Task 01 - Data Normalization & Z-Score

This report demonstrates data normalization techniques applied to a synthetic Sales dataset containing 10,000 rows.

Techniques Applied:

1   1. Min-Max Normalization: Scales values to a fixed range [0, 1]. Formula: (x - min) / (max - min)
2   2. Z-Score Normalization (Standardization): Transforms data to have mean = 0 and standard deviation = 1. Formula: (x - μ) / σ
3   3. Robust Normalization: Uses median and interquartile range to reduce the effect of outliers.
4   4. MaxAbs Normalization: Scales values by dividing by the maximum absolute value.

Mean Centering with Z-Score:

Performed using sklearn's StandardScaler with mean centering enabled. This ensures each feature has a mean of 0, which improves model performance in machine learning algorithms sensitive to scale.

Observations:

1   After Min-Max scaling, all values are between 0 and 1.
2   Z-Score normalization ensures each feature is centered and standardized.
3   Robust scaling reduces the effect of outliers compared to Min-Max scaling.
4   MaxAbs scaling works well for sparse datasets by scaling between -1 and 1.