Forecasting Future Educational Costs



Aikaterini Antoniou

EC Utbildning

Projekt i Data Science

2024-10

# Abstract

This study investigates the forecast of births and educational costs for primary and secondary education in Sweden over a 10-year period (2025–2035). Various predictive models, including ARIMA, SARIMA and LSTM, were employed to produce reliable estimates of birth numbers and total per-child costs. Additionally, a comparative analysis of these models was conducted to determine the most suitable model for this application. Cost calculations based on adjusted 2022 costs were integrated to assess future educational funding needs. This analysis offers guidance for future educational resource planning, aiding in financial decision-making for education per region.

Abbreviations and Terms  
  
ARIMA – Auto-Regressive Integrated Moving Average

SARIMA – Seasonal Auto-Regressive Integrated Moving Average

LSTM – Long Short-Term Memory

MLP – Multi-Layer Perception

GBR – Gradient Boosting Regression

RF – Random Forest

ADF Test – Augmented Dickey-Fuller Test

MSE – Mean Squared Error

MAE – Mean Absolute Error

RMSE – Root Mean Squared Error

SEK – Swedish Krona

Grundskola – Primary Education

Gymnasieskola – Secondary Education

Table of Contents

[Abstract 2](#_Toc181662458)

[Abbreviations and Terms 3](#_Toc181662459)

[1 Introductions 1](#_Toc181662460)

[1.1 Purpose of this study 1](#_Toc181662461)

[2 Theory 2](#_Toc181662462)

[2.1 Time Series Models 2](#_Toc181662463)

[2.1.1 ARIMA 2](#_Toc181662464)

[2.1.2 SARIMA 2](#_Toc181662465)

[2.2 Neural Networks 2](#_Toc181662466)

[2.2.1 LSTM 2](#_Toc181662467)

[2.2.2 MLP 2](#_Toc181662468)

[2.3 Machine Learning Approaches 2](#_Toc181662469)

[2.3.1 Linear Regression 2](#_Toc181662470)

[2.3.2 Random Forest 3](#_Toc181662471)

[2.3.3 GBR 3](#_Toc181662472)

[2.4 Statistical Analysis Tools 3](#_Toc181662473)

[2.4.1 ADF Test 3](#_Toc181662474)

[3 Method 4](#_Toc181662475)

[3.1 Data Collection 4](#_Toc181662476)

[3.1.1 Description of Sources 4](#_Toc181662477)

[3.1.2 Data Cleaning and Preparation 4](#_Toc181662478)

[3.2 Model Development 4](#_Toc181662479)

[3.2.1 LSTM 4](#_Toc181662480)

[3.2.2 ARIMA & SARIMA 4](#_Toc181662481)

[3.3 Forecasting Models 5](#_Toc181662482)

[3.4 Educational Cost Calculation 5](#_Toc181662483)

[3.5 Comparative Analysis 5](#_Toc181662484)

[4 Results and Discussion 6](#_Toc181662485)

[4.1 Presentation of Results 6](#_Toc181662486)

[4.1.1 Birth Predictions 6](#_Toc181662487)

[4.2 Analysis of Educational Costs 7](#_Toc181662488)

[4.2.1 Projected Educational Costs (Grundskola and Gymnasieskola) 7](#_Toc181662489)

[4.2.2 Total Educational Costs and Trends 7](#_Toc181662490)

[4.3 Interpretation and Discussion 7](#_Toc181662491)

[4.3.1 Interpretation of Forecasts and Model Evaluation 7](#_Toc181662492)

[4.3.2 Limitations and Suggestions for Improvement 8](#_Toc181662493)

[4.3.3 Suggested Strategies for Resource Allocation 8](#_Toc181662494)

[5 Conclusion 9](#_Toc181662495)

[6 Appendix A 10](#_Toc181662496)

[References 44](#_Toc181662497)

# Introductions

In recent years, accurately forecasting future educational expenses has become critical for government agencies, educational institutions and policymakers. With changing demographic trends, including birth rates and population distributions, predicting educational needs and associated costs is vital to ensure resources are appropriately allocated and future generations are well-supported. This study examines how predictive models, using machine learning and statistical techniques, can offer insights into future educational costs at different levels of schooling, particularly in primary (Grundskola) and secondary (Gymnasieskola) education.

By leveraging birth rate projections and education cost data, this project aims to provide a comprehensive forecast of educational costs for the years 2025 to 2035. The integration of LSTM, SARIMA, and Random Forest models enables us to analyze long-term trends in population growth and align them with anticipated educational expenditures. This predictive analysis helps address critical questions for policymakers regarding budget planning and the efficient allocation of resources.

## Purpose of this study

The purpose of this report is to develop a predictive model to estimate future educational expensesby addressing two key questions:

-How can population forecasts be effectively integrated with historical educational cost data to estimate future expenditures?

-What are the expected trends in educational costs across different levels of schooling over the next decade?

# Theory

## Time Series Models

### ARIMA

The ARIMA model is a statistical time series model particularly suited for forecasting data with linear trends. This model combines autoregressive terms, moving averages and differencing to manage trends and seasonality. In this study, ARIMA was applied to forecast educational expenditures based on historical birth rates and cost data, leveraging its linear assumption for trend prediction. (Saadeddin, Z. (2024, September 9). ARIMA for time series forecasting: A complete guide. DataCamp. Retrieved from https://www.datacamp.com/tutorial/arima)

### SARIMA

SARIMA extends ARIMA by incorporating seasonal components, which account for periodic fluctuations within the data. This model is beneficial when seasonality is evident, as in birth rates that may display annual or quarterly patterns. SARIMA enabled a more nuanced analysis of birth rate data, capturing both short-term and seasonal variations. (Brownlee, J. (2019, August 21). A gentle introduction to SARIMA for time series forecasting in Python. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/sarima-for-time-series-forecasting-in-python/)

## Neural Networks

### LSTM

LSTM networks, a variant of recurrent neural networks, are designed for long-term dependencies within sequential data, making them effective for time series prediction. In this study, LSTM was implemented to forecast birth rates based on historical sequences, maintaining relationships over extended periods and yielding robust predictions. (Brownlee, J. (2022, August 7). Time series prediction with LSTM recurrent neural networks in Python with Keras. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/; Brownlee, J. (2020, August 28). How to develop LSTM models for time series forecasting. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/)

### MLP

The MLP is a basic form of neural network consisting of layers with interconnected nodes (neurons). Initially, MLP was tested for birth rate prediction to assess its effectiveness in this domain, though it was ultimately not selected for final analysis due to its limitations in handling sequential dependencies inherent in time series data. (Brownlee, J. (2020, August 28). How to develop multilayer perceptron models for time series forecasting. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/how-to-develop-multilayer-perceptron-models-for-time-series-forecasting/)

## Machine Learning Approaches

### Linear Regression

Linear regression served as an initial, simple model for predicting educational expenses. This model is based on the assumption of a linear relationship between the predictor and the target variable. It provided a baseline for comparison against more complex, nonlinear models used in subsequent stages. (Frost, J. (n.d.). How to choose between linear and nonlinear regression. Statistics by Jim. Retrieved from https://statisticsbyjim.com/regression/choose-linear-nonlinear-regression/)

### Random Forest

Random Forest, an ensemble learning method, was explored for its ability to handle feature importance and improve prediction accuracy through the aggregation of multiple decision trees. Although useful for preliminary insights, it was ultimately not chosen for final forecasting due to its reduced performance in comparison to LSTM and ARIMA in time series contexts. (Brownlee, J. (2020, November 1). Random forest for time series forecasting. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/random-forest-for-time-series-forecasting/)

### GBR

Gradient Boosting Regressor was employed to enhance model accuracy by sequentially minimizing error. Coupled with GridSearch, this method enabled the optimization of hyperparameters, contributing to a better understanding of the data. However, its forecasting performance did not surpass the selected ARIMA and LSTM models. (Brownlee, J. (2021, March 19). How to use XGBoost for time series forecasting. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/xgboost-for-time-series-forecasting/; Rosebrock, A. (2021, May 24). Grid search hyperparameter tuning with scikit-learn and GridSearchCV. PyImageSearch. Retrieved from https://pyimagesearch.com/2021/05/24/grid-search-hyperparameter-tuning-with-scikit-learn-gridsearchcv/)

## Statistical Analysis Tools

### ADF Test

The ADF test was utilized to assess the stationarity of time series data, ensuring that the data met the stationarity assumption critical for ARIMA and SARIMA models. This test is fundamental for determining appropriate differencing parameters and optimizing model performance, particularly for time series with potential trends or seasonality. (Machine Learning Plus. (n.d.). Augmented Dickey Fuller test (ADF test) – Must read guide. Retrieved from https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/)

# Method

In this section, the methods used for forecasting birth rates and calculating educational costs are outlined. Various predictive models were employed, including ARIMA, SARIMA, MLP, Random Forest, Gradient Boosting, and LSTM, to capture potential trends in birth rates and evaluate future educational expenses in Sweden.

## Data Collection

The data collection phase was essential for the successful development of our predictive model. The data used in forecasting educational expenses were gathered from various reliable sources, primarily the Swedish Statistics Agency (SCB Statistical Database). This data included historical records of births, migration, and mortality for each region, along with the educational costs associated with primary (Grundskola) and secondary (Gymnasieskola) education up to the year 2022. SQLite was chosen as the database solution for this project due to its simplicity and efficiency in handling small to medium datasets. It enabled seamless integration of various data sources, such as regional birth data, population data, and mortality statistics. The structure of SQLite allowed for efficient querying and combination of tables, ensuring data coherence and accessibility for analysis, especially for time-series and regional data modeling.

### Description of Sources

The datasets, stored in CSV format, contained information from the years prior to 2022 and included details such as birth counts, population distributions and educational costs per age group and school level. Each dataset was tailored to fit the requirements of the analysis and forecasting models.

### Data Cleaning and Preparation

Initially, the data underwent cleaning to ensure accuracy and consistency, including handling any missing values or outliers that might impact model quality. Data normalization was also performed, especially necessary for neural network processing, as the varying scales could introduce errors. The cleaned and normalized data were saved locally, ready for input into the models.

## ModelDevelopment

The model development process involved selecting and tuning various forecasting algorithms to achieve optimal accuracy in predicting births and, subsequently, educational expenses. For our data analysis, three main models were applied: LSTM, ARIMA, and SARIMA, with additional approaches tested for comparative purposes.

### LSTM

The LSTM model, a type of recurrent neural network, was chosen for its ability to handle long-term sequences and detect relationships within historical birth data. The data was preprocessed through normalization to fit the LSTM’s scale, and the sequence was split into windows, allowing the network to train on data sequences.

The LSTM was trained with initial parameters, including the number of epochs and batch size. After training, the model was employed to forecast births over the next decade (2025–2035). The LSTM provided accurate forecasts by leveraging distant correlations for greater precision.

### ARIMA & SARIMA

The ARIMA model, a statistical method, was used for birth forecasting by analyzing linear relationships in the historical data. The ARIMA model was initially chosen as it suits data that exhibit stability after differencing. Various parameter values (p, d, q) were tested based on the results of the Augmented Dickey-Fuller Test to ensure model suitability.

The SARIMA model extended ARIMA by incorporating seasonal elements, which is useful when time series data show seasonal fluctuations. The SARIMA model was also applied to the birth data analysis, with model parameters (P, D, Q, S) selected for optimal performance. After comparison, the ARIMA model demonstrated the best performance and was chosen as the primary model for forecasting.

## Forecasting Models

To ensure robust and accurate predictions, different forecasting models were developed and optimized.

* **ARIMA and SARIMA**: These models were applied to forecast birth rates by using the historical birth data. ARIMA (Auto-Regressive Isntegrated Moving Average) is a standard time series model suitable for univariate forecasting. SARIMA (Seasonal ARIMA) includes seasonal parameters to capture periodic patterns. The optimal parameters (p, d, q) for ARIMA and (P, D, Q, s) for SARIMA were selected through grid search methods to minimize the Mean Squared Error (MSE).
* **MLP (Multi-Layer Perceptron)**: MLP, a type of artificial neural network, was used to model complex relationships in the birth rate data. The model was trained using historical birth data, with hidden layers optimized to enhance prediction accuracy. A grid search was performed to determine the optimal structure and learning rate.
* **Random Forest**: This ensemble learning method was also implemented for birth rate prediction. By combining multiple decision trees, Random Forest provided a robust model capable of handling non-linear relationships. Hyperparameters such as the number of trees, maximum depth, and minimum samples split were optimized to reduce prediction error.
* **Gradient Boosting**: Another ensemble method, Gradient Boosting, was used to predict birth rates by minimizing errors in each iteration. The learning rate, number of estimators, and maximum depth were optimized through hyperparameter tuning.
* **LSTM (Long Short-Term Memory)**: As a recurrent neural network model, LSTM is suitable for sequential data like time series. The model was trained on birth rate data to capture long-term dependencies and trends. Multiple configurations were tested to ensure effective learning.

## Educational Cost Calculation

The projected birth rates were then used to calculate the total educational costs for Grundskola and Gymnasieskola from 2025 to 2035. For each predicted birth rate, the total cost was calculated by multiplying the cost per child by the projected number of births. These projections provided insights into the future financial requirements for each educational level.

## Comparative Analysis

To assess the accuracy and reliability of each model, the results were compared based on Mean Squared Error (MSE). This comparison allowed for the identification of the most accurate model for forecasting future birth rates, facilitating more precise calculations of educational costs.

# Results and Discussion

## Presentation of Results

Using SQLite, data from diverse sources, including birth rates and regional population demographics, were combined to generate comprehensive datasets that informed the predictions in this study. By linking tables through common identifiers like region and year, critical relationships between birth rates and population data were established, facilitating the calculation of forecasts and trends in educational expenditures.

### Birth Predictions

Birth predictions for the period 2025–2035 were conducted using ARIMA, SARIMA and LSTM models. These predictions are analyzed and visualized through graphs, with each model presenting its own estimates for future birth rates over the next decade.

Each forecasting model provided a different outlook on future birth rates, as demonstrated by the respective MSE values:

ARIMA and SARIMA: The ARIMA model predicted a gradual decline in birth rates over the forecast period, while SARIMA provided similar results but with adjustments for seasonal variations. SARIMA generally yielded a lower MSE, indicating slightly better predictive accuracy.

MLP and Random Forest: The MLP model captured complex, non-linear trends but had a higher MSE compared to SARIMA, suggesting potential overfitting. The Random Forest model produced a moderate MSE, performing better than MLP in capturing birth trends but still lagging behind SARIMA.

Gradient Boosting and LSTM: Gradient Boosting achieved a lower MSE than Random Forest, indicating its strength in handling sequential errors and making accurate forecasts. LSTM, while capturing long-term dependencies, showed slightly higher MSE compared to Gradient Boosting, possibly due to the small sample size of the dataset.

The model with the lowest MSE, SARIMA, was chosen as the most reliable forecast for the purposes of this study.

#### ARIMA Model

The ARIMA model predictions are based on linear correlations in historical birth data, focusing on identifying trends and patterns. ARIMA produced satisfactory results in linear analysis, with forecasts that are stable and align with the general trend of increasing or decreasing births.

#### SARIMA Model

Incorporating seasonal parameters, the SARIMA model offered predictive capabilities for data with clear seasonal fluctuations. The SARIMA forecasts were more accurate than ARIMA in periods with seasonal variations, though they showed greater deviation in areas without strong seasonal trends.

#### LSTM Model

The LSTM, as a neural network capable of handling long-term data relationships, produced forecasts that were exceptionally smooth and stable, though less affected by seasonal fluctuations. While the data sequence provided high accuracy in stable trends, the model was limited in detecting changes arising from nonlinear and seasonal shifts.

#### ModelComparison

To assess model accuracy, metrics such as Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) were calculated for ARIMA, SARIMA, and LSTM. The comparison was based on the following observations:

Performance in Stable Trends: Both ARIMA and LSTM performed satisfactorily in trends without strong seasonal fluctuations, while SARIMA excelled in accuracy with seasonal trends.

Seasonal Variations: SARIMA performed best when there were significant seasonal fluctuations, while ARIMA maintained more stable forecasts overall.

Smoothness and Consistency: The LSTM provided the smoothest predictions, though without high sensitivity to seasonal changes, making it the least suitable for data with frequent variations over short time periods.

## Analysis of Educational Costs

### Projected Educational Costs (Grundskola and Gymnasieskola)

The analysis of educational costs for the years 2025–2035 revealed distinct trends for both Grundskola (primary education) and Gymnasieskola (secondary education). For modeling the projected educational expenses, birth rates per region were utilized and unit costs per student at each educational level. This approach allowed us to calculate total expenses for each period, factoring in the adjustments in per-student costs provided in the data.

The cost prediction graphs (for each educational category and region) indicate specific upward or steady trends. For example:

Grundskola: In most regions, costs show a steady increase, aligned with the rising number of births and the influx of new students. The ARIMA model, chosen for its accuracy in predictions, along with per-student costs, confirms this trend as increases in costs correspond to a general rise in student population.

Gymnasieskola: Similarly, Gymnasieskola costs demonstrate an upward trend over time, although the growth rate varies across regions. Projections indicate that this cost might increase more slowly in some regions, suggesting a smaller rise in the number of students advancing to secondary education.

### Total Educational Costs and Trends

Overall, the results indicate that national-level educational expenses for Grundskola and Gymnasieskola are expected to follow an upward trajectory, although this growth is relatively gradual. The factors influencing this trend include:

Demographic Changes: Population growth leads to higher student enrollment at educational levels.

Regional Differences: Certain regions exhibit varying growth rates in costs, reflecting either population growth or decline in specific areas.

These trends are essential for policymakers as they provide a clear indicator of future financial needs in educational resources. By emphasizing both the increase in students and per-student costs, our analysis suggests strategic insights for the efficient allocation of educational resources at both national and regional levels.

## Interpretation and Discussion

### Interpretation of Forecasts and Model Evaluation

The analysis indicates that the selected models, such as ARIMA and LSTM, were effective in forecasting educational costs and student enrollment. Key insights include:

Forecast Accuracy: Time series models like ARIMA proved especially useful for birth predictions, offering accurate trends over short periods.

Model Differences: While the LSTM model showed promise for capturing long-term trends, ARIMA was more effective for short-term forecasts and seasonal fluctuations, making it the primary choice for this study.

Using the SARIMA model predictions, the educational costs for Grundskola and Gymnasieskola were calculated for each year from 2025 to 2035. The projections indicate:

Grundskola: Costs are expected to gradually increase, with some fluctuation due to variations in predicted birth rates. This trend aligns with anticipated inflation and adjustments in educational resource allocation.

Gymnasieskola: Similar toGrundskola, costs are projected to increase over time. However, the growth rate in Gymnasieskola costs appears slightly steeper, suggesting that secondary education may require more resources per student in the coming years.

### Limitations and Suggestions for Improvement

The study's limitations include reliance on demographic data that may not fully capture future trends. Demographic shifts or migration waves might not be adequately represented.

For enhancing the study, the following suggestions are proposed:

Frequent Data Updates: Regularly updated data will increase forecast accuracy.

Advanced Forecasting Techniques: Integrating additional models, such as SARIMA or hybrid models, can improve accuracy.

Strategic Resource Storage: Flexible strategies for allocating resources by region are recommended to ensure that high-growth areas can meet increased demand.

### SuggestedStrategies for ResourceAllocation

The projected increase in educational costs, particularly for Gymnasieskola, underscores the need for budgetary adjustments and long-term planning in educational institutions. The gradual rise in costs necessitates proactive resource allocation to meet future demands, especially in high-growth regions as indicated in the regional analysis.

To ensure optimal use of the results, the following strategies are recommended:

Adjusting Educational Budgets Annually: Budgets can be adapted annually, taking into account updated projections for educational costs and student enrollment.

Regional Adjustments: Resource allocation should reflect growth trends by region to address the needs of high-growth areas effectively.

Adopting Flexible Prediction Models: Implementing multiple models and integrating forecasts, such as ARIMA and LSTM, will enhance accuracy and adaptability across regions.

# Conclusion

The predictive analysis presented in this report offers a comprehensive view of the anticipated educational costs in Sweden over the coming decade, leveraging advanced time series and machine learning models to estimate both student enrollment and associated expenses. By integrating birth rate data and educational expenditure trends through models like ARIMA, SARIMA, and LSTM, the study provides an informed projection of the financial resources required for primary and secondary education. The findings indicate a steady rise in educational costs, driven by demographic patterns and growing demand in certain regions. While the ARIMA model proved the most suitable for short-term predictions, the flexibility and potential of LSTM for long-term insights underscore the value of combining different approaches for a robust forecasting strategy.

Additionally, this study highlights the importance of adaptive budgeting strategies and resource allocation aligned with demographic changes. Implementing a data-driven approach to adjust educational budgets and address regional demands will be crucial for sustaining quality education in the future. Though limitations exist due to potential demographic shifts and the inherent uncertainties in forecasting, the framework established here offers policymakers a proactive tool for financial planning. Through continued data refinement and potential model enhancements, this approach can evolve to further support efficient educational resource management, aligning fiscal strategies with evolving demographic needs.

# Appendices

Εικόνα που περιέχει κείμενο, γραμμή, διάγραμμα, γράφημα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γράφημα, γραμμή, στιγμιότυπο οθόνης

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, στιγμιότυπο οθόνης, παράλληλα

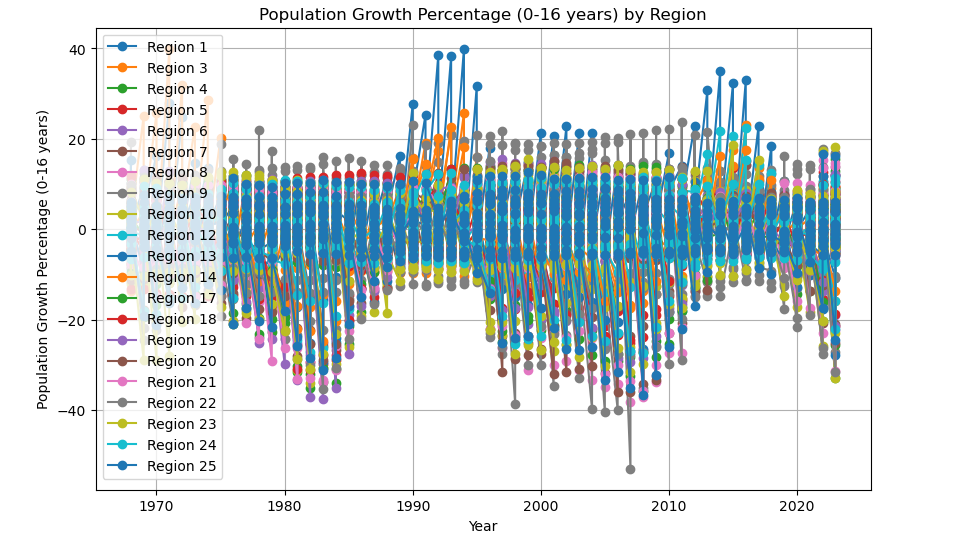
Περιγραφή που δημιουργήθηκε αυτόματα

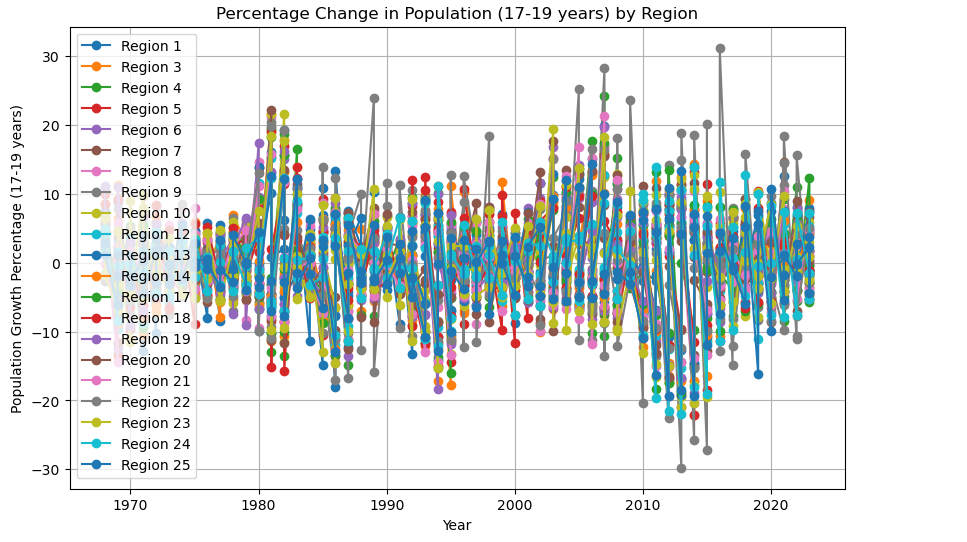
Εικόνα που περιέχει κείμενο, διάγραμμα, γραμμή, γράφημα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα





Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμμή, γράφημα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, διάγραμμα, γράφημα, στιγμιότυπο οθόνης

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, διάγραμμα, γράφημα, γραμμή

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, διάγραμμα, γραμμή

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γράφημα, γραμμή

Περιγραφή που δημιουργήθηκε αυτόματαΕικόνα που περιέχει κείμενο, διάγραμμα, γράφημα, γραμμή

Περιγραφή που δημιουργήθηκε αυτόματα  
Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, λογισμικό, γραμματοσειρά

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, λογισμικό

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, λογισμικό

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, ορθογώνιο παραλληλόγραμμο

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, οθόνη, λογισμικό

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, διάγραμμα, γραμμή, γράφημα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, λογισμικό

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, διάγραμμα, γράφημα, γραμμή

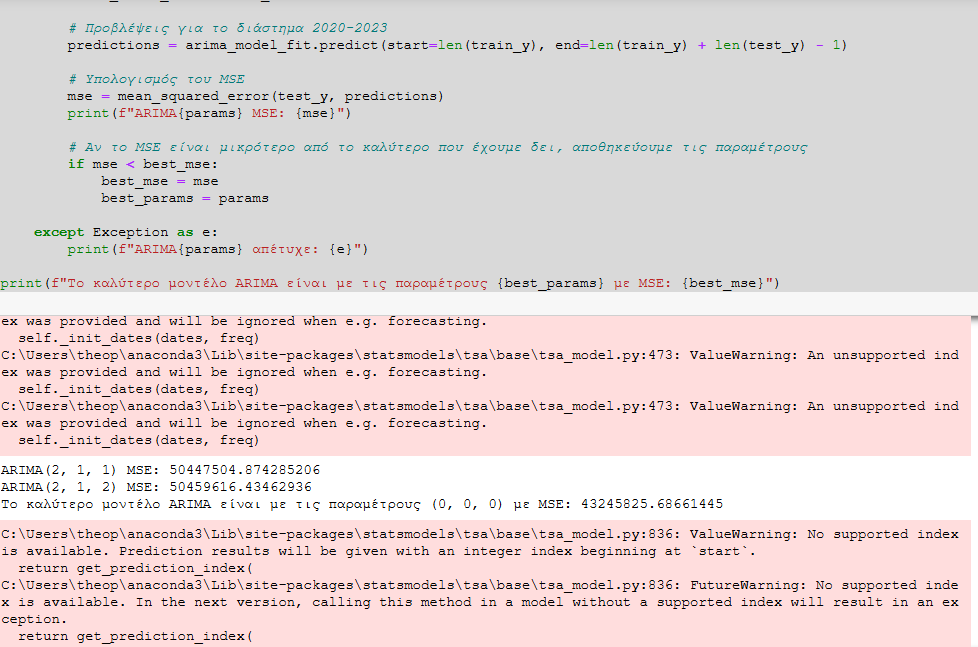
Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, λογισμικό, ιστοσελίδα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης

Περιγραφή που δημιουργήθηκε αυτόματα



Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, αριθμός, γραμματοσειρά

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, ηλεκτρονικές συσκευές, στιγμιότυπο οθόνης, ιστοσελίδα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, διάγραμμα, γράφημα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, ηλεκτρονικές συσκευές, στιγμιότυπο οθόνης, λογισμικό

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, αριθμός

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, αριθμός, γραμματοσειρά

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, λογισμικό, ιστοσελίδα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, λογισμικό, ιστοσελίδα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, λογισμικό, ιστοσελίδα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, λογισμικό, ιστοσελίδα

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, λογισμικό

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, λογισμικό

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, αριθμός

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, παράλληλα, γραμμή

Περιγραφή που δημιουργήθηκε αυτόματα

Total cost per year (grundskola&gymnasieskola)

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, αριθμός, γραμματοσειρά

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, διάγραμμα, γράφημα, γραμμή

Περιγραφή που δημιουργήθηκε αυτόματα

Average cost per child (Grundskola&Gymnasieskola) per year

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμμή, γράφημα

Περιγραφή που δημιουργήθηκε αυτόματα

Comparison of birth predictions and educational cost per year

Εικόνα που περιέχει γραμμή, κείμενο, στιγμιότυπο οθόνης, παράλληλα

Περιγραφή που δημιουργήθηκε αυτόματα

Comparison of birth predictions and educational cost per year

Εικόνα που περιέχει κείμενο, διάγραμμα, γραμμή, γράφημα

Περιγραφή που δημιουργήθηκε αυτόματα

Comparison of birth predictions and educational cost per year for region 1

Εικόνα που περιέχει κείμενο, διάγραμμα, γραμμή, γράφημα

Περιγραφή που δημιουργήθηκε αυτόματα

Total comparison of birth predictions and educational cost per year

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμμή, γράφημα

Περιγραφή που δημιουργήθηκε αυτόματα

Prediction Comparison for ARIMA & LSTM

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, αριθμός, λογισμικό

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, αριθμός, λογισμικό

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, αριθμός, γραμματοσειρά

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, αριθμός, γραμματοσειρά

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμματοσειρά, αριθμός, γραμμή

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, στιγμιότυπο οθόνης

Περιγραφή που δημιουργήθηκε αυτόματα

Predictions for educational cost 2025 – 2035

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμμή, παράλληλα

Περιγραφή που δημιουργήθηκε αυτόματα

Grundskola cost per region (2025 – 2035)

Εικόνα που περιέχει κείμενο, γραμμή, στιγμιότυπο οθόνης, γράφημα

Περιγραφή που δημιουργήθηκε αυτόματα

Percentage increase in educational cost per year

Εικόνα που περιέχει κείμενο, γραμμή, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματαComparison between grundskola&gymnasieskola costs (2025 – 2035)

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γράφημα, διάγραμμα

Περιγραφή που δημιουργήθηκε αυτόματα

Relation between population & educational cost per region (2025 – 2035)

# References

1.Brownlee, J. (2019, August 21). A gentle introduction to SARIMA for time series forecasting in Python. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/sarima-for-time-series-forecasting-in-python/

2.Brownlee, J. (2020, August 28). How to develop LSTM models for time series forecasting. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/

3.Brownlee, J. (2020, August 28). How to develop multilayer perceptron models for time series forecasting. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/how-to-develop-multilayer-perceptron-models-for-time-series-forecasting/

4.Brownlee, J. (2020, November 1). Random forest for time series forecasting. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/random-forest-for-time-series-forecasting/

5.Brownlee, J. (2021, March 19). How to use XGBoost for time series forecasting. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/xgboost-for-time-series-forecasting/

6.Brownlee, J. (2022, August 7). Time series prediction with LSTM recurrent neural networks in Python with Keras. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/

7.Frost, J. (n.d.). How to choose between linear and nonlinear regression. Statistics by Jim. Retrieved from https://statisticsbyjim.com/regression/choose-linear-nonlinear-regression/

8.Machine Learning Plus. (n.d.). Augmented Dickey Fuller test (ADF test) – Must read guide. Retrieved from https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/

9.Rosebrock, A. (2021, May 24). Grid search hyperparameter tuning with scikit-learn and GridSearchCV. PyImageSearch. Retrieved from https://pyimagesearch.com/2021/05/24/grid-search-hyperparameter-tuning-with-scikit-learn-gridsearchcv/

10.Saadeddin, Z. (2024, September 9). ARIMA for time series forecasting: A complete guide. DataCamp. Retrieved from https://www.datacamp.com/tutorial/arima