**SWIFT MART SALES DATA ANALYSIS**

[Introduction 2](#_Toc52795747)

[Data Description & Data Dictionary 2](#_Toc1171764132)

[Data Dictionary 3](#_Toc606637960)

[Data Description 3](#_Toc2010047090)

[Objective of the Project 6](#_Toc93067757)

[Client Requirement and Suggestions 6](#_Toc328174457)

[Project Outline 6](#_Toc429886893)

## Introduction

We have access to a comprehensive sales dataset from the renowned e-commerce site in United States, "Swift mart." This data set comprises 24 variables and encompasses 9,994 rows, providing a wealth of information regarding **sales,** **profitability,** and **discounts** associated with each **product**.

Within the dataset, we have valuable details such as the **sub-category** and **category** of each product, enabling us to discern patterns and trends across different product lines. Additionally, we have information on the geographical aspects of each order, including the **location of purchase**, **order date**, **order ID**, **ship date**, and **ship mode** used to deliver the products to customers.

To ensure personalized service, Swift mart has recorded the **names of customers,** allowing for individualized interactions and tailored marketing approaches. The data set also includes the information on **segment type of customers.**

**The primary objective of analyzing this data set is to assist Swift mart in optimizing their sales and profitability.** Specifically, they aim to identify which products would benefit from discontinuing discounts and, conversely, pinpoint products that could potentially generate higher sales and profit by increasing the offered discounts. By understanding the relationship between discounts, sales, and profitability, Swift mart can strategize effectively and maximize their overall revenue.

Through comprehensive data analysis techniques and employing statistical modeling, we can uncover meaningful insights from this dataset. These insights will guide Swift mart in making informed decisions regarding discount strategies, thereby driving sales growth and ultimately maximizing their profit.

## Data Description & Data Dictionary

There are 24 variables and 9994 observations.

### Data Dictionary

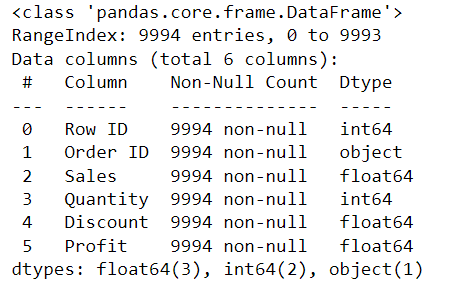
|  |  |  |
| --- | --- | --- |
| **Variable** | **Data Type** | **Description** |
| Row ID | int | Unique identifier for each row |
| Order ID | object | Unique identifier for each order |
| Order Date | date time | Date of the order |
| Ship Date | date time | Date of shipping |
| Ship Mode | object | Shipping mode for the order |
| Customer ID | object | Unique identifier for each customer |
| Country/Region | object | Country or region of the customer |
| Postal Code | float | Postal code associated with the customer's address |
| Product ID | object | Unique identifier for each product |
| Customer Name | object | Name of the customer |
| Segment | object | Segment/category to which the customer belongs |
| City | object | City where the customer is located |
| State | object | State where the customer is located |
| Region | object | Region where the customer is located |
| Category | object | Category to which the product belongs |
| Sub-Category | object | Sub-category to which the product belongs |
| Product Name | object | Name of the product |
| Sales | float | Sales amount for the product |
| Quantity | int | Quantity of products sold |
| Discount | float | Discount applied to the product |
| Profit | float | Profit earned from the product |

### Data Description

We have divided the dataset into two parts which are sales data and order data.

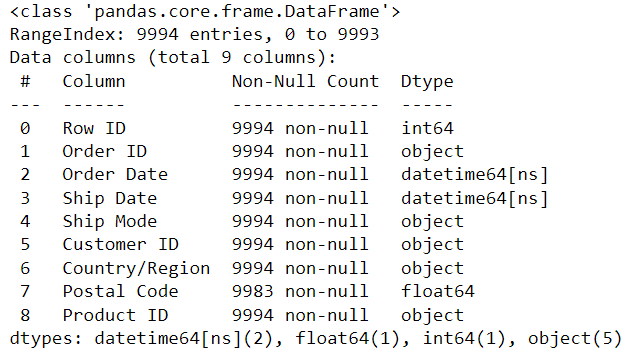
Further we have divided order data in to 4 parts. Orders, Location, Product and Customer

**Sales data**

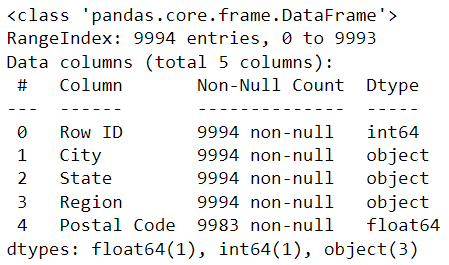


**Order data**

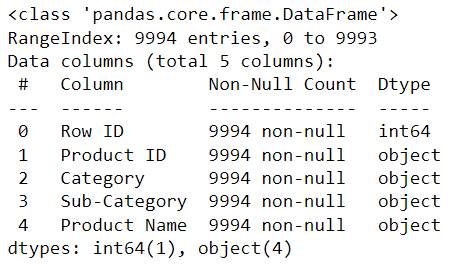
**Orders sheet**



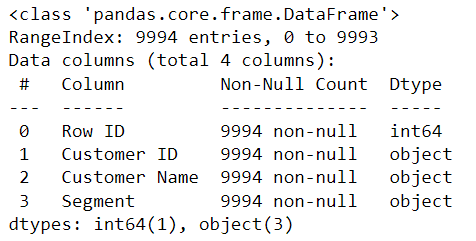
**Location sheet**



**Product sheet**



**Customer sheet**



Objective of the Project

To give an overview of the key sales metrics, analyze customer data and product data to find out profitable products and segments and decide on discounts to maximize sales and profitability.

## Client Requirement and Suggestions

* Create Multiple dashboards
* Use sheet swapping
* Make it interactive
* Create a data story of the visuals
* Think and come up with ideas what can be good for the client which will be beneficial for them.
* Create subscription and alerts in Tableau server

## Project Outline

1. Understand the data
2. Cleaning the data
3. Create the BRD
4. Data connectivity
5. EDA
6. Identify the problem
7. Find out the insights by analyzing data
8. Create layouts for dashboard
9. Complete the dashboard fitting in the right chart at right order to convey the right message to client.
10. Present to client

## Project Planning

We will create 3 dashboards

1. Overview
   * Provide an overview of key metrics like total sales, total profit, and total orders.
   * Show summary statistics such as average sales, average profit, and average discount.
   * Include charts or tables to display sales and profit trends over time (e.g., line chart or bar chart).
   * Use filters to allow users to drill down into specific regions, categories, or sub-categories.
2. Customer Analysis
   * Analyze customer-related metrics such as customer loyalty, sales per customer, and customer segments.
   * Include a customer segmentation analysis to identify the most valuable customer segments.
   * Visualize customer demographics (e.g., city, state, region) using maps or charts.
   * Show customer-specific metrics like average order size, average discount per customer, or top customers by sales/profit.
   * Enable interaction with customer data by filtering based on customer name or ID
3. Product performance analysis
   * Analyze product-related metrics such as sales, profit, quantity sold, and discounts.
   * Show product categories and sub-categories using a tree map or bar chart.
   * Display top-selling products or best-performing categories using charts or tables.
   * Include a product performance analysis to identify high-profit or low-profit products.
   * Enable interaction with product data by filtering based on product ID or name.

We will create a sheet swapping functionality to allow users to switch between different views of customer data in Customer Analysis dashboard

* Start by creating multiple worksheets with different visualizations based on the customer data. For example:
* **Customer Segmentation**: A bar chart showing the distribution of customers across different segments.
* **Sales by Customer Demographics**: A map or chart displaying sales based on the customer's location (city, state, or region).
* **Top Customers**: A table or chart highlighting the top customers based on sales or profit.

## Challenges

When we are combining two datasets using relationship, there are limitations in using table calculations.

We were trying to write a calculated field