

# Data Analysis Report

## Data Validation

The dataset contains **15000 rows and 8 columns** before cleaning and validation. All of the columns have been validated against the criteria in the dataset table:

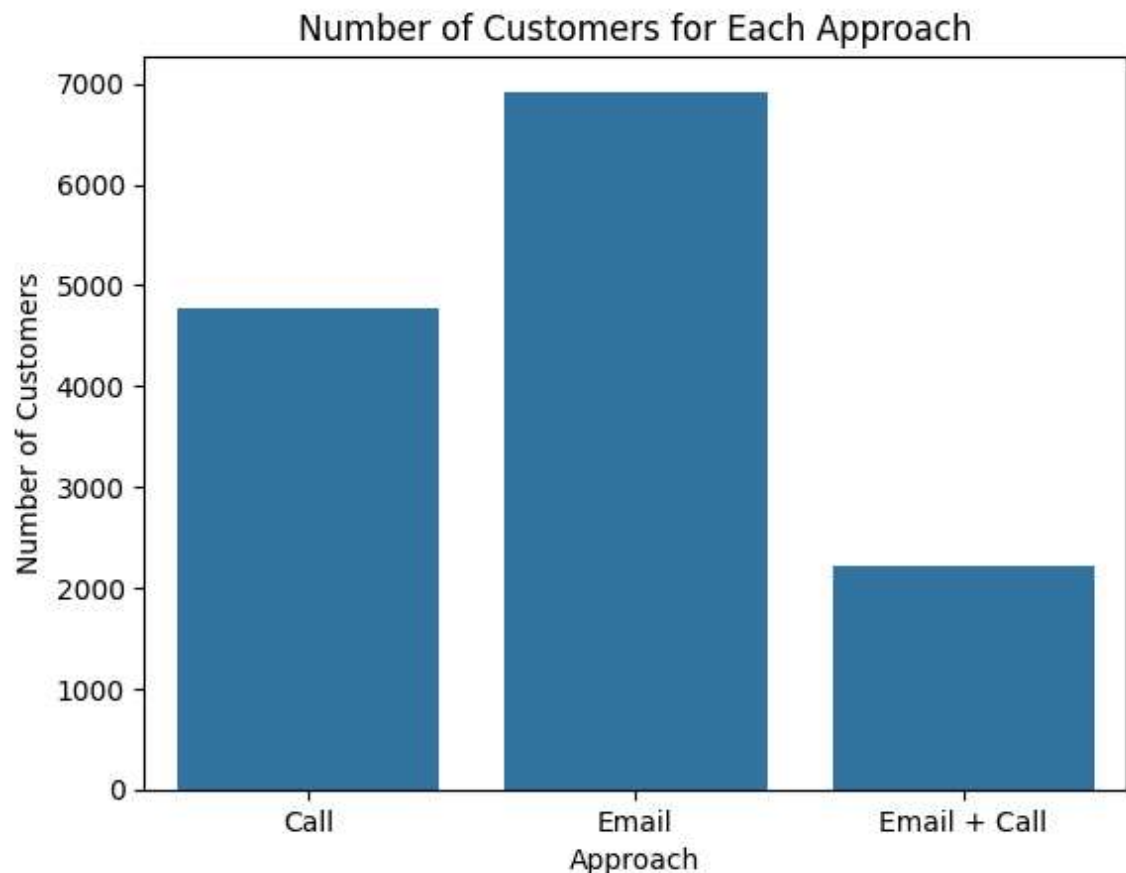
- **sales\_method** : 5 categories without missing values this is against the dataset table inspection there were 2 Typographical errors resulting in inconsistent categories necessary cleaning tasks have been done and the **sales\_method** column has been validated as the description.
- **customer\_id** : 15000 unique ID's without missing values, same as the description. No cleaning is needed.
- **state** : 50 states without missing values, same as the description. No cleaning is needed.
- **week** : 6 week categories without missing values, same as the description. No cleaning is needed.
- **nb\_sold** : numeric values, Based on the summary statistics, the data in this column appears to have no outliers. The values range from 7 to 16 with a mean of 10.008, indicating a relatively consistent distribution. same as the description. No cleaning is needed.
- **revenue** : numeric values, 1074 missing values, Based on the analysis, there appears to be no significant relationship between the missing data and other categorical variables in the dataset. This suggests the data is Missing Completely at Random (MCAR) Given this and its relatively small volume, imputation has been selected as the preferred method for handling the missing values. This approach preserves sample size while minimizing potential bias, ensuring robust subsequent analyses.
- **years\_as\_customer** : numeric values without missing values, it reflects the number of years a customer has been purchasing from the company since 1984, with a max of there were 2 entries with values over 40 so Due to the lack of additional context and the minimal impact on the dataset, removing the two rows with inaccurate values was done. column has been validated as the description.
- **nb\_site\_visits** : numeric values without missing values, Based on the summary statistics, the data for this column appears to have no outliers. The values range from 12 to 36 with a mean of 24.9, indicating a relatively consistent distribution. same as the description. No cleaning is needed.

After the data validation, the dataset contains **13924 rows and 8 columns** without missing values.

# Exploratory Analysis

## How many customers were there for each approach

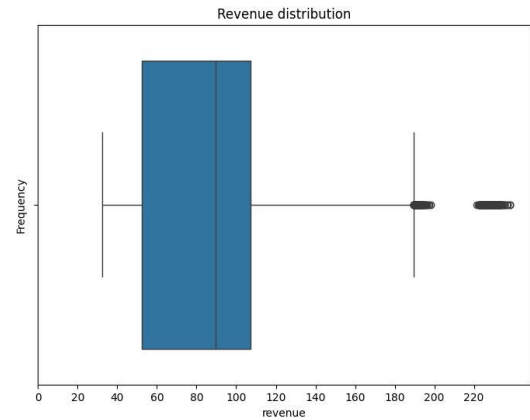
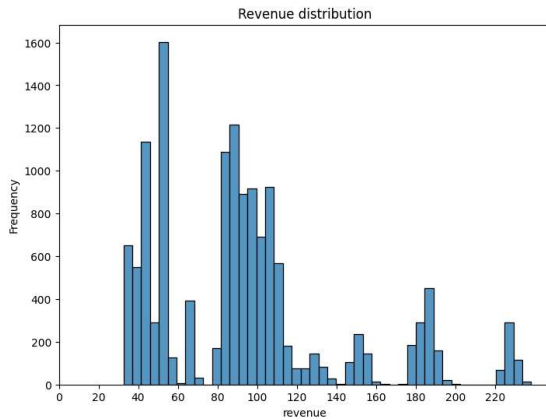
Email was the most used customer acquisition method, followed by phone calls, with hybrid approaches demonstrating the lowest amount of customers.



## What does the spread of the revenue look like overall? And for each method?

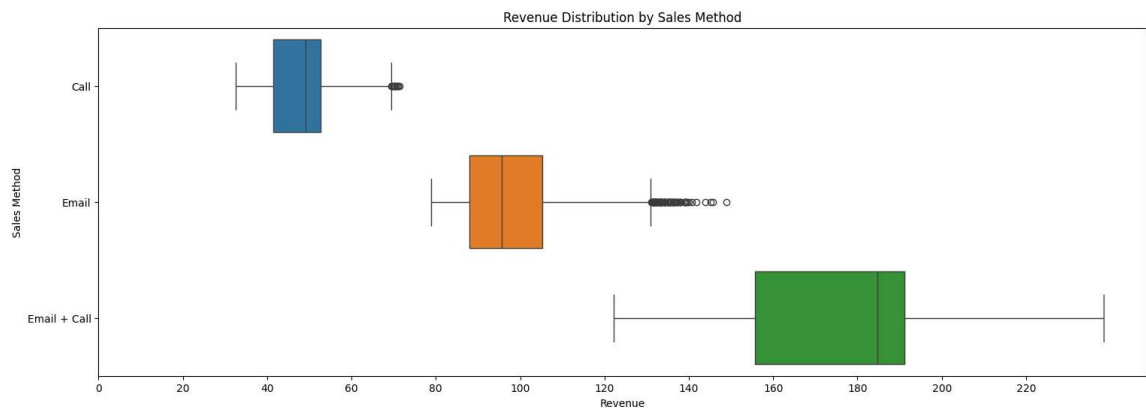
### Spread of the revenue overall

Our revenue data shows a wide range, with most earnings falling between 40,000 and 120,000. However, we have some standout performers bringing in up to \$220,000. This spread indicates diverse market performance.



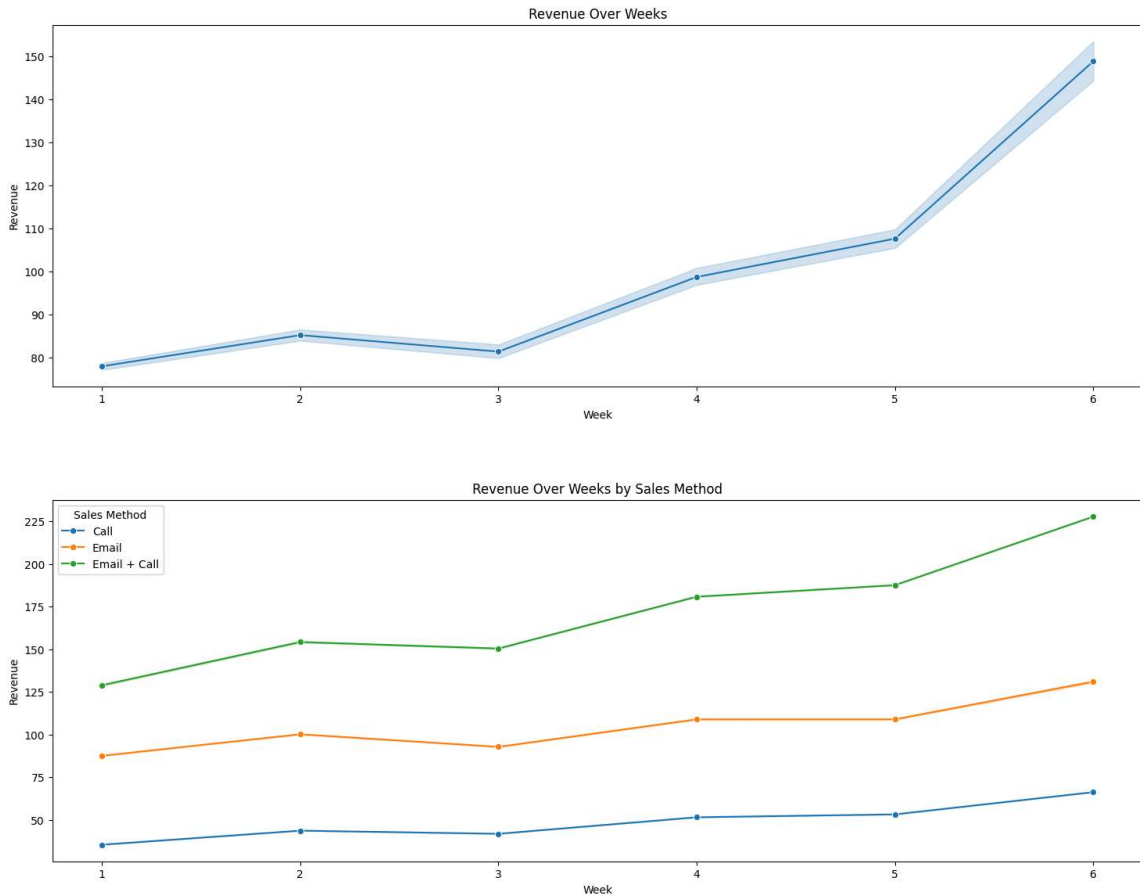
## The spread of the revenue for each method?

Our sales data reveals interesting patterns across different methods. Cold calling yields consistent but lower revenues. Email marketing shows better overall performance with some standout successes. However, our combined email and call approach is the clear winner, generating the highest average sales and greatest potential for big deals. While it's the most variable, it's also proving to be our most lucrative strategy. This suggests we should focus on refining and expanding our integrated email and call campaigns to maximize our revenue potential.



## Was there any difference in revenue over time for each of the methods

Revenue overall and from all sales methods increased over the six weeks period. The "Email + Call" method consistently generated the highest revenue, showing the most significant growth. The "Email" method also saw positive growth, while the "Call" method had the slowest increase. These findings suggest a combined sales approach is most effective, but there's potential for improvement in email and call strategies.

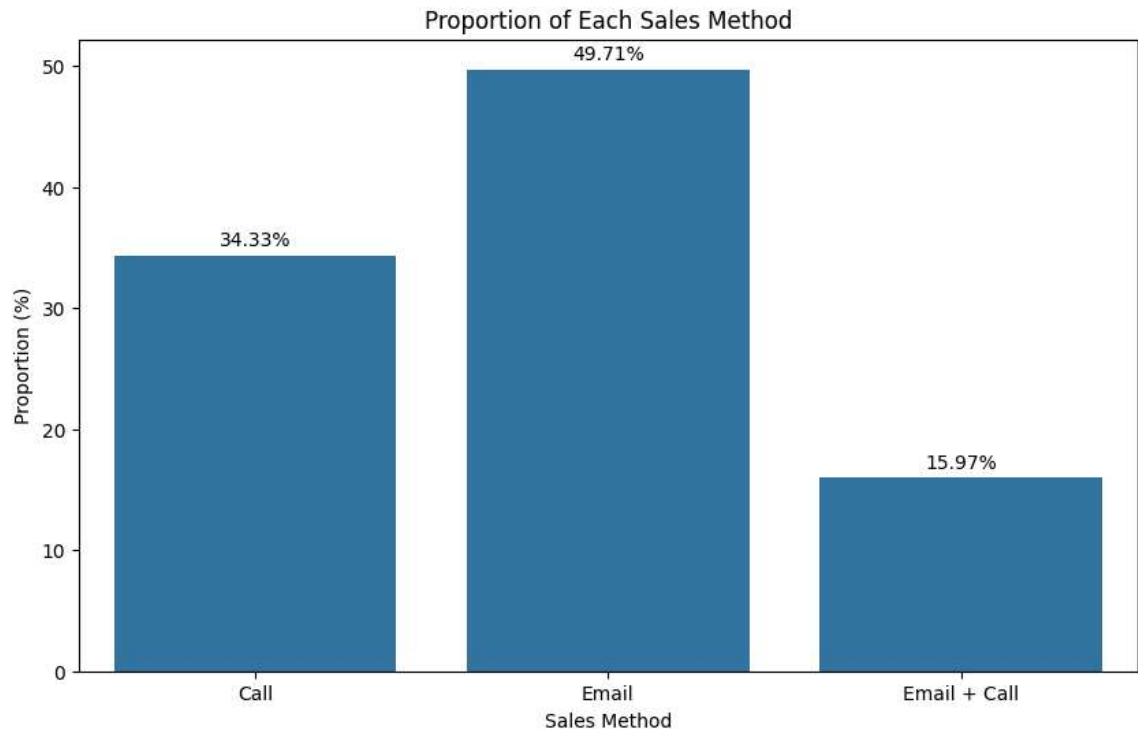


## Recommend sales approach

Based on the data analysis, the sales team should **prioritize the email + call approach**. This method consistently generates the highest revenue while maintaining a relatively low customer interaction time of 10 minutes. While the email-only approach also yields positive results, the incremental time investment for a brief follow-up call is justified by the potential for increased revenue. Conversely, the call-only method demonstrates lower revenue generation and a significantly higher time commitment of approximately 30 minutes per customer, making it less efficient.

## Business Metrics

Given the company's objective to sustain high revenue while adopting innovative sales strategies, the **"email + call" approach percentage is a key business metric**. This metric currently stands at **15.97%** of total sales. This method demonstrates high revenue growth while maintaining efficient resource allocation and time efficiency, aligning with the company's goals. A strategic priority is to increase the utilization of this sales tactic. Additionally, **overall revenue remains a crucial performance indicator** for assessing overall company performance.



## Final summary and recommendations

### Final Summary:

The analysis focused on evaluating different sales methods to determine the most effective strategy for maximizing revenue while minimizing customer interaction time. The dataset underwent thorough validation, with key findings highlighting the following:

#### Sales Methods Evaluation:

- The "email + call" method consistently generated the highest revenue with the low customer interaction time, making it the most efficient and effective approach.
  - The "email-only" method also showed positive results, though slightly less effective than the "email + call" combination.
  - The "call-only" method was the least efficient, requiring more time and generating lower revenue.
- Business Metrics:**
- The percentage of sales using the "email + call" approach is a critical business metric, currently standing at 15.97% of total sales.
  - Overall revenue remains a key performance indicator for the company.

### Recommendations:

**Prioritize the "Email + Call" Sales Method:**

- Encourage the sales team to focus on the "email + call" approach to optimize both revenue and time efficiency.
- Consider training and resources to further improve the effectiveness of this method.

**Monitor and Increase the Use of the "Email + Call" Approach:**

- Set targets to increase the utilization of this sales method, given its proven effectiveness.
- Regularly track the percentage of sales using this approach as a key business metric.

**Optimize Resource Allocation:**

- Reduce reliance on the "call-only" method, as it is less efficient and more time-consuming.
- Allocate resources to support the more effective sales strategies, such as "email + call" and "email-only" methods.