**Report on**

**Sales Analysis on supermarket grocery dataset**

CodeClause Data Science Internship – September 2022

Dataset report

The dataset used for experiments is totally fictional and is NOT owned by and collected from any original establishment of business [1]. The dataset represents the region-wise sales of grocery products in Tamil Nādu state, India. An individual sale is separated by a unique order ID where each sale is categorized in major category and later in subcategory. The dataset has quantified columns – sales, discount, and profit. Sales column represents the amount of sales, discount column represents the rate of discount, and profit is the amount earned as a profit per sale. An associated date of sale is also recorded. Following is the systematic report on the grocery sales dataset.

Dataset cleaning policy:

* Order ID and customer name are not considered for exploratory data analysis as well as for further procedures because it has no apparent impact on sales. Hence, removed from dataset.
* The whole data is collected for only single State – Tamil Nādu. Hence, the state column is removed because all rows contain only single variation of state name.
* The quantified columns like sales amount, discount, and profit amount are only considered till 2 decimal places. No further precision is considered.

**A. Basic Information about dataset**

Similarly, you can run *basic\_information\_utility.py* to print following information on console

**Unique major categories of sales**

|  |  |  |  |
| --- | --- | --- | --- |
| Eggs, Meat & Fish | Snacks | Beverages | Food Grains |
| Oil & Masala | Bakery | Fruits & Veggies |  |

**Unique subcategories of sales**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Category** | **Subcategory products** | | | |
| **Eggs, Meat & Fish** | Eggs | Mutton | Chicken | Fish |
| **Snacks** | Chocolate | Noodles | Cookies |  |
| **Beverages** | Health Drinks | Soft Drinks |  |  |
| **Food Grains** | Organic Staples | Atta & Flour | Dals & Pulses | Rice |
| **Oil & Masala** | Masala | Edible Oil & Ghee | Spices |  |
| **Bakery** | Breads & Buns | Biscuits | Cakes |  |
| **Fruits & Veggies** | Fresh Vegetable | Organic Fruits | Organic Vegetable | Fresh Fruits |

**Unique regions of sales**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Central | East | West | South | North |

**Other information**

|  |  |
| --- | --- |
| **Total number of records:** | 9994 (No Null values) |
| **The range of dates between which all sales took place:** | 03/01/2015 to 30/12/2018 |
| **The maximum price of sales:** | 2500.00 |
| **The minimum price of sales:** | 500.00 |
| **The minimum discount:** | 10% |
| **The maximum discount:** | 35% |
| **The minimum profit:** | 25.00 |
| **The maximum profit:** | 1120.00 |