Dear Harold,

Gala Groceries, a technology-led grocery chain in the United States, is asking Cognizant for assistance on how to better stock the items they sell. I have completed exploratory data analysis on the csv file Gala Groceries provided.

Below are my findings from the analysis:

* Gala Groceries provided a 7829 row file with 9 features for one week of data and one store. The features included in the file are transaction id, timestamp, product id, customer type, product type, category, unit price, quantity, and total.
* The file contained 300 product ids and 22 product categories. Each item in the file had a unique transaction id. The file did not contain any multiple item groupings for an order. There are four payment types and five customer types. Unit price ranged from $0.19 to $23.99. The timestamp in the file was 3/1/22 9am to 3/7/22 7pm.
* The most popular categories are fruit and vegetables for all days except Monday based on total quantity sold. Fruit and packaged goods are the most popular categories on Monday. The overall order requests for fruit and vegetables exceed the other categories by over 66%.
* Tuesday and Wednesday have the highest value for total quantity sold. The total quantity sold on these days exceeds the other days by at least 200.
* 9am, 4pm, and 11pm are the hours with the highest total quantity sold. 11pm has the highest total quantity sold.

Recommendations:

* The problem statement from the client is too broad. We need to narrow the focus in order to deliver a valuable end product.
* The client needs to provide more rows of data. Data for multiple years, months, days, hours, stores, etc.
* The client should provide additional features such as store, city, state, product name, product supplier, etc. The client should also provide data to determine how frequent an item is out-of-stock, when it is out-of-stock, and how long it is out-of-stock.

Best regards,

Sarah Pfeiffer