

Product Sales Analysis Report

Introduction

This report outlines the data validation and cleaning procedures applied to the sales dataset, explores customer engagement through various sales methods, and provides recommendations to optimize revenue generation strategies.

1. Data Validation and Cleaning Steps

1.1 Dataset Overview

The initial examination of the dataset was conducted using `data.info()`, providing insights into data types, non-null counts, and overall structure.

1.2 Missing Values Check

A thorough investigation of missing values was performed using `data.isnull().sum()`. The analysis revealed that approximately 7.16% of the revenue column contained missing data. To address this, mean imputation was employed to fill the gaps:

```
data['revenue'] = data['revenue'].fillna(data['revenue'].mean())
```

[Copy](#)

1.3 Duplicate Entries Check

To ensure data integrity, duplicate entries were checked with `data.duplicated().sum()`, identifying any potential redundancy in the dataset.

1.4 Standardizing Sales Methods

The `sales_method` column was scrutinized for inconsistencies. Variations in naming conventions were standardized using the `replace` method to unify labels. For example, variations like 'em + call' were standardized to 'Email + Call', enhancing consistency across the dataset.

1.5 Anomaly Detection

Anomalies were identified through checks for invalid values across several key columns:

- Negative values in `week`, `nb_sold`, `revenue`, and `nb_site_visits` were flagged as illogical.
- The `years_as_customer` column was validated against a plausible range based on the company's establishment year (1984).
- Rows containing invalid `years_as_customer` values were filtered out, ensuring data validity:

```
invalid_years_mask = (data['years_as_customer'] < 0) |  
                    (data['years_as_customer'] > (current_year - 1984))  
data_cleaned = data[~invalid_years_mask]  
data_cleaned.reset_index(drop=True, inplace=True)
```

[Copy](#)

1.6 Saving Cleaned Data

The validated dataset was saved as `clean_product_sales.csv`, making it available for subsequent analysis.

2. Exploratory Analysis

2.1 Customer Count per Sales Method

The analysis revealed:

- **Email:** Engaged 7,465 customers, marking it as the most effective method.
- **Call:** Engaged 4,961 customers, showing significant personal interaction but requiring more resources.
- **Email + Call :** Engaged 2,572 customers, indicating limited effectiveness in maximizing reach.

2.2 General Revenue Distribution

Findings included: The most common **revenue range** is around **100**, with a **right-skewed distribution** indicating a small number of high-revenue outliers. **Significant clusters** of customers were identified in the **50-75 and 30-40 revenue ranges**.

2.3 Revenue Distribution by Sales Method

The comparative analysis of revenue by sales method indicated:

- **Call :** Lowest median revenue at approximately 50.
- **Email :** Moderate revenue with a median of 100 and a small interquartile range (IQR).
- **Email + Call :** Highest median revenue close to 175, exhibiting greater variability and potential for higher returns.

2.4 Revenue Over Weeks by Sales Method

Key observations:

- **Email** method generated the highest initial revenue but experienced a sharp decline.
- **Email + Call** demonstrated steady growth, surpassing Email by week 5.
- **Call** method consistently underperformed across all weeks.

3. Metrics Definition

3.1 Average Revenue per Customer

Importance: This metric provides insight into how much revenue each customer contributes on average. Monitoring average revenue helps the business identify trends over time and assess the effectiveness of different sales strategies.

Monitoring Strategy: Track this metric weekly to understand fluctuations and correlate them with marketing campaigns, sales methods, or seasonal variations.

3.2 Revenue per Sales Method

Importance: Understanding revenue generated by each sales method allows the business to allocate resources efficiently. By identifying the most lucrative methods, the business can optimize its sales strategy and improve overall profitability.

Monitoring Strategy: Analyze this metric monthly to evaluate the effectiveness of each sales method and make data-driven decisions on where to focus sales efforts and resources.

3.3 Average Site Visits per Customer

Importance: This metric indicates customer engagement and interest. By monitoring site visits, the business can assess whether increased visits correlate with higher revenue, thus providing insight into customer behavior and preferences.

Monitoring Strategy: Monitor this metric weekly alongside average revenue to identify any shifts in customer engagement that may affect sales performance.

4. Final Summary and Recommendations

4.1 Summary

Based on the comprehensive analysis of the sales methods employed for the new product line, the following summary have been drawn:

The **Email** method is the most efficient in terms of broad outreach and consistent revenue generation, engaging the highest number of customers. The **Call** method, while fostering engagement, requires significant resources and may not yield proportional revenue increases. Its effectiveness needs further evaluation. The **Email + Call** hybrid approach shows potential for higher sales but requires strategic targeting of leads to justify the additional resources involved.

4.2 Recommendations

1. Prioritize Email:

- Leverage the Email method for broader outreach, utilizing automated tools to manage large volumes effectively.
- Consider enhancing the content and targeting of email campaigns to further improve engagement and revenue.

2. Evaluate Call Strategy:

- Conduct a cost-benefit analysis of the Call method, focusing on high-value leads to ensure efficient use of resources.
- Explore the feasibility of training sales personnel to maximize the effectiveness of calls and improve conversion rates.

3. Test Email + Call:

- Implement the hybrid Email + Call approach on a limited scale to assess its viability. Target high-potential leads to measure the effectiveness of combining both methods.
- Monitor the results closely to determine if the increased effort translates into significant revenue gains.

4. Monitor Metrics Closely:

- Establish a routine for monitoring the three defined metrics: Average Revenue per Customer, Revenue per -Sales Method, and Average Site Visits per Customer.
- Use these insights to adjust strategies in real-time, ensuring that the business remains responsive to customer behaviors and market trends.

5. Consider Regional Analysis:

- Although the average revenue across states shows little variation, further regional analysis may uncover other factors affecting revenue. This insight could inform tailored sales strategies that cater to specific regions.

Conclusion

In conclusion, our analysis highlights critical insights into customer engagement strategies, revealing that the Email method engages 7,465 customers and generates consistent revenue, while the Call method shows lower performance with a median revenue of approximately 50. By implementing targeted recommendations, such as prioritizing Email outreach and evaluating the hybrid Email + Call approach, the business can optimize resource allocation and enhance customer interactions. This data-driven strategy establishes a framework for ongoing performance monitoring, positioning the organization to leverage real-time insights for sustained growth in a competitive market.