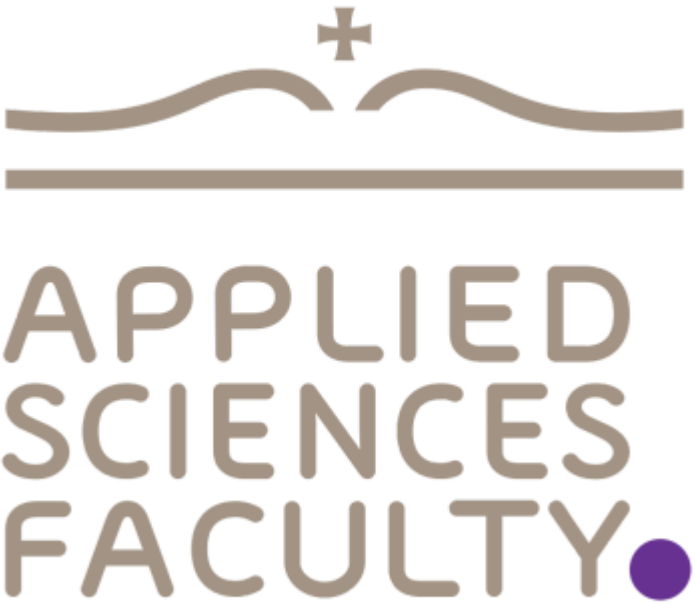


Final Project

The Dream of Alchemists and Scrooge McDuck: Revealing the Secrets of Gold

By Nataliia Skutar

Introduction

Topic and motivation

Gold, the gleaming champion of precious metals, has fascinated people across time. Its brilliance and allure are unmatched, captivating hearts with its radiant beauty. From delicate adornments to shimmering treasures, gold embodies elegance and luxury.

Beyond its captivating appearance, gold holds a deeper significance. It represents wealth, prosperity, and stability. In times of uncertainty, it stands as a steadfast refuge, offering a sense of security and assurance. A timeless symbol of value, gold has stood the test of time, enduring as a precious asset. This precious metal symbolises wealth and growth, acting as a hedge against inflation and market fluctuations. Investors seek its golden embrace, finding solace in its enduring worth. For centuries, gold has been a trusted refuge, a tangible representation of financial security.

Gold is a core of the world's financials and economy. Its prices impact every aspect of our everyday lives. Knowing and understanding the secrets of the gold market is our key to success. Therefore, I decided to dig into its depths.

The aim and the tasks

The main goal of my research is to gather, analyse, and predict gold prices. I want to understand what influences the prices of gold, what general trends or tendencies they have, what events have an impact on them and how the average person should deal with all that information in the proper way to reach success.

The tasks I set myself are to gather as much information on gold prices as I can, research the world of the golden market and its news, analyse all that data with appropriate methods, make some meaningful conclusions and then try to predict the gold prices.

Hypotheses and objectives description

I had a few assumptions, based on the general understanding of the gold significance and simple live observations before I started my research:

- Gold prices are steadily rising nowadays
- Gold prices should have some pattern
- Worlds news affect gold prices a lot
- The war in Ukraine particularly has an impact on gold prices
- It should not be extremely hard to predict gold prices if I have enough data and assuming no shocking events in the nearest future.

Therefore testing these points became my research objectives.

Literature review

These are the sources that I used or just useful if you want to understand the general topic better:

- One of the very few websites where you can download the gold price data, and not only look at general plots and today's numbers
- Same thing, but for money exchange rates, which I figured I needed during the work process
- A website with news and almost weekly updates and recaps on the gold market
- Cool article for general understanding
- And one more
- To understand the war dummy variable that will occur here later
- A good website that I found, but didn't use in my research
- And one more
- Finally, my GitHub

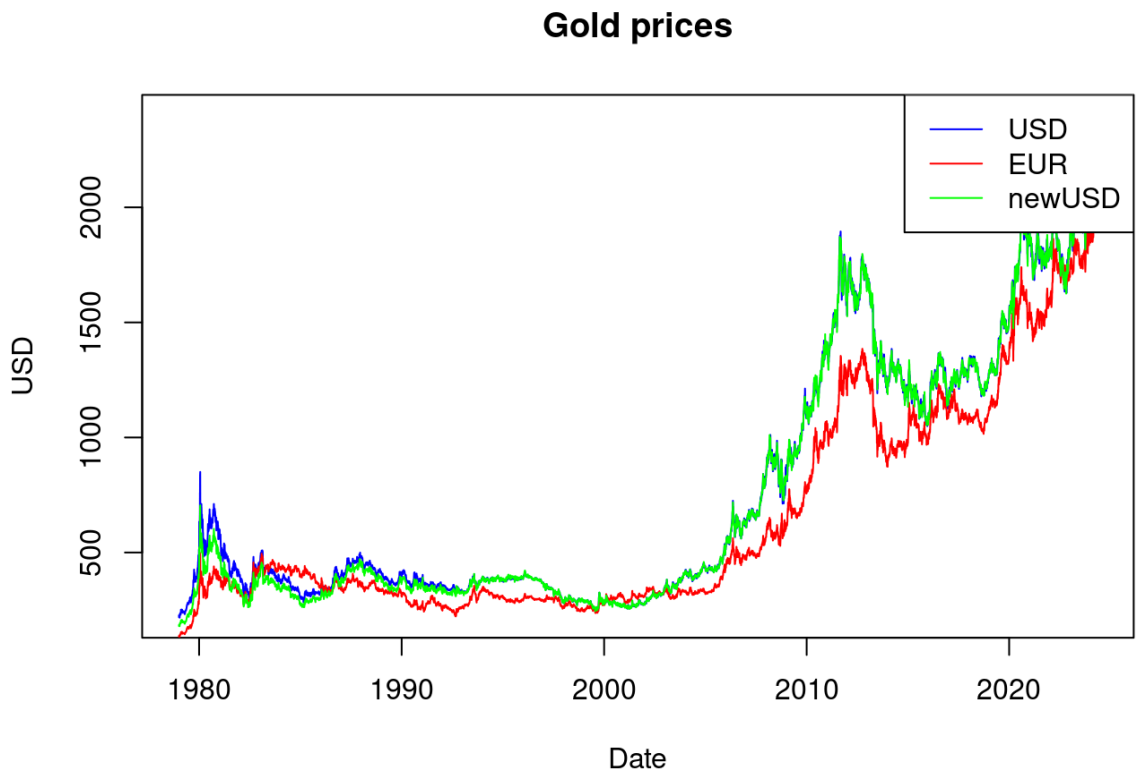
Data description

I wanted to use as much data as I could and for it to be collected as often as possible. The oldest data I found is from 45 years ago, starting from 01.01.1979 and up to last week

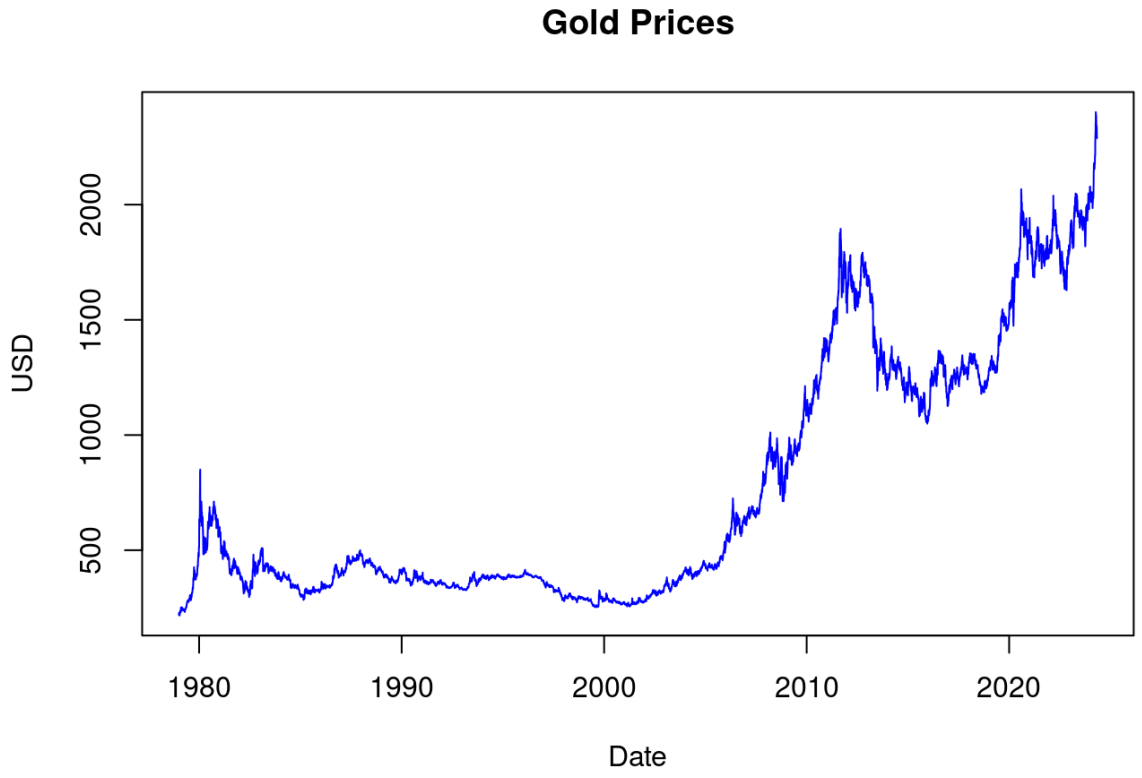
03.05.24. The data was gathered every working day because the gold commodity exchange is closed on the weekends, and no trades are happening. So, the data I started with looked like this:

	USD	EUR	JPY	GBP	CAD	CHF	INR	CNY	TRY	SAR	IDR	AED	THB	VND	EGP	KRW	RUB	ZAR	AUD
12/29/1978	226.0	137.1	#N/A	110.7	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
1/1/1979	226.0	137.1	#N/A	110.7	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
1/2/1979	226.8	137.3	43,164.9	111.5	263.7	359.6	1,792.9	#N/A	#N/A	735.6	138,160.5	851.3	4,454.6	#N/A	#N/A	107,470.0	#N/A	194.6	193.6
1/3/1979	218.6	134.0	43,717.9	108.0	264.4	365.9	1,802.2	#N/A	#N/A	739.4	138,877.0	855.7	4,477.7	#N/A	#N/A	108,027.4	#N/A	192.2	193.0
1/4/1979	223.2	136.8	43,674.9	110.7	264.1	366.4	1,811.7	#N/A	#N/A	743.4	139,616.3	860.3	4,501.5	#N/A	#N/A	108,602.5	#N/A	193.0	194.6
1/5/1979	225.5	138.4	44,582.5	111.6	268.4	373.7	1,843.6	#N/A	#N/A	756.4	142,069.1	875.4	4,580.6	#N/A	#N/A	110,510.4	#N/A	196.4	197.9
1/8/1979	223.1	136.4	44,436.2	110.2	269.5	372.6	1,841.3	#N/A	#N/A	750.9	141,870.9	869.7	4,574.2	#N/A	#N/A	110,356.3	#N/A	197.2	197.9
1/9/1979	224.0	137.3	44,045.6	111.0	266.3	372.8	1,822.5	#N/A	#N/A	743.3	140,446.7	860.9	4,528.3	#N/A	#N/A	109,248.4	#N/A	195.0	195.9
1/10/1979	220.7	135.5	43,366.4	109.9	261.8	365.9	1,803.5	#N/A	#N/A	735.6	138,981.0	852.0	4,481.0	#N/A	#N/A	108,108.3	#N/A	192.0	192.7
1/11/1979	220.7	135.9	43,770.6	110.2	262.5	370.9	1,814.6	#N/A	#N/A	740.1	139,833.8	857.2	4,508.5	#N/A	#N/A	108,771.7	#N/A	192.4	193.8
1/12/1979	217.6	134.1	42,837.1	109.0	257.2	364.6	1,782.6	#N/A	#N/A	727.1	137,371.8	847.1	4,429.1	#N/A	#N/A	106,856.6	#N/A	187.1	189.8

There was the column Date and many other columns with prices in different currencies. At first, I wanted to choose a few of the most essential currencies and derive different models for them. Then I realised that the exchange rate depends a lot on gold prices, so probably in different countries the prices on gold commodity exchange should be similar. To check, I chose two main currencies, the American dollar (USD) and the Euro (EUR), and found data on exchange rates of USD/EUR from 1979. It was gathered weekly, not daily so I extended it for further usage. Using that information and some math I combined the datasets and created a new column newUSD which shows the gold prices in USD derived from EUR and exchange rates.

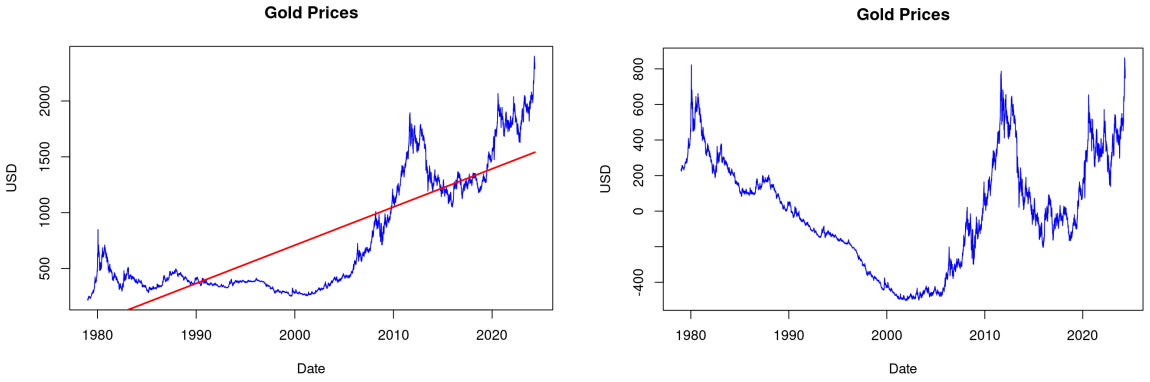


As you can see from the plot, I was right, so there was no sense in considering more than one currency. Hence, I ended up with just one column USD in my data and 11830 rows. Here is a clearer plot of data:



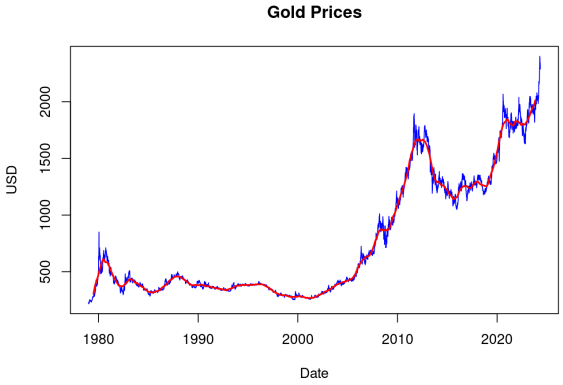
Data analysis

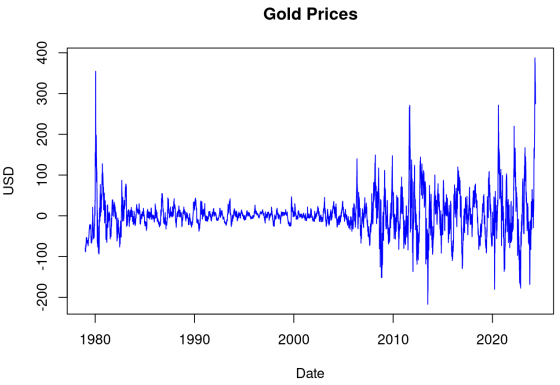
At first, I wanted to see if there was a clear trendline as I expected. Thats what I've got:



Then I tried to remove the trend but found out that the detrended data didn't look good.

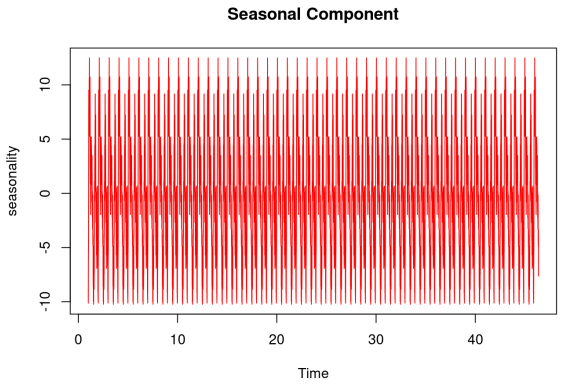
So I discarded my first assumption that gold prices have stable growth, but there could still be a non-linear trend.

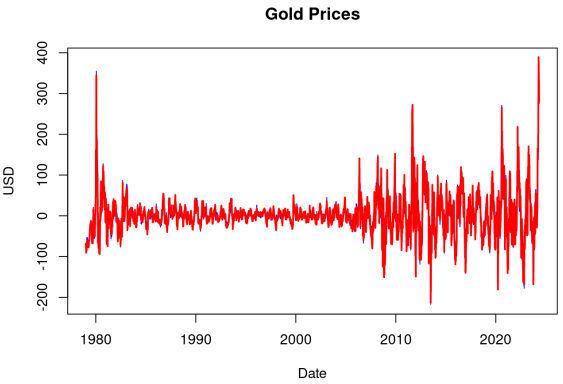




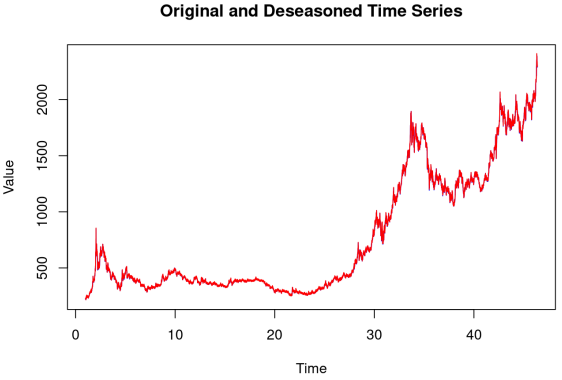
This detrending worked much better, and the data seems pretty stationary now.

Then I tried to find a seasonal component, but it seemed to be very small:





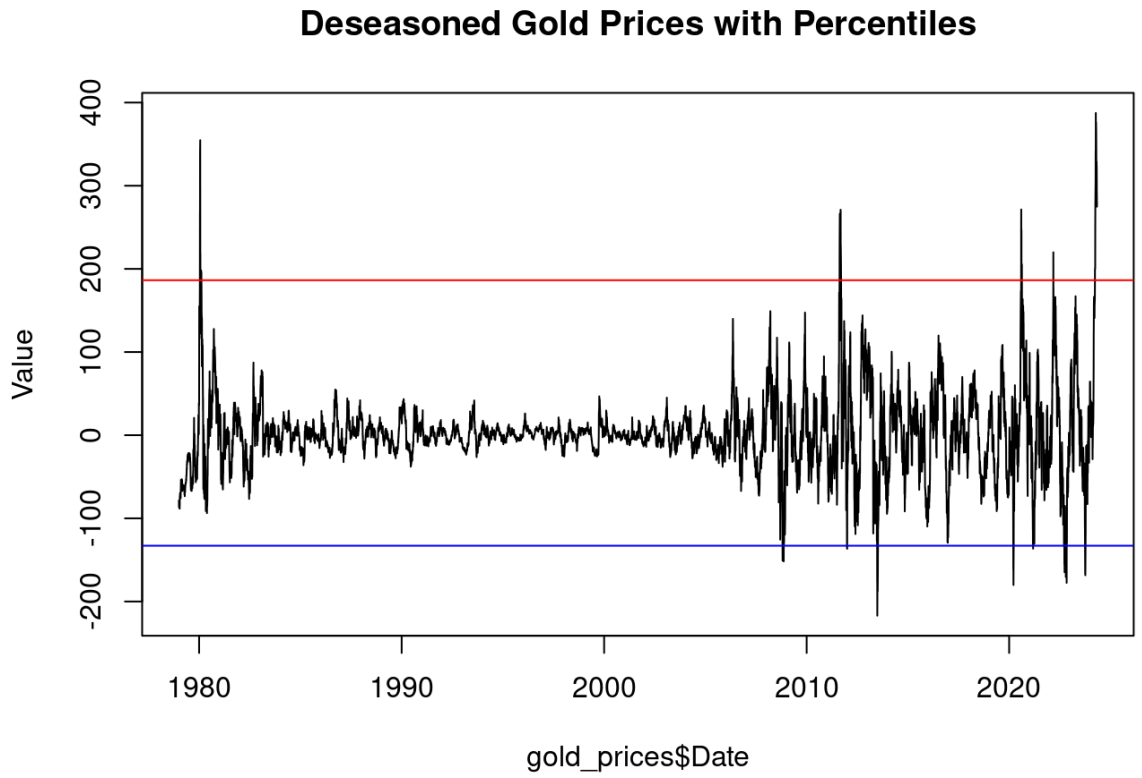
The difference appeared so insignificant that it's even hard to see it (you should zoom a lot and you'll see that the red deseasoned data is on top of the blue start data). It seemed strange because I really believed there should be some pattern, so I decided to check one more time on the not detrended data:



I even tried two different methods, but as you can see, the results were similarly disappointing. So I concluded that there is very insignificant seasonality I shouldn't consider it later; therefore, I returned to just detrended data.

Methodology

As you can see there are a lot of ups and downs in the data and deseasoning didn't help. Thus, I assumed this was due to some events in the world. I needed to consider only the most important ones so as not to overload the future model.



There are four significant positive spikes and one negative. I spend a lot of time going through all 12 thousand rows of my data, trying to notice where these anomalies start and end. When I got the approximate dates I searched the internet to understand the precise events.

So that's the list of events that I created:

1. Dates: 25.12.79 - 21.01.80

   Reason: USSR invaded into the Afghanistan

2. Dates: 20.05.2011 - 06.09.2011

Reason: There is no information about any event that happened, it seems to be only because of the world's global economic problems

3. Dates: 12.04.2013 - 28.06.2013

   Reason: Big gold crash that happened because of the news that Cyprus agreed to sell all their gold reserves

4. Dates: 11.05.20 - 18.08.20

   Reason: Global pandemic of Covid-19

5. Dates: 04.03.2024 - up to present

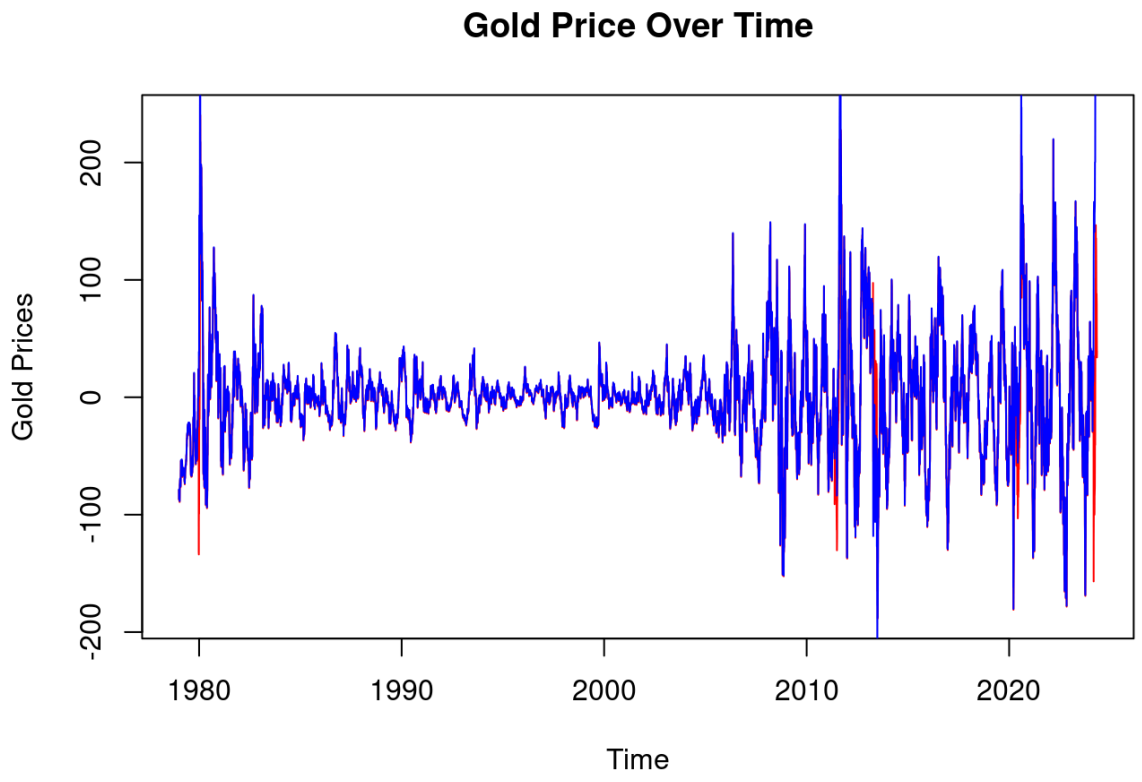
   Reason: I searched through the internet a lot, but no one actually knows the reason why prices suddenly started growing so fast

One of the smaller spikes in the plot is also connected to the full-scale russian invasion to Ukraine. It's not as significant as others, but it's important specifically for us.

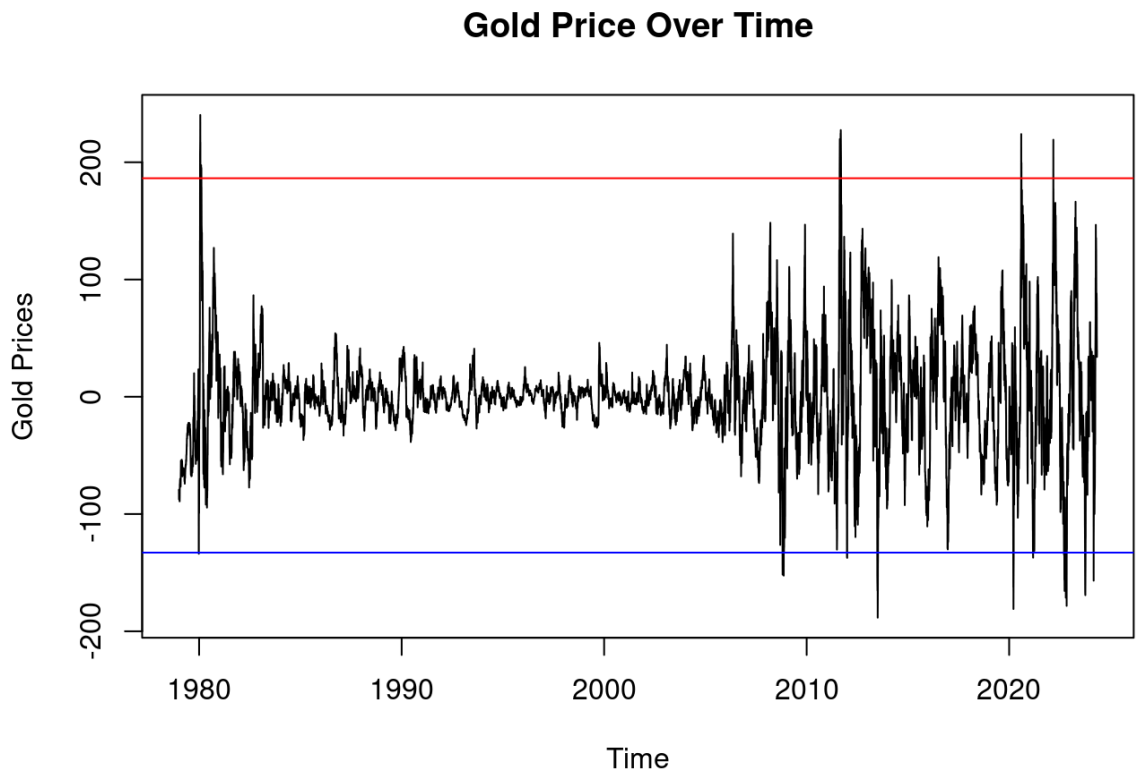
6. Dates: 22.02.22 - 08.03.22

   Reason: Full-scale russian invasion to Ukraine

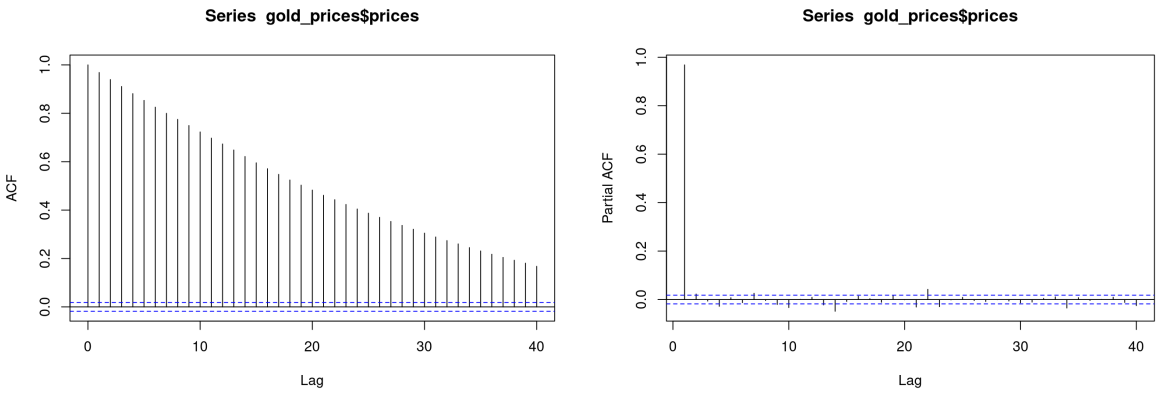
I added 6 corresponding dummy variables to the data and made a simple model to diminish their impact.



Here is how the data looks after that:



The next step is to build an ARIMA model. To specify it, at first, I needed to do an autocorrelation analysis.

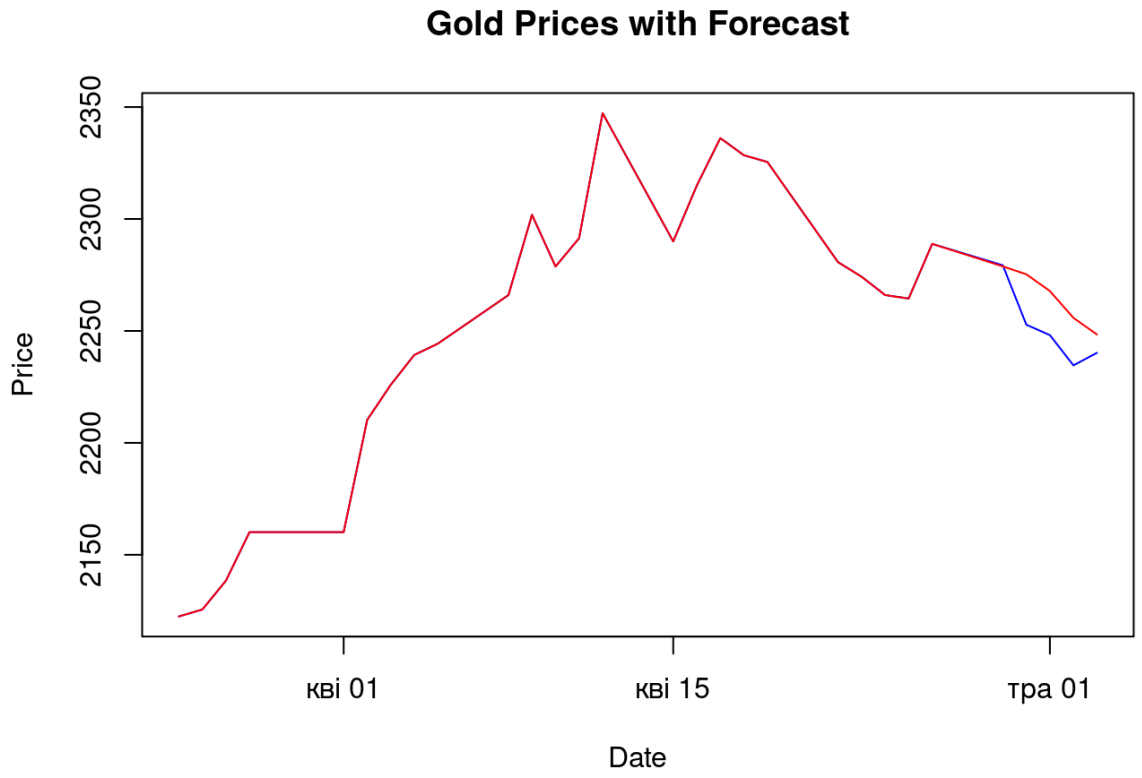


From these plots, we can see that the value of the AR process has a big range, while the value for the MA process could only be 1.

Simply sorting through the possible options and comparing its AICs I figured out that the best-fitted model is ARIMA(14,1, 1).

Results

ARIMA model helped me predict the values for this week:



Now, let's google the price for gold for 10.05.2024:

It’s important to remember, that at the end of the day, the price will be lower again and closer to yesterday's, so about 2320$, and my model predicted that the price is 2207.47 $. The difference is only about 5%.

Conclusions

This project was an excellent journey for me, and I hope for you too. I managed to fulfil my starting aim and tasks successfully. I had a few assumptions on the beginning so let’s go through all of them.

I was right about general gold price growth, but the trend appeared to be not even close to linear. On the other hand, I was totally wrong about patterns in gold prices; there appears to be almost no seasonality at all.

The world's events definitely affect gold prices but not as much, and as long as I thought, even the most important ones keep an impact only for a couple of months. Also, I was disappointed about so short-term effect of russia’s full-scale invasion in Ukraine.

In total can say that the mechanisms of prediction appeared to be a lot harder than I expected, but I believe I received good results.