1. Unsupervised Clustering model

   I've used K-means clustering algorithm with the value of k i.e., clusters as 4. I experimented with various other values too. The main objective of my analysis was to get information on which people use web to order and which use the in-store service. So, I gathered, combined as used different parameters for the same. The example that I illustrated in the notebook is the Income of the customer and how long since the customer has been registered with us. We can see clearly that the users have been divided into four groups:

   1. Those who have high income and are registered with us for long (old users) and use the web more often.
   2. Those who have less income and are new users and hence use the web less often.
   3. Those who have high income but are relatively new and use web comparatively more often.
   4. Those who are having less income and are old users and use web less often.

   In a similar fashion different segments can be found out from the data. We can find people using stores physically. Apart from that, the parameters for people's financials and demographics can also be swapped to see their effect. Total number of people in the family and Age along with recency can also be used and plotted to get valuable insights from the data.

2. Classification Model

   Here I've used XGBoost as the classification model. The target parameter was the campaign to which people responded and bought the product. So, there were 6 classes and one case where none of them were chosen. As the data was imbalanced and biased towards the last case, the data had to be normalized. As XGBoost uses Decision Tree as the base model, the parameters like number of nodes, depth, etc. was fine tuned and the accuracy received was around 68.8%.

   The suggestion for the next round of campaign should be providing discounts for a particular item at higher rates at a particular platform to increase volume of sales. For example, the sale of meat was most likely concentrated in in-store shopping, so providing incentives and discounts there will lead to better results rather than using resources everywhere. Such patterns can be observed and used for better results. Apart from that, using customer details pattern to predict what the customer is more likely to buy and recommending such items can also be performed.

   I don't know much about SHAP analysis. I have read about it but I don't have a clear idea about it and due to less time available, I couldn't research well on this topic. So I can't comment on this part.

   Thank You.
   Shivam Thakur