Point of Sale (POS)-Based Labor Planning: A Data Warehousing Approach to Enhance Retail Decision-Making

Akshay Sodha, Esha Aggarwal, Tanvi Pagrut, Priya Govindarajulu

Abstract—**Retail store cashier desks are not just for transactions; they are useful to customers as well. Quick and smooth checkouts aren't just key to customers' satisfaction, they also help with planning labor. Our project uses Point of Sale (POS) data from retail stores. We have a goal to identify when stores are busiest and which cashiers are top performers during peak hours. Using SQL operations and Python to manipulate data, we want to give stores insights on the busiest shopping times and how to plan work shifts, and this is done by integrating this data into a data warehouse.**

The main expected outcomes include:

- **Identifying Peak Store Hours: Assessing the busiest times for customers.**
- **Transaction Analysis: Enhance checkout efficiency through insights into cashier operations.**
- **Labor Allocation Strategy: Suggestions for effective staffing distribution during peak and off-peak hours, taking into account the transaction-processing efficiency of cashiers.**

I. Introduction

Using Point of Sale (POS) data, this research enhances decision-making in the retail optimization space. By analyzing data with SQL and Python, we aim to identify high footfall shopping times and give recommendations for great cashier performance, which will help with labor planning. The cashier desk is an essential component of customer service, our goal is to promote operational efficiency and improve customer satisfaction.

II. Motivation

Customer satisfaction in the retail industry is largely determined by the experience at the registers. Having worked in the retail sector, one of our team members often observed situations when there were either too few workers at the registers during periods of high customer traffic or too many during slower periods. Motivated by these direct observations, the project's goal is to solve labor allocation problems. Our goal is to create a data-driven strategy that will assist retail stores in optimizing their staffing schedule, guaranteeing a satisfying shopping experience for customers.

III. Literature Review

There have been several changes in the retail sector with the introduction of data-driven decision-making, majorly around Point of Sale (POS) systems. One of the significant studies in this field is by Chuang, Oliva, and Perdikaki, where they dived deep into "Traffic-Based Labor Planning in Retail Stores" [Chuang, H. H.-C. et al., 2016]. Since labor costs are such a huge part of retail expenses, their main focus was to find out how store performance is affected by staffing levels.

In addition to examining the required number of employees, they also took into account the impact that varying employment levels had on sales. This was especially notable because they developed a formula that illustrated the relationship between traffic, or customers entering the store, labor, or workers, and sales using actual data from a retail chain. They discovered that even in situations when there was a significant fluctuation in the number of consumers entering the store, their approach to sales prediction was nearly perfect.

Expanding on this, Smirnov and Huchzermeier's paper, which was published in the 2020 issue of the European Journal of Operational Research (Vol. 287, Issue 2, pages 668–681), made another interesting contribution. They presented an integrative approach designed specifically for labor planning in systems where client load or influx directly affects service times. They developed a model to forecast consumer arrivals by combining machine learning approaches with time-series forecasting. Their methodology may result in a significant 4.4 percent reduction in staffing expenses, meaning that the restaurant's profitability before interest and taxes could potentially increase by 13 percent. This was demonstrated in a real-world application within a big worldwide fast-food chain. The improvement of checkout procedures at stores is another topic that is receiving attention.

Customer satisfaction increases the likelihood that a customer will return, therefore an easy checkout process is essential. This is the main focus of our research, which involves using POS data to identify the best staffing schedules and improve the checkout procedure. We believe that our findings will greatly help retail stores with an innovative approach.

IV. Methodology

1. Data Extraction from Couchbase NoSQL Database (Operations Data): We established connections to the Couchbase NoSQL database, extracted the relevant Operations data using appropriate queries and transformed the extracted data into a usable format.

2. Data Extraction from Cloud-Stored Excel Files (Transactions Data): We accessed cloud-based storage containing key Excel files, extracted transaction data from these files and ensured compatibility with subsequent processing steps.

3. Extracting Cloud Files using ETL (Using Apache NiFi): Utilized Apache NiFi for efficient Extract, Transform, Load

(ETL) operations, set up data flows to extract data from cloud-stored Excel files and transformed the data during the extraction process.

4. Extracting Operations Data into Python Tool (Jupyter Notebook): We developed a Jupyter Notebook for extracting operations data, used appropriate libraries to connect to Couchbase and retrieved relevant information. This ensured seamless integration with the overall workflow.

5. Data Cleaning and Pre-processing (Using Pandas): We used Pandas for full data cleaning like handling missing values, correcting errors and ensuring data integrity and preprocess the data to make it suitable for further analysis.

6. Data Modeling - ER Diagram for Star Schema: We designed an Entity-Relationship(ER) diagram for the star schema and identified key entities, attributes, and relationships in the data model.

7. Data Warehouse Design in MySQL: We implemented the designed star schema in a MySQL database, created tables based on the ER diagram, made sure all the tables are normalized. We have made sure to set up appropriate constraints for relational database management.

8. ETL to Load Data into MySQL Data Warehouse (Using Python Pandas): We have developed ETL pipelines in Python using Pandas, extracted data from various sources, transformed it based on the star schema, and loaded it into the MySQL Data Warehouse.

9. Data Analysis and Visualization (Using Matplotlib): We have analyze the data within the MySQL Data Warehouse and utilized Matplotlib for creating insightful visualizations and analysed them to identify patterns, trends, and actionable recommendations.

10. Labor Planning Recommendations: We have derived labor planning insights based on the data analysis, proposed staffing strategies for peak and non-peak hours and provided 5 labor recommendations for optimizing workforce deployment.

11. Report Compilation: We have compiled all of our findings, methodology details, and visualizations into a comprehensive report. We have tried to clearly present labor planning recommendations and the rationale behind them and ensured the report is accessible and understandable for stakeholders.

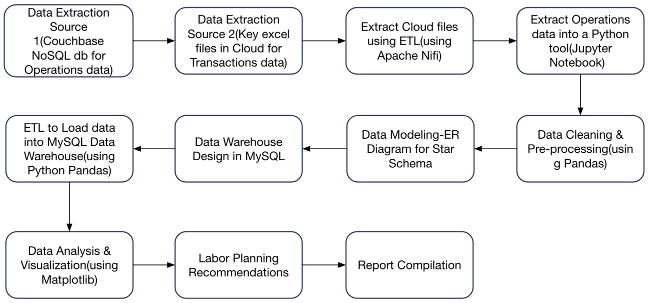

Data Extraction Source 1(Couchbase NoSQL db for Operations data)
Data Extraction Source 2(Key excel files in Cloud for Transactions data)
Extract Cloud files using ETL(using Apache Nifi)
Extract Operations data into a Python tool(Jupyter Notebook)
Data Cleaning & Pre-processing(using Pandas)
Data Modeling-ER Diagram for Star Schema
Data Warehouse Design in MySQL
ETL to Load data into MySQL Data Warehouse(using Python Pandas)
Data Analysis & Visualization(using Matplotlib)
Labor Planning Recommendations
Report Compilation

Fig. 1. Visual representation of the methodology steps.

V. Dataset Description

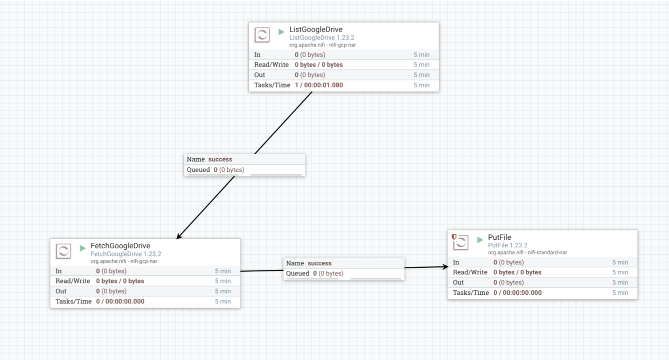
The datasets hold important data on two main topics: transactions and cashier operations. All the key features have been identified and the data covers three periods of time, each roughly two weeks long: December 2017, February 2019, and March to April 2019. A major point to consider is the new Polish regulation of 2018 that prohibited shopping on certain Sundays. This change affected the typical 7-day shopping pattern until the end of 2017, leading to supermarkets adjusting their timings, especially on Fridays and Saturdays. The 2019 datasets cover the effects of this regulation by including data from both working and non-working Sundays.

VI. System Design for Data Sourcing

The team decided to create a Polyglot Persistence environment to implement and show how the data sources can be from different systems or databases for the data warehouse. The dataset has been downloaded from https://www.mdpi.com/2306-5729/4/2/67. There are 6 CSV files in the dataset, of which 3 are for the Transaction data and 3 are for the Cashier Operations data.

The Couchbase database has been created and loaded with operations data. The data attribute is converted to a string and then the entire CSV file is converted into JSON format. Then the document for Operations data is loaded into the Couchbase database. Now this database is acting as one source database for the Datawarehouse.

The transaction data files were moved to the cloud, to create a second data source for the Datawarehouse. We have used Google Drive as a cloud resource. The process flow was created using the Apache NIFI ETL tool to extract the transaction files to the staging server, the local system. List Google Drive process lists all files in the shared folder, fetch Google Drive process retrieves the files, and put file process adds the retrieved file to the local system. This flow is automated and if new files are added, the data will be moved to the staging server.


```
ListGoogleDrive
ListGoogleDrive 1.23.2
org.apache.nifi - nifi-gcp-nar
In          0 (0 bytes)           5 min
Read/Write  0 bytes / 0 bytes     5 min
Out         0 (0 bytes)           5 min
Tasks/Time  1 / 00:00:01.080      5 min

Name    success
Queued  0 (0 bytes)

FetchGoogleDrive
FetchGoogleDrive 1.23.2
org.apache.nifi - nifi-gcp-nar
In          0 (0 bytes)           5 min
Read/Write  0 bytes / 0 bytes     5 min
Out         0 (0 bytes)           5 min
Tasks/Time  0 / 00:00:00.000      5 min

Name    success
Queued  0 (0 bytes)

PutFile
PutFile 1.23.2
org.apache.nifi - nifi-standard-nar
In          0 (0 bytes)           5 min
Read/Write  0 bytes / 0 bytes     5 min
Out         0 (0 bytes)           5 min
Tasks/Time  0 / 00:00:00.000      5 min
```

Fig. 2. The ETL process flow integrating data sources into the staging server.

VII. System Design and Implementation

A. Data Cleaning and Preprocessing

Data has been collected from different sources as discussed and cleaned. The data has been checked for null values, duplicate rows, data attributes and has been converted to pandas datatime format. We have both Transactions and Cashier Operations data that are not related to each other and only connected using operator ID. Also, the tranID is not unique on both sets of data. We have used the hash function in the Python pandas' tool to combine three columns(begin time, tranID, and operator ID) to create a unique transaction key for all data in both Transaction and Cashier Operations logs.

B. Data Modeling

We have decided to create a star schema design in the MySQL database to implement our Datawarehouse concept. The ER diagram shown below has been created with five dimensions table and one fact table. The fact table has the aggregates from both the Transaction and Operations data.

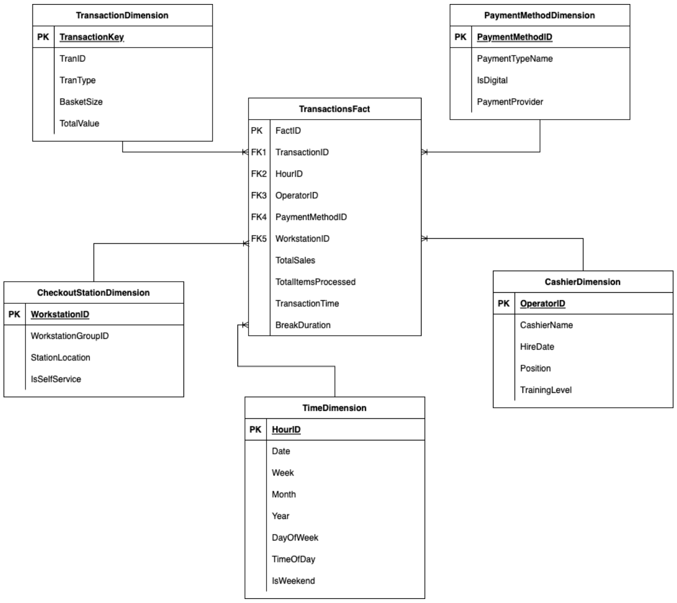


Fig. 3. Entity-Relationship Diagram of the Star Schema

The Time dimension table is extracted from the 'begin time' attribute. The Cashier dimension holds details about the operator id, name, and employee details. The Payment dimension holds payment details like cash or card transactions. The Transaction dimension holds transaction type and details. The Checkout dimension contains the workstation details.

This entire Datawarehouse is designed to support our analysis for this project and the dimensions can be changed or added in real-time based on the business requirement.

C. Datawarehouse Implementation

We have created a new database in MySQL and created two staging tables to load our cleaned Transactions and Operations data. We used the transaction key as a unique key to handle this data. We have created a dimensions table in the star schema and data loaded from the staging tables. Then the fact has been created with all the foreign keys referencing the dimension tables. The data lock occurred when we tried to combine both transaction and break time analysis facts in one insert SQL query. Hence, two more intermediate staging tables were created to load the transaction data facts and operations facts. Then, the final insert happened to load the data into the fact tables.

The SQL queries used to query the Datawarehouse, and the retrieved data have been stored in the pandas data frame. The tools pandas and matplotlib were used to create a visualization from the data generated from querying the Datawarehouse.

VIII. Analytics and Performance Evaluation

Due to the lack of credible data sources for a few factors that might affect transactions such as Holiday or Not and Cashier performance such as Training Level, Employee Performance Analysis in this project was measured using 3 metrics - Total Break Time, Total Basket Size handled, and Average Transaction Times. Nonetheless, such fields are included in the data warehouse model even though they just hold null values in order to imitate a real-life scenario where the data for these fields is available and can be used for analysis.

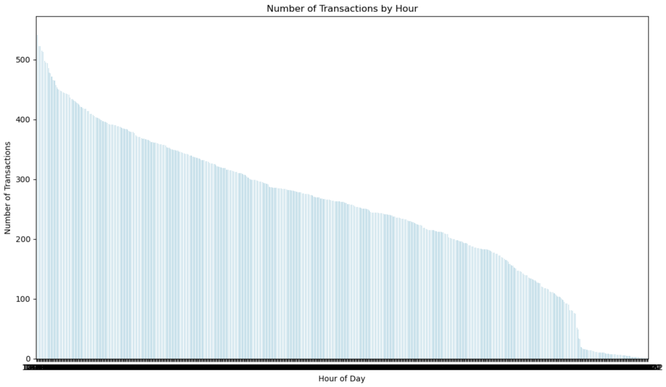


Fig. 4. Transaction Volume Throughout the Day

Insights: The number of transactions peaked in the initial hours of the store opening and experienced a gradual decrease throughout the day as the day progresses towards the end of operating hours. It's evident that the Peak Hours of Customer traffic are the first few hours after the store opening and the least customer traffic is in the final hours before the store closes.

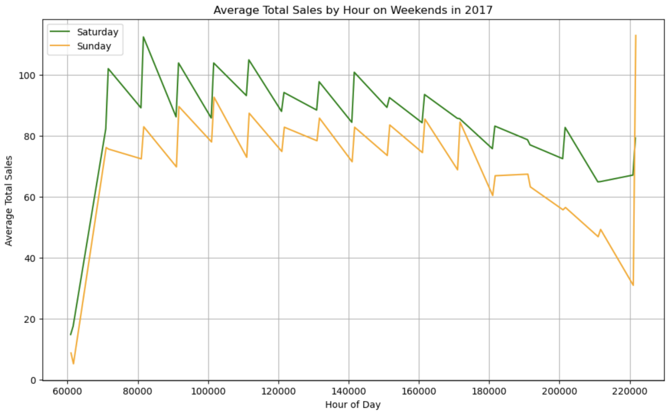


Fig. 5. Weekend Sales Patterns: Saturday vs. Sunday

Insights: As noticed before in transaction volume throughout the day, the sales for both days on the weekend peaked in the initial operating hours of the store- 8AM - 9AM on Saturdays 10 AM - 11AM on Sundays and decreased by the end of the day. The decrease was more significant on Sunday which is usually due to the time spent by customers for leisure from evening to night, typically 6PM -10 PM. For both days, the sales were good in the 3 hours following peak sales timings with only a few minor dips followed by major dips as the day progressed once those 3 hours finished.

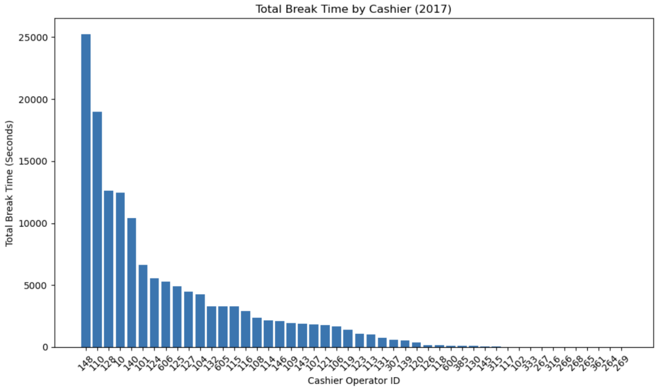


Fig. 6. 2017 Cashier Break Times Distribution

Insights: It can be noticed that 6 operators tend to take longer breaks than the rest of the operators. This can be due to longer experience or operational issues.Most cashiers lie on the right end of distribution indicating that most of them follow the break timings allotted to them.

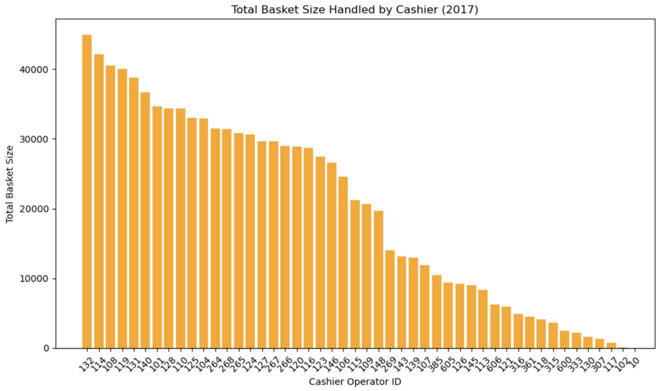


Fig. 7. Basket Size Handling Capacity per Cashier in 2017

Insights: Cashier with operator ID 132 handled the biggest basket size followed by 5 other cashiers.It is clear that most of the cashiers are experienced enough to handle moderate basket size excluding a few who handle big or small basket sizes.

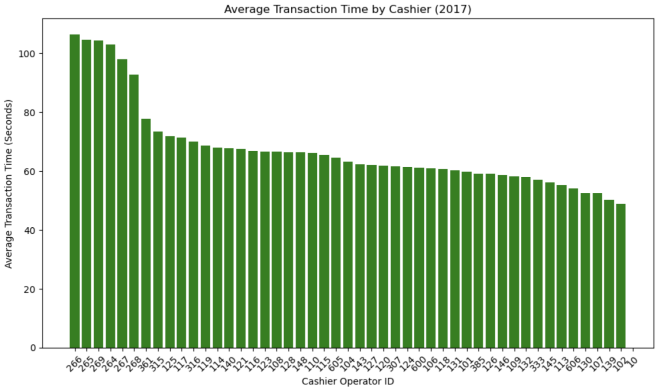


Fig. 8. Efficiency Analysis: Average Transaction Time per Cashier for 2017

Insights: Higher transaction time can be either due to great service or inefficient use of time.We can see that most of the cashiers are quick because they have shorter average transaction times excluding 6 cashiers that have high average transaction times.

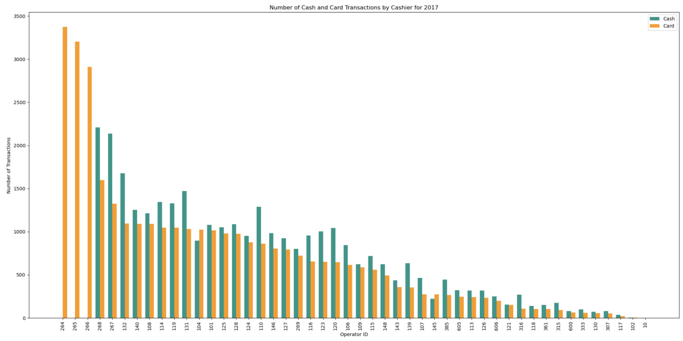


Fig. 9. Cashier Transaction Type Breakdown in 2017

Insights: All cashiers seem to handle both cash and card transactions except 3 operators who dealt only with cards throughout the year.It can be noticed that excluding a few, all cashiers handling both card and cash transactions seem to deal more cash than card transactions.

Employee Performance Analysis with Visualizations: Operator ID 264,265 and 266 had significantly higher Average Transaction times than the rest of the operators. This can be because they only handled card transactions and there might be a cause causing card transactions consuming more time than the cash ones. They also have low total break hours which means they do not take long breaks and spend most of their time on the registers when in store. Operator ID 132 dealt with the biggest basket size in 2017 while having a shorter average transaction time which sheds light on the efficiency of the operator in terms of handling several products quickly. This can be due to the fact that the operator majorly dealt cash transactions which seem to be quicker than card transactions as mentioned before. Operator also does not have significantly high break times. Operator ID 114 and 108 follow similar patterns.

Operator 145 has a short average transaction time but that does not necessarily mean that the operator is fast and agile as the basket size handled by the operator is low. But because the operator dealt slightly more cards than cash transactions,it may suggest that the cashier is indeed quick at their job. Operator has low break hours suggesting that the employee might be working on a part time basis leading to the small basket size. Operator 110 had a moderately high basket size along with a moderate average transaction time but significantly high break hours meaning that the operator does a decent job when at the register but spends a very high amount of time not being there. Operator 606 has a short average transaction time but that can be because a small basket size was handled with most transactions being cash (faster) while having high break hours indicating less overall throughput by the operator and bad utilization of time.

IX. Labor Recommendation

On the basis of all insights above, here are a few recommendations to plan labor:

A. Break Time Management

Implement policies that ensure cashiers are not crossing the limits of their respective allotted break times.

B. Training for Better Transactions

Cashiers with longer transaction times must be provided more training to improve the speed without compromising quality, while taking inputs from cashiers with shorter transaction times, with an emphasis on handling cash transactions.

C. Staffing during Peak Hours

Schedule more cashiers during peak hours to ensure enough staff is present to deal with high customer traffic. This staff should ideally have cashiers who have low transaction times and can handle the biggest basket sizes to reduce wait times.

D. Staffing during Comparatively Inactive Hours

Staffing can be reduced during less active hours. Eg: Evening to Night Operating Hours on Sundays.

E. Weekend Staffing

For the 3 hours following peak timings, schedule enough cashiers to handle the new incoming traffic along with the residual traffic immediately after the peak hours on the weekends.

F. Saturday Staffing

Schedule an adequate number of skilled operators on Saturday mornings and late afternoons to handle peaks in sales.

G. Sunday Staffing

Less staff is required on Sundays as compared to Saturdays due to fewer sales overall comparatively. So, staff size can be reduced. Schedule more operators to handle increases in sales in late afternoon or early evenings and on mornings during peak hours.

H. Incentive Programs

Enforce incentive programs that reward cashiers who follow break times properly, handle big basket sizes, and deal with a high number of transactions quicker than others.

X. Key Learnings from the Project

A. Data Management in Fact Tables

Challenge: Loading and managing multiple datasets in a fact table as part of a star schema.
Learning: We learned to manage and calculate different aggregates for each dataset. We also learned that sometimes dependencies need to be handled otherwise deadlocks can be caused. Learning to manage the data inside the Fact tables is very important for successful data warehousing and analytics, and to ensure that the data is correctly integrated.

B. Handling Data Uniqueness and Cleaning

Challenge: Ensuring uniqueness in transaction IDs, which were not unique in the initial dataset.
Solution: We came up with a way to combine various data columns to get distinct transaction IDs. This strategy guaranteed the accuracy of our analytics by not only resolving the immediate issue but also improving our team's proficiency in data preparation and cleaning.

C. ETL Process and Cloud Integration

Challenge: Setting up and configuring ETL (Extract, Transform, Load) processes in a cloud environment using Apache NiFi.
Outcome: We gained hands-on expertise with cloud-based ETL operations as a result of this assignment. We now have a better grasp of the challenges and specifications associated with integrating cloud data and carrying out activities in a cloud environment.

D. Data Type Conversion and Integrity

Challenge: Converting data types, specifically date-time formats, while maintaining data integrity.
Outcome: We learned from this exactly how crucial it is to convert all data types into the right formats to maintain the overall quality, integrity, and dependability of the data.

XI. Project Limitations

A. Limited Scope of Data Attributes

Situation: The dataset that was used contained little information about the operator and the type of payment. It did not originally have all the attributes that it should have in the real world.
Approach: In order to replicate a real-time data environment, we added these missing attributes to the dimensions and marked them as null.
Limitation: This approach exposed a critical limitation in the dataset's scope. While it was sufficient for our project's scope project, it did not fully represent a more comprehensive, real-world scenario.

B. Potential for Extending the Data Warehouse

Current Focus: The main goal of the data warehouse's design was to fit all the particular requirements of our analysis project.
Future Possibilities: This data warehouse could be developed to cover more varied dimensions in a real-world application, such as product information, promotions, customer demographics, etc.
Insight: This suggests that the current design is perfect for the project as it has been especially curated but has limitations in terms of scalability and adaptability for broader business intelligence applications.

XII. Innovation and Significance

In our project, we're making retail employee scheduling more detailed and informative by considering factors like transaction performance, punctuality, and consistency. We have carefully matched employee schedules with busy hours, ensuring the right staff is in the right place to manage workloads effectively and avoid burnout. We are also providing innovative labor recommendations such as Incentive Programs, Workload Distribution, Continuous Learning Culture, and Role Specialization. All these recommendations not only serve to benefit the customers with greater satisfaction but also encourage the employees to be as productive and efficient as they possibly can.

This approach not only improves customer happiness by providing prompt service but also creates a supportive work environment and fosters a healthy motivation for the employees to be and give their best.

Our focus on employees and customers puts our solution at the forefront of retail innovation.

XIII. Conclusion

During this project, the team explored how database theories can find application in the real time retail sector. We focused on the area of staff and resource scheduling and tried to optimize in these fronts. The real world application of enhancing efficiency of the staff and HR is the main focus of the project.

This involved deep analysis of sales data through advanced SQL and python. We also sifted through huge quantities of data to identify the time periods of high customer influx and rush. Alongside, we identified the cashiers who performed better during these rush hours. Using these insights, we aim to empower the retailers who seek to fine-tune their HR strategies to enhance their customer services.

Throughout the project, we faced several challenges of using the sales data to make decisions on planning the labor. In order to make well defined decisions it is necessary to draw meaningful insights from the data and utilize it accordingly. We learnt about the power of data driven insights and optimizations to be planned as per these insights.

XIV. Future Work

We plan to extend this project to add a variety of data elements and features. By including product specific trends in the sales and the impact of different promotional events(such as discounts, holidays etc), we will try to strengthen our insights, and further optimize operations. We also plan to learn and implement Machine learning, and use the technology to then forecast the future trends based on the previously collected historic data. We will therefore be able to predict the future rush hours, and strategize based on this information.

Along with this, we also plan to include the various employee features such as their work schedules, language skills and metrics to refine the analysis. Through this we will aim to deploy the right staff based on regional considerations and expected customers. For example- In the retail domain dealing with some regional condiments(such as Mexican), it would be beneficial to deploy a person who can understand that regional language.

Final Reflection: through this project, not only do we focus on applying the academic learning to real world problem solving, but also to optimize HR and staffing methods. This can be scaled across different industries as well and will help incentivize both, the customers and the staff.

References

[1] T. Antczak and R. Weron, ”Point of Sale (POS) Data from a Supermarket: Transactions and Cashier Operations,” *Data*, vol. 4, no. 2, p. 67, 2019. [Online]. Available: https://doi.org/10.3390/data4020067

[2] H. H.-C. Chuang, R. Oliva, and O. Perdikaki, ”Traffic-Based Labor Planning in Retail Stores,” *Production and Operations Management*, vol. 25, no. 1, pp. 96–113, 2016. [Online]. Available: https://doi.org/10.1111/poms.12403

[3] D. Smirnov and A. Huchzermeier, ”Analytics for labor planning in systems with load-dependent service times,” *European Journal of Operational Research*, vol. 283, no. 1, 2020. [Online]. Available: https://doi.org/10.1016/j.ejor.2020.04.036

[4] R. Kimball and M. Ross, *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, 3rd ed. Wiley, 2013.

Contributor Roles Taxonomy (CRediT)

Contributions to the project were as follows:

- **Conceptualization**: Priya Govindarajulu, Akshay Sodha, Esha Agarwal, Tanvi Pagrut
- **Methodology**: Priya Govindarajulu, Akshay Sodha, Esha Agarwal, Tanvi Pagrut
- **Software**: Priya Govindarajulu, Akshay Sodha, Esha Agarwal, Tanvi Pagrut
- **Validation**: Priya Govindarajulu, Akshay Sodha, Esha Agarwal, Tanvi Pagrut
- **Formal Analysis**: Priya Govindarajulu, Akshay Sodha, Esha Agarwal, Tanvi Pagrut
- **Investigation**: Priya Govindarajulu, Akshay Sodha, Esha Agarwal, Tanvi Pagrut
- **Resources**: Priya Govindarajulu, Akshay Sodha, Esha Agarwal, Tanvi Pagrut
- **Data Curation**: Priya Govindarajulu, Akshay Sodha, Esha Agarwal, Tanvi Pagrut
- **Writing – Original Draft Preparation**: Priya Govindarajulu, Akshay Sodha, Esha Agarwal, Tanvi Pagrut
- **Writing – Review & Editing**: Priya Govindarajulu, Akshay Sodha, Esha Agarwal, Tanvi Pagrut
- **Visualization**: Priya Govindarajulu, Akshay Sodha, Esha Agarwal, Tanvi Pagrut

Appendix

TABLE I: Project Evaluation Criteria

Criteria	**Link/Comment**	**Pts**
Presentation Skills	Includes time management	5
Code Walkthrough	Slide for code walkthrough and demo included for presentation	3
Discussion / Q&A	Slide for Q&A added, time allocated for questions	4
Demo	Demo source code created for presentation- In Jupyter notebook, we will establish a connection to the MySQL server, use our data warehouse database, execute SQL query for analysis, and create visualization from SQL query.	3
Version Control	Publicly accessible Git/GitHub repository. GitHub Repository	3
Significance to the real world	Included in the report and ppt under Innovation and Significance topic.	5
Lessons learned	Included in the report and presentation under the topic Key Learnings and Limitations.	5
Innovation	Data-driven approach for retail labor planning at a granular level. Slide included in the presentation.	5
Teamwork	The team connected using zoom meetings every week for follow-up. The screenshot attached.	5
Technical difficulty	Included the difficulties faced by the team under the same Key Learnings and Limitations in report and ppt	4
Practiced pair programming	Used Google Colab for pair programming. Google Colab Notebook	2

Table I continued from previous page

Criteria	**Link/Comment**	**Pts**
Practiced agile / scrum	Evidence submitted, Trello board for agile methodology: Trello Board	3
Used Grammarly / other tools for language	The team used Grammarly app while writing the content. Aiming for a score of 90 before adding the content to the report.	2
Slides		5
Report	Format, completeness, language, plagiarism, TurnItIn processing	7
Used unique tools	LaTeX report written in Overleaf. Prezi for presentation.	5
Performed substantial analysis using database techniques	SQL for analysis queries, visualizations with pandas and matplotlib, labor recommendations provided.	3
Used a new database or data warehouse tool not covered in class	Couchbase database used as a source database for the Datawarehouse.	3
Used appropriate data modeling techniques	ER diagram for the star schema modeled in Datawarehouse.	5
Used ETL tool	Files extracted from Google Drive using Apache NIFI to the staging server.	1
Demonstrated how Analytics support business decisions	Labor recommendation created based on analysis. Detailed in ppt and report.	3
Used RDBMS	MySQL for Datawarehouse with SQL queries for analysis.	1
Used Datawarehouse	Datawarehouse concept implemented using star schema in MySQL server.	1
Includes DB Connectivity / API calls	Jupyter notebook submitted with DB connectivity for Couchbase and MySQL using Python.	1
Used NoSQL	Couchbase NoSQL database used for operations data.	1
Total Points		**85**