Comprehensive Analysis of Supermarket Sales Trends: Insights and Visualizations

Name : Abhishek Buneti **Student ID : 23080067**

Email ID : ab24aft@herts.ac.uk **Link : Git Hub**

The supermarket sales data enables an understanding of branches, cities, and product lines, including unit price, quantity, payment type, and customer demographics. Based on the report, statistical summarisation and visualisation techniques involving pie plots, scatter plots, violin plots, and correlation heatmaps help to identify the most significant sale trends across different branches and product lines. This process helps to understand whether the dataset has been imported properly in the Jupyter Notebook, which helps to conduct further research.

	Branch	City	Customer type	Gender	Product line	Unit price	Quantity	Tax 5%	Total	Date	Payment	cogs	gross margin percentage	gross income	Rating
count	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000	1000.000000	1000.00000	1.000000e+03	1000.000000	1000.00000
mean	0.988000	1.008000	0.499000	0.499000	2.452000	55.672130	5.510000	15.379369	322.966749	2019-02-14 00:05:45.600000	1.001000	307.58738	4.761905e+00	15.379369	6.97270
min	0.000000	0.000000	0.000000	0.000000	0.000000	10.080000	1.000000	0.508500	10.678500	2019-01-01 00:00:00	0.000000	10.17000	4.761905e+00	0.508500	4.00000
25%	0.000000	0.000000	0.000000	0.000000	1.000000	32.875000	3.000000	5.924875	124.422375	2019-01-24 00:00:00	0.000000	118.49750	4.761905e+00	5.924875	5.50000
50%	1.000000	1.000000	0.000000	0.000000	2.000000	55.230000	5.000000	12.088000	253.848000	2019-02-13 00:00:00	1.000000	241.76000	4.761905e+00	12.088000	7.00000
75%	2.000000	2.000000	1.000000	1.000000	4.000000	77.935000	8.000000	22.445250	471.350250	2019-03-08 00:00:00	2.000000	448.90500	4.761905e+00	22.445250	8.50000
max	2.000000	2.000000	1.000000	1.000000	5.000000	99.960000	10.000000	49.650000	1042.650000	2019-03-30 00:00:00	2.000000	993.00000	4.761905e+00	49.650000	10.00000
std	0.817633	0.820127	0.500249	0.500249	1.715412	26.494628	2.923431	11.708825	245.885335	NaN	0.830475	234.17651	6.131498e-14	11.708825	1.71858

The descriptive statistics of key sales metrics of supermarkets involve an average unit price of $55.67 and total sales of about $322.97. Such analysis helps explore price trends and customer behaviour, helping one make informed decisions about inventory matters. Understanding sales dynamics enhances strategies as revenue and customer satisfaction are boosted significantly.

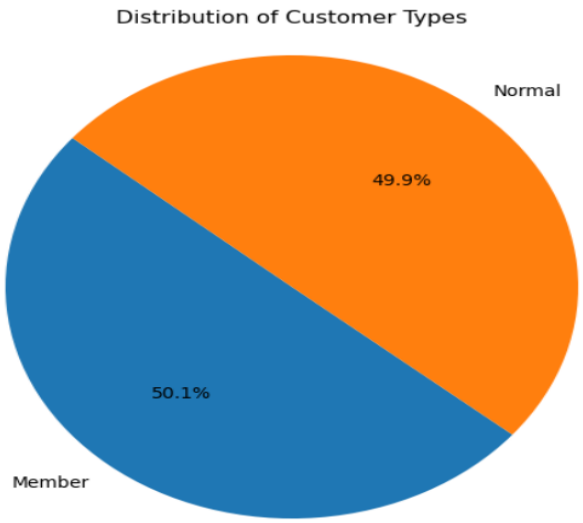

Figure 1：Customer type distribution

The pie chart in ***Figure 1*** shows a nearly equal distribution of customers, with Members at 50.1% and Normal at 49.9%. The pie chart helps to explore the distribution of different features with the percentage, offering in-depth visualisation (Roberts and Laramee, 2018). This enables customer marketing campaigns and promotions that would drive loyalty among members while driving more normal customers. With the help of the figure, it can be seen that the supermarket offers its services to all kinds of customers, which helps to expand its business and customer base properly, leading to higher sales and profits

.

Figure 2: Unit Price vs. Total Sales

The scatter plot for Unit price versus Total sales is quite compelling because the two variables are positively correlated, implying that if unit prices rise, total sales tend to increase as the supermarket might provide good quality products, which shows that even if the price is high, customers are buying the products. A scatter plot is the most efficient and widely used procedure or has a high degree of overlap, which offers a significant portion of data values (Nguyen *et al.*, 2020). ***Figure 2*** shows excellent value for strategy setting in terms of prices so businesses can optimise such strategies to obtain higher revenue despite customer interest in existing products.

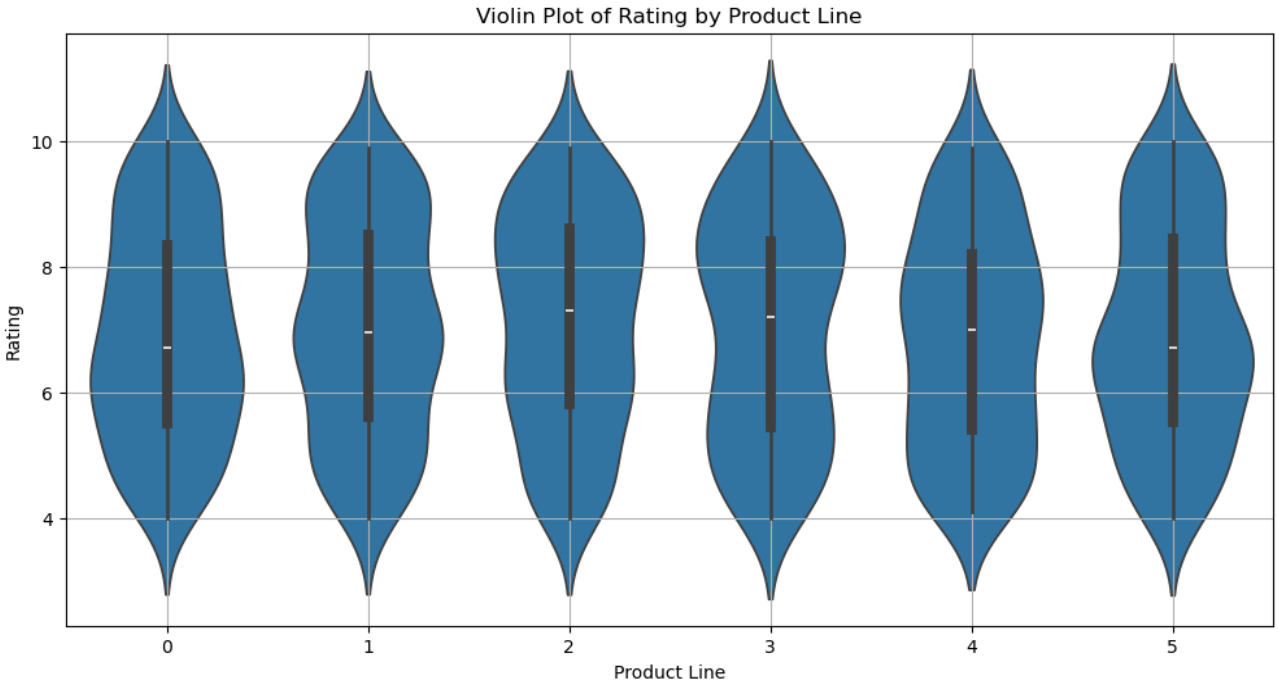

Figure 3: Rating by Product Line

The violin plot of rating by product line also gives an excellent view of how customers feel about each product line. A violin plot is widely used as it helps to compare distributions, show peaks in the data, and combine elements (Thrun, Gehlert, and Ultsch, 2020). Insights from ***Figure 3*** can be used in inventory and marketing decisions, which may improve customer engagement and sales performance.