Full Report of progress in Finance data science

Author: Adrián Portomarin Castaño

Date: September 2025

1. Introduction

This document summarizes the progress of one intensive week of work exploring financial data science and quantitative finance. The focus was on learning through practice, using Python and Jupyter Notebooks, while combining financial intuition with mathematical and statistical reasoning. Each section corresponds to a key concept or method applied during the week, accompanied by practical interpretations and critical reflections. The style of this report is intentionally interpretative, reflecting the process of learning and analysis rather than purely technical description.

2. Sharpe Ratio (static vs rolling)

The Sharpe Ratio was the first performance measure explored. It evaluates the relationship between risk and returns by dividing the mean excess return of an asset by its volatility. The daily Sharpe ratio for the S&P500 over a 60-day window was approximately 0.27, meaning that for every unit of risk taken, the investor was compensated with 0.27 units of excess return. When annualized, the Sharpe ratio reached 4.3, which is unrealistically high given the small dataset. This highlighted the limitation of annualizing short-term samples, as volatility does not scale linearly with time.

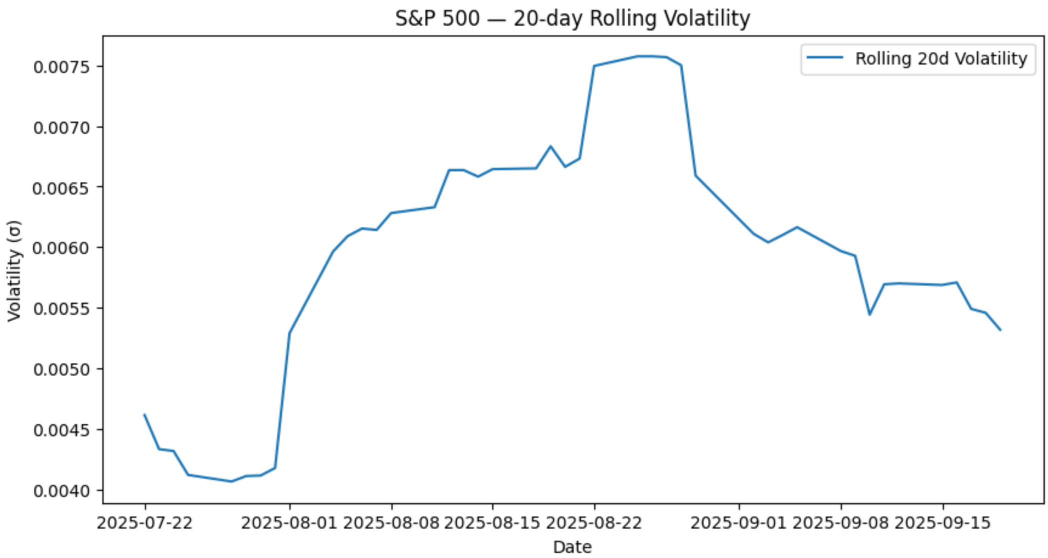
The introduction of the rolling Sharpe ratio provided a more realistic perspective. Unlike the static measure, the rolling Sharpe recognizes that both risk and return vary through time. The rolling version displayed peaks and troughs, even becoming negative around specific dates. This dynamic approach captures market regimes more effectively than a single static number, making it more useful in practice.

3. Rolling Volatility

Rolling volatility was calculated using a 10-day window. It illustrated how the “noise” or standard deviation of returns is not constant but fluctuates over time. The curve resembled a sharp alpine peak: starting low, climbing rapidly in mid-August, and then easing again. This pattern reflected market turbulence during the analyzed period.

Mathematically, rolling volatility is just the moving standard deviation of daily returns over a fixed horizon. Economically, it represents the instability of prices and can be linked to investor fear and uncertainty. For a risk manager, spikes in rolling volatility are warning signs, as they often precede or accompany stress events. The graph 1 shows how SP500 volatility fluctuated in the 20 days before the analysis through the Rolling Volatility measure. There bigger uncertainty moment was almost at the half of the period, showing than the prices were being able to change drastically over the average daily returns already calculated.

Graph 1



Source: Made by the author in Jupyter Lab with data from yfinance

4. Drawdowns

The drawdown analysis measured the distance between the current cumulative return and the historical peak. Half of the period showed drawdowns equal to zero, meaning that new highs were reached. The other half displayed various depths of decline, with a maximum drawdown of -2.37%. This coincided with the earlier spike in volatility, showing the consistency of different risk measures.

Interpretation was straightforward: a drawdown of -2.37% indicated that the investor, at the worst point during this horizon, was down 2.37% compared to their previous wealth maximum. Graphically, every time the drawdown touched zero, it meant that the market had reached a new historical high at that date.

5. Stress Testing (Shock scenarios)

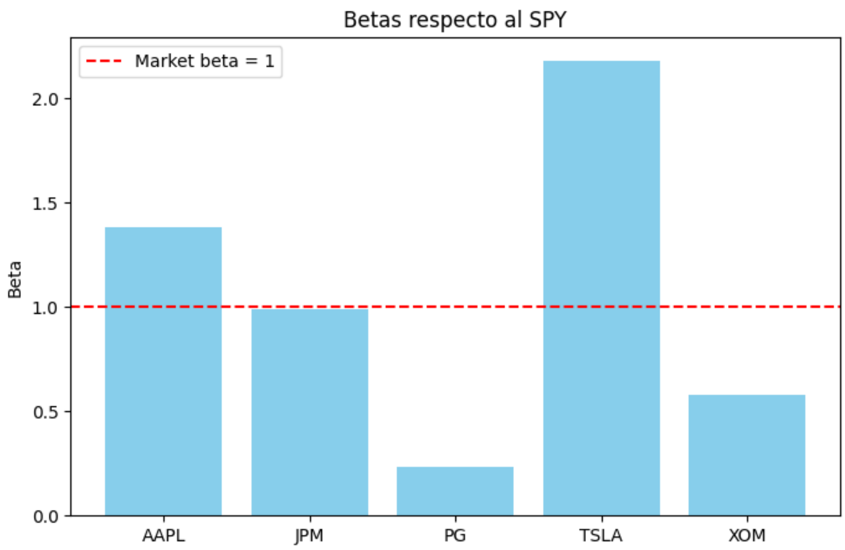
Stress testing simulated how portfolios react to sudden shocks. In the first scenario, a 10% shock was applied mid-period (day 30), leading to a maximum drawdown of -10.68% compared to -2.38% in the normal scenario. Recovery was slow, as the stressed portfolio consistently underperformed the normal one until it could eventually climb back to zero drawdown. This exercise illustrated the fragility of financial systems: a single unexpected event can widen losses dramatically and prolong recovery. The stress test also connected directly to risk management practice, where institutions regularly simulate shocks (political, macroeconomic, or credit-related) to evaluate resilience.

6. Correlation and Beta

Correlation matrices and beta coefficients were analyzed across a set of assets (AAPL, JPM, PG, XOM, TSLA, and SPY). Results showed high correlation of AAPL and JPM with the market, Tesla displaying amplified sensitivity, and PG showing very low correlation and beta.

In the Graph 2 we can see how each asset is intensively related with the SP500 through the Betas providing an intuitive reading: JPM moved almost one-to-one with SPY ($\beta \approx 1$), Tesla amplified market movements by more than double ($\beta \approx 2.18$), while PG barely moved in sync ($\beta \approx 0.24$). This reflected the economic nature of the firms: technology and financials tend to move with market cycles, while consumer staples like PG are defensive and more resilient. A network-style diagram of correlations illustrated these relationships visually, making interconnections between assets easier to grasp.

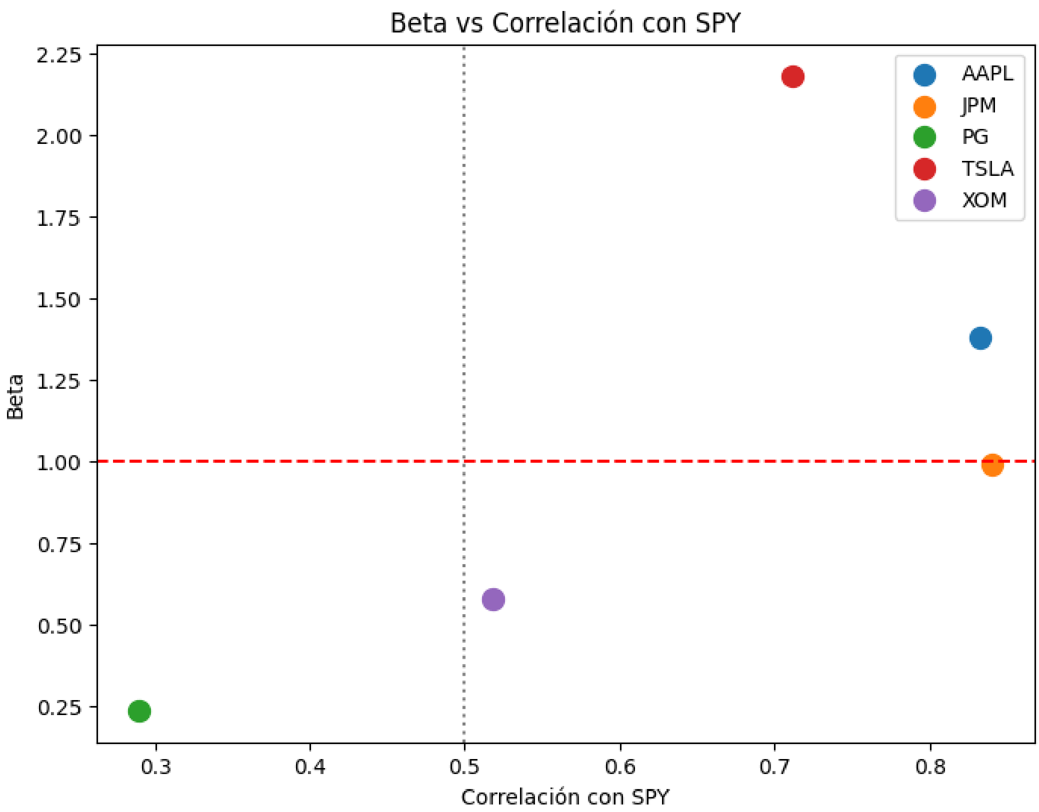
Graph 2



Source: Made by the author in Jupyter Lab with data from yfinance

While Graph 3 puts the relationship between beta and correlation, letting the assets in differente squares, classifying them if they behave the same direction as the market (SP500) and how strong they react in comparison to it. Like this we can define assets for different aims at the time of building a portfolio and adapt to each profile.

Graph 3



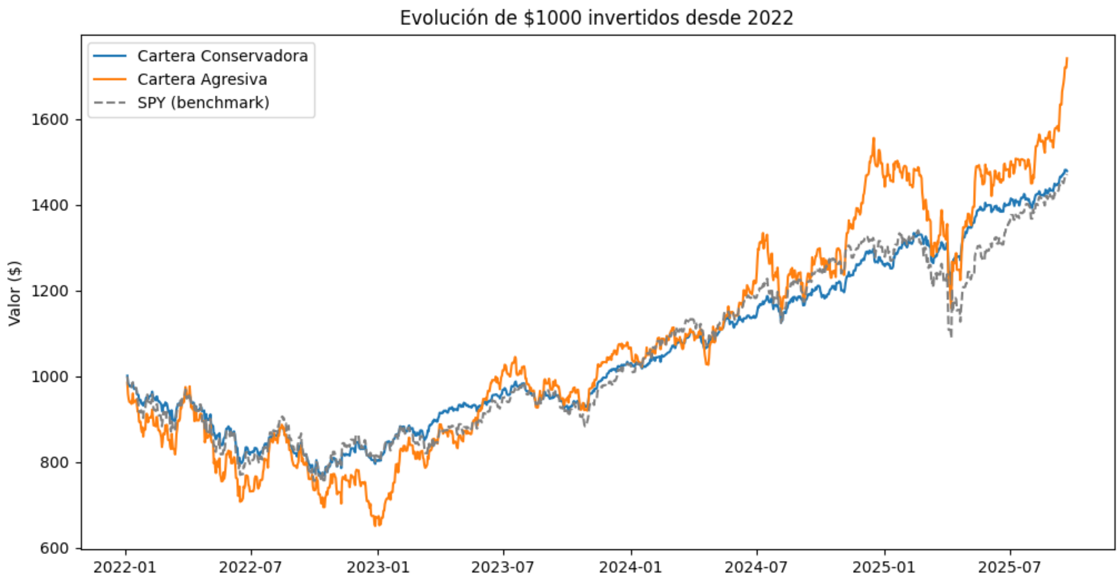
Source: Made by the author in Jupyter Lab with data from yfinance

7. Portfolio Construction (Conservative vs Aggressive)

Two multi-asset portfolios were constructed: one conservative and one aggressive. The conservative portfolio included a higher weight in bonds and defensive assets, while the aggressive one leaned toward equities, technology, and Bitcoin.

Performance analysis showed the conservative portfolio producing returns close to the market but with lower volatility and beta ≈ 0.53, consistent with its defensive nature. The aggressive portfolio achieved a 74.2% return from 2022 to 2025, with a beta of 1.13, meaning it moved slightly more than the market. Despite higher volatility, the aggressive portfolio successfully beat the market, showing the reward of taking extra risk. Graph 4 shows this evolution of both portfolios and the evolution of the market.

Graph 4



Source: Made by the author in Jupyter Lab with data from yfinance

8. Shock Scenarios on Portfolios

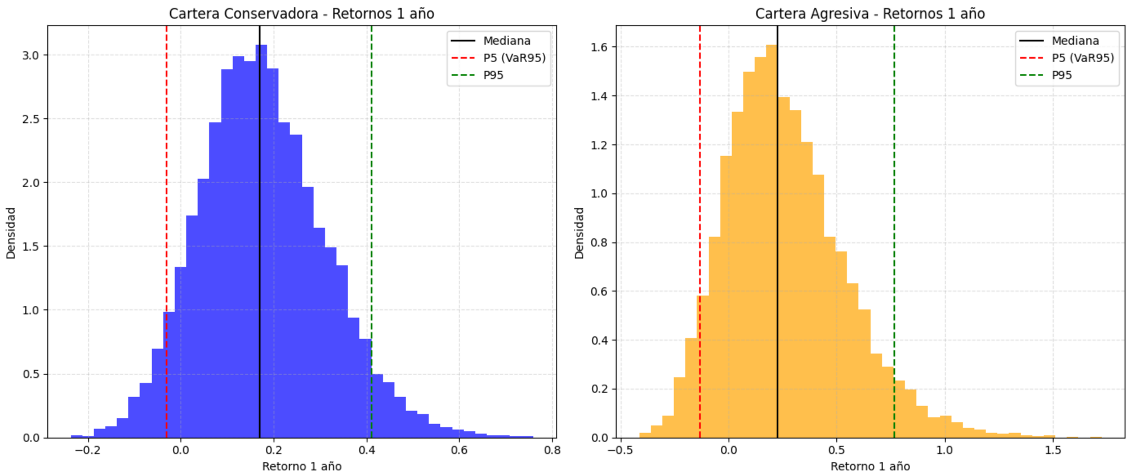
Applying shocks (such as the November 2024 "Trump scenario") allowed comparison between the market and portfolios. The conservative portfolio fell less and recovered more quickly, while the aggressive portfolio suffered larger drawdowns and a slower path to recovery. This difference highlighted the role of diversification and asset selection in cushioning against systemic shocks.

9. Value-at-Risk (VaR) and Conditional VaR (CVaR)

Both historical and parametric VaR were computed at 95%. For the aggressive portfolio, VaR was -2.04%, while CVaR reached -3.05%. This meant that in 95% of days, the portfolio would not lose more than 2.04%, but within the worst 5% of days, the average loss was -3.05%. The Graph 5 is a 2-graph distribution of returns in 1 year, blue for conservative portfolio and orange for aggressive portfolio, red marked with the VaR.

Comparing the historical vs parametric methods revealed that, in calm portfolios, both give similar results. However, for riskier portfolios, historical CVaR showed heavier tails, capturing extreme losses more accurately. This reinforced the intuition that financial losses are not normally distributed, and tail events matter more than the center of the distribution.

Graph 5



Source: Made by the author in Jupyter Lab with data from yfinance

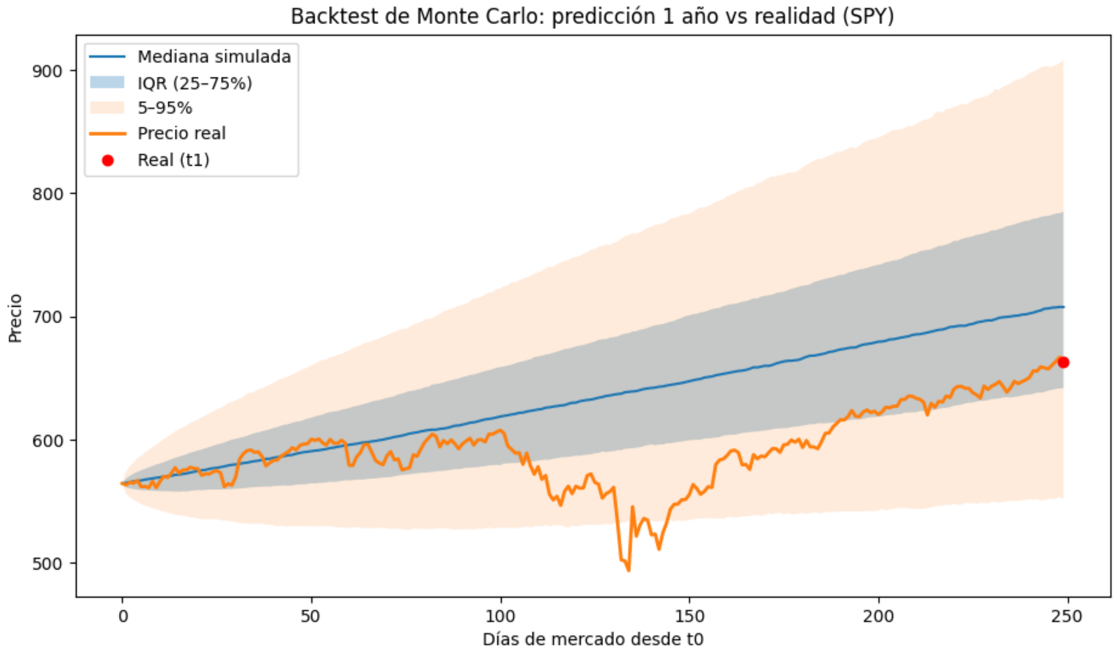
10. Monte Carlo Simulation (SPY and portfolios)

Monte Carlo simulations were applied to forecast future prices. For the SPY, 1,000 simulated paths over a one-year horizon showed a median return of +25%, with only 2% chance of losing more than 10% and 60% chance of gaining more than 20%. The results were optimistic, though subject to model assumptions and ignoring macro shocks.

Applied to portfolios, the conservative portfolio showed narrower bands (P5 = -3.12%, P95 = +41.17%), while the aggressive one displayed a much wider cone (P5 = -13.15%, P95 = +76.66%). This illustrated the core of portfolio design: narrower risk bands provide stability, wider bands allow for higher but riskier potential growth.

Finally, Monte Carlo was validated against reality by comparing forecasts from September 2024 with actual prices in September 2025. The real SPY price fell inside the [P5, P95] band, specifically around P33, and only 6.7% away from the simulated median. This confirmed both the usefulness and the limitations of Monte Carlo: while it cannot predict exact outcomes, it frames a range of plausible futures. This is showed in the Graph 6, in which besides the area of simulations, we can follow the red line to see the real evolution of market.

Graph 6



Source: Made by the author in Jupyter Lab with data from yfinance

11. Conclusions

Throughout the week, the work progressed from single-asset measures (returns, volatility, Sharpe) to advanced portfolio risk methods (VaR, CVaR, Monte Carlo, stress testing). The process combined economic interpretation with mathematical/statistical grounding and coding in Python. Key insights include:

- Risk and return are inseparable; higher returns come with wider tails and worse "storm days."
- Static measures (Sharpe, volatility) are limited; rolling measures and stress tests capture dynamics better.
- Portfolios behave differently under shocks; defensive ones recover faster, aggressive ones amplify losses.
- Monte Carlo is not a crystal ball but a probabilistic lens, framing what "could" happen.
- Comparing simulations with reality confirms the value of probabilistic modeling, while reminding us of the impossibility of exact predictions.

This project provided both the technical exposure (Python, Jupyter, yfinance, pandas, matplotlib) and the conceptual grounding (risk measures, portfolio construction, probabilistic modeling) that constitute the essence of financial data science