

Investment Recommendation System

Providing Personalized Financial Insights

Capstone project.

By

Vinitha Sampathkumar

Date: 02/11/2024

Table of Contents

Content	Page Number
Problem Statement Project Context and Objective Problem Definition Desired Outcome	4
Industry/Domain Overview and Relevance	5
Stakeholders Key Stakeholders and Expectations	5
Business Question Core Question and Value	5
Data Question Data Needs and Requirements	5
Data Sources, Attributes, and Quality	5
Data Science Process	6
Data Analysis Pipeline and EDA Highlights	6-9
Modeling Models, Features, and Metrics	10-13
Outcomes Key Findings	14
Implementation Deployment and User Experience	14
Data and Business Answers	14
Response to Stakeholders	14
End-to-End Solution	14
References	14

Introduction:

Household Savings Ratio



Source: Australian Bureau of Statistics



32.2%

Nearly a third have less than \$5,000 in savings; one-in-six have less than \$1,000.



27.5%

The percentage of people who don't budget at all, jumping to 31.9% of Baby Boomers.



43.2%

The proportion of renters to have 1 month's income or less in savings, twice as likely as homeowners (19.6%).



51.4%

More than half admit to spending money on 'vices' such as alcohol, fast food, smoking/vaping, or drugs.

This infographic sheds light on savings and budgeting habits in Australia, crucial for financial stability and investment readiness:

1. **Low Savings (32.2% have <\$5,000):** A substantial portion of Australians have minimal savings, with one in six holding under \$1,000. Without an emergency fund, many lack the financial stability needed to invest, highlighting the importance of building liquidity before pursuing investment opportunities.
2. **Lack of Budgeting (27.5%):** A large percentage, especially among Baby Boomers, do not budget. Without budgeting, people may overspend, limiting their ability to save and invest. Consistent budgeting supports disciplined saving, which is essential for achieving investment goals.
3. **Renters' Limited Savings (43.2%):** Renters are twice as likely as homeowners to have less than one month's income saved, indicating greater financial vulnerability. This makes it harder for renters to save for investments or homeownership. Renters may benefit from a savings strategy and lower-risk investments to build financial stability over time.
4. **Spending on Vices (51.4%):** Over half of Australians report spending on alcohol, fast food, or smoking, reducing their ability to save or invest. Redirecting these expenses toward savings or investments could significantly improve long-term financial health.
5. **Household Savings Ratio:** The household savings ratio peaked in 2020 due to reduced spending and stimulus support during COVID-19, but has since declined. The current low-savings rate limits disposable income available for investments, underscoring the need for disciplined savings.

These insights suggest that many Australians need to improve budgeting, saving, and spending habits to support investment and wealth-building. Enhancing financial literacy around budgeting and disciplined saving could empower more Australians to allocate funds toward investments aligned with their financial goals.

1. Problem Statement

- **Problem/Opportunity:** Many individuals lack personalized guidance in making investment decisions. This project aims to automate investment recommendations based on user profiles.
- **Value of Addressing the Problem:** Financial planning enhances individual economic security, leading to better financial outcomes.
- **Current State:** Limited accessibility to personalized investment advice.
- **Desired State:** A user-friendly, automated recommendation system providing real-time investment insights.
- **Previous Research:** Reference similar systems in financial planning and their success or limitations.

How can personalized investment recommendations be automated to aid individual financial decision-making?



2. Industry/Domain

- **Industry/Domain:** Financial Investment and Advisory.
 - **Current State:** Increasing challenges from startups offering automated financial advice.
 - **Industry Value Chain:** Data sourcing, user profiling, recommendation, and real-time updates.
 - **Key Concepts:** Investment diversification, personalized finance, stock and commodity markets.
 - **Relevance to Other Industries:** Can extend to industries like banking, financial consulting, and wealth management.
-

3. Stakeholders

- **Identified Stakeholders:** Individual investors, financial advisors, banks, and fintech startups.
 - **Stakeholders' Interest:** To provide accessible and accurate financial advice, improving user investment confidence.
 - **Expectations:** Accurate, real-time, and tailored investment recommendations.
-

6. Data

- **Source:** Financial data sourced from yfinance API, proprietary user data.
- **Volume/Attributes:** Stock (US, AUS), Gold (2022-2024), Real Estate (2022 Q1-2024 Q1); demographics and financial behaviour.
- **Reliability:** Data is regularly updated; sourced from reputable financial platforms.
- **Quality of Raw Data:** Subject to data cleaning and validation.
- **Ongoing Availability:** Data can be accessed in real-time through API.

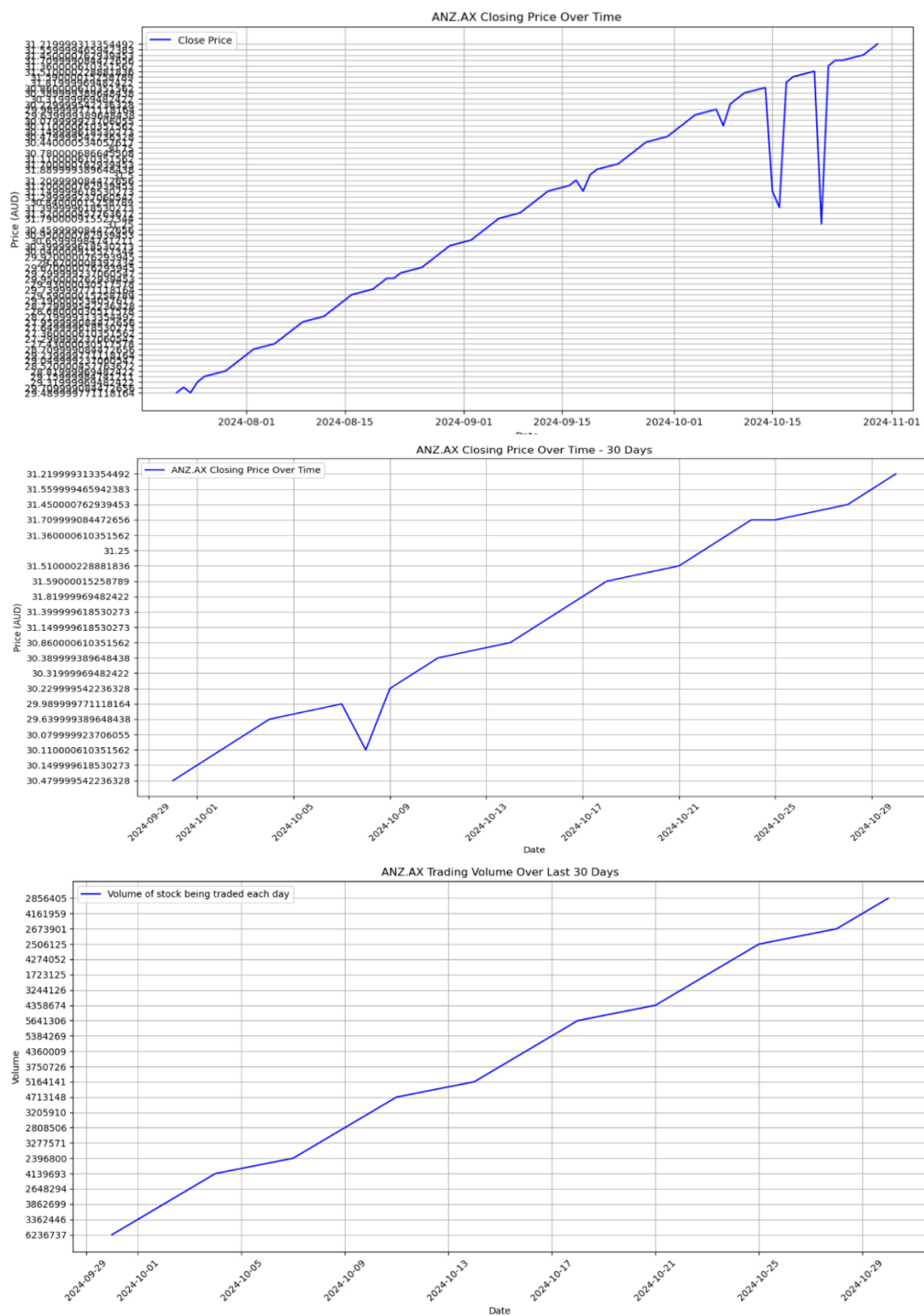


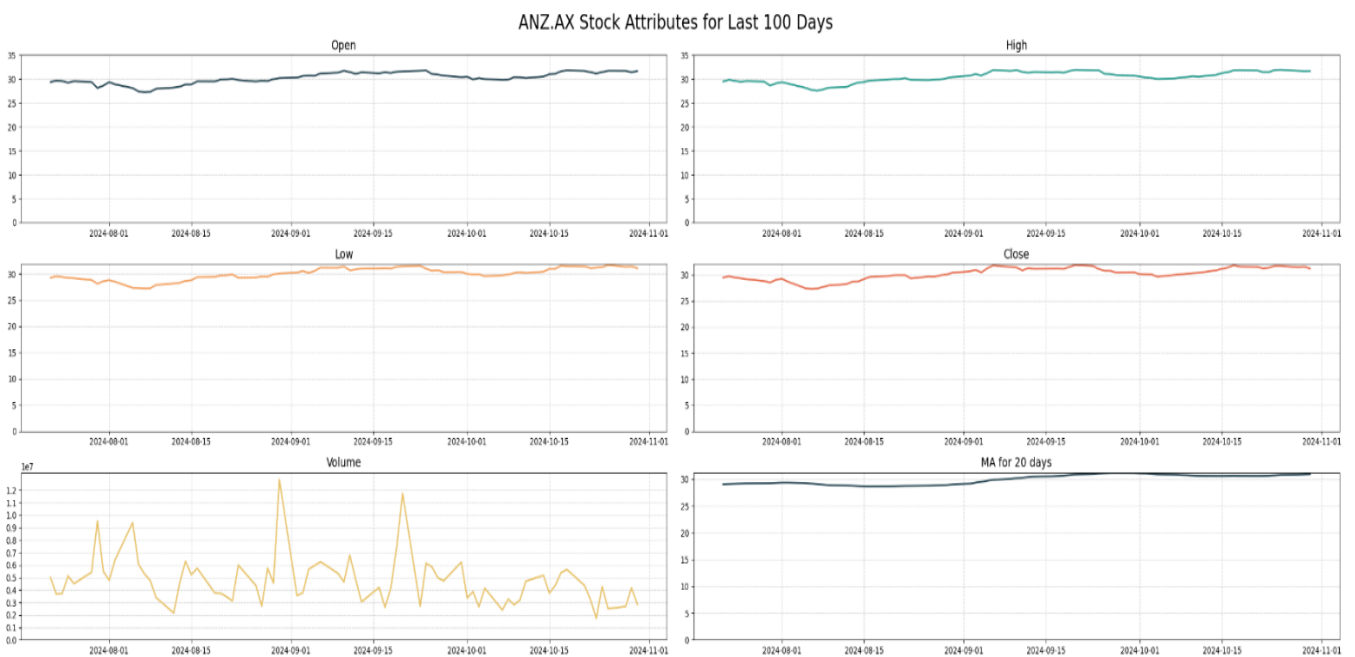
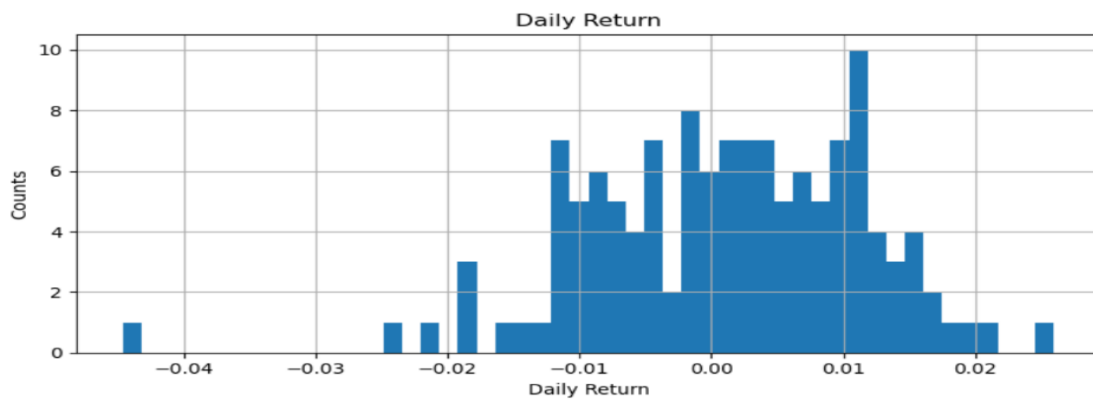
7. Data Science Process

7.1 Data Analysis

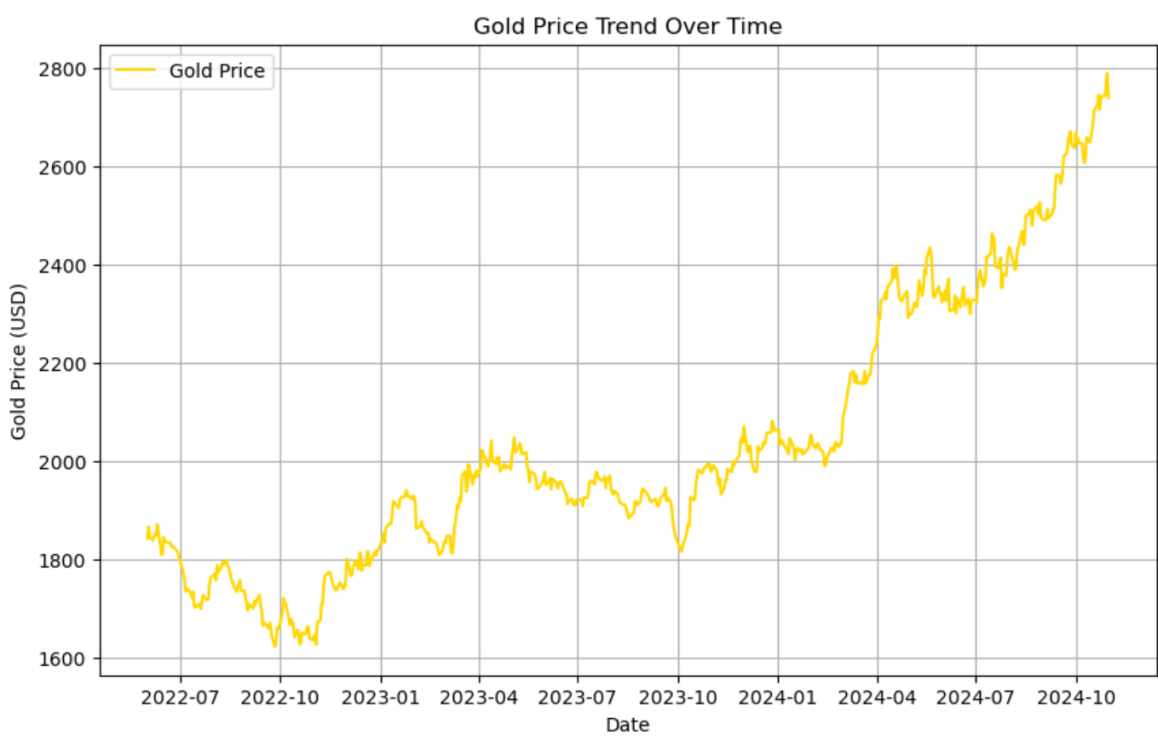
- **Data Pipeline:** Implemented a reusable pipeline for future data processing.
- **Highlights of EDA:** Identified trends in stock, gold, and real estate markets.
- **Intermediary Data Structures:** User profile segments for recommendations.

Stock Recommendation EDA:

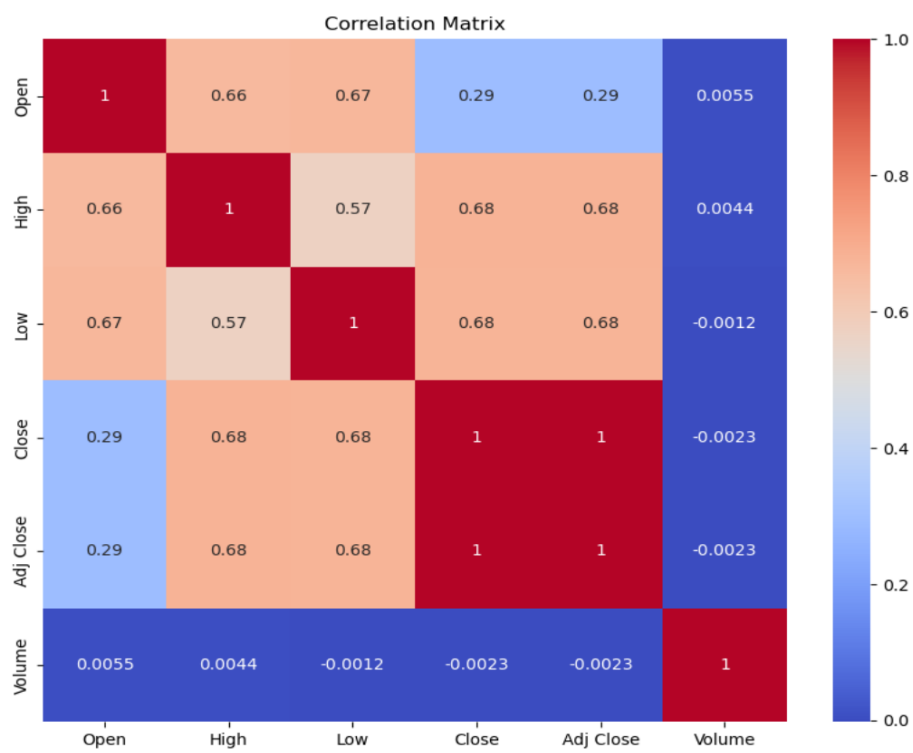




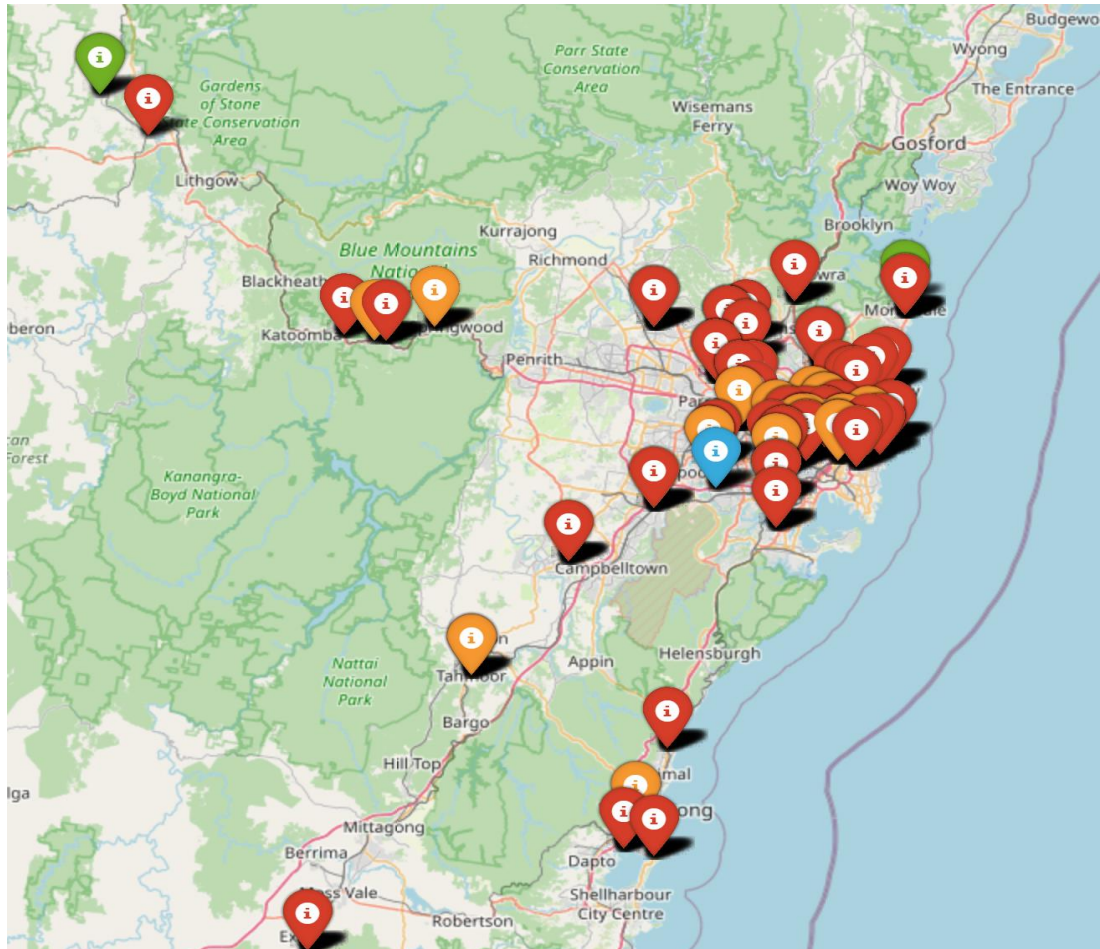
Gold Recommendation EDA:



Used returns (percentage change) instead of raw prices for correlation analysis



Real-estate Recommendation EDA:



7.2 Modeling

- **Features:** Investment goals, demographics, market trends.
- **Feature Selection:** Selected based on financial relevance; engineered returns for better prediction.
- **Models Used:**
 - Stock: Linear Regression, RandomForestRegressor, LSTM.
 - Gold: Linear Regression, Prophet.
 - Real Estate: Linear Regression.
- **Training Time:** Reasonable for each model; optimized through efficient feature engineering.
- **Tools Used:** Streamlit, Python, yfinance, and relevant machine learning libraries.
- **Performance Metrics:** RMSE, MAE, and accuracy based on each investment type.
- **Model Selection:** The model with the best performance for each asset class.

Stock Recommendation:

Models:

1. Linear Regression
2. RandomForestRegressor + Gridsearch
3. LSTM

Gold Recommendation:

Models:

1. Linear Regression
2. Prophet

Real-estate Recommendation:

Model:

1. Linear Regression

Stock Recommendation:

Linear Regression

Mean Absolute Error (MAE): 0.6793948279560473
 Mean Squared Error (MSE): 0.734262860734546
 R-squared (R^2): 0.36662203996947373

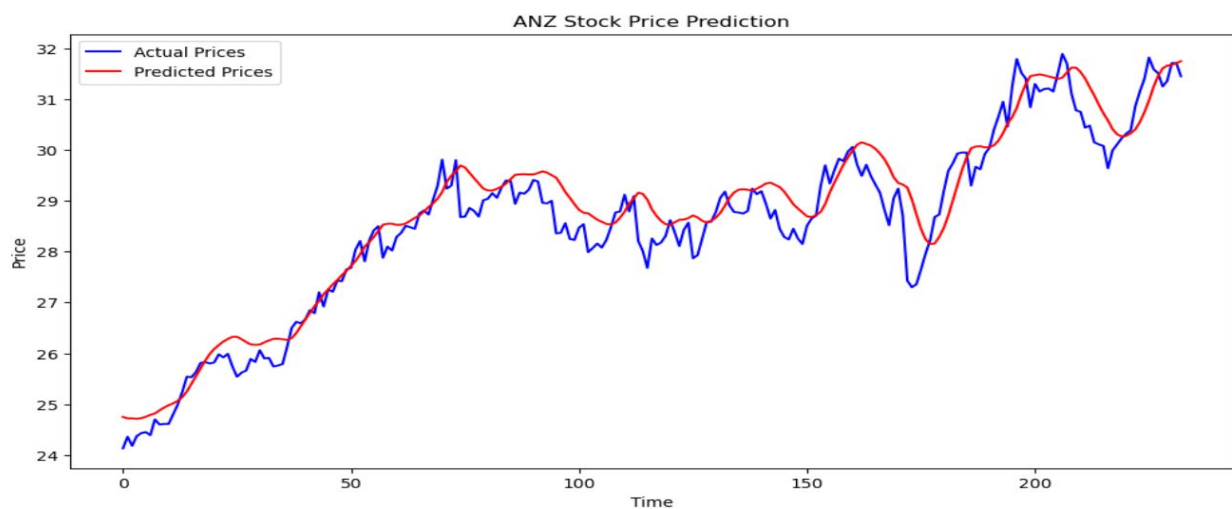
RandomForestRegressor + Grid search

Mean Absolute Error (MAE): 0.516178674954627
 Mean Squared Error (MSE): 0.6658547583362627
 R-squared (R^2): 0.42563113148642606

Long short-term memory

Mean Absolute Error (MAE): 0.3928506057149863
 Mean Squared Error (MSE): 0.2559799158422873
 Root Mean Squared Error (RMSE): 0.5059445778366315

Best Model Actual vs predicted Stocks



	Actual Price	Predicted Price	Signal
213	30.150000	31.072933	Buy
214	30.110001	30.890718	Buy
215	30.080000	30.722231	Buy
216	29.639999	30.580843	Buy
217	29.990000	30.413321	Hold
218	30.110001	30.309423	Hold
219	30.230000	30.266340	Hold
220	30.320000	30.278135	Hold
221	30.389999	30.329885	Hold
222	30.860001	30.405394	Hold
223	31.150000	30.548460	Hold
224	31.400000	30.743200	Sell
225	31.820000	30.970787	Sell
226	31.590000	31.240734	Hold

Gold Recommendation:

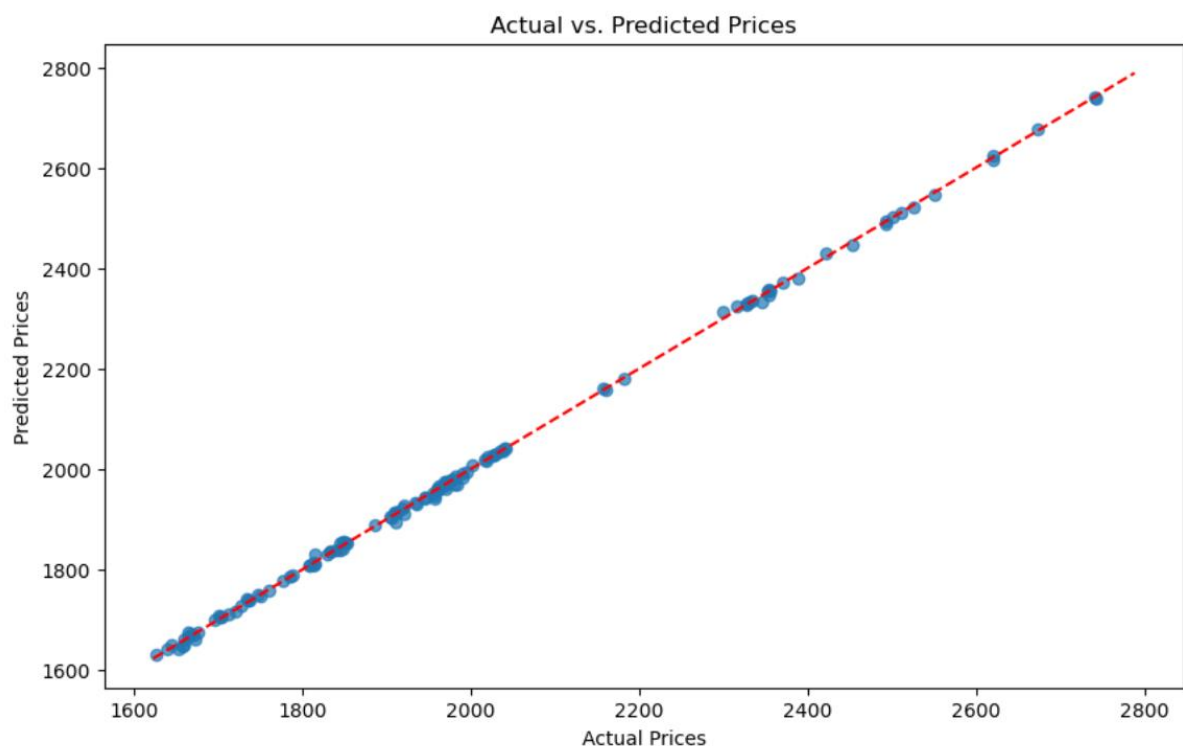
Linear Regression

Mean Squared Error: 30.724012592997564
R² Score: 0.99959085160806

Prophet

MAPE: 2.71%
MAE: 64.85

Best Model Actual vs predicted Gold price



	Actual Price	Predicted Price	Decision
0	1970.500000	1961.093459	Sell
1	1945.500000	1940.971893	Sell
2	2354.800049	2358.412549	Buy
3	2493.399902	2487.876903	Sell
4	1760.800049	1757.693298	Sell

Real-estate Recommendation:

Linear Regression

Mean Squared Error: 37306.83

Actual_Median_Sales_Price	Predicted_Median_Sales_Price
890	916.217876
1103	1006.057187
600	655.741625
1155	1121.321913
2100	1763.861558
1030	792.852349
1598	1078.992454
800	898.904119
920	1094.687567
1046	1136.104838
549	548.636128

8. Outcomes

- **Key Findings:** Personalized recommendations enhance user investment confidence.
 - **Conclusion:** Each model accurately predicts respective investment outcomes, providing actionable insights.
-

9. Implementation

- **Production Considerations:** Regular data updates, UI/UX design for ease of use, real-time updates.
-

10. Data Answer

- **Answer to Data Question:** Satisfactory; high confidence in recommendation accuracy.
-

11. Business Answer

- **Answer to Business Question:** Satisfactory; high confidence that recommendations meet user expectations.
-

12. Response to Stakeholders

- **Message and Recommendations:** Automated investment system provides tailored financial insights, improving investment decisions.
-

13. End-to-End Solution

- **Solution Summary:** A Streamlit web app that personalizes financial advice based on real-time data and user inputs.
-

14. References

- **Data & Code:** Main items include Jupyter notebooks for each asset class, datasets, and exported models.
- **Resources:** Python libraries (e.g., yfinance, pandas), algorithms (Linear Regression, LSTM, Prophet), and Streamlit framework.

Disclaimer

The contents of this document do not constitute investment advice. All information provided is for educational and informational purposes only and should not be considered as a substitute for professional financial advice. Readers should seek independent financial advice tailored to their individual circumstances.