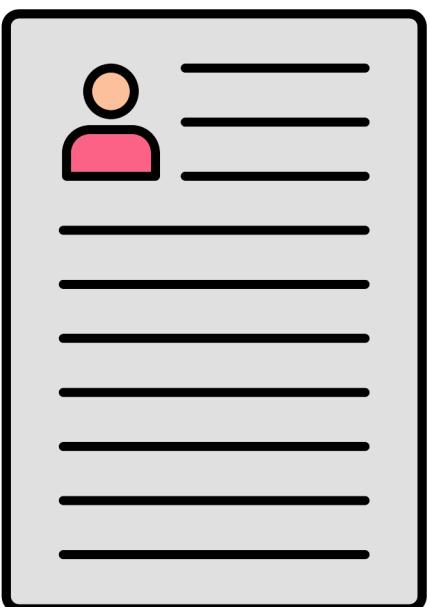


REALTIME STOCK MARKET DATA STREAMING AND PORTFOLIO OPTIMIZATION



BIG DATA ANALYSIS PROJECT

Team Members
Zane Falcao (9603)
Alroy Pereira (9631)



FEATURES

01. CLASSIFICATION & ASSET ALLOCATION

02. CUSTOMIZE PORTFOLIO PLAYGROUND

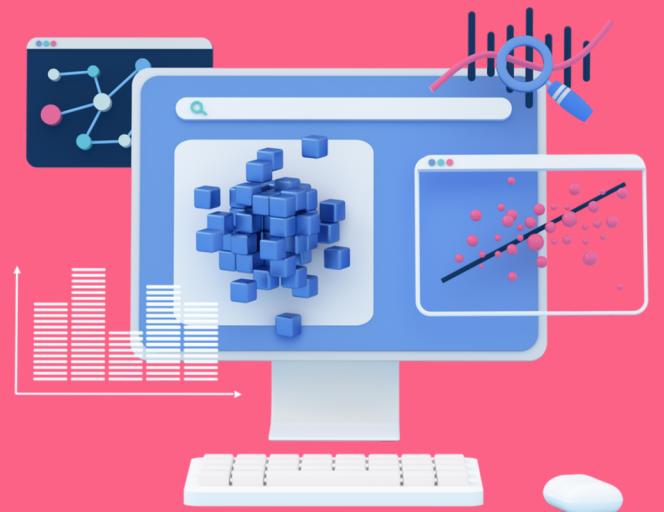
03. REAL-TIME DATA STREAMING

04. PORTFOLIO REBALANCING



MONTE CARLO SIMULATION

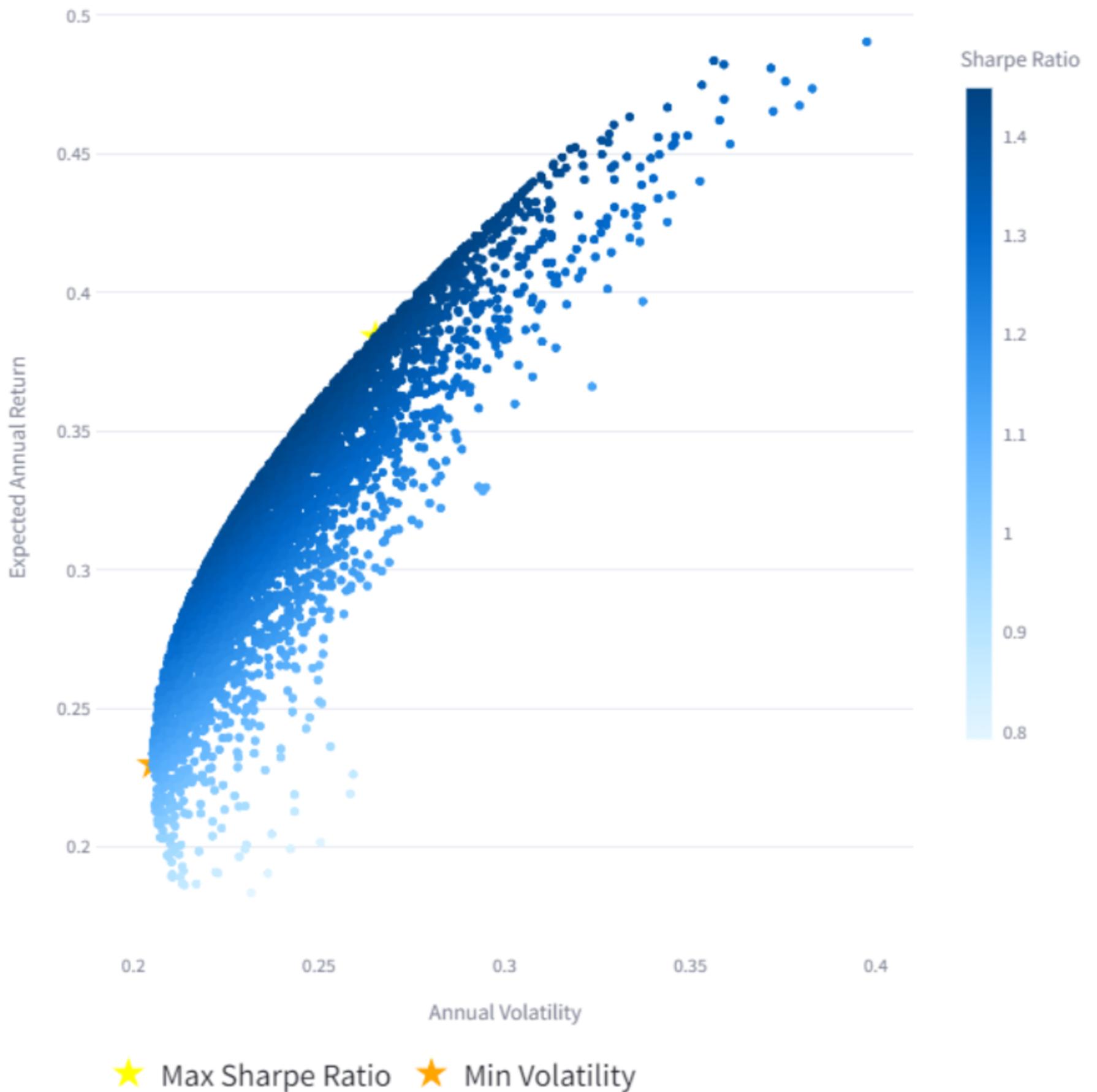
- Monte Carlo uses random sampling to estimate outcome distributions.
- Commonly used in finance to evaluate portfolio returns and risks.
- Generates random samples of possible stock price paths.
- Estimates expected return and risk from samples.
- Repeated sampling produces a distribution of outcomes.
- Provides insights for informed investment decisions.
- Stress-testing by simulating adverse conditions (e.g., recessions, crashes), helping assess how portfolios might perform under extreme market stress.



MONTE CARLO SIMULATION

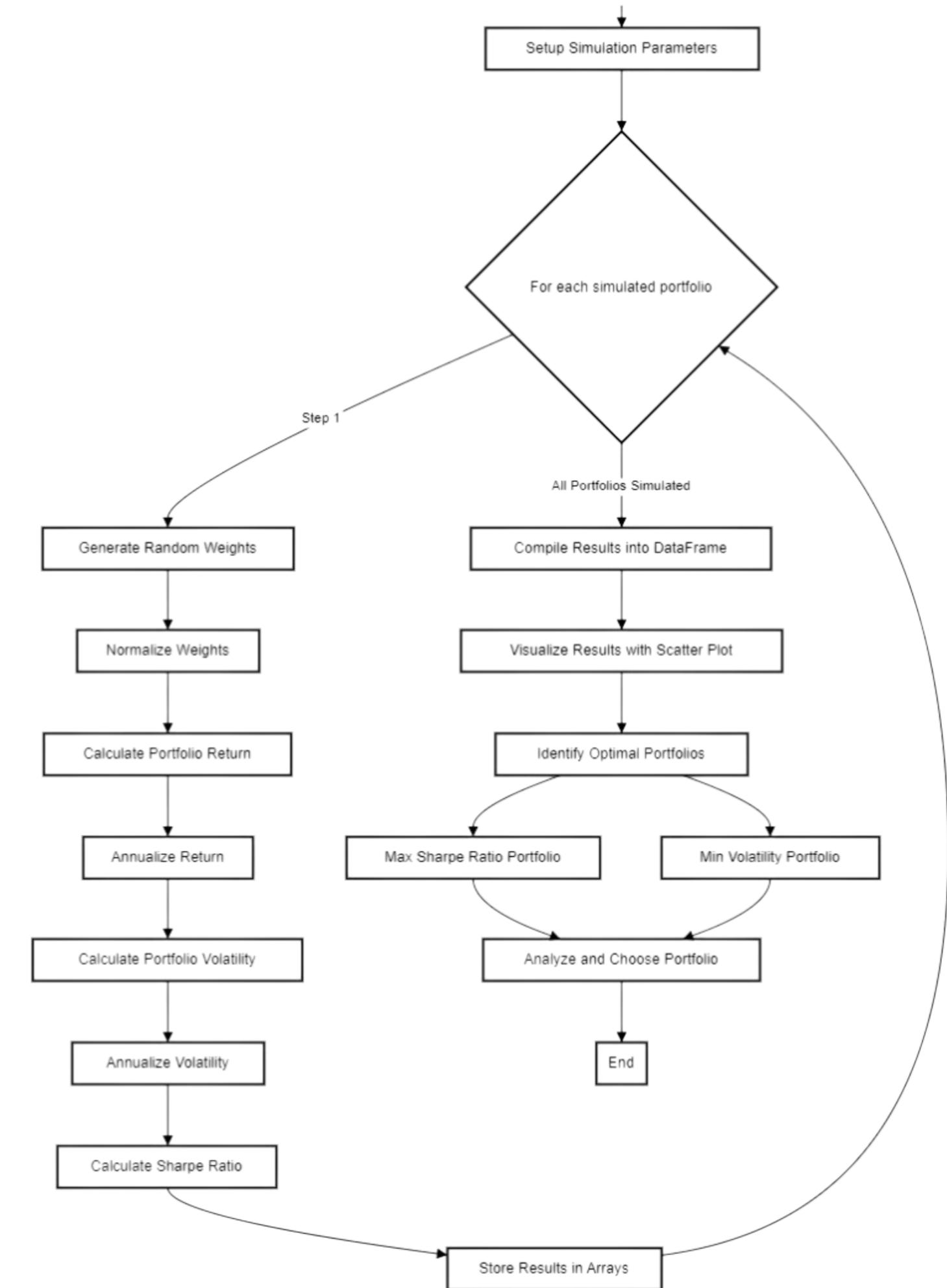
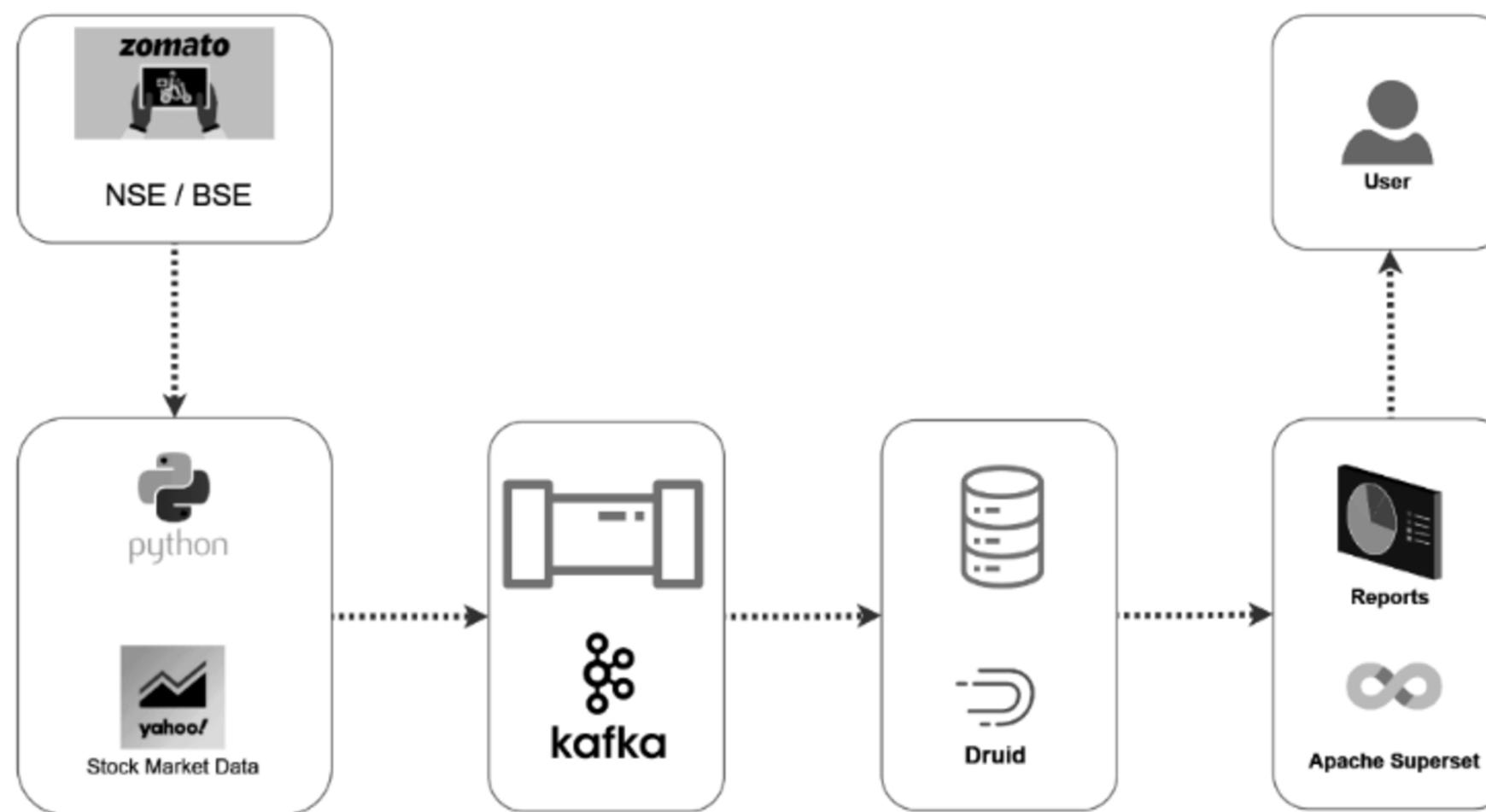


Monte Carlo Simulation

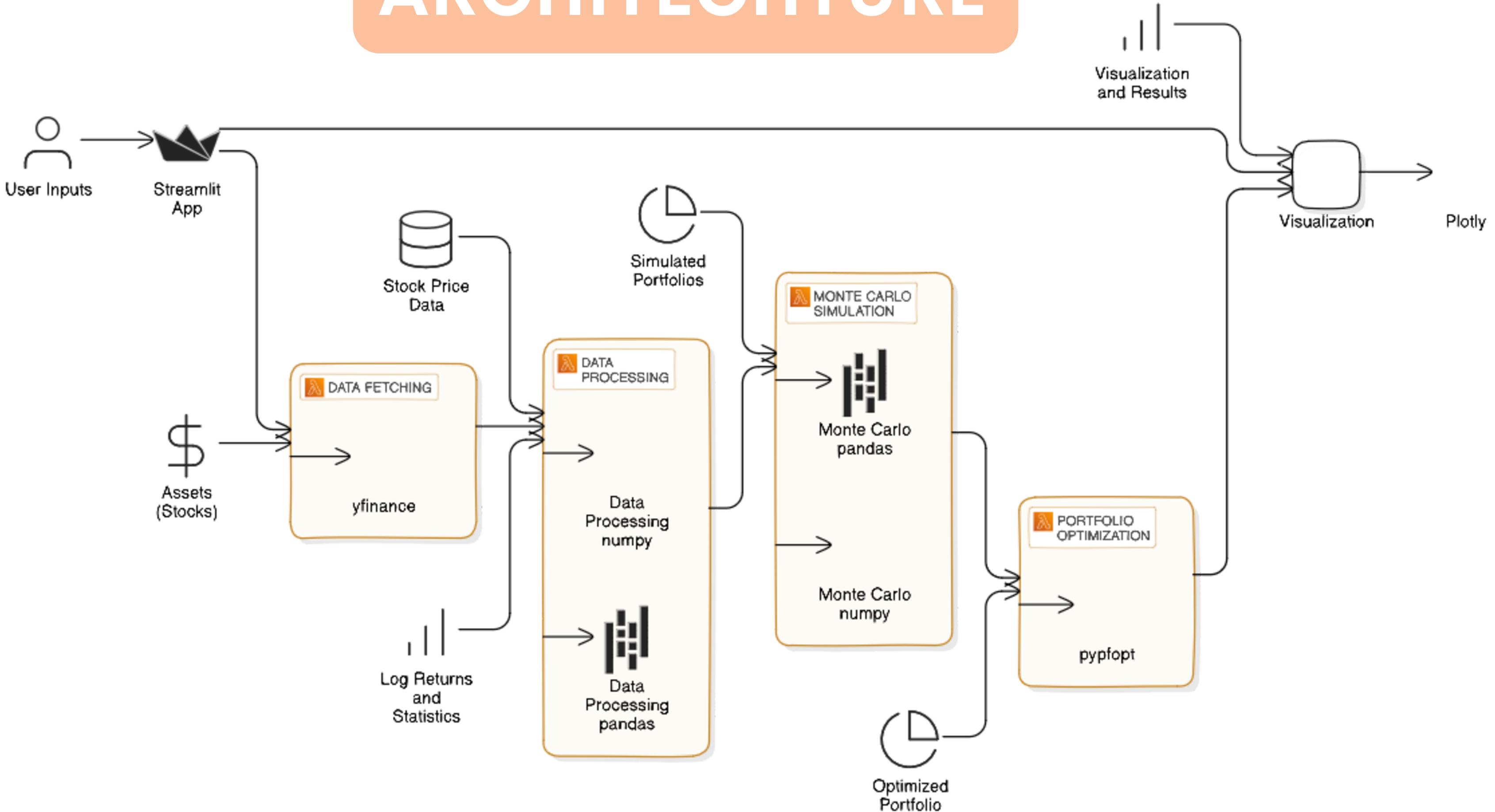


<https://youtu.be/7ESK5SaP-bc?t=65>

FLOWCHART



ARCHITECHTURE



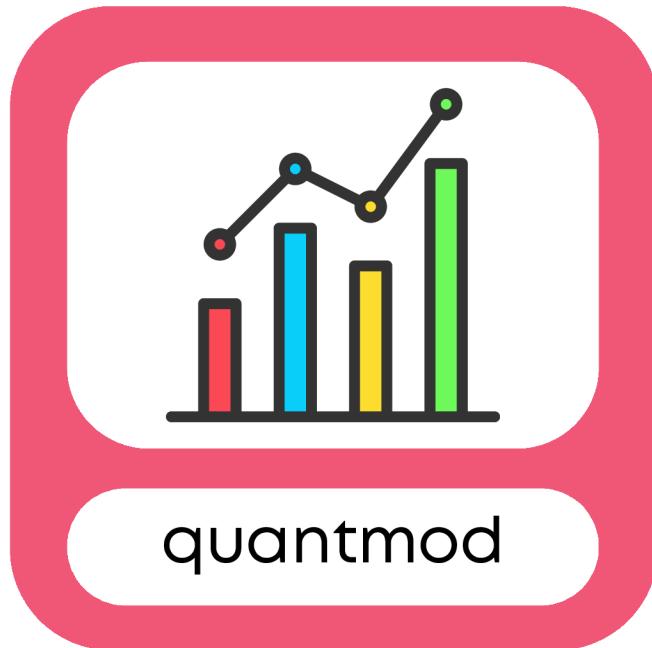
TECH STACK



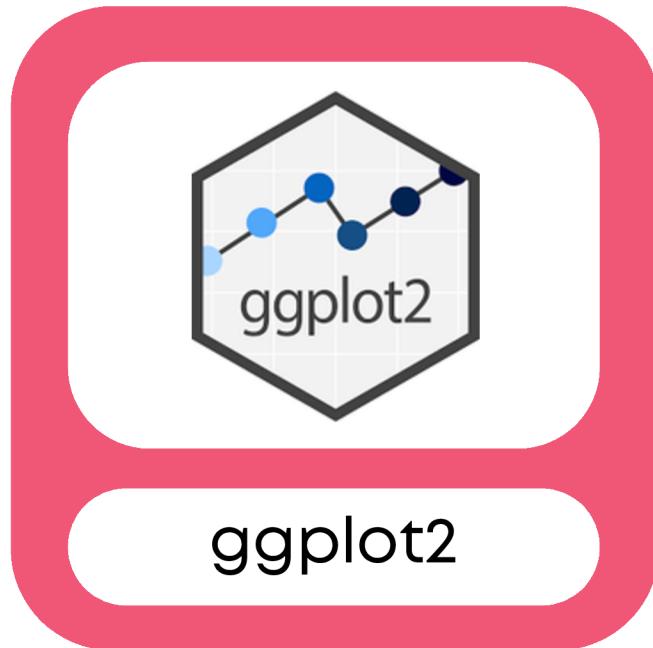
R



Shiny



quantmod



ggplot2



tidyverse



plotly

• • •
• • •
• • •
• • •
• • •

• • •
• • •
• • •
• • •
• • •

TECH STACK



REFERENCES

- Kim, D., & Lin, H. (2023). "Integrating Apache Kafka and Druid for High-Throughput Financial Data Analytics." *International Journal of Big Data and Finance*, 12(3), pp. 188-199.
- Chen, X., Wu, Q., & Zhang, Y. (2023). "A Real-Time Approach to Financial Data Visualization Using Apache Superset and Machine Learning Models." *Journal of Financial Computing and Big Data*, 17(2), pp. 345-360.
- Yao, J., & Li, W. (2022). "Optimizing Retirement Portfolios with Monte Carlo Simulations and Big Data Techniques." *Journal of Financial Engineering and Data Science*, 8(4), pp. 250-268.
- Miller, P., & Sun, T. (2022). "The Use of Distributed Data Storage and Processing for Financial Analytics in High-Frequency Markets." *IEEE Transactions on Financial Computing Systems*, 11(1), pp. 85-97.
- Huang, R., & Singh, A. (2021). "Big Data Architecture for Real-Time Financial Market Analysis and Decision Making." *Journal of Applied Data Science in Finance*, 6(2), pp. 410-426.

CONCLUSION

This project demonstrates the effectiveness of combining real-time stock analysis with retirement portfolio optimization. By utilizing a blend of big data technologies—Apache Kafka, Apache Druid, and Apache Superset—with portfolio analytics in R, the system meets its objectives of real-time data processing, effective data storage, and insightful visualization. With further developments, such as adding more data sources and enhanced visualization options, this system could serve as a comprehensive tool for financial data analysis and retirement planning. The successful integration of real-time data processing and portfolio management methodologies showcases a practical approach to applying big data analytics in the financial domain.

THANK YOU!