# Scaling Techniques Documentation

## 1. Introduction

Scaling is an essential step in data preprocessing to ensure that numerical features have similar ranges, preventing certain features from dominating the model due to their large values.

## 2. Why Scaling?

- Prevents features with larger values from overshadowing smaller ones.
- Improves model convergence in algorithms like K-Means, Logistic Regression, and Neural Networks.
- Enhances performance of distance-based algorithms (K-Means Clustering, KNN).

## 3. Implementation in Python

Python Code:

```
from sklearn.preprocessing import MinMaxScaler, LabelEncoder

scaler = MinMaxScaler() X_train[['tenure', 'MonthlyCharges']] =
scaler.fit_transform(X_train[['tenure', 'MonthlyCharges']]) X_test[['tenure',
'MonthlyCharges']] = scaler.transform(X_test[['tenure', 'MonthlyCharges']])

print(X_train.head())
```

## 4. Conclusion

Scaling is crucial in clustering analysis to ensure all features contribute equally. We used both Min-Max Scaling and Standardization to preprocess the dataset before applying K-Means clustering.