Factors affecting the Prices and Forecasting Model for Prediction of the Prices of Gold in India

Pritee Singh
Department of Statistics, Institute of Science, Nagpur.
priteesingh25@gmail.com

Abstract:

This study aims to examine the factors that influence gold prices in India and proposes a forecasting model for predicting future trends in gold pricing. Several variables that affect the price of gold in the Indian market have been identified. To analyze these factors, a multiple regression model was applied, along with the Coefficient of Determination (R^2), to understand the impact of various elements such as exchange rates, inflation rates, crude oil prices, and interest rates on gold prices. The dataset was first transformed to ensure stationarity, and lag values were examined using Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. To evaluate the accuracy of the forecasting model, various measures were used, including the Mean Absolute Percentage Error (MAPE), Bayesian Information Criterion (BIC), Mean Absolute Error (MAE), and the R^2 value. After conducting these tests, it was concluded that the ARIMA (Autoregressive Integrated Moving Average) model provided the most accurate forecasts.

Key Words: *Forecasting model, Prediction, Gold price, Coefficient of Determination, Autocorrelation plot, Partial autocorrelation plot*

1. Introduction:

Gold is widely regarded as a symbol of economic wealth across the globe, and its price plays a key role in determining the financial stability of a country. In India, gold is often used as a hedge against inflation, and its price fluctuates based on market conditions and economic demands. During periods of economic downturns, gold also serves as a safe-haven asset, providing security to investors when other assets might be underperforming.

1.1 Factors that affect the gold prices

Several key factors affect the price of gold, both in India and internationally:

a) Exchange Rate: When the exchange rate of the Indian Rupee (INR) against major currencies like the US Dollar (USD) fluctuates, it directly impacts the price of gold in India. If the value of the Rupee weakens relative to the USD, the price of gold tends to rise as it becomes more expensive in local currency terms.

b) Inflation: Gold is often seen as a store of value during inflationary periods. As inflation increases, the demand for gold typically rises as investors seek to protect their wealth from the erosion of purchasing power.

c) Crude Oil Prices: Gold and crude oil prices are often linked, as both are considered crucial commodities in the global economy. When crude oil prices rise, there is often an increase in the cost of goods and services, which may drive up gold prices as a hedge against rising living costs.

d) Interest Rates: Interest rates have an inverse relationship with gold prices. When

interest rates increase, the opportunity cost of holding gold (which does not earn interest) rises, leading to reduced demand for gold. Conversely, when interest rates decrease, gold becomes more attractive to investors.

e) BSE Sensex Index: The BSE Index, commonly referred to as the S&P BSE Sensex, serves as the primary benchmark index for the Bombay Stock Exchange (BSE) in India. There is a noted inverse correlation between the Sensex and gold prices. Typically, when investors anticipate an upward trend in the stock market, they are more inclined to allocate their funds to stocks, hoping to capitalize on rising prices in the future. This shift in investment strategy results in decreased demand for gold, leading to a decline in its prices. Thus, it can be concluded that an inverse relationship exists between the prices of gold and the Sensex.

1.2 Objective of Study

- To identify various factors affecting gold prices in the Indian Economy.
- To develop a forecasting model to predict price of gold in India.

2. Review of Literature

The chapter literature review is based on previously conducted studies related to gold price and various factors affecting it. Also it consists of conducted studies related to the forecasting of gold prices. Following are some studies regarding to it.

Dr. Sindhu (2013) examined a study on impact of selected factors on the price of Gold. Exchange rate of US dollar with INR, Crude oil prices, repo rate and inflation rate were taken into consideration. Each factor is studied with price of gold. The research findings indicated that there is an inverse relationship between the US$ and gold prices. The crude oil prices have an impact on the gold prices. Gold prices and repo rates are interdependent. Gold prices and inflation have positive association.

Siti Nurulhuda Ibrahim (2014) investigated the determinants of gold prices in Malaysia. Multiple Linear Regression Model was used to determine the significant relationship between dependent and independent variables. She used the data for 10-year period which are from 2003 until 2012. The price of gold was the dependent variable and crude oil prices, inflation rates and exchange rates were independent variables. The empirical results have revealed that there is negative and significant association between inflation rates and exchange rates on gold prices, while a crude oil price is positively significant.

Samveg Patel (2013) investigated the relationship between the inflation, exchange rate and gold price in India by using monthly data from January 1991 to October 2012. He used the Johansen's Cointegration Test in order to establish relationship between these factors.

R Baber, G Thomas (2013) in their study "Factors affecting Gold prices: a case study of India" list different factors affecting the gold price in India and gives special emphasis on rise in gold price in the decade from 2002 to 2012.

Banhi Guha and Gautam Bandyopadhyay (2016) examined the Gold Price Forecasting Using ARIMA Model. This study gives an inside view of the application of ARIMA time series model to forecast the future Gold price in Indian browser based on past data from November 2003 to January 2014 to mitigate the risk in purchases of gold. Hence, to give guideline

for the investor when to buy or sell the gold.

Deepika M G, Gautam Nambiar & Rajkumar M (2012) has tried to study the forecasting of gold price through ARIMA model & Regression but their finding suggests that suitable model was not identified to forecast Gold price through ARIMA Model Hence Regression analysis was carried out in the later part of their study.

3. Data Collection and Research Methodology:

3.1 Population and Sample

Year	Gold Prices	Exchange Rate (₹ per US dollar)	CPI (in % p.a.)	BSE Index	Interest Rate	Crude oil Prices (per Barel)
1990-91	3451.52	17.4992	8.97%	1048.29	14%	387.5
1991-92	4297.63	22.6890	13.48%	1908.85	13%	461.2
1992-93	4103.66	25.9206	9.83%	2615.37	11%	586.5
1993-94	4531.87	31.4458	7.32%	3346.06	10%	496.8
1994-95	4667.24	31.3742	10.28%	3926.9	11%	525.3
1995-96	4957.60	32.4232	9.96%	3110.49	13%	584.2
1996-97	5070.71	35.4294	9.43%	3085.2	12.75%	749.2
1997-98	4347.07	36.3196	6.48%	3658.98	11.75%	645.4
1998-99	4268.17	41.2677	13.13%	3055.41	11%	524.9
1999-00	4393.56	43.0485	3.32%	5005.82	10.25%	945.4
2000-01	4473.60	44.9401	3.93%	3972.12	9.75%	1283.5
2001-02	4579.12	47.1857	4.32%	3262.33	8.25%	1098.4
2002-03	5332.36	48.5993	3.98%	3377.28	5.88%	1331.3
2003-04	5718.95	46.5819	3.85%	5838.96	5.38%	1335.7
2004-05	6145.38	45.3165	3.83%	6602.69	6.25%	1853.1
2005-06	6900.56	44.1000	4.41%	9397.93	6.75%	2528.2
2006-07	9240.32	45.3070	6.73%	13786.91	8.13%	2869.8
2007-08	9995.62	41.3485	6.14%	20286.99	8.75%	3239.7
2008-09	12889.74	43.5049	9.11%	9647.31	8.13%	3757.2
2009-10	15756.09	48.4049	12.38%	17464.81	7.25%	3310.8
2010-11	19227.08	45.7262	10.46%	20509.09	8.63%	3854.4
2011-12	25722.42	46.6723	8.38%	15454.92	8.18%	5134.4
2012-13	30163.93	53.4376	10.84%	19426.71	8.75%	5604.6
2013-14	29190.39	58.5978	9.68%	21170.68	8.80%	6283.4
2014-15	27414.55	61.0295	6.26%	27499.42	8.38%	5065.3
2015-16	26534.26	64.1519	5.66%	26117.54	7.15%	2997.9
2016-17	29665.28	67.1953	4.14%	26626.46	6.63%	3211.5
2017-18	29300.08	65.1216	3.79%	34056.83	6.50%	3591.5
2018-19	31193.41	68.3895	4.85%	36068.33	6.25%	4679.9
2019-20	37017.91	70.4203	7.66%	41253.74	5.75%	4322.6
2020-21	45032.31	74.6400	5.43%	47751.33	5.40%	3386.7

3.2 Tools and Techniques of Data Collection

The data used for this analysis has been gathered from a variety of reliable sources. The time period considered for the analysis spans from 1990 to 2021. For forecasting the gold prices on a monthly basis, monthly data of gold prices was utilized.

The data related to gold prices reflects the average price of gold (per 10 grams) in domestic and foreign markets and it obtained from the RBI Handbook of Statistics on Indian Economy available on

the official website of the Reserve Bank of India. Along with it the exchange rate which is expressed as average rupees per unit of U.S. dollar and Interest rates (expressed in percent per annum) have also been adopted from the RBI Handbook of Statistics on Indian Economy.

Additionally, data on the Consumer Price Index (CPI) for industrial workers was sourced from the Labour Bureau of the Government of India. The BSE Sensex index data was collected from the official website, www.bseindia.com. The monthly average prices of crude oil, measured in hundreds of rupees per barrel, were obtained from the respective online platforms.

3.3 Statistical tools and Forecasting Models

A range of statistical methodologies were utilized to examine the multifaceted factors driving gold price fluctuations. This analysis enabled the identification of crucial variables and the development of a predictive model for forecasting gold prices.

3.3.1 Linear Regression Model

Linear regression attempts to model the relationship between two variables one is the dependent variable Y and other one is explanatory (or independent) variable X. The regression model is given by

$Y_i = \beta_0 + \beta_i X_i + €_i \quad i = 1, 2, n$

Where,

Y_i = The i th observation on the dependent variable,

X_i = The i th observation on the independent variable,

β_0 = Intercept β_i = Slope coefficient

3.3.1.1 Simple Regression Model

The regression model in which the dependent variable is related to a single explanatory variable is known as simple regression model and it is expressed as

$Y = \beta_0 + \beta_1 X + €$

Where,

Y = Dependent variable

X = Independent variable

β_0 = Intercept

β_1 = Slope coefficient which measures unit change in mean value of Y as result of unit change in X.

€ = Error term or the residual for the given observation

The slope β_1 is the change in Y for a given one unit change in X. The slope can be positive, negative or zero, calculated as

$$\beta_1 = \frac{Cov(X,Y)}{Var(X)} = \frac{\sum_{i=1}^{N}(yi-\bar{y})(xi-\bar{x})}{\sum_{i=1}^{N}(xi-\bar{x})^2}$$

The intercept β_0 is the line of intersection with the Y-axis at X=0. The intercept can be positive, negative or zero, calculated as,

$$\beta_0 = \bar{y} - \beta_1 \bar{x}$$

If there is no reason between the dependent and an independent variable, the slope coefficient b would be zero (0). A zero slope indicates that there is no change in Y for a given change in X.

This model has been used as a facet to multiple regression model to explore the impact of each individual independent variable on the dependent variables.

3.3.1.2 Multiple Regression Model

The regression model in which the dependent variable is related to two or more explanatory variables is known as a multiple regression model and is expressed as

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_n X_n + €$

Where n denotes number of observations; the remaining symbols denote the same variables as explained above. The degrees of freedom for the test of a slope coefficient are N-k-1, where n is the number of observations in the sample and k is the number of independent variables.

3.3.2 Coefficient of Determination (R^2)

The coefficient of determination R^2, is the percentage of variation in the dependent variable explained by the independent variables. It is a measure that reveals about the 'goodness of fit' of the regression line to the observed data values. The value of R^2 lies between 0 and 1. The value of $R^{2\ near}$ 1 indicates that the regression line fits the data well, whereas an R2 closer to 0 indicates that a regression line does not fit the data well. It is defined as

$$R^2 = \frac{\text{Explained Variation}}{\text{Total Variation}}$$

$$= \frac{\text{Total Variation} - \text{Unexplained Variation}}{\text{Total Variation}}$$

$$= \frac{\text{SSRegression}}{\text{SSTotal}}$$

$$R^2 = \frac{\sum_{i=1}^{N} \frac{(\hat{y}-\bar{y})^2}{k}}{\sum_{i=1}^{N} \frac{(\hat{y}-\bar{y})^2}{n-k-1}}$$

Where,

$\hat{y}$ = The predicted value of the dependent variable i.e gold price

$\bar{y}$ = the mean of Y

n = Number of observations

K = Number of independent variables

SSTotal= Total sum of squares or total variation of actual Y value

SSRegression = Explained sum of squares or explained variation of Y values

3.3.3 Adjusted R^2

It is a modified version of R-squared for the number of independent variables in a model. Unlike R^2, which is a non-decreasing function of the number of explanatory variables in the model, adjusted R^2 increases only if a newly added explanatory variable enhances the model above what would be obtained by probability and decreases if a newly added explanatory variable enhances the model less than what is predicted by chance. The adjusted R-squared can turn out to be negative, in which case, its value is taken as zero. It is defined as

$$R^2_{adj} = 1 - \left[\frac{(1-R^2)(n-1)}{n-k-1}\right]$$

Where,

R^2 = Coefficient of determination

k = Number of Independent variables

N = Number of observations

3.3.4 Correlation Analysis

Correlation is a statistical tool that helps to measure and analyze the degree of linear relationship between two variables. Correlation analysis deals with the association between two or more variables. The coefficient of correlation is a measure of linear relationship between two variables. The value of the correlation coefficient lies between -1 to 1

3.3.5 F-test

This has been used to test the overall significance of a linear regression model as to its relevance in the context. The hypotheses formed for this purpose are as follows:

H0 = β1= β2=……. βn=0 vs

H1= Not all slope coefficients are simultaneously zero

The F-test is a measure of the overall significance of the estimated regression and therefore a test of significance of $R2$. The test statistic F is calculated as

$$F = R^2/(k-1)\ (1-R^2)/(n-k)$$

Where, R2 = Coefficient of determination, n = number of observations, k = number of variables

Test Rule: If calculated F –Value > tabulated F-Value, the null hypothesis is rejected.

	Critical t value = 1.697 At 0.05 level of Significance		Critical F value = 4.183 At 0.05 level of significance	
Independent Variables	Calculated t-value	Test of significance	Calculated F value	Test of significance
Exchange Rate	9.096	Significant	82.73	Significant
CPI	-0.53	Insignificant	0.289	Insignificant
BSE Index	15.73	Significant	247.6	Significant
Interest Rate	-3.75	Significant	14.11	Significant

3.3.6 T-test

T-test is used as a hypothesis testing tool that allows testing of an assumption applicable to a population. In regression analysis, the t-test is used to conduct hypothesis test on regression coefficients obtained in linear regression to determine whether there is a significant linear relationship between the dependent variable Y and an independent variable X. We state the hypotheses as follows

$$H0 : \beta 1 = 0 \text{ and } H1 : \beta 1 \neq 0$$

The test statistic is a t-statistic defined by the following equation:

$$t = \frac{\beta}{SE} \sim t\ df$$

Test Rule: If calculated value of t statistic exceeds tabulated t-value, null hypothesis is rejected and the slope of sample regression line, β is said to be statistically significant at a particular level of significance, which is generally taken as 5%.

3.3.7 Variance Inflation Factor

The Variance Inflation Factor (VIF) measures the severity of multicollinearity in Regression analysis. This test is performed to explore the possibility of multicollinearity in a set of multiple regression variables. VIF measures how much the variance of an independent variable is inflated by its correlation with other independent variables in the model. It is calculated as 1/Tolerance = (1/1-Ri ^{2}) and is always greater than or equal to 1.

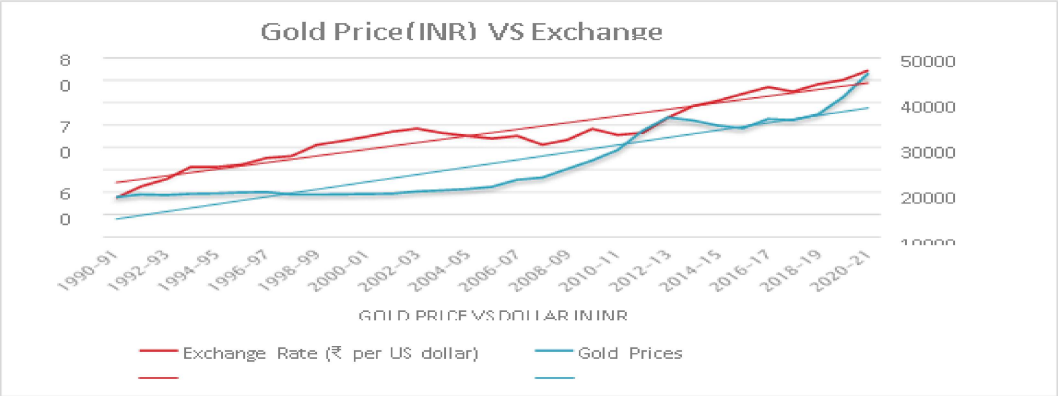
1. Statistical Analysis

4.1 Analysis for factor affecting gold prices determination

Result of Linear Regression:

Crude oil Prices	7.773	nificant	60.42	Significant

a) Gold Price vs Exchange Rate:



Hypothesis Assumed (H0): Gold Prices do not depend upon Exchange rate.

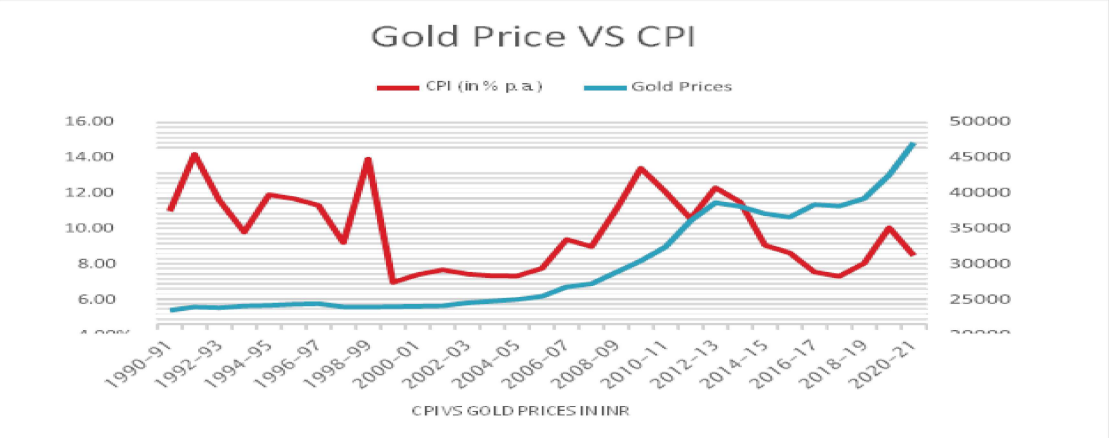
The regression equation is given by:

Y=-20271.1 + 748.6X

The regression coefficient β = 748.6 in respect of exchange rate which implies that an increase in exchange rate by 1 results in an increase in gold price by 748.6. Also the value of R2 is 0.74 which implies approximately 74% variation in gold prices accounted for with Dollar value. From t value, t=9.096 which is more than the tabulated value t=1.697(table 2) it can predict that there is a relation between Exchange rate and gold prices. Therefore, the hypothesis is rejected that gold prices do not depend upon the Exchange rate.

Independent Variable	Intercept α	Slope coefficient β	S.E	Coefficient of Correlation	R^2	Adj R^2
Exchange Rate	-20271.1	748.6	82.3	0.860	0.7404	0.7315
CPI	17687	-4065.6	75587	-0.099	0.0098	-0.0242
BSE Sensex	1.905e+03	9.005e-01	5.723e-02	0.946	0.8952	0.8915
Interest Rate	40201	-2899.8	77184	-0.572	0.3274	0.3042
Crude Oil Price	577.2924	5.7105	0.7347	0.821	0.6757	0.6645

b) Gold Price vs CPI:



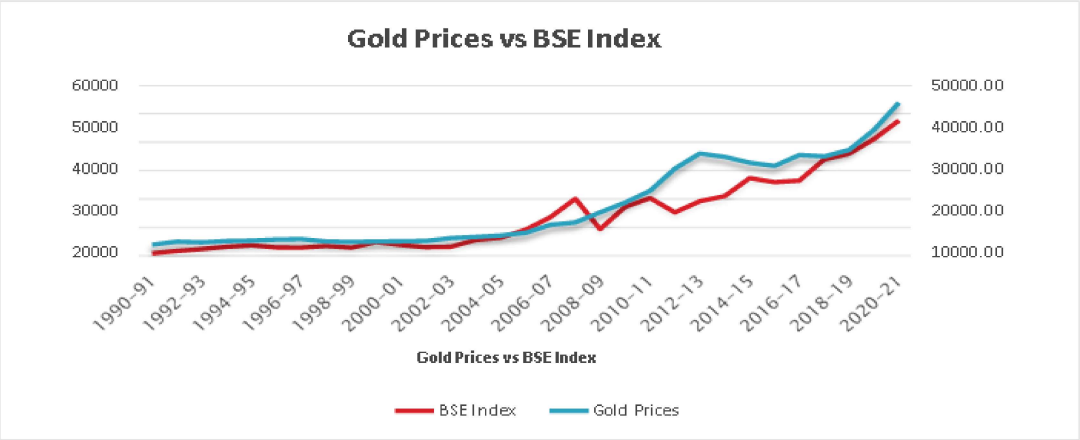
Hypothesis Assumed (H0): Gold prices do not depend on CPI

The Regression equation is given by:

Y= 17687 - 406.56X

The correlation coefficient between Inflation (CPI) on gold prices is very low. Also the correlation coefficient is -0.099 by the negative value it means if one factor is increasing the other is decreasing, it is clear from the fig. The R^2 = 0.0098 which is almost zero and also statistically insignificant as revealed by the F-test. The calculated F-value 0.2893 does not exceed the critical F-value of 4.183 at 5% level of significance, this indicating that the regression model is irrelevant. This implies that there is no any effect of CPI on gold prices.

C) Gold Price vs BSE Index



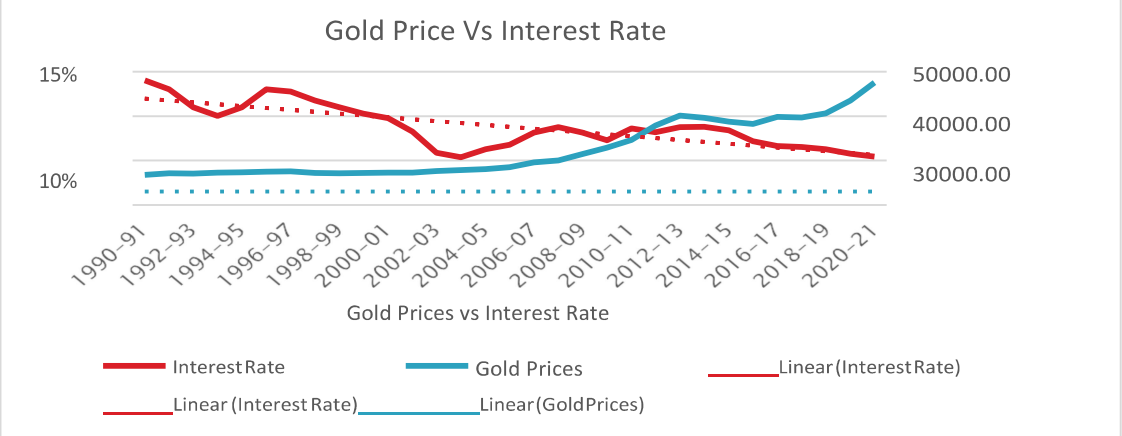
Hypothesis Assumed (H0): Gold prices do not depend on BSE Index.

The Regression equation is given by:

Y= 1905 + 0.9005X

The regression coefficient β = 0.9 which is nearly 1 in respect of BSE Sensex implies that an increase in BSE Sensex by 100 base results in an increase in gold price by ₹900. Also the value of R2 is 0.89 which implies approximately 89% variation in gold prices accounted for with BSE index. The calculated t-value = 15.736 exceeds the tabulated t-value of 1.697. Thus it can be predicting that there is a relation between Gold Price and BSE Index. Thus it suggests that the positive impact of BSE Sensex on gold prices is statistically significant at 5% level of significance.

d) Gold Price Vs Interest Rate

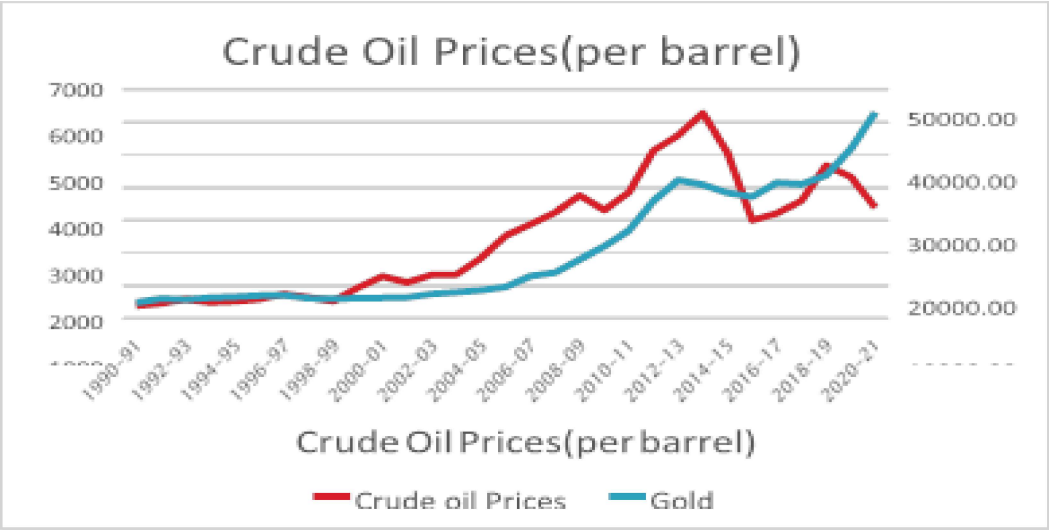


Hypothesis Assumed (H0): The Interest rate does not affect the gold prices.

The Regression equation is given by:
Y= 40200.1 – 2899.81*X

The correlation coefficient between Interest Rate and gold prices reveals that gold prices are negatively correlated with interest rate, as their coefficient of correlation is, -0.572 but not fairly high. The adjusted R2 = 0.3274 is statistically significant as per the F-test (Table 2). The calculated F-value 14.11 exceeds the critical F-value of 4.183, with (1, 29) d.f., at 5% level of significance, indicating that the regression model is relevant. The regression coefficient β= –2899.81 implies that an increase in interest rate by 1% results in a decrease in gold price by 2900. The t value is -3.757 which is negative but mod of t- value is greater than tabulated t value and negative sign only shows the inverse correlation with prices. This leads to the conclusion that the negative impact of interest rates on gold prices is statistically significant at 5% level of significance.

a) Gold Price Vs Crude Oil Price



Hypothesis Assumed (H0): The Crude oil prices does not affect the gold prices.

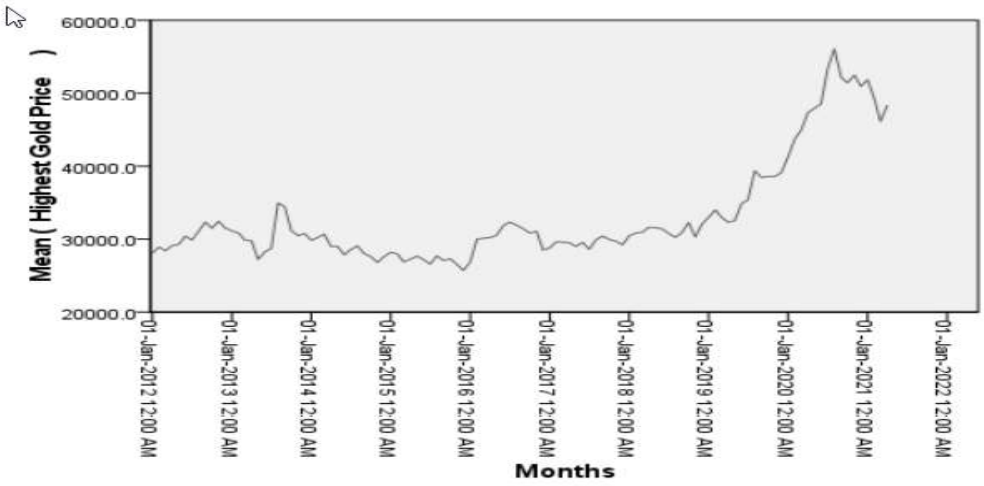
The regression equation is given by:

Y= 577.2924 + 5.7105X

The regression coefficient β= 5.710 implies that an increase in crude oil price by 100 results in an increase in gold price by 571.05. Also the value of R2 is 0.67 which implies approximately 67% variation in gold prices accounted for with Crude Oil Prices. Also t=7.773 which is more than the tabulated value t=1.697 (table 2) it can predict that there is a relation between Crude oil and gold prices. Therefore, the hypothesis is rejected that gold prices do not depend upon the Crude Oil Prices. Thus it suggests that the positive impact of crude oil prices on gold prices is statistically significant at 5% level of significance.

Forecasting

ARIMA model is used for analysing a times series which shows non stationarity. The following graph shows our data is based on time series analysis which is increasing as time increases and it shows non stationarity.



Non stationarity was measured through sequence charts and The Augmented Dickey Fuller test. The ADF was checked and concluded p value by Python Jupyter Notebook.

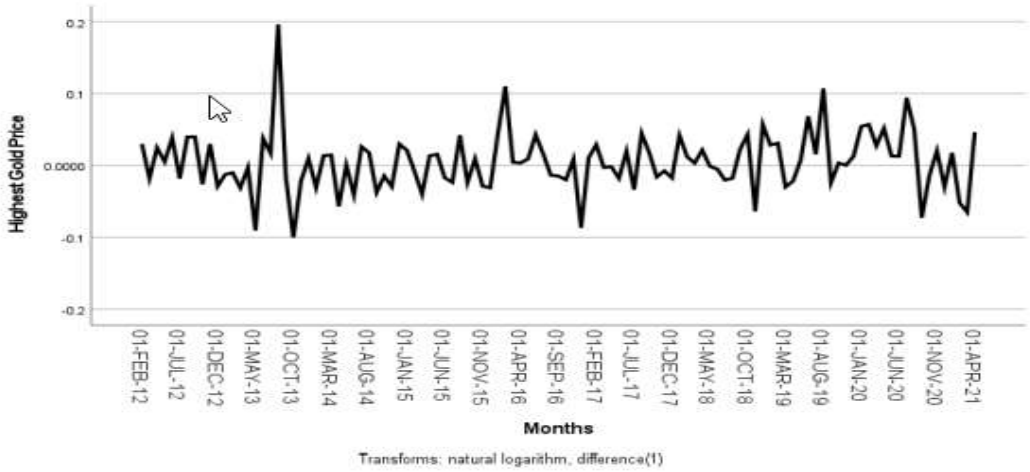
Result of Dickey Fuller Test7

p value 0.947261 #lags used 0.000000

Number of Observations used 111.000000

dtype: float64

Non stationarity can be removed by differencing the time series. The following graph is the first order difference of the above graph. A stationary process has the property that the mean, variance and autocorrelation structure do not change over time.



Above graph shows the fluctuation around mean i.e 0 which shows that our data is stationary. Stationarity can be defined in precise mathematical terms, but for our purpose we mean a flat looking series, without trend, constant variance over time, a constant auto correlation structure over time and no periodic fluctuations.

We checked the data for stationarity by checking for autocorrelation and partial autocorrelation as well as through the ADF test. The result of Augmented Dickey fuller

test is calculated in Python Jupyter notebook and given by,

Result of Dickey Fuller Test
Test statistic -1.029061e+01
p value 3.578278e-18 #lags used
0.000000e+00
Number of Observations used
1.100000e+02
dtype: float64

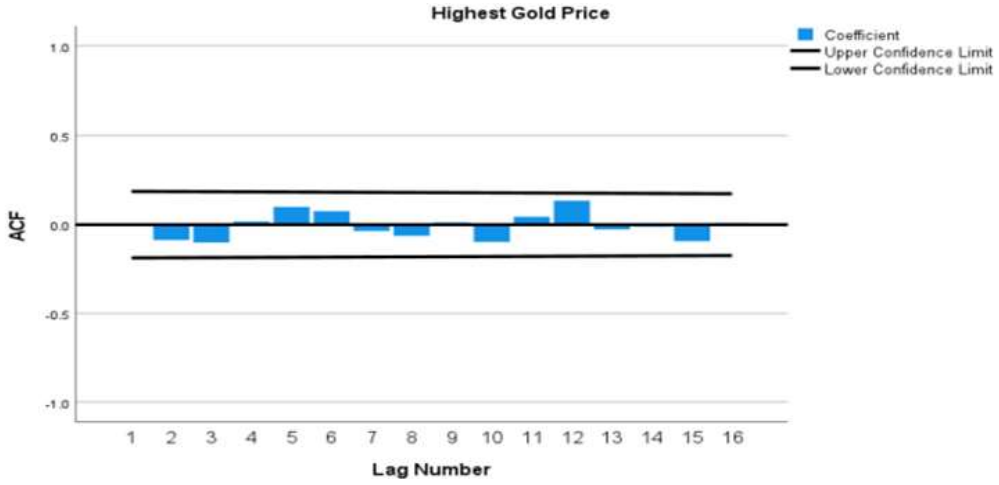
In this data, we can clearly see the p-value is less than 0.05., which shows the data is stationary.

After checking the stationarity by ADF test, ACF and PACF plots must obtain to check the stationarity and lag values present in the data. In ACF and PACF plots make sure many of the lag values may not cross the level of significance, if any lag value crosses the level at maximum then it is taken as our order of Autocorrelation in PACF and order of moving averages in ACF. ARIMA model has three model parameters AR(p), I(d) and MA(q) all combined to form ARIMA (p, d, q) model where,

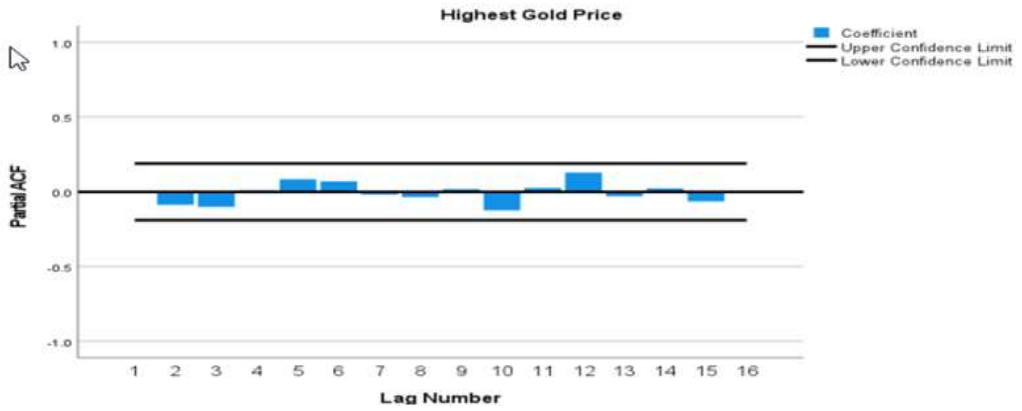
p = order of autocorrelation d = order of integration (differencing) q = order of moving averages

A. Autocorrelation Plot

It is defined by ACF = corr (X_t , X_{t+k}) i.e. relationship between each other. Here X_t is the current observation and X_{t+k} is observation after k period. It ranges from -1 to +1.



A. Partial Auto-Correlation Plot



From the ACF and PACF plot it is clear that there is no lag values outside the critical region. Hence, we conclude that the value of p and q.

Selection of Appropriate ARIMA (p, d, q) Model

Model for non-seasonal series is called Autoregressive Integrated Moving Average Model, denoted by ARIMA (p, d, q). Here "p" the order of autoregressive part, "d" indicates the order of differencing, and "q" indicates the order of moving average part. In general, a series which is stationary after being differenced "d times" is said to be integrated of order "d", denoted by I(d). If the original series is stationary, d=0 and the ARIMA models reduce to ARMA models. In our case we have d=1. To get the appropriate numbers for "p" (in AR) and "q" (in MA) in the model, we should check the Correlogram after first difference in time series. Since there are no significant spikes of ACF and PACF, the residuals of the selected ARIMA model are white noise and there is no need for further consideration of one more AR(p) and MA(q). To choose one best ARIMA model amongst a numerous combinations present, the following criterions are used.

a. Comparatively low Normalized Bayesian Information Criterions (BIC).
b. Comparatively high adjusted R-square (R2).
c. Root Mean Square Error (RMSE) should be relatively low.
d. Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) should be low.

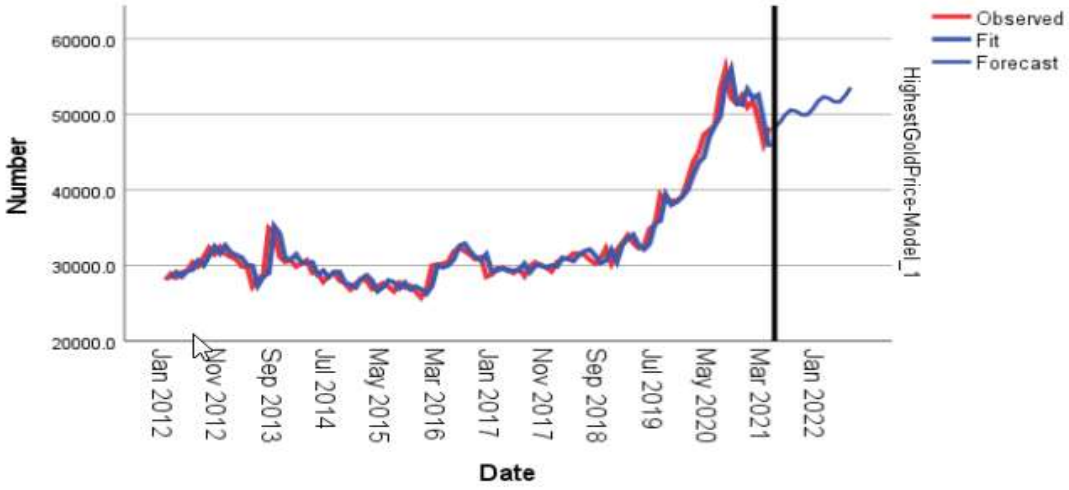
ARIMA	RMSE	MAE	MAPE	Normalised BIC	Adjusted R2
(0,1,0)	**1459.133**	**1023.094**	**2.990**	**14.614**	**.959**
(1,1,0)	1465.505	1023.351	2.991	14.665	.959
(1,1,1)	1470.296	1021.697	2.988	17.383	.959
(2,1,0)	1462.270	1012.065	2.956	17.571	.96
(2,1,1)	1470.368	1009.276	2.943	14.756	.96
(1,1,2)	1471.788	1011.639	2.950	14.758	.959
(2,1,2)	**1438.414**	**987.079**	**2.888**	**14.755**	**.962**

From the table above first model i.e (0,1,0) and (2,1,2) showing good results. (2,1,2) shows more precise results than (0,1,0).

After checking the robustness of the statistics given in the above Table, it is found that only ARIMA (2,1,2) model convinces all the norms (Normalised BIC, RMSE, MAE, MAPE and the relatively high Adjusted R2 values), hence this model is considered to be the best predictive model, which is used to forecast the future values of the time series.

FORECASTING USING SELECTED ARIMA (p, d, q) MODEL

The present study is based on monthly data on the highest gold prices covering the period from January 2012 to April 2021, having a total number of 112 observations. Of which, the period from January 2021 to July 2022, having 19observations are used for forecasting length.



Month	Predicted Values	Actual Values
Jan-2021	52152.7	51842
Feb-2021	52588.5	49209
Mar-2021	49057.4	46119
Apr-2021	45652.8	48307
May-2021	48268.7	49150
Jun-2021	48965.7	49674
Jul-2021	50580.9	
Aug-2021	50416.2	
Sep-2021	49951.5	
Oct-2021	49944.6	
Nov-2021	50691.3	
Dec-2021	51741.7	
Jan-2022	52321.4	
Feb-2022	52125	
Mar-2022	51653.3	
Apr-2022	51680.3	
May-2022	52478.4	
Jun-2022	53556.1	

5. CONCLUSION

In our study we tried to identify the impact of various factor on gold price and try to build a forecasting model which help in predicting gold prices in upcoming months. From analysis, we understand that impact of Exchange Rate, BSE Sensex and Crude oil price on gold price is positive and statistically significant. Whereas the individual impact of inflation and interest rate on gold price is negative and insignificant. The data was transformed and using ADF test the stationarity was tested. The lag values were concluded by ACF and PACF plots. Taking the possible neighbor values from resulting model and Using mean absolute percentage error (MAPE), Bayesian Information Criterion(BIC), Mean Absolute Error(MAE) and the value of R2 as the forecasting accuracy measure. The study concludes that ARIMA (2,1,2) model is more appropriate model from all other models to forecast the monthly gold prices.

Reference:

1. Sindhu, D. (2013). *A study on the impact of select factors on the price of gold*.
2. Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2015). *Time series analysis: Forecasting and control*.
3. Gupta, S. C. (2014). *Fundamentals of mathematical statistics*.
4. Guha, B., & Bandyopadhyay, G. (2016). *Gold price forecasting using the ARIMA model*.
5. Sharma, A. M., & Baby, S. (2015). *Gold price forecasting in India using ARIMA modelling*.
6. Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*.
7. https://towardsdatascience.com/detecting-stationarity-in-time-series-data-d29e0a21e638
8. https://www.analyticsvidhya.com/blog/2016/02/time-series-forecasting-codes-python

***