1. **Project Overview**

As part of this project, I have analysed the sales of a reputed superstore in USA and derived meaningful insights in-order to understand the strengths & weaknesses of the business. This document covers an overview of the dataset utilized and the various analytical insights that will be derived from it.

1. **Overview of the dataset**

It is critical to understand the column descriptions and the structure of dataset prior to deriving any insights.

|  |  |
| --- | --- |
| **ColumnName** | **Description** |
| Row ID | Unique ID for each row |
| Order ID | Unique Order ID for each Customer |
| Order Date | Order Date of the product |
| Ship Date | Shipping Date of the Product |
| Ship Mode | Shipping Mode specified by the Customer |
| Customer ID | Unique ID to identify each Customer |
| Customer Name | Name of the Customer |
| Segment | The segment where the Customer belongs |
| Country | Country of residence of the Customer |
| City | City of residence of of the Customer |
| State | State of residence of the Customer |
| Postal Code | Postal Code of every Customer |
| Region | Region where the Customer belong |
| Product ID | Unique ID of the Product |
| Category | Category of the product ordered |
| Sub-Category | Sub-Category of the product ordered |
| Product Name | Name of the Product |
| Sales | Sales of the Product |
| Quantity | Quantity of the Product |
| Discount | Discount provided |
| Profit | Profit/Loss incurred |

1. **List of insights to derive**

(As an initial step, import the csv file containing the data into SQL server table in-order to be able to run SQL queries)

* The available shipping modes for the order delivery
* Identity the key columns in the data set
* Check if there are duplicate records for the key columns. If yes, clean up the data set by remove duplicate rows for the key columns
* Identify the top performing states with respect to the profits
* Total sales for each state
* Which sub-categories have the most sales
* Average days it takes to ship a product for each state
* Identify the states with the quickest shipping