University of Hertfordshire UH

School of Physics, Engineering and Computer Science

MSc Data Science Project
7PAM2002-0509-2024
Department of Physics, Astronomy and Mathematics

Data Science FINAL PROJECT REPORT

Project Title:

PREDICTING UK AVERAGE HOUSE PRICES USING ARIMA, SARIMA, AND LSTM MODELS

Student Name and SRN:

https://github.com/AdedayoDanielAkinseli/predictionusingARIMASARIMALSTM

ADEDAYO DANIEL AKINSELI 23100060

Supervisor: Dr. Robert M Yates

Date Submitted: 28th August, 2025

Word Count: 5543

DECLARATION STATEMENT

This report is submitted in partial fulfilment of the requirement for the degree of Master of Science **in Data Science** at the University of Hertfordshire.

I have read the detailed guidance to students on academic integrity, misconduct and plagiarism information at [Assessment Offences and Academic Misconduct] and understand the University process of dealing with suspected cases of academic misconduct and the possible penalties, which could include failing the project or course.

I certify that the work submitted is my own and that any material derived or quoted from published or unpublished work of other persons has been duly acknowledged. (Ref. UPR AS/C/6.1, section 7 and UPR AS/C/5, section 3.6)

I did not use human participants in my MSc Project.

I hereby give permission for the report to be made available on module websites provided the source is acknowledged.

Student Name printed: ADEDAYO DANIEL AKINSELI

Student Name signature 

Student SRN number: 23100060

UNIVERSITY OF HERTFORDSHIRE

SCHOOL OF PHYSICS, ENGINEERING AND COMPUTER SCIENCE

ABSTRACT

The main objective of this dissertation was to evaluate the predictive performance of ARIMA (Autoregressive Integrated Moving Average), SARIMA (Seasonal Autoregressive Integrated Moving Average), and LSTM (Long-Short Term Memory) models in forecasting the average house price in the United Kingdom using Government data. Using this univariate dataset, the models were trained on 80% data, and the remaining 20% was tested. Results show the ARIMA model performed poorly by failing to capture the non-linear and seasonal dynamics in the housing market, yielding high errors and a negative R-squared value. SARIMA model improved its accuracy by incorporating seasonality, yielding an R-squared value of 0.8574, though it struggles with long-term dependencies. LSTM outperformed both models, achieving the highest accuracy with R-squared values of 0.8827 and having lower error metrics. LSTM demonstrates its ability to capture non-linear and long-term dependencies. However, LSTM model limitations, which include its reliance on historical data, lack of external economic factors, high computational costs, and restricted interpretability, challenge its forecasting ability. This study highlights the potential of deep learning models like LSTM in offering a reliable framework for forecasting long-term UK housing prices and recommends integrating a multivariate model for future research.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 BACKGROUND

After the War, the post war(1945-1951) period saw a golden age of homeownership, the idea of owning a House was part of the national psyche, The Labour Government built more than a million homes, 80% was Council Housing, which was largely used to replace those bombed during the Cold War, The Housing Construction Surge continued when the Conservative return to power(1951-1955). It reached its highest peak in the 1970s, with more than 300,000 houses built. The average price then was around £10,000, equivalent to approximately £164,000 in 2025. In 1980, Conservative Prime Minister Margaret Thatcher championed the implementation of the Right to Buy Scheme. The policy granted Council tenants a minimum of three years' opportunity to buy houses at a generous discount. This policy aimed to create a property-owning democracy, fulfilling one of the key promises of the Conservative Party. For over 40 years, the policy has earned £47 billion for the Treasury, which is a decrease in revenues compared to the Treasury's annual budget of £137.7 billion. Following the introduction of this policy, the construction of Housing started to decline. UK social housing has been steadily decreasing. In 1973, council housing accounted for nearly 30% of the total number of houses in the United Kingdom. 50 years later, the percentages have dropped to over 6% as privately owned and rented homes make up the majority of the market. Why?, The transformation to the new landlord age began slowly with the 1986 part-privatisation of building societies into banks and the introduction of housing benefits, which were paid directly to landlords. It was shown by Dorling (2015, pp. 2) that "this transformation began under Mrs Thatcher's tenure of office". It added that "Private renters' tenancies were also made insecure by Acts of Parliament passed while she was Prime Minister, especially the 1988 Housing Act". Due to the slowdown in housing construction, rising population growth, and successive governments' reluctance to commit to large-scale housebuilding, **House prices** have increased at a significantly faster rate than wages. A declining proportion of money spent on housing, on rents and mortgages, is being used to maintain housing. The author explained how citizens with higher incomes were investing in buy-to-let mortgages in an effort to secure financial gains in the future. "Housing prices can rise beyond many people's ability to ever pay for them because of the anticipated future profits from renting them in perpetuity", according to Dorling (2015, pp. 2)

The 2008 financial crisis resulted from the crash of the US housing market, which was fueled by risky lending practices such as subprime mortgages to borrowers with bad credit. These faulty loans caused significant losses for worldwide banks, resulting in a severe credit constraint. Financial institutions either failed, received government bailouts, or suffered significant restructuring. The crisis triggered a global recession, which had a significant impact on economies around the world, including the United Kingdom. Dorling (2015, pp. 5) argues that "before the 2008 financial crash, the slow transformation of the British housing landscape had been so gradual that few noticed the change in the direction of travel. Mortgage holding had been falling since the mid-1990s, but this was initially thought to be due to demographic, not economic trends. In hindsight, it was the beginning of a trend that has been greatly accelerated by the outcome of the banking crash". In the United Kingdom, Miles & Monro (2021) report that average real house prices are now significantly three times higher than the same period in 1980, while accounting for inflation with the consumer price index.

Following years, the UK Housing Markets have undergone considerable volatility caused by economic cycles, interest rate fluctuations, Inflation, and political events such as BREXIT, Immigration, etc. Predicting Financial and economic-related time series data is one of the major priorities of the Government, Investors, Developers, and Individual buyers. Time series forecasting has long been used to predict housing patterns. (Sirisha, Belavagi, and Attigeri, 2022) noted that time series forecasting has become important in recent years. They explain further that it is widely employed in fields such as banking, industry, healthcare, and meteorology. They argue that evaluating financial data is critical for both online and offline businesses since it allows for a comprehensive understanding of sales performance, profits, and losses, and helps in the prediction of future values.

Traditional Statistical models such as AutoRegressive Integrated Moving Average (ARIMA) and Seasonal AutoRegressive Integrated Moving Average (SARIMA) are frequently utilized in this field. The ARIMA is made up of three components: The autoregression(**AR**), which predicts the current values based on the past values, think of Autoregressive as Sales of a Retail Store, if the Sales were high yesterday, they might be high today, because customers tend to shop around the same time this week. Integration or Differencing(**I**) use for making dataset stationary, because most dataset have trends and seasonality pattern, these patterns will be discussed further in the another section and Moving Average(**MA**) unlike Autoregression, it predict the current values using the past forecast errors, the core idea is that it looks at residual from past predictions and uses it to correct the current value. These models anticipate future prices based on previous values and lagged errors, assuming linear relationships. According to Szostek et al.(2024, pp. 1-2), "forecasting models operate in various ways, analyzing both single time series and complex multidimensional dependencies, such as atmospheric and seasonal conditions, to predict the future".

However, real-world time series, particularly in housing, frequently exhibit non-linear, non-stationary, and seasonal patterns. These traits call into question the performance of the traditional statistical models. Sezer, Gudelek, and Ozbayoglu (2020) report that Deep Learning has emerged as the most effective prediction method in the Machine Learning field, with a wide range of applications in recent years. Sirisha, Belavagi, and Attigeri (2022) discuss that Neural Networks are effective for time series analysis, as the Network can capture sequential dependencies inherent in the time series data.

Recent advances in deep learning, particularly Long Short-Term Memory (LSTM), provide alternatives that can learn and extract important information from extended sequences and complex patterns without requiring prior feature engineering. The LSTM model is designed to learn from these sequences over time by maintaining memory of previous data points, which helps to improve the accuracy of the predictions for data with time-based patterns. The LSTM has a specialised unit called Memory cells, which allows the model to retain information for longer periods. This architecture addresses the limitations of the Recurrent Neural Network(**RNN),** which struggles to maintain long-term dependencies due to a proclivity to forget previous information.

This Dissertation will focus on comparing the predictive performances of ARIMA, SARIMA, and LSTM in forecasting the average house price in the United Kingdom from 1990 to 2024. The Study examines this period from 1990 to ensure the use of reliable and consistent datasets, which coincide with substantial structural changes in the UK housing market, and improve the relevance and accuracy of forecasting models.

1.2 RESEARCH QUESTIONS

1. To what extent can the average house price in a region be predicted over time using ARIMA and SARIMA models?

2. Does using Long Short-Term Memory (LSTM) networks improve the prediction accuracy of housing values when compared to traditional time series models?

The goal of the project is to show which model provides the most reliable forecast for informed housing-related economic planning.

2. LITERATURE REVIEW

The section provides a Comprehensive review of the existing literature on House Prices in the United Kingdom and the application of ARIMA, SARIMA, and LSTM models in time series forecasting. Barrell, Kirby, and Whitworth (2011, pp. 1) note that “the housing market plays a fundamental role in the economy, and its functioning affects both consumer welfare and economic stability”. The chart below, according to the same authors(Barrell, Kirby, and Whitworth 2011, pp. 1), illustrates that “historically, spikes in the growth rates of real and nominal house prices have been followed by periods of readjustment, often necessitating declines in real house prices for several years”.

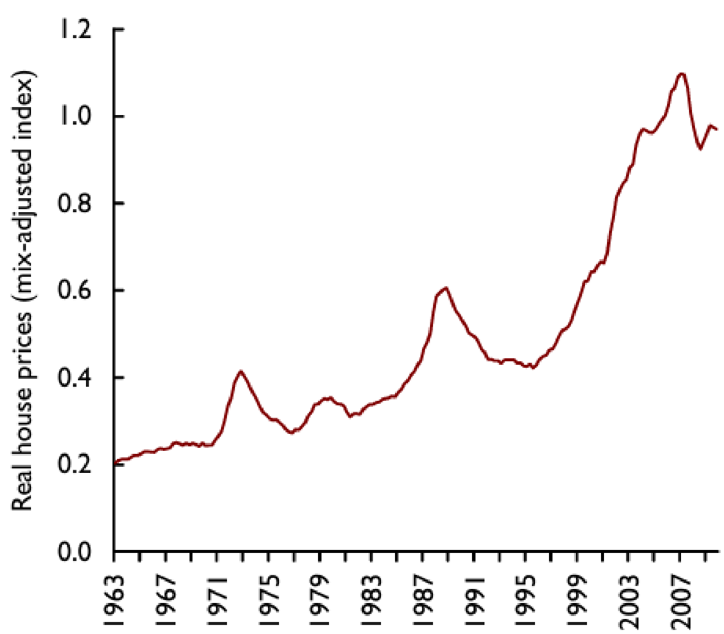

FIGURE 1: *Real house prices in the UK Adapted from Barrell, Kirby & Whitworth (2011, pp. 1)*.

Miles and Monro (2021, pp. 3) state that, “House prices in most developed economies, have risen greatly over the past few decades”, and Miles and Monro (2021, pp. 1) add that “the price rises have been greatest in the United Kingdom, where real house prices have risen more than three and a half times since the 1970s, substantially outpacing real income growth”.

According to Vishwas and Patel (2020, pp. 1-2), “Mathematical and statistical analysis performed on data to find hidden patterns and meaningful insight is called Time series analysis”. They continued that “a time series contains data points that increase, decrease, or otherwise change in chronological order over a period. A time series that incorporates the records of a single feature or variable is called a univariate time series. If the records incorporate more than one feature or variable, the series is called a multivariate time series. In addition, a time series can be designated in two ways: continuous or discrete (Vishwas and Patel, 2020, pp. 1-2). One of the main applications of time series analysis is time series forecasting, according to Joseph (2022). Sirisha, Belavagi, and Attigeri (2022, pp. 1-2) state that “for this effective analysis, the statistical methods Autoregressive Integrated Moving Average (**ARIMA**) and Seasonal ARIMA models (**SARIMA**), and the deep learning method- Long Short-Term Memory (**LSTM**) Neural Network model in time series forecasting have been chosen”.

They explained how the dataset must be stationary, especially for **ARIMA,** but not **for SARIMA** and **LSTM**. They used the ARIMA, SARIMA, and LSTM models to predict profit sales of the Dataset, which contains one million sales records, recorded from over 46 years ago. Fattah et al. (2018, pp.1) applied the ARIMA model in Forecasting demand, and they review that Demand forecasting is critical for inventory management, as stock inventory levels depend completely on demand forecasting. They demonstrate that historical demand data can be used to forecast future demand and how these forecasts influence the supply chain. Demand forecasting is the process of predicting future demand for a product or service by analyzing past data, market trends, and other relevant factors.

Deretić et al. (2022, pp. 6) discussed the SARIMA Modelling Approach for Forecasting of Traffic Accidents. They explained that the SARIMA model is one of the most effective linear models for seasonal time series forecasting.

Popesu (2020, pp.1) reports two case studies, in which a multiplicative SARIMA model and an intervention model were applied to time series data: one representing the number of fatal traffic accidents in the United States, and the other capturing road traffic accidents resulting in death or serious injury in the United Kingdom, before and after the Arab embargo was imposed in November 1973.

Song et al. (2024) wrote a review on the development of deep learning-based forecasting models between 2014 to 2024. They provide evaluations of how deep learning models perform in time series forecasting in both univariate and multivariate cases across multiple applications. The authors evaluate several methods to improve the effectiveness of long-term trend time series forecasting and examine the various loss functions used in these models."They methodically analyze these methods' performance in univariate and multivariate time series forecasting tasks in a variety of fields. They discuss the strengths and limitations of various algorithms from multiple perspectives, examine their ability to capture various types of time series information, such as trend and season patterns, and compare methods for improving the computational efficiency of these models"(Song et al., 2024, pp. 1).

Atef et al. (2021, pp. 1) stated that "a time series forecasting framework that uses deep learning to predict the environmental attributes that affect olive fruit farming. In this framework, it consists of a data preprocessing phase and a prediction phase. In the prediction phase, both the Long-Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) deep learning approaches were used to develop two models for predicting the environmental attributes. This framework's performance was evaluated using real-life agriculture data collected for twenty years from a Spanish olive grove".

Sarkar and De Bruyn (2021) demonstrate that long-short-term memory (LSTM) neural networks, which use raw data as input, can predict customer behaviors with high accuracy. Sarkar and De Bruyn (2022, pp. 1-2) stated that "LSTM neural networks are excellent candidates for modeling customer behavior using panel data in complex environments (e.g., direct marketing, brand choices, clickstream data, churn prediction)". They elaborate on how the Recurrent Neural Network (RNN) struggles in capturing long-term dependencies.

3. METHODOLOGY

3.1 DATASET SOURCE

The dataset was downloaded from the United Kingdom government website, whose main source is the HM Land Registry for England and Wales, Registers of Scotland, and His Majesty's Revenue and Customs Stamp Duty Land Tax data for the Northern Ireland House Price Index. The Dataset has been assessed against the Quality Assurance of Administrative Data (**QAAD**) toolkit.
The dataset consists of over 100,000 entries of data ranging over a period of 59years(1965 - 2024). It includes multiple columns, which are :

1. Datetime
2. Region Name, e.g, United Kingdom, England, Scotland, Wales, and Northern Ireland, Belfast, Mid Ulster, Derry City and Strabane, East Midlands, etc.
3. Area code
4. Average price
5. Price Index
6. Property Type – detached, semi-detached, terraced, flat, or all property types
7. Buyer status – first-time buyers or former owner-occupiers
8. Funding status – cash and mortgage
9. Property status – new builds and existing builds

3.1.1 DATA PREPROCESSING

For the preparation of time series data, the required libraries were imported into the notebook. The Dataset was loaded using the CSV from Pandas.

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
import scipy.stats as ss
from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score
```

```
from statsmodels.tsa.arima.model import ARIMA
```

```
from statsmodels.tsa.statespace.sarimax import SARIMAX
```

```
from sklearn.preprocessing import MinMaxScaler, StandardScaler
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM, Dense, Dropout, InputLayer
from tensorflow.keras.callbacks import ModelCheckpoint, EarlyStopping
from tensorflow.keras.losses import MeanSquaredError
from tensorflow.keras.metrics import RootMeanSquaredError
from sklearn.metrics import mean_squared_error
from tensorflow.keras.optimizers import Adam
```

Figure 2: *Imported libraries from the Notebook*

The date column was converted into a datetime object by using the parse_dates. It was then converted into a Dataframe index, and also ensured the date conforms to UK-style dayfirst format. Missing values were handled, and unimportant columns were dropped. Several duplicates were present in the dataset, but were not removed because the duplicates only indicate that there are some months in which the Average Price and Index price did not change. Data Filtering was an important part of the Preprocessing Stage. Four regions, namely England, Wales, Scotland, and Northern Ireland, were extracted for further visualization. The main region United Kingdom, which will be the focus point later in the search, was also extracted. This was univariate time series data. The Region Name column was dropped, leaving just one observation(Column) in each extracted data.

3.2 DATA VISUALIZATION AND INTERPRETATION

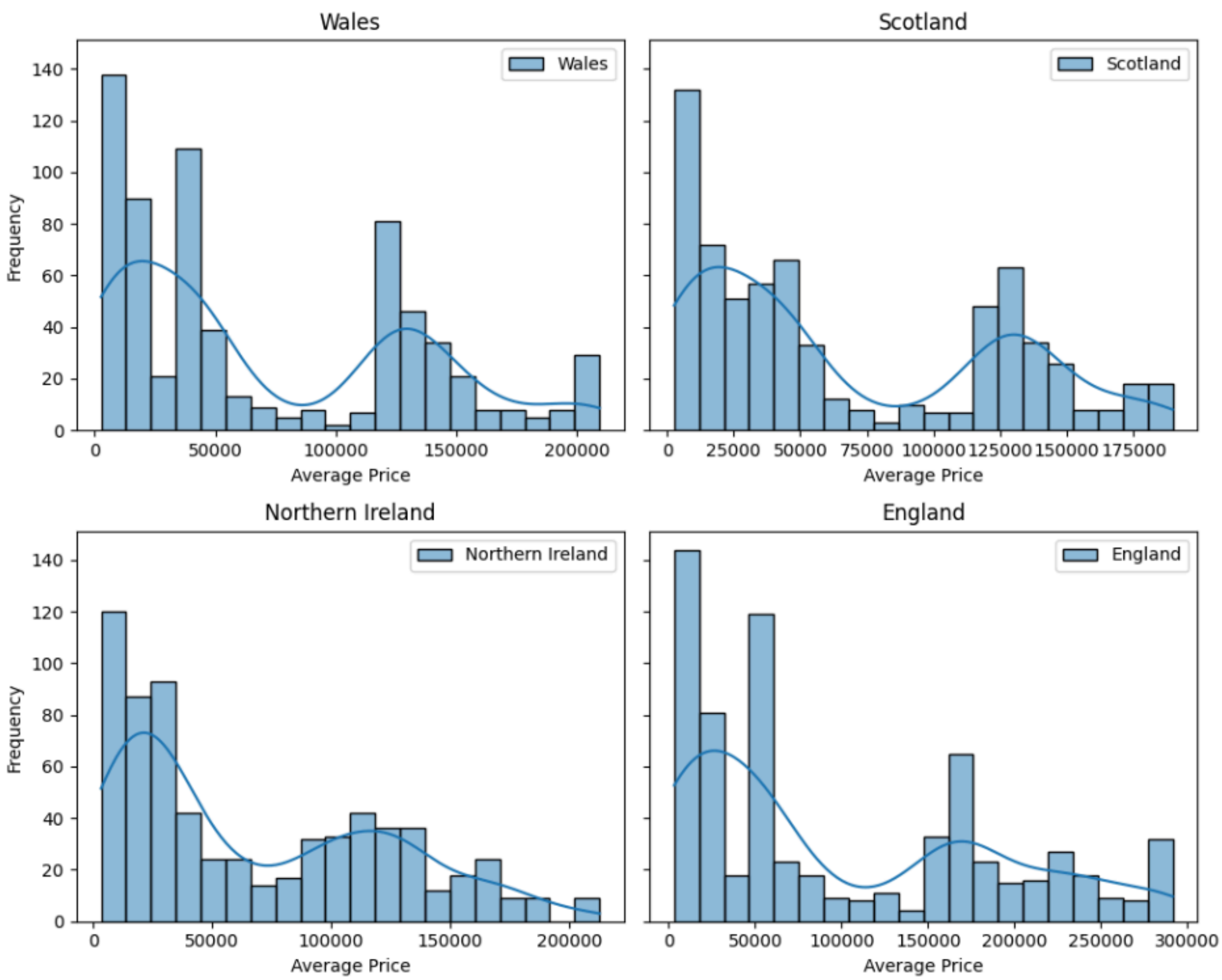

Figure 3: *Histogram showing the distributions of average house prices in four regions*

Figure 3 shows frequency distributions of average house prices in Wales, Scotland, Northern Ireland, and England for the recorded period. The x-axis displays the average price (GBP), while the y-axis depicts the frequency (number of monthly records) for each price range. Each diagram includes a Kernel Density Estimate (KDE) that depicts the smoothed probability distribution.

3.2.1 INTERPRETATION OF GRAPH

- WALES

  The distribution shape is a bimodal distribution with two distinct peaks. The first peak is around £0-£50,000, reflecting prices from previous decades (1970s-1990s). The second peak is around £125,000-£150,000, which corresponds to recent years. Prices ranged from around £20,000 in the early period to more than £200,000 in subsequent years. This change in mode over time implies significant long-term inflation in house prices.

- SCOTLAND

  The distribution shape is a bimodal distribution; however, the first mode is more significant than the second. The first peak is between £0 and £50,000, showing that lower past prices dominate the data. The second peak is over £125,000, attributed to recent market conditions. Maximum values reach over £180,000. Scotland's distribution has a less severe right tail than England's, reflecting its lower overall price ceiling and relatively stable growth trend.

- NORTHERN IRELAND

  The distribution is Bimodal, with a greater spread in the mid-range values. The first peak was between £0 and £50,000, which represented market values before 2000. The second peak is about £100,000-£125,000, indicating the post-crash recovery period. A wider right-hand tail as a result of the 2005-2007 price boom, which exceeded £200,000, followed by a dramatic post-2008 decline. Northern Ireland's price variability over time is greater than that of other locations, consistent with its boom-bust cycle.

- ENGLAND

  The distribution is a Bimodal Distribution. The first peak is around £0–£50,000, which represents historical pricing. The second peak is about £200,000, reflecting recent high prices. The longest right tail, which may reach about £300,000. In recent decades, England has maintained consistently higher price levels and larger dispersion.

3.2.2 GENERAL OBSERVATION OF THE HISTOGRAMS

All four regions show bimodal distributions; early prices form the first mode, while more recent, higher prices form the second mode. The spread and tail length are the smallest in Scotland and the longest in England. Northern Ireland has unique mid-range clustering as a result of its market collapse and delayed recovery following 2008.

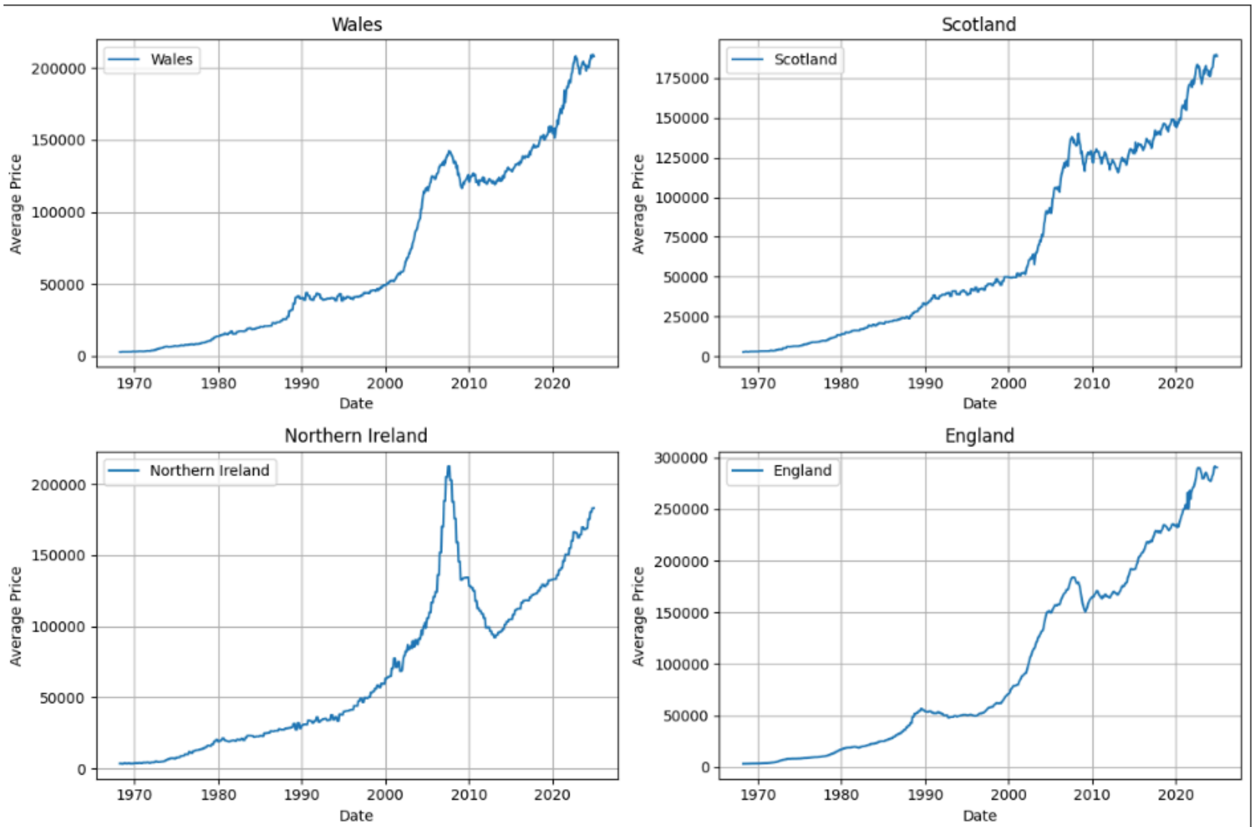

Figure 4: *Scatterplots showing average housing prices in the four regions*

The graph above displays four time series plots of average housing prices in Wales, Scotland, Northern Ireland, and England. The x-axis indicates the date (roughly from 1965 to 2024), and the y-axis shows the average price in GBP.

- WALES

  The graph shows a trend that gradually increases from the late 1960s until roughly 2000, followed by a more rapid climb till 2008. The Impact of the 2008 Financial Crisis saw a significant slowdown in growth and a short-term drop between 2008 and 2009.
  Post-crisis recovery shows a consistent growth from 2013 onwards, accelerating sharply after 2020, reaching £200,000 by 2024.

- SCOTLAND
  The graph shows a comparable ascending long-term trajectory trend. Growth was moderate from 1990 until 2003, after which it accelerated to reach its peak in 2008.
  After the 2008 Financial Crisis, prices remained stable. After 2013, there was a significant increase in costs, which by 2024 reached £180,000.

- NORTHERN IRELAND

  The trend is different from other regions due to high volatility. It shows an exceptional spike from 2005 to 2007, with prices tripling in a short period, peaking at roughly over £200,000. Significant drop after 2008 due to the 2008 Financial Crisis, Slower than other UK nations, but approaching £170,000 by 2024.

- ENGLAND

  The graph shows that England has the highest prices among regions. The trend gradually rises until the middle of the 1990s, then sharp gains until 2008. The 2008 Financial Crisis caused a noticeable decline that was followed by a robust rebound. After 2020, there was a significant growth that would reach £300,000 by 2024.

3.2.3 GENERAL OBSERVATION OF THE SCATTERPLOTS

The Global Financial Crisis of 2008 was reflected in all countries. The post-2020 trend acceleration was most likely affected by pandemic-related housing market dynamics, low interest rates, and supply restrictions. England has the highest average prices, followed by Wales and Scotland, while Northern Ireland is recuperating from a more severe post-crisis decline. Northern Ireland experiences the most extreme fluctuations in prices, whilst Scotland has the most consistent price growth trend.

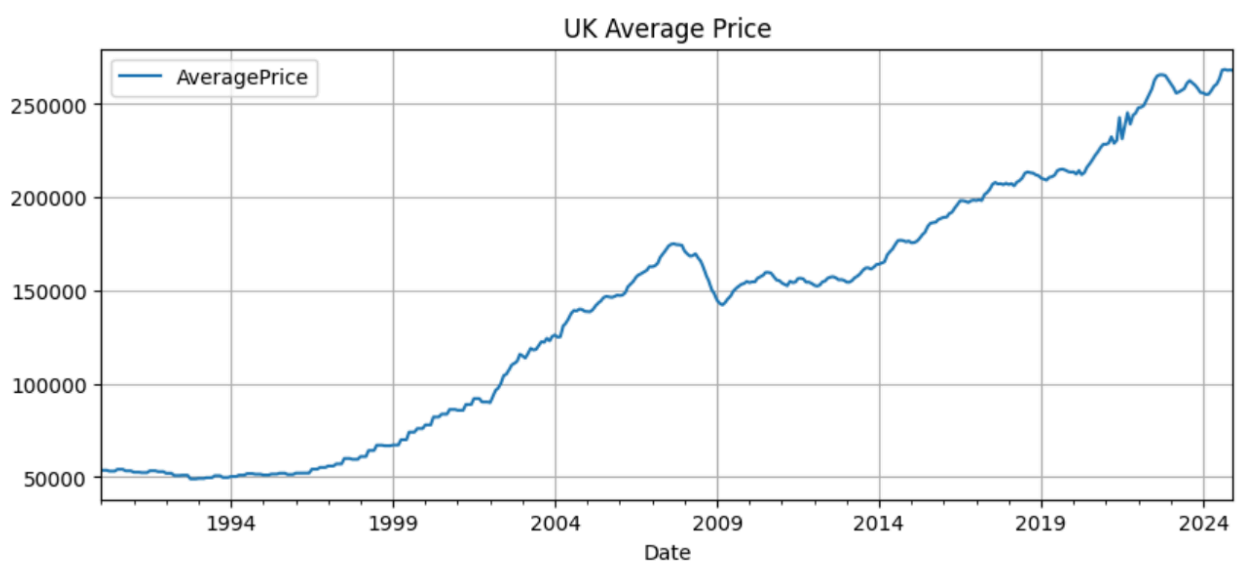

Figure 5: *UK average price from 1990 - 2024*

This graph depicts both long-term structural inflation in the UK housing market and short-term shocks resulting from macroeconomic events.

3.3 ARIMA MODEL DEVELOPMENT

The dataset undergoes a Time series decomposition to better understand the variety of patterns exhibited. There are three time series patterns, which are Trends, Seasonality, and Cycles.

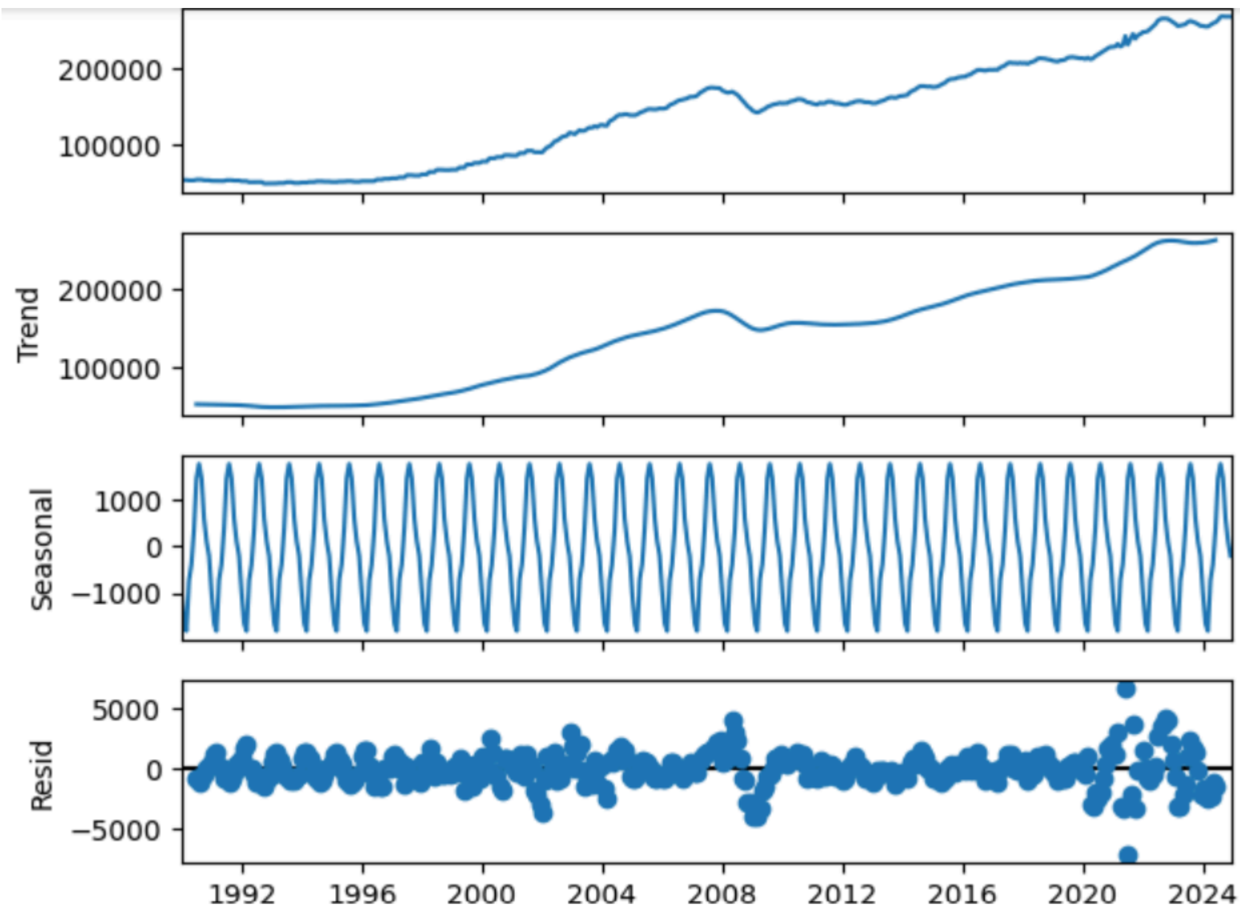

Figure 6: *Time Series Decomposition from the notebook*

The Stationarity of the dataset was evaluated using two methods: The Rolling Statistics and the Augmented Dickey-Fuller (ADF) Test. The rolling statistics use the rolling mean, which gives a visual check of trends, and the rolling standard deviation evaluates the changes in volatility. The ADF test provided more statistical confirmation. Both method demonstrates that the original series dataset was not stationary.

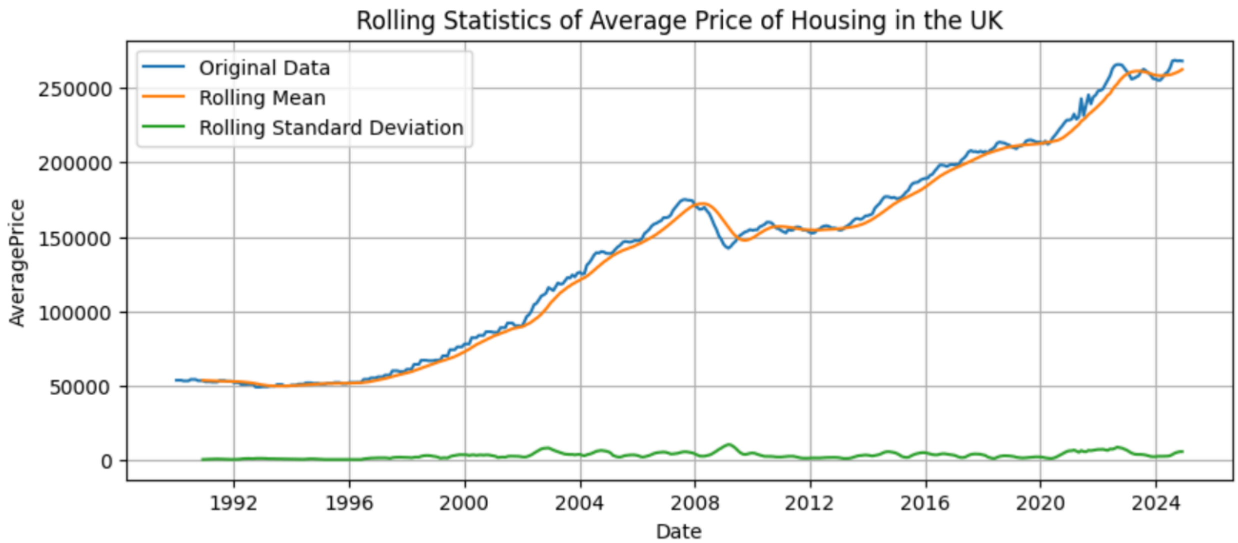

Figure 7: *Rolling Statistics of Average Price of Housing in the UK*

To address this, two transformations were used. First, the moving average method was used, which involved subtracting the rolling mean period 12 from the series and removing missing values. The ADF test of this altered series proved stationarity.

```
Results of Dickey-Fuller Test:
Test Statistic                  -2.991024
p-value                          0.035741
#Lags Used                      16.000000
Number of Observations Used    392.000000
Critical Value (1%)             -3.447142
Critical Value (5%)             -2.868941
Critical Value (10%)            -2.570713
```

Figure 8: *Result showing the Stationarity from the Moving Average method*

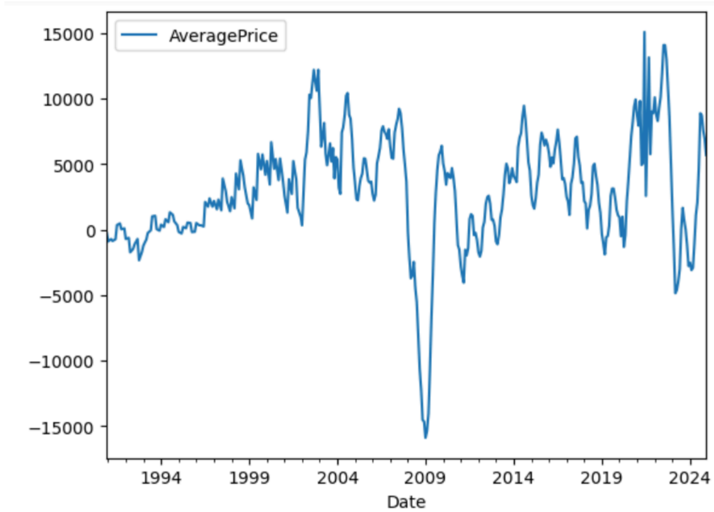

Figure 9: *Graph of the first Transformation using the Moving Average Method*

From the Graph, visually, it shows that there is still little trend exhibited in the graph, indicating the data have not fully undergone stationarity. First Order Differencing was introduced. This method calculates the differences between each value and the previous value. This is defined as:

$$Y'[t] = y[t] - y[t-1]$$

This method erased the trend while stabilizing the mean. After clearing the null values, the ADF test revealed that the differenced series was stationary.

```
Results of Dickey-Fuller Test:
Test Statistic                  -5.574449
p-value                          0.000001
#Lags Used                      15.000000
Number of Observations Used    403.000000
Critical Value (1%)             -3.446681
Critical Value (5%)             -2.868739
Critical Value (10%)            -2.570605
```

Figure 10: *Result showing the stationarity from first-order differencing*

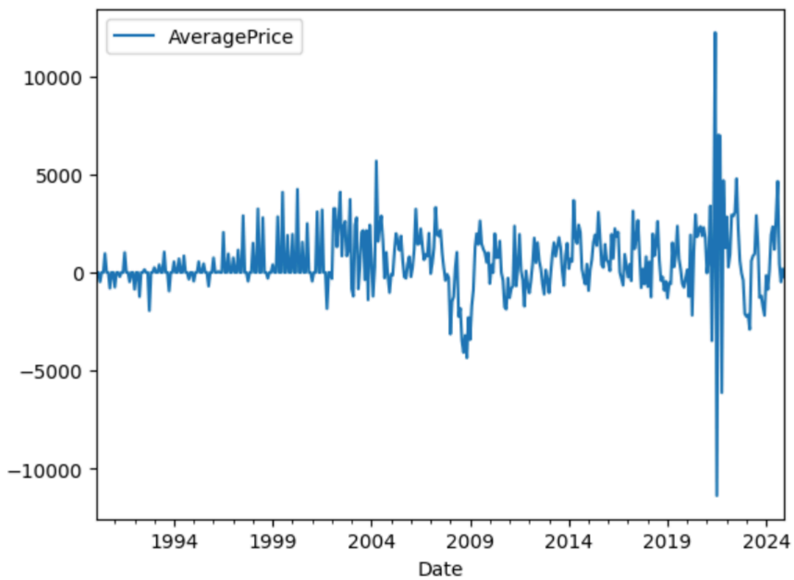

Figure 11: *Graph of the First Order Differencing*

3.3.1 PACF AND ACF GRAPH

The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) were implemented to examine the correlations between the series and its lagged values, which helped guide the selection of appropriate ARIMA parameters.

The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) were utilized to determine the ARIMA model parameters(**p, d,** and **q)**. p(Autoregressive), which is the number of lagged observations in the model, d is differencing, and q indicates Moving average. The ACF graph determines if the model is a Moving Average(q), while the PACF determines if the model is autoregressive.

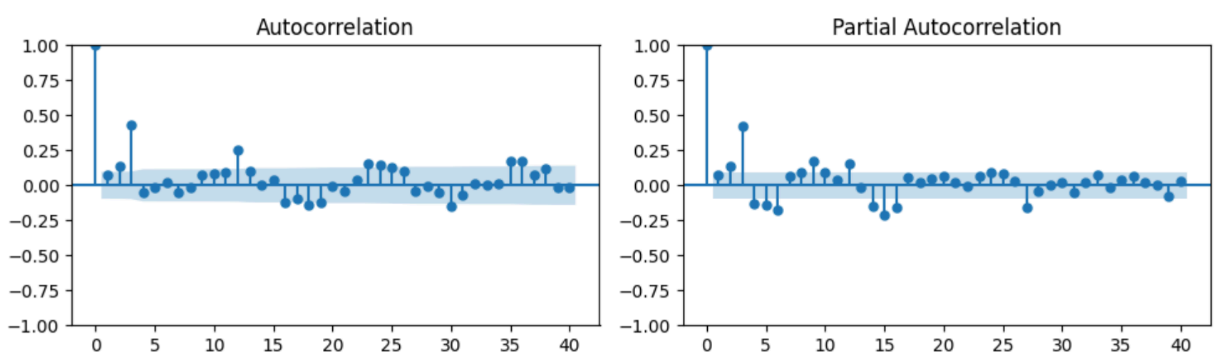

Figure 12: *Graph of PACF and ACF*

From the graphs above, the ACF graph shows a slow exponential decay while the PACF graph shows some significant spikes, most significant at lag 3, which suggests, based on the characteristics shown by both graphs it is an autoregressive model (AR). The suitable model parameters selected were (0,1,3). The parameter was evaluated together with other selected parameters such as (2,1,2),(3,1,2),(2,1,3),(1,1,1) using AIC. AIC (Akaike Information Criterion) is particularly useful for determining the ARIMA model. The model with the lowest AIC is usually selected as the suitable parameter.

The data was divided into training and testing subsets. A common strategy in time series forecasting is to utilize the majority of the data for training while keeping a smaller amount for model validation. In this study, 80% of the observations were used to train the model, while the remaining 20% served as unseen test data. This split allowed the model to learn from historical trends while also being evaluated for its ability to generalize to new data. The Arima model was fitted to the training data for forecasting. The result of the model was discussed in the Results Section

3.4 SARIMA MODEL

The Arima model produced unacceptable results during a Model Evaluation. The study moved on to Seasonal Autoregressive Integrated Moving Average (SARIMA). SARIMA, unlike ARIMA, accounts for the seasonal effect in time series analysis, which makes it suitable for this dataset. The model was defined as SARIMA(p, d, q)(P, D, Q)m, with m indicating a 12-month seasonal cycle. The non-seasonal parts are the (p, d, q) while the seasonal parts are (P, D, Q)m

The seasonal part of an AR or MA model will be seen in the seasonal lags of the PACF and ACF (Hyndman & Athanasopoulos, 2018). From the Graph of the ACF and PACF, it shows that ACF shows a slow decay at each lag m and PACF shows a spike at lag 12, which suggests that the appropriate model for (P, D, Q)m with be (0,1,1)12. The model was fitted to the already split train model to make a prediction. After Model Evaluation, the all dataset was refitted into the SARIMA model structure, then a forecast for the next two years was generated.

3.5 LONG SHORT-TERM MEMORY(LSTM)

LSTM is a specialised form of recurrent network that is designed to capture the long-term dependencies in sequential data. Unlike traditional ARIMA and SARIMA models, which are linear, LSTMs can model complex nonlinear patterns, making them appropriate for forecasting housing price dynamics featuring long-range temporal dependencies.

3.5.1 DATA PREPROCESSING

The dataset was normalised using a MinMaxScaler to rescale values into [0,1]. This step ensured that the LSTM's activation functions (ReLU and tanh) performed effectively, while also increasing convergence during training.

3.5.2 PREPARATION OF A WINDOW SEQUENCE

The dataset for the LSTM model was prepared using a sliding window technique. A customised function (date_time_series) was used to produce input-output sequences, with each input sequence consisting of a fixed number of previous observations and the output being the next value in the series. In this study, the input sequences consist of a 12-month window, and the model was trained to forecast the next month's property price based on the past twelve months' values.

The generated sequences were then divided into training and testing sets, with 80% for training and 20% for testing, while ensuring chronological order to eliminate look-ahead bias. Finally, the input data was transformed into a three-dimensional format (samples, timesteps, and features) as required by Keras' LSTM architecture. The number of features in this study was one, as the model was applied to a univariate series of average housing prices. This preprocessing step allowed the LSTM to accurately capture temporal dependencies in the data.

3.5.3 MODEL ARCHITECTURE

The LSTM model was built using the Keras Tensorflow backend. A layer was created on the model. Each layer plays a different role in the model.

1. The first layer: the brain, which helps to understand the pattern
2. The second layer: another brain, which gives the final idea of the output
3. The third layer: the dense layer turns the complex pattern into a decision
4. The fourth layer: the dropout layer, which prevents overfitting in the model
5. The final layer: a simple dense layer with one neuron, which makes the final predictions
6. The model compilation gives the model instructions on how to learn. Adam is an optimiser that helps models to adjust their learning.

3.5.4 HYPERPARAMETER TUNING

The initial LSTM architecture consisted of a single hidden layer and a simple neural network configuration. To increase performance, various hyperparameters were carefully tuned. Model complexity and overfitting were balanced by adjusting the number of hidden layers, experimenting with different configurations of units per layer (256, 128, and 64), and the dropout rate was refined gradually from 0.5 to 0.1. Iterative experimentation also helped to optimize the learning rate, batch size, and number of epochs(from 100 to 20). Early stopping was applied to terminate training once validation loss stopped improving, ensuring the model did not overfit to the training set. A validation split was applied manually to monitor performance on unseen data, and data shuffling was disabled to preserve temporal order(time series data). This adjusting approach improved the model's ability to capture temporal dependencies while preserving generalisation to previously unknown data.

3.5.5 FINAL MODEL ARCHITECTURE

```
# Build time series model
model = Sequential()

# First layer
model.add(LSTM(64, activation='relu', return_sequences=True, input_shape=((X_train.shape[1], X_train.shape[2]))))

# Second layer
model.add(LSTM(64, activation='relu'))

# Third layer - Dense layer
model.add(Dense(32, activation='relu'))

# Fourth layer - Dropout layer
model.add(Dropout(0.1))

# Final layer
model.add(Dense(1))
```

```
# compile the model
model.compile(loss=MeanSquaredError(), optimizer=Adam(learning_rate=0.001), metrics=[RootMeanSquaredError()])
```

Figure 13: *Showing the Model Architecture of LSTM from the Notebook*

The trained model was used to make predictions on the test data, and Model evaluation was done.

4. RESULTS AND ANALYSIS

4.1 RESULT AND ANALYSIS FOR ARIMA AND SARIMA MODEL

The table demonstrates how well the ARIMA and SARIMA models predicted UK average house prices. Four evaluation metrics were utilized: The Root Mean Squared Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared

1. Mean Squared Error (MSE): average of squared differences between predicted and actual values
2. The Root Mean Squared Error (RMSE): the square root of MSE, easily interpretable because it directly compares to the actual values
3. Mean Absolute Error (MAE): average of absolute differences between predicted and actual values
4. R-SQUARED: measures how well the model performed

MODEL	**RMSE**	**MSE**	**MAE**	**R-SQUARED**
ARIMA (0,1,3)	36389.33	1324183518.212	29021.94	-1.7253
SARIMA (0,1,1)12	8321.29	69243945	7031.68	0.8574 (86%)

Table 1: *Results showing the Model Evaluation of the ARIMA and SARIMA*

The ARIMA model generated significant error values and a negative R-squared, indicating a poor fit. This implies that the model is unable to effectively capture the patterns in the dataset and may not be appropriate for projecting house prices.

The SARIMA model performed significantly better than ARIMA. Lower RMSE, MSE, and MAE values indicate higher prediction accuracy. With an R-squared of 0.8574, the model explains nearly 86% of the variance in housing prices, indicating a strong fit. The incorporation of seasonal components enables SARIMA to capture monthly swings in housing prices precisely.

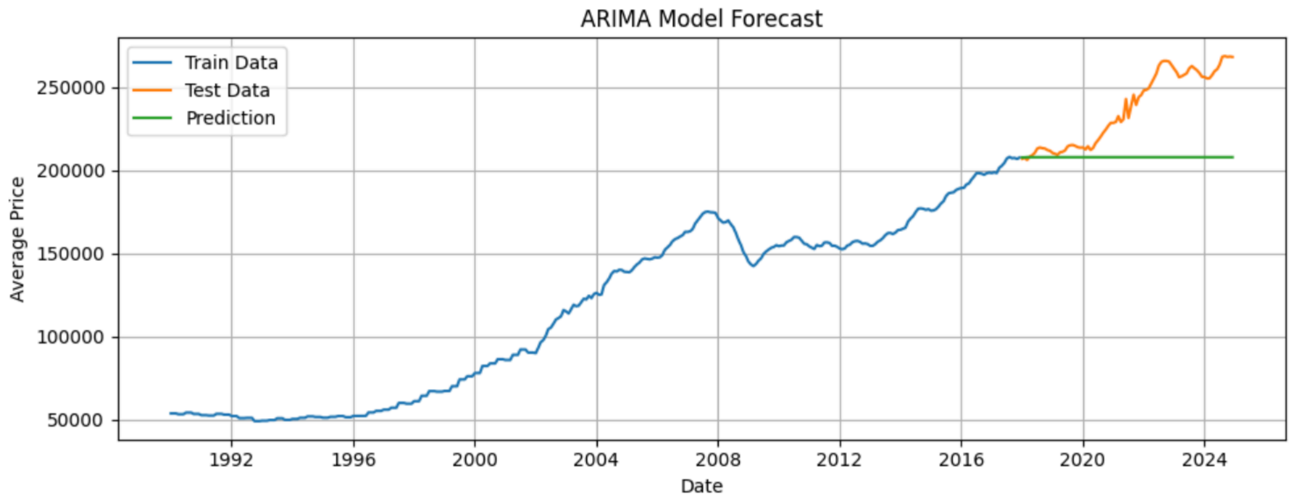

Figure 14: *Graph showing comparison of ARIMA model predictions and actual test data*

From Figure 14 graph, the ARIMA model fails to capture the pattern in the dataset. This shows that the model ignores the trends and seasonality which clearly present in the data, and predicts a flat forecast(green line).

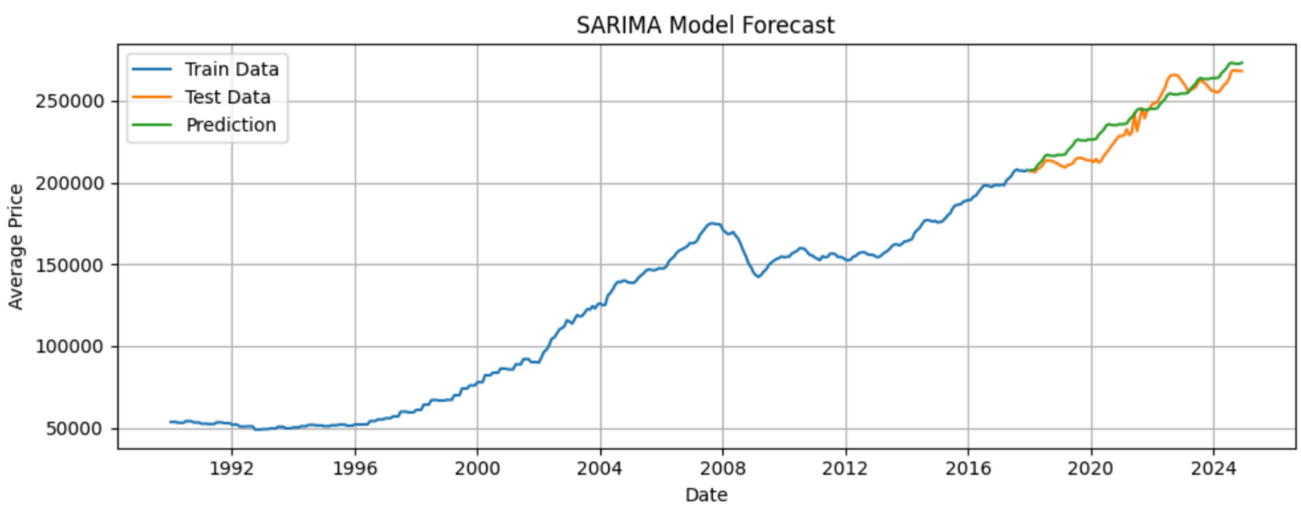

Figure 15: *Comparison of SARIMA model predictions and actual test data*

Based on the evaluation metrics, the model demonstrates satisfactory performance. However, visual analysis of the predictions reveals that, while the model typically follows the upward trend of the test data, it does not fully capture its specific patterns, indicating certain limitations in its predictive accuracy.

4.1.2 TWO-YEAR FORECASTING OF UK AVERAGE PRICE USING SARIMA

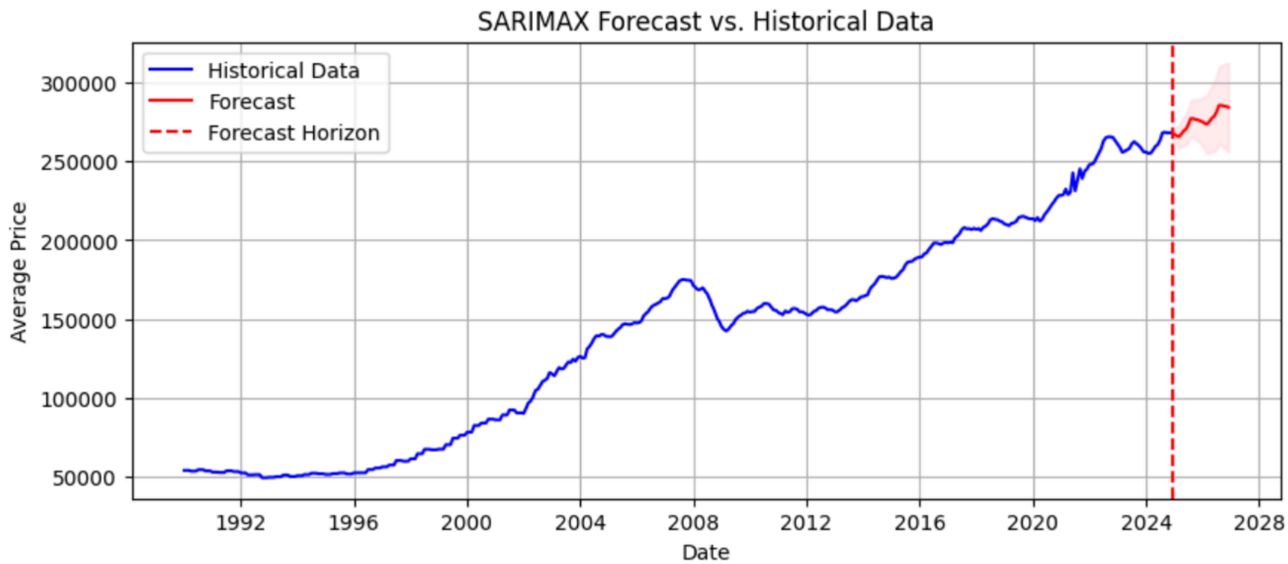

Figure 16: *SARIMA Model Forecast for Two Years*

Figure 16 depicts the SARIMA forecast vs historical UK property prices. The model accurately captures the overall upward trend and predicts continuous growth beyond the projected horizon. However, the projection is smoother than the actual record, showing a limited capacity to replicate short-term fluctuations. The widening confidence interval indicates increased uncertainty in longer-term projections.

4.2 RESULTS AND ANALYSIS OF LSTM MODEL

4.2.1 RESULT OF THE INITIAL PARAMETERS TESTING

MODEL TESTED CONFIGURATION	RMSE	R-SQUARED
Hiddenlayers(256,256) dropout(0.5) without early stopping	34939.35	-1.5654
Hiddenlayers(256,256) dropout(0.5) with early stopping	14701.42	0.5457
Hiddenlayers(128,128) dropout(0.2) with early stopping	6467.59	0.9140
Hidden layers(128, 64) dropout(0.2)with early stopping	5354.56	0.9583
Hidden layers(64,32) dropout(0.1)without early stopping	0.9135	6117.67

Table 2: *Showing results of the Initially Tested Configuration*

The evaluation of LSTM models for UK house price forecasting found that network architecture, dropout rate, and early stopping have a significant impact on performance. Large networks with substantial dropout performed poorly without early stopping. Smaller, moderately regularised networks with early stopping had the highest accuracy.

4.2.2 RESULT OF SELECTED PARAMETERS

MODEL	RMSE	MSE	R- SQUARED
LSTM layers(64, 64) dropout(0.1)with early stopping	7470.14	55803119.05	0.8827

Table 3: *Showing the result of the Selected Configuration*

The LSTM model with two layers of 64 neurons, a dropout rate of 0.1, and early stopping has an RMSE of 7,470.14, an MSE of 55803119.05, and R- SQUARED of 0.8827, indicating high prediction performance. The combination of the network size, low dropout, and early stopping is effective in balancing learning capacity with overfitting control, yielding a dependable LSTM model for forecasting UK average house prices.

4.3 TRAIN AND VALIDATION LOSS

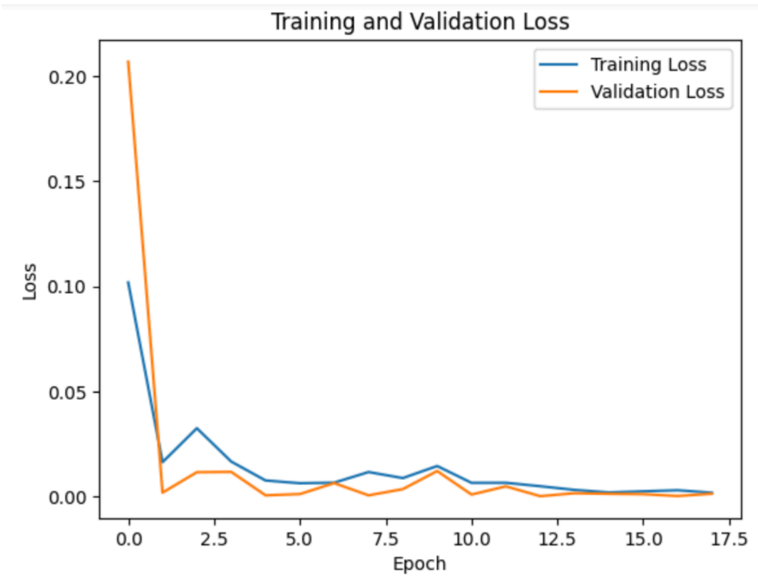

Figure 17: *Graph of the Train and Validation Loss*

The training loss decreases from its initial value of around 0.10 and stabilizes near zero. The validation loss follows a similar trajectory with slightly higher initial values. In general, the training and validation loss curves both exhibit a rapid decrease in the early epochs, followed by smooth convergence to zero. The close alignment of the curves during training demonstrates that the LSTM model successfully learns the temporal patterns in the data without overfitting.

4.4 GRAPH OF THE PREDICTED VALUES VS THE ACTUAL VALUES

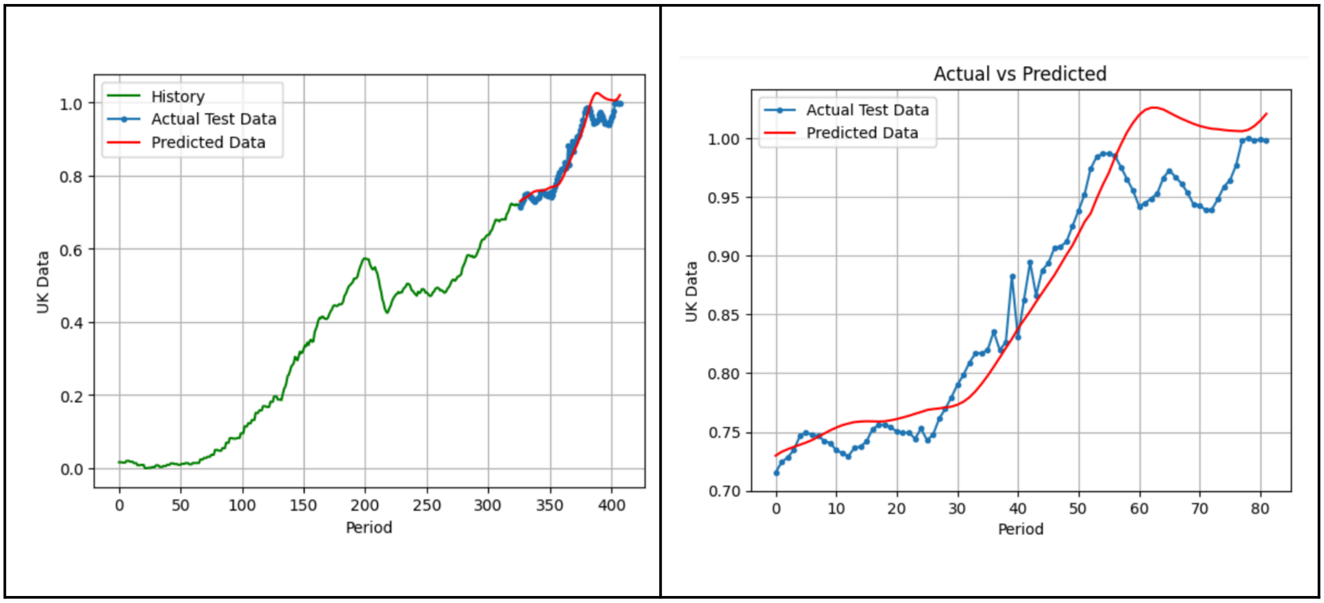

Figure 18: *Graphs showing comparison of the Predicted Data and Actual test data using the LSTM model, the left plot displays the entire data, comprising the historical data, test data, and the predicted values. The right plot concentrates on the test data and the predicted values*

Figure 18 shows that the predicted values (red line) correspond to the overall increasing trend of the actual test data (blue line), reflecting the long-term growth pattern of UK house prices.. The model smoothes out some short-term fluctuations, which is common in Neural Networks when using a one-step ahead forecasting. One-step-ahead forecasting uses past values from the data to make predictions. The model accurately predicts the flattening of housing values around the test period 50 - 80. This indicates that the LSTM correctly identified the declining growth momentum from the lagged input values and incorporated this knowledge into its forecasts. The close alignment between the two series demonstrates its ability to generalise underlying trends.

4.5 TWO-YEAR FORECASTING OF UK AVERAGE PRICE USING LSTM

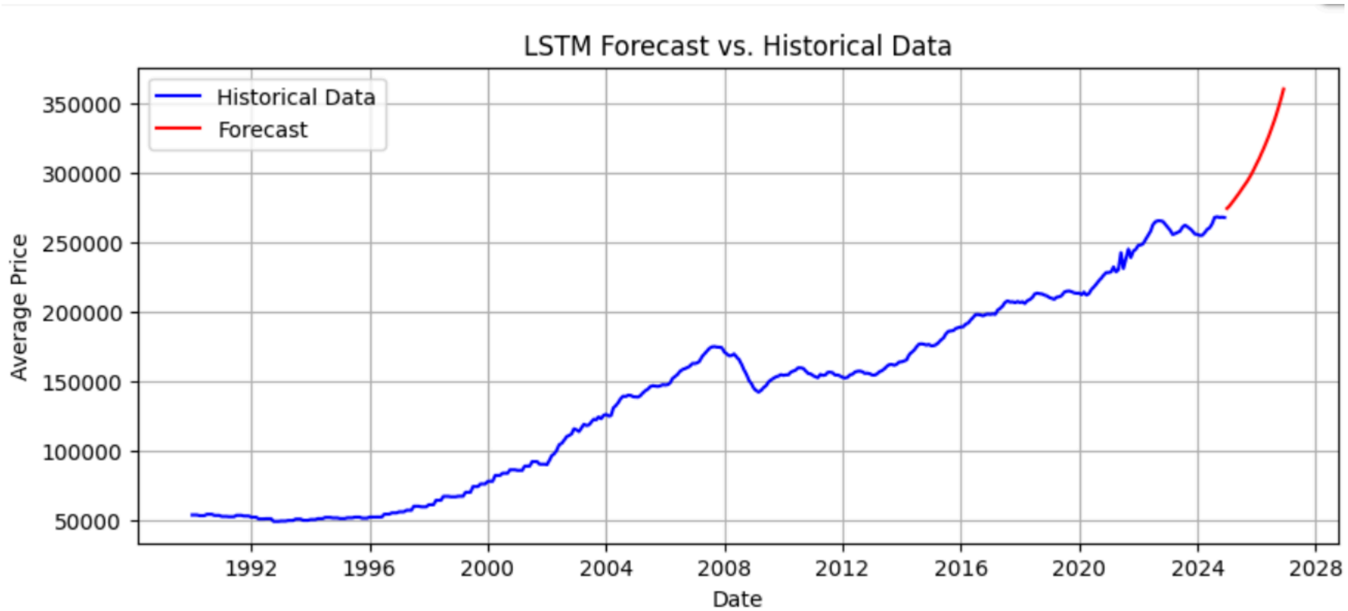

Figure 19: *Two-year forecast of UK house prices using the LSTM model*

Figure 19 shows that the forecast (represented by the red line) shows a continued upward trend, predicting a significant price increase beyond 2024. The LSTM model indicates that prices will keep rising and demand will remain strong in the coming years, as evidenced by the exponential growth pattern in the curve. The LSTM model applied recursive multi-step forecasting to make this prediction. The approach involves predicting the next step based on the last observed sequence. Each prediction is added to the input sequence and used iteratively to generate subsequent forecasts, allowing for multi-step predictions while capturing temporal dependencies. While the model effectively captures the long-term increasing trend, it can accrue error over longer periods. It does not account for potential external shocks, such as policy changes, interest rate fluctuations, and economic downturns.

5. DISCUSSION AND CONCLUSION

5.1 COMPARATIVE EVALUATION OF THE MODELS

ARIMA performed excellently at capturing short-term dynamics, but failed with nonlinear trends in housing data. Its reliance on linear assumptions reduced its predictive power for long-term forecasts. SARIMA improved upon ARIMA by incorporating seasonality. Although it captured cyclical swings more accurately, it fell short when long-term structural changes surfaced in the data. LSTM beat ARIMA and SARIMA in evaluation metrics (lower RMSE, MsE, and greater R^2). Its ability to understand non-linear and long-term dependencies allows it to accurately reproduce the growing trend of UK property prices.

MODEL	RMSE	MSE	R^2
ARIMA	36389.33	1324183518.21	-1.7253
SARIMA	8321.29	69243945.56	0.8574 (86%)
LSTM	7470.14	55803119.05	0.8827(88%)

Table 4: *Showing the comparison of the results model*

5.1.1 KEY LIMITATIONS OF TRADITIONAL MODELS(ARIMA AND SARIMA)

- Difficulty with non-stationary and non-linear series in the dataset
- Limited ability to detect abrupt structural changes, such as financial crisis, the COVID-19 crisis, in the data
- Requires manual parameter tuning, which may be subjective and a tedious process in implementing

5.1.2 KEY LIMITATIONS OF THE LSTM MODEL

The Long Short-Term Memory (LSTM) model accurately captures long-term trends and temporal relationships in historical housing price data. However, it shows limitations when confronted with external shocks such as policy interventions, interest rate increases, or general economic downturns.

The application of LSTM models to UK house price forecasting revealed several constraints. Their reliance on historical price data restricted their ability to predict shocks that are not present in the data. Developed as univariate models, LSTM forecasts are based on previous prices, excluding exogenous influences such as interest rates, inflation, and housing policy, which strongly shape affordability and demand. Training also required substantial data. For this dataset, the available data was not enough to achieve stable and reliable performance. The performance of LSTMs is heavily dependent on hyperparameter selection. As it was shown in the results section, hyperparameter optimisation was computationally intensive and time-consuming. Furthermore, the LSTM model's blackbox nature limits the interpretability of the trained data, unlike ARIMA and SARIMA, when the goal is not simply prediction, but also understanding the fundamental determinants of housing market behavior.

5.1.3 RESEARCH QUESTIONS

1. To what extent can the average house price in a region be predicted over time using ARIMA and SARIMA models

   The result shows SARIMA can predict the average house prices more effectively than ARIMA, due to the underlying data exhibiting a seasonality pattern. The model accuracy diminishes when confronted with long-term dependencies and nonlinear dynamics.

2. Does using Long Short-Term Memory (LSTM) networks improve the prediction accuracy of housing values when compared to traditional time series models?

   The Study shows that the LSTM model improves the prediction accuracy. Its ability to understand non-linear and long-term dependencies allows it to accurately reproduce the growing trend of UK property prices. However, the improvement comes with limitations.

6. CONCLUSION

This study aimed to assess the predictive effectiveness of classical statistical models (ARIMA and SARIMA) and a deep learning model (LSTM) in anticipating UK average house prices. The analysis indicated that, while ARIMA and SARIMA models are good at capturing linear patterns and seasonal components, their prediction accuracy is limited when the data contains strong nonlinear dynamics. In contrast, the LSTM model, with its capacity to capture complex temporal connections, provided predictions that more accurately reflected the long-term increasing trend.

However, the study also revealed that the LSTM's improved performance has significant limitations, which have been stated in the previous sections. The results from the LSTM should be interpreted and used with caution. The Forecast is exclusively based on historical price data, which does not factor in economic variables such as interest rates, inflation, and Housing policy, all of which have a substantial effect on the housing market. For the improvement of forecast accuracy, future studies should use multivariate models as well as external economic considerations. The adaptations of the hybrid model architecture approach, Babu and Reddy (2014) propose using the linear(ARIMA) and nonlinear model combination(Artificial Neural Network) for forecasting time series dataset, enhancing model interpretability(**LSTM black box**), using techniques such as SHAP(**SHapley Additive exPlanation**) and LIME(Local Interpretable Model AgnosticExplanation) to help understand the factors that influence predictions in a model, and help improve its output.

7. REFERENCES

Atef, M., Khattab, A., Agamy, E.A. and Khairy, M.M. (2021) ‘*Deep Learning Based Time-Series Forecasting Framework for Olive Precision Farming*’. IEEE, pp. 1062–1065

Babu, C.N. and Reddy, B.E. (2014) ‘A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data’, *Applied soft computing*, 23, pp. 27–38.

Barrell, R., Kirby, S., & Whitworth, R. (2011). REAL HOUSE PRICES IN THE UK. *National Institute Economic Review*, 216(1), F62-F68. https://doi.org/10.1177/0027950111411374 (Original work published 2011)

Barker, K. (2004). *Review of housing supply: delivering stability: securing our future housing needs: final report: recommendations*. London: HM Treasury.

Deretić, N., Stanimirović, D., Al Awadh, M., Vujanović, N. & Djukić, A. (2022). SARIMA Modelling Approach for Forecasting of Traffic Accidents. *Sustainability* [Preprint], 14(8), p.4403. https://doi.org/10.3390/su14084403

Dorling, D. (2015) ‘Policy, politics, health and housing in the UK’, *Policy and politics*, 43(2), pp. 163–180.

Fattah, J., Ezzine, L., Aman, Z., El Moussami, H. and Lachhab, A. (2018) ‘Forecasting of demand using ARIMA model’, *International journal of engineering business management*, 10, pp.1–9. doi:10.1177/1847979018808673

Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2

Joseph, M. (2022). *Modern time series forecasting with Python. Explore industry-ready time series forecasting using modern machine learning and deep learning*. Birmingham: Packt Publishing.

Miles, D. & Monro, V. (2021). UK house prices and three decades of decline in the risk-free real interest rate. *Economic Policy*, 36(108), pp.627–684. Available at: https://doi-org.ezproxy.herts.ac.uk/10.1093/epolic/eiab006

Popescu, T.D. (2020) ‘Time Series Analysis for Assessing and Forecasting of Road Traffic Accidents – Case Studies’, *WSEAS Transactions on Mathematics*, 19, pp. 294–302

Sarkar, M., & De Bruyn, A. (2022). LSTM Response Models for Direct Marketing Analytics: Replacing Feature Engineering with Deep Learning. *Journal of Interactive Marketing*, 53(1), 80-95. https://doi.org/10.1016/j.intmar.2020.07.002 (Original work published 2021)

Sezer, O.B., Gudelek, M.U. and Ozbayoglu, A.M. (2020) 'Financial time series forecasting with deep learning : A systematic literature review: 2005–2019', *Applied soft computing*, 90, p. 106181.

Sirisha, U.M., Belavagi, M.C. and Attigeri, G. (2022). Profit prediction using ARIMA, SARIMA and LSTM models in time series forecasting: A comparison. *IEEE Access*, 10, pp.124715-124727. https://doi.org/10.1109/ACCESS.2022.3224938

Song, X., Deng, L., Wang, H., Zhang, Y., He, Y. & Cao, W. (2024). Deep learning-based time series forecasting. *Artificial Intelligence Review*, 58, article no. 23. https://doi.org/10.1007/s10462-024-10989-8

Szostek, K., Mazur, D., Drałus, G. and Kusznier, J. (2024) 'Analysis of the Effectiveness of ARIMA, SARIMA, and SVR Models in Time Series Forecasting: A Case Study of Wind Farm Energy Production', *Energies* (Basel), 17(19), p. 4803.

Vishwas, B.V., Patel, A. (2020). Time-Series Characteristics. In: *Hands-on Time Series Analysis with Python*. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-5992-4_1

DATASET USED: Official UK Government Website (UK House Price Index: data downloads December 2024 - GOV.UK)