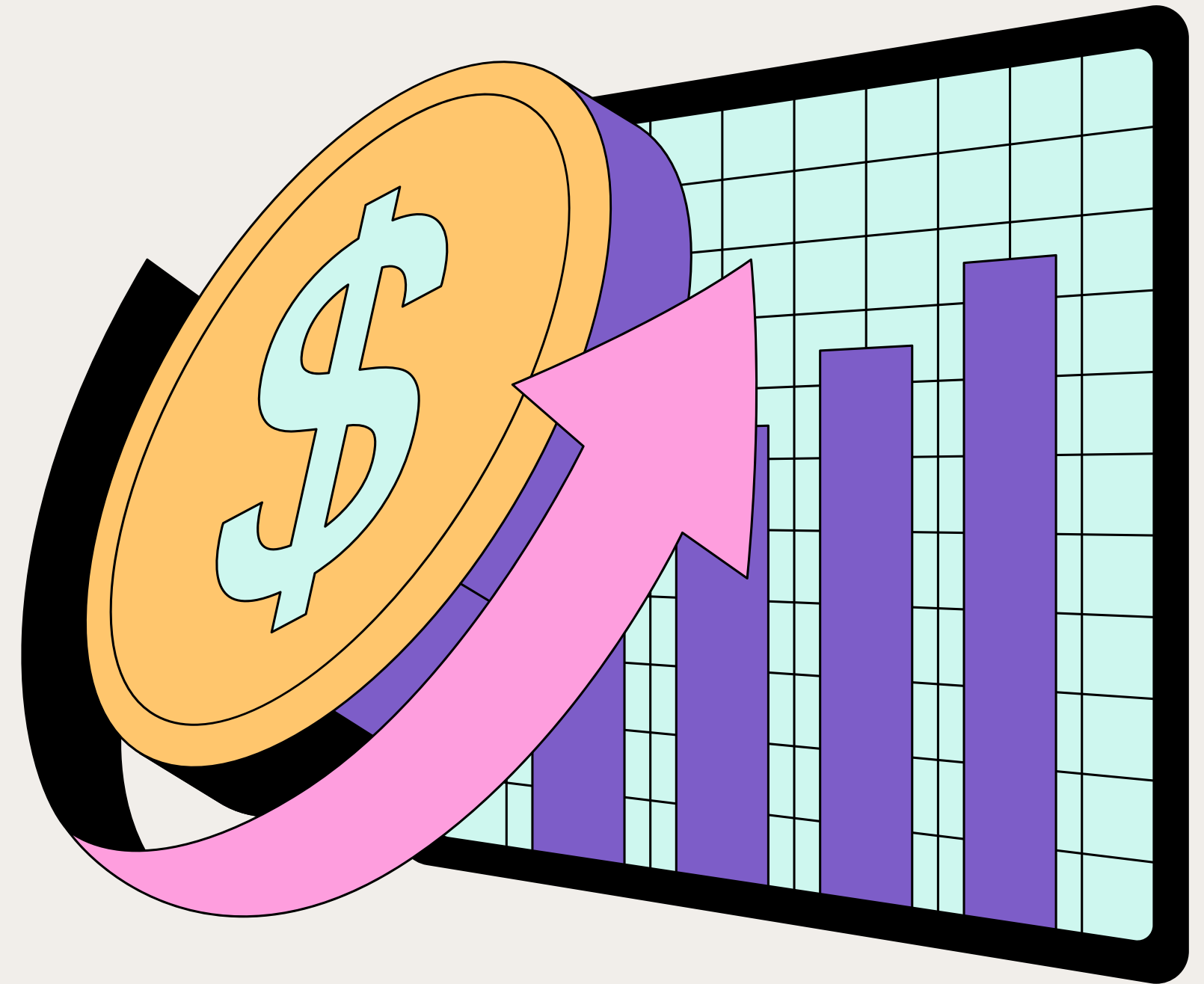


Understanding The

Quant Copilot!

TEAM: SAGORIKA, PRIYAL, BALAJI, SAIKIRAN, ROHITH

Powered by Databricks





Problem Statement

- Equity research is fragmented across fundamentals, macro, sentiment, and alternative data.
- PMs and analysts lose time stitching CSVs, SQL, and notebooks for basic answers.
- Retail investors rarely see factor exposures, red flags, or macro credit context.

We wanted a single place where you can ask a question about a stock and get a factor-aware answer grounded in data, not vibes.

[HOME](#)[ABOUT](#)[CONTENT](#)[OTHERS](#)

Project Overview

- AI-Powered Investment Assistant: Built on Databricks, answers real investor questions using historical financial data.
- Integrated Insights: Combines market trends, sentiment, and consumer signals into a single conversational interface.
- QuantCopilot for Stock Research: LLM-driven tool providing analysis from alternative data and macro spreads to plain-language investment advice.

Powered by Databricks



[HOME](#)[ABOUT](#)[CONTENT](#)[OTHERS](#)

Data Sources

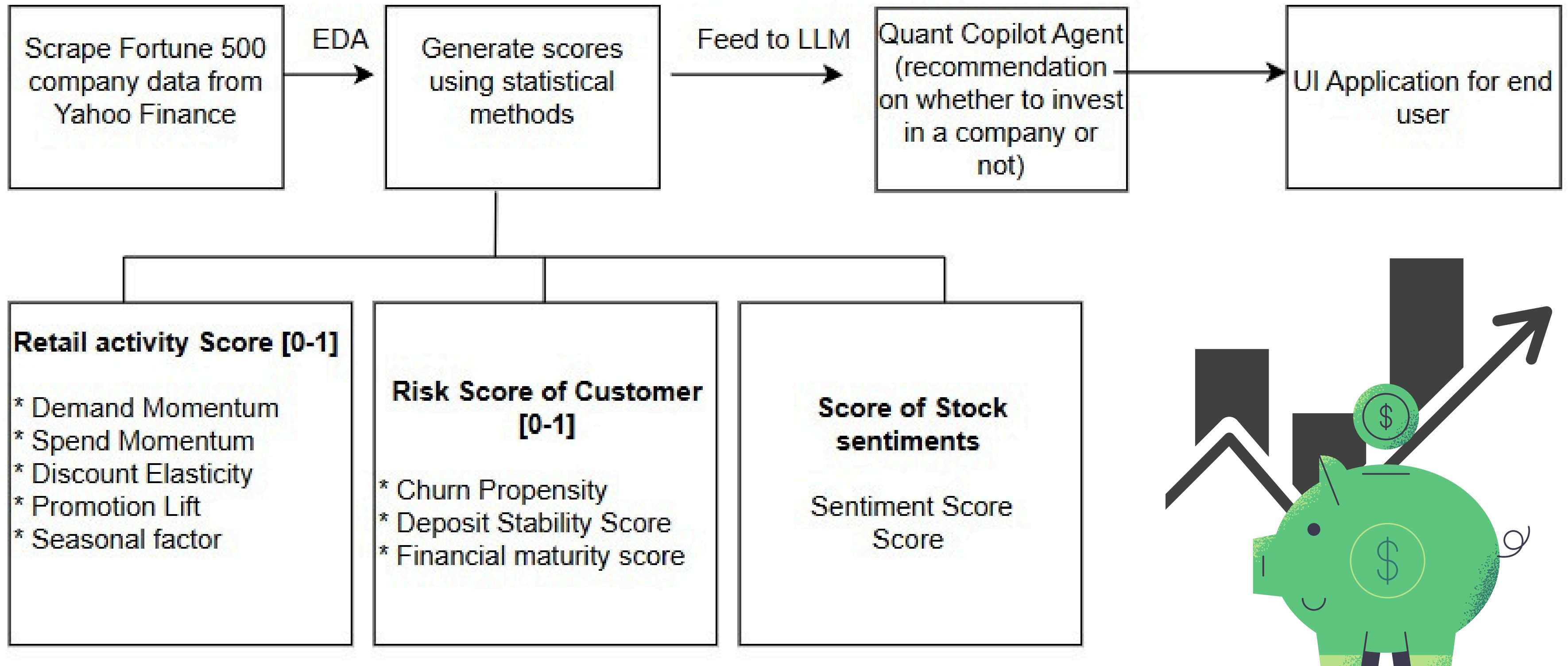
- Retail & Consumer Data: OpenDataBay Synthetic Retail Dataset & Investment Habit Analysis Dataset
- Market Sentiment: Stock Market Twitter Sentiment Dataset (Kaggle)
- Financial & Market Fundamentals: Yahoo Finance API, FRED API, Corporate Bond Yield Spreads (BAA–AAA)
- Fortune 500 company data scrapped using the Python API: yfinance

Powered by Databricks



Process Flow

Powered by Databricks



[HOME](#)[ABOUT](#)[CONTENT](#)[OTHERS](#)

Workflow Overview

- Data Collection: Scrape Fortune 500 company data from Yahoo Finance
- EDA & Feature Engineering: Generate scores using statistical methods
- Score Categories:
 - Retail Activity Score [0–1]: Demand Momentum, Spend Momentum, Discount Elasticity, Promotion Lift, Seasonality Factor
 - Customer Risk Score [0–1]: Churn Propensity, Deposit Stability, Financial Maturity
- Stock Sentiment Score: Sentiment Score from tweets & other sources
- LLM Advisory: Feed scores to LLM for investment guidance per company
- User Interface: Query the model via a chatbot



Powered by Databricks

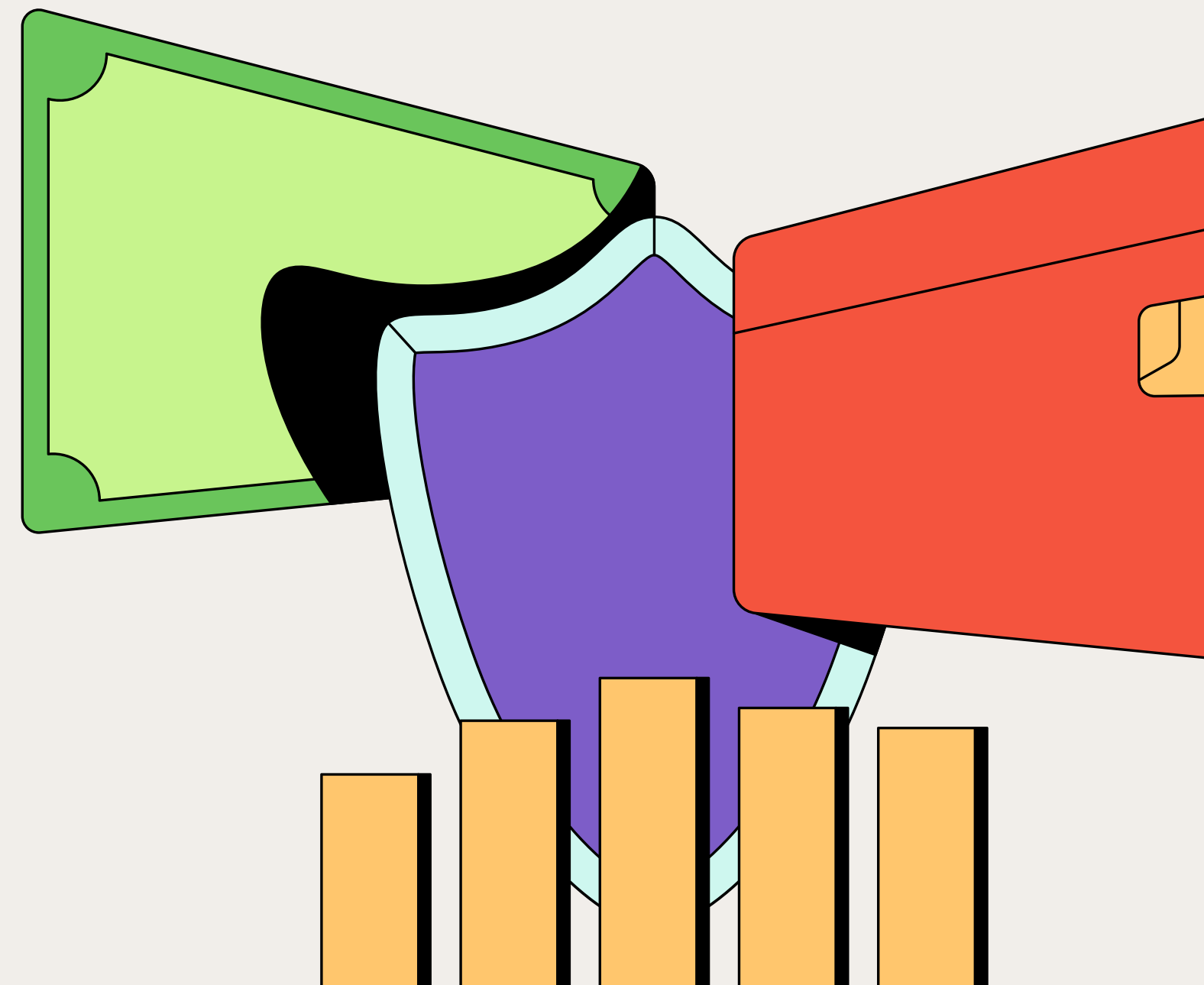
[HOME](#)[ABOUT](#)[CONTENT](#)[OTHERS](#)

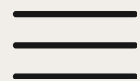
Single stock view

- User enters a ticker and natural language question.
- App fetches full context and calls the LLM.
- Output sections: high-level view, positives, risks, macro/scenarios, final verdict.
- Radar chart shows value, growth, quality, and inverted risk.

App UX: Single Stock

Powered by Databricks



[HOME](#)[ABOUT](#)[CONTENT](#)[OTHERS](#)

App UX – Screener and Scenario Lab

- Natural language to SQL screener on yfin_with_anomalies.
- LLM explains why top results match the screen.
- Scenario Lab creates bull/base/bear narratives using factor scores and spreads.
- Simple Monte Carlo chart shows intuitive path variability.

Powered by Databricks



[HOME](#)[ABOUT](#)[CONTENT](#)[OTHERS](#)

- User uploads daily prices across tickers.
- Computes returns and simple inverse variance weights.
- Shows approximate annual return and volatility, plus cumulative return chart

Portfolio Lab

Powered by Databricks

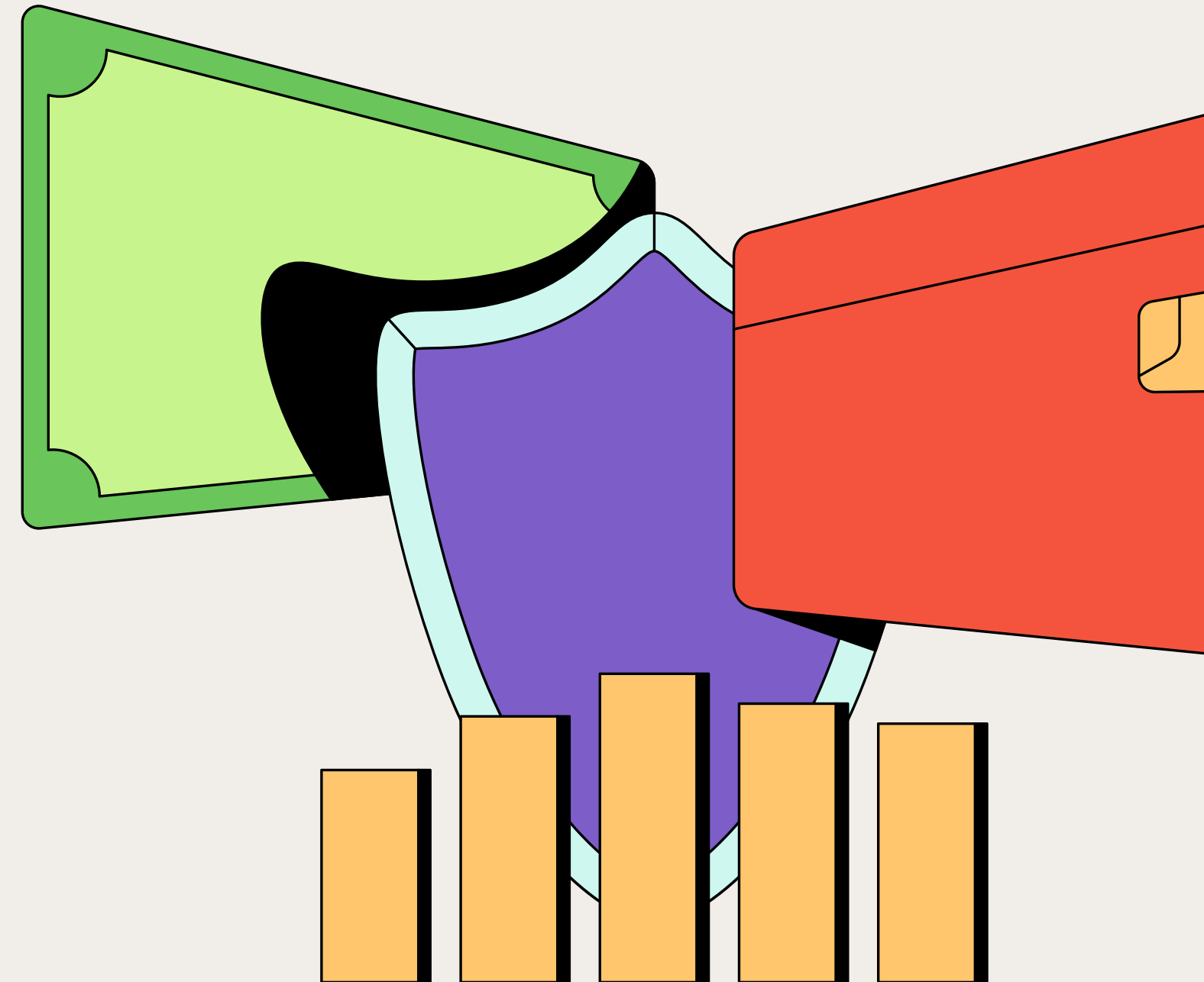
[HOME](#)[ABOUT](#)[CONTENT](#)[OTHERS](#)

What do you get?!

Value for real PMs

- Clear explanation of what drives signal for a stock.
- Early warning on leverage, volatility, and stress fragility.
- Macro credit context linked to company-specific risk.
- Fast ranking via natural language screens.
- Auto-generated memos and export to PDF.

Powered by Databricks



[HOME](#)[ABOUT](#)[CONTENT](#)[OTHERS](#)

Feedback

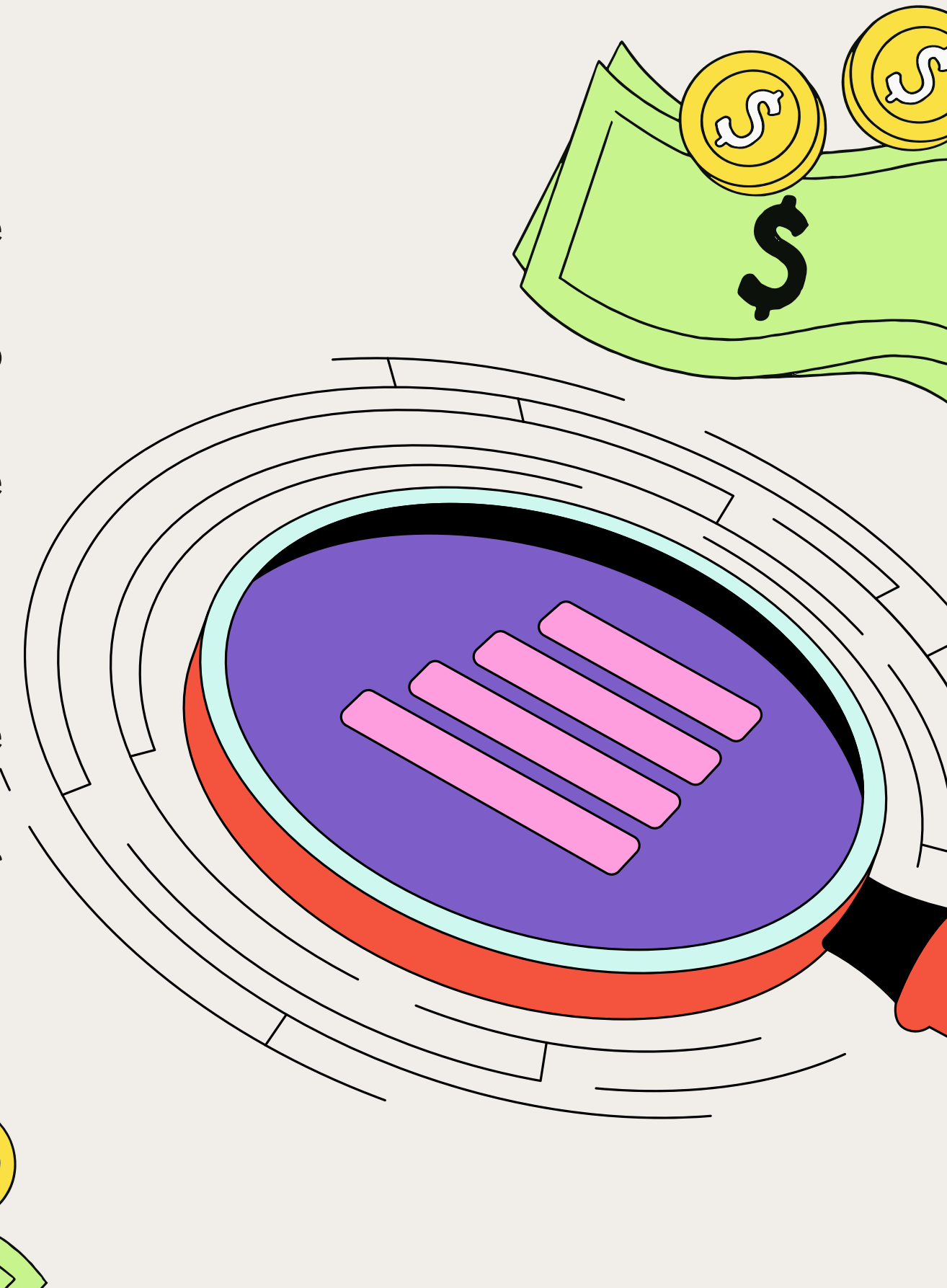
Pros

1. Easy to spin up notebooks and try out ideas fast, which helps when time is tight.
2. Built-in support for Spark, MLflow, and Delta Lake makes it simple to handle messy data or try quick ML experiments.
3. Collaboration is smooth since everyone can work in the same workspace without setting up anything locally.

Cons

1. Platform stopped working for 5-10 min at a stretch when multiple members were trying to share and work on the same workspace.
2. The interface has a learning curve if you haven't used Spark or Databricks before.

Powered by Databricks





[HOME](#)

[ABOUT](#)

[CONTENT](#)

[OTHERS](#)

Thank You!

Powered by Databricks

