Business Case

Purchasing management includes anticipating product sales. Forecasting future demand for goods, especially in the supermarket and grocery store industry, is critical in setting inventory stock levels. Overall profit might be harmed if commodities are not easily available or if supply exceeds demand. As a result, product sales forecasting can be crucial in reducing loss. Furthermore, as shops add new areas with unique needs, new products, ever-changing seasonal tastes, and unpredictable product promotion, the problem becomes much more complicated.

In this project, I'm attempting to forecast product sales for the supermarket "Corporation Favorita" which is located at Ecuador based on items, stores, transactions, and other dependent variables such as holidays and oil prices from previous years, to better ensure grocery stores to please customers by stocking just enough of the right products at the right time. In order to forecast sales, I have planned to compare different models like Linear Regression, Decision Tree, Random Forest etc..

Description of dataset

stores.csv

- Store metadata, including city, state, type, and cluster.
- cluster is a grouping of similar stores.

items.csv

- Item metadata, including family, class, and perishable.
- NOTE: Items marked as perishable have a score weight of 1.25; otherwise, the weight is 1.0.

transactions.csv

- The count of sales transactions for each date, store_nbr combination.

oil.csv

- Daily oil price. Includes values during both the train and test data timeframe.

  (Ecuador is an oil-dependent country and its economic health is highly vulnerable to shocks in oil prices.)

holidays_events.csv

- Holidays and Events, with metadata