📊 Financial Report Analyzer

Fast, AI-powered analysis of financial reports using PDF parsing, NLP compression, and interactive dashboards

---

🧠 Problem Statement

Annual and quarterly financial reports often exceed 100+ pages and contain dense financial and narrative information. Manually reviewing these documents is time-consuming and difficult, especially when comparing multiple companies across periods. Extracting risks, trends, sentiment, and key financial signals requires significant effort.

There is a need for a lightweight system that can quickly compress, analyze, and visualize financial reports to support faster understanding and comparison.

💡 Solution Overview

Financial Report Analyzer is a 1-day hackathon-style Python project that analyzes uploaded financial report PDFs and combines document parsing, NLP compression, and market data analytics into a single Streamlit dashboard.

The application:

• Extracts text and tables from uploaded reports
• Compresses long narratives using ScaleDown-style summarization
• Performs sentiment analysis on financial narratives
• Extracts and summarizes risk factors
• Detects financial trends using Yahoo Finance data
• Compares peer companies
• Analyzes earnings call transcripts
• Displays insights in an interactive dashboard

The system works without SEC APIs and relies on uploaded PDFs and Yahoo Finance structured data.

🔑 Key Features

• PDF financial report parsing
• Narrative compression (target ~80% reduction)
• Financial table extraction
• Risk factor detection and summarization
• Sentiment analysis
• Earnings transcript analysis
• Trend detection using Yahoo Finance
• Peer company comparison
• Rule-based alert generation
• Interactive Streamlit dashboard

⚙️ Tech Stack

• Python
• Streamlit
• pdfplumber
• camelot
• pandas
• yfinance
• transformers
• scikit-learn
• matplotlib / plotly

🏗️ System Architecture

User PDF Upload
→ Text & Table Extraction
→ Narrative Compression
→ NLP Analysis (Sentiment + Risk)
→ Market Trend Data (Yahoo Finance)
→ Peer Comparison
→ Alerts
→ Dashboard Visualization

🔁 Working Flow

1. User uploads a financial report PDF
2. System extracts raw text and tables
3. Narrative sections are chunked and summarized
4. Sentiment analysis runs on compressed content
5. Risk factor sections are detected and summarized
6. Historical stock data is fetched from Yahoo Finance
7. Peer tickers are analyzed and compared
8. Alert rules are evaluated
9. Insights are shown across dashboard sections

📊 Dashboard Sections

Filing Summary

• Compressed narrative summary
• Overall sentiment score
• Key extracted themes

Risk Factors

• Risk section summary
• Top risk keywords
• Highlighted major risks

Tables

• Extracted financial tables
• Structured dataframe previews
• Multi-page table support

Trends

• Historical stock price charts
• Moving averages

• Volatility indicators
• Trend direction signals

Peer Comparison

• Side-by-side ticker comparison
• Growth and volatility charts
• Relative performance metrics

Transcript Analysis

• Earnings call transcript input
• Auto-generated summary
• Sentiment score
• Theme keywords

Alerts

• Negative sentiment flags
• High volatility flags
• Risk-heavy language flags
• Downward trend signals

⚠️ Limitations

• Uses simplified NLP summarization models
• PDF parsing depends on document formatting
• Table extraction may fail on complex layouts
• Sentiment model is general-purpose, not finance-trained
• Risk detection is pattern-based, not fully semantic
• Alerts are rule-based, not predictive

🔮 Future Improvements

• Live filings API integration
• Finance-tuned NLP models
• Better semantic risk extraction
• Vector search across reports
• Multi-report comparison
• Advanced ratio analysis
• Background processing & caching
• Report-to-report difference detection