Developing a Business Analytics Model with Jupyter Notebook

A step-by-step plan for implementing the model:

Data Preprocessing:

- Load the dataset.
- Review the dataset to understand its structure, features, and target variable.
- Handle missing values: Identify any missing values and decide on a strategy to handle them (e.g., imputation or removal).
- Handle duplicates: Check for duplicate rows and decide whether to remove them.
- Split the dataset into features (X) and the target variable (y).

Model Development:

- Linear regression was selected as an appropriate model for the business analytics task
- Divide the dataset into training and testing sets to train and evaluate the model.
- Train the model on the training data.

Model Training:

- Fit the model to the training data.
- Monitor the training process to ensure convergence and avoid overfitting.

Model Evaluation:

- Evaluate the trained model's performance using appropriate evaluation metrics (e.g., mean squared error and R-squared).
- Visualize the model's predictions and compare them with the actual values.

- Interpret the model's coefficients or feature importance to gain insights into the underlying relationships.

Model Optimization:

- Experiment with different models or algorithms to improve performance.
- Fine-tune hyperparameters based on performance metrics and domain knowledge.
- Consider ensemble techniques or advanced models for better performance.

Documentation and Reporting:

- Document the entire process, including data preprocessing steps, model selection criteria, and evaluation results.