PROJECT CODES

(R-Programming language)

```
X=read.csv(file.choose())

X

YEARS=(X$YEARS)

YEARS

GDP_GROWTH_RATES=(X$GDP.GROWTH..RATES)

GDP_GROWTH_RATES

EXCHANGE_RATES=(X$EXCHANGE.RATES)

EXCHANGE_RATES

CPI_RATES=(X$CPI.RATES)

CPI_RATES

LENDING_RATES=(X$ LENDING.RATES)

LENDING_RATES

INFLATION_RATES=(X$INFLATION.RATES)

INFLATION_RATES

#Descriptive statistics

summary(X)

library(moments)

skewness(X)

kurtosis(X)

cor(X, method = c("pearson", "kendall", "spearman"))

sd(GDP_GROWTH_RATES)

sd(EXCHANGE_RATES)

sd(CPI_RATES)

sd(LENDING_RATES)

sd(INFLATION_RATES)

#Data visualizations

#line graph
```

```
plot(YEARS,GDP_GROWTH_RATES,type="l")

plot(YEARS,EXCHANGE_RATES,type="l")

plot(YEARS,CPI_RATES,type="l")

plot(YEARS,LENDING_RATES,type="l")

plot(YEARS,INFLATION_RATES,type="l")

#Histogram

hist(GDP_GROWTH_RATES)

hist(EXCHANGE_RATES)

hist(CPI_RATES)

hist(LENDING_RATES)

hist(INFLATION_RATES)

#Model fitting

model=lm(INFLATION_RATES~GDP_GROWTH_RATES+EXCHANGE_RATES+CPI_RATES+LENDING_RATES,
data=X)

model

summary(model)

plot(model)
```

RESULTS AND FINDINGS

4.1 Introduction.

This chapter involves examining and interpreting the data collected for our study. We will explore the findings and draw meaningful conclusions from the descriptive statistics, data visualizations and the fitted regression model.

4.2 Descriptive Statistics.

This is the summary and description of the main features of the variables such as the central Tendency, variability and distribution. These statistics provide an overview of the data and help identify patterns and relationships in the variables.

```
> X=read.csv(file.choose())
> X
   YEARS GDP.GROWTH..RATES EXCHANGE.RATES CPI.RATES LENDING.RATES INFLATION.RATES
1   1971             22.17           7.14      0.93          9.00             3.78
2   1972             17.08           7.14      0.98          9.00             5.83
3   1973              5.90           7.00      1.07          9.00             9.28
4   1974              4.07           7.14      1.26          9.50            17.81
5   1975              0.88           7.34      1.50         10.00            19.12
```

6	1976	2.15	8.37	1.68	10.00	11.45
7	1977	9.45	8.28	1.92	10.00	14.82
8	1978	6.91	7.73	2.25	10.00	16.93
9	1979	7.62	7.48	2.43	10.00	7.98
10	1980	5.59	7.42	2.77	10.58	13.86
11	1981	3.77	9.05	3.09	12.42	11.60
12	1982	1.51	10.92	3.73	14.50	20.67
13	1983	1.31	13.31	4.15	15.83	11.40
14	1984	1.76	14.41	4.58	14.42	10.28
15	1985	4.30	16.43	5.17	14.00	13.01
16	1986	7.18	16.23	5.30	14.00	2.53
17	1987	5.94	16.45	5.76	14.00	8.64
18	1988	6.20	17.75	6.47	15.00	12.26
19	1989	4.69	20.57	7.36	17.25	13.79
20	1990	4.19	22.91	8.67	18.75	17.78
21	1991	1.44	27.51	10.41	19.00	20.08
22	1992	-0.80	32.22	13.26	21.07	27.33
23	1993	0.35	58.00	19.35	29.99	45.98
24	1994	2.63	56.05	24.93	36.24	28.81
25	1995	4.41	51.43	25.31	28.80	1.55
26	1996	4.15	57.11	27.56	33.79	8.86
27	1997	0.47	58.73	30.69	30.25	11.36
28	1998	3.29	60.37	32.75	29.49	6.72
29	1999	2.31	70.33	34.63	22.38	5.74
30	2000	0.60	76.18	38.09	22.34	9.98
31	2001	3.78	78.56	40.27	19.67	5.74
32	2002	0.55	78.75	41.06	18.45	1.96
33	2003	2.93	75.94	45.09	16.57	9.82
34	2004	5.10	79.17	50.34	12.53	11.62
35	2005	5.91	75.55	55.53	12.88	10.31
36	2006	6.47	72.10	63.55	13.64	14.45
37	2007	6.85	67.32	69.75	13.34	9.76
38	2008	0.23	69.18	88.06	14.02	26.24
39	2009	3.31	77.35	96.19	14.80	9.23
40	2010	8.06	79.23	100.00	14.37	3.96
41	2011	5.12	88.81	114.02	15.05	14.02
42	2012	4.57	84.53	124.72	19.72	9.38
43	2013	3.80	86.12	131.85	17.31	5.72
44	2014	5.02	87.92	140.91	16.51	6.88
45	2015	4.97	98.18	150.19	16.09	6.58
46	2016	4.21	101.50	159.65	16.56	6.30
47	2017	3.84	103.41	172.43	13.67	8.01
48	2018	5.65	101.30	180.51	13.06	4.69
49	2019	5.11	101.99	189.97	12.44	5.24
50	2020	-0.27	106.45	200.23	12.00	5.40
51	2021	7.59	109.64	212.47	12.08	6.11
52	2022	4.85	117.87	228.74	12.34	7.66

```
> summary(X)
```

	GDP.GROWTH.RATES	EXCHANGE.RATES	CPI.RATES	LENDING.RATES	INFLATION.RATES
Minimum	-0.800	7.00	0.93	9.00	1.550
1st Quartile	2.270	14.13	4.473	12.40	6.253
Median	4.255	58.37	29.125	14.39	9.790
Mean	4.599	52.34	57.377	16.30	11.506
3rd Quartile	5.902	79.19	97.142	18.52	13.900
Maximum	22.170	117.87	228.740	36.24	45.980
S.D	3.875853	36.62541	67.61805	6.524063	7.880084
Skewness	2.30856016	0.04180871	1.1131215	1.42277171	2.02457811
Kurtosis	10.997722	1.513018	2.923248	4.464064	8.694079

Our summary statistics indicated that inflation hit its highest at 45.980% in 1993 with a standard deviation of 7.880084. However, only exchanges rates is approximately symmetric while the other variables are highly skewed. The standard deviation of exchange rates and CPI rates is relatively

higher compared other variables, this means that exchange rates and CPI rates observations are more deviated from their mean. All variables have positive kurtosis values this indicates that they have relatively higher peaks than the normal curve distribution.

```
>Correlation matrix
> cor(X, method = c("pearson","kendall","spearman"))
```

	YEARS	GDP.GROWTH.RATES	EXCHANGE.RATES	CPI.RATES	LENDING.RATES	NFLATION.RATES
YEARS	1.000000	-0.24601433	0.9714401	0.90129731	0.17071831	-0.2589895
GDP.GROWTH.RATES	-0.24601	1.00000000	-0.2087752	-0.0606365	-0.4103590	-0.3658451
EXCHANGE.RATES	0.971440	-0.20877515	1.0000000	0.88091150	0.21547913	-0.2672189
CPI.RATES	0.901297	-0.06063657	0.8809115	1.00000000	-0.0997056	-0.3269809
LENDING.RATES	0.170718	-0.41035904	0.2154791	-0.0997056	1.00000000	0.2615392
INFLATION.RATES	-0.25898	-0.36584513	-0.2672189	-0.3269808	0.26153916	1.0000000

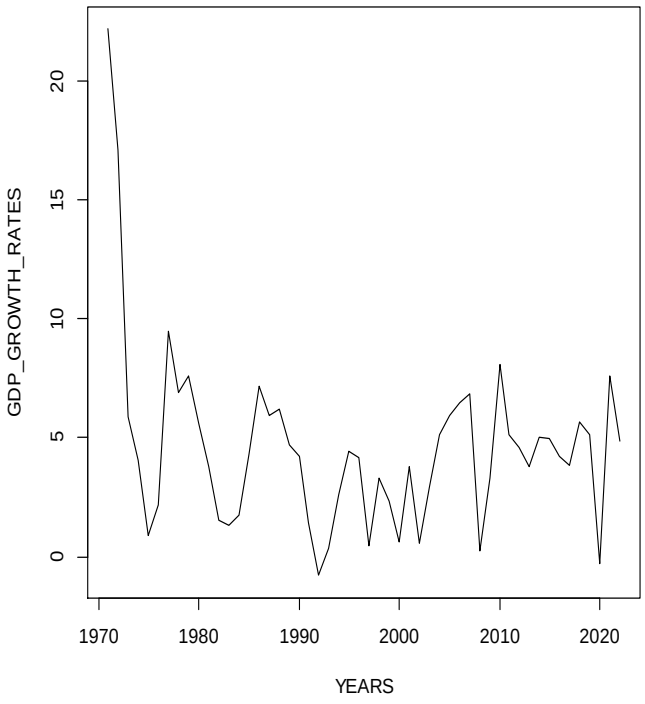
In our correlation matrix, the independent variables didn't show any serious relationships except for CPI rates and exchange rates with a positive correlation of 0.88091150. This means that CPI rates has a strong and positive relationship with exchange rates such that an increase in CPI rates will result to an increase in exchange rates and vice versa.

4.3 Data visualizations.

In our analysis we used a line graph to visualize the changes in the variables over the years. Also, a scatter plot was used to visualize the distribution, the spread of each variable and identify any outliers. Additionally, a histogram was used to show the shape of the distribution, identify its central tendency and observe any skewness.

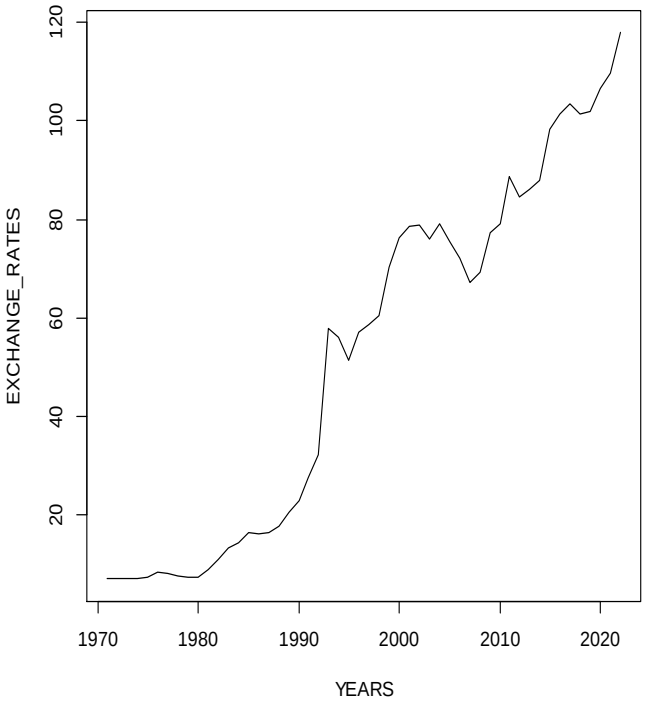
i) Line graph

```
> plot(YEARS,GDP_GROWTH_RATES,type="l")
```



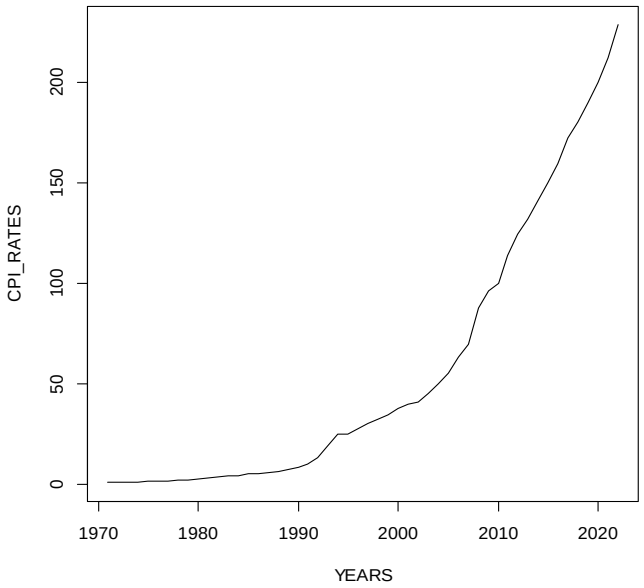
GDP growth rates seems to vary with time which shows fluctuations in economic activities.

```
> plot(YEARS,EXCHANGE_RATES,type="l")
```



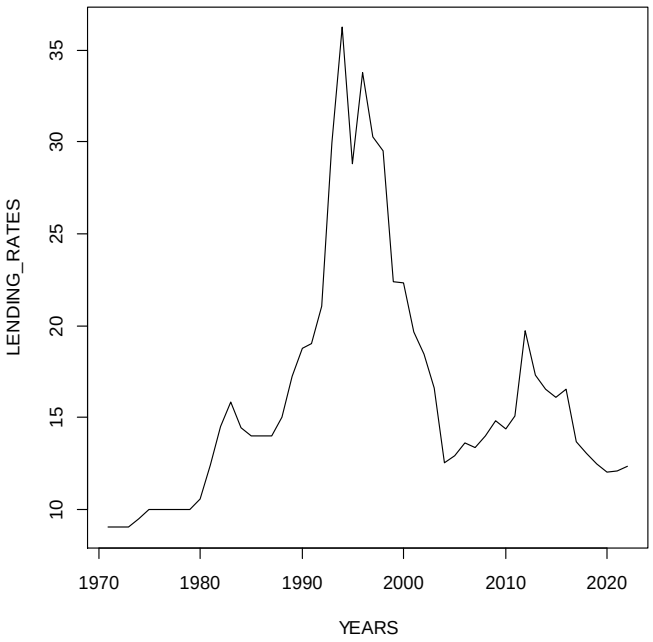
Exchange rates have a rising trend which indicates that the value of the Kenyan shilling is decreasing over time compared to the USD.

```
> plot(YEARS,CPI_RATES,type="l")
```



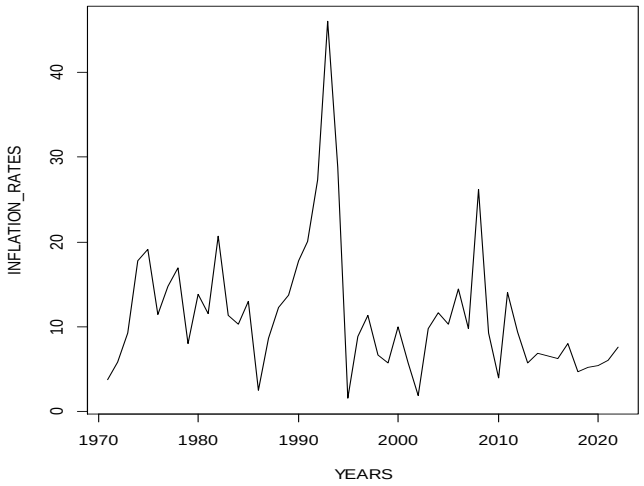
CPI rates trend indicates that the average prices of goods and services are gradually increasing over time.

```
> plot(YEARS,LENDING_RATES,type="l")
```



Lending rates trend indicates that higher lending rates cluster around the central year of our data (1995) with fewer rates at the extremes.

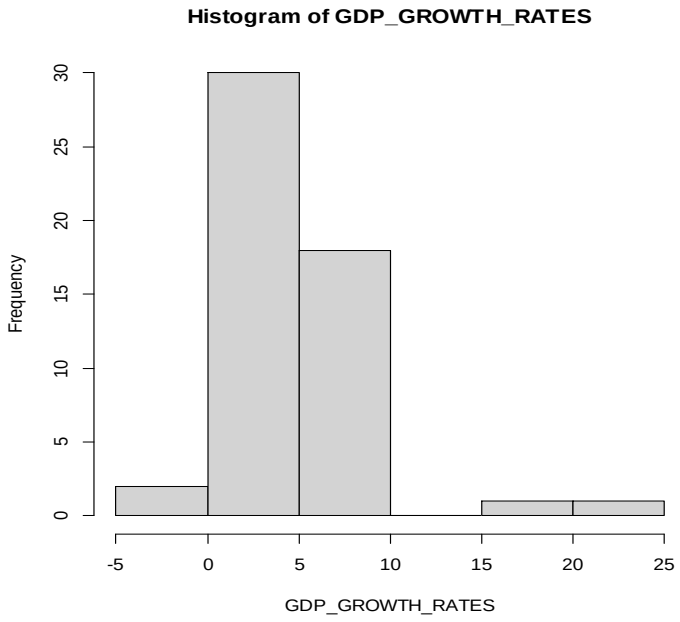
```
> plot(YEARS,INFLATION_RATES,type="l")
```



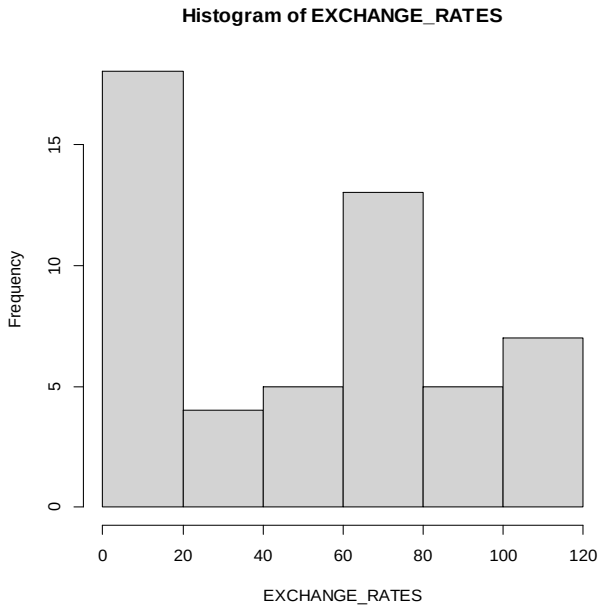
Inflation rates trend indicates that the rates vary with time which shows fluctuations prices of goods and services over the years.

ii) Histogram

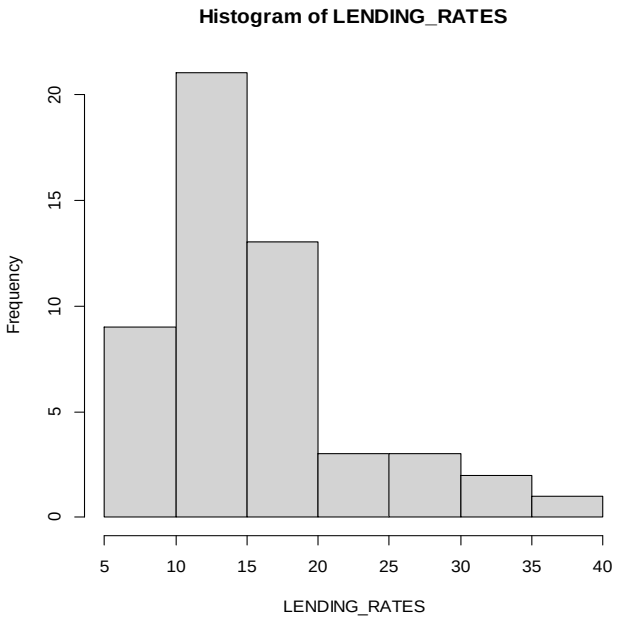
```
> hist(GDP_GROWTH_RATES)
```



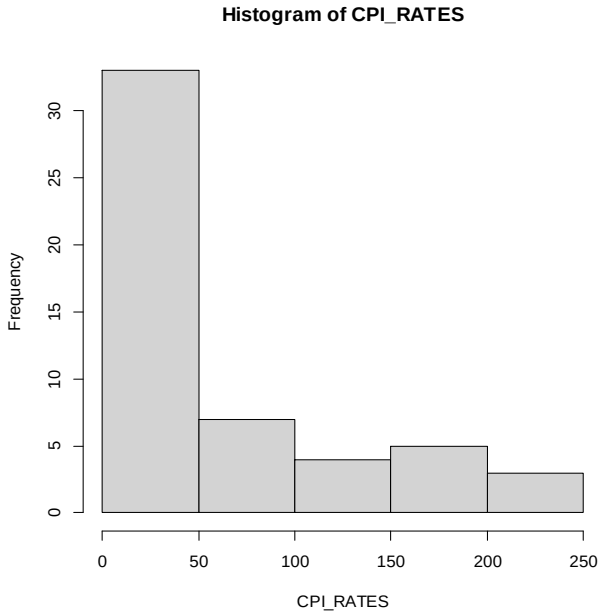

```
> hist(EXCHANGE_RATES)
```



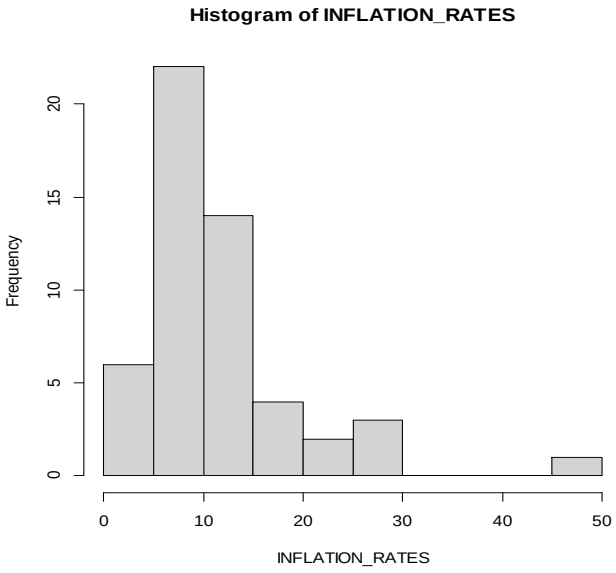
```
> hist(LENDING_RATES)
```




```
> hist(CPI_RATES)
```




```
> hist(INFLATION_RATES)
```



The histograms show that, except Exchange rates histogram which suggests that the exchange rates are scattered and random without any noticeable clustering, the other variables exhibit asymmetric properties. This can be translated to most of the data points being on the left side/lower values.

4.4 Multiple Linear Regression fitting.

A multiple linear regression was fitted using Inflation as the response variable and CPI rates, GDP growth rates, exchange rates and Lending rates as the predictor variables and the results were as follows;

```
>model=lm(INFLATION_RATES~GDP_GROWTH_RATES+EXCHANGE_RATES+CPI_RATES+LENDING_RATES,data=X)
>model

Call:
lm(formula=INFLATION_RATES~GDP_GROWTH_RATES+EXCHANGE_RATES+CPI_RATES+LENDING_RATES,data=X)

Coefficients:
     (Intercept)  GDP_GROWTH_RATES    EXCHANGE_RATES         CPI_RATES     LENDING_RATES
       15.329223          -0.747718          -0.097725           0.008431           0.260543
```

The multiple regression analysis conducted on the relationship between inflation rates and our predictor variables revealed that the intercept of the model is at 15.329223 which signifies the baseline inflation rate when all the predictor variables are neutral. Moving on to the coefficients of the independent variables, a notable inverse relationship emerges between GDP growth rates and inflation. Specifically, for every unit increase in GDP growth rates inflation rates are expected to decrease by approximately 0.747718 units holding other variables constant. Similarly, exchange rates exhibit a negative impact on inflation with a unit increase in exchange rates associated with a decrease in inflation rates by around 0.097725 units. Conversely, Lending rates demonstrate a positive association with inflation indicating that a unit increase in lending rates correspond to an increase in inflation rates by approximately 0.260543 units. Lastly, the coefficient for CPI rates is relatively small indicating that changes in Consumer Price index have a minor impact on inflation rates.

4.5 Model Validation.

i) Model summary

The summary statistics of the final regression model was conducted and the results were as follows;

```
> summary(model)
```

```
Call:
lm(formula=INFLATION_RATES~GDP_GROWTH_RATES+EXCHANGE_RATES+CPI_RATES+LENDING_RATES,dat
a=X)
Residuals:
     Min       1Q   Median       3Q      Max
-13.1728  -3.9984  -0.6969   3.0624  28.6037
Coefficients:

                  Estimate Std. Error t value Pr(>|t|)
(Intercept)      15.329223   3.758272   4.079 0.000174 ***
GDP_GROWTH_RATES -0.747718   0.277006  -2.699 0.009626 **
EXCHANGE_RATES   -0.097725   0.073373  -1.332 0.189321
CPI_RATES         0.008431   0.038890   0.217 0.829312
LENDING_RATES     0.260543   0.203805   1.278 0.207388
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 6.923 on 47 degrees of freedom
Multiple R-squared:  0.2887,      Adjusted R-squared:  0.2281
F-statistic: 4.769 on 4 and 47 DF,  p-value: 0.002599
```

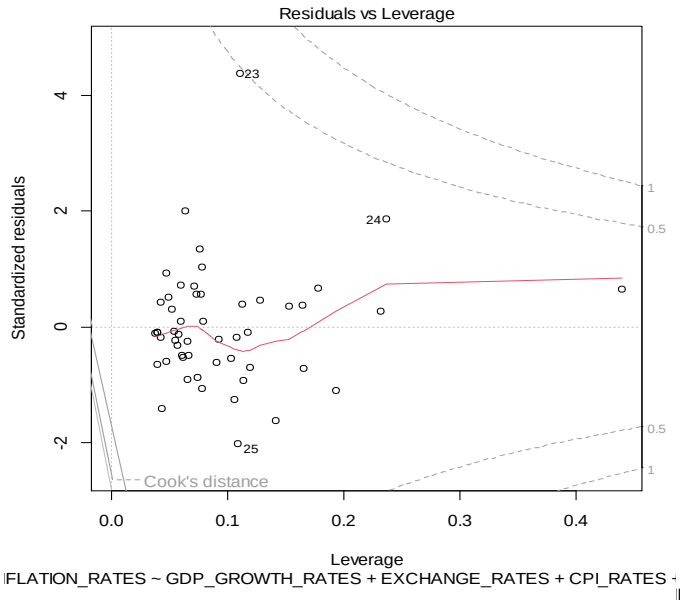
The model intercept (15.329223) and the coefficient of GDP growth rates (-0.747718) are statistically significant with a p-value of 0.000174 and 0.009626 respectively. This implies that the negative relationship between inflation rates and GDP growth rates is indeed realistic and a stronger economic growth tends to lead to lower inflationary pressures which aligns to economic theory. However, the coefficients for exchange rates, CPI rates and lending rates have p-values of 0.189321, 0.829312 and 0.207388 respectively. This shows that there is no enough evidence to conclude that there is any significant relationship between these variables and the inflation rates, that is, they do not have substantial influence on inflation rates in this model.

Overall, the model exhibits statistical significance as indicated by the F-statistic of 4.769 with a p-value of 0.002599 meaning that the included variables together contribute significantly to explain the variations in the inflation rates. Additionally the model has a Residual standard error of 6.923 which means that the average difference between the predicted values and the actual values of the dependent variable is relatively small. Lastly, the multiple R-squared of 0.2887 implies that the model explains only 28.87% of the variability in inflation rates.

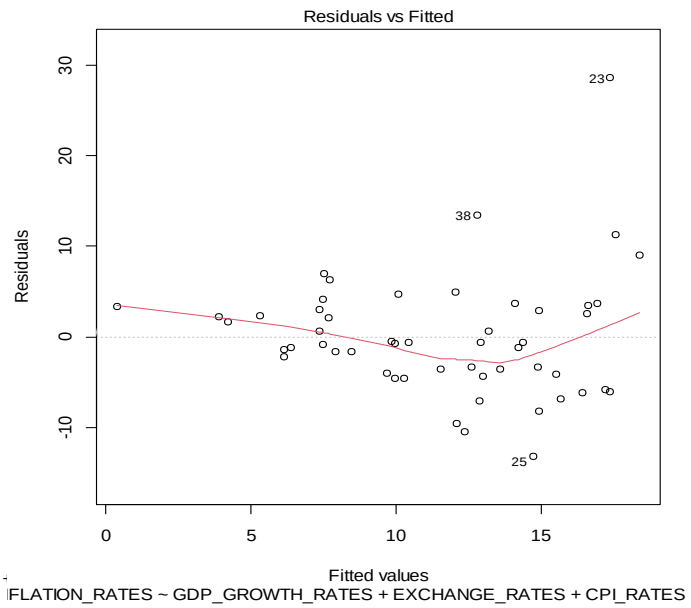
ii) Model Assessment

The model visualization plots are used to assess several assumptions of the multiple linear regression model and identify potential patterns in the model. These diagnostic plots are crucial for evaluating issues with the model such as heteroscedasticity, non-linearity and outliers.

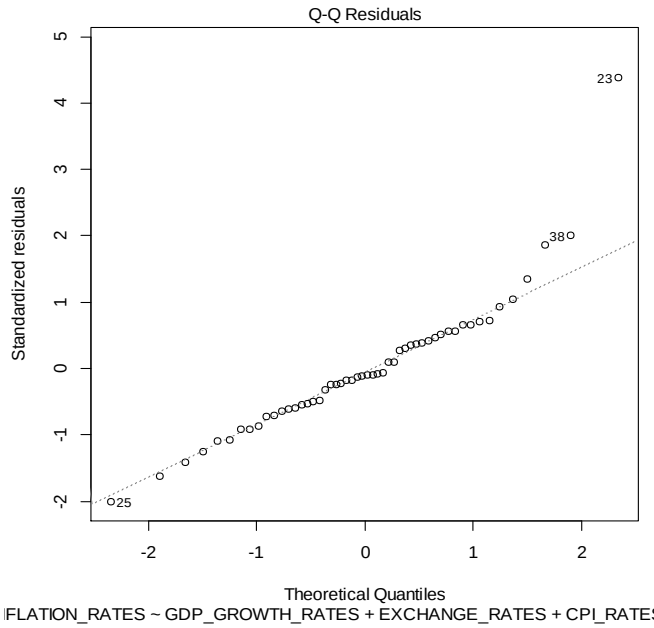
```
> plot(model)
```

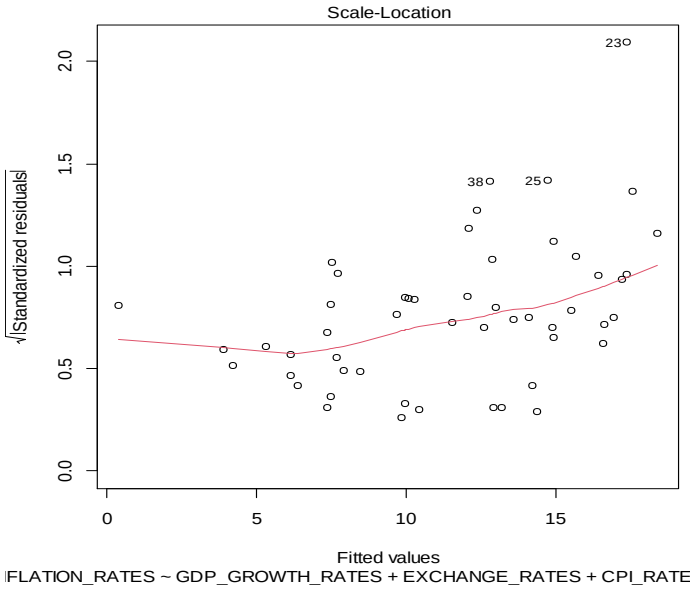


This plot is used to check for leverage points and influential points. Higher leverage points may affect the estimated coefficients if they are influential in the model. We use the cooks distance to determine which points are influential and needs further investigations. In our analysis, no data point lies outside the cooks distance.



This plot is used to assess the relationship between the fitted values and the predicted values. Since there is no discernible pattern in the data points, this suggests that the model adequately captures the assumption of linear relationship between variables.





This plot is used to assess the normality of residuals and check for tails in the model. A straight line indicates that the residuals follow a normal distribution while the points that deviate from the straight line indicates the tails. Our analysis revealed that the residuals follow a normal distribution and that our model isn't heavily tailed since almost all points are around the straight line.

The main purpose of scale location plot is to check for homoscedasticity. A random scatter of points around the middle line implies that the spread of residuals is approximately constant across the range of fitted value.