# KNIME - DBSCAN Clustering Documentation

## Introduction

This document provides insights into the role of KNIME in the hotel booking dataset analysis project, focusing on its application in performing DBSCAN clustering. KNIME served as a powerful tool for exploratory data analysis (Explore stage) within the SEMMA methodology.

## Task Accomplished with KNIME

## DBSCAN Clustering

- Objective:
    - Uncover hidden patterns and clusters within the dataset to derive valuable insights.
- Workflow:
    - Utilized KNIME to implement the DBSCAN clustering algorithm.
- Nodes Used:
    - DBSCAN (Density-Based Spatial Clustering of Applications with Noise) Node:
        - Configured the DBSCAN node to perform clustering based on density, allowing the identification of data points in dense regions.
    - Cluster Assigner Node:
        - Applied the Cluster Assigner node to assign each data point to a specific cluster.
- Configurations:
    - Adjusted parameters within the DBSCAN node to optimize clustering results.
- Results:
    - Identified and visualized clusters formed within the dataset based on cancellation patterns and other relevant factors.

## KNIME Workflow

## Task: DBSCAN Clustering

Node Configuration:

- Configured the DBSCAN clustering node in KNIME with appropriate settings.

Execution:

- Executed the DBSCAN clustering workflow to identify clusters within the hotel booking dataset.

Result Analysis:

- Analyzed the results, focusing on the formed clusters and their significance in understanding booking patterns.

## Conclusion

KNIME played a crucial role in the exploration phase of the SEMMA methodology by enabling DBSCAN clustering. The utilization of KNIME, specifically the DBSCAN node and Cluster Assigner node, allowed for a comprehensive analysis of hidden patterns and clusters in the dataset, providing valuable insights for further exploration and modeling stages.