HSBC Share Price Forecasting Using ARIMA Model

Project Overview

This project leverages the ARIMA(3,0,1) statistical model to forecast the future share prices of HSBC. By using historical price data, the study aims to provide actionable insights into market trends, assisting investors and financial analysts in making informed decisions.

Data Source

Source: Yahoo Finance

Details: Historical share prices of HSBC are meticulously gathered from Yahoo Finance. This data forms the backbone of our forecasting model, offering a detailed look at price fluctuations over the specified period.

Tools and Technologies

- **IBM SPSS:** Primary tool for statistical modeling and analysis.
- **Microsoft Excel:** Used for initial data preparation and cleaning.

Methodology

Data Collection

Data is sourced from Yahoo Finance, focusing specifically on the historical share prices of HSBC. This dataset includes critical market data necessary for a robust analysis.

Data Preparation

The dataset undergoes rigorous preprocessing to address missing values and outliers, ensuring the integrity and accuracy of our model.

Model Selection

The ARIMA(3,0,1) model is selected based on its suitability to address the non-stationary nature of financial time series data. Preliminary tests, including the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF), help determine the appropriate parameters.

Model Implementation

The model is implemented in SPSS, where extensive testing is conducted to refine the parameters and validate the model's effectiveness through various diagnostic checks.

Key Findings

- **Model Performance:** The ARIMA(3,0,1) model achieves an R-squared value of 0.907, indicating a strong predictive capability.
- **Forecast Accuracy:** With a Root Mean Square Error (RMSE) of 27.163, the model demonstrates high accuracy in forecasting future share prices.
- **Market Insights:** The analysis provides deep insights into how economic variables influence HSBC's share price, aiding strategic investment decisions.

Results

The ARIMA model successfully forecasts future share price movements with high accuracy. Detailed forecasts include visual representations like graphs and tables, offering both quantitative and qualitative insights.

Conclusion

The ARIMA(3,0,1) model proves to be a powerful tool in predicting financial market trends, specifically for HSBC's share prices. This project not only underscores the viability of ARIMA models in financial forecasting but also provides a strategic framework for investors looking to navigate market volatilities effectively.

Appendix

Result of ARIMA

Model Description

			Model Type
Model ID	Price	Model_1	ARIMA(3,0,1)(0,0,0)

Model Fit

Fit Statistic	Mean	SE	Minimum	Maximum	Percentile 5	10	25	50	75	90	95
Stationary R-squared	.907	.	.907	.907	.907	.907	.907	.907	.907	.907	.907
R-squared	.907	.	.907	.907	.907	.907	.907	.907	.907	.907	.907
RMSE	27.163	.	27.163	27.163	27.163	27.163	27.163	27.163	27.163	27.163	27.163
MAPE	4.903	.	4.903	4.903	4.903	4.903	4.903	4.903	4.903	4.903	4.903
MaxAPE	15.130	.	15.130	15.130	15.130	15.130	15.130	15.130	15.130	15.130	15.130
MAE	20.820	.	20.820	20.820	20.820	20.820	20.820	20.820	20.820	20.820	20.820
MaxAE	82.355	.	82.355	82.355	82.355	82.355	82.355	82.355	82.355	82.355	82.355
Normalized BIC	6.803	.	6.803	6.803	6.803	6.803	6.803	6.803	6.803	6.803	6.803

Model Description

			Model Type
Model ID	Price	Model_1	ARIMA(3,0,1)(0,0,0)

Model Statistics

Model	Number of Predictors	Model Fit statistics: Stationary R-squared	Normalized BIC	Ljung-Box Q(18): Statistics	DF	Sig.	Number of Outliers
Price -Model_1	0	.907	6.803	15.493	14	.345	0

ARIMA Model Parameters

					Estimate	SE	t	Sig.
Price -Model_1	Price	No Transformation	Constant		446.930	50.326	8.881	.000
			AR	Lag 1	.081	.100	.809	.420

			Estimate	SE	t	Sig.
		Lag 2	.760	.082	9.249	.000
		Lag 3	.076	.101	.752	.454
	MA	Lag 1	-.964	.039	-24.436	.000

Residual ACF

Model		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Price -Model_1	ACF	.012	.108	.064	-.086	.184	.123	-.108	.025	-.082	-.037	.134	.054	.039	.014	-.022	.038	-.042	-.031	-.052	-.034	.031	-.132	-.042	.032
	SE	.091	.091	.092	.093	.093	.096	.098	.099	.099	.099	.099	.101	.101	.101	.101	.101	.101	.102	.102	.102	.102	.102	.103	.104

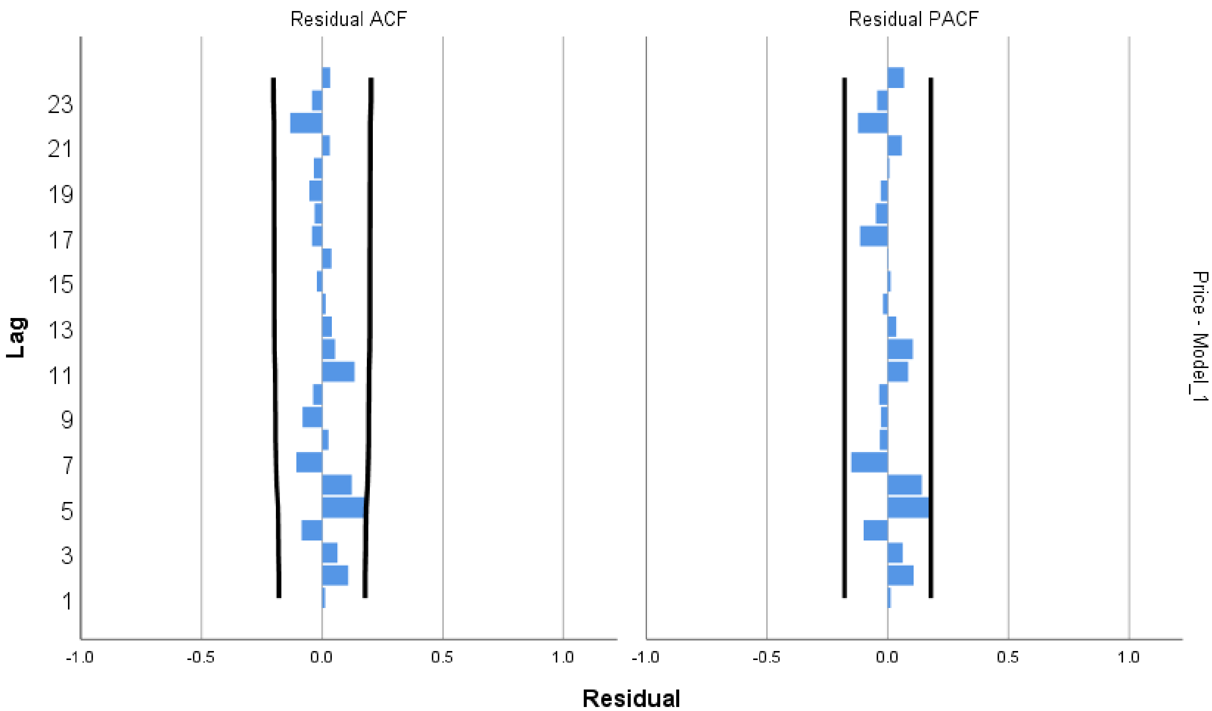
Residual PACF

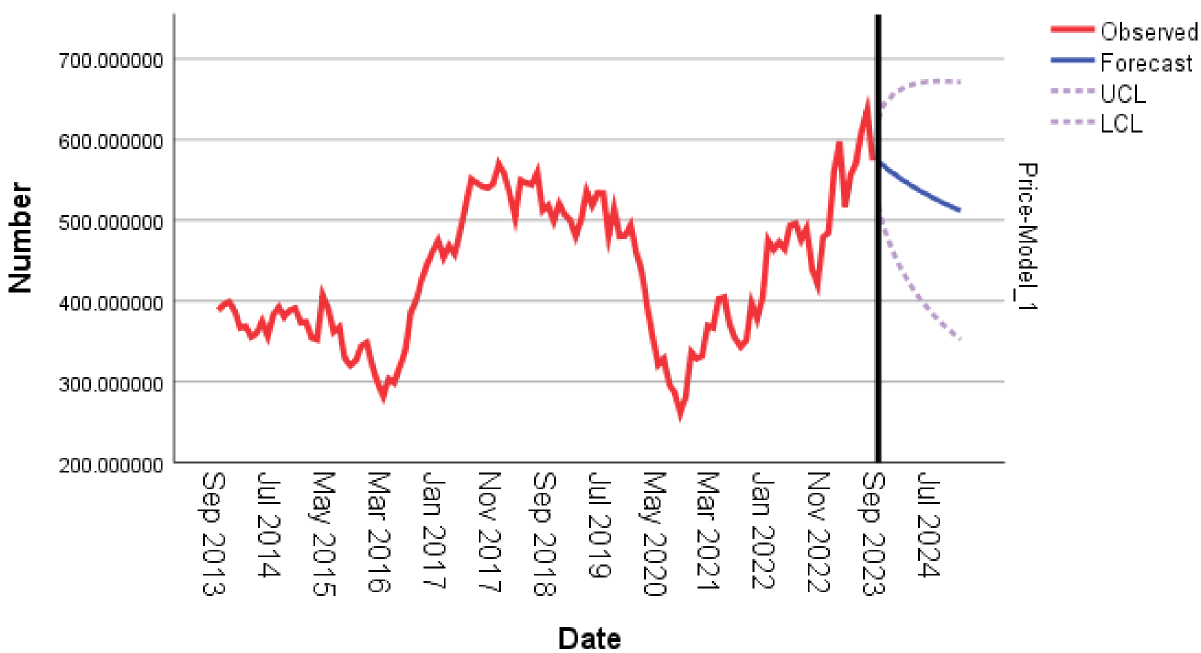
Model		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Price -Model_1	PACF	.012	.108	.063	-.100	.176	.141	-.151	-.032	-.027	-.035	.085	.104	.035	-.019	.012	.003	-.115	-.048	-.029	.007	.057	-.123	-.043	.069
	SE	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091	.091

Forecast

Model		Sep 2023	Oct 2023	Nov 2023	Dec 2023	Jan 2024	Feb 2024	Mar 2024	Apr 2024	May 2024	Jun 2024	Jul 2024	Aug 2024	Sep 2024	Oct 2024	Nov 2024	Dec 2024
Price -Model_1	Forecast	571.560826	567.763537	560.956294	557.343281	551.593298	547.870369	542.928689	539.267312	534.936391	531.432098	527.582513	524.282522	520.826970	517.750335	514.627594	511.777260
	UCL	624.288844	644.029839	649.257337	658.541139	661.145495	666.042202	667.428417	670.031979	670.657697	671.917128	672.020055	672.444238	672.176190	672.064887	671.533077	671.073942
	LCL	518.832809	491.497235	472.655252	456.145422	442.041101	429.698536	418.428961	408.502645	399.215086	390.947068	383.144972	376.120806	369.477751	363.435784	357.722111	352.480577

For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.





Model Description

Model Name		MOD_5
Series Name	1	Price

Transformation	None
Non-Seasonal Differencing	1
Seasonal Differencing	0
Length of Seasonal Period	No periodicity
Maximum Number of Lags	16
Process Assumed for Calculating the Standard Errors of the Autocorrelations	Independence(white noise)[a]
Display and Plot	All lags

Applying the model specifications from MOD_5

a. Not applicable for calculating the standard errors of the partial autocorrelations.

Case Processing Summary

		Price
Series Length		120
Number of Missing Values	User-Missing	0
	System-Missing	0
Number of Valid Values		120
Number of Values Lost Due to Differencing		1
Number of Computable First Lags After Differencing		118

Autocorrelations

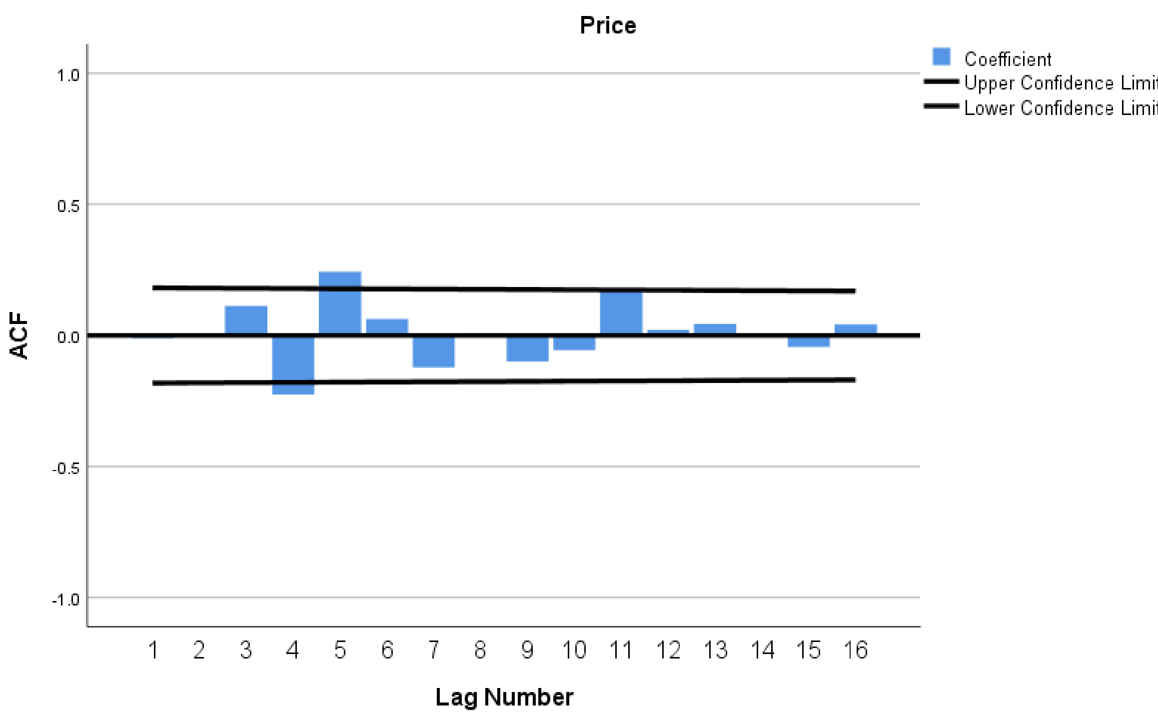
Series: Price

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	-.010	.091	.012	1	.913
2	.002	.090	.012	2	.994
3	.113	.090	1.609	3	.657
4	-.226	.089	7.993	4	.092
5	.244	.089	15.493	5	.008
6	.063	.089	16.000	6	.014
7	-.123	.088	17.942	7	.012
8	.008	.088	17.951	8	.022
9	-.100	.087	19.248	9	.023
10	-.056	.087	19.659	10	.033
11	.167	.087	23.372	11	.016
12	.022	.086	23.439	12	.024

13	.045	.086	23.714	13	.034
14	.001	.085	23.714	14	.050
15	-.044	.085	23.978	15	.065
16	.043	.085	24.231	16	.085

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.



Partial Autocorrelations

Series: Price

Lag	Partial Autocorrelation	Std. Error
1	-.010	.092
2	.002	.092
3	.113	.092
4	-.226	.092
5	.261	.092
6	.034	.092
7	-.084	.092
8	-.098	.092

9	.009	.092
10	-.079	.092
11	.131	.092
12	.052	.092
13	.074	.092
14	-.055	.092
15	.049	.092
16	-.039	.092