Product Sales Data Analysis Pens and Printers

(Data Analyst Practical Exam)

Biman Mondal

Nov. 18, 2025

Project Overview

- Process and analyze new product line of stationary sales data from last 6 weeks

- Quantify sales information
  - Across sales methods
  - Revenue over time

- Make recommendation
  - Sales metric to monitor
  - Evaluate sales methods

Dataset

Column Name	Details
week	Week sale was made, counted as weeks since product launch
sales_method	Character, which of the three sales methods were used for that customer
customer_id	Character, unique identifier for the customer
nb_sold	Numeric, number of new products sold
revenue	Numeric, revenue from the sales, rounded to 2 decimal places.
years_as_customer	Numeric, number of years customer has been buying from us (company founded in 1984)
nb_site_visits	Numeric, number of times the customer has visited our website in the last 6 months
state	Character, location of the customer i.e. where orders are shipped

```
   week  sales_method                           customer_id  nb_sold  \
0     2         Email  2e72d641-95ac-497b-bbf8-4861764a7097       10
1     6  Email + Call  3998a98d-70f5-44f7-942e-789bb8ad2fe7       15
2     5          Call  d1de9884-8059-4065-b10f-86eef57e4a44       11
3     4         Email  78aa75a4-ffeb-4817-b1d0-2f030783c5d7       11
4     3         Email  10e6d446-10a5-42e5-8210-1b5438f70922        9

      revenue  years_as_customer  nb_site_visits      state
0   99.674980                  0              24    Arizona
1  225.470000                  1              28     Kansas
2   52.550000                  6              26  Wisconsin
3  108.872129                  3              25    Indiana
4   90.490000                  0              28   Illinois
(15000, 8)
```
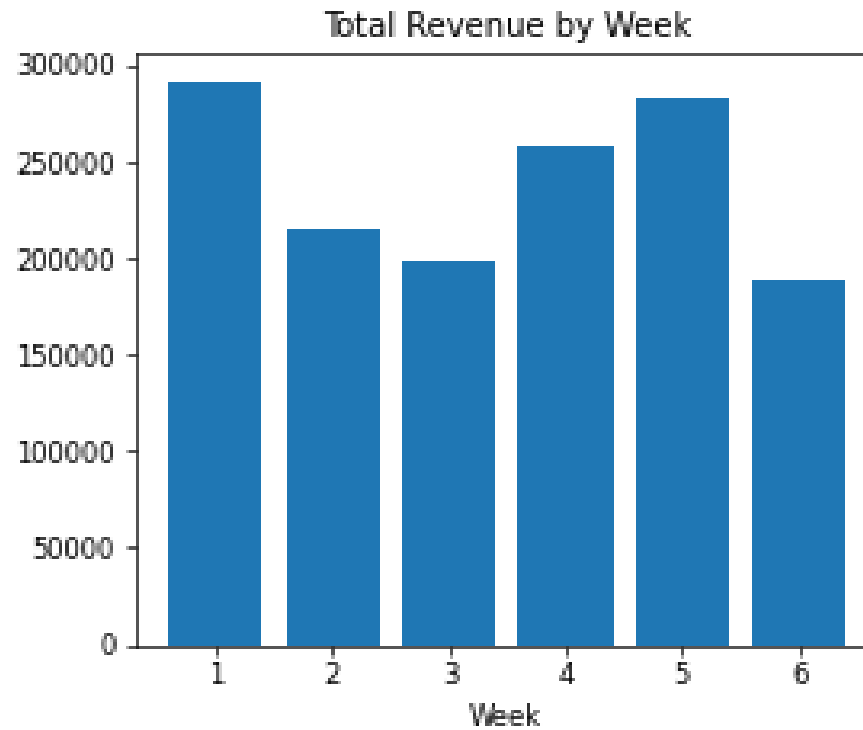
Dataframe includes 1500 rows and 8 columns – 5 continuous, 3 categorical.

Preprocess Dataset

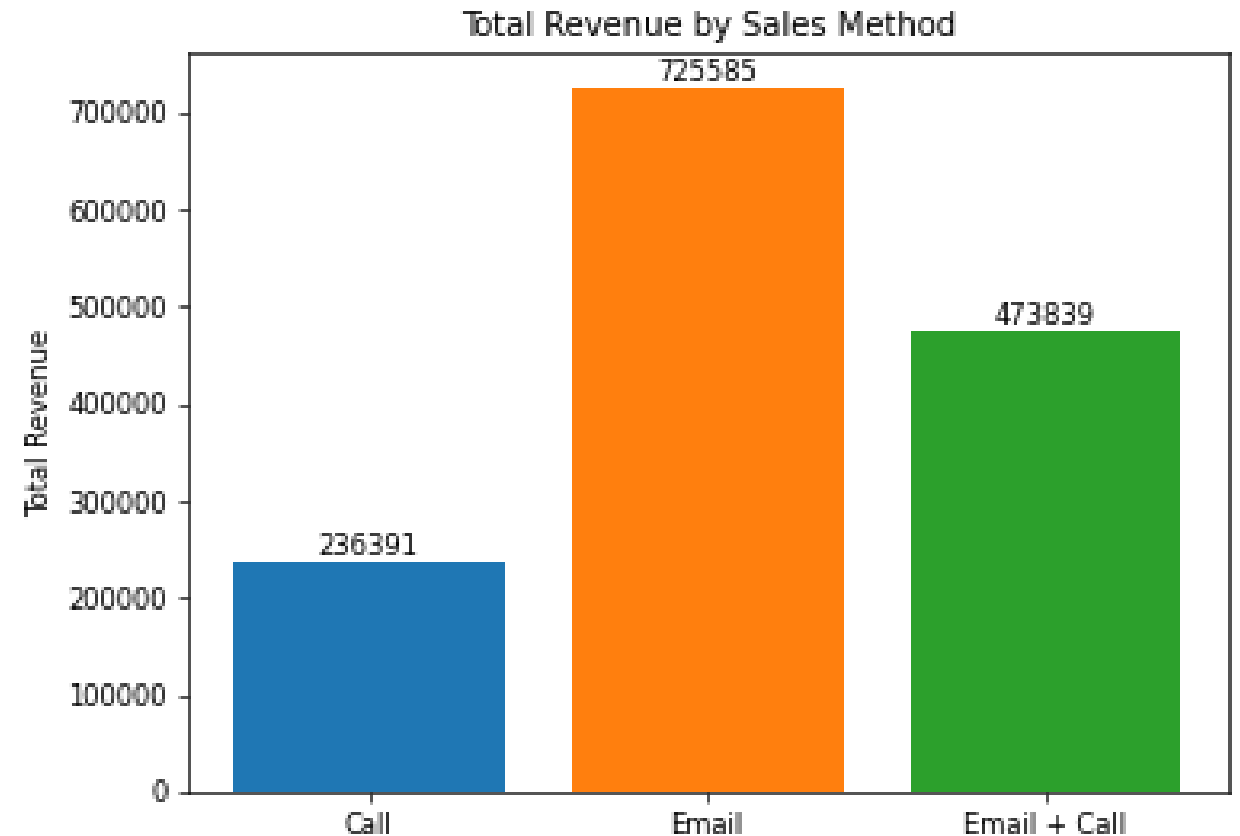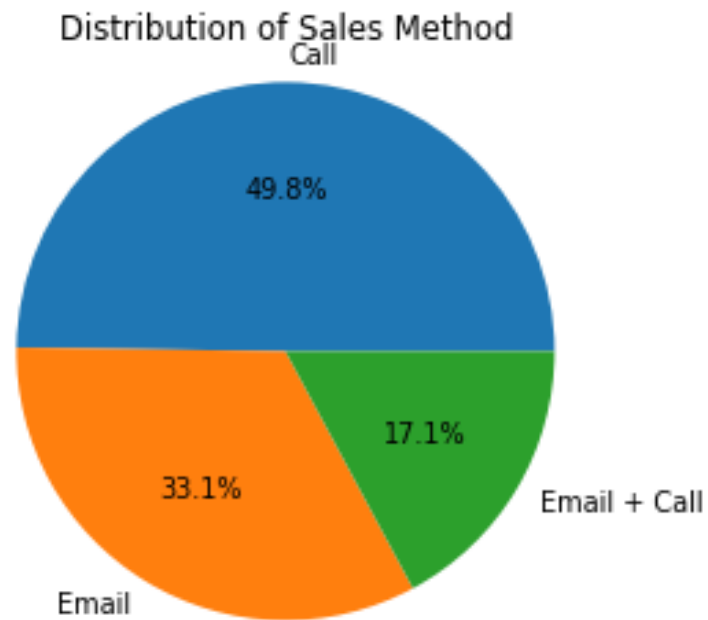| Columns	
week	Numerical . Ranges from 1-6.
sales_method	Should be 3 categories. Includes two extra categories.
customer_id	Unique text/number combination.
nb_sold	Numerical value. Ranges from 7-16. Average 10 items.
revenue	1074 rows missing. Replace using linear regression.
years_as_customer	Age of company 41 years. Remove years greater than 41.
nb_site_visits	Numerical value. Average ~25.
state	Convert to category type. All 50 states represented.

Sales data scrubbed to appropriately make prediction.
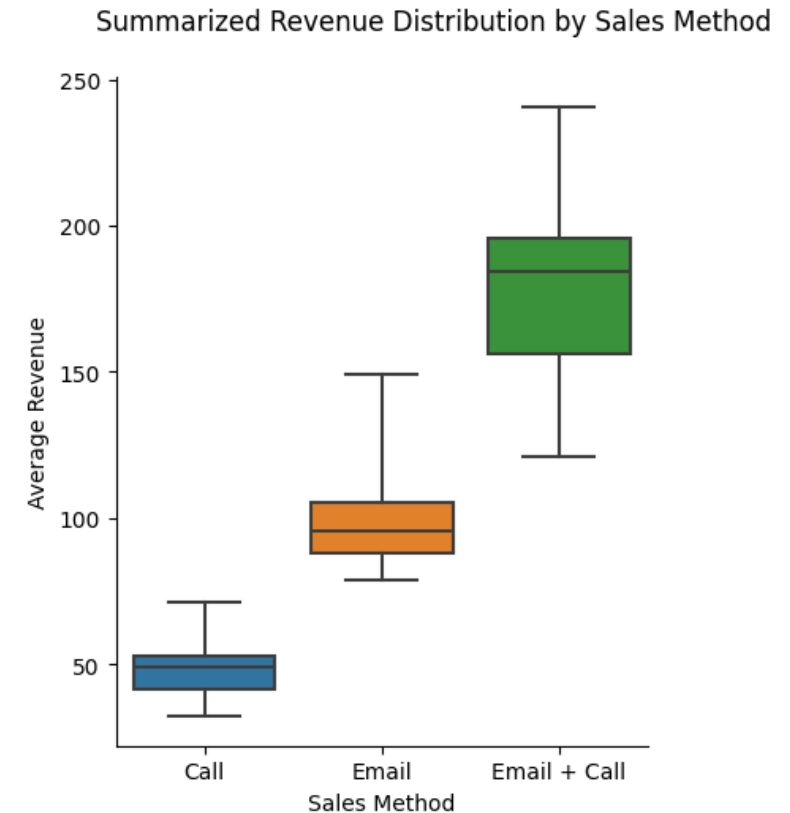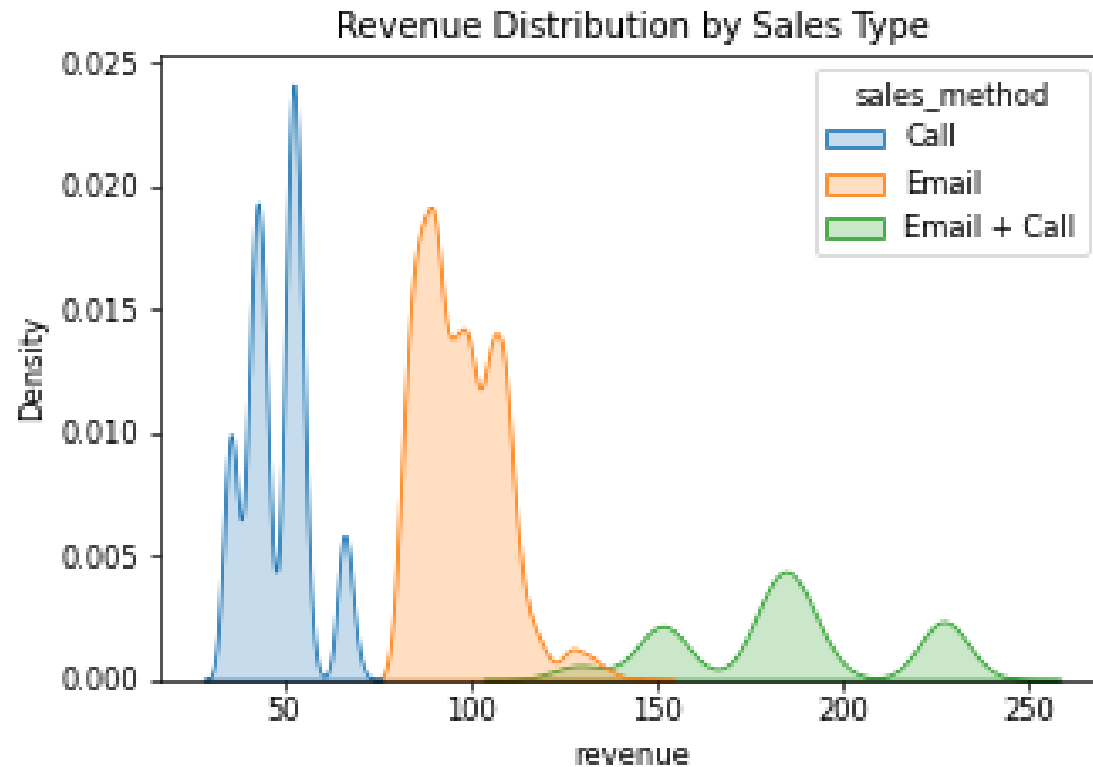
Revenue Over time

Average revenue 239k/week.
Revenue has lower revenue in last week.
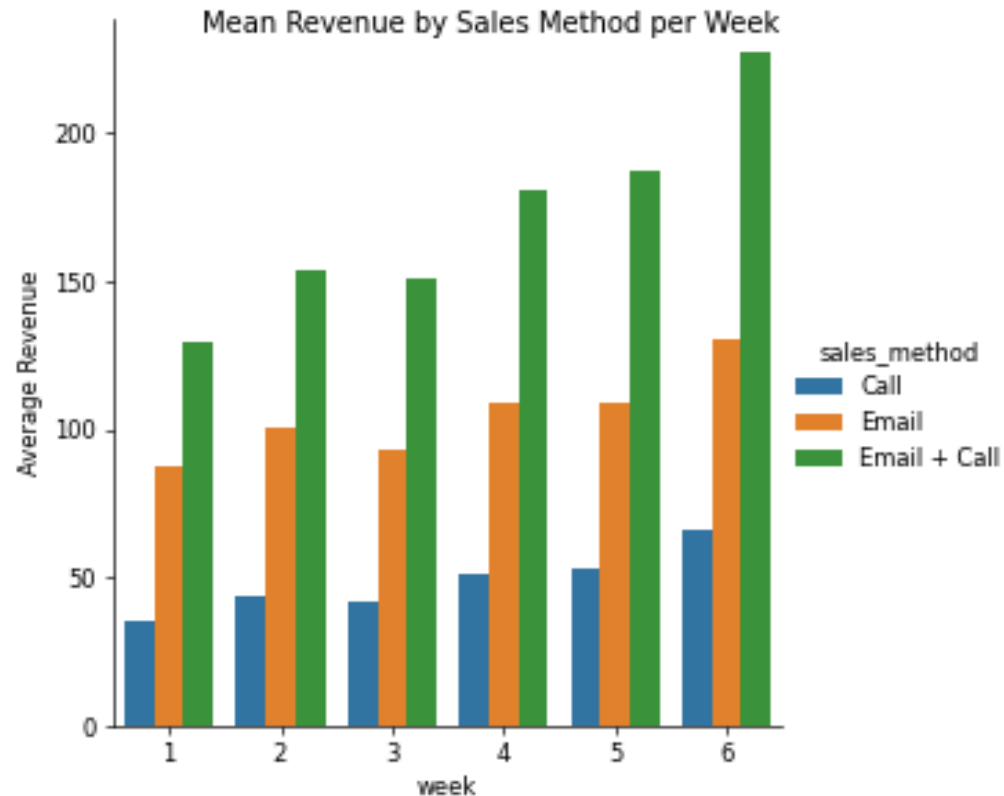
Revenue and Sales Method

Sales by 'Call' represents 50% overall but only 16% of revenue.

Revenue Distribution

The three different sales methods have different average revenues.

Metric to Monitor

Mean Revenue by Sales Method per Week

Sales Method	Average Revenue/Sale
Call	48.0
Email	97.0
Email + Call	184.0

Monitor average revenue per sale.

Summary & Recommendation

- Revenue of new product line has positive trend.

- Sales method affects revenue
  - Most revenue is generated by Email.

- Drop calling as sales method
  - Focus on Email and Email + Call

- Monitor average revenue per sale across weeks by sales method
  - Recommend target of $100/sale.