TECHNICAL ANALYSIS: ANALYSIS OF RETAIL DATASET.

<u>INTRODUCTION</u>

Businesses generally produce a large amount of data through their daily operations. Retail businesses based on their scale generate data from sales, expenditures, and other daily operations which when recorded provides the business on insight and steps that should be taken in other to improve the business. Examining this data aids in comprehending customer behavior, sales performance, and inventory management. This specific given dataset includes sales data, featuring details like order dates, product codes, sales amounts, and customer information. This report seeks to derive insights from this dataset.

<u>DATASET OVERVIEW:</u>

Dataset description

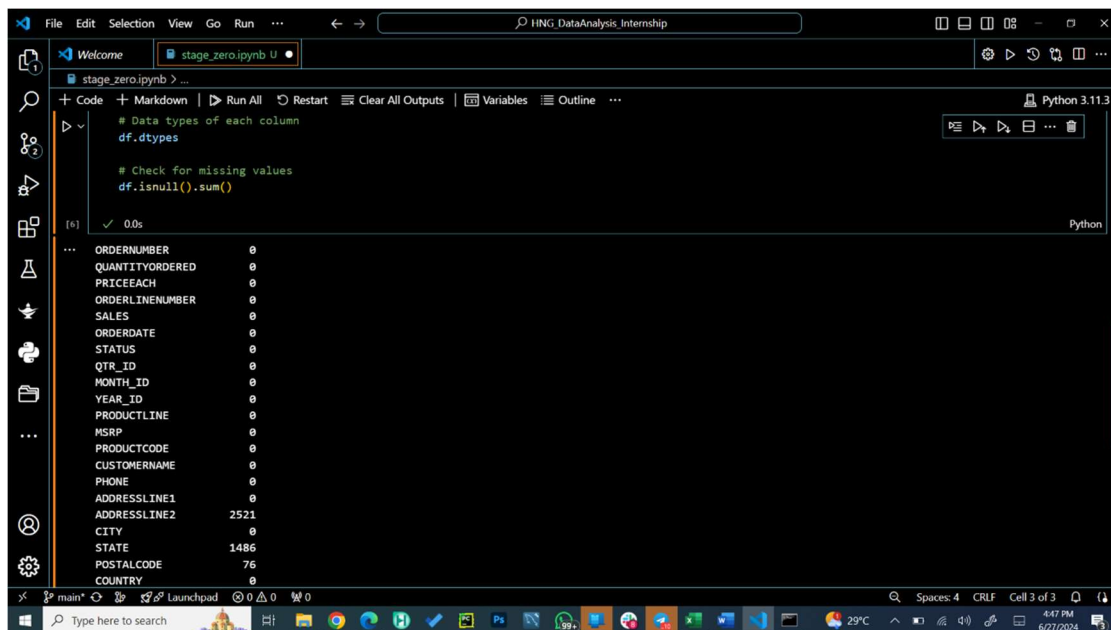The dataset description and their datatype :

- ORDERNUMBER: Order identification number (numerical)
- QUANTITYORDERED: Number of items ordered (numerical)
- PRICEEACH: Price of each item (numerical)
- ORDERLINENUMBER: Order line number
- SALES: Total sales amount (numerical)
- ORDERDATE: Date of order (Date)
- STATUS: Order status (categorical)
- QTR_ID: Quarter identifier (numerical)
- MONTH_ID: Month identifier (Date)
- YEAR_ID: Year identifier (Date)
- PRODUCTLINE: Product line category (categorical)
- MSRP: Manufacturer's suggested retail price (numerical)
- PRODUCTCODE: Product code (string)
- CUSTOMERNAME: Name of the customer (string)
- PHONE: Customer phone number (numerical)
- ADDRESSLINE1: Address 1 (string)

- ADDRESSLINE2: Address 2 (string)

- CITY: Customer city (string)

- STATE: Customer state (string)

- POSTALCODE: postal code (numerical)

- COUNTRY: country (string)

- TERRITORY: Sales territory (categorical)

- CONTACTLASTNAME: Contact person's last name (string)

- CONTACTFIRSTNAME: Contact person's first name (string)

- DEALSIZE: Size of the deal (categorical)

OBSERVATIONS

From the dataset with 2823 entries for several years, quarters, and months, the dataset offers a thorough understanding of sales over time. Some missing data in the [Address line 2] because its optional , [postal code] and [state] column. These missing figures, however, will not significantly affect the overall sales analysis or make it more difficult to draw insights from the dataset.

MISSING VALUES

Figure 1: python code for checking missing values


```python
import matplotlib.pyplot as plt

# Group by year and month
monthly_sales = df.groupby(['YEAR', 'MONTH'])['SALES'].sum().reset_index()

# Pivot for better plotting
monthly_sales_pivot = monthly_sales.pivot(index='MONTH', columns='YEAR', values='SALES')

# Plotting monthly sales trend
monthly_sales_pivot.plot(figsize=(12, 8))
plt.title('Monthly Sales Trend')
plt.xlabel('Month')
plt.ylabel('Sales')
plt.show()
```

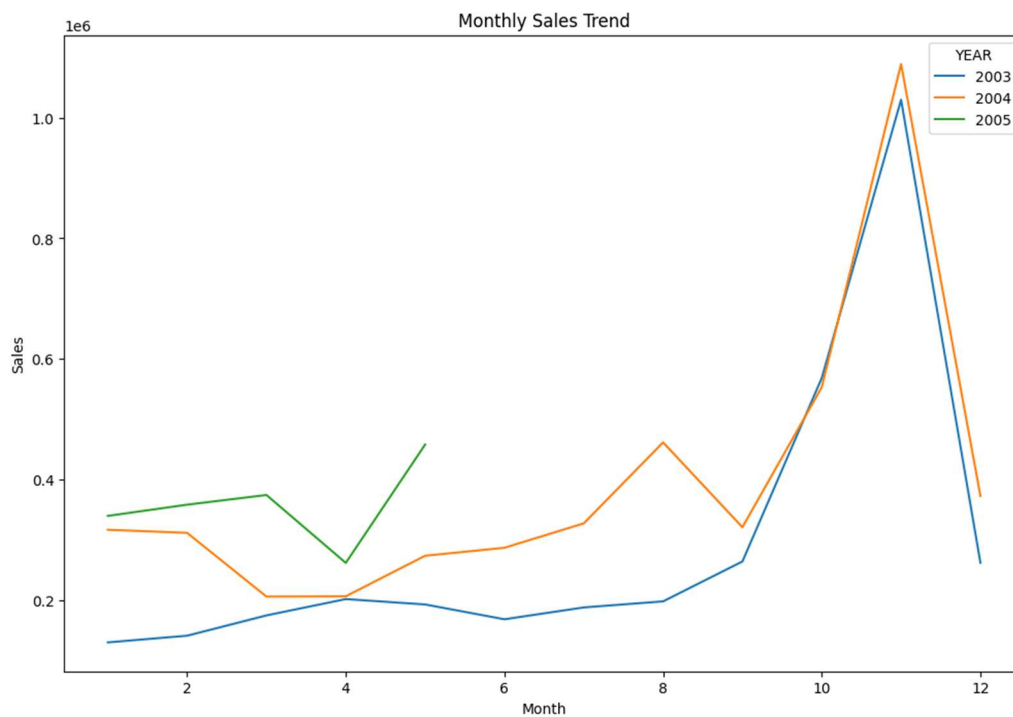Figure 2 python code for graph for monthly sales trend

SALES TREND

Figure 3 Line graph for monthly sales trend

A line graph depicts a monthly sales trend, though the specific months and units of measurement are not provided. The graph shows a trend, suggesting increasing sales over the time frame shown.

Figure 4: python code for simple bar chart shows sales based on type of products.
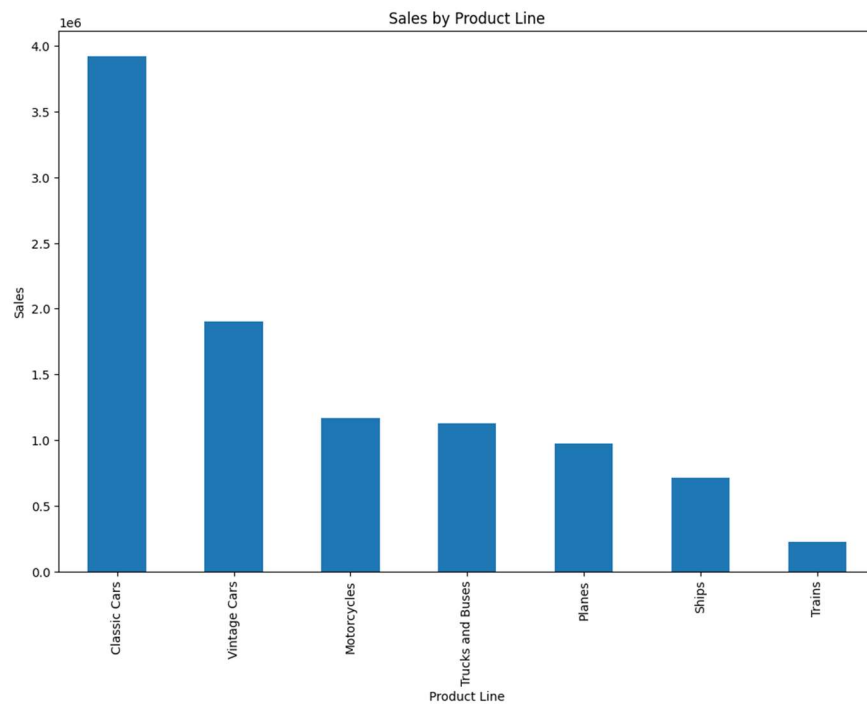
PRODUCT SALE TREND

Figure 5: Bar chart shows sales based on type of products

CONCLUSION

Retail companies generate a lot of data from their regular business dealings. Analyzing this data makes it easier to understand sales results, inventory, and consumer behavior. The given dataset includes sales data, featuring details like order dates, product codes, sales amounts, and customer information. This report seeks to derive insights from this dataset.

The scope of this report covers an examination of a retail sales dataset for insights concerning the trend in sales, product line performance, and customer contribution that would assist company make smart choices.