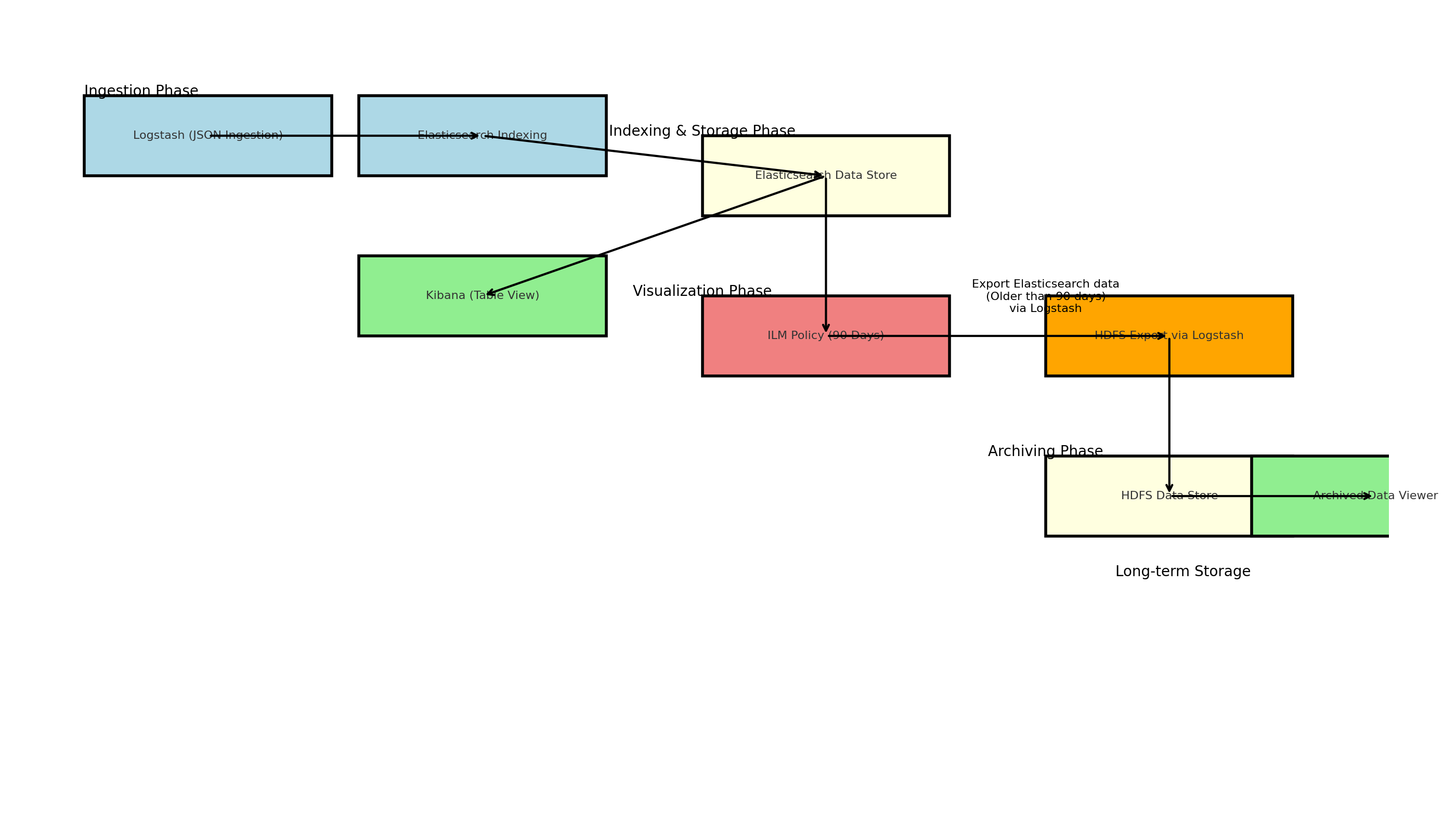
## Workflow Diagram

The diagram below illustrates the step-by-step workflow for this process:



## Diagram Description

1. \*\*Logstash (JSON Ingestion)\*\*: The process starts with Logstash ingesting JSON data from various sources such as files, APIs, or message queues. Logstash processes and formats the data before sending it to Elasticsearch.

2. \*\*Elasticsearch Indexing\*\*: The ingested data is indexed in Elasticsearch, making it searchable and available for further processing.

3. \*\*Elasticsearch Data Store\*\*: Indexed data is stored in Elasticsearch for easy querying and retrieval.

4. \*\*Kibana (Table View)\*\*: Kibana is used to visualize the data stored in Elasticsearch. It allows users to query the data using SQL and view it in a tabular format.

5. \*\*ILM Policy (90 Days)\*\*: An Index Lifecycle Management (ILM) policy is applied to manage the lifecycle of data. Data older than 90 days is marked for archival.

6. \*\*HDFS Export via Logstash\*\*: Using Logstash, the archived data is exported from Elasticsearch to HDFS. The HDFS output plugin writes the data in JSON format to the specified HDFS directory.

7. \*\*HDFS Data Store\*\*: The exported data is stored in HDFS for long-term storage and archiving.

8. \*\*Archived Data Viewer\*\*: Archived data can be retrieved and viewed using appropriate HDFS tools for analysis or restoration.

## Creating a Data Table Visualization in Kibana

A screenshot of a graph

Description automatically generated