XYZ ENTERPRIES SALES ANALYSIS

Project Overview: Sales Data Analysis (2014–2018)

This project focused on analyzing a comprehensive sales dataset comprising approximately 64104 records spanning from January 2014 to February 2018. The dataset was provided in six separate Excel sheets.

The initial phase involved data cleaning and preparation using Python in Jupyter Notebook. Tasks included importing multiple sheets, handling null values, merging datasets, and removing duplicate or redundant columns to ensure data integrity.

Following Python-based cleaning, the dataset was exported to Excel for additional data enrichment, such as creating calculated columns and organizing the data for ease of analysis.

The final dataset was then imported into Power BI, where interactive dashboards and visualizations were developed to analyze sales trends, identify key insights, and support data-driven decision-making.

Problem Statement:

The sales data provided was extensive, spanning multiple years and stored across six separate Excel sheets, making it difficult for management and board members to quickly understand sales performance for specific dates, products, or regions. This lack of clarity posed challenges in tracking trends, identifying key insights, and making timely, informed decisions during strategic meetings.

Tools Used:

- Python Pandas for data cleaning
- Excel for data enrichment
- Power BI for Data visualization
- Chat GPT for various uses

Process:

1) Got sales data of xyz on excel format as' Regional Sales Data Analysis', file got five sheets Sales Order, Customers, Regions, State Regions, Products and 2017 Budgets. Imported it to jupyter notebook for data cleaning

```
L]: import pandas as pd
    import numpy as np
```

```
2]: all_files=pd.read_excel('Regional sales Data analysis.xlsx',sheet_name=None)
```

2) Assigned unique Data Frame names to each sheet for organized data processing and efficient analysis.

```
df_sales=all_files['Sales Orders']
df_customers=all_files['Customers']
df_regions=all_files['Regions']
df_state=all_files['State Regions']
df_products=all_files['Products']
df_budget=all_files['2017 Budgets']
```

3) A detailed review of the Excel file revealed that the individual sheets were interconnected and contained complementary information. Therefore, the datasets were merged into a single consolidated table to support efficient analysis and meaningful business insights.

At first 'Sales Orders' and 'Customers' sheets that was given name as df_sales and df_customers were merged by left join where 'customer name index' and 'customer index' have same customer reference data

```
df=df_sales.merge(df_customers,how='left',left_on='Customer Name Index',right_on='Customer Index')
```

4) The previously merged dataset was further combined using a left join based on the *Product Description* fields. The join was performed between the Product Description Index and Index columns, as both contained identical state reference information.

```
df=df.merge(df_products,how='left',left_on='Product Description Index',right_on='Index')
```

5) The previously merged dataset was further combined using a left join based on the *Delivery Index* fields. The join was performed between the Delivery Region Index and id columns, as both contained identical Delivery reference information.

```
df=df.merge(df_regions,how='left',left_on='Delivery Region Index',right_on='id')
```

6) The previously merged dataset was further combined using a left join based on the *State Code* fields. The join was performed between the state_code and State Code columns, as both contained identical state reference information.

```
df=df.merge(df_state,how='left',left_on='state_code',right_on='State Code')
```

7) Redundant columns created during merging were identified and removed, retaining only one representative column.

```
unwanted_columns=['Customer Index','Index','id','State Code','State 1']
df=df.drop(columns=unwanted_columns,errors='ignore')
```

8)Gone through '2017 Budgets' sheet and found it contain one column 'Product Name' that was also in main sheet formed after merging previous data sheets. Change name of column to 'Product Name2', then merged it using left join base on Product Name. The join was performed between the 'Product Name' and

'Product Name2', as both contained identical Product reference information. Then redundant column 'Product Name2' was deleted

```
df_budget=df_budget.rename(columns={'Product Name' :'Product Name2'})
```

```
df=df.merge(df_budget,how='left',left_on='Product Name',right_on='Product Name2')
```

```
df=df.drop(columns='Product Name2')
```

9) Changed the name of columns to small case and removed spaces between words by '_'

```
df.columns=df.columns.str.lower().str.replace(' ','_')
```

```
df.columns
```

```
Index(['ordernumber', 'orderdate', 'customer_name_index', 'channel',
       'currency_code', 'warehouse_code', 'delivery_region_index',
       'product_description_index', 'order_quantity', 'unit_price',
       'line_total', 'total_unit_cost', 'customer_names', 'product_name',
       'name', 'county', 'state_code', 'state', 'type', 'latitude',
       'longitude', 'area_code', 'population', 'households', 'median_income',
       'land_area', 'water_area', 'time_zone', 'region', '2017_budgets'],
      dtype='object')
```

10) Budget values were applied only to 2017 records; for all other years, the budget field was set to blank.

```
df['budget'] = np.where(
    df['orderdate'].dt.year == 2017,
    df['budget'],
    np.nan
)
```

11) Cleaned data was saved back to excel file

```
]: df.to_excel('cleaned_data .xlsx')
```

12) Using power query I had added 3 columns 'Total_cost','Profit','Profit_margin(%), that will make analysis easier and deleted unrelated columns

13) Cleaned data were imported to power BI and dashboard was prepared

Key Findings:

1) Overall Business Performance:

- Total Revenue = 1.24 Billions

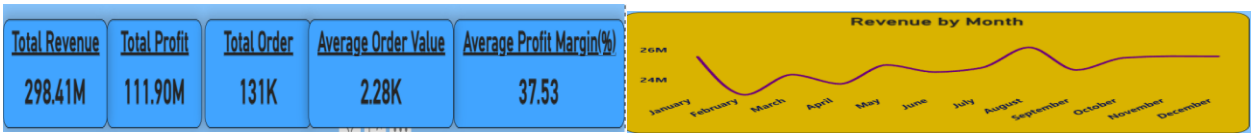
- Total Profit = 461.77
- Total orders = 541K
- Average Order Value(AOV)=2.28K
- Average Profit Margin= 37.36%

Business seems in healthy position with attractive profit margin, but as the company was focus on Wholesale, Distributor, and Export Business Average Order Value(AOV) of 2.28K is not so high, there is lot of room for improvement to increase it.

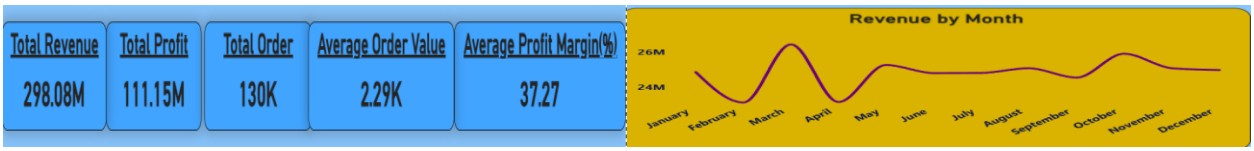
2) Sales Trend:

If we look at overall monthly trends it shows in Jan/Feb sales normally pick but if we go by monthly trends on yearly basis, it shows sales are flat no improvement in any of the KPI included above, all the indicators were flat2014:

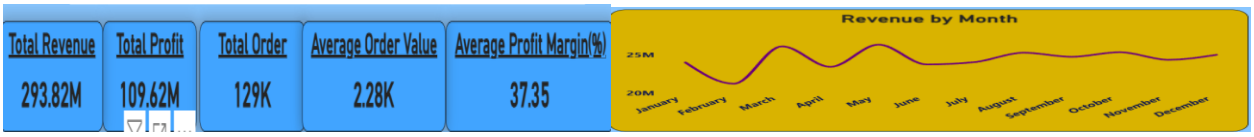
2015:



2016:



2017:



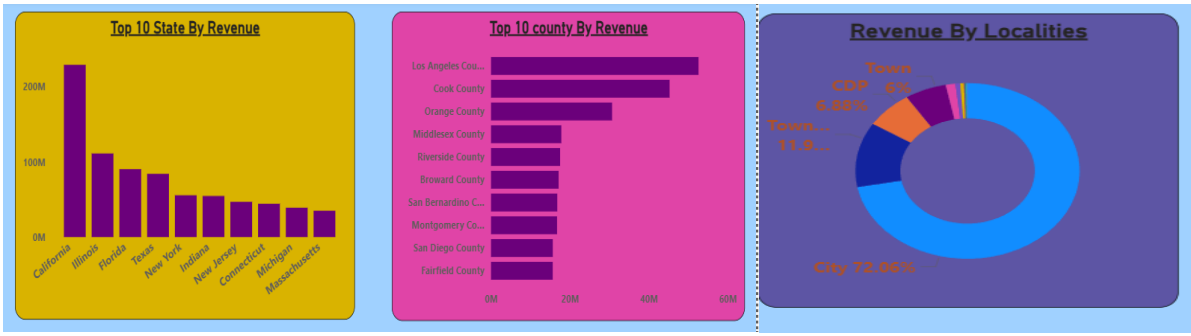
2018

2018 contain data of first 2 months only, It's Revenue, AOV, Average Profit Margin were same as previous year.

Marketing &Sales Department need proactive actions and need to add professional manpower to boost sales as margin was hefty, can run promotion campaign and discounts also to acquire customers.

3) Geographic Performance:

Sales were driven highly by states like California, Illinois, Florida and Texas and county like Los Angeles, Cook County, Orange County, Middlesex County, and cities carried 72.06% of total revenue, which means there is plenty of room to increase revenue by focusing on other localities, states and county



Run Promotional campaign and discounts to the suppliers of states and county that were contributing less in the revenue, we can give few extra margin to the distributors focusing on those areas, as we have healthy margin

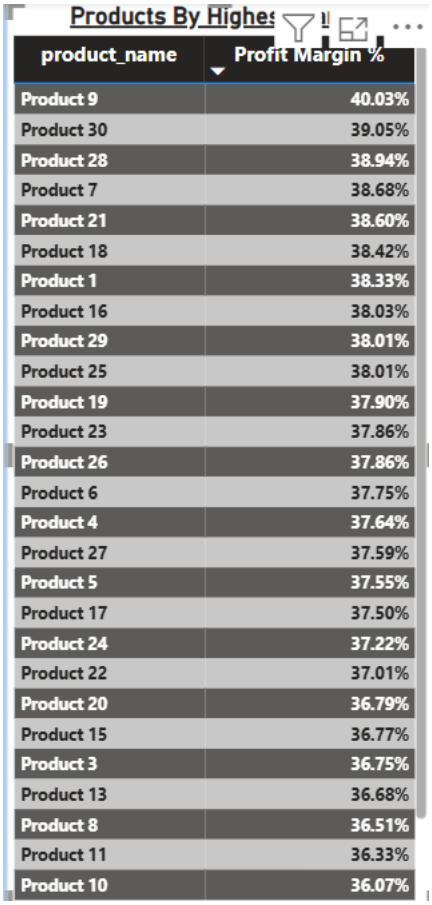
4) Channel & Operational Insights:

warehouses AXW291, GUT930 contribute high in revenue then it's obviously it contributes high on profit also. NXH382 is not far behind but there were lot of room for improvement for FLR025 to improve. Wholesale were contributing maximum for revenue and profit but export was disappointing

Suggestions to HR, hire professional to look overseas sales as it was flat from 4 years and no sign of improvement, there were plenty of room for increase export

5) Product performance:

Overall revenue was driven by few products, states and county. Almost all products contribute healthy profit margin



product_name	Profit Margin %
Product 9	40.03%
Product 30	39.05%
Product 28	38.94%
Product 7	38.68%
Product 21	38.60%
Product 18	38.42%
Product 1	38.33%
Product 16	38.03%
Product 29	38.01%
Product 25	38.01%
Product 19	37.90%
Product 23	37.86%
Product 26	37.86%
Product 6	37.75%
Product 4	37.64%
Product 27	37.59%
Product 5	37.55%
Product 17	37.50%
Product 24	37.22%
Product 22	37.01%
Product 20	36.79%
Product 15	36.77%
Product 3	36.75%
Product 13	36.68%
Product 8	36.51%
Product 11	36.33%
Product 10	36.07%

Conclusions:

Revenue, profits, profit margin was in excellent conditions but limited to few localities and products only, so It is recommended to hire professional, competent sales persons that will focus on export market, and run promotional discount and bundles offers to other localities to expand the market and acquire new customers