# Lookalike Model Report (Task 2)

## *Objective*

The goal of the Lookalike Model was to recommend 3 similar customers based on their profile and transaction history, using cosine similarity.

## *Approach*

1. **Data Merging**: Combined data from the `Customers`, `Transactions`, and `Products` datasets to get a comprehensive view of customer behavior.
2. **Feature Creation**:
   o Total transaction value (`TotalValue`).
   o Total quantity of items purchased (`Quantity`).
   o Number of distinct products purchased (`ProductID`).
3. **Data Normalization**: Used `StandardScaler` to normalize transaction data to a common scale.
4. **Cosine Similarity Calculation**: Calculated the cosine similarity between customers based on the selected features.
5. **Top 3 Similar Customers**: For each customer, the top 3 similar customers were selected based on the cosine similarity score.

## *Evaluation Metrics*

- **Cosine Similarity**: Measures how similar two customers are based on their transaction and product data.
- **Similarity Scores**: Ranges from 0 to 1, where higher values indicate more similarity.

## *Output*

- **Lookalike.csv**: A file containing the top 3 similar customers for each of the first 20 customers (CustomerID: C0001 to C0020), along with their similarity scores.

## *Conclusion*

- The model successfully identifies the most similar customers based on transaction behavior.
- Cosine similarity was an effective method to compute similarity.
- The results are saved in the **Lookalike.csv** file, providing a practical recommendation for each customer.

## *Future Considerations*

- Additional customer profile data (e.g., demographics) could improve recommendations.
- Explore faster algorithms for larger datasets.