**Exploratory Data Analysis of Monthly Gold Prices 1979 – 2021**

Kendridge Mothapo

10/08/2023

*source:* <https://www.kaggle.com/datasets/odins0n/monthly-gold-prices>

Acknowledgements*:* <https://www.gold.org/goldhub/data/gold-prices>

In the early 1500’s, the economic world was ruled by a school of thought named “The Mercantilist”. This group viewed gold as the most desirable form of wealth. Mercantilists tended to equate the wealth of a nation to the amount of gold and silver bullion it possessed. Trading of gold amongst nations flourished during this period. This was encouraged by the duty-free import on raw materials imposed by countries which followed the tenets of the Mercantilists.

Gold prices have been on the rise over the couple of decades. Countries which are lucky to be mining this rare commodity has seen an increase in the price in which they trade their gold. This is due to the increased demand on this rare material.

We have herein done some exploration on the monthly gold prices from January 1979 to 2021. Data exploration on structured data sets can be done on Python, Excel, SQL amongst others subject to preference and limitations. For the purposes of this data set, I have chosen to do exploration in SQL and visualized some of the results in Power Bi

The following Indicators are which we focused on.

* **Count** (Estimates the number of times a country traded gold during this period)
* **Average** (The average/mean is the SUM of all gold prices of a distinct country divided by its COUNT. This is used as a measure of dispersion.)
* **Standard Deviation** (This reflects how much gold prices are dispersed relative to the mean gold price. In essence, this measures the average amount of variability in our data set)
* **Minimum** (This estimates the minimum price each country has traded their gold during this period)
* **Lower**(q1)**, Median**(q2)**, Upper quartiles**(q3)(*q*1 represents the value that indicates the end of the first 25% of the data values; *q*2 represents the value that indicates the end of the second 25% (or the value which divides the data set into two equal parts, that is the median); and *q*3 is the value indicating the end of the third 25%. The middle 50% of the data lies between *q*1 and *q*3.
* **Maximum** (This estimates the maximum price each country has traded their gold during this period)

The following SQL queries are which were used to achieve the above set of indicators. A table containing results of all the indicators will follow the SQL queries.

--Exploratory Data Analysis of Monthly Gold\_Prices of 18 Countries/Nations over the period 1979-2021

--Retrieves all the columns on the Gold\_Prices table

SELECT \*

FROM Gold\_Prices

ORDER BY Date

--Gets the Data Type and information about the table

EXEC sp\_help Gold\_Prices

--Creates a Temporary Table which will be used for Data Exploration on the Gold\_Prices table

CREATE TABLE #temp\_Gold\_Prices (

Date nvarchar(100),

[United States(USD)] float,

[Europe(EUR)] float,

[Japan(JPY)] float,

[United Kingdom(GBP)] float,

[Canada(CAD)] float,

[Switzerland(CHF)] float,

[India(INR)] float,

[China(CNY)] float,

[Turkey(TRY)] float,

[Saudi Arabia(SAR)] float,

[Indonesia(IDR)] float,

[United Arab Emirates(AED)] float,

[Thailand(THB)] float,

[Vietnam(VND)] float,

[Egypt(EGP)] float,

[South Korean(KRW)] float,

[Australia(AUD)] float,

[South Africa(ZAR)] float

)

INSERT INTO #temp\_Gold\_Prices

SELECT \*

FROM Gold\_Prices

--Replacing all values which were automatically populated with '0' to NULL in the Temporary Table. This will help avoid dividing by Zero

--Only four(4) Countries/Nations had rows with '0' values

UPDATE #temp\_Gold\_Prices

SET [China(CNY)] = NULL

WHERE [China(CNY)] = '0'

UPDATE #temp\_Gold\_Prices

SET [Turkey(TRY)] = NULL

WHERE [Turkey(TRY)] = '0'

UPDATE #temp\_Gold\_Prices

SET [Vietnam(VND)] = NULL

WHERE [Vietnam(VND)] = '0'

UPDATE #temp\_Gold\_Prices

SET [Egypt(EGP)] = NULL

WHERE [Egypt(EGP)] = '0'

--Retrieves the updated Gold\_Prices Temporary table which is now populated with NULL where values equalled 0

SELECT \*

FROM #temp\_Gold\_Prices

--Count of Monthly Gold Prices for each Country/Nation

SELECT COUNT([United States(USD)]) as 'USD', COUNT([Europe(EUR)]) as 'EUR', COUNT([Japan(JPY)]) as 'JPY', COUNT([United Kingdom(GBP)]) as 'GBP',

COUNT([Canada(CAD)]) as 'CAD', COUNT([Switzerland(CHF)]) as 'CHF', COUNT([India(INR)]) as 'INR', COUNT([China(CNY)]) as 'CNY', COUNT([Turkey(TRY)]) as 'TRY',

COUNT([Saudi Arabia(SAR)]) as 'SAR', COUNT([Indonesia(IDR)]) as 'IDR', COUNT([United Arab Emirates(AED)]) as 'AED', COUNT([Thailand(THB)]) as 'THB',

COUNT([Vietnam(VND)]) as 'VND', COUNT([Egypt(EGP)]) as 'EGP', COUNT([South Korean(KRW)]) as 'KRW', COUNT([Australia(AUD)]) as 'AUD', COUNT([South Africa(ZAR)]) as 'ZAR'

FROM #temp\_Gold\_Prices

--Monthly Mean/Average Price for each Country/Nation over the period 1979-2021

SELECT AVG([United States(USD)]) as 'USD', AVG([Europe(EUR)]) as 'EUR', AVG([Japan(JPY)]) as 'JPY',

AVG([United Kingdom(GBP)]) as 'GBP', AVG([Canada(CAD)]) as 'CAD', AVG([Switzerland(CHF)]) as 'CHF',

AVG([India(INR)]) as 'INR', AVG([China(CNY)]) as 'CNY', AVG([Turkey(TRY)]) as 'TRY',

AVG([Saudi Arabia(SAR)]) as 'SAR', AVG([Indonesia(IDR)]) as 'IDR', AVG([United Arab Emirates(AED)]) as 'AED',

AVG([Thailand(THB)]) as 'THB', AVG([Vietnam(VND)]) as 'VND', AVG([Egypt(EGP)]) as 'EGP', AVG([South Korean(KRW)]) as 'KRW',

AVG([Australia(AUD)]) as 'AUD', AVG([South Africa(ZAR)]) as 'ZAR'

FROM #temp\_Gold\_Prices

--The Standard Deviation for the Monthly\_Gold\_Prices over the period 1979-2021

SELECT STDEV([United States(USD)]) as 'USD', STDEV([Europe(EUR)]) as 'EUR', STDEV([Japan(JPY)]) as 'JPY',

STDEV([United Kingdom(GBP)]) as 'GBP', STDEV([Canada(CAD)]) as 'CAD', STDEV([Switzerland(CHF)]) as 'CHF',

STDEV([India(INR)]) as 'INR', STDEV([China(CNY)]) as 'CNY', STDEV([Turkey(TRY)]) as 'TRY',

STDEV([Saudi Arabia(SAR)]) as 'SAR', STDEV([Indonesia(IDR)]) as 'IDR', STDEV([United Arab Emirates(AED)]) as 'AED',

STDEV([Thailand(THB)]) as 'THB', STDEV([Vietnam(VND)]) as 'VND', STDEV([Egypt(EGP)]) as 'EGP', STDEV([South Korean(KRW)]) as 'KRW',

STDEV([Australia(AUD)]) as 'AUD', STDEV([South Africa(ZAR)]) as 'ZAR'

FROM #temp\_Gold\_Prices

--The Minimum Gold Price for each Country/Nation over the Period 1979-2021

SELECT MIN([United States(USD)]) as 'USD', MIN([Europe(EUR)]) as 'EUR', MIN([Japan(JPY)]) as 'JPY',

MIN([United Kingdom(GBP)]) as 'GBP', MIN([Canada(CAD)]) as 'CAD', MIN([Switzerland(CHF)]) as 'CHF',

MIN([India(INR)]) as 'INR', MIN([China(CNY)]) as 'CNY', MIN([Turkey(TRY)]) as 'TRY',

MIN([Saudi Arabia(SAR)]) as 'SAR', MIN([Indonesia(IDR)]) as 'IDR', MIN([United Arab Emirates(AED)]) as 'AED',

MIN([Thailand(THB)]) as 'THB', MIN([Vietnam(VND)]) as 'VND', MIN([Egypt(EGP)]) as 'EGP', MIN([South Korean(KRW)]) as 'KRW',

MIN([Australia(AUD)]) as 'AUD', MIN([South Africa(ZAR)]) as 'ZAR'

FROM #temp\_Gold\_Prices

--The LOWER QUARTILE(25%) Gold\_Price of each Country/Nation from 1979-2021

SELECT TOP 1

percentile\_cont(0.75) within group(order by [United States(USD)] desc) over( ) as 'USA 25%',

percentile\_cont(0.75) within group(order by [Europe(EUR)] desc) over( ) as 'EUR 25%',

percentile\_cont(0.75) within group(order by [Japan(JPY)] desc) over( ) as 'JPY 25%',

percentile\_cont(0.75) within group(order by [United Kingdom(GBP)] desc) over( ) as 'GBP 25%',

percentile\_cont(0.75) within group(order by [Canada(CAD)] desc) over( ) as 'CAD 25%',

percentile\_cont(0.75) within group(order by [Switzerland(CHF)] desc) over( ) as 'CHF 25%',

percentile\_cont(0.75) within group(order by [India(INR)] desc) over( ) as 'INR 25%',

percentile\_cont(0.75) within group(order by[China(CNY)] desc) over( ) as 'CNY 25%',

percentile\_cont(0.75) within group(order by [Turkey(TRY)] desc) over( ) as 'TRY 25%',

percentile\_cont(0.75) within group(order by [Saudi Arabia(SAR)] desc) over( ) as 'SAR 25%',

percentile\_cont(0.75) within group(order by [Indonesia(IDR)] desc) over( ) as 'IDR 25%',

percentile\_cont(0.75) within group(order by [United Arab Emirates(AED)] desc) over( ) as 'AED 25%',

percentile\_cont(0.75) within group(order by [Thailand(THB)] desc) over( ) as 'THB 25%',

percentile\_cont(0.75) within group(order by [Vietnam(VND)] desc) over( ) as 'VND 25%',

percentile\_cont(0.75) within group(order by [Egypt(EGP)] desc) over( ) as 'EGP 25%',

percentile\_cont(0.75) within group(order by [South Korean(KRW)] desc) over( ) as 'KRW 25%',

percentile\_cont(0.75) within group(order by [Australia(AUD)] desc) over( ) as 'AUD 25%',

percentile\_cont(0.75) within group(order by [South Africa(ZAR)] desc) over( ) as 'ZAR 25%'

FROM #temp\_Gold\_Prices

--The MEDIAN(50%) Gold\_Price of each Country/Nation from 1979-2021

SELECT TOP 1

percentile\_cont(0.5) within group(order by [United States(USD)] desc) over( ) as 'USA 50%',

percentile\_cont(0.5) within group(order by [Europe(EUR)] desc) over( ) as 'EUR 50%',

percentile\_cont(0.5) within group(order by [Japan(JPY)] desc) over( ) as 'JPY 50%',

percentile\_cont(0.5) within group(order by [United Kingdom(GBP)] desc) over( ) as 'GBP 50%',

percentile\_cont(0.5) within group(order by [Canada(CAD)] desc) over( ) as 'CAD 50%',

percentile\_cont(0.5) within group(order by [Switzerland(CHF)] desc) over( ) as 'CHF 50%',

percentile\_cont(0.5) within group(order by [India(INR)] desc) over( ) as 'INR 50%',

percentile\_cont(0.5) within group(order by [China(CNY)] desc) over( ) as 'CNY 50%',

percentile\_cont(0.5) within group(order by [Turkey(TRY)] desc) over( ) as 'TRY 50%',

percentile\_cont(0.5) within group(order by [Saudi Arabia(SAR)] desc) over( ) as 'SAR 50%',

percentile\_cont(0.5) within group(order by [Indonesia(IDR)] desc) over( ) as 'IDR 50%',

percentile\_cont(0.5) within group(order by [United Arab Emirates(AED)] desc) over( ) as 'AED 50%',

percentile\_cont(0.5) within group(order by [Thailand(THB)] desc) over( ) as 'THB 50%',

percentile\_cont(0.5) within group(order by [Vietnam(VND)] desc) over( ) as 'VND 50%',

percentile\_cont(0.5) within group(order by [Egypt(EGP)] desc) over( ) as 'EGP 50%',

percentile\_cont(0.5) within group(order by [South Korean(KRW)] desc) over( ) as 'KRW 50%',

percentile\_cont(0.5) within group(order by [Australia(AUD)] desc) over( ) as 'AUD 50%',

percentile\_cont(0.5) within group(order by [South Africa(ZAR)] desc) over( ) as 'ZAR 50%'

FROM #temp\_Gold\_Prices

--The UPPER QUARTILE(75%) Gold\_Price of each Country/Nation from 1979-2021

SELECT TOP 1

percentile\_cont(0.25) within group(order by [United States(USD)] desc) over( ) as 'USA 75%',

percentile\_cont(0.25) within group(order by [Europe(EUR)] desc) over( ) as 'EUR 75%',

percentile\_cont(0.25) within group(order by [Japan(JPY)] desc) over( ) as 'JPY 75%',

percentile\_cont(0.25) within group(order by [United Kingdom(GBP)] desc) over( ) as 'GBP 75%',

percentile\_cont(0.25) within group(order by [Canada(CAD)] desc) over( ) as 'CAD 75%',

percentile\_cont(0.25) within group(order by [Switzerland(CHF)] desc) over( ) as 'CHF 75%',

percentile\_cont(0.25) within group(order by [India(INR)] desc) over( ) as 'INR 75%',

percentile\_cont(0.25) within group(order by[China(CNY)] desc) over( ) as 'CNY 75%',

percentile\_cont(0.25) within group(order by [Turkey(TRY)] desc) over( ) as 'TRY 75%',

percentile\_cont(0.25) within group(order by [Saudi Arabia(SAR)] desc) over( ) as 'SAR 75%',

percentile\_cont(0.25) within group(order by [Indonesia(IDR)] desc) over( ) as 'IDR 75%',

percentile\_cont(0.25) within group(order by [United Arab Emirates(AED)] desc) over( ) as 'AED 75%',

percentile\_cont(0.25) within group(order by [Thailand(THB)] desc) over( ) as 'THB 75%',

percentile\_cont(0.25) within group(order by [Vietnam(VND)] desc) over( ) as 'VND 75%',

percentile\_cont(0.25) within group(order by [Egypt(EGP)] desc) over( ) as 'EGP 75%',

percentile\_cont(0.25) within group(order by [South Korean(KRW)] desc) over( ) as 'KRW 75%',

percentile\_cont(0.25) within group(order by [Australia(AUD)] desc) over( ) as 'AUD 75%',

percentile\_cont(0.25) within group(order by [South Africa(ZAR)] desc) over( ) as 'ZAR 75%'

FROM #temp\_Gold\_Prices

--The Maximum/Outlier Gold\_Price for each Country/Nation over the Period 1979-2021

SELECT MAX([United States(USD)]) as 'USD', MAX([Europe(EUR)]) as 'EUR', MAX([Japan(JPY)]) as 'JPY',

MAX([United Kingdom(GBP)]) as 'GBP', MAX([Canada(CAD)]) as 'CAD', MAX([Switzerland(CHF)]) as 'CHF',

MAX([India(INR)]) as 'INR', MAX([China(CNY)]) as 'CNY', MAX([Turkey(TRY)]) as 'TRY',

MAX([Saudi Arabia(SAR)]) as 'SAR', MAX([Indonesia(IDR)]) as 'IDR', MAX([United Arab Emirates(AED)]) as 'AED',

MAX([Thailand(THB)]) as 'THB', MAX([Vietnam(VND)]) as 'VND', MAX([Egypt(EGP)]) as 'EGP', MAX([South Korean(KRW)]) as 'KRW',

MAX([Australia(AUD)]) as 'AUD', MAX([South Africa(ZAR)]) as 'ZAR'

FROM #temp\_Gold\_Prices

DROP TABLE #temp\_Gold\_Prices

The following Tables contains the results of the set of indicators from the above SQL queries.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **United States(USD)** | **Europe**  **(EUR)** | **Japan**  **(JPY)** | **United Kingdom(GBP)** | **Canada**  **(CAD)** | **Switzerland**  **(CHF)** |
| **Count** | 511 | 511 | 511 | 511 | 511 | 511 |
| **Mean** | 693.7039 | 572.9094 | 82947.7855 | 450.1446 | 843.5041 | 826.1233 |
| **Std deviation** | 473.9370 | 381.4214 | 44427.4261 | 346.7455 | 546.9960 | 365.5337 |
| **Min** | 233.7 | 144.8 | 27937.5 | 116.2 | 267.1 | 379.3 |
| **25%** | 354.2 | 306.95 | 43095.75 | 218.4 | 454.3 | 501.65 |
| **50%** | 415.1 | 364.1 | 70455.9 | 256.4 | 532.3 | 698.2 |
| **75%** | 1166.65 | 929.25 | 124847.45 | 742.15 | 1326.95 | 1126.95 |
| **Max** | 1964.9 | 1661.7 | 207845 | 1497.1 | 2632 | 1787.5 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **India**  **(INR)** | **China**  **(CNY)** | **Turkey**  **(TRY)** | **Saudi Arabia**  **(SAR)** | **Indonesia**  **(IDR)** | **United Arab Emirates(AED)** |
| **Count** | 511 | 439 | 464 | 511 | 511 | 511 |
| **Mean** | 33124.2863 | 4883.2845 | 1739.8759 | 2580.5323 | 6207426.6022 | 2549.2014 |
| **Std deviation** | 36068.2221 | 3298.6485 | 3058.6491 | 1789.3698 | 7326262.3223 | 1740.3986 |
| **Min** | 1840.8 | 821.1 | 0.1 | 745.2 | 139739.3 | 856.1 |
| **25%** | 6273.75 | 2171.35 | 2.475 | 1311.35 | 697973.6 | 1297.55 |
| **50%** | 13027.2 | 3287.8 | 446.05 | 1502.6 | 2536802.7 | 1528.9 |
| **75%** | 61884.75 | 7991.05 | 2595.025 | 4375.95 | 12021637.05 | 4285 |
| **Max** | 146999.1 | 13714 | 16105.4 | 7369.2 | 28687538.6 | 7217.4 |

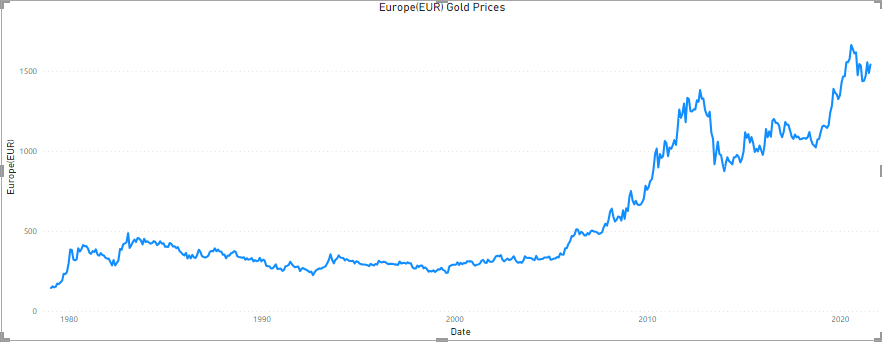
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Thailand**  **(THB)** | **Vietnam**  **(VND)** | **Egypt**  **(EGP)** | **South Korean**  **(KRW)** | **Australia**  **(AUD)** | **South Africa**  **(ZAR)** |
| **Count** | 511 | 380 | 392 | 511 | 511 | 511 |
| **Mean** | 21659.5327 | 15126332.9889 | 6759.5776 | 717691.5170 | 869.8753 | 5950.7951 |
| **Std deviation** | 15607.6553 | 12765711.5358 | 8061.7218 | 579586.1635 | 573.9375 | 7352.3626 |
| **Min** | 4511.3 | 1571727.9 | 870.1 | 108235.8 | 197.2 | 194.5 |
| **25%** | 9612.4 | 4090357.325 | 1176.125 | 301251.75 | 474.55 | 972.5 |
| **50%** | 12140.5 | 7502083.8 | 2578.9 | 374979.6 | 554.1 | 1864.7 |
| **75%** | 38774.35 | 27547811.1 | 9149.475 | 1304939.05 | 1372.6 | 9549.35 |
| **Max** | 61270.5 | 45539502.6 | 31379.5 | 2340981.6 | 2739.9 | 33459.8 |

The following visuals from PowerBI represent gold prices of different countries in this period (1979-2021).

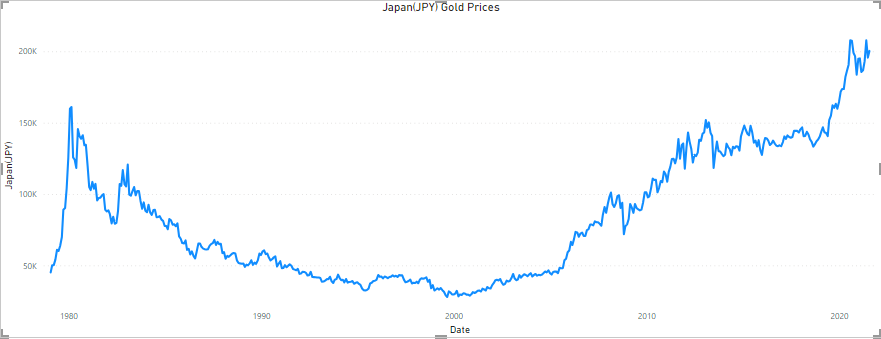
**United States (USD)**



**Europe (EUR)**

****

**Japan (JPY)**

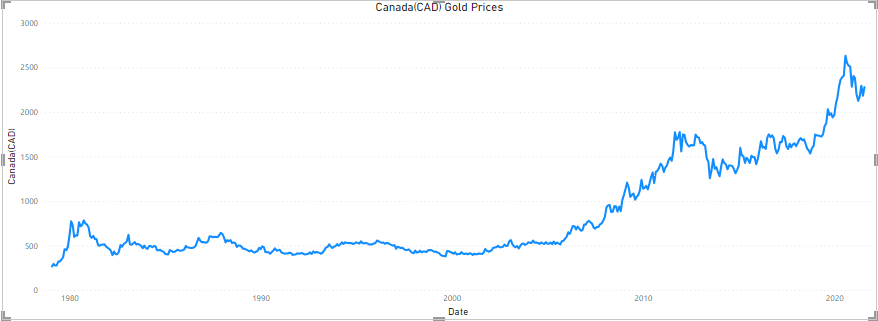
****

**United Kingdom (GBP)**

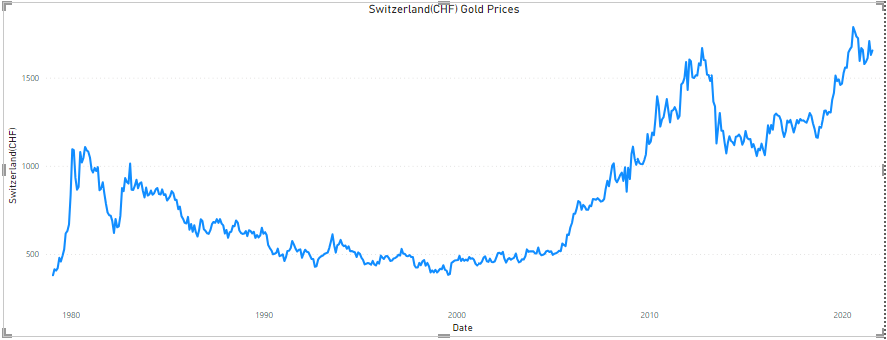
**A graph showing a line

Description automatically generated**

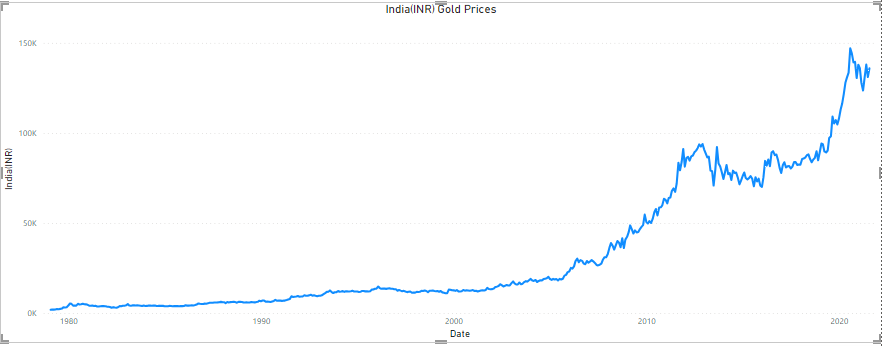
**Canada (CAD)**

****

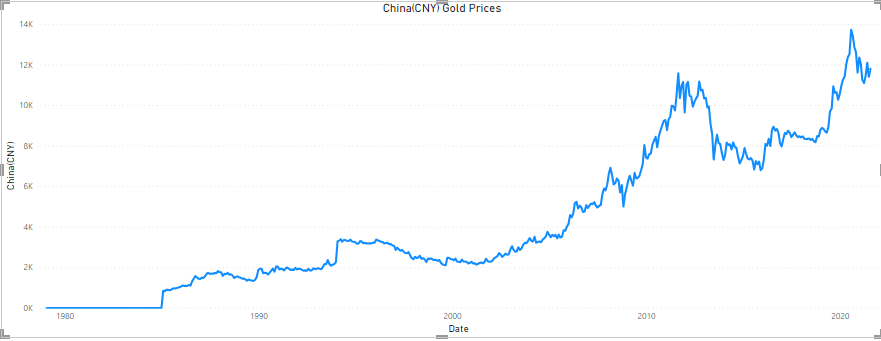
**Switzerland (CHF)**

****

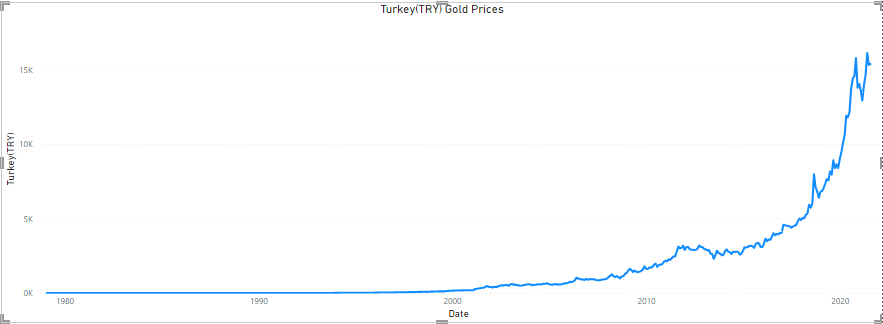
**India (INR)**

****

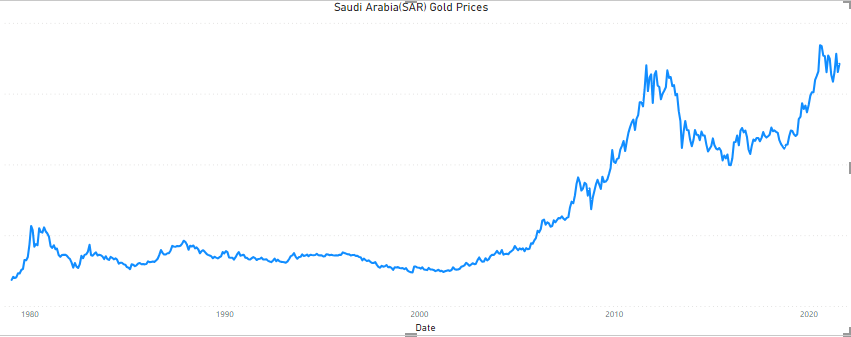
**China (CNY)**

****

**Turkey (TRY)**

****

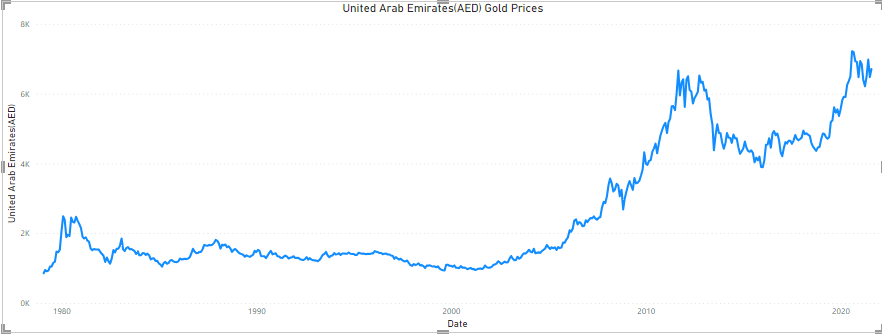
**Saudi Arabia (SAR)**

****

**Indonesia (IDR)**

****

**United Arab Emirates (AED)**

****

**Thailand (THB)**

**A graph showing a line

Description automatically generated**

**Vietnam (VND)**

**A graph showing a line

Description automatically generated**

**Egypt (EGP)**

**A graph showing a line

Description automatically generated**

**South Korean (KRW)**

**A graph showing a line

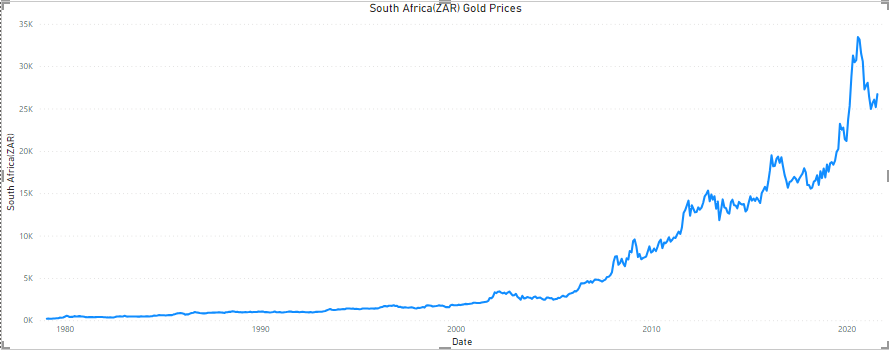
Description automatically generated with medium confidence**

**Australia (AUD)**

**A graph showing a line

Description automatically generated with medium confidence**

**South Africa (ZAR)**

****

**Special Focus: South Africa (ZAR) Gold Prices**

From the above visual, we can see an upward curve in the gold prices of South Africa from the year 2000 until 2020.