

CHAPTER I

INTRODUCTION

1.1 Background

The development of the coffee industry in Indonesia in recent years shows that coffee is no longer viewed merely as a consumption commodity, but has also become part of people's lifestyle. This change has encouraged the emergence of various local coffee brands that rely on quick-service concepts, digital channels, app-based promotions, and outlet expansion in various regions. One of the local coffee brands that is widely known by the public is Kopi Kenangan. In the increasingly competitive food and beverage (F&B) industry, understanding changes in public attention toward a brand becomes important as part of market trend observation.

Amid the increasing activities of society in digital spaces, online search data can be used as one source of information to read the tendency of public attention toward a keyword, topic, or brand. One service that provides such data is Google Trends. Google Trends displays search interest for a keyword in the form of a relative index normalized by time and region. The displayed values range from 0 to 100, where 100 indicates the highest relative popularity point in the selected period and region. Therefore, Google Trends data cannot be interpreted as the absolute number of searches [16].

In the context of digital marketing research and information-search behavior, online search data can provide an initial picture of public attention or interest in a particular product, brand, or phenomenon. However, the use of search data must be carried out carefully because search activity does not always end in a purchase or transaction. Therefore, Google Search Trends for the keyword “Kopi Kenangan” in this study is used as a proxy or early indicator of public search interest, not as sales data, number of consumers, transaction volume, or actual demand for Kopi Kenangan products [1], [16].

The use of Google Trends in forecasting research has developed in various fields. Google Trends data has been used as a supporting indicator in forecasting waste generation volume, tourist visits, tourism demand, and economic phenomena [7]-[11]. Research also shows that Google Trends data may contain noise, low values, and sampling variation, so preprocessing is needed before it is used in a forecasting model [12]. These findings reinforce that Google Trends can be a

relevant digital data source as long as its interpretation is limited to search interest and is not equated with actual sales measures.

In the context of consumer products, Rambung, Wahyuningsih, and Dani (2024) applied Singular Spectrum Analysis (SSA) and a hybrid SSA-ARIMA model to forecast Google Trends search trends for the chocolate keyword “Silverqueen” in Indonesia [1]. That study is closely related to this research because both use Google Trends for consumer-product brands in Indonesia. The difference is that this study focuses on the keyword “Kopi Kenangan” and uses SSA as the main forecasting method.

Google Trends data has the characteristics of a time series because its values are observed sequentially over time. A time series may contain trend patterns, seasonal patterns, fluctuations, and random components or noise. In brand search data such as “Kopi Kenangan”, changes in the search index can be influenced by various external factors such as promotions, digital campaigns, seasonal moments, social media issues, or changes in public behavior. However, these factors are not included as variables in this study because the research focuses on univariate forecasting of the Google Search Trends index.

Singular Spectrum Analysis was selected because it is a nonparametric method capable of decomposing a time series into main components, such as trend, periodic components, and noise. SSA does not require a specific parametric model form at the beginning, making it suitable for fluctuating index data. Several studies show that SSA can be applied to rainfall data, passenger numbers, composite indices, and Google Trends search trends [1]-[6].

This study also considers that the SSA model requires comparison models so that the evaluation results have a clearer context. Therefore, in addition to testing various SSA parameter combinations, this study uses naive, seasonal naive, moving average, and simple ARIMA grid baselines. These baselines do not change the focus of the study, but are used as benchmarks to assess whether SSA produces lower error than simpler models.

Based on the description above, the research gap lies in the application of SSA to forecast Google Search Trends for the keyword “Kopi Kenangan” in Indonesia using research data up to June 2026, a more complete evaluation, comparison baselines, and additional walk-forward validation. With clear interpretive limitations, the results of this study are expected to provide an academic

overview of the tendency of public search-index interest in Kopi Kenangan without making excessive claims regarding sales or actual demand.

1.2 Problem Formulation

Based on the background above, the problem formulations in this study are as follows:

1. What are the characteristics of the Google Search Trends time series for the keyword “Kopi Kenangan” during the period January 2021 to June 2026 in terms of trend patterns, monthly variation, and fluctuations?
2. What are the results of applying the Singular Spectrum Analysis method to Google Search Trends data for the keyword “Kopi Kenangan”, from the decomposition process to time series reconstruction?
3. What SSA parameter combination produces the best model based on the evaluation metrics MAE, MSE, RMSE, MAPE, sMAPE, and MASE?
4. How does SSA perform compared with the naive, seasonal naive, moving average, and simple ARIMA-grid baselines?
5. What is the form of the Google Search Trends index forecast for the keyword “Kopi Kenangan” for the next 12 months based on the best SSA model?

1.3 Research Objectives

The research objectives are arranged to answer the problem formulations in a directed manner. The objectives of this study are as follows:

1. To describe the characteristics of the Google Search Trends time series for the keyword “Kopi Kenangan” during the period January 2021 to June 2026.
2. To prepare the Google Trends data processing workflow so that it is ready to be analyzed using SSA.
3. To apply SSA to Google Search Trends data using several configurations of window length and number of reconstruction components.
4. To evaluate SSA accuracy using MAE, MSE, RMSE, MAPE, sMAPE, and MASE and compare it with simple baselines.
5. To produce a 12-month forecast of the Google Search Trends index and interpret the results as a tendency of public search interest.

1.4 Research Benefits

This study is expected to provide theoretical and practical benefits. These benefits remain limited to the context of forecasting digital search indices, not direct sales decision-making.

1.4.1 Theoretical Benefits

1. To enrich studies on the application of SSA to Google Trends-based time series data, particularly in the context of F&B brands in Indonesia.
2. To provide an example of using Google Trends as a proxy for public search interest with explicit interpretive boundaries.
3. To demonstrate the importance of comparison baselines, walk-forward validation, and additional metrics in index-based forecasting research.

1.4.2 Practical Benefits

1. To provide an initial overview of the tendency of public searches for the keyword “Kopi Kenangan” over time.
2. To provide an example of technical stages for data acquisition, preprocessing, SSA modeling, model evaluation, and interpretation of forecasting results.
3. To serve as an initial reference for researchers or data practitioners who wish to develop a digital search-trend monitoring system based on open data.

1.5 Problem Limitations

In order to keep the study focused and avoid claims that exceed the data, the problem limitations are set as follows:

1. The data used are Google Search Trends for the keyword “Kopi Kenangan”.
2. The analyzed data period is limited to January 2021 to June 2026.
3. The data collection region is limited to Indonesia and the search type is Web Search.
4. The final data frequency used is monthly. If the initial data have a higher frequency, the data are aggregated into monthly averages.
5. This research is univariate because it uses only one main variable, namely the Google Search Trends index.

6. The main method used is SSA. Naive, seasonal naive, moving average, and ARIMA baselines are used only as performance comparisons, not as the main focus of the research.

7. This study does not use sales data, transaction volume, promotions, social media, product prices, or number of outlets as external variables.

8. The forecasting results are interpreted as the tendency of the public search index, not as predictions of sales or actual demand for Kopi Kenangan products.