**Phase 1**

Step 1: Data Collection and Preparation

- Gather historical sales data, which should include information about time periods (e.g., daily or monthly sales), product details, pricing, promotions, and any other relevant variables.

- Ensure that your dataset is clean, handle missing values, outliers, and format it properly for analysis.

Step 2: Exploratory Data Analysis (EDA)

- Conduct EDA to understand your data better. Visualize your data, compute summary statistics, and analyze correlations.

- Identify trends, seasonality, and patterns in the historical sales data.

Step 3: Feature Engineering

- Create new features or transform existing ones that can help capture important information for sales prediction. Consider the following features:

- Lag features: Past sales data for the same or related products.

- Time-related features: Day of the week, month, year, and holidays.

- Product-related features: Pricing, product category, promotions, and inventory levels.

Step 4: Data Splitting

- Split your dataset into training and validation (and possibly testing) sets. Typically, you can reserve the most recent data for validation and testing.

- Ensure that your data is time-ordered, so you train on earlier data and validate on more recent data.

Step 5: Model Selection

- Choose an appropriate predictive modeling technique. In this case, time series forecasting models like SARIMA (Seasonal AutoRegressive Integrated Moving Average) or Prophet are often suitable choices.

- Consider trying machine learning models like XGBoost or deep learning models like LSTM if your data has complex patterns.

Step 6: Model Training

- Train your selected model using the training dataset. Pay attention to the hyperparameters of the chosen model.

- If using a time series model, consider tuning hyperparameters like seasonality, trend order, and differencing.

Step 7: Model Evaluation

- Assess the model's performance on the validation dataset using relevant evaluation metrics for sales forecasting, such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), or Mean Absolute Percentage Error (MAPE).

- Compare the model's performance to a baseline model for context.