STUDENT ACADEMIC PERFORMANCE FORECASTING ANALYSIS REPORT

Objective: To analyze the academic performance (GPA) of a representative student (ID: 1001) and forecast their performance for the next three semesters using various time series forecasting models. The goal is to identify the most accurate model and generate a reliable future projection.

1. Executive Summary

This report details the analysis and forecasting of a student's GPA based on simulated historical data. A time series for a single student (ID: 1001) was constructed from 18 semesters (Spring 2020 - Fall 2025). The dataset was split, using the first 15 semesters (Spring 2020 - Fall 2024) for training and the last 3 semesters (Spring 2025 - Fall 2025) for model validation.

Three forecasting models were evaluated: ARIMA, Seasonal ARIMA (SARIMA), and Holt-Winters Exponential Smoothing. The models' performance was assessed using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE).

The **ARIMA model** was found to be the most accurate for this specific student's data, achieving the lowest error metrics across all categories. Based on this best-performing model, a final forecast for the next three semesters (Spring 2026 - Fall 2026) was generated, predicting a consistent GPA of approximately **3.05**.

2. Methodology

The analysis followed a structured workflow:

2.1 Data Simulation & Preparation:

A synthetic dataset of 500 students' GPAs from 2020-2025 was generated. The data for the first student (ID: 1001) was isolated and structured into a quarterly time series.

2.2 Train/Test Split:

The time series of 18 data points was split into a training set (first 15 semesters) and a test set (last 3 semesters) to simulate a real-world forecasting scenario.

2.3 Model Training & Forecasting:

Three distinct forecasting models were trained on the 15-semester training data:

- **ARIMA (AutoRegressive Integrated Moving Average):** A non-seasonal model. The notebook attempted to automatically determine the optimal parameters (auto_arima) but defaulted to a simple order due to the limited data.

- **SARIMA (Seasonal ARIMA):** A seasonal extension of ARIMA, using a seasonal period of 3 to account for the Spring/Summer/Fall academic cycle.
- **Holt-Winters Exponential Smoothing:** A model that explicitly captures trend and seasonality. An additive seasonal component was used.

2.4 Performance Evaluation:
The 3-semester forecasts from each model were compared against the actual held-out test data. The error metrics (RMSE, MAE, MAPE) were calculated to objectively determine the best model.

2.5 Final Forecasting:
The best-performing model was re-trained on the *entire* 18-semester dataset to generate a final 3-semester forecast (Spring 2026 - Fall 2026).

3. Model Performance Analysis & Results

The performance of each model on the test data is summarized in the table below. Lower values indicate better accuracy

Model	RMSE	MAE	MAPE
ARIMA	**0.1795**	**0.1611**	**5.50%**
SARIMA	0.1992	0.1637	5.55%
Holt-Winters (add)	0.2435	0.2074	6.95%

3.1 ARIMA:
The ARIMA model provided the most accurate forecast for this student, exhibiting the lowest error across all three metrics. Its MAPE of 5.50% suggests a strong predictive capability. The model's forecast appeared as a flat line, indicating it may have essentially predicted the historical mean for the future.

3.2 SARIMA:
The SARIMA model performed similarly to ARIMA but with slightly higher RMSE. This could indicate that the seasonal component was either not very strong in the test period or was not captured perfectly with the limited data.

3.3 Holt-Winters:
This model had the highest error rates. Its forecast seemed to attempt to extrapolate a downward trend, which did not align well with the actual test values, leading to its lower accuracy.

3.4 Conclusion of Model Selection:
Based on its superior performance on the validation set, the **ARIMA model was selected as the best model** for generating the final forecast.

4. Final Forecast

The selected ARIMA model was re-trained on all 18 historical data points for student 1001. The forecast for the next three semesters (Spring 2026 - Fall 2026) is presented below.

Semester	Date	Forecasted GPA
Spring 2026	2026-03-01	3.053
Summer 2026	2026-07-01	3.053
Fall 2026	2026-11-01	3.053

4.1 Interpretation:
The ARIMA model forecasts a stable GPA of approximately **3.05** for the next three semesters. This flat forecast suggests that based on the historical pattern, the model does not detect any significant upcoming upward or downward trend in the student's academic performance. This projection should be viewed as a central estimate, with the understanding that actual future performance will likely vary around this mean.

5. Limitations & Considerations

The findings and forecast in this report should be interpreted with the following limitations in mind:

- **Simulated Data:** The analysis was conducted on **simulated, not real, student data**. While the simulation aimed to mimic real-world patterns, the results are purely illustrative and cannot be generalized to actual students.

- **Limited Data:** A time series of 18 data points is relatively small for robust time series modeling, especially for models with seasonal components. The performance of the models could change significantly with a longer historical record.
- **Single Student Focus:** The analysis was performed for only one student. A model that performs best for this student may not be optimal for another with a different academic trajectory.
- **Model Simplification:** The auto_arima function selected a simple (0,0,0) order for the non-seasonal ARIMA model, which essentially means it predicted a constant mean. This is a plausible outcome for a stable series but may not capture more complex underlying patterns.

6. Conclusion

This analysis successfully built and evaluated three time series forecasting models to predict a student's GPA. On the simulated data for student 1001, the ARIMA model proved to be the most accurate, achieving a MAPE of 5.50% on the validation set. Its final forecast projects a stable GPA of 3.05 for the subsequent three semesters. This exercise demonstrates a robust framework for academic performance forecasting that can be applied to real student data to provide early warnings and support data-driven academic advising.