STATISTICAL DATA ANALYSIS PROJECT

A COMPARATIVE STUDY OF INFLATION RATES IN INDIA, CHINA, USA, GERMANY, AND JAPAN

GROUP MEMBERS:

KUNTAL DUTTA(A24007)

JAYANT KUMAR(A24005)

INTRODUCTION

- IMPORTANCE OF ANALYZING INFLATION RATES:
  - INFLATION IMPACTS THE COST OF LIVING, ECONOMIC STABILITY, AND GLOBAL TRADE.
  - COMPARING INFLATION RATES ACROSS COUNTRIES HELPS UNDERSTAND ECONOMIC HEALTH.
- COUNTRIES COMPARED:
  - INDIA, CHINA, USA, GERMANY, JAPAN
- PROJECT OBJECTIVES:
  - ANALYZE INFLATION TRENDS OVER TIME FOR THE SELECTED COUNTRIES.
  - UNDERSTAND THE RELATIONSHIPS BETWEEN INFLATION AND OTHER FACTORS
  - UNDERSTAND THE DIFFERENCES BETWEEN COUNTRIES' INFLATION RATES.

DATA OVERVIEW

- **DATASET INCLUDES:**
  - INFLATION RATE
  - COUNTRIES (INDIA, CHINA, USA, GERMANY, JAPAN)
  - YEARS (PERIOD OF THE DATASET)
  - GDP AND GDP GROWTH RATE
  - GDP PER CAPITA AND GDP PER CAPITA GROWTH RATE
  - UNEMPLOYMENT
- **PREPROCESSING STEPS:**
  - SET DATE AS AN INDEX TO HANDLE TIME SERIES DATA.
  - HANDLE MISSING DATA THROUGH INTERPOLATION OR CLEANING.

METHODOLOGY

- **ANALYSIS METHODS:**
  - EXPLORATORY DATA ANALYSIS (EDA) FOR DESCRIPTIVE STATISTICS AND UNDERSTANDING DATA.
  - TIME SERIES ANALYSIS OF INFLATION TRACK CHANGES OVER TIME.
  - CORRELATION BETWEEN INFLATION AND OTHER VARIABLES
  - ANOVA TEST
  - MULTIVARIATE REGRESSION
- **VISUALIZATIONS:**
  - LINE PLOTS TO SHOW INFLATION RATES OVER TIME FOR EACH COUNTRY.

```
datas.head()
```

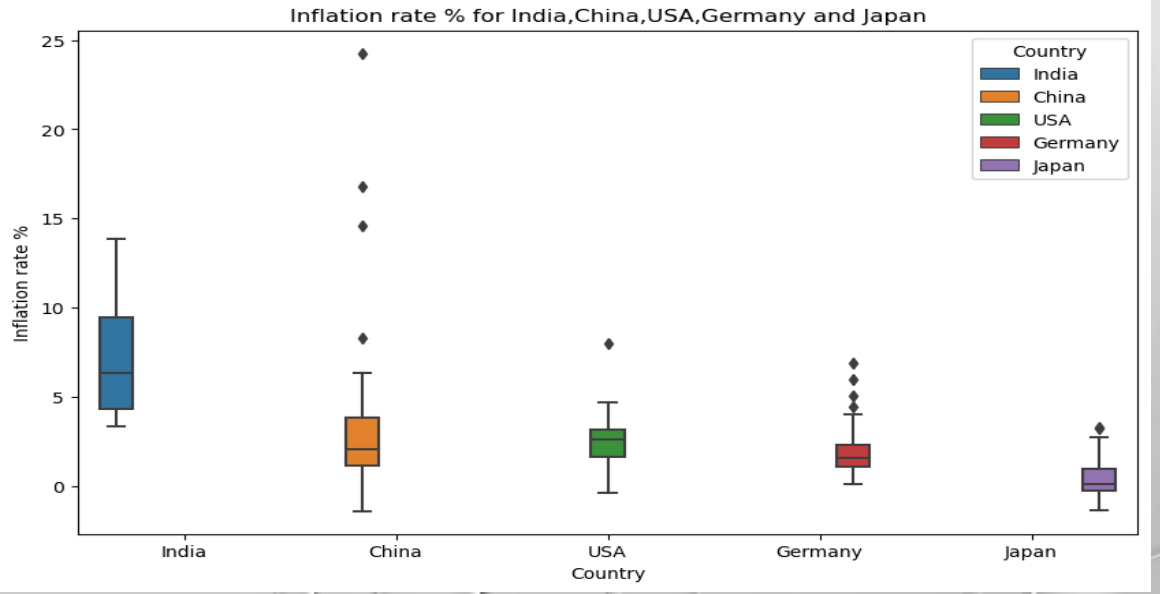
	Country	Year	GDP	GDP_growth	GDP_per_capita	GDP_per_capita_growth	Unemployment	Inflation
0	India	1991	270105341879	1.06	303.85	-1.05	6.85	13.87
1	India	1992	288208070278	5.48	317.56	3.32	6.85	11.79
2	India	1993	279295648983	4.75	301.50	2.63	6.86	6.33
3	India	1994	327274843459	6.66	346.23	4.53	6.83	10.25
4	India	1995	360281909643	7.57	373.63	5.45	6.99	10.22

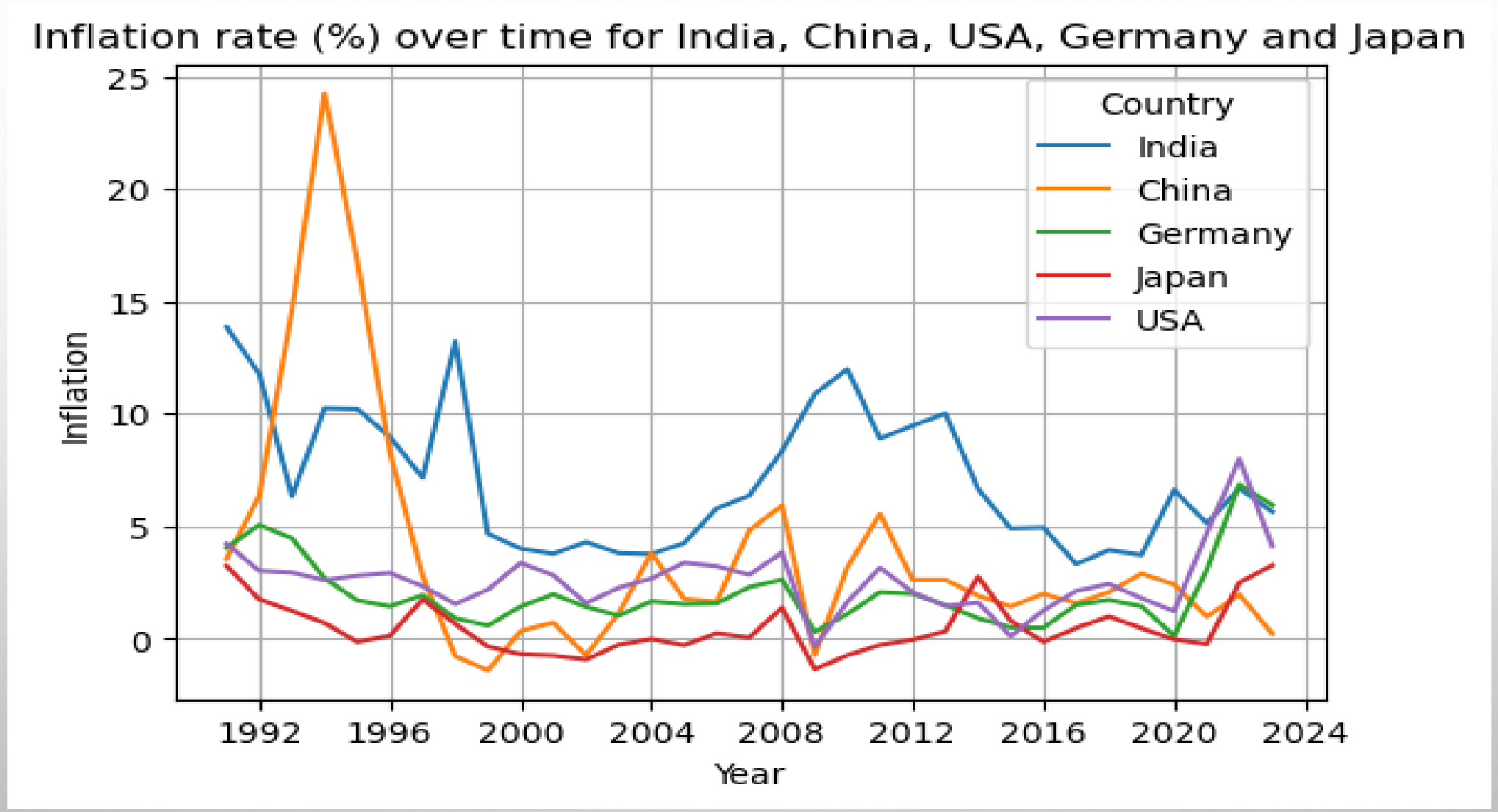
```
datas['Year']=datas['Year'].astype(str)
datas.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 165 entries, 0 to 164
Data columns (total 8 columns):
 #   Column                 Non-Null Count  Dtype  
---  ------                 --------------  -----  
 0   Country                165 non-null    object 
 1   Year                   165 non-null    object 
 2   GDP                    165 non-null    int64  
 3   GDP_growth             165 non-null    float64
 4   GDP_per_capita         165 non-null    float64
 5   GDP_per_capita_growth  165 non-null    float64
 6   Unemployment           165 non-null    float64
 7   Inflation              165 non-null    float64
dtypes: float64(5), int64(1), object(2)
memory usage: 10.4+ KB
```

```
datas.describe()
```

	GDP	GDP_growth	GDP_per_capita	GDP_per_capita_growth	Unemployment	Inflation
count	1.650000e+02	165.000000	165.000000	165.000000	165.000000	165.000000
mean	5.979318e+12	3.935030	25592.358424	3.250848	5.576303	3.208788
std	5.854453e+12	3.880645	20779.616905	3.633568	2.140721	3.676989
min	2.701053e+11	-5.780000	301.500000	-6.670000	2.060000	-1.400000
25%	2.071324e+12	1.490000	1915.550000	1.150000	4.000000	0.990000
50%	4.082469e+12	2.970000	29967.710000	2.390000	4.980000	2.190000
75%	7.551546e+12	6.990000	41103.260000	6.090000	7.650000	4.120000
max	2.736094e+13	14.230000	81695.190000	13.640000	11.170000	24.260000




Inflation rate (%) over time for India, China, USA, Germany and Japan
Inflation
Year
Country
India
China
Germany
Japan
USA
25
20
15
10
5
0
1992
1996
2000
2004
2008
2012
2016
2020
2024

INFLATION TREND VISUALIZATION

- ***INDIA'S*** INFLATION RATE FLUCTUATED SIGNIFICANTLY, WITH A PEAK BETWEEN 1991-1995, REACHING AROUND 15%.
- AFTER 2008, AGAIN INDIA SAW A SIGNIFICANT RISE IN INFLATION WITH SOME LARGE SPIKES, PEAKING AT AROUND 10% BEFORE STABILIZING.
- THE ***USA*** SHOWS A RELATIVELY STABLE INFLATION RATE, MOSTLY RANGING BETWEEN 0% TO 5%.
- POST-2020, THERE IS A VISIBLE INCREASE, POSSIBLY DUE TO ECONOMIC SHOCKS FROM THE PANDEMIC AND OTHER GLOBAL EVENTS, PUSHING THE RATE UP TO 6-7%.

- ***GERMANY*** HAS CONSISTENTLY MAINTAINED A LOW AND STABLE INFLATION RATE, TYPICALLY UNDER 5%.
- A NOTICEABLE SLIGHT RISE OCCURRED POST-2020, LIKELY DUE TO THE PANDEMIC, REACHING 6-7%.
- ***JAPAN*** HAS CONSISTENTLY MAINTAINED VERY LOW INFLATION RATES, OFTEN NEAR OR BELOW 0%, INDICATING DEFLATION IN SOME YEARS.
- A SMALL RISE IS OBSERVED POST-2020, BUT IT REMAINS ONE OF THE LOWEST INFLATION RATES COMPARED TO THE OTHER COUNTRIES.
- ***CHINA'S*** INFLATION RATE WAS NOTABLY HIGH IN THE EARLY 1990S, PEAKING AT AROUND 25% BETWEEN 1993-1994.
- AFTER 1995, IT SHARPLY DROPPED, REMAINING RELATIVELY LOW (<5%) FOR MOST OF THE PERIOD.

KEY OBSERVATIONS

- COUNTRIES WITH **HIGHER** INFLATION:
  - E.G., INDIA AND CHINA MIGHT HAVE HAD HIGHER INFLATION RATES AT CERTAIN POINTS.
- COUNTRIES WITH **LOWER OR MORE STABLE** INFLATION:
  - E.G., GERMANY AND JAPAN MIGHT SHOW STABLE, LOWER INFLATION TRENDS.
- **COMPARATIVE INSIGHTS**:
  - INDIA'S INFLATION MAY BE MORE VOLATILE COMPARED TO THE USA OR GERMANY.

STATISTICAL ANALYSIS

- **CORRELATION ANALYSIS:**
  - CORRELATION BETWEEN INFLATION RATES AND OTHER VARIABLES.
  - FOR EXAMPLE, CORRELATION COEFFICIENTS SHOWING THE RELATIONSHIP BETWEEN INFLATION AND (GDP GROWTH, GDP PER CAPITA)
- **KEY FINDINGS:**
  - STRONG OR WEAK CORRELATIONS BETWEEN INFLATION RATES AND OTHER VARIABLES.
  - INSIGHTS INTO WHETHER GLOBAL EVENTS INFLUENCE INFLATION ACROSS COUNTRIES SIMILARLY.

```
Correlation between Inflation and other variables for China:
GDP                     -0.317132
GDP_growth               0.500853
GDP_per_capita          -0.319977
GDP_per_capita_growth    0.469739
Unemployment            -0.486196
Name: Inflation, dtype: float64
```

```
Correlation between Inflation and other variables for Japan:
GDP                     -0.479318
GDP_growth               0.167636
GDP_per_capita          -0.451394
GDP_per_capita_growth    0.183630
Unemployment            -0.680317
Name: Inflation, dtype: float64
```

```
Correlation between Inflation and other variables for Germany:
GDP                      0.071275
GDP_growth               0.178252
GDP_per_capita           0.056956
GDP_per_capita_growth    0.097216
Unemployment            -0.274571
Name: Inflation, dtype: float64
```

```
Correlation between Inflation and other variables for USA:
GDP                      0.164113
GDP_growth               0.215529
GDP_per_capita           0.170454
GDP_per_capita_growth    0.242643
Unemployment            -0.358626
Name: Inflation, dtype: float64
```

```
Correlation between Inflation and other variables for India:
GDP                     -0.291148
GDP_growth              -0.117789
GDP_per_capita          -0.280437
GDP_per_capita_growth   -0.165859
Unemployment            -0.070655
Name: Inflation, dtype: float64
```

ANNOVA TEST

WE WILL USE ONE-WAY ANOVA TO COMPARE THE MEANS OF INFLATION RATES BETWEEN DIFFERENT COUNTRIES (CHINA, INDIA, USA, GERMANY AND JAPAN) TO SEE IF THERE'S A SIGNIFICANT DIFFERENCE IN INFLATION RATES AMONG THESE COUNTRIES.

H0: THERE IS NO STATISTICALLY SIGNIFICANT DIFFERENCE IN MEANS OF INFLATION RATES AMONG THE COUNTRIES.

H1: AT LEAST ONE COUNTRY HAS STATISTICALLY SIGNIFICANT DIFFERENCE IN MEANS OF INFLATION RATES.

FROM THE P-VALUE OF THE ABOVE HYPOTHESIS TESTING, WE'RE ABLE TO REJECT THE NULL HYPOTHESIS AT 0.05 SIGNIFICANCE LEVEL AND CONCLUDE THAT THERE IS SOME STATISTICALLY SIGNIFICANT DIFFERENCE IN MEANS OF INFLATION RATES AMONG THE COUNTRIES

```
from scipy.stats import f_oneway
India_inflation = datas[datas['Country']=='India']['Inflation']
China_inflation = datas[datas['Country']=='China']['Inflation']
Germany_inflation = datas[datas['Country']=='Germany']['Inflation']
Japan_inflation = datas[datas['Country']=='Japan']['Inflation']
USA_inflation = datas[datas['Country']=='USA']['Inflation']

f_statistics, p_value = f_oneway(India_inflation,China_inflation,Germany_inflation,Japan_inflation,USA_inflation)
print("F-statistics:", f_statistics)
print("P_value:", p_value)
```

```
F-statistics: 22.72350230786072
P_value: 7.034646464221198e-15
```

MULTIPLE LINEAR REGRESSION

Regression for China:

OLS Regression Results

Dep. Variable:	Inflation	R-squared:	0.447
Model:	OLS	Adj. R-squared:	0.345
Method:	Least Squares	F-statistic:	4.372
Date:	Thu, 26 Sep 2024	Prob (F-statistic):	0.00481
Time:	12:17:10	Log-Likelihood:	15.186
No. Observations:	33	AIC:	-18.37
Df Residuals:	27	BIC:	-9.394
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.5997	0.357	-1.682	0.104	-1.331	0.132
GDP	13.1324	11.939	1.100	0.281	-11.364	37.629
GDP_growth	9.9107	5.327	1.860	0.074	-1.020	20.841
GDP_per_capita	-53.2489	50.738	-1.049	0.303	-157.355	50.857
GDP_per_capita_growth	-9.1520	5.364	-1.706	0.099	-20.157	1.853
Unemployment	0.4501	1.120	0.402	0.691	-1.847	2.748

Omnibus:	15.659	Durbin-Watson:	1.189
Prob(Omnibus):	0.000	Jarque-Bera (JB):	18.258
Skew:	1.348	Prob(JB):	0.000108
Kurtosis:	5.451	Cond. No.	2.63e+03

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 2.63e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Regression for Germany:

OLS Regression Results

Dep. Variable:	Inflation	R-squared:	0.233
Model:	OLS	Adj. R-squared:	0.091
Method:	Least Squares	F-statistic:	1.644
Date:	Thu, 26 Sep 2024	Prob (F-statistic):	0.182
Time:	12:17:10	Log-Likelihood:	49.971
No. Observations:	33	AIC:	-87.94
Df Residuals:	27	BIC:	-78.96
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.2637	0.127	2.077	0.047	0.003	0.524
GDP	10.2854	7.520	1.368	0.183	-5.145	25.715
GDP_growth	0.5316	0.559	0.951	0.350	-0.615	1.678
GDP_per_capita	-2.6202	1.907	-1.374	0.181	-6.532	1.292
GDP_per_capita_growth	-0.4051	0.547	-0.741	0.465	-1.527	0.716
Unemployment	-0.0604	0.062	-0.966	0.343	-0.189	0.068

Omnibus:	2.868	Durbin-Watson:	0.656
Prob(Omnibus):	0.238	Jarque-Bera (JB):	2.184
Skew:	0.630	Prob(JB):	0.335
Kurtosis:	2.976	Cond. No.	1.01e+03

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.01e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Regression for India:

OLS Regression Results

Dep. Variable:	Inflation	R-squared:	0.551
Model:	OLS	Adj. R-squared:	0.468
Method:	Least Squares	F-statistic:	6.623
Date:	Thu, 26 Sep 2024	Prob (F-statistic):	0.000384
Time:	12:17:10	Log-Likelihood:	36.797
No. Observations:	33	AIC:	-61.59
Df Residuals:	27	BIC:	-52.62
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.2290	0.372	3.300	0.003	0.465	1.993
GDP	-69.7229	15.085	-4.622	0.000	-100.674	-38.771
GDP_growth	-0.6142	4.294	-0.143	0.887	-9.425	8.196
GDP_per_capita	298.6876	62.128	4.808	0.000	171.212	426.164
GDP_per_capita_growth	0.3114	4.385	0.071	0.944	-8.685	9.308
Unemployment	-1.1671	0.288	-4.050	0.000	-1.758	-0.576

Omnibus:	3.238	Durbin-Watson:	2.113
Prob(Omnibus):	0.198	Jarque-Bera (JB):	2.346
Skew:	0.651	Prob(JB):	0.309
Kurtosis:	3.115	Cond. No.	5.97e+03

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 5.97e+03. This might indicate that there are
strong multicollinearity or other numerical problems.

Regression for Japan:

OLS Regression Results

Dep. Variable:	Inflation	R-squared:	0.589
Model:	OLS	Adj. R-squared:	0.513
Method:	Least Squares	F-statistic:	7.753
Date:	Thu, 26 Sep 2024	Prob (F-statistic):	0.000125
Time:	12:17:10	Log-Likelihood:	70.149
No. Observations:	33	AIC:	-128.3
Df Residuals:	27	BIC:	-119.3
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.2302	0.102	2.256	0.032	0.021	0.440
GDP	0.8062	6.233	0.129	0.898	-11.983	13.595
GDP_growth	-0.5230	0.543	-0.964	0.344	-1.637	0.591
GDP_per_capita	-0.6304	2.414	-0.261	0.796	-5.584	4.323
GDP_per_capita_growth	0.5893	0.552	1.068	0.295	-0.543	1.721
Unemployment	-0.2393	0.087	-2.746	0.011	-0.418	-0.061

Omnibus:	4.983	Durbin-Watson:	1.594
Prob(Omnibus):	0.083	Jarque-Bera (JB):	3.442
Skew:	0.709	Prob(JB):	0.179
Kurtosis:	3.701	Cond. No.	1.49e+03

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.49e+03. This might indicate that there are
strong multicollinearity or other numerical problems.

Regression for USA:

OLS Regression Results

Dep. Variable:	Inflation	R-squared:	0.157
Model:	OLS	Adj. R-squared:	0.001
Method:	Least Squares	F-statistic:	1.007
Date:	Thu, 26 Sep 2024	Prob (F-statistic):	0.433
Time:	12:17:10	Log-Likelihood:	51.153
No. Observations:	33	AIC:	-90.31
Df Residuals:	27	BIC:	-81.33
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.0158	0.276	0.057	0.955	-0.550	0.582
GDP	-0.9516	1.655	-0.575	0.570	-4.348	2.445
GDP_growth	0.0496	1.961	0.025	0.980	-3.974	4.073
GDP_per_capita	1.1162	1.904	0.586	0.563	-2.791	5.023
GDP_per_capita_growth	0.0201	1.935	0.010	0.992	-3.950	3.990
Unemployment	-0.0767	0.070	-1.088	0.286	-0.221	0.068

Omnibus:	13.401	Durbin-Watson:	1.071
Prob(Omnibus):	0.001	Jarque-Bera (JB):	15.845
Skew:	1.097	Prob(JB):	0.000362
Kurtosis:	5.590	Cond. No.	428.

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

CONCLUSION

- **INFLATION TRENDS:**
  - INDIA AND CHINA SHOW HIGHER INFLATION RATES OVER TIME.
  - USA AND GERMANY HAVE RELATIVELY STABLE INFLATION.
  - JAPAN STAND OUT WITH CONSISTENTLY LOW INFLATION RATES.
- **DIFFERENCES IN INFLATION RATES ACROSS COUNTRIES:**
  - INFLATION BEHAVIOR VARIES ACROSS COUNTRIES DUE TO FACTORS LIKE ECONOMIC POLICIES, GLOBAL TRADE, AND EXTERNAL SHOCKS.
  - GLOBAL EVENTS (E.G., THE 2008 FINANCIAL CRISIS AND THE COVID-19 PANDEMIC) IMPACTED INFLATION DIFFERENTLY ACROSS COUNTRIES, WITH NOTABLE SPIKES POST-2020.
- **CORRELATION:**
  - THE CORRELATION ANALYSIS REVEALED SIGNIFICANT RELATIONSHIPS BETWEEN INFLATION AND OTHER ECONOMIC FACTORS LIKE GDP GROWTH AND GDP PER CAPITA.
  - THESE CORRELATIONS INDICATE HOW INFLATION IS INTERTWINED WITH BROADER ECONOMIC PERFORMANCE, INFLUENCING KEY MACROECONOMIC INDICATORS.
- **REGRESSION:**
  - SIGNIFICANT PREDICTORS: THE MODEL IDENTIFIES SIGNIFICANT FACTORS AFFECTING INFLATION, SUCH AS GDP GROWTH AND UNEMPLOYMENT.

REFERENCES

- **DATA SOURCES:**
  - WORLD BANK, INTERNATIONAL MONETARY FUND (IMF), NATIONAL STATISTICAL AGENCIES (E.G., INDIA'S MINISTRY OF STATISTICS).

THANK YOU