End of course project

About the Dataset

Data source

https://www.kaggle.com/datasets/knightbearr/sales-product-data?select=Sales_August_2019.csv

Context

Sales analytics is the practice of generating insights from sales data, trends, and metrics to set targets and forecast future sales performance. Sales analysis is mining your data to evaluate the performance of your sales team against its goals. It provides insights about the top performing and underperforming products/services, the problems in selling and market opportunities, sales forecasting, and sales activities that generate revenue.

Content

- Order ID - An Order ID is the number system that Amazon uses exclusively to keep track of orders. Each order receives its own Order ID that will not be duplicated. This number can be useful to the seller when attempting to find out certain details about an order such as shipment date or status.
- Product - The product that has been sold.
- Quantity Ordered - Ordered Quantity is the total item quantity ordered in the initial order (without any changes).
- Price Each - The price of each product.
- Order Date - This is the date the customer is requesting the order be shipped.
- Purchase Address - The purchase order is prepared by the buyer, often through a purchasing department. The purchase order, or PO, usually includes a PO number, which is useful in matching shipments with purchases; a shipping date; billing address; shipping address; and the requested items, quantities, and price.

Tasks:

Step 1: Create a copy of the January and February data sets.

Step 2: Data Cleaning:

1. Explore the January data sales and remove blank columns
2. Separate the date and time into separate columns
3. Separate the purchase address into PO Box, City, and address
4. Standardize the columns to the correct number or text formats

Step 3: Creating a macro:
Create a macro that applies all the steps done on the January copy data to the February copy data.

Step 4: Use Power Query to combine and transform all datasets into one.

Step 5: Data exploration

1. Total Revenue
2. Total Number of orders
3. Total units sold
4. Average revenue per month

5. Average number of orders per month
6. Average units sold per month
7. Average daily units sold
8. Average daily revenue
9. What was the best month for sales? How much was earned that month?
10. What city had the highest number of sales? How much was it?
11. What product sold the most? Why do you think it sold the most?

BONUS QUESTIONS

12. Sales by Day of the Week: Analyze sales performance by day of the week and identify the most and least profitable days.
13. Sales by Time of Day: Explore sales based on the time of day and determine the peak hours.
14. Cross-Selling Opportunities: Identify product combinations frequently purchased together.

Step 6: Creating a dashboard

1. Create a dashboard in excel showing the sales for the month of January using pivot tables and pivot charts.
2. Import the combined data into power bi and create a dashboard for the sales, including necessary slicers and visualizations.