Objective of the Problem

It is a problem of customer analysis for retail optimization in this question the company is grouping customers based on their behavioural data (e.g., monthly spending, product returns, session duration, coupon usage) they are doing it to identify pattern of their customers. So that they can have a targeted market strategy but they are taking 15 behavioural metrics per customer which is impractical so we have to reduce it to meaningful data while minimising data loss and also with that they want to quantify the error introduced by the approximation.

Approach Used for solving the Problem

For solving this problem we have used is Principal Component Analysis (PCA) using basic linear algebra operations.

1. Firstly we have centered the data for stability.
2. Then we have used **SVD Decomposition** for finding the directions for which the data varies the most.
3. We have used **Projection to map** the high-dimensional data to top 2-3 components.
4. We have used **Reconstruction** to rebuild the original data from reduced form.
5. Lastly we have used **Error Evaluation** for measuring how much information is lost.

Experience and Learnings

From this assignment ,we gained the following learnings:
1. Understanding PCA without any machine learning libraries which help our understanding of linear algebra behind dimensionally reduction.

2. We have used SVD to compute principal directions , a core technique in linear algebra that decomposes data into orthogonal bases and singular values.
3. We learned how to determine the number of dimensions to keep based on cumulative variance, ensuring we retain most of the meaningful information.
4. We gained insight into the trade-off between compression and accuracy by measuring reconstruction error.