Practical exam for Data Analysis

Jean-Baptiste GOMEZ

Pens and Printers Sales

01

Data Presentation

In this project, we examine the sales performance of a new office stationery line using three sales methods: Emails, Calls, and a combination approach termed “Email + Call”. The analysis utilizes a dataset from our sales tracking system with 15,000 observations across 8 key variables—sales method, customer ID, units sold, revenue, customer relationship duration, site visit frequency over the past six months, and customer geographical location. Collected over the first six weeks post-launch, this data helps assess the effectiveness of each sales strategy in driving engagement and revenue, providing insights to optimize our sales tactics in an increasingly digital marketplace.

Data Validation

02

In our analysis aimed at enhancing product sales, we conducted a comprehensive review of a dataset featuring 15,000 customers. Key steps included confirming the absence of duplicates and ensuring data integrity across all columns.

- We addressed 1074 missing values in the revenue column by removing them to maintain data quality.
- Additionally, we streamlined the sales_method column by correcting duplicate values, reducing the unique sales method categories from five to three by amending entries mislabeled as "em + call" and "email".
- We also identified and removed outliers in the years_as_customer column, specifically entries claiming 47 and 63 years of loyalty, which exceeded the store's 40-year operational period since its 1984 inception.
- Post-cleanup, the validated dataset now contains 13,926 rows across 8 columns, with all remaining columns verified as clean and requiring no further adjustments.

Exploratory Analysis

03

1. How many customers were there for each approach?

The bar chart depicting sales methods reveals that **Email** was the most successful method, achieving 6,921 sales. It was followed by the **Call** method, which secured 4,780 sales, and the combined "**Email + Call**" method, which accounted for 2,223 sales. Notably, sales via Email constituted approximately 50% of the total sales.

Exploratory Analysis

2.1 What does the spread of revenue look like overall?

The boxplot chart displaying the distribution of revenue shows that the median value is slightly under 100.00. The interquartile range, highlighting the middle 50% of revenues per customer, lies approximately between 50.00 and 100.00. Furthermore, any values exceeding 190.00 are deemed outliers.

The histogram reveals two sales peaks at approximately $50 and $90, with the highest customer concentrations at these points, 2100 and 2600 customers respectively.Then the spread of the revenue on this histogram displays multi-modal distribution.

Exploratory Analysis

2.2 What does the spread of revenue look like for each sales method?

Revenue distribution by sales method reveals distinct differences: the **Email + Call** method has a wide spread and the highest median at approximately $190. The **Call** method shows a narrower spread with a median of around $50. The **Email** method, with a spread similar to the **Call** method, has a higher median near $90.

The chart at left shows that the **Email** sales method generated the most revenue, accounting for approximately 51.4% of the total. **Email + Call** contributed around 31.2%, while **Calls** produced the least, making up about 17.4% of total revenue.

Exploratory Analysis

3. Was there any difference in revenue over time for each of the methods?

The line graph data indicates that the **Email** sales method initially peaked in revenue at about $230,000 in Week 1 but then declined by over 58% to approximately $99,000 by Week 3. In contrast, the **Email + Call** method started lower but eventually surpassed other methods by the end of the observation period, suggesting it as the most effective strategy. This method showed consistent growth, peaking in Week 6, indicating potential for continued strong performance. Meanwhile, the **Call-only** method exhibited minimal changes in revenue, reinforcing that strategies either solely using **Email** or combining **Email** and **Calls** might be more effective in engaging customers.

Recommendations

07

4. Based on the Data, which method would you recommend we continue to use?

- The **Call** strategy yields the least favorable results, with low ROI and a high time demand of about 30 minutes per interaction.
- Integrating **emails** with **calls** could enhance the effectiveness of the standalone calling approach, as emails require minimal effort yet provide valuable product insights.
- The **Email + Call** strategy is **highly effective**, displaying an upward trend in both average and total weekly revenue, and is more time-efficient with an average call duration of just 10 minutes.
- Given its efficiency and effectiveness, **Email + Call is recommended as the primary sales strategy**, with **Email** alone as a **secondary option**.
- Moving away from the **Call** approach is advised due to its poor performance and significant time consumption.
- Setting the recent sales performance of the **Email + Call** method as a benchmark, an increase in the percentage of new products sold over the next six months would indicate progress towards enhancing sales and customer creativity.

Business Metric to Monitor

08

			mean
sales_method	Call	Email	Email + Call
week			
1	35.350243	87.496401	128.895649
2	43.602606	100.141243	154.246959
3	41.759528	92.762573	150.423272
4	51.447061	108.877141	180.822633
5	53.146087	108.892313	187.588343
6	66.166089	130.976519	227.770635

			mean			median
sales_method	Call	Email	Email + Call	Call	Email	Email + Call
week						
1	35.350243	87.496401	128.895649	35.195000	86.535000	128.720000
2	43.602606	100.141243	154.246959	43.460000	99.250000	154.170000
3	41.759528	92.762573	150.423272	41.470000	91.800000	149.930000
4	51.447061	108.877141	180.822633	51.270000	108.090000	180.550000
5	53.146087	108.892313	187.588343	52.970000	108.020000	187.195000
6	66.166089	130.976519	227.770635	65.970000	130.450000	227.405000

How should the business monitor what they want to achieve?

1. **Prioritize Performance Metrics:** Focus on tracking average weekly revenue generated by each sales method as the primary performance indicator to optimize revenue and profit growth.

2. **Superiority of Email + Call:** Data shows that the **Email + Call** method consistently yields higher average weekly revenues compared to other methods, indicating its effectiveness in boosting sales.

3. **Detailed Revenue Analysis:** Weekly average revenues for the **Email + Call** method are detailed, ranging from $128.89 in week 1 to $227.77 in week 6, highlighting its upward revenue trend despite occasional dips.

4. **Continuous Monitoring and Adjustment:** Regular analysis of this key metric enables timely strategic adjustments, ensuring sales strategies align with company objectives and market demands.

Final Summary and Conclusion

In the coming weeks, we recommend focusing on several strategic actions to boost sales:

1. Monitoring Key Metrics

2. Launching a Loyalty Campaign

3. Maintaining Awareness of Marketing Trends

4. Enhancing Data Collection for In-depth Analysis

To conclude, adopting the **Email + Call** method as a multi-channel approach yields a higher return on time invested and significantly enhances operational efficiency and revenue generation. It is advisable to make follow-up calls a day or two after sending product information emails, as experience shows that calls delayed by a week result in reduced engagement and recall. Prompt follow-ups effectively leverage timing to boost sales conversion rates, positioning the team to excel and outperform the competition with these focused strategies.

Thank for your attention!!!