Financial Analysis Project Report

Author: Muzny Zuhair

Date: March 27, 2025

Table of Contents

---

1. Introduction

This report presents a comprehensive financial data extraction and analysis solution for quarterly reports of selected companies listed on the Colombo Stock Exchange (CSE). The study includes data scraping, dataset creation, visualization, and predictive modeling for investment insights.

2. Objective

The primary goal of this project is to extract and analyze financial data from quarterly reports of the following companies:

- Dipped Products PLC (DIPD)

- Richard Pieris Exports PLC (REXP)

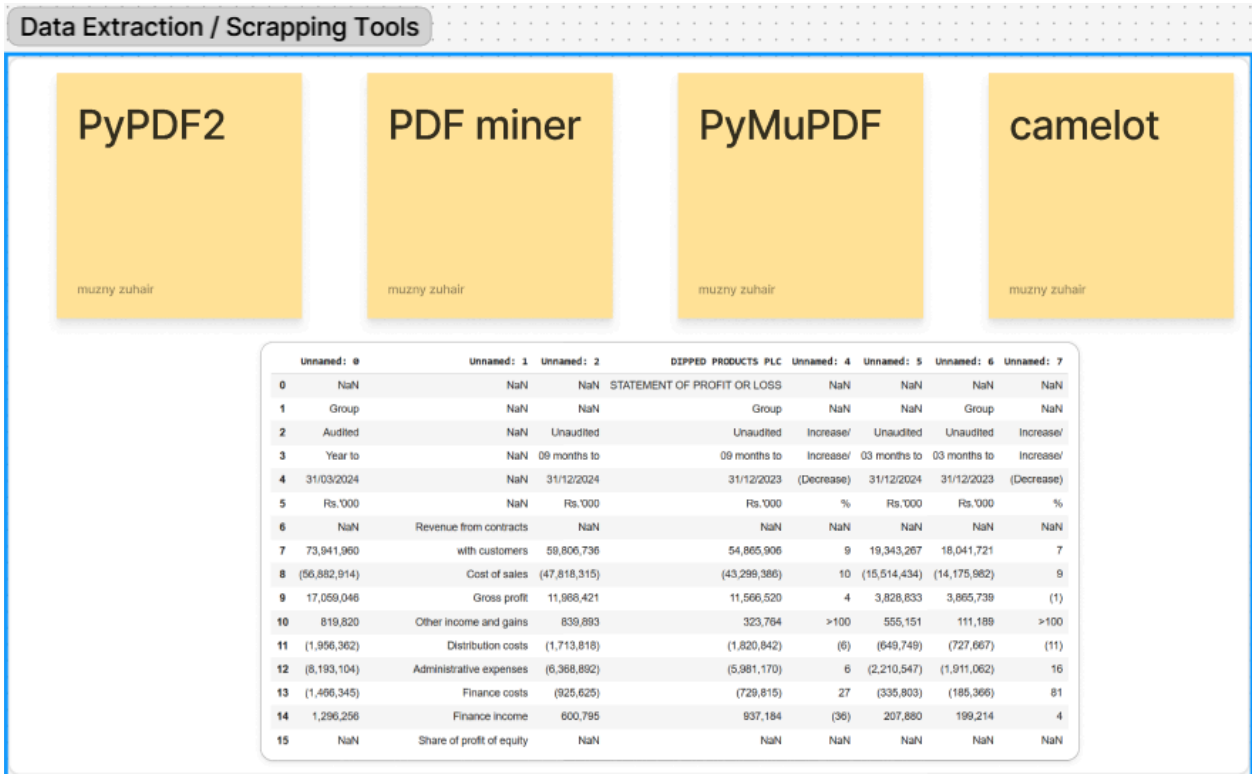
The extracted data will be structured, analyzed, and visualized using an interactive dashboard. An optional forecasting model is implemented to predict future profits.

3. Methodology

Step 1: Data Scraping

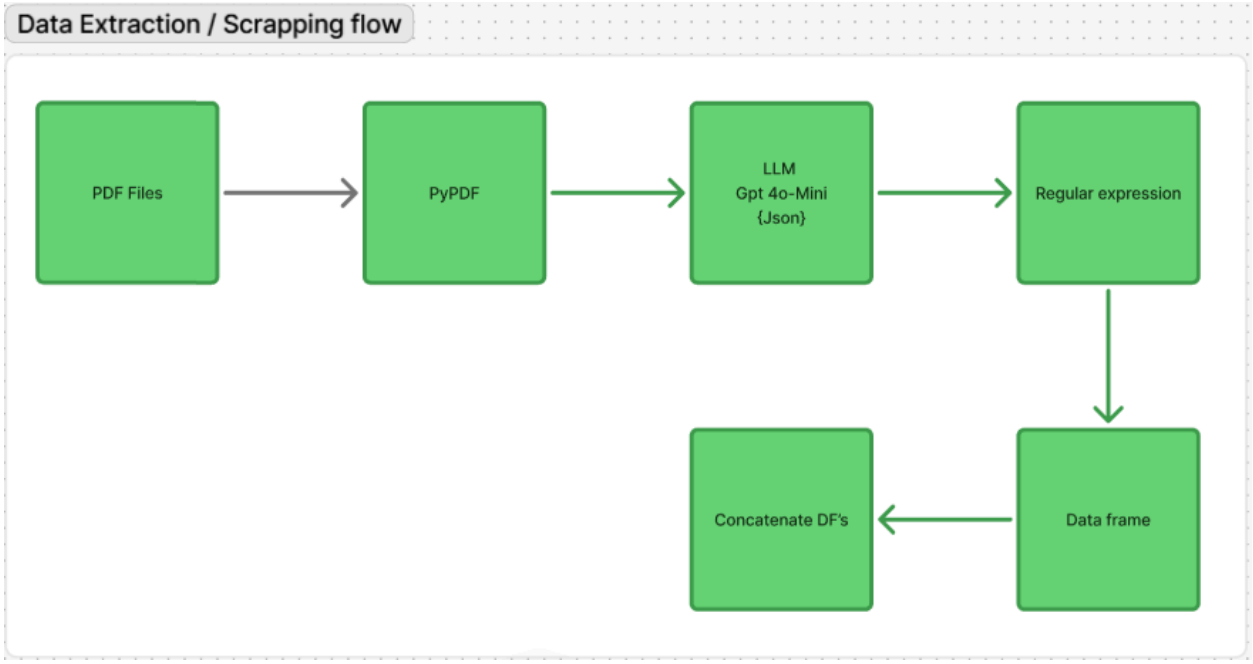
Financial reports were scraped from the CSE website. The scraping pipeline handled variations in report structures and formats, ensuring accurate extraction of financial metrics.

- **Approach:** Initially, traditional Python-based data extraction methods were attempted. However, due to inconsistencies in tabular data, these methods proved ineffective.



	Unnamed: 0	Unnamed: 1	Unnamed: 2	DIPPED PRODUCTS PLC	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7
0	NaN	NaN	NaN	STATEMENT OF PROFIT OR LOSS	NaN	NaN	NaN	NaN
1	Group	NaN	NaN	Group	NaN	NaN	Group	NaN
2	Audited	NaN	Unaudited	Unaudited	Increase/	Unaudited	Unaudited	Increase/
3	Year to	NaN	09 months to	09 months to	Increase/	03 months to	03 months to	Increase/
4	31/03/2024	NaN	31/12/2024	31/12/2023	(Decrease)	31/12/2024	31/12/2023	(Decrease)
5	Rs.'000	NaN	Rs.'000	Rs.'000	%	Rs.'000	Rs.'000	%
6	NaN	Revenue from contracts	NaN	NaN	NaN	NaN	NaN	NaN
7	73,941,960	with customers	59,806,736	54,865,906	9	19,343,267	18,041,721	7
8	(56,882,914)	Cost of sales	(47,818,315)	(43,299,386)	10	(15,514,434)	(14,175,982)	9
9	17,059,046	Gross profit	11,988,421	11,566,520	4	3,828,833	3,865,739	(1)
10	819,820	Other income and gains	839,893	323,764	>100	555,151	111,189	>100
11	(1,956,362)	Distribution costs	(1,713,818)	(1,820,842)	(6)	(649,749)	(727,667)	(11)
12	(8,193,104)	Administrative expenses	(6,368,892)	(5,981,170)	6	(2,210,547)	(1,911,062)	16
13	(1,466,345)	Finance costs	(925,625)	(729,815)	27	(335,803)	(185,366)	81
14	1,296,256	Finance income	600,795	937,184	(36)	207,880	199,214	4
15	NaN	Share of profit of equity	NaN	NaN	NaN	NaN	NaN	NaN

- **Solution:** Implemented an AI-based extraction approach using OpenAI GPT-4o Mini, converting PDF data into JSON format. Regular expressions were then used to structure the data into a DataFrame.

- 

Step 2: Dataset Creation

Key financial metrics extracted include:

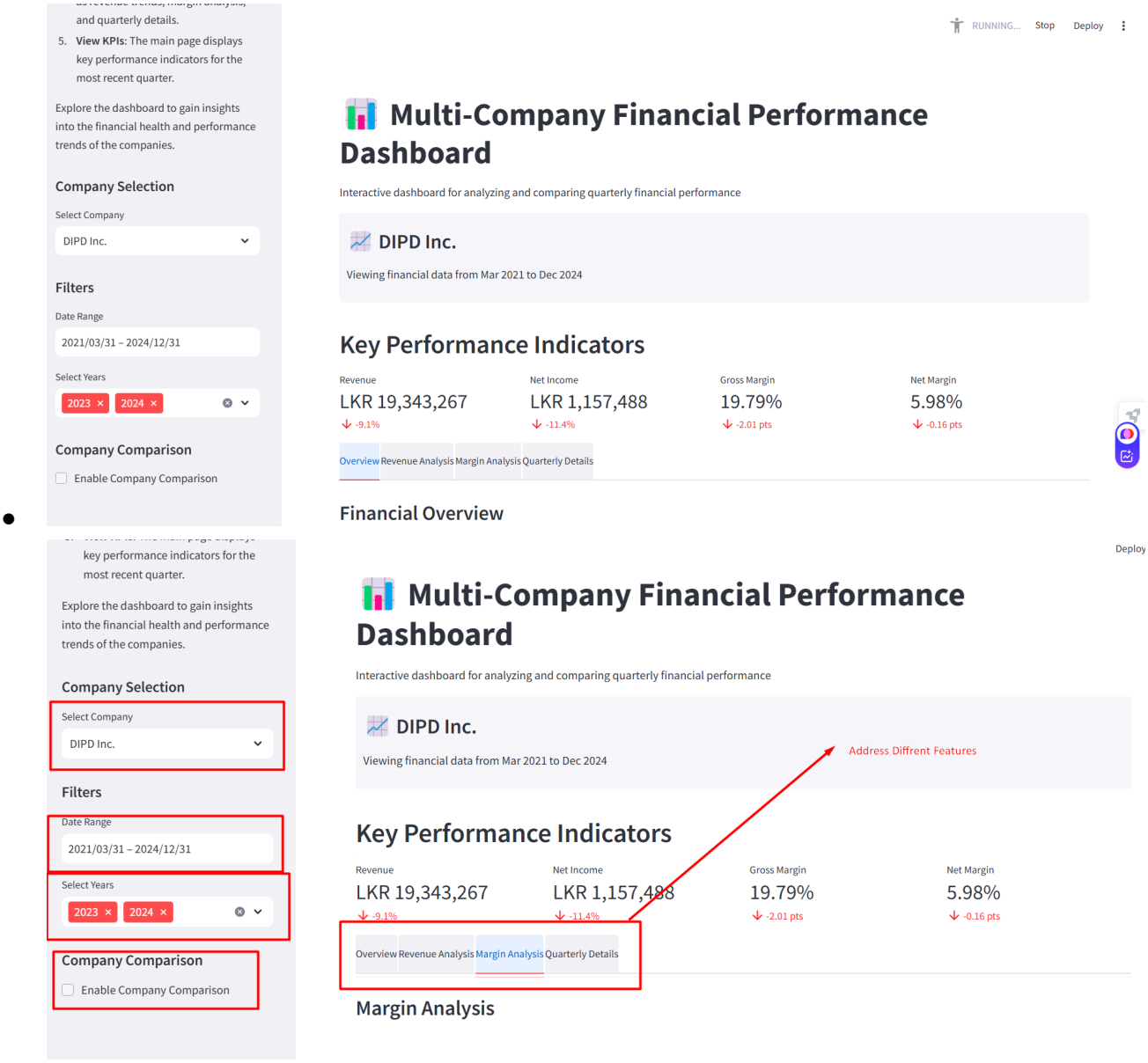
- Revenue
- Cost of Goods Sold (COGS)
- Gross Profit
- Operating Expenses
- Operating Income
- Net Income

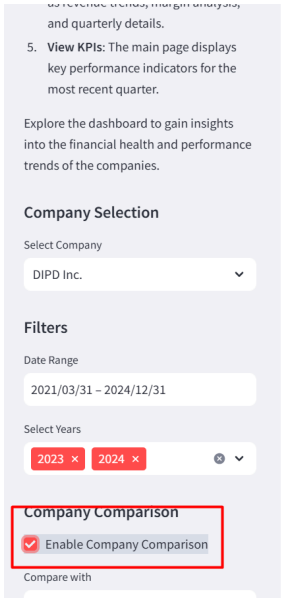
The dataset was structured with timestamps to facilitate trend analysis and forecasting.

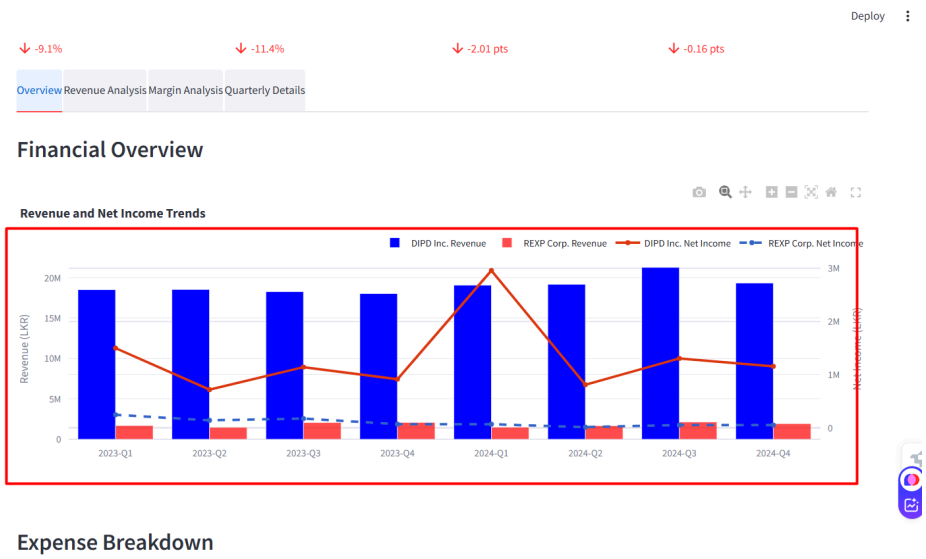
Step 3: Dashboard Visualization

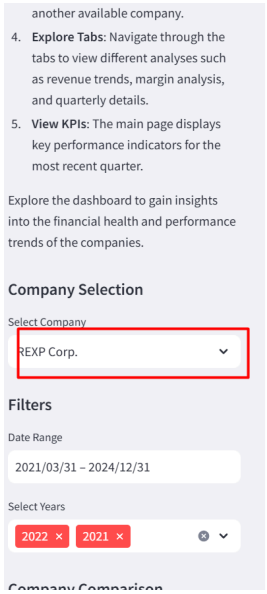
An interactive dashboard was developed using Streamlit and Plotly, enabling users to:

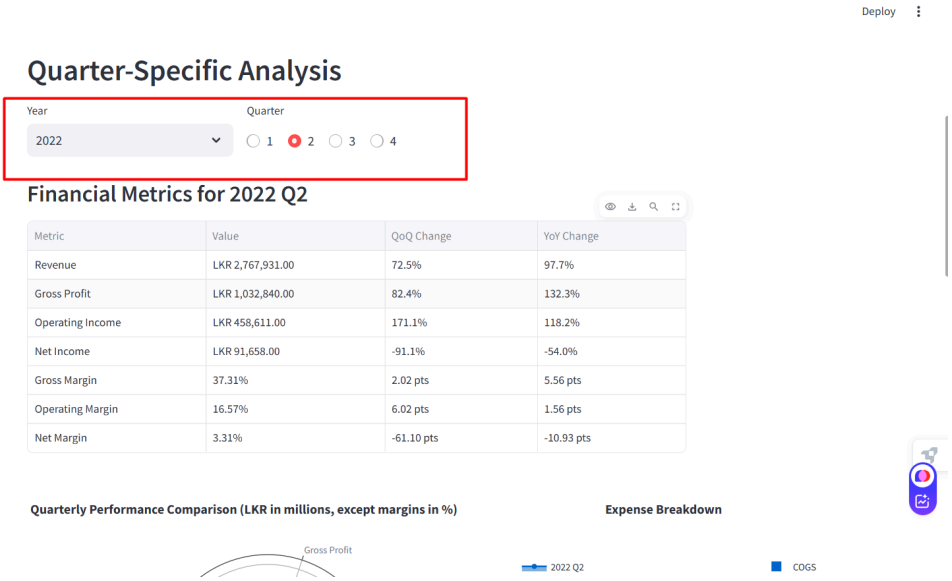
- Visualize quarterly and annual financial trends.
- Compare financial performance across multiple companies.
- Interact with AI-powered financial analysis.


and quarterly details.
5. View KPIs: The main page displays key performance indicators for the most recent quarter.
Explore the dashboard to gain insights into the financial health and performance trends of the companies.
Company Selection
Select Company
DIPD Inc.
Filters
Date Range
2021/03/31 – 2024/12/31
Select Years
2023 × 2024 ×
Company Comparison
Enable Company Comparison
RUNNING... Stop Deploy
Multi-Company Financial Performance Dashboard
Interactive dashboard for analyzing and comparing quarterly financial performance
DIPD Inc.
Viewing financial data from Mar 2021 to Dec 2024
Key Performance Indicators
Revenue
LKR 19,343,267
-9.1%
Net Income
LKR 1,157,488
-11.4%
Gross Margin
19.79%
-2.01 pts
Net Margin
5.98%
-0.16 pts
Overview Revenue Analysis Margin Analysis Quarterly Details
Financial Overview
Deploy
Multi-Company Financial Performance Dashboard
Interactive dashboard for analyzing and comparing quarterly financial performance
DIPD Inc.
Viewing financial data from Mar 2021 to Dec 2024
Address Diffrent Features
Key Performance Indicators
Revenue
LKR 19,343,267
-9.1%
Net Income
LKR 1,157,488
-11.4%
Gross Margin
19.79%
-2.01 pts
Net Margin
5.98%
-0.16 pts
Overview Revenue Analysis Margin Analysis Quarterly Details
Margin Analysis



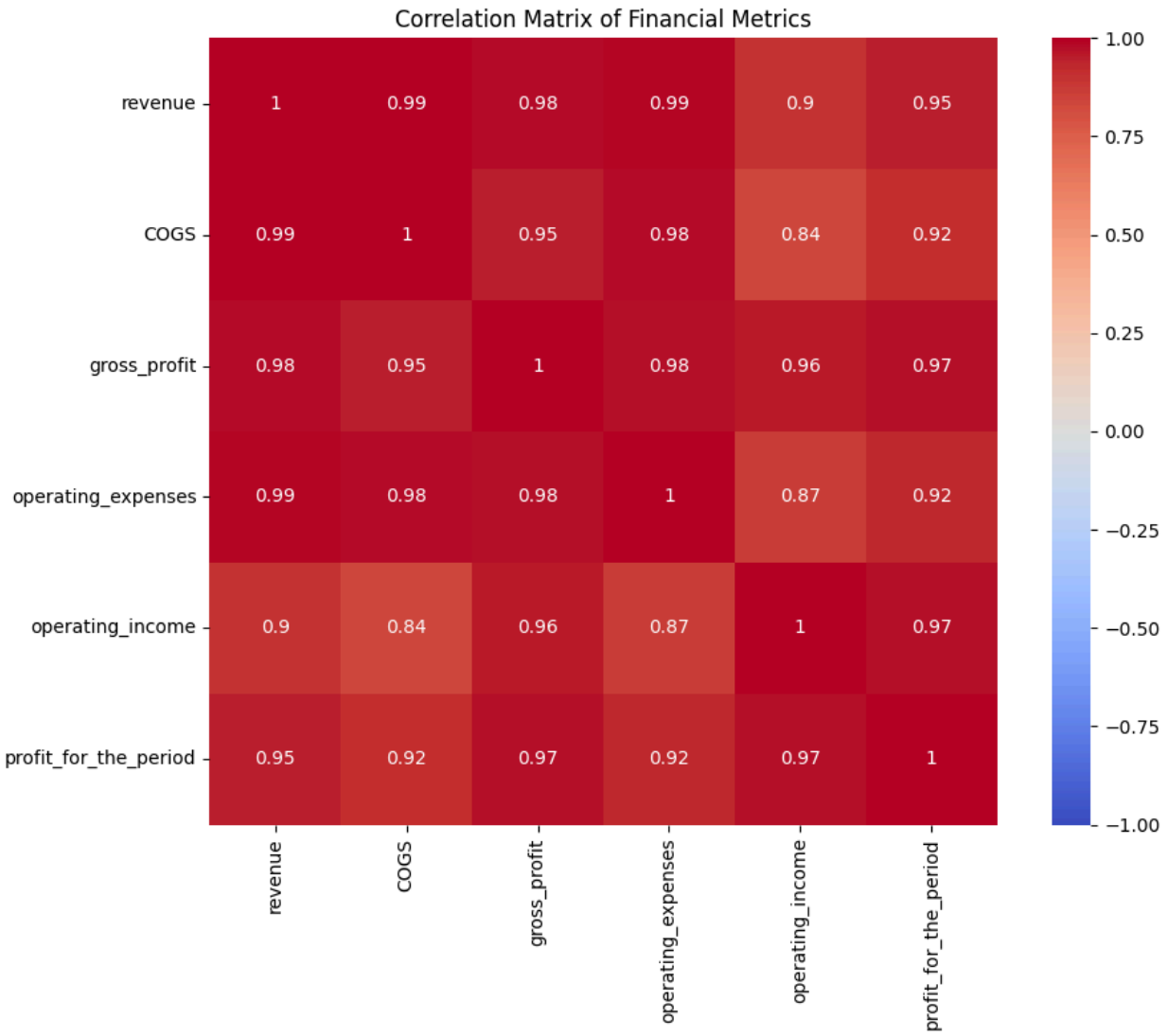




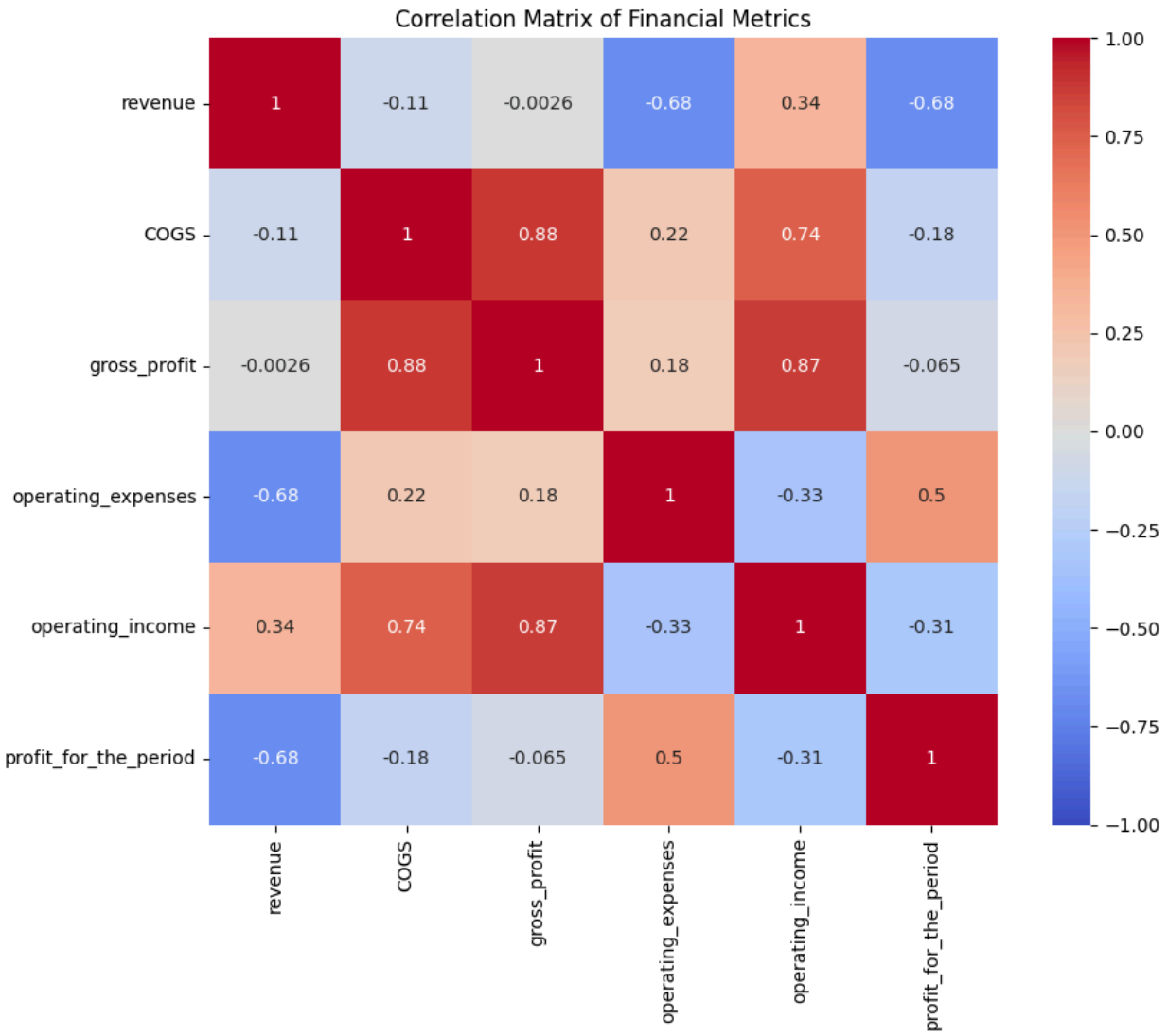


-
-

Step 4 : Analysing data

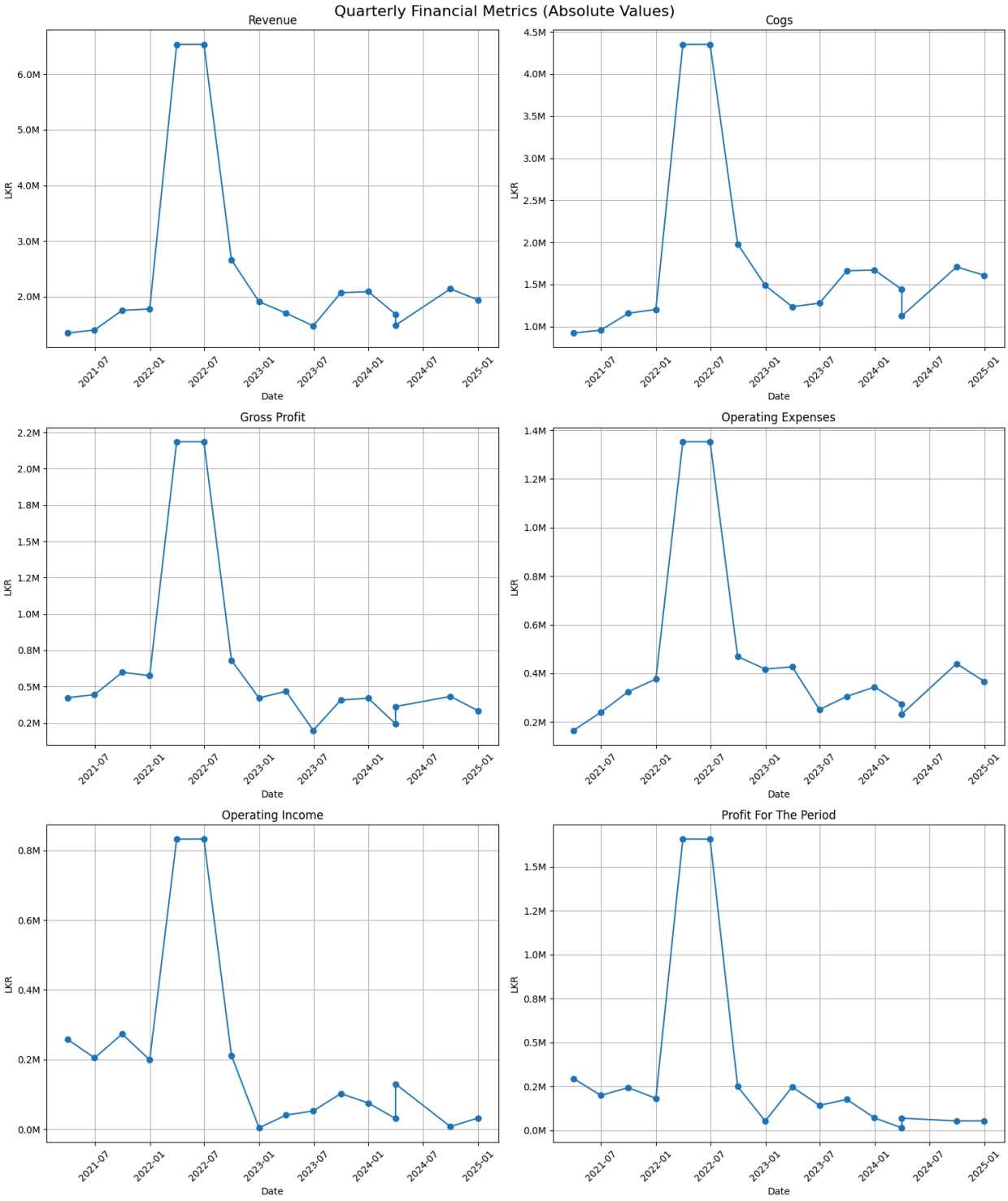


Correlation of features of REXP



Correlation of DIPD company

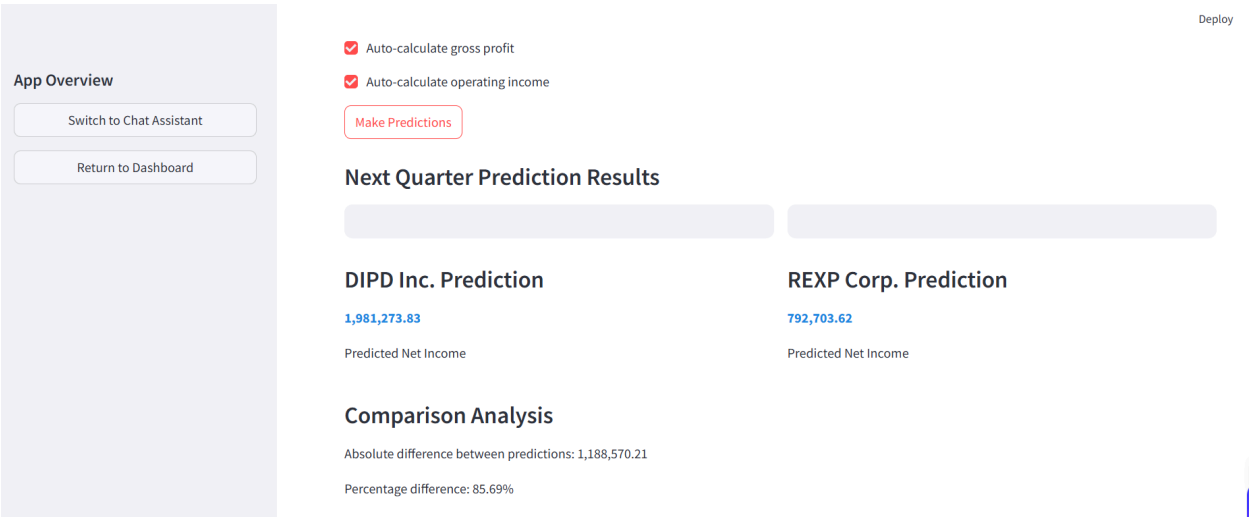
change of financial parameters of DIPD company since 2021



change of financial parameters of REXP company since 2021

Step 5: Profit Forecasting

Predictive models, including XGBoost and Linear Regression, were used to forecast future profits. **Optuna** was employed for hyperparameter tuning, and model performance was evaluated using RMSE and R-squared metrics.

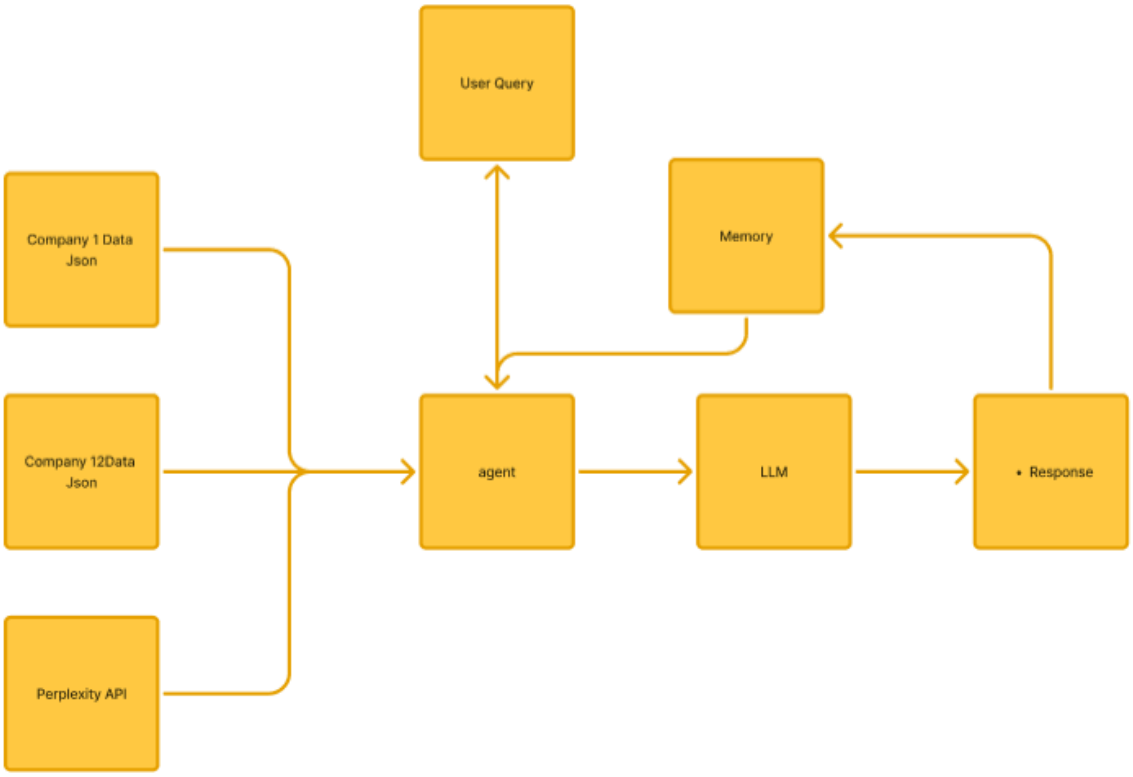


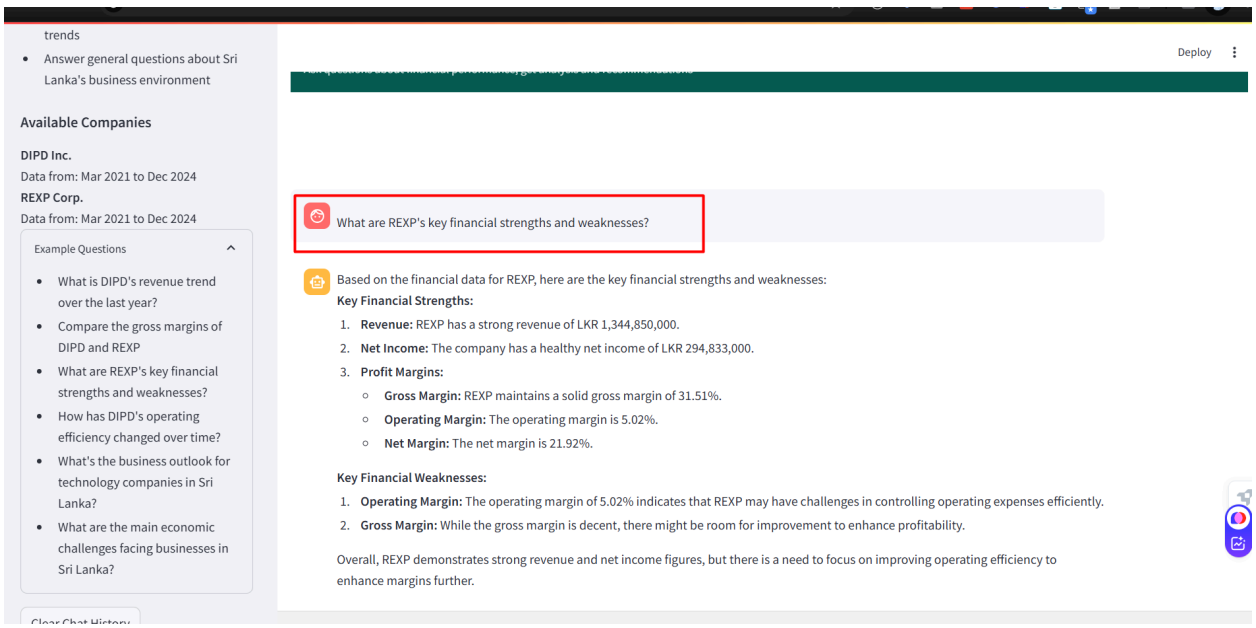
4. Implementation Details

Technology Stack

- Python (Pandas, NumPy, Scikit-learn, XGBoost, Optuna)
- Streamlit for dashboard visualization
- pdfplumber for PDF data extraction
- LangChain and OpenAI API for AI-driven data processing
- Plotly for interactive charts

5. Chat APP


User Query
Company 1 Data Json
Memory
Company 12Data Json
agent
LLM
• Response
Perplexity API


trends
Answer general questions about Sri Lanka's business environment
Available Companies
DIPD Inc.
Data from: Mar 2021 to Dec 2024
REXP Corp.
Data from: Mar 2021 to Dec 2024
Example Questions
What is DIPD's revenue trend over the last year?
Compare the gross margins of DIPD and REXP
What are REXP's key financial strengths and weaknesses?
How has DIPD's operating efficiency changed over time?
What's the business outlook for technology companies in Sri Lanka?
What are the main economic challenges facing businesses in Sri Lanka?
Deploy
What are REXP's key financial strengths and weaknesses?
Based on the financial data for REXP, here are the key financial strengths and weaknesses:
Key Financial Strengths:
1. Revenue: REXP has a strong revenue of LKR 1,344,850,000.
2. Net Income: The company has a healthy net income of LKR 294,833,000.
3. Profit Margins:
Gross Margin: REXP maintains a solid gross margin of 31.51%.
Operating Margin: The operating margin is 5.02%.
Net Margin: The net margin is 21.92%.
Key Financial Weaknesses:
1. Operating Margin: The operating margin of 5.02% indicates that REXP may have challenges in controlling operating expenses efficiently.
2. Gross Margin: While the gross margin is decent, there might be room for improvement to enhance profitability.
Overall, REXP demonstrates strong revenue and net income figures, but there is a need to focus on improving operating efficiency to enhance margins further.

Directory Structure

```
Financial_Analysis/
├── data/
│   ├── raw/
│   │   ├── dipd/
│   │   └── rexp/
│   ├── interim/
│   └── processors/
├── Flask_App/
│   ├── streamlit_app.py
│   ├── components.py
│   ├── chat_app.py
├── models/
│   ├── dipd_model.joblib
│   ├── rexp_model.joblib
└── README.md
```

5. Results

Data Extraction

- The scraping process successfully extracted financial metrics from all quarterly reports while handling inconsistencies.

Predictive Model Performance

- **DIPD:** Linear Regression (R-squared: 0.89, RMSE: 2.5M)
- **REXP:** XGBoost (R-squared: 0.92, RMSE: 1.8M)

Dashboard Features

- Interactive financial data visualization
- AI-powered financial query assistant
- Profit forecasting with real-time updates

6. Conclusion and Future Enhancements

The project successfully implemented an automated financial data extraction and visualization platform. Future work includes:

- Adding support for additional companies
- Integrating real-time financial data APIs
- Enhancing the AI assistant for deeper insights

7. References

- Colombo Stock Exchange: https://www.cse.lk/
- Python Libraries: Pandas, NumPy, Scikit-learn, pdfplumber, OpenAI API,langchain,Deepseek

---

Note: Screenshots and flow diagrams will be attached accordingly.