

Sector-Sensitivity Monitor for Index Funds

Department of Applied Data Science, San Jose State University

Disha Rawat, Naga Sai Haritha Gottumukkala, Rahul Mohan Devi, Ruchika Chaurasia

Professor: Dr. Keeyong Han

Department of Applied Data Science, San José State University

DATA 226 – Data Warehouse and Pipeline (Section 23), Fall 2025

Professor: Dr. Keeyong Han

Proposal

Title: Building a Sector-Sensitivity Monitor for Index Funds: Quantifying How Each Sector Drives Market Volatility

Abstract

This project will build a data pipeline to measure how each sector actually influences the movement and volatility of the S&P 500 ETF (SPY). While the sector weights of SPY are publicly available, they only show what the fund owns, not how each sector impacts its day-to-day behavior. Using Airflow, Snowflake, and dbt, we will automatically collect daily data for SPY and the eleven sector ETFs from Yahoo Finance/ finnhub, store it in Snowflake, and calculate rolling sector betas that show which sectors contribute most to overall volatility and market movement. Visualization will be created using Tableau, Microsoft Power BI, or Preset to display changing sector sensitivities and diversification trends. Once the core system is complete, we may extend it with forecasting or cross-fund comparisons to demonstrate how sector influence evolves over time.

Dataset

- Historical Source: [Yahoo Finance/Finnhub](#) — daily adjusted close prices for SPY and the eleven GICS sector ETFs (XLK, XLF, XLY, XLI, XLE, XLB, XLV, XLU, XLC, XLP, XLRE), converted into daily returns for modeling.
 - Real-Time Source: [Yahoo Finance/Finnhub API](#) — used to refresh SPY and sector data automatically through Airflow scheduling.
 - Why Important: Although SPY's sector composition is known, this dataset enables us to analyze each sector's real behavioral contribution to volatility, revealing how market concentration and diversification change dynamically over time.
-

Reference Links

- [Yahoo Finance API/ FinnHub API](#)