Capstone Project - Module 2

General Question

1. What are the business context and assumptions that can be gained from the database?
2. How much is there in the database? Expand the table and describe

SQL

1. For an analysis that is focused on the 'Products' feature of the data, which tables are needed for this analysis? Can they be combined into one table?
2. How much category does the products have?
3. How much is the total quantity for each product?

Data Manipulation

1. Are there any anomalies in the data? If yes, what are they and how do you solve it?
2. Are there any outliers in the data?
3. What are the variables that uses the datetime type?
4. How do you see the selling frequency of each product? If a new data frame is required, describe.

Data Visualization & Statistics

Focus analysis: Products

For Visualizations:

1. What country does most customer come from? From that country, which city?
2. What kind of products is frequently bought from the region from num 1? Which category?
3. Are there any loyal customers?
4. What is the most bought product out of all products?
5. How frequent is the top selling product bought over the years?

For Statistics:

1. Given two hypotheses, where:
   a. H0: the mean of category Meat/Poultry's unit price is equal to the mean of category Seafood's unit price
   b. Ha: the mean of category Meat/Poultry's unit price is not equal and/or greater or less than category Seafood's unit price

   How do you prove these hypotheses? Describe.

2. Does the reorder level of a product affect the discontinuation of the product?