Problem Statement:

A FastMoving Consumer Goods (FMCG) company, let's call it XYZ, is looking for a solution to accurately forecast demand for new product innovations. The company has a dataset of all their Innovation products from the last 3 years, which includes Product, Time, UPC, Category, Sub Category, Dollar Sales, Unit Sales, Avg Prc/Unit, NonPromoted Avg Prc/Unit, Promoted Avg Prc/Unit, Distribution Index, % Stores, and Unit Velocity. The company wants you to build an Innovation Forecasting Model that can accurately predict the sales for a new product for 52 Weeks, **Banana Flavoured Oatmeal**, which is going to be launched next year at a price point of $3.49 and will be on promotion :

- Week 6 to Week 9 @ $2.49
- Week 12 to Week 16 @ Buy 1 get 1 Free.
- Remaining Weeks will be at the Regular price.

Data Dictionary:
- Product: The name of the product.
- Time: The time period of the sales data.
- UPC: The Universal Product Code of the product.
- Category: The category of the product.
- Sub Category: The subcategory of the product.
- Dollar Sales: The total sales revenue in dollars.
- Unit Sales: The total sales volume in units.
- Avg Prc/Unit: The average price per unit.
- NonPromoted Avg Prc/Unit: The average price per unit for nonpromotion periods.
- Promoted Avg Prc/Unit: The average price per unit for promotion periods.
- Distribution Index: A measure of the product's distribution in stores.
- % Stores: The percentage of stores carrying the product.
- Unit Velocity: The rate at which the product is selling in units per week.

Expected Solution:
You are expected to develop an Innovation Forecasting Model that can accurately predict the sales for the new product, Banana Flavored Oatmeal, based on the historical data for the company's Innovation products. The model should take into account various factors such as the product's price point, promotion period, category, and subcategory, Distribution, Velocity etc.

You should use machine learning algorithms such as Linear Regression, Random Forest, XGBoost, Deep Learning or any other algorithm, to build the model. The model should be fine tuned to optimize its performance and accuracy.

You should also use feature engineering techniques to identify the most important variables that impact sales, such as distribution index, % stores, and unit velocity, Price ,Holidays , Seasonality etc. The model should be able to accurately predict the sales for the new product based on its price point, promotion period, category, and subcategory, Distribution, etc.

Conclusion:

The Innovation Forecasting Model developed by the you should be highly accurate in predicting the sales for the new product, Strawberry Flavored Oatmeal. The model should take into account various factors such as the product's price point, promotion period, category, and subcategory, to deliver reliable sales forecasts.

The model's ability to identify the most important variables that impact sales, such as distribution index, % stores, and unit velocity, makes it a valuable tool for the FMCG industry. By leveraging the power of data science and machine learning, FMCG companies can make more informed decisions about their product portfolio and optimize their sales and profitability.

Expected Solution

You are expected to submit the following deliverables:

- **Solution Pitching Deck:** The deck should include a clear explanation of the modeling principles used, the solutions proposed, and the accuracy achieved by the Innovation Forecasting Model. Youshould highlight any key insights and trends in the data and provide visualizations to support their findings.

- **Excel Output**: You should provide an Excel output that includes the predictions for the sales of the new product, Strawberry Flavoured Oatmeal. The output should provide details on the predicted sales volume, revenue, and profitability based on the product's price point, promotion period, category, and sub-category. The output should also include accuracy metrics, such as mean absolute error or root mean squared error, to evaluate the performance of the model.

- **Code**: You should provide well-documented code that can be used to replicate the Innovation Forecasting Model. The code should include details on the machine learning algorithms used, feature engineering techniques applied, and any other relevant modeling details. The code should also include instructions on how to run the model and generate the Excel output.

Overall, the expected solution should demonstrate the your expertise in developing an accurate and reliable Innovation Forecasting Model. The solution should be well-organized, easy to understand, and clearly demonstrate the participant's ability to apply machine learning and data science techniques to solve complex problems in the FMCG industry.