Information Sciences
Rochester Institute of Technology
Rochester, NY 14586
dk8499@rit.edu

***Abstract*—Dining experiences play an important role in shaping students overall campus journeys in the huge expanse of student life. Using Data Science and Analytics, this project will uncover student dining patterns on campus, shedding light on food preferences, potential meal combinations, and crucially, peak dining hours. In the short term, this exploration seeks to enhance the overall student experience and to ensure efficient resource utilization.**

A dynamic model can intelligently recommend food items to promote based on a noticeable dip in CrossRoads sales during specific times. This approach also helps to minimize food waste, an issue that impacts both the environment and the economy.

An integral part of the project is the Microsoft Fabric platform, which is a friendly analytics ecosystem. The OneLake is a data reservoir and is the heart of the system. Data stored in this OneLake is accessible across all the platforms in the Microsoft Fabric suite. Using the SQL Endpoint, relationships between the tables are established. The data science platform sets the stage for rigorous analysis, often involving Peak time analysis, Food preference analysis, and Market basket analysis. These analyses culminate in a PowerBI report, visualizing insights that have the potential to change campus dining by balancing student satisfaction with institutional resource optimization.

Index Terms - Microsoft Fabric; Peak time analysis; Food preference analysis; Market basket analysis

I. Introduction

Campuses are constantly crowded with students from different cultures. One location, the dining hall, is overflowing with activity in addition to the classrooms and dormitories. It's more than simply a quick supper; it's a place where flavours converge, tales are spoken, and comfort food feels like a hug after a hard day. However, the management has a lot to think about in the background: What foods are popular? How much food needs to be maintained on hand? What time is rush hour? We're diving into these issues with Data Science & Analytics in the hopes of improving the eating experience for students and educating those in control.

A. Why Examine Dining Patterns?

The difficult game of efficiency in operation, which every institution aspires to master, is at the core of this research. Dining, which is so central to student life, offers a wealth of insights that are just waiting to be explored.

By diving deeply into these patterns [1], the university administration can analyze food habits. Inventory management is one of these insight's most practical implementations. When management is aware of the foods that are popular with students, they may make sure that the components are constantly available. This precise balance ensures that neither an excess nor a shortage could leave a customer unhappy.

This naturally brings up the urgent issue of food waste in complex settings like universities. However, with the data on dining habits at their fingertips, the institution is able to personalize servings and alter food production methods in response to demand in the present. This reduces waste while also promoting sustainable dining, which is crucial in the modern world.

However, the consequences extend beyond the kitchen's doors. The routines of student life can also be understood through these patterns. This information is helpful if, for example, there is a recurring spike in student attendance at the dining hall during particular times. The management may then make sure there are enough workers available during these busy periods, ensuring prompt service and content for students.

In conclusion, exploring the confusing world of eating habits is more than just a theoretical exercise. It's a method for institutions to operate more efficiently, to promote sustainability, and, most crucially, to keep the campus's students, who make up its very existence, happy and fed.

B. The Bigger Picture

A university serves as more than simply a place to learn; it also provides food for thousands of people every day. We don't just optimise the eating experience to reduce costs. It's all about being accountable, minimizing waste, and making sure that every student can find something delicious on the menu.

C. Actions From Observations:

Here's an example. Sales at CrossRoads occasionally decline. To understand this better, I analyze specific data points like customer footfall, purchase patterns, and menu preferences. The insights derived from these analyses are then visualized in PowerBI [2], providing management with clear, data-driven metrics [3]. This approach moves beyond mere speculation, allowing for informed decision-making that not only meets student needs effectively but also optimizes revenue management

In order to enhance the campus dining experience, my approach has been to implement a streamlined process for the collection and analysis of student dining data. The first step in this process involves automating the retrieval of data from secure servers to ensure a consistent and reliable flow of information. Subsequently, the gathered data undergoes a meticulous cleaning process to guarantee its quality before being transformed [4] into a more accessible format. This prepared data is then organized into a structured and orderly format, which facilitates the application of sophisticated analytical techniques aimed at revealing meaningful patterns and trends. The culmination of this process is the creation of visual reports that encapsulate these insights. These reports enable management to make informed decisions, aimed at improving the overall student dining experience on campus.

The study is focused on improving campus dining management by using data analysis techniques. I will be using Market Basket Analysis, Peak Time Analysis, and Food Preference Analysis to gather information and create actionable insights. The goal is to create PowerBI reports that will help management make informed decisions and improve the student dining experience and efficiency of campus dining services.

II. Background

A. The Bigger Picture:

A seamless dining experience is part of the broader student experience. If a college can optimize dining, it demonstrates attention to detail, responsiveness to student needs, and an overall commitment to enhancing campus life, which can contribute to overall student satisfaction.

B. The Smaller Picture:

The objective of this project is to create detailed PowerBI reports that utilize Market basket analysis, Peak time analysis, and Food preference analysis to provide actionable insights for campus dining management. These reports will enable management to make informed decisions based on visualized and interpreted data, ultimately improving the quality of dining services offered to students. By leveraging data analysis techniques, this project contributes to the advancement of dining management practices on college campuses.

C. Technological Infrastructure and Tools:

Central to this system is integrating Microsoft Fabric and its suite of tools, which work in tandem to streamline data processes from collection to visualization. Power Automate serves as the automation engine, transferring data from the SFTP server to a secure SharePoint location [5] to initiate data flow within the system. Data automation is crucial in maintaining up-to-date information, ensuring that decision-makers have access to the latest data without delay.

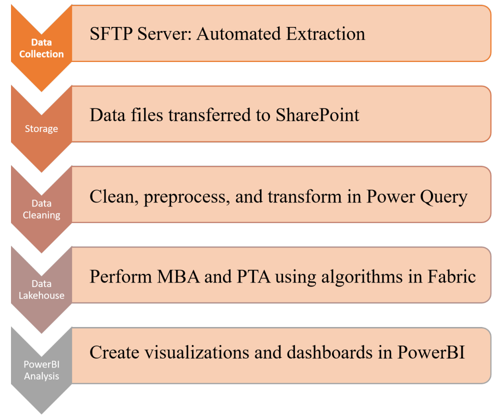

Fig. 1. Project Flow

D. Data Preparation and Cleansing:

Power Query [6] is implemented through Microsoft Fabric to extract data from SharePoint [5], clean essential data, and prepare it in a structured CSV format. This is critical to ensure that the data feeding into analyses is accurate, clean, and representative of real-world scenarios.

E. Algorithm Implementation and Data Analysis:

In the Microsoft Fabric data lakehouse, Market basket analysis, Peak time analysis, and Food preference analysis algorithms are utilized to process structured data and extract valuable insights into student behavior and dining trends. These insights are then incorporated into structured tables to facilitate data visualization and reporting in subsequent stages.

F. Data Visualization and Reporting:

PowerBI plays a role in using structured and cleansed data to produce intuitive and informative visuals [2] for data visualization and reporting. The reporting tools provided by PowerBI are leveraged to deliver a clear and engaging representation of the findings, making it easier for management to digest complex information and make informed decisions.

G. End Goal:

The use of advanced data analysis techniques such as Market basket analysis [7], Peak time analysis, and Food preference analysis in the Microsoft Fabric data lakehouse [8] project has the ultimate goal of equipping dining services management with comprehensive PowerBI reports. These reports provide insights into various aspects of campus dining, allowing management to make informed decisions, optimize dining services for efficiency and student satisfaction, and improve the overall campus experience.

III. Methodology

A. Data Acquisition and Preparation

Historical Data Retrieval and Processing: Our project starts by obtaining a large dataset spanning the last three years, consisting of 60 compressed XML files stored on an SFTP server. To manage this data effectively, we use Microsoft

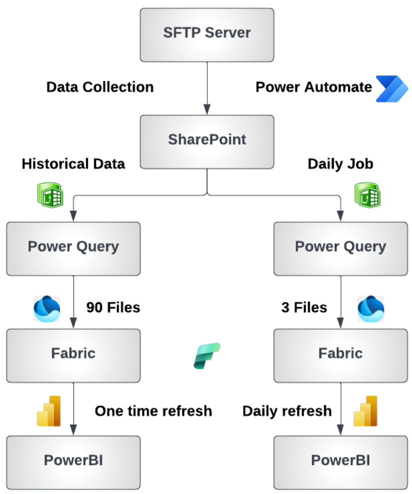

Fig. 2. Flow Diagram for Data Extraction

Power Automate to automatically extract the files and move them to a specific location in SharePoint. After the successful transfer, as shown in Fig.2. Power Query is utilized for two purposes. Firstly, it works as a data cleansing tool to ensure that the historical data is free of inconsistencies and redundancies. Secondly, it creates a schema that caters to the complex requirements of 31 distinct tables. Then, it generates corresponding CSV files for each. Once these CSV files are ready, they are seamlessly incorporated into the tables within Microsoft Fabric's Data Lakehouse [9]. This integration marks the conversion of raw data into a structured format, which can be further analyzed.

Automation of Daily Data Ingestion: Every day, three new files are added to the SFTP server. Power Automate then extracts these files and automatically transfers them to SharePoint. Subsequently, these files undergo pre-processing and cleaning before being sent to the Microsoft Fabric Data Lakehouse tables. The Market Basket Analysis (MBA) and Peak Time Analysis (PTA) algorithms are run daily on this exported data, which in turn refreshes the visuals in the PowerBI report. The current setup within Microsoft Fabric allows for easy automation with a few clicks. However, the full automation of this system is in progress. Once this is finalized, it will enable a system that is user-friendly and allows even less experienced users to generate new reports easily

B. Data Relationship Establishment - E-R Diagram

An Entity-Relationship (E-R) diagram has been created to visually depict the complex framework of our campus stores' transactional data. This diagram captures the architecture of our database, showcasing the interplay between 31 interconnected tables. Out of these, four are fact tables, which serve as the centerpiece for our transactional data analysis. The fact tables represent the essence of business processes and are supported by various dimension tables, which provide additional context to the data to give it meaning. These dimension tables contain descriptive attributes related to our fact tables, allowing for multidimensional analysis. This design enables us to not only understand individual transactions but also to explore trends and patterns over time, across different dimensions such as time, item categories, and store locations. The E-R diagram is a comprehensive representation of the transactional data collected from all campus stores, including sales, inventory, staffing, and customer interactions, all of which are interconnected to provide a holistic view of the campus dining ecosystem.

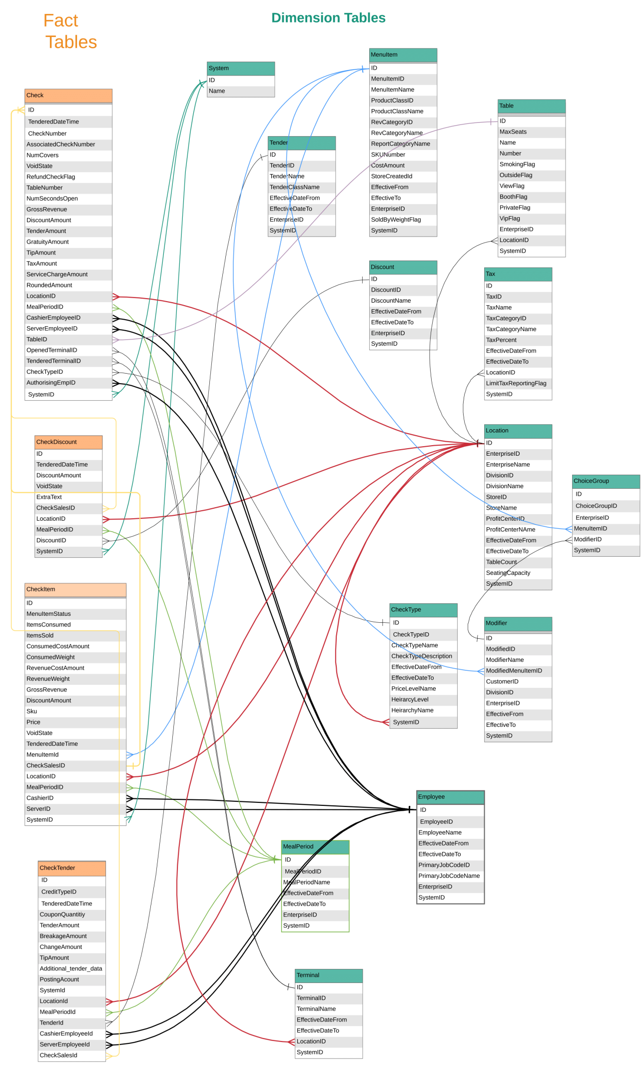

Fig. 3. Entity Relationship Diagram

The visual illustration presented in Fig.3. acts as a guide for navigating the intricate data connections that exist within our database. It is a crucial aid for individuals with or without technical expertise to understand the organization of

```
# Generating Frequent Itemsets with Apriori Algorithm
frequent_itemsets = apriori(basket_pd, min_support=0.01,
                            use_colnames=True)

# Deriving Association Rules with a Minimum Lift of 1
rules = association_rules(frequent_itemsets,
metric="lift", min_threshold=1)
```

Fig. 4. Market Basket Analysis - Apriori Algorithm

the data, which leads to better comprehension and informed decision-making. By using this diagrammatical representation, we establish a foundation for the following phases of data analysis and extraction of insights.

C. Market Basket Analysis

Market Basket Analysis is a technique used in data mining to discover links between sets of items that are frequently purchased together. This approach can be useful in campus dining by identifying groups of menu items that tend to be ordered together, thereby assisting with menu planning and promotional efforts. The first step in the process is to gather transactional data, which in this scenario refers to the individual items on each customer's receipt. To get the data ready for the MBA algorithm, it must be processed and transformed, and the code snippet plays a role in each of these stages.

1) **Data Preparation:** The transaction data is loaded, null values are removed, and the data is grouped by check number. Each menu item is pivoted into a new column, and the count of each item per check is binary-encoded (1 for present, 0 for absent), preparing it for the Apriori algorithm.
2) **Apriori Algorithm:** The apriori function generates frequent itemsets. These are groups of items that appear together in transactions with a frequency above the specified min-support threshold.
3) **Key Terms of MBA:**
   Antecedents are items that, when found in a transaction, will influence the probability of the presence of other items, known as consequents.
   Consequents are items that appear frequently in combination with antecedents.
   Support is the proportion of transactions that include the itemset (both antecedent and consequent).
   Confidence is a measure of the likelihood that an itemset is purchased once the antecedent items are purchased. It's the ratio of the support of the itemset to the support of the antecedents.
   Lift indicates the strength of a rule over the random co-occurrence of the antecedent and consequent, calculated by dividing the rule's confidence by the support of the consequents.
4) **Result Interpretation:** The association rules function extracts the rules from the frequent itemsets, considering a minimum lift of 1, which suggests at least a neutral (if not positive) relationship between antecedent and consequent. The rules are refined to display the actual menu items, joining the results with a menu item table to replace item IDs with names for better readability.
5) **Actionable Insights:** Finally, the rules are stored in a database for further analysis, potentially influencing menu design, cross-selling strategies, and customer experience enhancements.

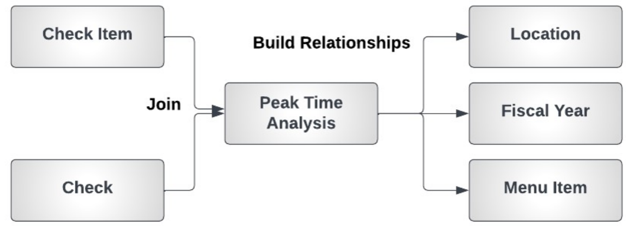

Fig. 5. Peak Time Analysis

```sql
SELECT C.TenderedDateTime AS HourOfDay,
SUM(NULLIF(TRY_CAST(C.GrossRevenue AS DECIMAL(15, 5)), 0)) AS TotalRevenue,
SUM(NULLIF(TRY_CAST(C.DiscountAmount AS DECIMAL(15, 5)), 0)) AS TotalDiscount,
SUM(NULLIF(TRY_CAST(CI.ItemsSold AS INT), 0)) AS TotalItemsSold,
COUNT(C.Covers) AS TotalCovers,
C.LocationID, C.MealPeriodID, COUNT(*) AS TotalTransactions
```

Fig. 6. SQL Query attributes

D. Peak Time Analysis and Food Preference Analysis

Understanding peak activity times helps optimize resource allocation, while food preference analysis identifies top-selling and least-selling items. Below are the steps

1) We started with two primary data sources: the "Check" and "Check Item" tables from our DataModelPOS database.
2) Designed SQL query for analysis, including attributes like revenue, discount, items sold, covers, location, meal period, and transactions as shown in Fig.6. Ensured data accuracy with type casting (revenue, discount, and items sold). Standardized "HourOfDay" column for consistent date and time representation
3) Linked "Check" and "Check Item" tables based on "CheckNumber" and "CheckSalesID". Created "PTA" (Peak Time Analysis) table for data insertion. Enriched analysis with location names, Fiscal Years and Menu Items details by joining "PTA" with "Location", "Fiscal Year"and "Menu Item" tables as shown in Fig.5.
4) Also determined the most and least-selling items to understand customer preferences. Identified the lead-selling item contributing to high revenue.

The combined analysis of peak times and food preferences provides actionable insights to enhance the restaurant's performance. It guides resource allocation, menu optimization, and marketing strategies to boost customer satisfaction and revenue.

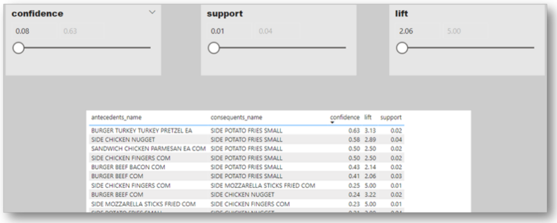

Fig. 7. Food Combinations PowerBI visual

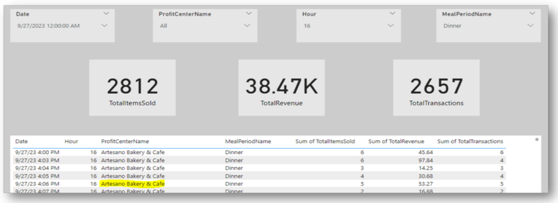

Fig. 8. PowerBI visuals for Total Revenue, Most selling items and Total Transactions

IV. Evaluation

A. Evidence Gathered

A finalized PowerBI report with comprehensive visuals [2] detailing student dining patterns, preferences, and peak hours.

B. Market Basket Analysis

In the Market Basket Analysis, visual representations were utilized to identify popular item combinations, thereby offering insights for refining the menu and developing targeted promotional strategies. As illustrated in Fig.7, data from the previous week was analyzed for a specific location i.e.Cross Roads. This analysis revealed a notable combination: the Turkey Burger with a small side of potato fries exhibited a high confidence and lift value. Specifically, a 63% likelihood exists that a customer purchasing a Turkey Burger will also buy potato fries. Based on these findings, it is recommended that management maintain an increased stock of potato fries and consider promotional discounts for this combination, optimizing inventory and potentially boosting sales.

C. Peak Time Analysis and Food Preference Analysis

The PowerBI report generated provides insights into peak dining hours, top-selling items, and least popular items at specific locations and during different time periods. Integrating the Peak Time Analysis table with Fiscal Year data allows for comparative analysis across different time frames. As demonstrated in Fig.8, this integration facilitates calculations such as Total Revenue or identification of best-selling items for a specific location during a Fiscal Year. It also enables comparisons between the current and previous fiscal years, or between similar days (e.g., Mondays) across semesters. These analytical capabilities assist management in tracking revenue performance for specific items, enabling more informed decision-making regarding menu and operational adjustments.

V. Conclusion

Our project's outcome is an interactive PowerBI dashboard that serves as the conclusive output, encapsulating the core of our data-powered endeavor to enhance the campus dining experience. This dashboard precisely presents our analytical pursuits, featuring detailed trends in student dining behavior through Market Basket Analysis, Peak Time Analysis, and Food Preference Analysis. Every visual element provides practical suggestions that uncover not only the frequency and timing of students' dining habits but also the interdependencies between their food preferences.

This analysis gives us a comprehensive understanding of the campus dining environment, allowing the management to make informed and strategic decisions [10]. As we continue, the results of this project will guide future improvements, ensuring that the dining experience for students remains engaging and adaptable to the evolving needs of campus life.

VI. Future Work

The future work focuses on streamlining data management, enabling efficient handling from raw data processing to in-depth SQL analysis. It envisages expanding data integration by incorporating additional sources like Kronos for workforce insights, on-campus housing data for a broader view of student life, and MyCourses data to correlate academic schedules with dining habits. Additionally, the plan includes leveraging Microsoft Fabric's capabilities for comprehensive campus-wide data integration and promoting cross-departmental data sharing and analysis, thereby broadening the utility of the established system

References

[1] S. Agarwal, "Data mining: Data mining concepts and techniques," in *2013 International Conference on Machine Intelligence and Research Advancement*, 2013, pp. 203–207.

[2] G. B. Mandava, "Analysis and design of visualization of educational institution database using power bi tool," 10 2018.

[3] D. Power, "Understanding data-driven decision support systems," *IS Management*, vol. 25, pp. 149–154, 03 2008.

[4] A. Raj, J. Bosch, H. H. Olsson, and T. J. Wang, "Modelling data pipelines," in *2020 46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, 2020, pp. 13–20.

[5] P. Ted and L. Daniel, *Inside microsoft windows sharepoint services 3.0*. Microsoft Press, 2007.

[6] E. Kajáti, M. Miškuf, and P. Papcun, "Advanced analysis of manufacturing data in excel and its add-ins," in *2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2017, pp. 000 491–000 496.

[7] S. M. Laksmana, M. and R. Maulana, "Product recommendations using market basket analysis with fp-growth and clustering techniques," in *In 1st Australian International Conference on Industrial Engineering and Operations Management*, 2022.

[8] N. G. Kuftinova, O. I. Maksimychev, A. V. Ostroukh, A. V. Volosova, and E. N. Matukhina, "Data fabric as an effective method of data management in traffic and road systems," in *2022 Systems of Signals Generating and Processing in the Field of on Board Communications*, 2022, pp. 1–4.

[9] R. Hai, C. Koutras, C. Quix, and M. Jarke, "Data lakes: A survey of functions and systems," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 12, pp. 12 571–12 590, 2023.

[10] Z. Ren, “Delivering a comprehensive bi solution with microsoft business intelligence stack,” in *2010 International Conference on Challenges in Environmental Science and Computer Engineering*, vol. 2, 2010, pp. 278–281.