Short Plan for Preparing the Group Presentation

Our group is following a structured and phased plan to prepare the Checkout Abandonment Risk Analytics presentation. This plan ensures effective collaboration, clear task allocation, and alignment between data analysis, business insights, and final recommendations.

1. Defining the Business Problem

We are first clarifying the business problem related to high checkout abandonment in e-commerce and its impact on revenue. The group is focusing on predicting abandonment during the checkout session rather than reacting after abandonment occurs. This shared understanding guides all subsequent analysis.

2. Data Understanding and Preparation

We are reviewing the dataset and its star-schema structure, including the session-level fact table and the customer, product, device, and date dimension tables. At this stage, we are ensuring all team members understand the variables, data scope, and data quality before moving forward.

3. Exploratory Data Analysis (EDA)

We are conducting exploratory data analysis to identify key patterns associated with checkout abandonment. This includes analysing abandonment behaviour by device type, customer type, and checkout funnel stage. Insights from this step are informing our feature engineering and modelling approach.

4. Feature Engineering

Based on the findings from EDA, we are engineering additional features such as behavioural ratios, time-based indicators, and interaction variables. These features are being created to enhance model accuracy and support explainable predictions.

5. Model Development and Evaluation

We are developing and evaluating multiple machine learning models using a consistent framework. The data is being split into training and testing sets, and models such as Logistic Regression, Random Forest, and XGBoost are being compared using accuracy and ROC-AUC metrics. XGBoost is currently emerging as the strongest candidate model.

6. Model Interpretation and Explainability

To ensure transparency and business interpretability, we are analysing feature importance and applying SHAP techniques. This allows us to explain which factors contribute most to checkout abandonment and why the model makes specific predictions.

7. Business Impact and Simulation

We are estimating the business value of the model by calculating revenue at risk and simulating the potential impact of real-time intervention strategies. This step helps connect analytical outcomes with measurable business benefits.

8. Recommendations and Implementation Roadmap

Based on the analytical results, we are formulating targeted intervention strategies for different risk segments. In parallel, we are developing a phased implementation roadmap to demonstrate how the solution can be deployed in practice.

9. Presentation Preparation and Team Coordination

The presentation content is being divided among group members according to analysis stages.