Product Sales Analysis

## 1. Data Validation & Cleaning

To ensure reliable analysis, each column in the dataset was validated and cleaned as follows:

| Column	Description	Cleaning/Validation Steps
week	Week number since product launch	Validated as integer; values range from 1–6 (as expected from a 6-week pilot).
sales_method	Method used to sell to the customer	Standardized values by stripping spaces, converting to lowercase, and title casing. Fixed typos (em + call → Email And Call).
customer_id	Unique ID per customer	Verified uniqueness. Dropped duplicates where present.
nb_sold	Number of new products sold	Checked for non-negative integers.
revenue	Revenue from sales	Identified 1,074 missing values. Retained rows with available values; imputation not appropriate here.
years_as_customer	Tenure with company	Removed unrealistic values over 40 (company founded in 1984).
nb_site_visits	Visits to website in last 6 months	Filled any missing values with 0. Validated as non-negative integers.
state	Customer state	Validated 50 distinct US state entries.

2. Exploratory Data Analysis (EDA)

Q1: How many customers per approach?

Number of Customers per Sales Method
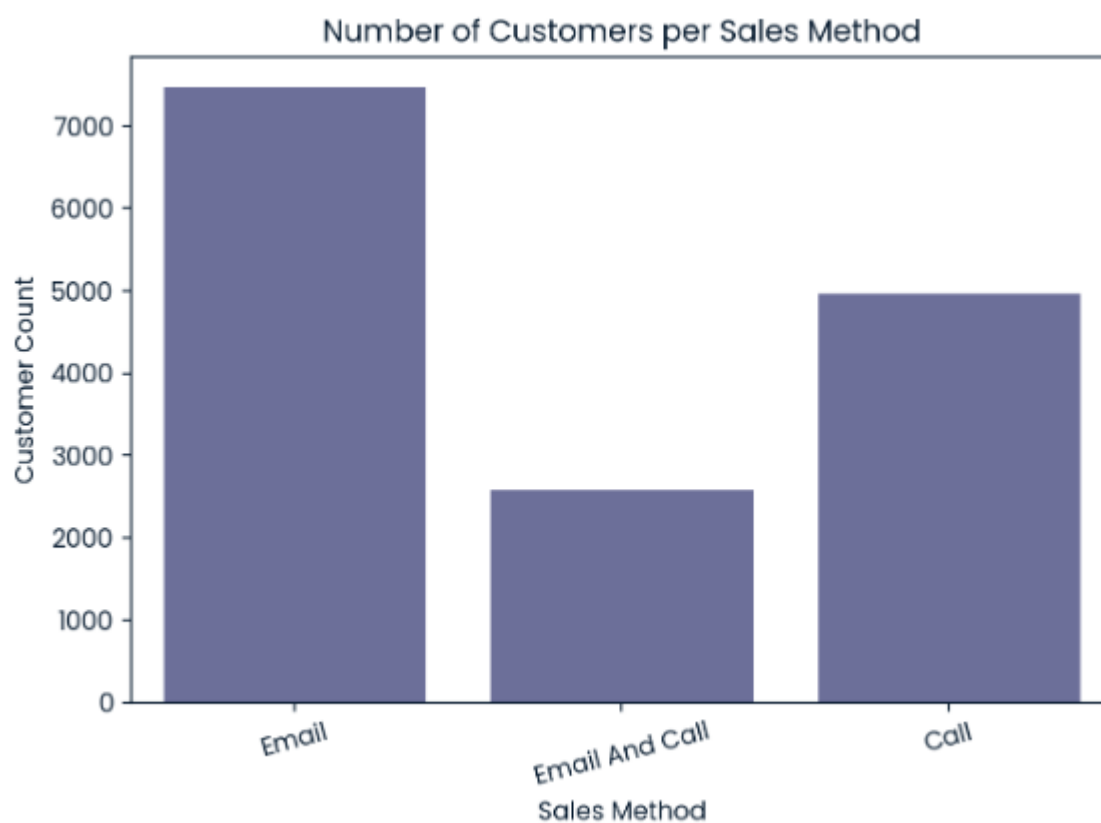
- **Email:** ~7,465 customers

- **Call:** ~4,961 customers

- **Email + Call:** ~2,549 customers

Observation: Email was the most scalable outreach method.

Q2: What is the overall revenue spread?

Overall Revenue Distribution
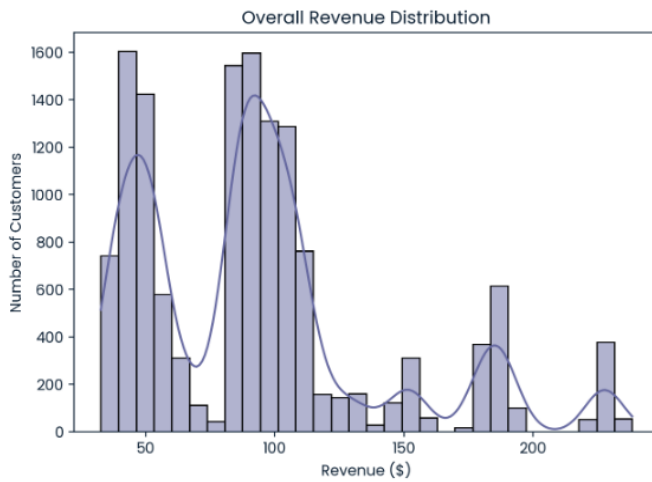
- Most customers generate less than $200.
- The distribution is right-skewed with some high-revenue outliers.


Revenue Spread by Sales Method
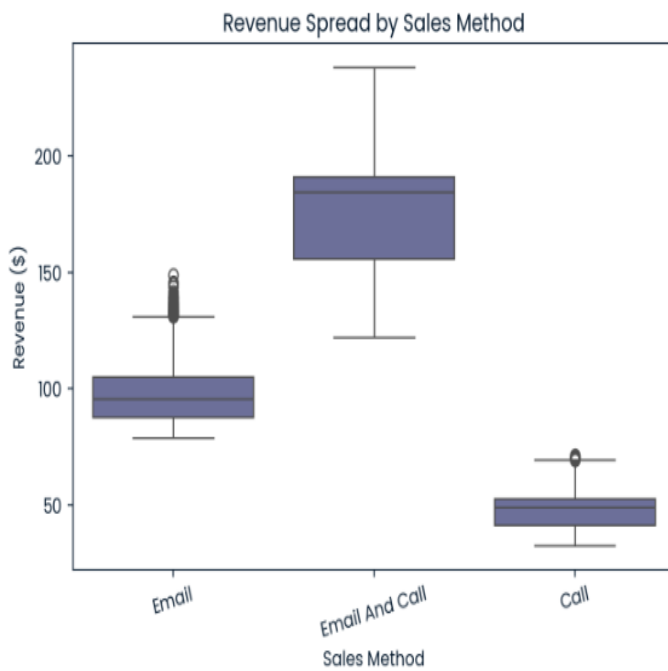
- Email + Call generates the highest median revenue.
- Email is more consistent with fewer extreme values.
- Call has a tighter, lower range

Q3: Revenue per method?

**Average Weekly Revenue by Sales Method**
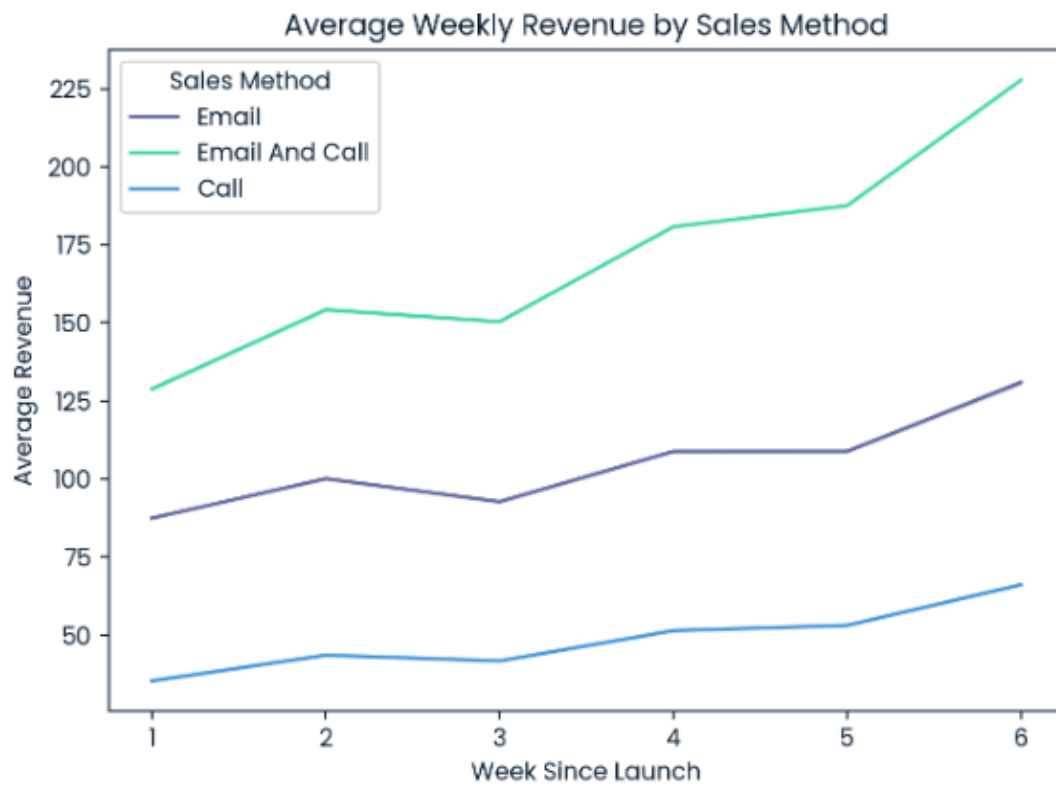
- Email + Call shows consistently higher weekly performance.
- Email is stable across all weeks.
- Call lags behind both.

Q4: Revenue over time by method?

Revenue per Hour of Effort by Sales Method
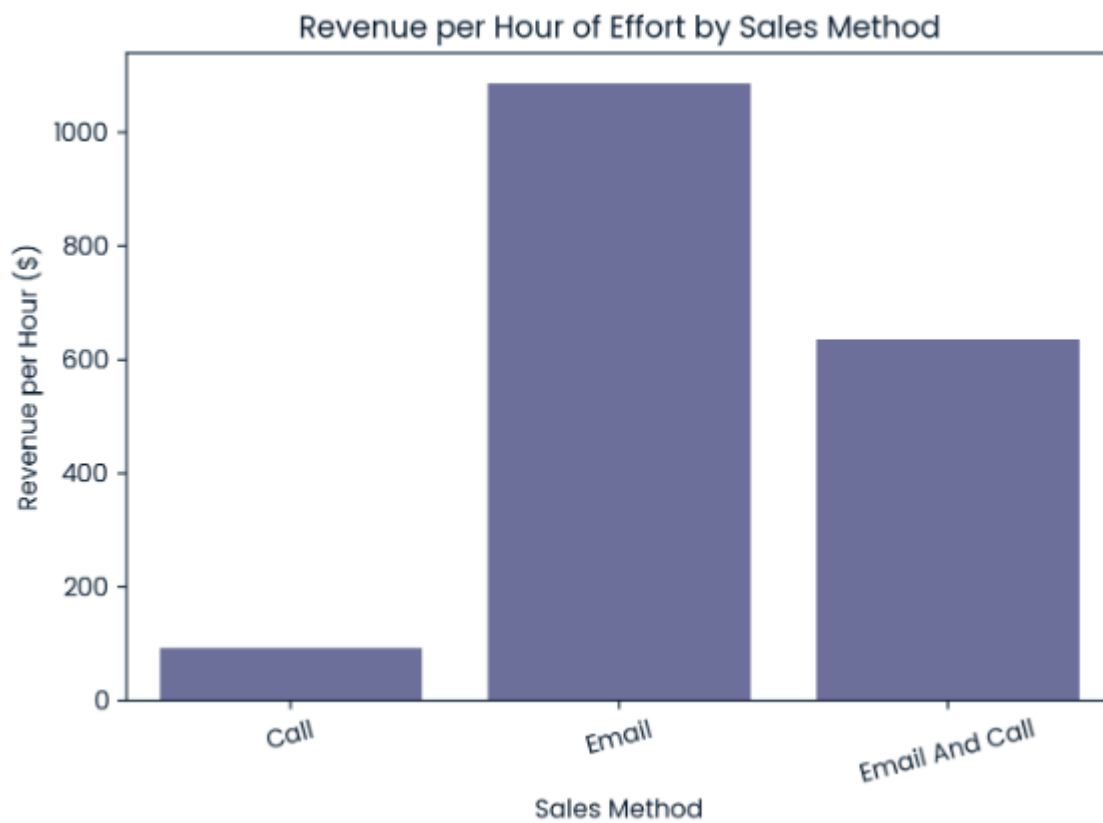
- Email is clearly the most efficient.
- Email + Call performs well but is more time-intensive.
- Call has the lowest efficiency.

Additional Insight: Customer Loyalty by Method

Customer Loyalty by Sales Method
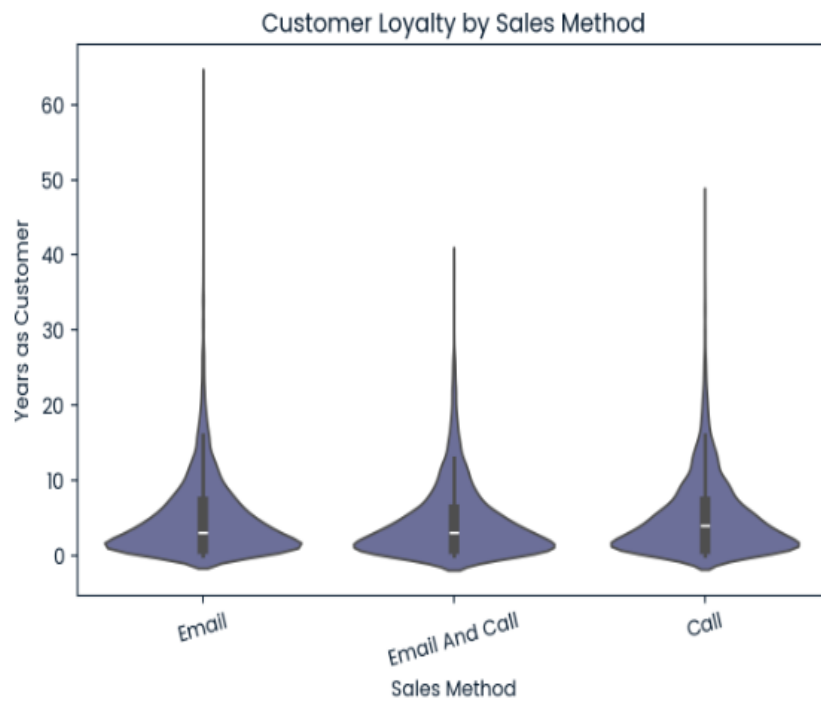
- Email and Email + Call seem to attract longer-tenured customers.
- Call is more frequently used on newer customers.

Revenue per Hour of Effort

This metric accounts for *both* effectiveness and cost (effort/time).

Effort Assumptions:

Method	Estimated Time per Customer
Email	5 minutes (0.083 hr)
Call	30 minutes (0.5 hr)
Email + Call	10 minutes (0.25 hr)

Initial Values from Dataset:

Sales Method	Avg Revenue	Total Effort (Hours)	Overall Revenue/Hour
Email	$97.13	619.68	$1,084.95
Call	$47.60	2,481.00	$91.72
Email And Call	$183.65	643.00	$634.92

4. Summary and Recommendations

Summary:

After cleaning and validating the dataset, we analyzed the performance of three sales methods used to promote a new product line. The key metric defined was **Revenue per Hour of Effort**, which balances sales effectiveness against time investment.

Key findings include:

- **Email** campaigns resulted in the highest **efficiency**, generating over **$1,080 per hour** of effort. With minimal time investment per customer, this method is ideal for large-scale outreach.

- **Email And Call** delivered the **highest average revenue per customer**, but required more time. It achieved a solid **$634 per hour**, making it a good option for high-value leads.

- **Call-only** outreach performed the worst, with a low **$91 per hour**, due to high effort and comparatively low revenue.

These insights help differentiate between **scalable**, **targeted**, and **inefficient** strategies.

Recommendations:

- **Use Email as the Primary Channel**
  It offers the best return on effort, and is ideal for reaching a large number of customers efficiently.
- **Deploy Email And Call Strategically**
  Reserve this hybrid approach for key accounts or high-potential customers where deeper engagement could yield strong results.
- **Phase Out Call-Only Campaigns**
  The high time cost and low return make this method inefficient. Limit its use to exceptional cases where personal outreach is essential.
- **Adopt Revenue per Hour as a Core KPI**
  This metric provides a meaningful measure of sales strategy efficiency and should be monitored weekly to guide future decisions.
- **Ensure Consistent Data Entry**
  Standardize sales method naming in CRM systems to avoid fragmentation (e.g., "Email + Call" vs. "Em + Call") and ensure accurate performance tracking.
- **Explore Customer Segmentation for Deeper Insights**
  Future analysis could evaluate sales performance by state, tenure, or website engagement level to refine targeting strategies further.