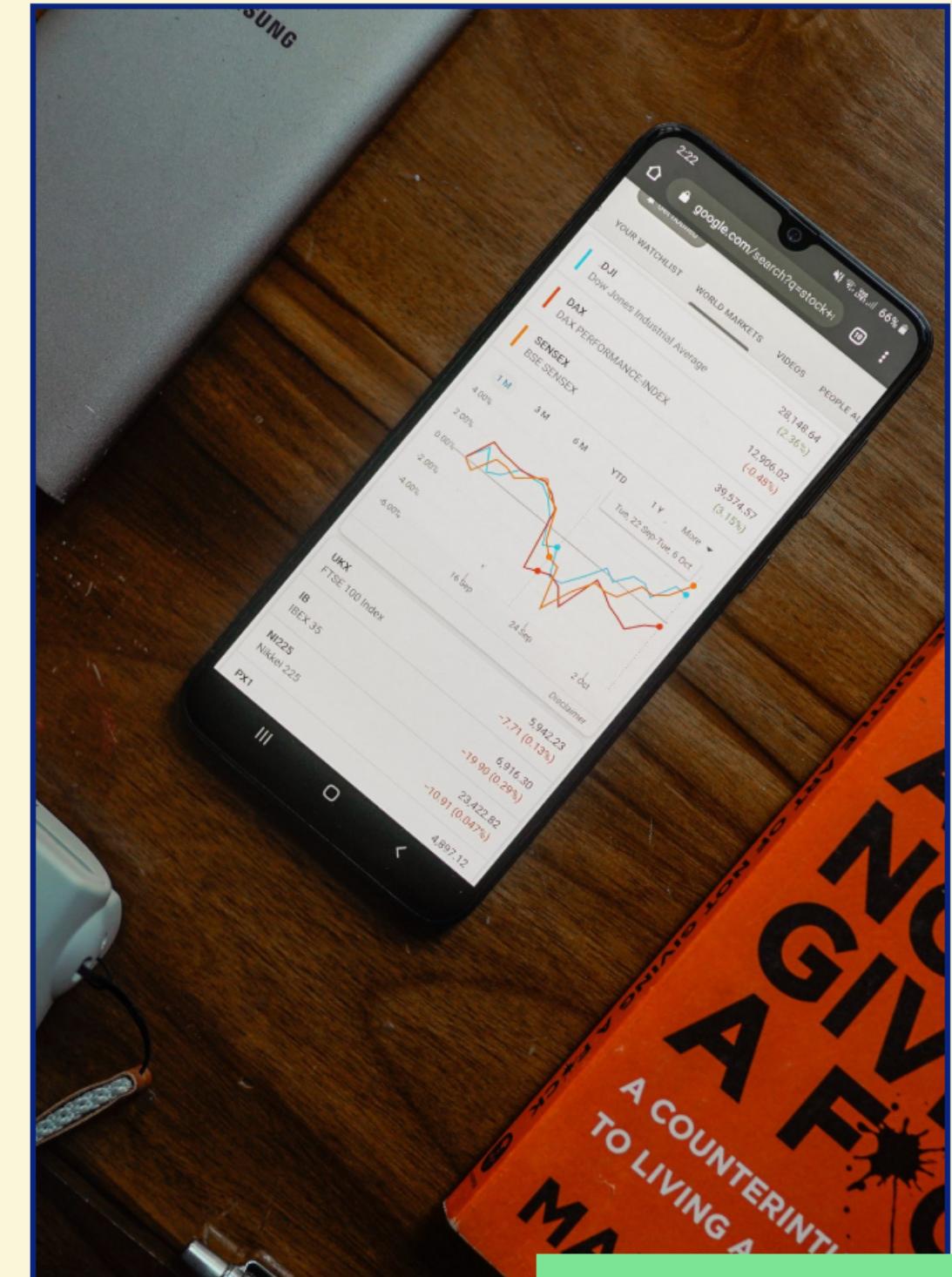


Stock Market Analysis Overview

Exploring the convergence of Hive technology and web integration for effective stock market analysis.

Harshwardhan Jain, Raghav Sharma, Mehul Miglani



Stock Market Data Integration

■ Stock Market: Massive Data

The stock market generates vast amounts of data, necessitating advanced tools for effective analysis.

■ Traditional Systems: Lagging

Conventional data processing systems struggle with the volume and speed of stock market data, leading to inefficiencies.

■ Hadoop + Hive: Data Analysis

Utilizing **Hadoop** and **Hive** allows for scalable data storage and efficient querying of stock market data.

■ Django + React: Web Access

Employing **Django** for server-side logic and **React** for the front end enhances user interaction with stock market data.

■ Data-Driven Decisions

Integrating advanced technologies supports more informed, data-driven decisions in stock market investments.

Scalable Stock Analysis with Hive

■ **Hive enables scalable stock analysis**

Leveraging Hive allows for handling large datasets efficiently, essential in stock market analysis.

■ **Queries extract trends from big data**

Advanced querying capabilities enable the extraction of meaningful trends from vast amounts of stock data.

■ **Website interface shows query + results + chart**

A user-friendly web interface displays queries alongside their results and visual representations, enhancing user experience.

Stock Market Analysis Objectives

1

Load stock data

Ingest stock market data into HDFS for analysis.

2

Analyze using HiveQL

Leverage HiveQL to perform comprehensive data analysis.

3

Visualize results

Create visualizations to represent analysis outcomes effectively.

Tools for Stock Market Analysis

1 Hadoop for Big Data Processing

Utilized for scalable storage and processing of large datasets in stock market analysis.

2 Hive for Data Querying

Provides a SQL-like interface for querying and managing large datasets stored in Hadoop.

3 HDFS for Distributed Storage

Enables high-throughput access to application data with a distributed file system.

4 Django as API Backend

Framework for developing the backend, facilitating robust API integration for the application.

5 React + Tailwind for UI

Combines React for dynamic user interface with Tailwind for responsive design and styling.

6 Python for Data Visualization

Used for creating charts and visual representations of stock market data analysis.

Integrating Hive for Stock Analysis



Upload data to HDFS enables storage

The first step involves uploading stock market data to HDFS for efficient storage and processing.



Create Hive tables for data structure

Next, Hive tables are created to define the structure of the uploaded stock data, making it queryable.



Run queries to extract insights

Queries are executed on the Hive tables to extract relevant stock market insights and trends from the data.



Django API serves data to front end

A Django API is deployed to serve the processed data to the front-end applications for visualization.



React shows results in a user-friendly format

Finally, React is used to display the results of the queries in a user-friendly interface, allowing users to interact with the data.

Comprehensive Stock Dataset Overview

Dataset

Description

StockPrices.csv

- 850K+ rows, prices, volume

StockCompanies.csv

- 500+ rows, sector, industry

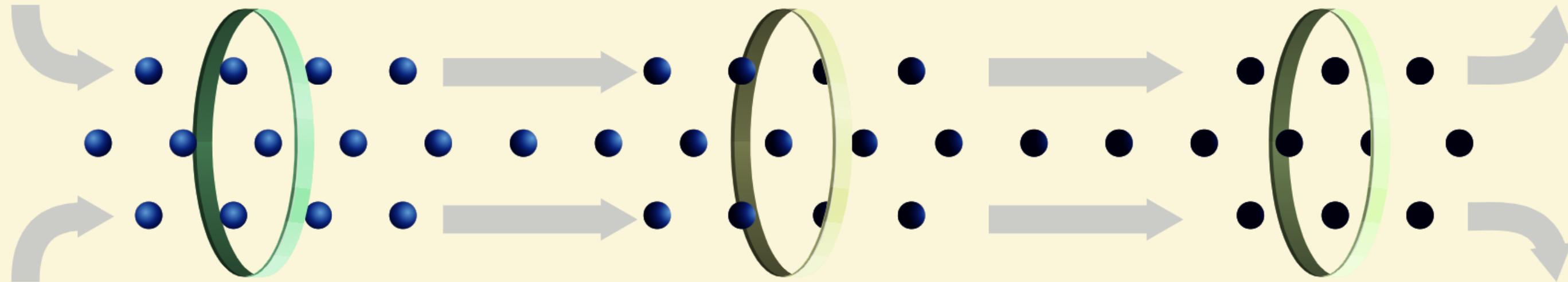
Linked by

stock_id

Data Loading Process Overview

HDFS Data Upload

`hadoop fs -put` command is used to upload data files to the HDFS (Hadoop Distributed File System). This is a crucial step in preparing data for analysis.



Hive Table Creation

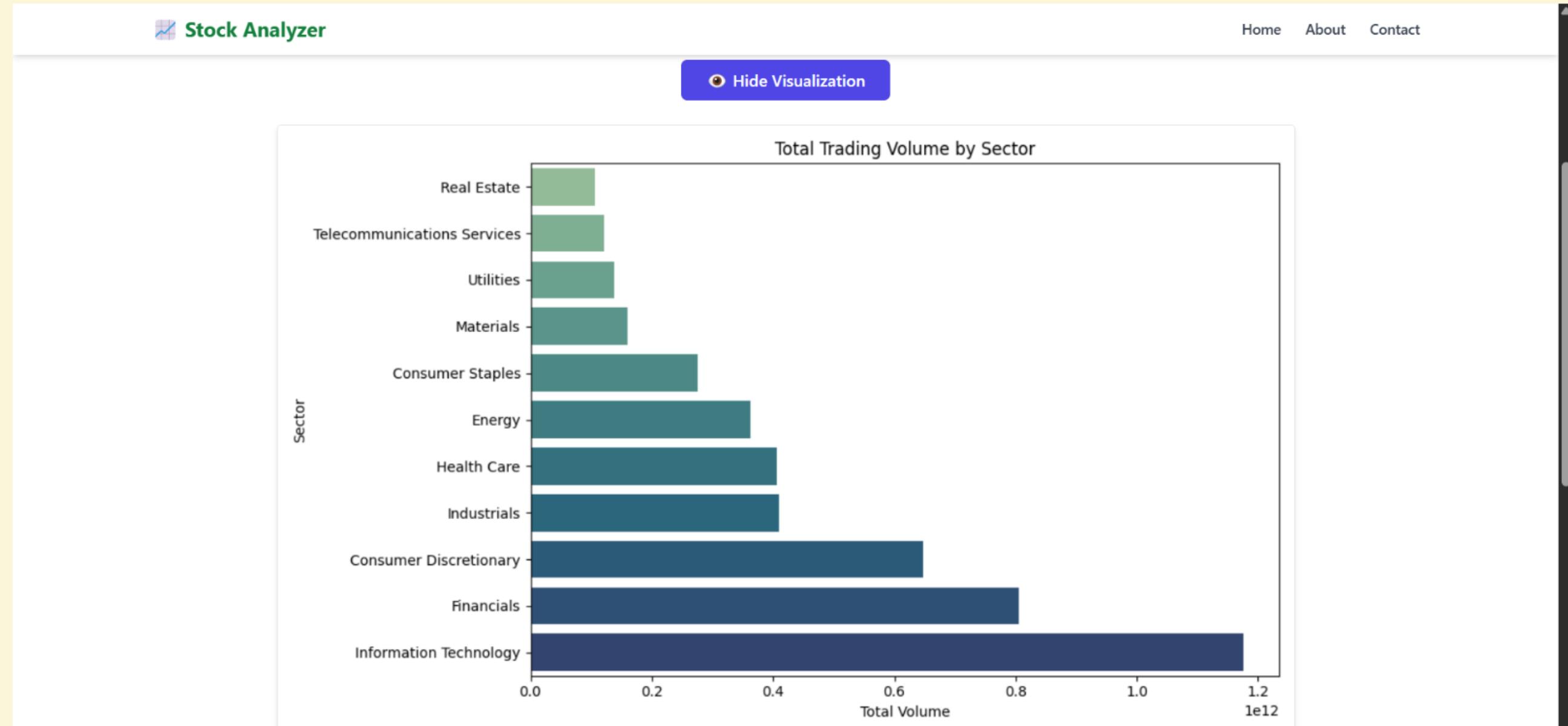
Two main Hive tables have been created: **stock_prices** and **stock_companies**. These tables serve as the primary structure for storing our stock market data.

Data Format Specification

The data is loaded into Hive in **TEXTFILE** format and is **comma-delimited**. This format is essential for efficient processing and querying.

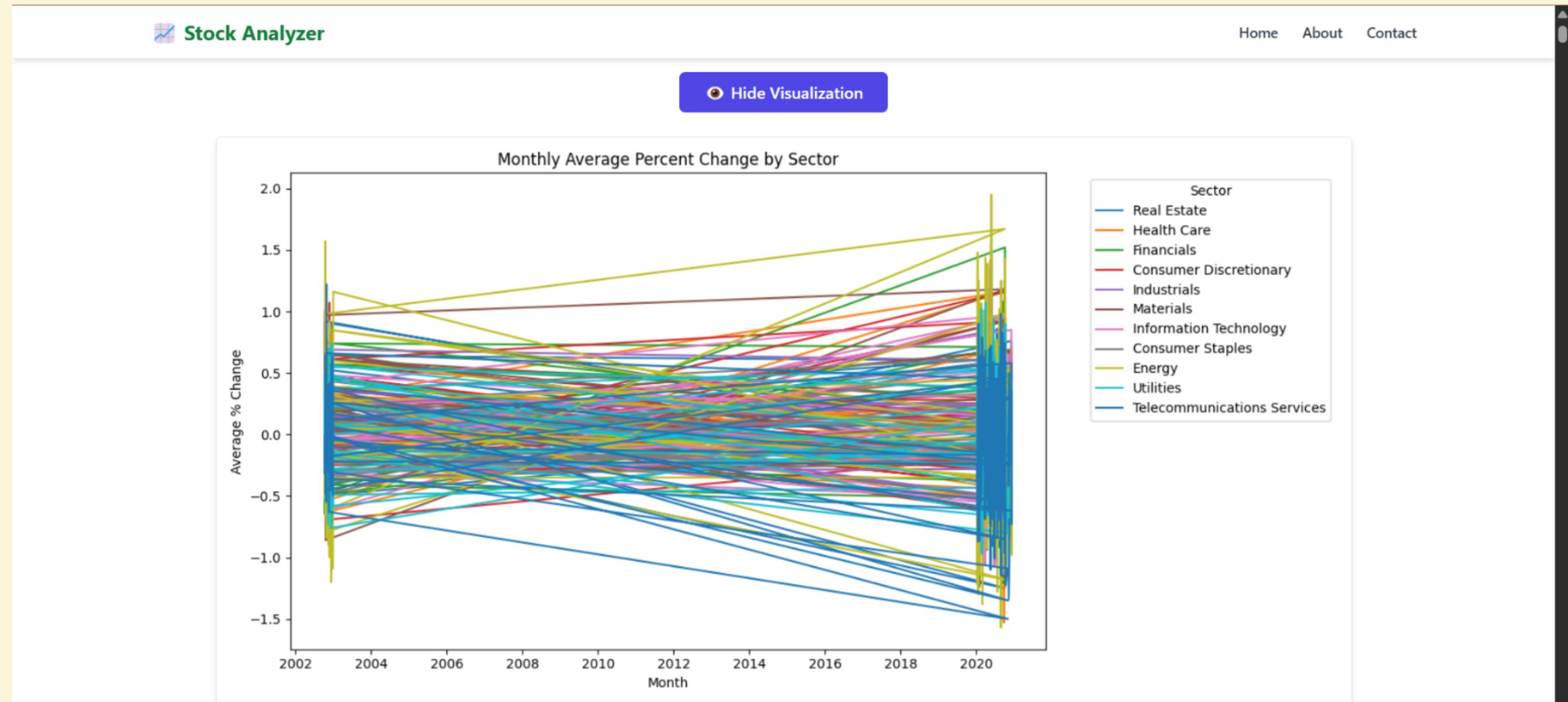
Analyzing Stock Market with Hive

Volume Analysis: Sector-wise totals - Analyzes total trading volume across different market sectors.



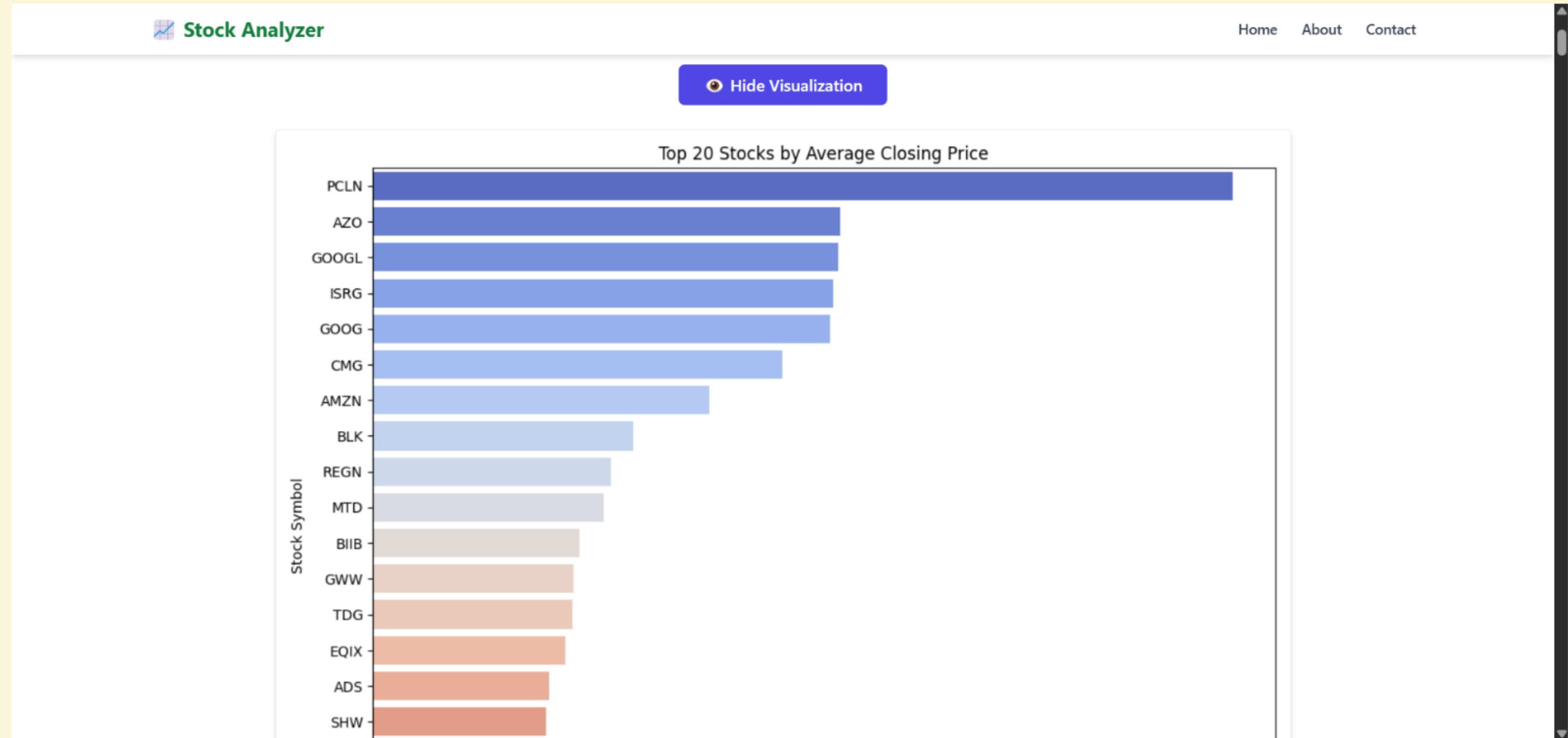
Analyzing Stock Market with Hive

Gainers Overview: Top percentage changes - Identifies stocks with the highest percentage gains.



Analyzing Stock Market with Hive

Close Prices: Average price trends - Tracks average closing prices over a defined period.



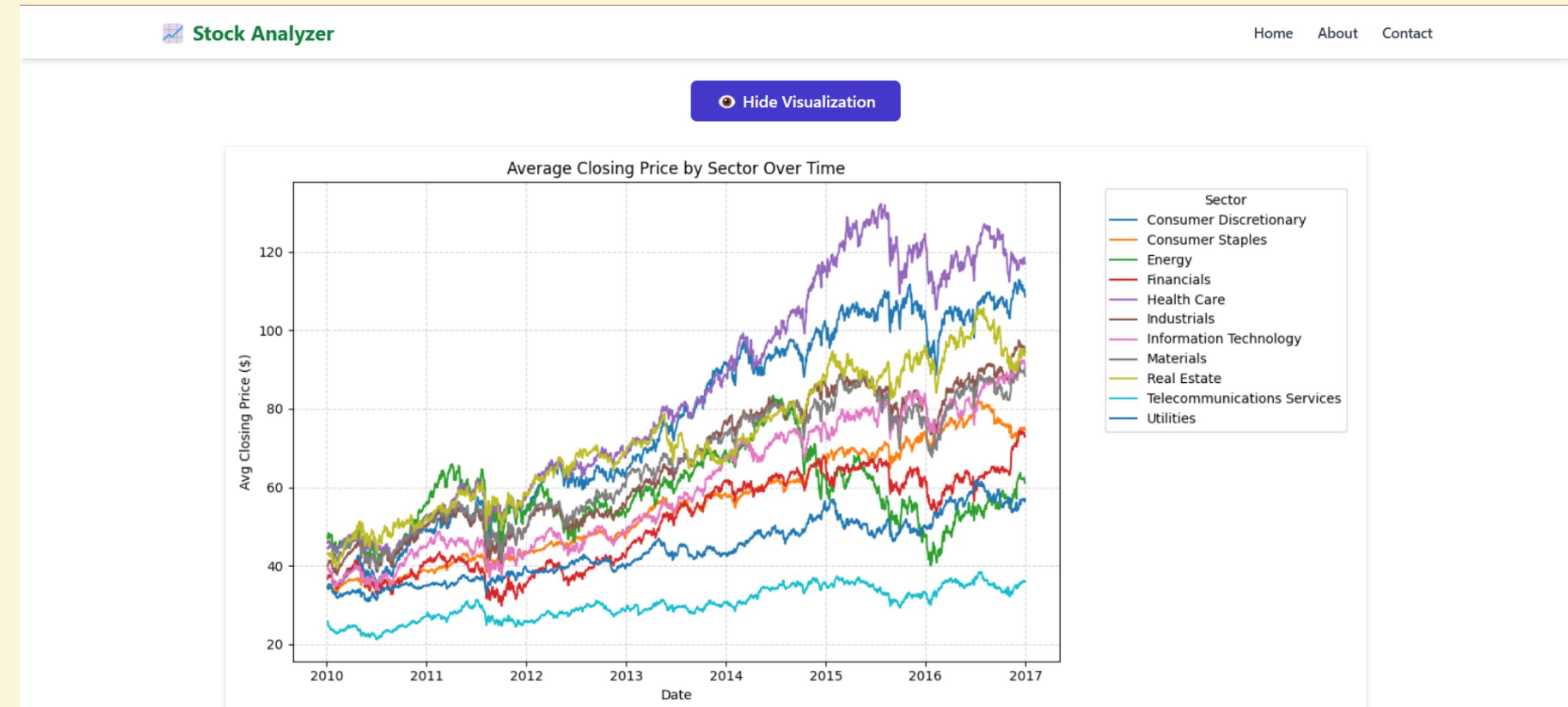
Analyzing Stock Market with Hive

Volatility Metrics: Daily price ranges - Measures daily price movements to assess market volatility.



Analyzing Stock Market with Hive

Meta Queries: Joins and counts - Utilizes Hive for complex joins and aggregations in stock data.



Query Examples for Stock Analysis

	Description	Example Query
Price Retrieval	Fetches current stock prices	<code>SELECT * FROM stock_prices WHERE date = CURRENT_DATE;</code>
Historical Data	Retrieves historical stock data	<code>SELECT * FROM stock_prices WHERE date BETWEEN '2023-01-01' AND '2023-12-31';</code>
Volume Analysis	Analyzes trading volume	<code>SELECT stock_symbol, SUM(volume) FROM stock_trades GROUP BY stock_symbol;</code>
Price Comparison	Compares prices between two stocks	<code>SELECT a.stock_symbol, a.price, b.price FROM stock_prices a JOIN stock_prices b ON a.stock_symbol = 'AAPL' AND b.stock_symbol = 'GOOGL';</code>

Integrating Django with React



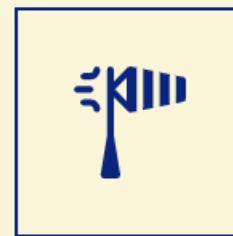
Django as Backend Framework

Utilizes Django to serve static files such as .txt, .csv, and .png from designated folders, ensuring efficient data handling.



React for Frontend Development

Employs React for building dynamic routes and reusable components, enhancing user interface interactivity.



Tailwind CSS for Styling

Integrates Tailwind CSS to create a visually appealing and responsive user interface, streamlining the design process.



Web Integration Benefits

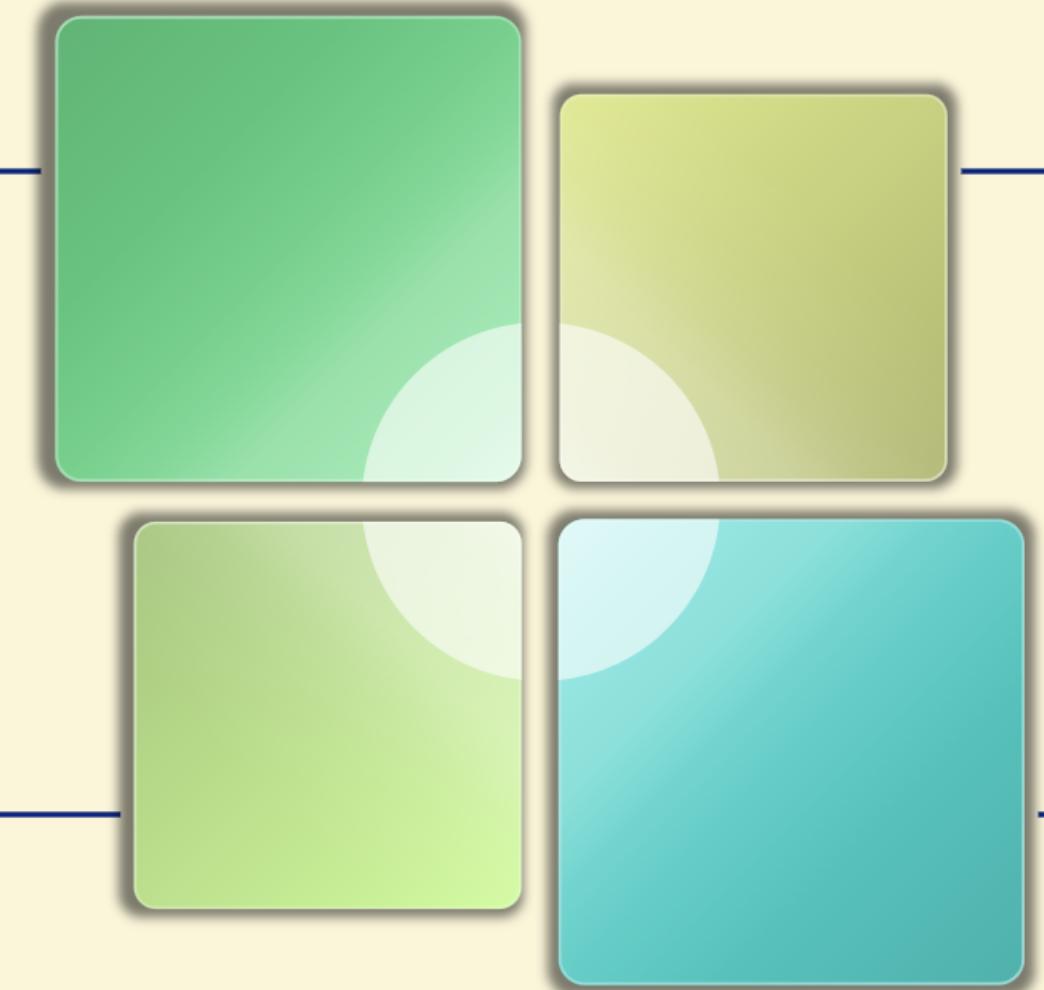
Combining Django and React provides a robust solution for handling complex data interactions and improving performance.

Web UI Preview for Stock Analysis

X

Homepage Structure

The **Homepage** displays a comprehensive **category list** to navigate through stock market sectors.



User Interaction

Users can **click** on categories to **show Hive queries**, facilitating targeted data analysis.

Data Output Visualization

The system generates an **output table** alongside a **plot chart**, presenting data in an informative manner.

Hive Query Efficiency

Utilizing **Hive queries** enhances the efficiency of data retrieval for stock analysis, ensuring quick access to insights.

TECH INSIGHTS

Analyzing Tech Sector Insights

Exploring the tech sector's peak at \$7.33E8, volatility trends, and sector patterns illustrated through detailed charts.



A close-up photograph of a person's hand holding a black picture frame. The frame is tilted slightly, revealing a white rectangular card inside. The card has a thin green border at the top and bottom. The text is centered on the card.

STOCK MARKET

Optimizing Stock Market Insights

Explore further optimizations in queries, enhance data visualization techniques, and investigate additional datasets for broader insights.