**DATA DESCRIPTION**

1. **Automobile Sales Dataset**

The dataset contains Sales data of an Automobile company. In the Automobile dataset, there exist 20 variables and 2747 observations. The dataset is saved as **“autosale.xlsx”**. The variable information is as follows:

|  |  |
| --- | --- |
| **Column Name** | **Description** |
| **ORDERNUMBER** | This column represents the unique identification number assigned to each order. |
| **QUANTITYORDERED** | It indicates the number of items ordered in each order. |
| **PRICEEACH** | This column specifies the price of each item in the order. |
| **ORDERLINENUMBER** | It represents the line number of each item within an order. |
| **SALES** | This column denotes the total sales amount for each order, which is calculated by multiplying the quantity ordered by the price of each item. |
| **ORDERDATE** | It denotes the date on which the order was placed. |
| **DAYS\_SINCE\_LASTORDER** | This column represents the number of days that have passed since the last order for each customer. It can be used to analyze customer purchasing patterns. |
| **STATUS** | It indicates the status of the order, such as "Shipped," "In Process," "Cancelled," "Disputed," "On Hold," or "Resolved." |
| **PRODUCTLINE** | This column specifies the product line categories to which each item belongs. |
| **MSRP** | It stands for Manufacturer's Suggested Retail Price and represents the suggested selling price for each item. |
| **PRODUCTCODE** | This column represents the unique code assigned to each product. |
| **CUSTOMERNAME** | It denotes the name of the customer who placed the order. |
| **PHONE** | This column contains the contact phone number for the customer. |
| **ADDRESSLINE1** | It represents the first line of the customer's address. |
| **CITY** | This column specifies the city where the customer is located. |
| **POSTALCODE** | It denotes the postal code or ZIP code associated with the customer's address. |
| **COUNTRY** | This column indicates the country where the customer is located. |
| **CONTACTLASTNAME** | It represents the last name of the contact person associated with the customer. |
| **CONTACTFIRSTNAME** | This column denotes the first name of the contact person associated with the customer. |
| **DEALSIZE** | It indicates the size of the deal or order, which are the categories "Small," "Medium," or "Large." |

1. **Global Country Information**

This comprehensive dataset provides a wealth of information about **all countries worldwide**, covering a wide range of indicators and attributes. It encompasses demographic statistics, economic indicators, and much more. In the dataset, there exist 14 variables and 195 observations. The dataset is saved as **“world-data-2023.xlsx”**. The variable information is as follows:

|  |  |
| --- | --- |
| **Column Name** | **Description** |
| **Country** | Name of the country. |
| **Latitude** | Latitude coordinate of the country's location |
| **Longitude** | Longitude coordinate of the country's location |
| **Birth Rate** | Number of births per 1,000 population per year |
| **CO2-Emission** | Carbon dioxide emissions in tons |
| **CPI** | Consumer Price Index, a measure of inflation and purchasing power. |
| **Gasoline Price** | Price of gasoline per liter in local currency |
| **GDP** | Gross Domestic Product, the total value of goods and services produced in the country |
| **Life expectancy** | Average number of years a newborn is expected to live |
| **Population** | Total population of the country |
| **Tax Revenue(%)** | Tax revenue as a percentage of GDP |
| **Total tax rate** | Overall tax burden as a percentage of commercial profits. |
| **Unemployment rate** | Percentage of the labor force that is unemployed |
| **Urban\_population** | Percentage of the population living in urban areas |