



Programming for Artificial Intelligence Individual Assessment Scenario

Scenario for Task 2 – Supermarket Dataset

You have been contracted by a supermarket chain to design a Python-based data structure and algorithm to analyse **customer purchasing patterns** from their transaction records. The dataset you are provided with contains thousands of market basket transactions, where each transaction is a list of items purchased together by a customer during a single shopping visit.

The supermarket is interested in efficiently answering queries such as:

- “Which items are most frequently bought together with bread?”
- “What are the top three most common product bundles across all transactions?”
- “How can we quickly identify whether two items are often co-purchased?”

To support these requirements, your task is to:

1. Data Structure Design

- Choose an appropriate data structure to model the supermarket transactions as a network of items. For example, you might choose to represent the items and their co-occurrences using a **graph**, where:
 - Each node represents a grocery item (e.g., bread, milk, eggs).
 - An edge between two nodes represents that the items were bought together, with edge weights indicating frequency.
- This structure must allow efficient queries such as finding item associations, updating the network with new transactions, and retrieving frequent co-purchases.

2. Algorithm Implementation

- Implement an algorithm to process the dataset into your chosen structure.
- Then implement one or more algorithms that operate on the structure, for example:
 - **Graph traversal/search** (e.g., BFS/DFS) to identify related items.
 - **Frequent pair/itemset mining**
 - **Sorting and ranking algorithms** to identify the top product bundles.

3. Application-Specific Requirements and Extensions

- Visualise the item relationship graph, highlighting the strongest associations.
- Use filters to focus on certain items/itemsets
- Or implement a recommendation-style query: given an item (or itemset), return the top items likely to be bought with it.

This scenario reflects a real-world **market basket analysis** use case, widely applied in online and offline retail for product placement, promotions, and recommendation systems. Your



implementation should focus on efficient data structures and algorithms rather than on external machine learning libraries. To answer this question (Task2) you must only use the data structures and algorithms that you learnt during this module (not other methods which might be out of the scope of this module; if you are not sure, you can double-check with your module tutors.)