Hackathon Data Problem

Description: One of India's top most cosmetic product company has a wide range of sales spread across different parts of the country. The given data is the sales record of approximately **3 years**. Analyze the data for some **good actionable insights** which can increase their demand planning, warehouse planning, pricing of products which in turn can increase their revenue.

For some **quick ideas** below are the areas you can search for to build your analyses:

a. Get the top revenue contributing products.
b. Continuing with the above point, know where else their sales can be improved.
c. Analyse for their trend and seasonality factors which can give some insights about the product's sales behavior.
d. Analyse the products which cannibalize each other's sales.

With the above mentioned points as your **baseline for analyses**, build your hypotheses around to get some good actionable insights. Participants need not follow the above mentioned points strictly, but they can assume freedom in building their hypotheses for getting the insights.

Common Data Problems:

1. Real-world data always has a considerable noise which needs to be dealt with. The given data is no exception.
2. Identify the noise present and try to come up with a valid solution to rectify it.
3. Need not to be explained, but to mention you need to clean the noise before proceeding with the analyses.

Data Dictionary:

- ParentSKU: Parent Stock Keeping Unit (Primary Product ID)
- Site ID: Sales Depots.
- Category Name ID: Product's Category ID.
- Qty: Number of units sold.
- Price: Price of each unit sold.

Evaluation Parameters:

- Approach to the hypotheses.
- Impact of hypotheses.
- Cleaning the noise in the data.
- Data wrangling.
- Cleanliness in plots i.e, reduce the noise and highlight the important aspects.
- Code efficiency.
- Structuring the modules of the code.