K-Means Clustering Analysis Report

Project Title: Audience Segmentation Based on Competitor Traffic Data

Objective: To identify high-opportunity keywords for targeted marketing campaigns using K-Means clustering.

---

1. Introduction

The purpose of this analysis was to study competitor keyword traffic and keyword difficulty data to understand audience preferences and uncover actionable insights. Using K-Means clustering, the goal was to segment the keywords into meaningful groups that can inform marketing strategies.

---

2. Methodology

- **Dataset:** Competitor traffic and keyword performance data.
- **Selected Features:**
    - Traffic (number of visits generated by each keyword)
    - Keyword Difficulty (competitiveness of the keyword)
- **Tools Used:** Python (Pandas, Scikit-Learn, Matplotlib)
- **Approach:**
    - Data preprocessing and scaling
    - Application of K-Means clustering algorithm
    - Silhouette analysis to determine the quality of clustering

---

3. Results

- **Optimal Number of Clusters:** 2 (based on highest Silhouette Score)
- **Silhouette Score:** 0.514 (good cluster separation)

Based on KMeans clustering of Traffic and Keyword Difficulty, **two distinct audience segments** were identified, enabling targeted marketing campaigns toward high-opportunity keywords.

Cluster Profiling:

Cluster	Average Traffic	Average Keyword Difficulty	Description
0	High	Low	High Opportunity Keywords
1	Lower	Higher	Competitive or Low-Traffic Keywords

Interpretation:

- **Cluster 0** consists of keywords with higher traffic and lower competition — ideal for content marketing and paid ad focus.
- **Cluster 1** includes more competitive keywords with lower traffic — requiring strategic investment or alternative approaches.

---

4. Conclusion

This analysis successfully segmented keywords into actionable groups. Marketers can prioritize keywords in Cluster 0 to maximize reach and ROI while planning carefully for Cluster 1 keywords.

---

5. Recommendations

- Focus campaign budgets and SEO efforts on Cluster 0 keywords.
- Perform additional content differentiation for Cluster 1 keywords.
- Regularly refresh the clustering analysis as new traffic and difficulty data emerge.

Prepared by: Anuradha Menon

Internship Project: Website Traffic Analysis for Audience Preference Identification

Date: 27/04/2025