**Project Title:**

**Market Basket Analysis – Purchase Pattern Analytics**

**Prepared By:**

**Shilpa Mane**

**Timeline:**

**4 Weeks (W1–W4)**

**1. Project Overview**

* **Objective**: To identify customer purchasing patterns using Market Basket Analysis and provide actionable business insights.
* **Tools Used**:
  + Power BI (for data cleaning, visualizations)
  + Python (Apriori Algorithm)
  + Excel (initial data inspection)
  + Orange3 (early exploratory analysis)

**2. Data Cleaning & Preparation**

**Initial Dataset Fields**:  
BillNo, ItemName, Quantity, Price, CustomerID, Country, PresentDate

**Cleaning Steps**:

* Removed null values and duplicates
* Converted date fields to datetime
* Created a standardized product name column (CleanItem)
* Removed records with Quantity ≤ 0
* Filtered out single-item transactions for co-occurrence logic
* Created derived tables: ItemListA, ItemListB, ProductPairs

**3. Exploratory Data Analysis (EDA)**

* Top selling items identified by frequency
* Total unique products and transactions analyzed
* Used Power BI visuals:
  + Bar chart: Most purchased items
  + Table: Top customers by purchase volume
  + Slicers for interactive filtering by country/date

**4. Apriori Algorithm Implementation**

**Purpose**: Find frequent itemsets and generate association rules.

**Implementation Steps**:

* Used Python
* Required Python environment set with necessary libraries (mlxtend, pandas)
* Script included:
  + Transaction formatting
  + Apriori frequent itemset generation
  + Rule creation using confidence and lift

**5. Visualizations**

* **Heatmap**: Product co-occurrence matrix
* **Network Graph**: Association rules visualization (using Network Navigator in Power BI)
* **KPI Cards**: Total Products, Transactions
* **Charts**:
  + Frequent items
  + Top product pairs
  + Lift & confidence visuals

**6. Key Insights**

* Products frequently purchased together (Lift: 3.2)
* sales involved multi-item purchases
* Certain item groups performed better in specific countries
* Some products act as strong lead-ins for upselling

**7. Recommendations**

* Implement bundling strategies for top product pairs
* Improve placement for frequently co-purchased products
* Launch targeted promotions using high-lift associations
* Use insights for optimizing stock and warehouse planning

**8. Final Deliverables**

* ✔️ Cleaned and Transformed Dataset (CleanItem column, filtered transactions)
* ✔️ Power BI Dashboard (interactive, KPI, slicers, graphs)
* ✔️ Apriori Algorithm Implementation in Python
* ✔️ Network Visual showing product relationships
* ✔️ Final Presentation Slides (PPTX)
* ✔️ This Final Documentation Report (PDF/Word)