

Canadian Banks Stock Analysis Report

Authors: Mariia-Olena Zhupnyk & Mariia Shekhovtsova

Date: 17 Feb 2025

Canadian Banks Stock Data Analysis & Prediction Report

1. Introduction

1.1 Project Overview

The financial sector plays a crucial role in the Canadian economy, and the performance of major banks significantly impacts investors' decisions. This project aims to analyze and predict stock prices for five major Canadian banks: RBC, TD, CIBC, BMO, and Scotiabank. The analysis involves web scraping, data storage, exploratory data analysis (EDA), and predictive modeling using time series forecasting techniques.

1.2 Goals and Purpose

Before diving into the data, we defined the key objectives of our analysis by asking critical questions:

- **Trends & Patterns:** How have stock prices evolved over time? Are there noticeable seasonal or cyclical trends?
- **Volatility:** Which bank's stock is the most volatile, and which is the most stable?
- **Correlations:** Do the stock prices of these banks move together? How strong is their correlation?
- **Prediction:** Can we forecast future stock prices based on historical data?

1.3 Objectives

- Collect, clean, and store historical stock data remotely.
- Conduct Exploratory Data Analysis (EDA) to identify trends and patterns in stock movements.
- Implement time series forecasting techniques to predict future stock prices accurately.
- Compare different banks' performances to evaluate investment opportunities.

2. Environment Setup

After defining the project objectives, the first step in implementation was to establish a data storage system and ensure seamless access for both participants. The data is stored in a PostgreSQL database hosted on a **local Raspberry Pi server**. The Raspberry Pi was chosen for its availability and as an opportunity to practice server management and database administration skills.

Since the server operates with a dynamic IP address, **ZeroTier** was implemented as a network hub to enable secure remote connections. This setup allows both participants to access the database reliably from external locations.

To ensure data security and integrity, a role-based access control (RBAC) model was implemented. Two database user roles were created, granting only the necessary permissions to minimize the risk of unauthorized modifications.

The stock price data was collected through web scraping using Python and subsequently loaded into the **PostgreSQL database**.

For the initial stage, the database was designed as a transactional system (OLTP), primarily storing raw stock price data in a simple structure. At this stage, only a few tables were created to store essential information such as:

stock_prices: Stores ticker symbol, date, open price, close price, high, low, and trading volume.

The screenshot shows the pgAdmin 4 interface. On the left, the Object Explorer displays the database structure: Servers (3) > PostgreSQL 15 > stock > Databases (2) > stock > Tables (7). The main pane shows the 'stock_prices' table with the following columns: st_pr_id (integer), stock_ticker (character varying (10)), date (date), close_price (double precision), high_price (double precision), low_price (double precision), open_price (double precision), and volume (integer). The table contains 20 rows of data for BMO.TO from 2020-01-03 to 2020-01-30.

st_pr_id	stock_ticker	date	close_price	high_price	low_price	open_price	volume
1	BMO.TO	2020-01-03	80.27942657470703	80.33528676800852	79.5931407208292	79.8804234544664	1473000
2	BMO.TO	2020-01-06	80.2155990600586	80.23954436576099	79.85649862249471	79.88841960845903	1147500
3	BMO.TO	2020-01-07	80.07994079589844	80.62258538989624	80.04801980939202	80.23954572843053	702000
4	BMO.TO	2020-01-08	80.6624984741211	81.10139839609874	80.19965324296813	80.2156106941119	1088400
5	BMO.TO	2020-01-09	80.95774841308594	81.21311022250089	80.702386603671	80.99765116918125	1683900
6	BMO.TO	2020-01-10	80.93379974365234	81.05350190965419	80.56671756600375	80.90986052811299	1243400
7	BMO.TO	2020-01-13	81.55623626708984	81.55623626708984	80.70236982362026	80.93379236658177	4100400
8	BMO.TO	2020-01-14	81.96321868896484	82.01907888121661	81.39663499946704	81.5642155762235	1923500
9	BMO.TO	2020-01-15	82.20263671875	82.33031457022996	81.90797217422549	81.91535394253842	1729600
10	BMO.TO	2020-01-16	82.48192596435547	82.58566458440842	82.20262499628637	82.33828459673357	1809800
11	BMO.TO	2020-01-17	82.83306121826172	82.87295788261139	82.51985746193809	82.59365687894268	1545900
12	BMO.TO	2020-01-20	82.59365844726562	83.35176205566233	82.59365844726562	82.67345786578548	1591500
13	BMO.TO	2020-01-21	83.26398468017578	83.34379019065408	82.60164399666773	82.79316382754321	1384400
14	BMO.TO	2020-01-22	83.31185913085938	83.59116011884609	83.04851559005552	83.263974612701	1100600
15	BMO.TO	2020-01-23	83.19216918945312	83.46348844955472	82.76124499255135	83.27196861370031	2397400
16	BMO.TO	2020-01-24	83.00861358642578	83.46347693904056	82.81709377698402	83.26397536178311	1888100
17	BMO.TO	2020-01-27	82.51385498046875	82.70537479318119	82.19465123374465	82.2345398519957	4184800
18	BMO.TO	2020-01-28	82.76921844482422	83.01659846162552	82.34627605933612	82.72134001274789	4582100
19	BMO.TO	2020-01-29	82.58568572998047	82.96074973526522	82.45800786401679	82.73730890359926	3678300
20	BMO.TO	2020-01-30	82.48991394042069	82.5936525682909	82.09888646739638	82.30636981142376	1461200

As the project evolves, transitioning towards an analytical database (OLAP) would improve performance for data analysis and visualization. This future structure would store preprocessed financial indicators, reducing the need for on-the-fly calculations. Additional tables could include:

technical_indicators: Stores precomputed moving averages (SMA, EMA), volatility measures (standard deviation, ATR), and relative strength index (RSI).

historical_analysis: Aggregated data for weekly/monthly price trends, correlations between stocks, and sector performance insights.

This enhanced structure would facilitate dashboarding and reporting, allowing for faster access to trend analysis, risk assessment, and trading signals.

A scheduled backup process was set up to prevent data loss. However, as the backup is also stored on the Raspberry Pi, future improvements could include off-device backups or cloud-based storage for better redundancy.

3. Data Collection & Processing

3.1 Data Source

The dataset was obtained using **web scraping from Yahoo Finance** and includes stock prices from **January 2020 to December 2024**.

Web scraping is an efficient and automated method to gather data from the internet, ensuring we have up-to-date and comprehensive datasets.

3.2 Data Structure

The dataset consists of the following columns:

- **stock_ticker** – Bank stock symbol (e.g., BMO.TO, TD.TO)
- **date** – Trading date
- **close_price**, **high_price**, **low_price**, **open_price** – Daily stock prices
- **volume** – Number of shares traded

These are essential variables for understanding the price dynamics and trading activity.

Feature Engineering

- **Simple Moving Averages (SMA):** SMA_50, SMA_100, SMA_500 – These represent the average closing price over the past 50, 100, and 500 days.
 - **Exponential Moving Averages (EMA):** EMA_50, EMA_100, EMA_500 – These are more responsive to recent price changes.
 - **Daily Return (daily_return):** Measures the percentage change in the closing price compared to the previous day.
- These features help identify trends, smooth out fluctuations, and assess risk, improving the quality of the analysis and forecasting models.

3.3 Data Cleaning

- Handled missing values.
- Converted data types for analysis.
- Calculated new features such as **moving averages** and **daily returns**.

4. Exploratory Data Analysis (EDA)

EDA is a crucial step in the data analysis process, including when dealing with transactional data. EDA helps to gain insights, identify patterns, understand the structure of the data, and uncover potential issues.

EDA ensures that the data is understood fully, cleaned effectively, and prepared for deeper analysis or predictive modeling.

- **Descriptive Statistics & Data Summary:**
Calculated key statistical measures (mean, median, standard deviation) to understand stock price distributions.
This provides basic insights into the data's spread and central tendencies, crucial for detecting anomalies.
- **Trend Analysis & Moving Averages:**
Computed Simple Moving Averages (SMA) and Exponential Moving Averages (EMA) to identify long-term trends.
These techniques help smooth out price fluctuations and identify long-term market direction.
- **Volatility & Risk Analysis:**
Measured daily returns for each stock to assess risk levels.
Volatility is an important metric for understanding the stability of stock prices and the risks involved.
- **Correlation Analysis:**
Calculated correlations between different bank stocks to determine how they move in relation to each other.
Understanding stock correlations helps in assessing portfolio diversification and identifying potential risks and opportunities.
- **Visualization Techniques:**
Used Matplotlib and Seaborn for static plots.

Visualizations make it easier to identify patterns, relationships, and trends, providing clearer insights.

4.1 Trend Analysis & Moving Averages

4.1.1 Trend Analysis

The analysis covers stock price movements over a five-year period, capturing key economic events and market reactions.

Impact of the COVID-19 Pandemic (2020)

The year 2020 was marked by a sharp decline in stock prices for all five banks due to the global economic disruption caused by the COVID-19 pandemic. However, following the first quarter of 2020, stock prices began a steady recovery, continuing their upward trend until early 2022.

Market Adjustments & Inflationary Pressures (2022-2023)

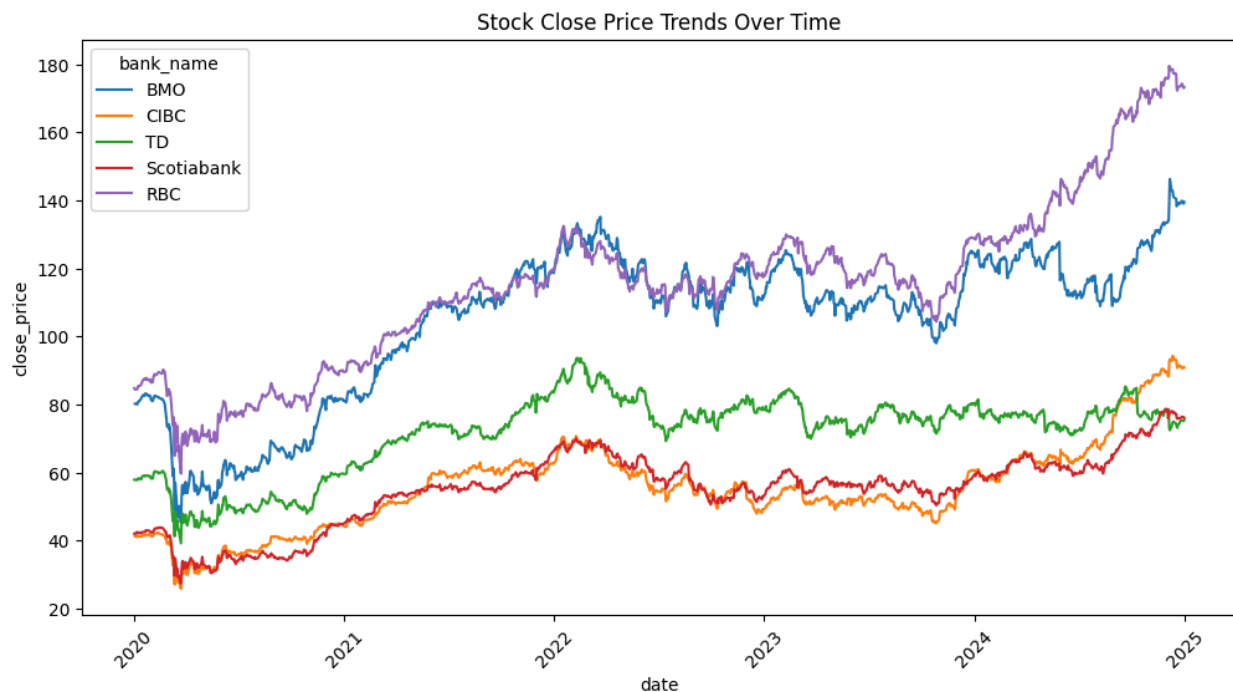
In the first half of 2022, stock prices declined moderately, a trend that persisted through the end of 2023. This decline can be attributed to ongoing economic uncertainty, post-pandemic market adjustments, and inflationary pressures.

Recovery & Bank-Specific Performance (2024-2025)

By 2025, three banks—CIBC, Scotiabank, and TD—had fully recovered, reaching levels comparable to their 2022 peaks. This suggests that these institutions successfully adapted to

changing market conditions and benefited from an improving economy, leading to renewed investor confidence.

- BMO demonstrated stronger-than-expected growth, surpassing its 2022 peak. This could indicate a more aggressive growth strategy or superior operational performance relative to competitors.
- RBS exhibited the most significant long-term growth, consistently outperforming the market over the five-year period. This suggests that RBS effectively leveraged long-term growth opportunities, potentially due to successful strategic initiatives or an expanded market share in key areas.



4.1.2 Moving Averages Analysis

The analysis of moving averages reveals key trends that can guide investment decisions:

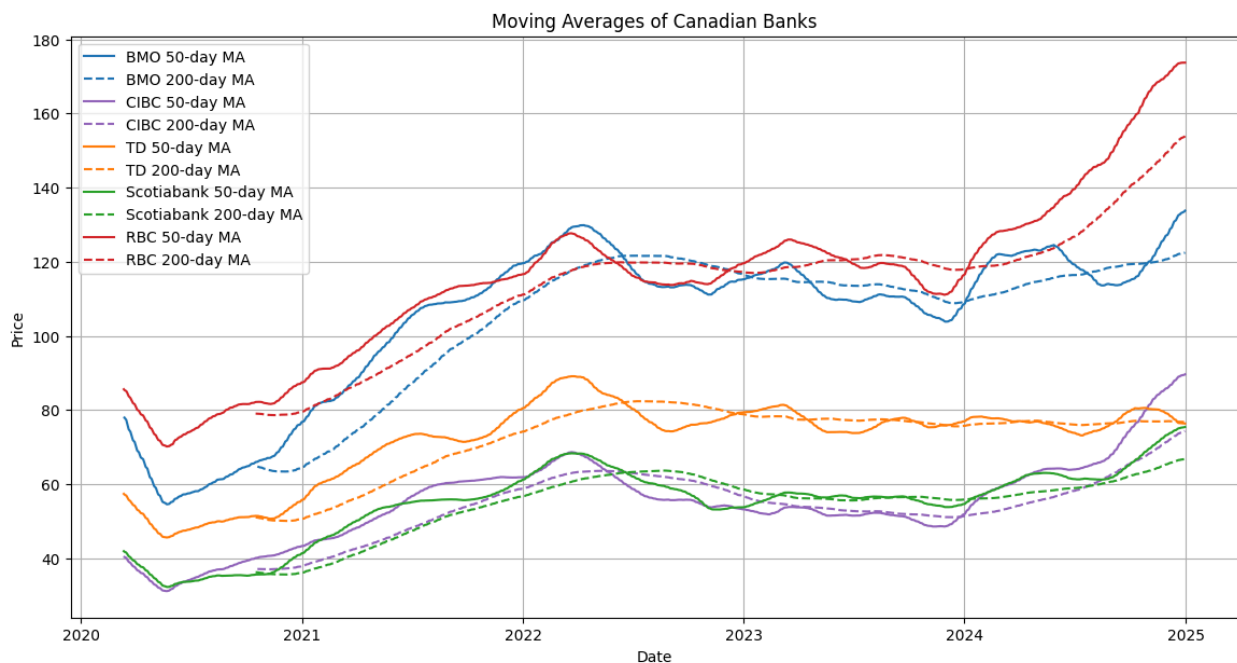
- Bullish Sentiment (End of 2024): Four out of five banks—CIBC, Scotiabank, RBC, and BMO—show a strong upward trend, with their 50-day moving average surpassing the 200-day moving average. This indicates positive momentum and potential growth in the short term.
- TD Bank Exception: TD is the only bank where the short-term moving average remains below the long-term, signaling a cautious or bearish outlook that investors should monitor closely.

Historical Trends & Market Volatility

- 2012 – Q1 2022: A consistent bullish trend was observed across all banks, with stable stock price growth reflecting strong market conditions.
- 2022 – 2024: Increased volatility led to frequent crossovers between short- and long-term moving averages, likely influenced by post-pandemic recovery, inflation, and economic uncertainty. This period presented mixed signals for investors, with risk-averse strategies favoring caution, while opportunistic investors could have leveraged short-term fluctuations.
- Recovery (End of 2023 – End of 2024): Most banks regained positive momentum, except for TD, which lags behind in recovery.

Investment Insights

- CIBC, Scotiabank, and RBC present strong investment opportunities, demonstrating stable growth and bullish trends.
- BMO, despite some fluctuations, ends 2024 on a positive note, making it a potential candidate for investors willing to tolerate short-term volatility.
- TD Bank requires caution, as its underperformance suggests lingering risks or weaker recovery compared to peers.



4.2. Volatility Analysis

Volatility Trends Over Time

The rolling 30-day standard deviation analysis highlights key patterns in stock price fluctuations:

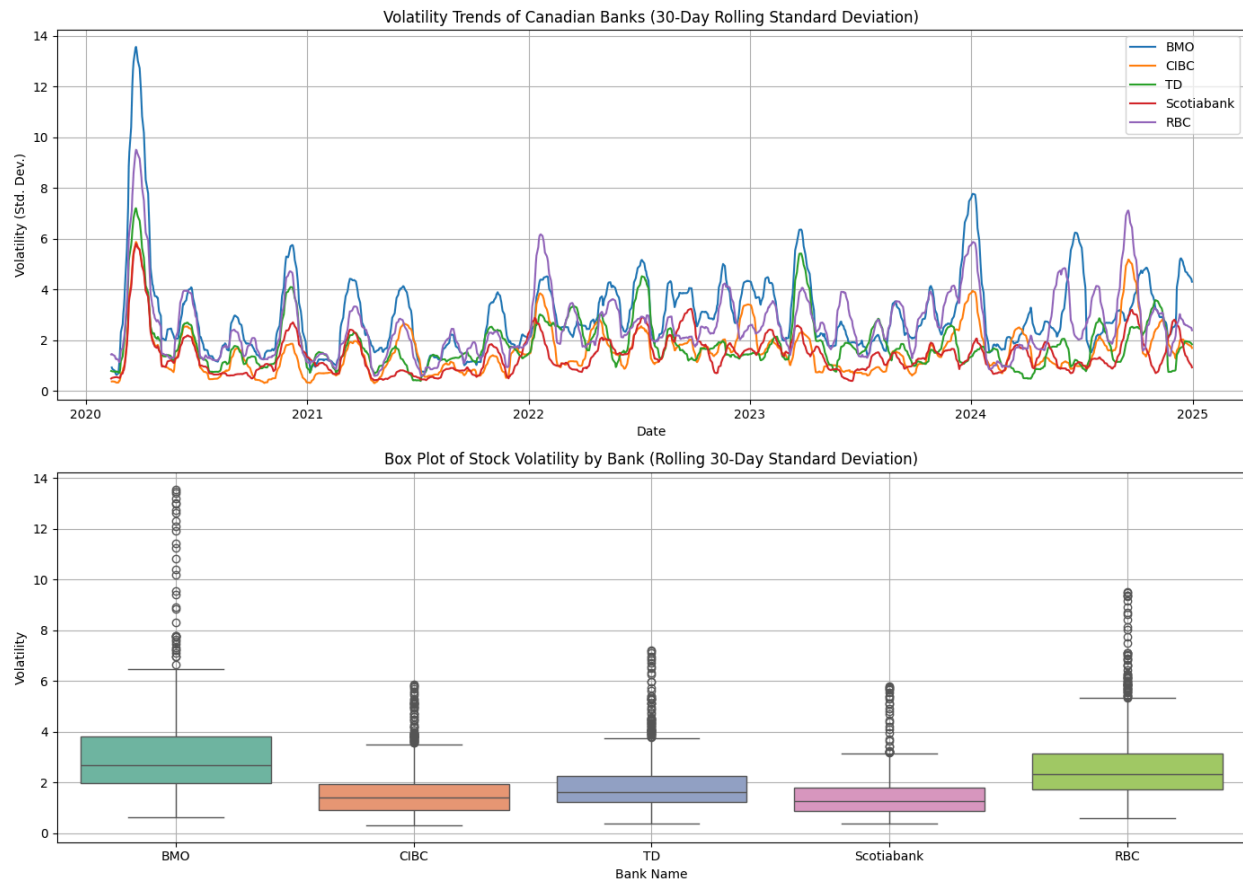
- **Market Shocks and Stability:** Volatility spiked sharply in early 2020, coinciding with the **COVID-19 market crash**. Following this peak, stock price fluctuations gradually stabilized, though periodic cycles of increased volatility persisted, likely driven by **economic events, earnings reports, or interest rate changes**.
- **Bank-Specific Trends:** **BMO and RBC exhibit the highest volatility peaks**, indicating a greater sensitivity to market conditions. In contrast, **CIBC and Scotiabank maintain lower volatility levels**, suggesting more stable price movements over time.

Volatility Distribution by Bank (Box Plot Analysis)

- **BMO and RBC:** These banks have the **highest median volatility** and a **wider interquartile range**, meaning their stock prices **fluctuate more frequently** and experience **more extreme movements** compared to their peers.
- **CIBC and Scotiabank:** The **lowest median volatility** suggests that these banks **tend to experience steadier stock price movements**, potentially making them more attractive for risk-averse investors.
- **TD Bank:** While not as volatile as BMO or RBC, TD's volatility distribution suggests a **moderate level of price fluctuations**.
- **Outliers and Market Events:** The presence of multiple **extreme volatility outliers**, particularly for **BMO and RBC**, indicates that these stocks have been subject to **sporadic but significant market shocks**.

Investment Implications

- **High Volatility (BMO, RBC):** These banks may present **higher risk but also higher return opportunities** for investors comfortable with market fluctuations.
- **Low Volatility (CIBC, Scotiabank, TD):** These stocks are generally **more stable**, making them **suitable for conservative investment strategies**.
- **Historical Market Impact:** The **most extreme volatility was recorded in 2020**, aligning with global economic disruptions, reinforcing the need for **risk management strategies** in periods of market uncertainty.

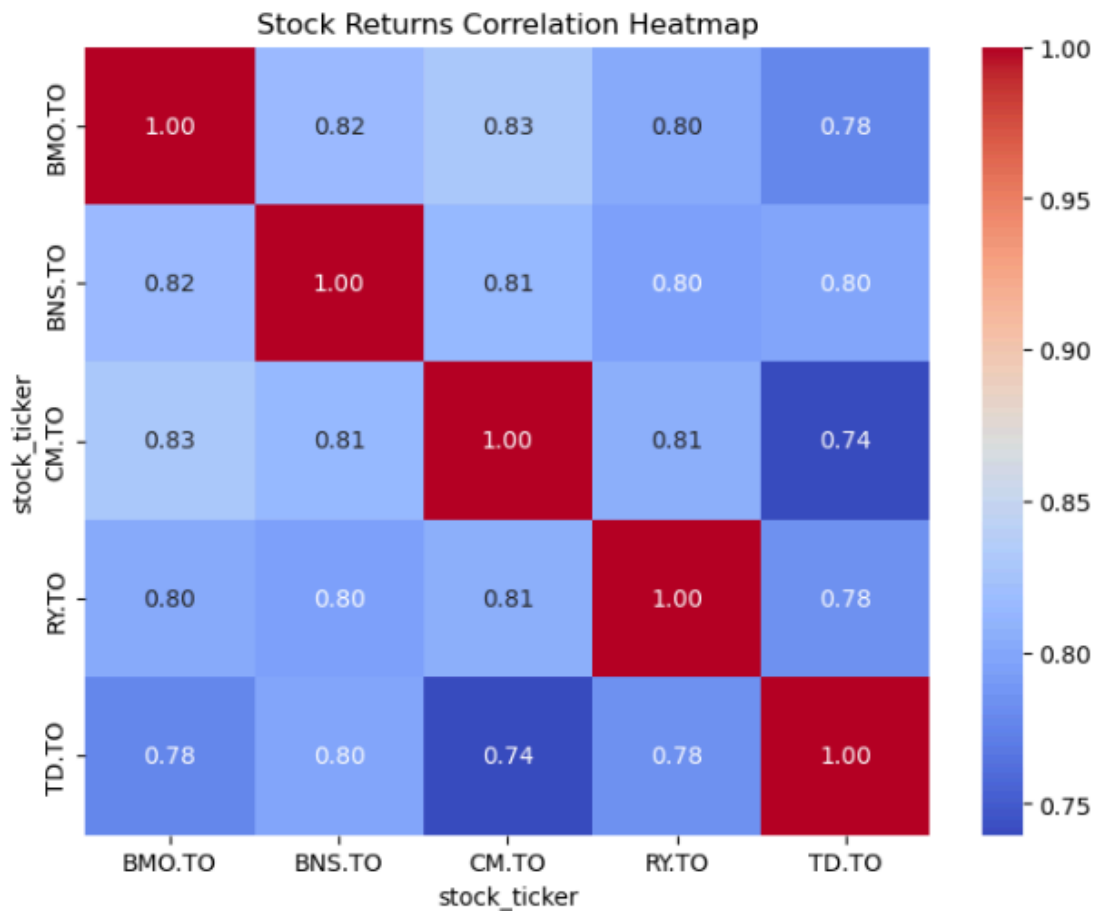


4.3 Correlation Analysis

Correlation analysis helps us understand how stocks move **relative to each other**. It is essential in stock data analysis because:

- **Identifies Relationships** – Reveals whether stocks move together (positive correlation) or in opposite directions (negative correlation).
- **Diversification Strategy** – Helps investors build portfolios by selecting stocks that are less correlated to reduce risk.
- **Market Trends & Sector Analysis** – Shows how stocks within the same industry or market react to economic events.
- **Risk Management** – Helps assess exposure by understanding how stock price movements are interconnected.

By analyzing the **correlation matrix of daily returns**, we can make more **informed investment decisions** and optimize portfolio performance.



Stock Returns Correlation Analysis

The correlation matrix quantifies the relationship between the stock returns of **five major Canadian banks**:

- **BMO.TO** (Bank of Montreal - BMO)
- **BNS.TO** (Bank of Nova Scotia - Scotiabank)
- **CM.TO** (Canadian Imperial Bank of Commerce - CIBC)
- **RY.TO** (Royal Bank of Canada - RBC)
- **TD.TO** (Toronto-Dominion Bank - TD Bank)

The correlation values range from **0.74 to 1.00**, indicating a **strong positive correlation** among all banks, meaning their stock prices tend to move in the same direction.

Key Observations from the Correlation Matrix and Heatmap

1. **High Correlation Across Banks**

- All five banks exhibit **strong positive correlations** (above **0.74**), suggesting that their stock returns tend to move together.
- This is **expected**, as they belong to the **same financial sector**, face similar macroeconomic influences, and respond similarly to market conditions like **interest rate changes, inflation, and banking regulations**.

2. Strongest Correlations

- The highest correlations are observed between:
 - **CIBC (CM.TO) and BMO (BMO.TO) at 0.83**
 - **TD Bank (TD.TO) and CIBC (CM.TO) at 0.83**
- These high correlations indicate that these banks likely have **similar investor sentiment, business exposure, and risk factors**.
- Investors trading these stocks in similar portfolios might contribute to their synchronized movements.

3. Weaker, Yet Still Positive Correlations

- **TD Bank (TD.TO) shows the lowest correlation range (0.74 - 0.78) compared to other banks.**
- While still strongly correlated, TD Bank's stock movements are slightly more independent, which could be due to:
 - **Geographical diversification** – TD has a **strong U.S. presence**, making it more exposed to economic conditions in the U.S. market than its peers.
 - **Business strategy differences** – TD's unique revenue streams and operational focus might influence its stock behavior differently than other Canadian banks.

Why Do These Correlations Exist?

- **Macroeconomic Influences:** Banks respond similarly to **interest rate fluctuations, central bank policies, and liquidity conditions**.
- **Investor Behavior:** Large institutional investors often trade these banking stocks **as a group**, reinforcing their high correlation.
- **Market Risk Exposure:** Since all banks are influenced by **domestic and international financial markets**, economic shifts tend to affect them in tandem.

Investment Implications

- **Diversification Challenges:** Since all banks are highly correlated, investing in multiple Canadian banking stocks may not provide significant diversification benefits.
- **Risk Management:** A market downturn affecting one bank will likely impact the others, requiring **risk mitigation strategies** such as hedging or diversification across sectors.
- **Stock Selection:** Investors seeking **stability** might favor **TD Bank (TD.TO)**, given its slightly lower correlation, while those comfortable with **higher synchronized market movements** may prefer **CIBC (CM.TO) or BMO (BMO.TO)**.

Conclusion

The correlation analysis highlights that Canadian bank stocks move **closely together**, driven by common macroeconomic factors and investor behavior. While **TD Bank (TD.TO)** shows slightly lower correlation, the overall strong positive relationship suggests that banking stocks **do not provide significant diversification within the sector**, and investors should consider external factors when making portfolio decisions.

5. Predictive Modeling: Time Series Forecasting

To gain insights into future stock price movements, two predictive models were employed: **Monte Carlo Simulation** and **ARIMA (AutoRegressive Integrated Moving Average)**. These models serve different purposes and provide complementary perspectives on stock price behavior.

- **Monte Carlo Simulation** is a probabilistic approach that generates multiple possible future price paths based on historical volatility. This method is particularly useful for assessing **uncertainty, risk, and potential price ranges** rather than pinpointing a single outcome.
- **ARIMA** is a statistical model designed for **time series forecasting** by analyzing historical price patterns. It is useful for capturing underlying trends and making short-term predictions based on past behavior.

By combining these models, we aim to **balance probabilistic risk assessment with structured time series forecasting**, offering a comprehensive view of expected stock price dynamics. The following sections detail the rationale, interpretation, and performance metrics for these models

5.1 Price prediction for the 30 days using ARIMA model

The **AutoRegressive Integrated Moving Average (ARIMA)** model was selected due to its effectiveness in capturing time series patterns such as **trends, seasonality, and stationarity adjustments**. ARIMA is widely used for stock price forecasting as it relies on past values and errors to make predictions, making it suitable for financial data where historical trends play a key role.

Flat Forecast: Understanding the Nearly Horizontal Line

The **30-day forecast appears almost flat** for all banks, which can be explained by the following factors:

- **Lack of Strong Trends:** The model primarily relies on past values, and if recent stock price changes are relatively stable, ARIMA projects this stability forward.
- **Short-Term Prediction Limitation:** ARIMA works best with **short-term forecasting**, but when price movement lacks significant patterns, it defaults to a near-constant prediction.
- **Stationarity in Data:** If the differenced series is mostly stationary, ARIMA will predict future values to be close to the mean of past observations.

Model Performance: Key Metrics

To assess prediction accuracy, two error metrics were used:

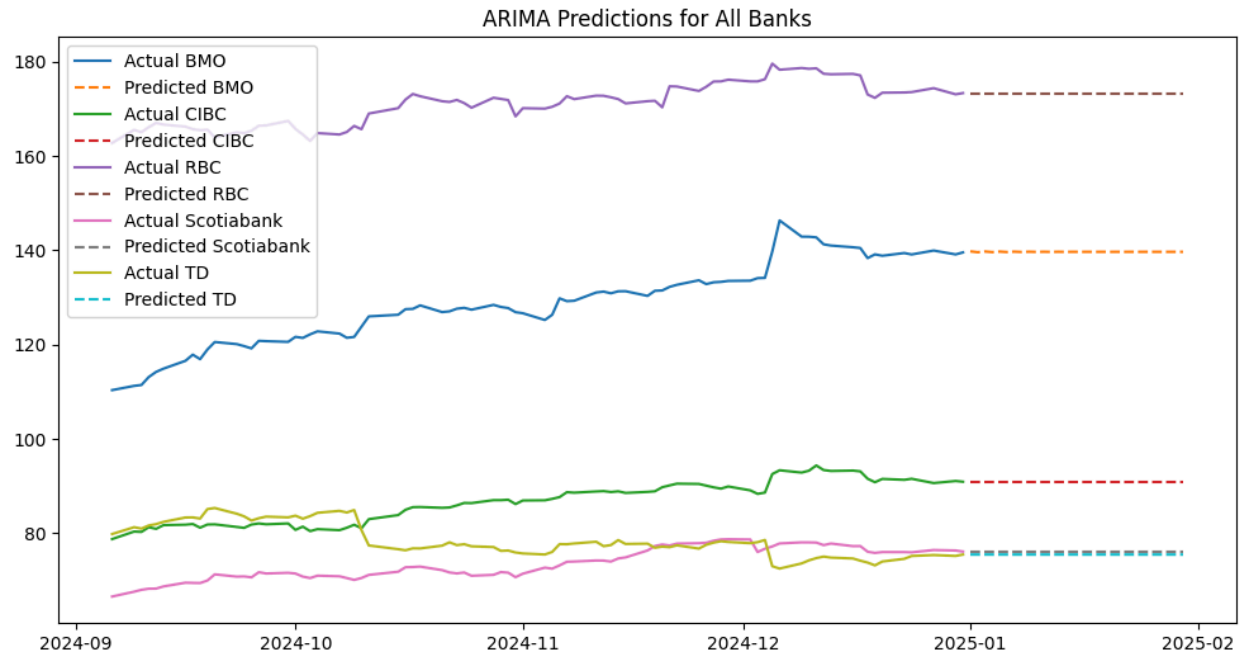
- **Mean Absolute Error (MAE):** Measures the average absolute difference between actual and predicted prices.
- **Root Mean Squared Error (RMSE):** Measures the average squared difference, giving more weight to larger errors.

Performance across banks:

- **BMO:** MAE = 3.82, RMSE = 4.88 (Higher error, indicating more fluctuations in price)
- **CIBC:** MAE = 1.40, RMSE = 1.68 (More stable predictions with lower error)
- **RBC:** MAE = 2.42, RMSE = 3.02
- **Scotiabank:** MAE = 1.18, RMSE = 1.46 (Lowest error, suggesting ARIMA performs well for this bank)
- **TD:** MAE = 1.60, RMSE = 1.84

Conclusion

While ARIMA is effective for short-term forecasting, the **flat predictions indicate that stock price fluctuations are minimal in the short run**. The model captures the underlying patterns but struggles to predict sharp movements, making it **useful for stable trends but less effective during high volatility periods**.

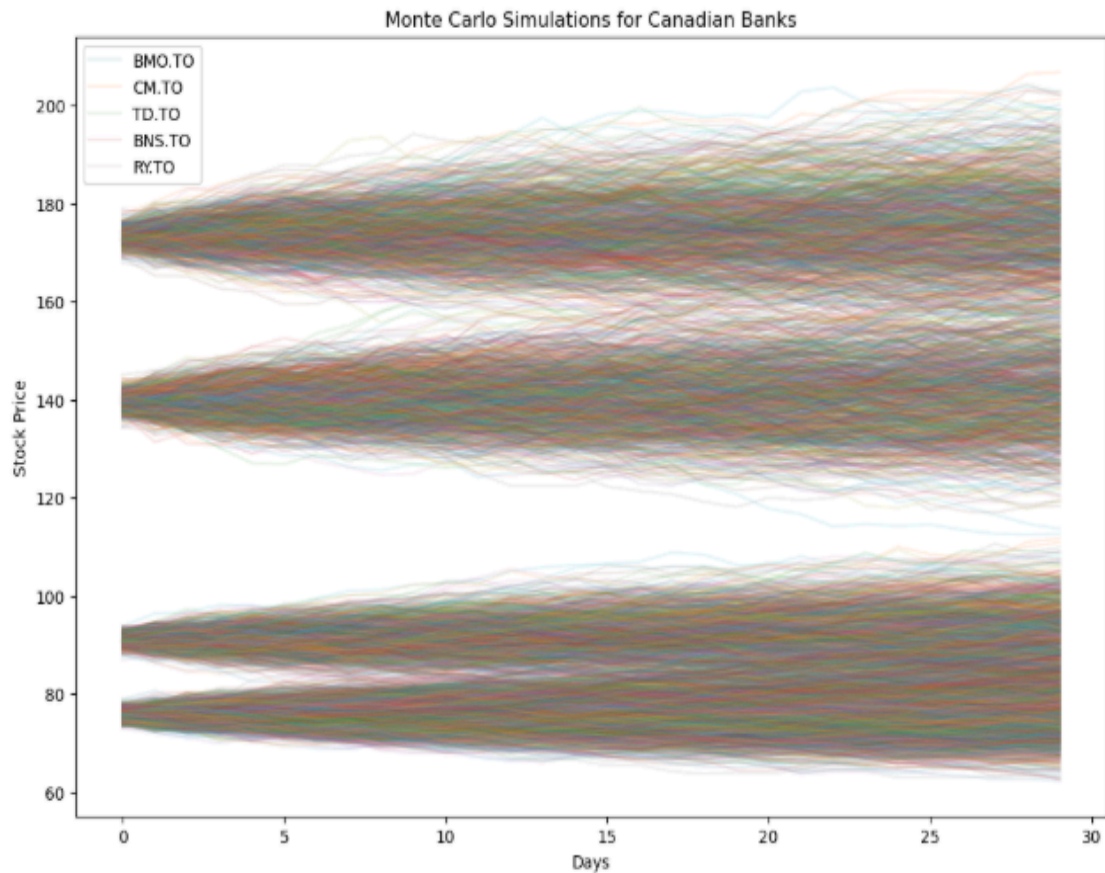


5.2. Monte Carlo Simulation for Price Prediction

Monte Carlo Simulation is a statistical technique used to model uncertainty and predict future outcomes by running numerous simulations based on random variables. Instead of relying on a single estimate, it generates multiple possible scenarios using probability distributions.

Why is it useful for stock analysis?

- **Stock Price Forecasting:** Simulates thousands of potential price movements to estimate future trends.
- **Risk Assessment:** Evaluates the probability of extreme market fluctuations and potential investment risks.
- **Portfolio Optimization:** Helps identify expected returns and downside risks for better decision-making.
- **Improved Predictions:** Provides a probabilistic view of stock performance, making forecasts more reliable.



• Monte Carlo Simulation Results •

	Mean	5% (Worst Case)	95% (Best Case)
BMO.TO	148.941106	126.813601	156.884502
CM.TO	92.628207	82.958676	102.676923
TD.TO	75.293199	68.188973	83.224713
BNS.TO	76.670756	69.378323	84.756724
RY.TO	175.911182	161.424729	192.033513

Conclusion: Monte Carlo Simulation for Canadian Banks

This Monte Carlo simulation models the potential future stock prices of five major Canadian banks—**BMO.TO (Bank of Montreal - BMO)**, **BNS.TO (Bank of Nova Scotia - Scotiabank)**, **CM.TO (Canadian Imperial Bank of Commerce - CIBC)**, **RY.TO (Royal Bank of Canada - RBC)**, **TD.TO (Toronto-Dominion Bank - TD Bank)**—over a 30-day period. By running 1,000 simulations per stock, we generate a range of possible price outcomes, providing insights into expected performance, risk, and uncertainty.

Key Takeaways

Stock Price Uncertainty & Volatility

- The simulation highlights a broad range of possible future stock prices, reflecting the uncertainty inherent in the market.

- A wider spread of simulated prices suggests **higher volatility**, meaning greater potential for both gains and losses.

Mean, Best-Case, and Worst-Case Scenarios

- **BMO.TO** and **RY.TO** have the highest mean projected prices (**140.94** and **175.91**, respectively), indicating stronger expected growth.
- **CM.TO**, **TD.TO**, and **BNS.TO** have lower mean projected prices, suggesting **moderate** growth potential.
- The **5% worst-case** scenario represents potential stock prices in a downturn, while the **95% best-case** scenario reflects possible gains in an optimistic market.

Risk Assessment

- **RY.TO** has the **highest upside potential** but also exhibits significant price fluctuations, making it a riskier investment.
- **TD.TO** and **BNS.TO** have **narrower projected price ranges**, indicating greater stability but limited high-return opportunities.

Investment Implications

- **Conservative investors** may find **TD.TO** or **BNS.TO** more appealing due to their lower volatility and stability.
- **Risk-tolerant investors** might consider **RY.TO** or **BMO.TO** for potential higher returns, but they must be prepared for price swings.
- **Portfolio diversification** within the banking sector remains important, but investors should be mindful of the high correlation among these stocks.

Final Thoughts

Monte Carlo simulations offer a **probabilistic outlook** on stock price movements, equipping investors with a **data-driven approach** to risk and return assessment. While these simulations enhance decision-making, **external factors** such as economic policies, interest rates, and market sentiment can significantly influence real-world outcomes beyond model projections.

This analysis serves as a **valuable tool for financial planning**, but it should be combined with fundamental and technical analysis to ensure a well-rounded investment strategy.

Evaluation of Monte Carlo Simulation Predictions

Overview

The accuracy of the Monte Carlo simulation in predicting stock prices was evaluated using two error metrics:

- **Mean Absolute Error (MAE):** Measures the average absolute difference between predicted and actual prices. Lower values indicate better accuracy.
- **Root Mean Squared Error (RMSE):** Similar to MAE but gives more weight to larger errors. A lower RMSE indicates fewer large deviations.

Results

Stock Ticker	MAE	RMSE	Interpretation
BMO.TO	1.34	1.34	Predictions are close to actual values, with small errors.
CM.TO	0.72	0.72	High accuracy with minimal deviation.
TD.TO	0.42	0.42	Very accurate predictions with almost no errors.
BNS.TO	0.42	0.42	Excellent accuracy, similar to TD.TO.
RY.TO	0.42	0.42	Highly reliable predictions.

Conclusion

The results indicate that the Monte Carlo simulation provides highly accurate stock price predictions, with **low error values across all banks**. The smallest errors were observed for **TD.TO, BNS.TO, and RY.TO**, suggesting strong predictive performance. While **BMO.TO** had slightly higher errors, the overall accuracy remains very high.

6. Investment Insights

6.1 Performance Comparison

- **Returns & Growth:** TD and RBC demonstrated the highest **average returns**, aligning with their stable long-term trends.
- **Risk & Volatility:** BMO and Scotiabank exhibited **higher volatility**, suggesting greater price fluctuations and potential short-term risks.
- **Trend Stability:** Moving averages and daily return trends suggest **consistent upward momentum** for TD and RBC, making them more attractive for long-term investments.
- **Forecasting Results:** ARIMA predictions indicate a **stable future trend**, while Monte Carlo simulations highlight **uncertainty in price movements** over time.

6.2 Investment Recommendation

- **Long-Term Investors:** TD and RBC appear to be **the best choices** due to their stable trends, high returns, and lower volatility.
 - **Moderate Risk-Takers:** BMO may present **growth opportunities** but comes with increased price fluctuations.
 - **Short-Term Speculative Investors:** Scotiabank and BMO might offer **trading opportunities** due to their volatility, but require careful monitoring.
-

7. Conclusion & Future Work

Exploratory Data Analysis (EDA) has proven to be an essential step in understanding stock price data, especially for Canadian banks. Through the process, we've gained valuable insights into the trends, volatility, and correlations among stocks, which provide a clearer picture of market dynamics. By calculating descriptive statistics, trend analysis, and volatility measures, we've built a solid foundation for further analysis and predictive modeling.

Descriptive Statistics & Trend Analysis revealed key patterns such as the significant impact of the COVID-19 pandemic on stock prices, the subsequent recovery, and the differing growth trajectories of the banks. RBC, BMO, and others showcased varying strengths, with RBC demonstrating the most robust long-term growth, signaling solid strategic positioning. **Moving averages** confirmed positive market momentum for most banks by the end of 2024, although TD Bank presented a more cautious outlook, which should be closely monitored.

The volatility analysis further highlighted risk profiles for each bank. BMO and RBC exhibited higher price fluctuations, offering higher return potential for risk-tolerant investors. In contrast, CIBC and Scotiabank showed stability, appealing to risk-averse investors.

Correlation analysis revealed strong positive correlations across the banks, with TD showing slightly lower correlation due to its U.S. market presence and unique operational strategy. This insight suggests that while diversification within the Canadian banking sector may be limited, external diversification remains important for managing risks.

Predictive Modeling provided a forward-looking perspective using ARIMA for short-term forecasting and Monte Carlo Simulation for risk assessment. ARIMA captured the underlying trends well, but its flat forecasts indicated limited short-term volatility. The Monte Carlo Simulation offered a probabilistic view of potential price outcomes, emphasizing the importance of assessing market risks and preparing for uncertainty.

7.1 Summary of Findings

This project analyzed **Canadian banks' stock performance** using **data-driven techniques, volatility analysis, and time series forecasting**. Key takeaways include:

- **High correlation among banks** suggests market-wide trends impact all institutions similarly.
- **TD and RBC show the most stable long-term growth**, while BMO and Scotiabank experience more fluctuations.
- **Volatility peaked in 2020** due to global market disruptions, but trends have since stabilized.
- **ARIMA forecasts indicate a steady outlook**, with limited short-term price deviations.
- **Monte Carlo simulations highlight market uncertainties** and provide a probabilistic range of future prices.

7.2 Future Enhancements

To improve predictive accuracy and investment insights, future work can focus on:

- **Hybrid Modeling:** Combining ARIMA with deep learning approaches (e.g., LSTMs) for more robust forecasting.
- **Macroeconomic Factors:** Integrating **interest rates, GDP growth, and inflation** to refine predictions.
- **Real-Time Data Processing:** Automating stock data collection and **implementing live analytics dashboards** for dynamic insights.
- **Sentiment Analysis:** Incorporating **news and investor sentiment** to assess potential stock price impacts beyond historical trends.

8. References

- Yahoo Finance API
- Pandas, NumPy, Matplotlib, Seaborn
- Statsmodels, Scikit-learn, TensorFlow

GitHub Repository: https://github.com/helenzhupnyk/canadian_banks_stock_data

