Social Media Engagement Analysis

Dataset Characteristics

The dataset I selected for my project is the *Social Media Engagement Report* from Kaggle. It contains information about user interactions with social media posts from different platforms such as Twitter, Instagram, LinkedIn and Facebook.

(The dataset has an AS index as you called it, way bigger than you recommended and I'm aware of that. I may do a reduction of the dataset if I deem it necessary.)

The main variables of the dataset are:

Platform: The social media site where the content was posted.

Post Type: Type of content (video, image, link, etc.).

Likes, Comments, Shares: Variables indicating user engagement.

Impressions and Reach: How many times the post was viewed by unique users.

Engagement Rate: A metric indicating the "effectiveness" of the post.

Sentiment: Categorization of post sentiment (positive, neutral, negative).

Audience Demographics: Includes age, gender, location, and interests.

General Idea

The primary objective of this study is to analyse engagement trends on social media platforms to understand which factors get more user interaction. This analysis aims to identify which types of posts, content strategies, and user demographics contribute to higher engagement levels.

Intended User

The one who would benefit the most from this project would be marketing analysts and social media managers, because they can get insights in effective engagement strategies and which social media platform prioritize for their strategies.

It can be beneficial too for researchers or data analysts exploring online interaction trends.

Used Analytics

The used analytics will be:

Exploratory Data Analysis (EDA): Summary statistics and engagement trends.

Dimensionality Reduction: PCA to cluster posts based on engagement metrics. I use PCA because the dataset is too big so it's the most efficient.

Network Analysis: Graph visualization of influencer-audience relationships.

Sentiment Analysis: Correlating sentiment polarity with engagement performance.

Relation to the Visual Analytics Cycle

First, the dataset is pre-processed by cleaning, normalizing, and applying dimensionality reduction (PCA) to extract relevant patterns.

Then, various visual representations, such as bar charts, heatmaps, network graphs, and scatter plots, are used to analyse engagement trends, influencer impact, and sentiment correlation. These visualizations help generate actionable knowledge, identifying key factors that drive user interaction.

The cycle includes a feedback loop, where the analysis is continuously refined by adjusting data processing, optimizing visual clarity to improve insight generation. This iterative approach ensures a comprehensive and dynamic understanding of social media engagement trends.

User Interface Mockup

A preliminary UI would include:

A dashboard displaying engagement trends, filters for selecting platform, range of time (or select date), sentiment, and influencer ID. As well as interactive charts (bar charts, heatmaps, network graphs).