DS Portfolio Project Proposal:
Pricing Financial Options Using the Binomial Tree and Black–Scholes Models

Daron Baltazar

February 6, 2026

1 Project Overview

Financial derivatives, particularly stock options, play a central role in modern financial markets. An option's value depends on the uncertain future movement of the underlying asset, making option pricing a natural application of probabilistic modeling, numerical methods, and data-driven computation.

The goal of this project is to build a complete computational workflow for pricing European call options using the Cox–Ross–Rubinstein (CRR) binomial tree model, validate the numerical results against the Black–Scholes closed-form solution, and analyze the convergence and accuracy of the discrete-time approximation. This project frames option pricing as a data science workflow by combining real-world data acquisition, statistical parameter estimation, numerical simulation, and model validation.

This project blends mathematical modeling, financial theory, and computational implementation in Python, creating a reproducible data science pipeline that connects real market data to theoretical pricing models.

2 Motivation and Research Questions

Option pricing provides an ideal case study for numerical modeling because:

- A closed-form theoretical solution exists (Black–Scholes).

- A discrete numerical approximation exists (binomial tree).

- The relationship between the two can be tested empirically.

Research Questions:

1. How accurately does the binomial tree approximate the Black–Scholes price?

2. How does the number of binomial steps affect pricing error?

3. How sensitive is option pricing to volatility estimation?

4. What tradeoff exists between computational cost and accuracy?

3 Data Sources

3.1 Market Data

- Stock price data retrieved using the `yfinance` Python API

- Historical returns used to compute realized volatility

- Risk-free interest rates obtained from the FRED database

All data used in this project are publicly available, ensuring the study can be fully reproduced by other researchers.

3.2 Volatility Estimation

- Historical volatility from log returns

- Implied volatility from backwards engineering Black-Scholes

4 Methodology

This project focuses on European call options, which can only be exercised at maturity and therefore provide a clean comparison between the binomial and Black–Scholes frameworks.

4.1 Binomial Stock Price Tree

Time is divided into steps:
$$dt = \frac{T}{N}$$

Up/down factors:
$$u = e^{\sigma\sqrt{dt}}, \quad d = \frac{1}{u}$$

Stock lattice:
$$S_{j,i} = S_0 u^j d^{i-j}$$

4.2 Backward Induction

Payoff at maturity:
$$V_T = \max(S_T - K, 0)$$

Risk-neutral pricing:
$$V = e^{-rdt}[pV_{up} + (1-p)V_{down}]$$

4.3 Black–Scholes Benchmark

$$C = S_0 N(d_1) - Ke^{-rT} N(d_2)$$

5 Preliminary Implementation Results

A prototype pricing pipeline has already been implemented in Google Colab. The notebook constructs a Cox–Ross–Rubinstein binomial tree, computes the option price via backward induction, and compares the result to the Black–Scholes benchmark.

5.1 Example Pricing Output

Using market data for Starbucks (SBUX) and representative parameters, the notebook produced:

- Binomial price: 6.2614

- Black–Scholes price: 6.3382

- Absolute error: 0.0769

- Relative error: 1.21%

This confirms that even with a modest number of steps ($N = 12$), the binomial model already produces a close approximation to the continuous-time solution.

5.2 Convergence Behavior

The notebook also evaluates binomial prices across increasing step sizes $N = \{5, 10, 25, 50, 100, 200, 500\}$. The convergence plot shows the binomial price approaching the Black–Scholes value as N increases, confirming the theoretical expectation that the CRR model converges to the Black–Scholes model as the time step approaches zero.
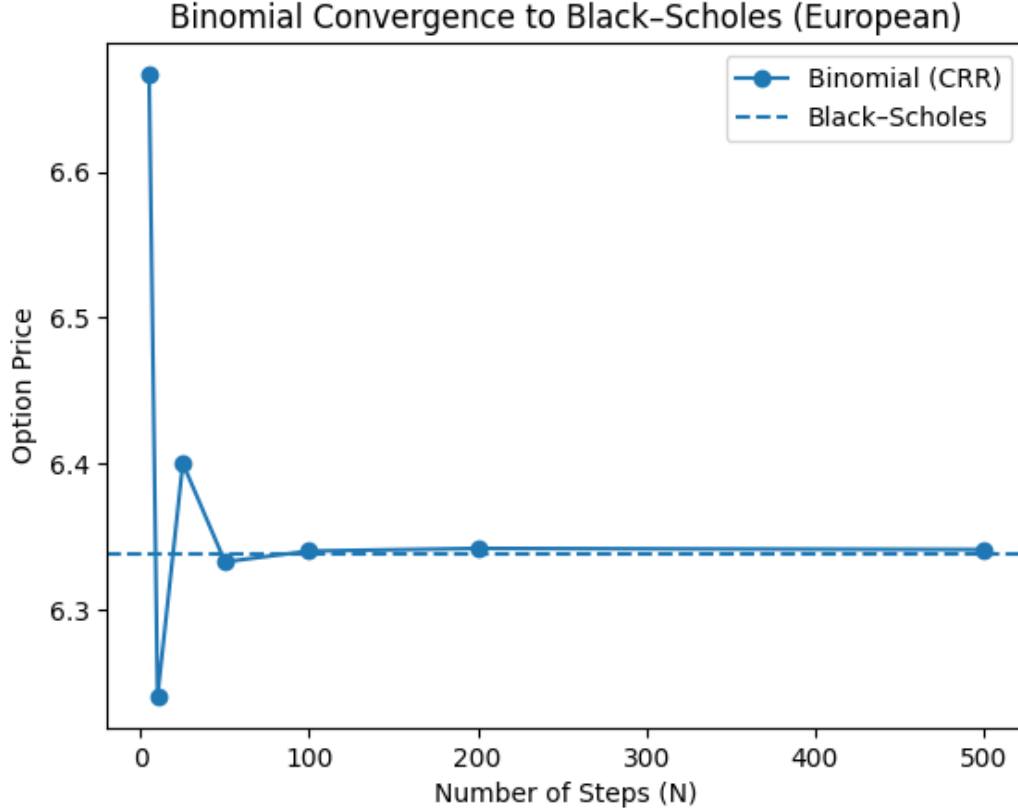
Figure 1: Observed convergence of binomial prices to the Black–Scholes benchmark.

This convergence property is central to validating the binomial model as a numerical approximation of the continuous-time diffusion process.

6 Literature Review

The theoretical foundation for this project draws from three key works in financial mathematics and computational finance.

6.1 Black–Scholes (1973)

Black and Scholes (1973) introduced the first closed-form solution for European option pricing under the assumption that stock prices follow a geometric Brownian motion. Their model assumes:

- Continuous trading

- No arbitrage opportunities

- Constant volatility and interest rates

The Black–Scholes equation provides a continuous-time benchmark that serves as the gold standard for validating numerical pricing methods.

6.2   Cox–Ross–Rubinstein (1979)

Cox, Ross, and Rubinstein developed the binomial tree model as a discrete-time alternative to Black–Scholes. Their work demonstrated that:

- A simple up/down lattice can replicate option payoffs

- Risk-neutral valuation eliminates the need for expected returns

- The binomial model converges to Black–Scholes as the number of time steps increases

This convergence result provides the theoretical justification for using binomial trees as a numerical pricing method.

6.3   Modern Computational Extensions

Recent work in computational finance emphasizes the importance of numerical methods, reproducibility, and data-driven calibration. Master's thesis work by Vidic highlights how binomial and Black–Scholes models remain essential tools for education, prototyping, and financial engineering workflows. Together, these works establish the theoretical and computational foundation for this project.

7   Planned Experimental Analysis

The final project will extend the prototype notebook into a structured experimental study.

7.1   Experiment 1: Convergence Analysis

Measure pricing error as a function of binomial step size:

$$\text{Error}(N) = |C_{binomial}(N) - C_{BS}|.$$

This experiment will quantify the tradeoff between computational cost and accuracy.

7.2   Experiment 2: Volatility Sensitivity

Compare option prices under:

- Historical volatility

- Implied volatility (when available)

This experiment will illustrate how volatility estimation influences pricing outcomes.

7.3   Experiment 3: Computational Efficiency

Measure runtime as a function of tree depth N to illustrate the computational scaling of the binomial method.

8 Reproducibility and Tools

All experiments will be implemented in Python using a reproducible Google Colab notebook. The project will rely on:

- `numpy` and `pandas` for numerical computing

- `matplotlib` for visualization

- `scipy` for statistical functions

- `yfinance` for market data acquisition

The final repository will include:

- A fully executable notebook

- Clear documentation and usage instructions

This ensures that all results can be replicated by other students.

9 Project Timeline

Window	Milestone / Work Product	Output
Late Jan – Early Feb	Finalize proposal and baseline parameters (S, K, r, T, σ) and clean notebook structure	Proposal PDF + cleaned Colab notebook
Feb 10 – Feb 23	Implement full CRR workflow (tree construction, payoff computation, backward induction), add visualization tools	Stock tree + option tree visuals
Feb 24	Mini Lecture (Practical): Binomial Tree (CRR)	1-page cheat sheet + demo notebook
Feb 25 – Mar 16	Run experiments: convergence analysis (N sweep), runtime scaling, volatility sensitivity	Figures + results tables
Mar 17	Mini Lecture (Theory): No-arbitrage and risk-neutral valuation	Resource page + activity prompts
Mar 18 – Early Apr	Write paper draft (methods, results, discussion) and refine visualizations	Draft academic report
Early – Mid Apr	Final revisions, literature review polishing, GitHub packaging and reproducibility checklist	Final paper + GitHub repository
Apr 22	Final presentation (SPARK Symposium)	Slides/poster + presentation

Table 1: Planned project timeline and deliverables.

10 Final Expected Contributions

This project will demonstrate how:

- A theoretical finance model becomes a data science workflow

- Numerical methods approximate continuous-time models

- Real market data can be integrated into academic modeling

The result will be a reproducible computational study linking theory, data, and implementation.

11   Significance and Impact

Option pricing sits at the intersection of finance, mathematics, and data science. By developing a fully reproducible workflow that connects real market data to foundational financial theory, this project demonstrates how theoretical models translate into practical computational tools.

The project is particularly valuable for:

- Students learning quantitative finance and numerical methods

- Data scientists interested in financial modeling workflows

- Demonstrating reproducibility and model validation practices

More broadly, this work illustrates how discrete numerical methods converge to continuous-time stochastic models, providing a concrete example of how mathematics, programming, and real-world data interact in modern data science.

12   Conclusion

This proposal outlines a data-driven study of option pricing using both discrete and continuous-time financial models. By combining real market data, numerical experimentation, and theoretical validation, the project demonstrates how financial mathematics can be implemented as a reproducible data science workflow. The completed study will provide empirical evidence of binomial model convergence, illustrate the role of volatility estimation, and highlight the tradeoff between computational cost and pricing accuracy.

13   Reproducibility Plan

All code, data acquisition steps, and experiments will be fully reproducible. The project will provide:

- A public GitHub repository containing all code

- A Google Colab notebook with executable cells and documentation

- Clear instructions for reproducing figures and experiments

This ensures the results can be verified and extended by other students.

References

1. Black, F., & Scholes, M. (1973). *The Pricing of Options and Corporate Liabilities.* Journal of Political Economy, 81(3), 637–654.

2. Cox, J., Ross, S., & Rubinstein, M. (1979). *Option Pricing: A Simplified Approach.* Journal of Financial Economics, 7(3), 229–263.

3. Vidic, I. (Year). *Binomial and Black–Scholes Option Pricing Models.* Master's Thesis.