Finlatics
ENABLING INSIGHTS

Facebook Marketplace Case Project

Analyze social media engagement data for Thai fashion and cosmetics brands to optimize customer interaction strategies.

A case study carried out by Aryan Singhal.

OUTLINE

- Introduction
- Data cleaning
- Key visualizations
- K-Means Clustering

Introduction

Summary of Dataset

The dataset offers insights into social media engagement for Thai fashion and cosmetics sellers, tracking reactions, comments, and shares. Analyzing post types and timing helps businesses understand and optimize customer engagement strategies.

Column Descriptions

- **status_id:** Post identifier.
- **status type:** Post type.
- **status_published:** Date and time.
- **num_reactions:** Total reactions.
- **num_comments:** Total comments.
- **num_shares:** Total shares.
- **num_likes:** ‘Like’ reactions.
- **num_loves:** ‘Love’ reactions.
- **num_wows:** ‘Wow’ reactions.
- **num_hahas:** ‘Haha’ reactions.

Data cleaning

Resolving issues of messy and dirty data

Issues with Dataset

Dirty Data

- **Validity issues** like wrong data-type assigned.

Messy Data

- **Unnecessary columns**

Key visualizations

Visualizations from the dataset

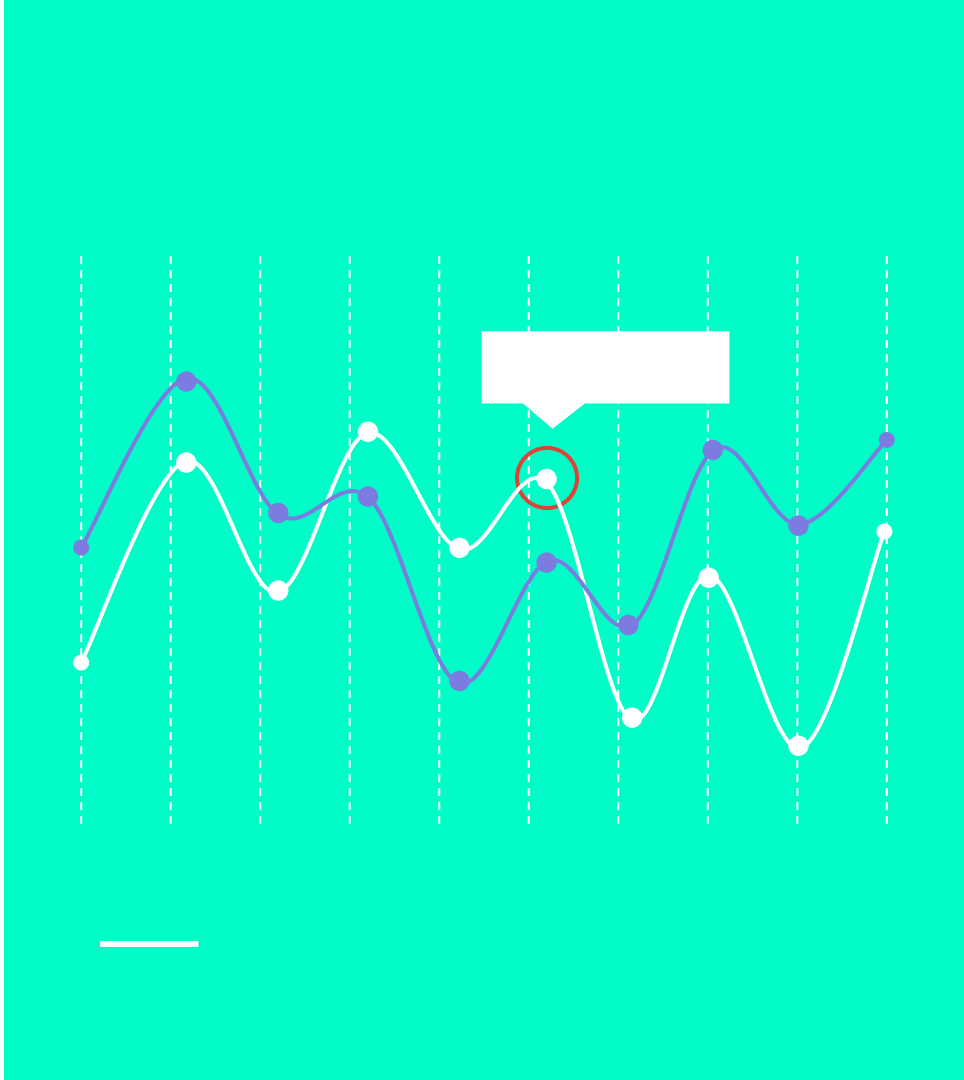


Chart 1: Reactions vs Upload Time

Reactions are generally spread throughout the day, with slightly higher concentrations in the early morning and late evening hours.

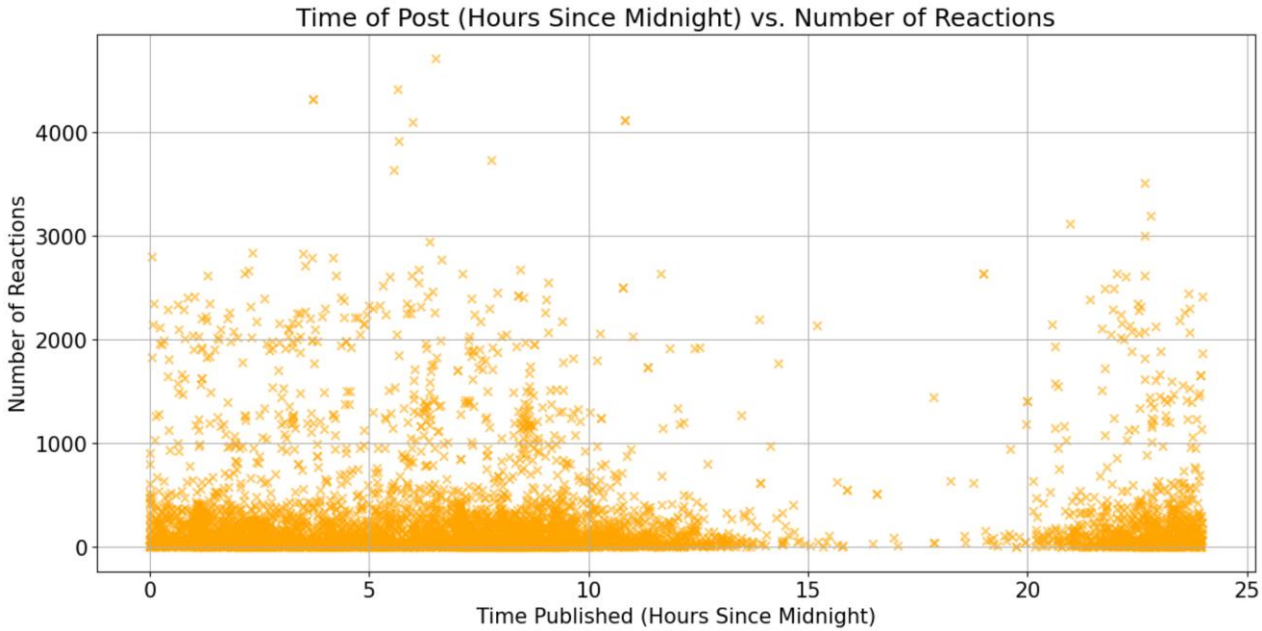


Chart 2: Correlation of Metrics

Reactions and likes are nearly perfectly correlated. Reactions and comments, and reactions and shares, start with strong correlations at low engagement, but both weaken as engagement increases.

Chart 3: Reactions vs Comments

Reactions and comments strongly correlate at low engagement levels, but this weakens as engagement rises, leading to a weak positive correlation overall.
(Pearson: 0.15, Spearman: 0.73)

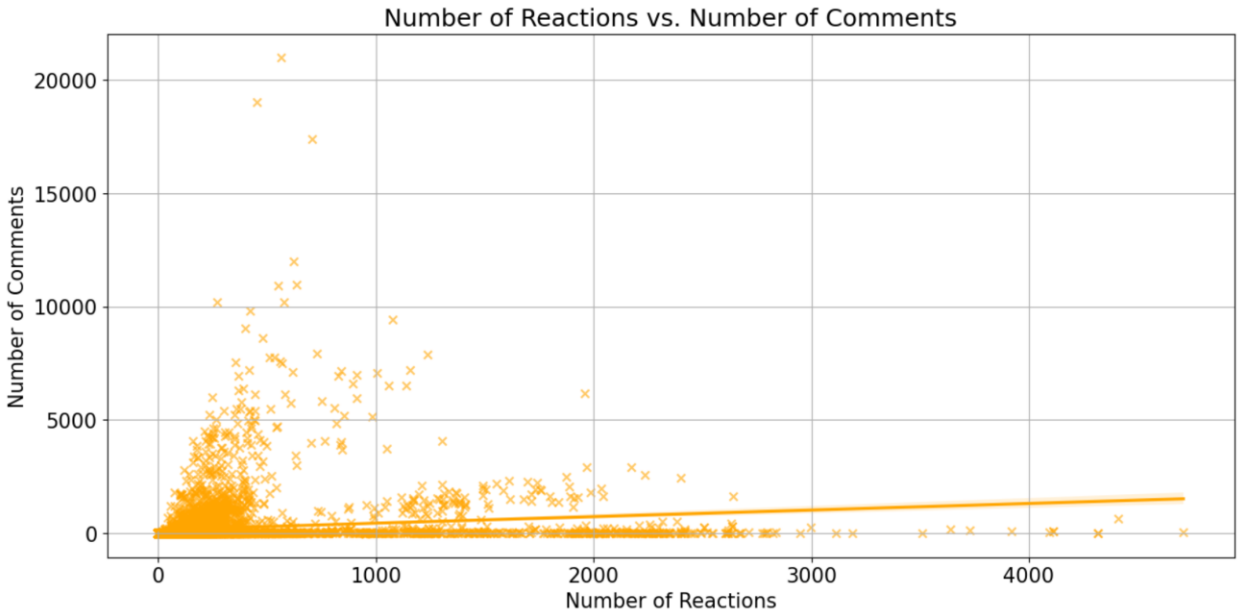


Chart 4: Reactions vs Likes

Reactions and likes have a nearly perfect linear relationship, increasing together at an equivalent rate.

(Pearson:0.99, Spearman:1)

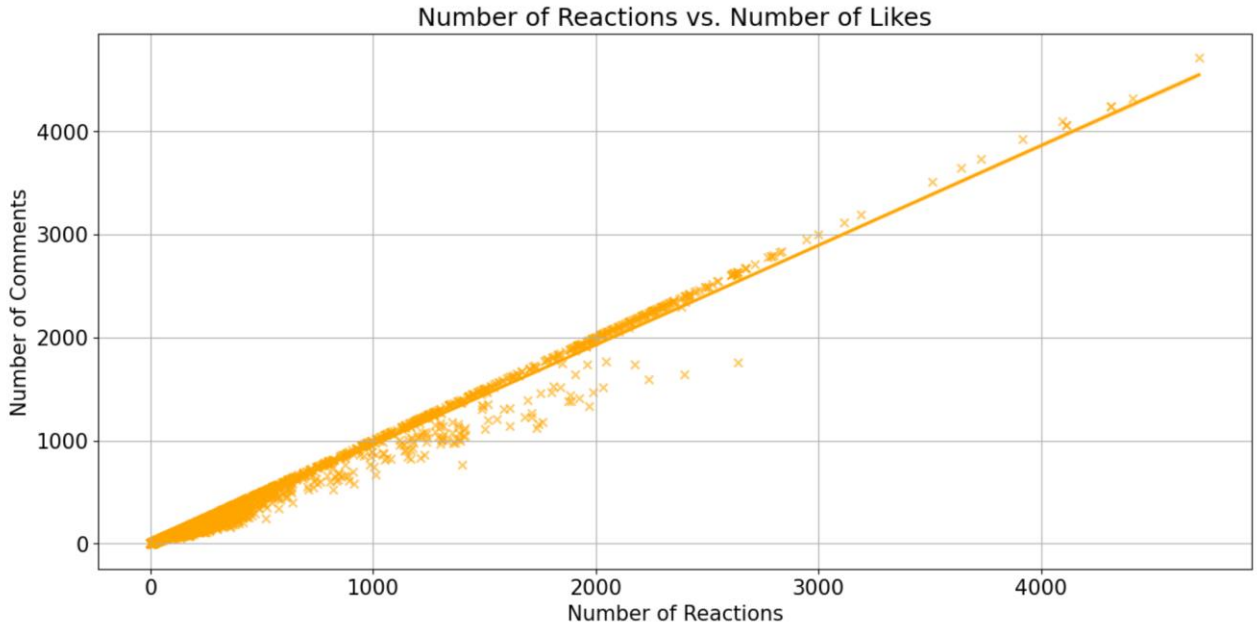


Chart 5: Reactions vs Shares

Reactions and shares are positively correlated at low engagement levels, but declines to a weak positive correlation as engagement increases

(Pearson: 0.25, Spearman: 0.56).

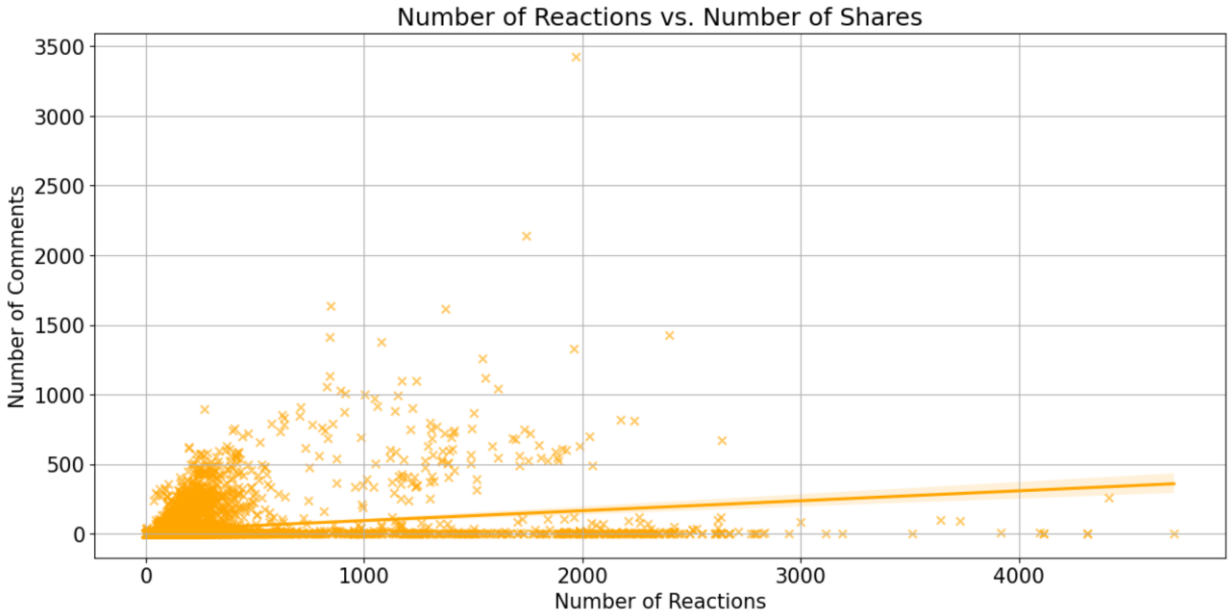


Chart 6: Post-Type Distribution

The dataset consists of 60.8% photos (4288 posts), 33.1% videos (2334 posts), 5.2% statuses (364 posts), and 0.9% links (63 posts).

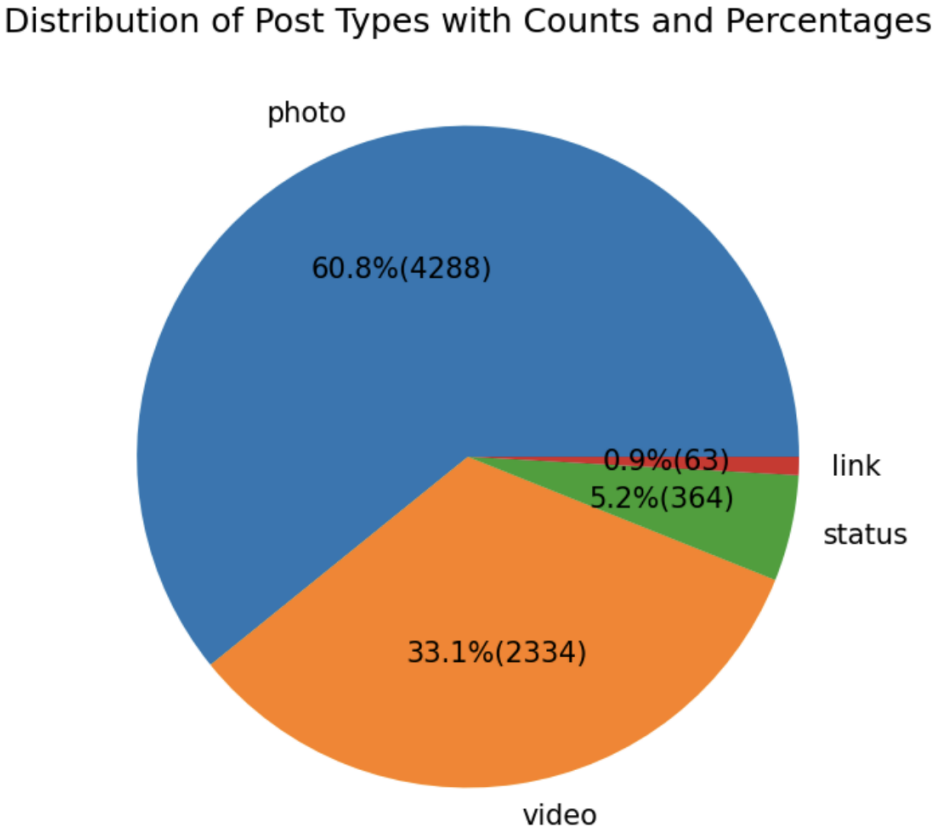
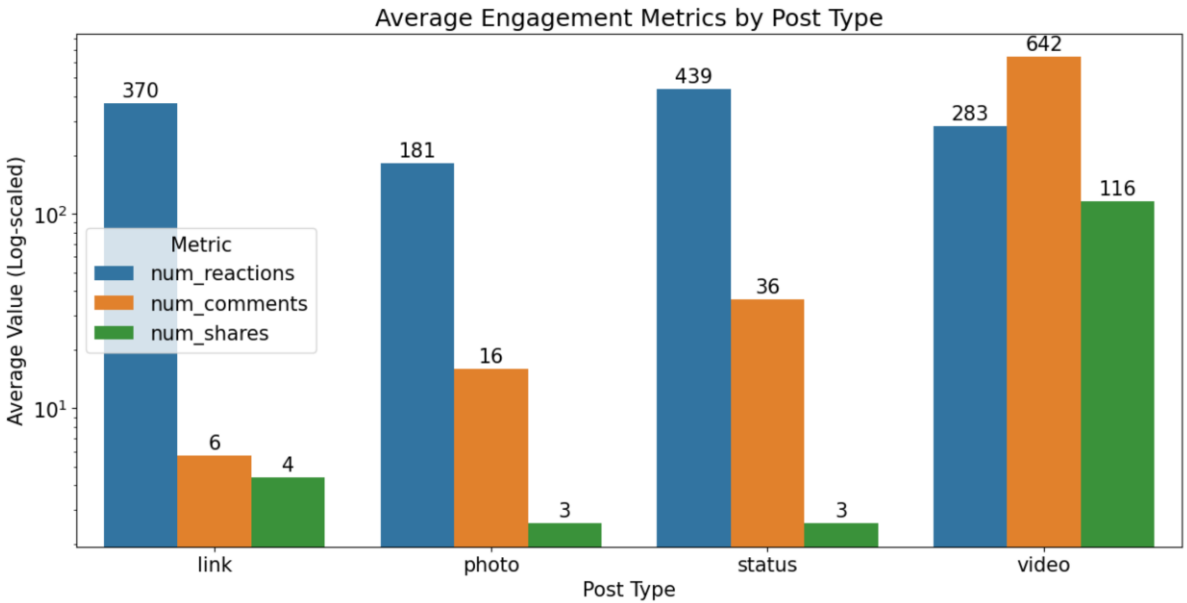
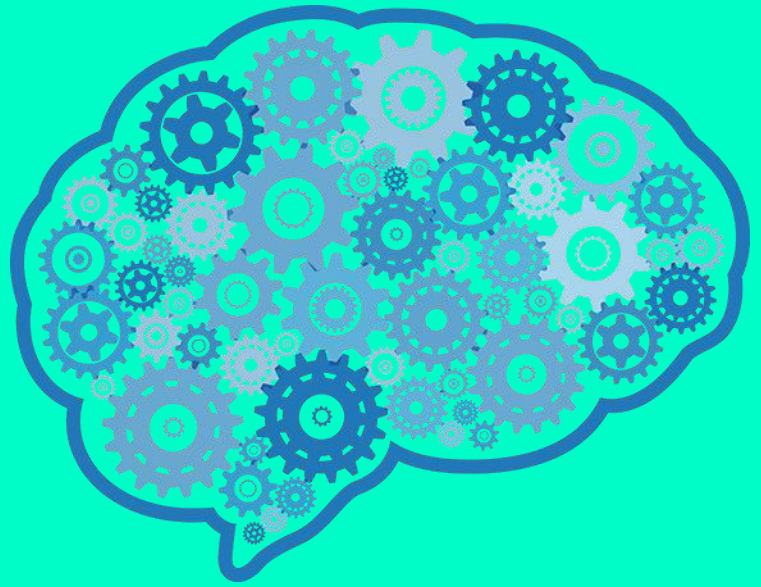


Chart 7: Average Metrics vs Post-Type

Status posts have the most reactions (439), while photos have the fewest (181). Videos get the most comments (642) and shares (116), while links have the least engagement with 6 comments and 4 shares.



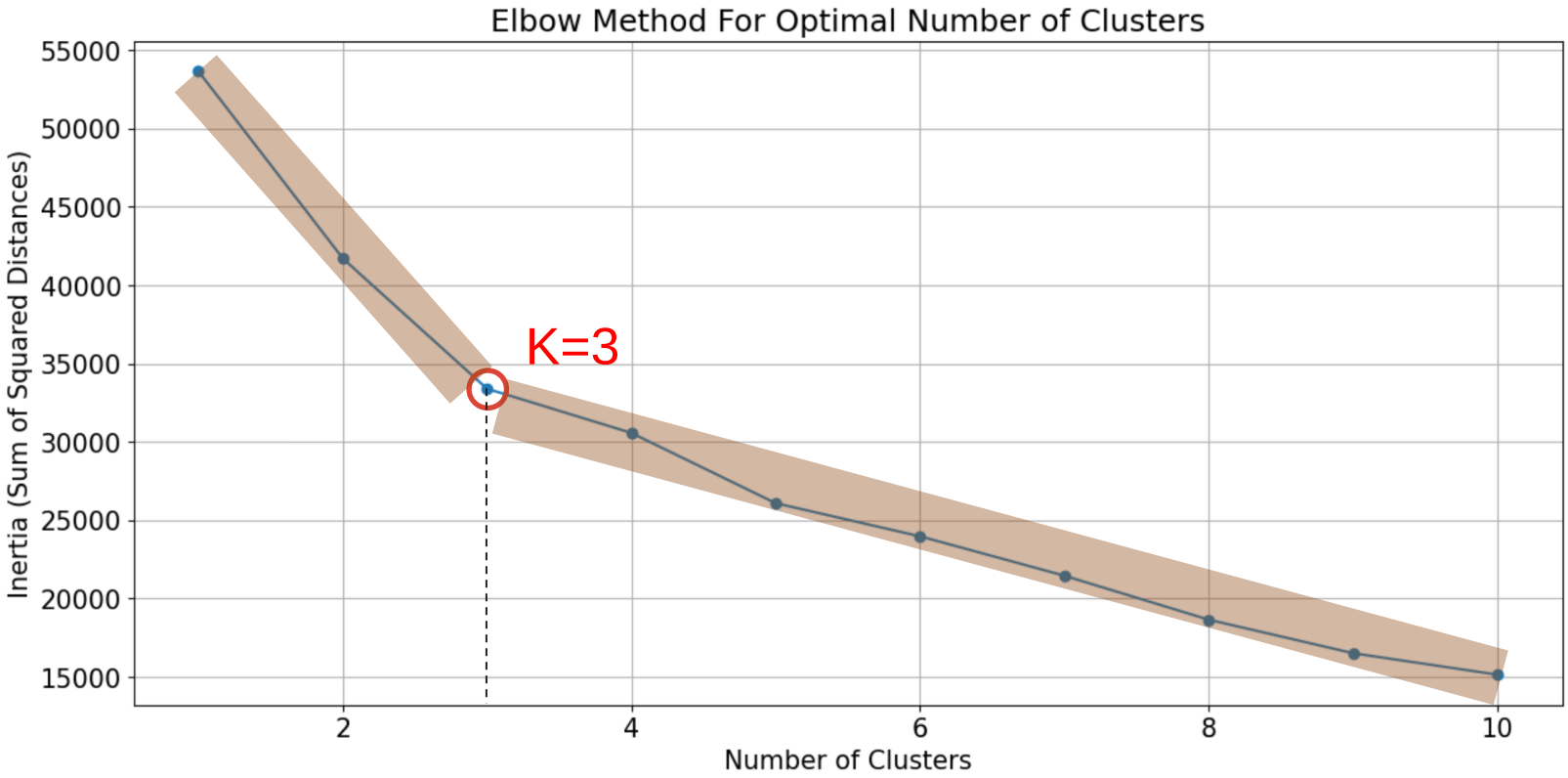
K-Means Clustering



Workflow

- Train-Test Split
- One-hot Encoding (Categorical)
- Data Scaling (Numerical)
- Fit K-Means Model
- Elbow Method for Optimal Clusters
- Refit K-Means with Optimal Clusters

Elbow Method to find optimal Clusters



THANK YOU!