Dear sir/madam,

This is Rupam kumari and I am one of the participants in Cognizant Data Science Virtual Internship program at forage. First of all, Thank you for providing professionals and non-professionals (including me) to learn about your company and how you process data for analysis and modelling for better evaluation. With this email, I performed Task 1 and I’m glad to acquire the knowledge that I got while performing EDA on the provided dataset.

The task included Exploratory Data Analysis on the provided Gala Grocery retail data and the major goal was to better stock the items that they sell in the market to earn more profits and gain more customers while not losing existed ones.

Summary findings:

1. Overall the dataset contains 9 predictors, where we have 3 categorical columns and the remaining columns are numerical.
2. The dataset contains no missing values(7829 rows in each column).
3. Most frequent customers are from:

* timestamp(2022-03-02 19:32:20) with count 2
* product\_id(ecac012c-1dec-41d4-9ebd-56fb7166f6d9) with count 114
* category (fruit) with count 998
* customer\_type (non-member) with count 1601
* payment\_type (cash) with count 2027.

1. Some of the products customer bought a lot like medicine, meat, seafood, beverages etc.

Altogether, my recommendations while analysing this data are:

1. Overall, the data seems biased.
2. We need more training data to generate a best model for predictors.
3. Need more data based on the weekdays and weekend shopping by customers.
4. Need more data based on different time zones(morning, afternoon, evening, night) to get a more clearer way of customer behaviour while purchasing stuffs from the grocery retail.
5. We can combine more frequent bought items along with less bought items to give some discount to customers so that they can also try new options in the provided retail store that can lead us to generate more sales and good customer loyalty in the case product performs good as per user requirement.

Best Regards,

RUPAM KUMARI