Strategic Decision Making with PowerBI Prof Arpit Yadav FINAL REPORT



"Data-Driven Insights for Blinkit: A Power BI

Dashboard Project"

Submittied By: Abhinav Tripathi

2023JULB01176

Table of Content

1. Problem Statement	3
2. Data Requirement	5
3. Data Collection	7
4. Data Validation	8
5. Data Cleaning	11
6. Tools and Technologies Used	11
7. Dashboard Analysis	19
8. Storytelling with Data	21
9. Conclusion	23



Data-Driven Insights for Blinkit: A Comprehensive Power BI Report

Business Context

Blinkit, a prominent player in the grocery retail industry, is committed to delivering exceptional customer experiences and operational efficiency. To achieve this, the company recognizes the pivotal role of data-driven decision-making. By harnessing the power of data analytics, Blinkit aims to uncover valuable insights that will drive strategic initiatives and optimize business performance.

Key Challenges and Objectives

Challenges:

1. Sales Optimization:

- o Identifying top-performing products and outlets.
- Understanding factors contributing to underperformance.

2. Customer Satisfaction:

- o Pinpointing factors influencing customer ratings.
- o Improving product visibility and stock availability.

3. Operational Efficiency:

- o Reducing wastage through inventory optimization.
- o Addressing regional disparities in outlet performance.

4. Strategic Planning:

- Analyzing historical trends to replicate successful strategies.
- o Predicting future market shifts to stay ahead of competitors.

Objectives:

1. **Revenue Growth:** Increase total sales revenue across all product categories and outlets.

- 2. **Customer Loyalty:** Enhance customer ratings to improve loyalty and reduce churn.
- 3. **Operational Excellence:** Optimize outlet operations to ensure profitability and efficiency.

Data-Driven Approach

To address these challenges and achieve its objectives, Blinkit will employ a comprehensive data-driven approach:

1. Data Collection and Integration:

- Gather relevant data from various sources, including sales transactions, customer feedback, inventory records, and outlet performance metrics.
- Ensure data quality and consistency through rigorous data cleaning and preprocessing techniques.

2. Exploratory Data Analysis (EDA):

- Utilize statistical analysis and visualization techniques to uncover patterns, trends, and anomalies in the data.
- Identify key performance indicators (KPIs) and metrics that will drive decision-making.

3. Predictive Analytics:

- Develop predictive models to forecast future sales, customer behavior, and inventory demand.
- Utilize machine learning algorithms to identify potential opportunities and risks.

4. Prescriptive Analytics:

- Employ optimization techniques to determine the best course of action in specific scenarios.
- Simulate different scenarios to evaluate potential outcomes and make informed decisions.

Expected Benefits

By implementing a data-driven approach, Blinkit can expect the following benefits:

- Enhanced Decision-Making: Data-driven insights will empower decision-makers to make informed choices.
- Improved Operational Efficiency: Optimized inventory management, reduced wastage, and streamlined operations.
- Increased Customer Satisfaction: Personalized product recommendations, improved stock availability, and enhanced customer experiences.
- Competitive Advantage: Proactive response to market trends and innovative strategies.
- **Revenue Growth:** Increased sales and market share.

Through the effective utilization of data and analytics, Blinkit can position itself as a leader in the grocery retail industry, delivering exceptional value to its customers and shareholders.

2. Data Requirement

Specific Data Needs:

1. Product Attributes:

- o Item Identifier: Unique tracking for each product.
- o **Item Type**: Category-level classification (e.g., Beverages, Fruits).
- Fat Content and Weight: Details influencing customer preference.

2. Outlet Attributes:

- Location Type and Size: For regional and demographic analysis.
- Establishment Year: Historical performance correlation.

3. Performance Metrics:

Sales: Revenue generated by products and outlets.

- o **Item Visibility**: Shelf placement and marketing impact.
- Customer Ratings: Direct customer feedback.

Use Cases for Each Attribute:

- Item Type: Identify which categories contribute most to sales.
- Outlet Size: Understand how outlet size influences profitability.
- Sales and Ratings: Correlate sales trends with customer satisfaction.

To effectively address its business challenges and achieve its objectives, Blinkit requires a comprehensive dataset that includes the following key attributes:

Product Attributes

- 1. **Item Identifier:** A unique identifier for each product to track its performance and inventory.
- 2. **Item Type:** Categorical classification of the product (e.g., Beverages, Fruits, Dairy).
- 3. **Fat Content:** Information about the product's fat content to cater to health-conscious consumers.
- 4. **Weight:** The weight of the product, which can influence pricing and inventory management.

Use Case: Analyzing sales trends by item type can help identify popular categories and inform inventory decisions.

Outlet Attributes

- 1. **Location Type:** Categorical classification of the outlet's location (e.g., urban, suburban, rural).
- 2. **Outlet Size:** Numerical representation of the outlet's size, which can impact sales potential and operational costs.
- 3. **Establishment Year:** The year the outlet was established, providing historical context for performance analysis.

Use Case: Understanding how outlet size and location influence sales and profitability can help optimize store expansion strategies.

Performance Metrics

- 1. Sales: The total revenue generated by a product or outlet.
- 2. **Item Visibility:** A measure of the product's visibility on shelves, which can impact sales.
- 3. **Customer Ratings:** Direct feedback from customers about their shopping experience.

Use Case: Correlating sales trends with customer ratings can help identify factors driving customer satisfaction and loyalty.

3. Data Collection and Understanding

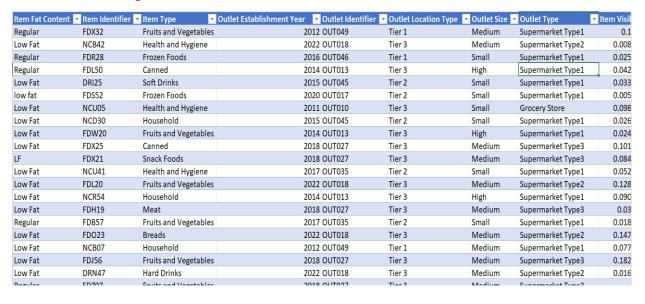
Dataset Overview:

- Size: $8,523 \text{ rows} \times 12 \text{ columns}$.
- Data Types:
 - Categorical: Fat Content, Item Type, Outlet Type.
 - o Numerical: Sales, Item Visibility, Weight.
 - o **Temporal**: Outlet Establishment Year.
- Coverage: Comprehensive representation of Blinkit's operations.

Initial Observations:

- 1. Missing weights for certain items.
- 2. Inconsistent entries for categorical values (e.g., "low fat" vs. "Low Fat").
- 3. Unbalanced data distribution:

Large outlets dominate sales, while smaller ones contribute



minimally.

Preliminary Insights:

- Sales Drivers: High sales correlate with large outlets and specific item types (e.g., snacks).
- **Visibility Impact**: Products with higher shelf visibility generally perform better.

4. Data Validation (Bias/Transparency/Reliability)

Validation Steps:

1. Completeness:

o Identified missing weights; flagged for imputation.

2. Accuracy:

- Verified numerical columns (e.g., sales ≥ 0).
- Cross-checked establishment years against plausible ranges.

3. Consistency:

Standardized categorical variables.

Removed duplicate records.

4. Bias Check:

o Ensured data is representative of all outlet types and locations.

Reliability Issues:

- Outlier sales in Tier 3 locations need further investigation.
- Visibility scores for some items appear inconsistent with sales figures.

Data validation is a critical step in the data analysis process. It ensures that the data is accurate, complete, consistent, and unbiased. This step helps to build trust in the insights derived from the data.

Validation Steps:

1. Completeness:

- o **Identify Missing Values:** Identify missing values, particularly in the "weight" column.
- Flag for Imputation: Mark missing values for appropriate imputation techniques (e.g., mean, median, mode, or more advanced methods).

2. Accuracy:

- Verify Numerical Values: Ensure that numerical columns like "sales" contain only positive values.
- Cross-Check Categorical Values: Verify that categorical variables like "outlet type" and "fat content" are accurate and consistent.
- Validate Temporal Data: Check if the "establishment year" for each outlet is within a reasonable range.

3. Consistency:

- Standardize Categorical Variables: Ensure consistent formatting for categorical variables (e.g., "Low Fat" vs. "low fat").
- Remove Duplicate Records: Eliminate any duplicate records to avoid data redundancy.

4. Bias Check:

- Representative Sample: Verify that the data is representative of all outlet types and locations, avoiding biases.
- o **Fairness and Equity:** Ensure that the data does not contain biases related to gender, race, or other sensitive attributes.

Reliability Issues:

1. Outlier Sales in Tier 3 Locations:

- o **Investigate Anomalies:** Conduct a deeper analysis to identify the reasons for unusually high or low sales in Tier 3 locations.
- Consider Data Quality Issues: Check for errors in data entry or data collection processes.
- Analyze External Factors: Explore external factors like economic conditions, local events, or competitive landscape that might influence sales.

2. Inconsistent Visibility Scores:

- Review Data Collection Methods: Assess the accuracy and reliability of the visibility scores.
- Cross-Reference with Sales Data: Analyze the correlation between visibility scores and sales to identify discrepancies.
- o Consider Subjective Bias: If visibility scores are based on human judgment, ensure consistency and minimize subjective bias.

5. Data Cleaning (Exploratory Data Analysis)

Tools Used:

1. Python:

- Imputation of missing weights using mean and median by item type.
- Z-score analysis to detect outliers in sales and visibility.

2. Power BI:

- Initial visualization of missing values and inconsistencies.
- o Preliminary correlations to identify key drivers.

3. **SQL**:

Used to query and structure raw data for further analysis.

EDA Highlights:

• Missing Value Imputation:

o Used mean weight for missing values grouped by item category.

• Outlier Detection:

o Sales outliers capped at the 95th percentile.

• Standardization:

 \circ Unified categorical variables (e.g., "low fat" → "Low Fat").

6. Tools Selection

Why Power BI for Blinkit's Data-Driven Transformation?

Power BI emerges as a powerful tool for Blinkit's data analysis and visualization needs due to several key advantages:

1. Interactivity:

• **Dynamic Filtering:** Users can easily filter data based on various parameters (e.g., product category, outlet location, time period) to gain specific insights.

• **Drill-Down Capabilities:** The ability to drill down into detailed data levels provides a granular view of performance metrics, enabling a deeper understanding of underlying trends.

2. Dynamic Dashboards:

- Real-Time Updates: Dashboards can be configured to display real-time data, ensuring that decision-makers have access to the latest information.
- **Data Source Integration:** Power BI seamlessly integrates with various data sources, including databases, spreadsheets, and cloud-based services, providing a unified view of business performance.
- Customizable Visualizations: Users can create custom visualizations to tailor the dashboard to their specific needs, enhancing data interpretation and decision-making.

3. Visualization Diversity:

- Rich Visualization Library: Power BI offers a wide range of visualization options, including charts, maps, and graphs, allowing for effective data storytelling.
- Interactive Visualizations: Users can interact with visualizations to explore data from different angles and uncover hidden patterns.
- **Data Storytelling:** Power BI enables the creation of compelling data stories by combining visualizations with narrative explanations, making complex data insights easily understandable.

By leveraging Power BI's capabilities, Blinkit can:

- Monitor Key Performance Indicators (KPIs): Track sales, customer satisfaction, and operational efficiency metrics.
- Identify Trends and Patterns: Uncover emerging trends and patterns in customer behavior, product performance, and outlet performance.
- Make Data-Driven Decisions: Support informed decision-making by providing actionable insights.
- Communicate Insights Effectively: Create visually appealing and interactive dashboards to share insights with stakeholders.

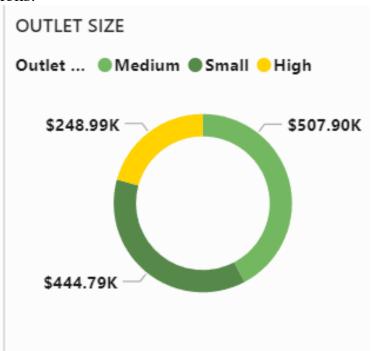
7. Graphs and Charts Used

Graph 1: Outlet Size

• **Description**: A donut chart shows revenue distribution among three outlet sizes: Small, Medium, and High.

• Insights:

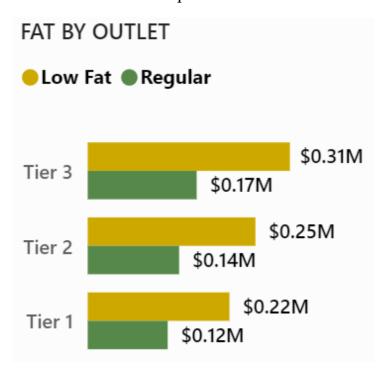
- High-sized outlets generate the highest revenue, contributing \$507.90K.
- o **Medium-sized outlets** follow, with \$444.79K in revenue.
- o Small-sized outlets contribute the least, with \$248.99K.
- Key Takeaway: Larger outlets (High and Medium) are more profitable, indicating they should be a priority for scaling operations.



Graph 2: Fat Content by Outlet Tier

- **Description**: A stacked bar chart compares revenue generated by Low Fat and Regular items across three outlet tiers.
- Insights:

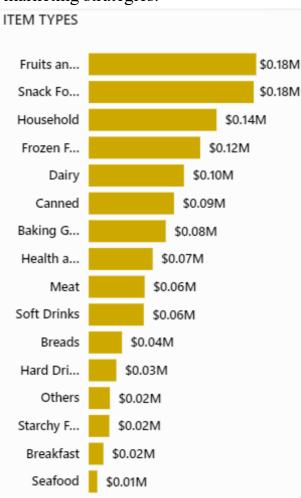
- Tier 3 outlets generate the most revenue for both Low Fat (\$0.31M) and Regular items (\$0.17M), with Low Fat items outperforming Regular items.
- o **Tier 2 outlets** follow, generating \$0.25M for Low Fat and \$0.14M for Regular items.
- o **Tier 1 outlets** have the lowest revenue, with \$0.22M from Low Fat and \$0.12M from Regular items.
- Key Takeaway: Low Fat items consistently outperform Regular items across all outlet tiers, suggesting a strong customer preference for health-focused products.



Graph 3: Item Types

- **Description**: A horizontal bar chart shows the revenue generated by different product categories.
- Insights:
 - The top contributors to revenue are Fruits and Vegetables and Snack Foods, both generating \$0.18M.

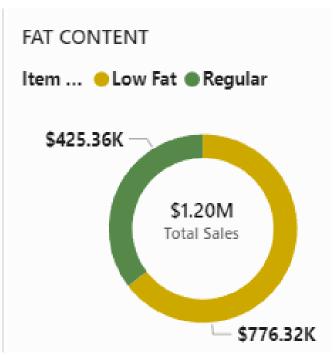
- Other significant contributors include Household items (\$0.14M) and Frozen Foods (\$0.12M).
- Categories like Seafood (\$0.01M), Breakfast items (\$0.02M), and Starchy Foods (\$0.02M) contribute the least.
- Key Takeaway: Focus on high-performing categories like Fruits and Vegetables, Snack Foods, and Household items to maximize revenue. Low-performing categories may require reevaluation or targeted marketing strategies.



Graph 4: Fat Content (Donut Chart):

- **Description**: The chart depicts the total sales split between "Low Fat" and "Regular" products.
- Insights:
 - "Regular" products contribute \$776.32K, making up approximately
 65% of total sales.

- "Low Fat" products contribute \$425.36K, accounting for 35% of total sales.
- Customers tend to prefer "Regular" products over "Low Fat"
 products, suggesting a focus on taste or other factors over health-



conscious choices.

Graph 5: Outlet Type (Table):

- **Description**: The table showcases metrics such as total sales, the number of items sold, average sales per item, average rating, and item visibility across different outlet types.
- Insights:
 - Supermarket Type1 dominates sales at \$787.55K, accounting for the majority of total sales. It also sells the highest number of items (5,577), indicating a higher footfall or better product variety.
 - Grocery Store contributes \$151.94K, ranking second but with a much smaller footprint compared to Supermarket Type1.

- Supermarket Types 2 and 3 have similar performance in sales (~\$131K), with slightly higher average sales per item in Supermarket Type2.
- o All outlets maintain a consistent average rating of 4, showing a uniform level of customer satisfaction.
- o **Item visibility** is highest in Supermarket Type1 (339), which could be driving its superior sales performance.

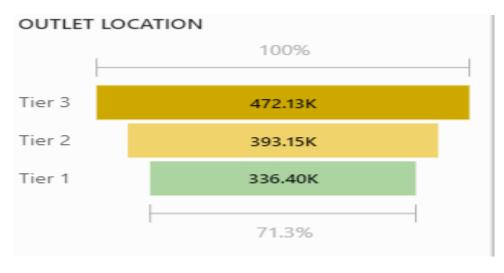
OUTLET TYPE					
Outlet Type	Total Sales	No of Items	Avg Sales	Avg Rating	Item Visibility
Supermarket Type1	\$787.55K	5577	\$141	4	339
Grocery Store	\$151.94K	1083	\$140	4	114
Supermarket Type2	\$131.48K	928	\$142	4	57
Supermarket Type3	\$130.71K	935	\$140	4	55

Graph 6: Outlet Location (Bar Chart):

• **Description**: This bar chart compares sales across outlets in Tier 1, Tier 2, and Tier 3 cities.

• Insights:

- o **Tier 3 locations** lead with \$472.13K in sales, indicating a strong customer base or fewer competitors in these areas.
- o **Tier 2 cities** follow with \$393.15K, showing moderate engagement and potential for growth.
- o **Tier 1 cities** lag at \$336.40K, which may suggest higher competition or lower purchasing power in these areas for these products.

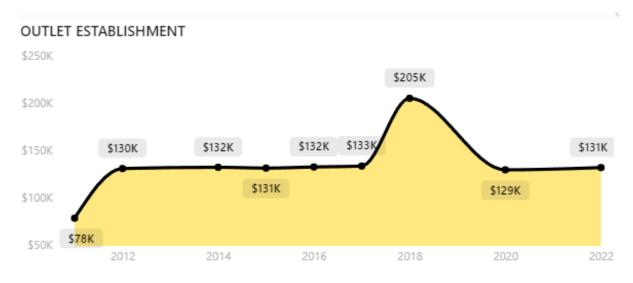


Graph 7:

This line graph illustrates the sales performance of an outlet over the years 2012 to 2022. The y-axis represents sales revenue in thousands of dollars, while the x-axis shows the years.

Insights:

- Sales Growth: The outlet experienced significant growth from 2012 to 2018, peaking at \$205K in 2018.
- **Stagnation and Decline:** After the peak, sales plateaued and then started declining, reaching \$131K in 2022.
- **Potential Factors:** Several factors could have contributed to this trend, such as changes in market competition, economic conditions, shifts in consumer preferences, or internal operational issues.



8. Dashboard Analysis



Overview of Dashboards:

1. Sales Performance:

o Metrics: Total sales, average sales per item/outlet.

o Insight: Large outlets and snacks dominate revenue.

2. Customer Insights:

o Focus: Customer ratings vs. sales.

o Insight: Products with lower visibility receive fewer ratings.

3. Outlet Efficiency:

o Metrics: Sales by location and outlet type.

o Insight: Tier 3 regions present growth opportunities.

Overall View

• **Totals:** The dashboard provides an overview of total sales (\$1.2M), the number of items sold (8523), average sales per item (\$141), and the average rating (3.9).

Outlet Establishment

- **Sales Trend:** The line graph shows a clear upward trend in sales from 2012 to 2018, with a peak of \$205K. However, after 2018, sales have stagnated and even declined slightly.
- **Potential Factors:** The decline in sales could be attributed to various factors, such as increased competition, changes in consumer preferences, or internal operational issues.

Outlet Size and Location

- **Size Distribution:** The pie chart reveals that most outlets are Mediumsized (47.21%), followed by Small (39.15%) and Large (13.64%).
- Location Impact: The bar chart shows that Tier 3 outlets contribute the most to total sales (\$444.79K), followed by Tier 2 (\$393.15K) and Tier 1 (\$336.40K). This suggests that outlets in Tier 3 locations are performing better.

Outlet Type

- **Sales Performance:** Supermarket Type1 dominates sales with \$787.55K, followed by Grocery Store (\$151.94K) and Supermarket Types 2 and 3 (\$131K+).
- Item Visibility: Supermarket Type1 also has the highest item visibility (339), which could be a contributing factor to its strong sales performance.

Fat Content

• Low-Fat Preference: The bar chart shows that low-fat items contribute more to total sales compared to regular items.

Additional Insights

- **Average Rating:** The average rating of 3.9 across all outlets suggests a generally positive customer experience.
- **No of Items:** The number of items sold varies across outlet types, with Supermarket Type1 having the highest number (5577).

Actionable Insights

- Focus on Tier 3 Locations: Given their higher sales contribution, prioritize expansion and marketing efforts in Tier 3 locations.
- Optimize Supermarket Type1: Analyze the factors contributing to Supermarket Type1's success and replicate them in other outlet types.
- **Promote Low-Fat Products:** Capitalize on the growing demand for low-fat products and increase their availability and visibility.
- Improve Operations: Address any operational inefficiencies that might be hindering sales growth.

Limitations

- **Data Period:** The analysis is based on a specific time period. Trends and performance may vary over different time frames.
- External Factors: The dashboard doesn't account for external factors like economic conditions, competitive landscape, or regulatory changes that could impact sales.

9. Storytelling (Business Impact)

Key Narratives:

1. Revenue Trends:

- o Products like snacks drive high sales.
- Opportunities in promoting low-fat items to health-conscious buyers.

2. Operational Focus:

o Large outlets in Tier 3 areas are pivotal for revenue growth.

3. Historical Learnings:

o Replicating 2018 strategies could sustain growth.

Analytical Storytelling for the Dashboard

Title: Unlocking Sales Potential: A Deep Dive into Outlet Performance

Executive Summary:

This dashboard provides a comprehensive view of our outlet performance, highlighting key trends and opportunities. While we've achieved significant sales growth, there's potential for further optimization. By focusing on specific outlet types, locations, and product categories, we can drive sustainable growth and enhance customer satisfaction.

Key Findings:

1. Sales Momentum:

- The outlet has experienced substantial sales growth from 2012 to 2018, peaking at \$205K.
- However, sales have plateaued since then, indicating a need for renewed growth strategies.

2. Outlet Performance:

- Supermarket Type1 is the top-performing outlet, contributing significantly to overall sales and item visibility.
- Tier 3 locations show promising potential, with higher sales compared to Tier 1 and 2.
- Medium-sized outlets dominate the market share, suggesting a preference for this size category.

3. Product Insights:

- Low-fat products are driving sales, indicating a growing healthconscious consumer base.
- Analyzing sales by category can help identify opportunities for product bundling and cross-selling.

Recommendations:

1. Strategic Focus on Tier 3 Locations:

- Invest in marketing and promotional activities in Tier 3 locations to capitalize on their growth potential.
- Conduct market research to identify local preferences and tailor product offerings accordingly.

2. Leverage Supermarket Type1's Success:

- Analyze the factors contributing to Supermarket Type1's high performance, such as inventory management, customer service, and marketing strategies.
- Replicate these best practices in other outlet types to drive sales growth.

3. Optimize Product Mix:

- Continuously monitor consumer preferences and adjust the product mix to prioritize high-demand, low-fat items.
- Consider introducing new products or product lines to cater to emerging trends and customer needs.

4. Enhance Customer Experience:

- Implement customer loyalty programs and personalized offers to foster long-term relationships.
- Train staff to provide excellent customer service and address customer concerns promptly.

By implementing these recommendations and closely monitoring key performance indicators, we can unlock the full potential of our outlets and achieve sustainable growth.

10. Conclusion

Actionable Recommendations:

1. Expand High-Performing Outlets:

Focus investments on large outlets in Tier 3 locations.

2. Optimize Product Strategy:

 Diversify offerings with innovative promotions for underperforming categories.

3. Enhance Customer Experience:

o Improve delivery, visibility, and quality for better ratings.

The provided dashboard offers valuable insights into the outlet's performance. By analyzing sales trends, outlet characteristics, and product preferences, we can identify key areas for improvement.

Key takeaways include:

- Sales Growth and Stagnation: The outlet's sales growth from 2012 to 2018 is encouraging. However, the subsequent plateau necessitates a renewed focus on growth strategies.
- Outlet Performance: Supermarket Type1 and Tier 3 locations emerge as top performers. Understanding the factors driving their success can inform strategic decisions.
- **Product Insights:** Low-fat products have significant potential, and analyzing sales by category can optimize product offerings.

Recommendations:

To capitalize on these insights and drive future growth, the following recommendations are proposed:

- Strategic Focus on Tier 3 Locations: Invest in marketing and tailor product offerings to local preferences.
- Leverage Supermarket Type1's Success: Replicate best practices in other outlet types.
- Optimize Product Mix: Prioritize high-demand, low-fat items and consider new product introductions.
- Enhance Customer Experience: Implement loyalty programs and focus on excellent customer service.

