Analysis, Trading Strategy, and Prediction of Gold price and influencing factors

Jovan Trajceski

Introduction

- Due to COVID-19, **gold** has become a safe-haven asset to buy
- We will explore gold relationship with other variables and trading strategy
- Prediction of future gold price

What we will cover?

This python project will walk you through:

- Exploratory Data Analysis
- Correlation between gold, silver, oil, S&P 500 and US dollar index
- Financial analysis
- Regression methodologies
- Gold trading strategy and backtesting
- Prediction of gold price
- Communication and alert system:
  - A telegram bot where you can check the real-time price of gold , and,
  - Receive an instant email notification upon subscription

Detailed Agenda

- PART 1: Load data and initial data preview
- PART 2: EDA (Exploratory Data Analysis)
- PART 3: Visualizing time series data
- PART 4: Financial analysis
- PART 5: Gold risk and volatility profile
- PART 6: Regression analysis - OLS
- PART 7: Gold trading strategy
- PART 8: Back-testing
- PART 9: Predicting gold prices via Facebook Prophet
- PART 10: Communication and alert system/s

Library used

- pandas_datareader
- datetime
- pandas
- Yahoo finance
- numpy
- matplotlib.pyplot
- statsmodels.api
- pandas tseries
- pandas.core tools
- telepot
- Time
- Sys
- matplotlib.patches
- matplotlib.collections
- seaborn
- matplotlib
- plotly.express
- sklearn.linear_model
- sklearn.model_selection
- fbprophet
- Smtplib
- Requests
- Json
- …etc

Data Preparation

Raw dataset

Ticker	Date	High	Low	Open	Close	Adj Close
GC=F	2001-11-01	279.700012	279.700012	279.700012	279.700012	279.700012
	2001-11-02	279.799988	279.799988	279.799988	279.799988	279.799988
	2001-11-05	278.500000	278.500000	278.500000	278.500000	278.500000
	2001-11-06	279.700012	279.700012	279.700012	279.700012	279.700012
	2001-11-07	281.000000	280.500000	280.500000	281.000000	281.000000

Data source: gold/oil/silver/S&P 500/US dollar index from 2001-2020 as per yahoo finance

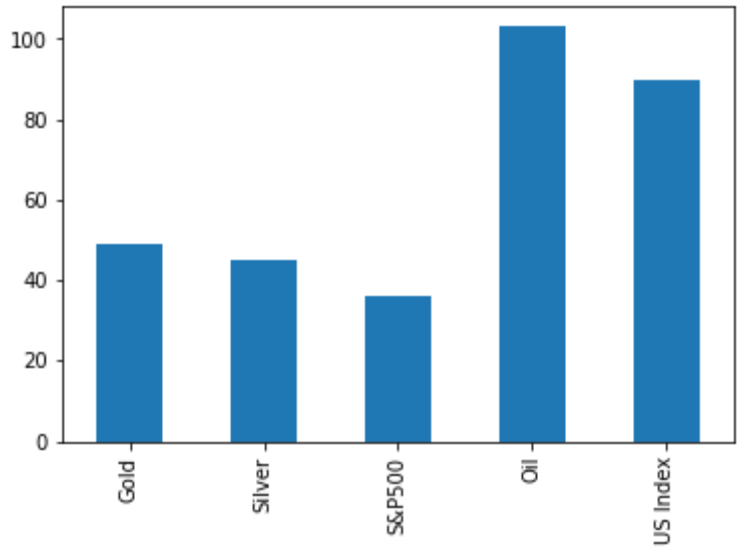
Data Preparation

Date	Gold	Silver	S&P500	Oil	US Index
2001-11-01	279.700012	4.196	1082.00	20.389999	114.459999
2001-11-02	279.799988	4.099	1089.50	20.180000	114.540001
2001-11-05	278.500000	4.087	1103.75	20.020000	114.540001
2001-11-06	279.700012	4.108	1121.00	19.920000	114.809998
2001-11-07	281.000000	4.118	1119.00	20.090000	114.980003

	Gold	Silver	S&P500	Oil	US Index
count	4728.000000	4732.000000	4741.000000	4674.000000	4687.000000
mean	1026.413261	15.989062	1647.822571	64.307441	88.852195
std	442.095322	8.117746	653.590428	25.424797	9.881474
min	272.200012	4.026000	676.000000	-37.630001	71.330002
25%	619.399994	11.029250	1145.000000	46.032499	80.704998
50%	1186.850037	15.863000	1383.500000	60.725000	87.599998
75%	1322.724976	18.362000	2079.750000	85.432503	96.274998
max	1888.699951	48.584000	3388.250000	145.179993	120.239998

Data source: gold/oil/silver/S&P 500/US dollar index from 2001-2020 as per yahoo finance.

Data Preparation

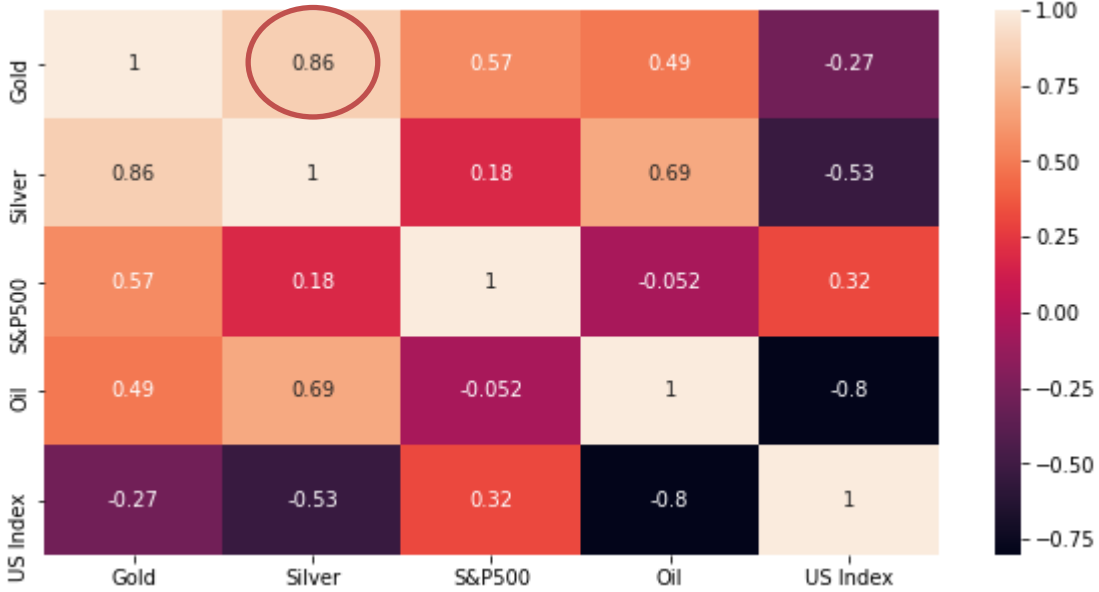


- ❑ Plot the NaN value(mostly US dollar index and oil have Nan value)
- ❑ Then propagate NaN values forward

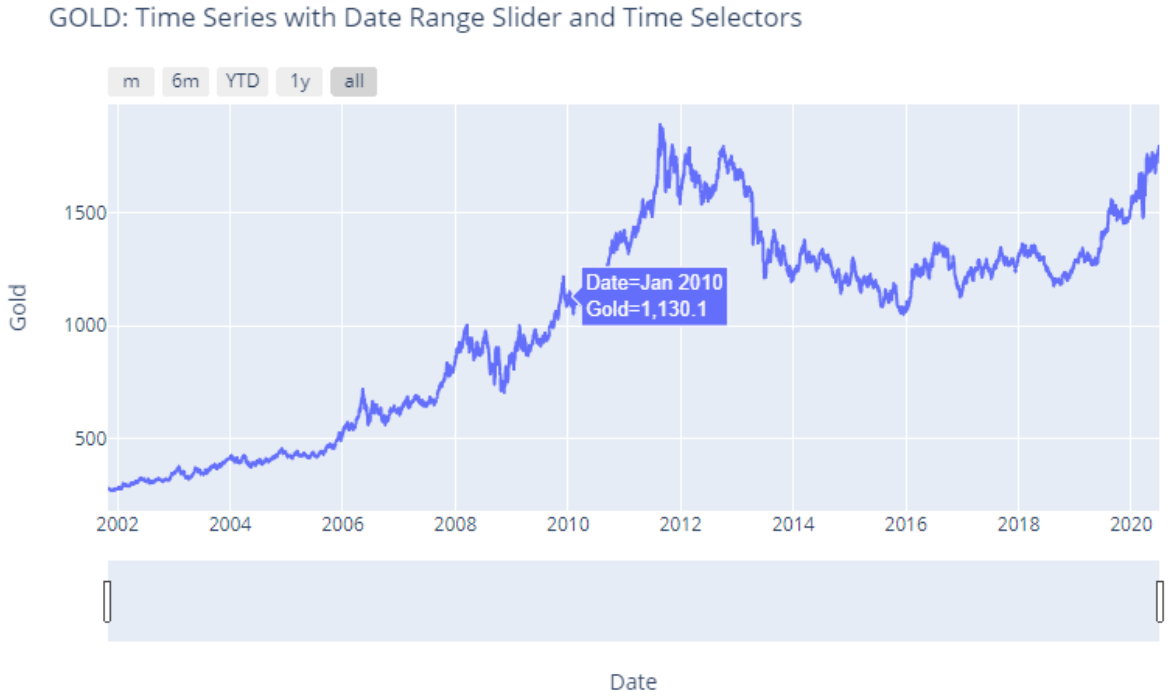
Data source: gold/oil/silver/S&P 500/US dollar index from 2001-2020 as per yahoo finance.

Exploratory Data Analysis

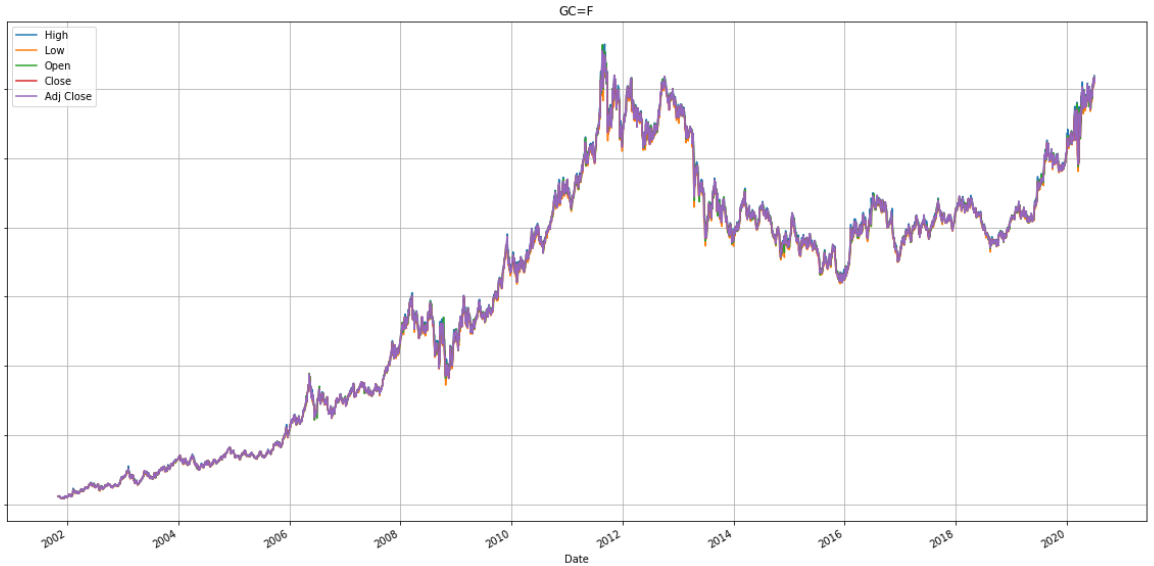
Correlation heatmap analysis



EDA-Interactive GOLD Chart

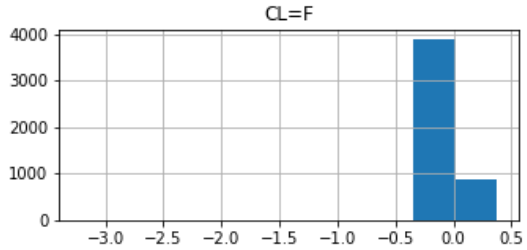


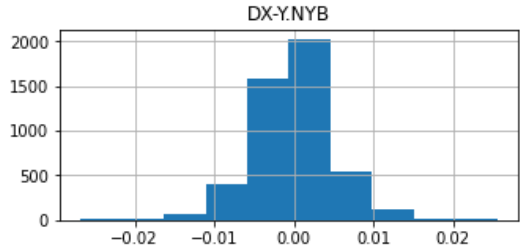
Visualizing Time Series Data

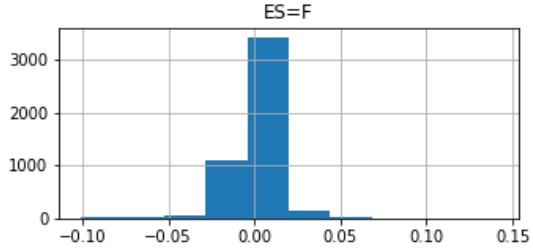


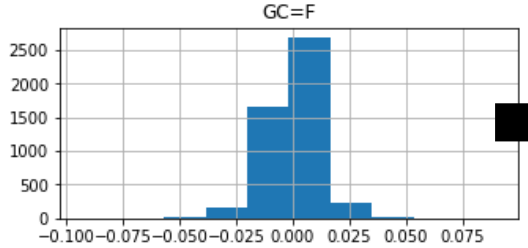
Financial Analysis

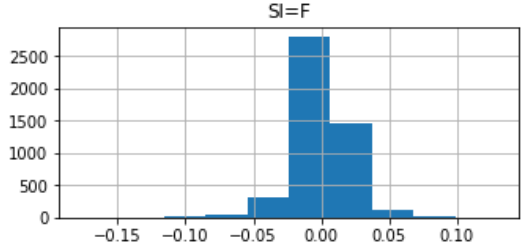
-Daily Changes and Distribution











- Gold: The distribution looks very symmetrical and normally distributed

Financial Analysis

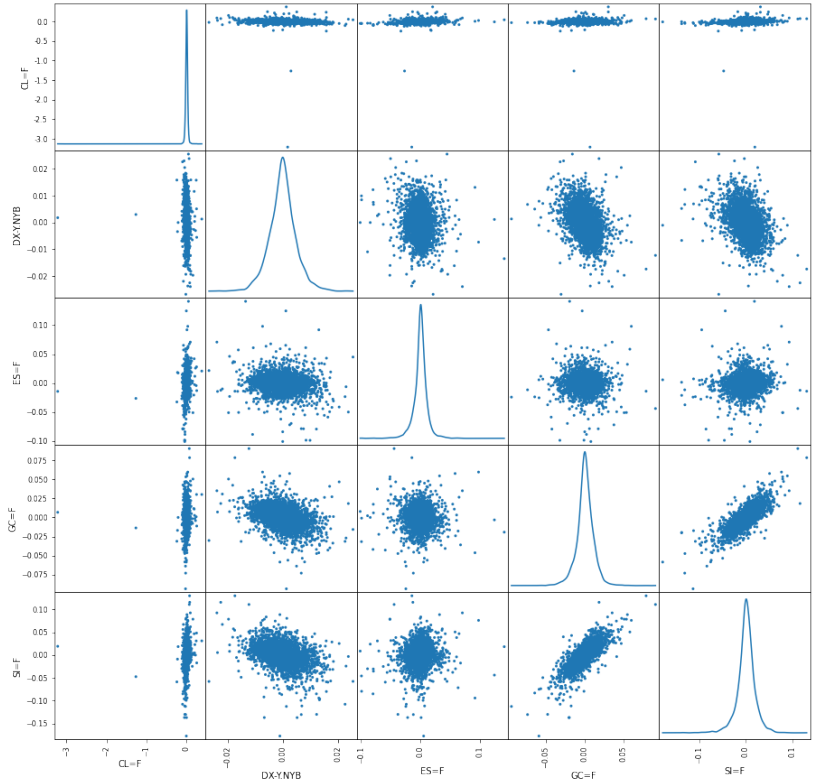
-Return Comparison

Daily cumulative returns

Monthly cumulative returns

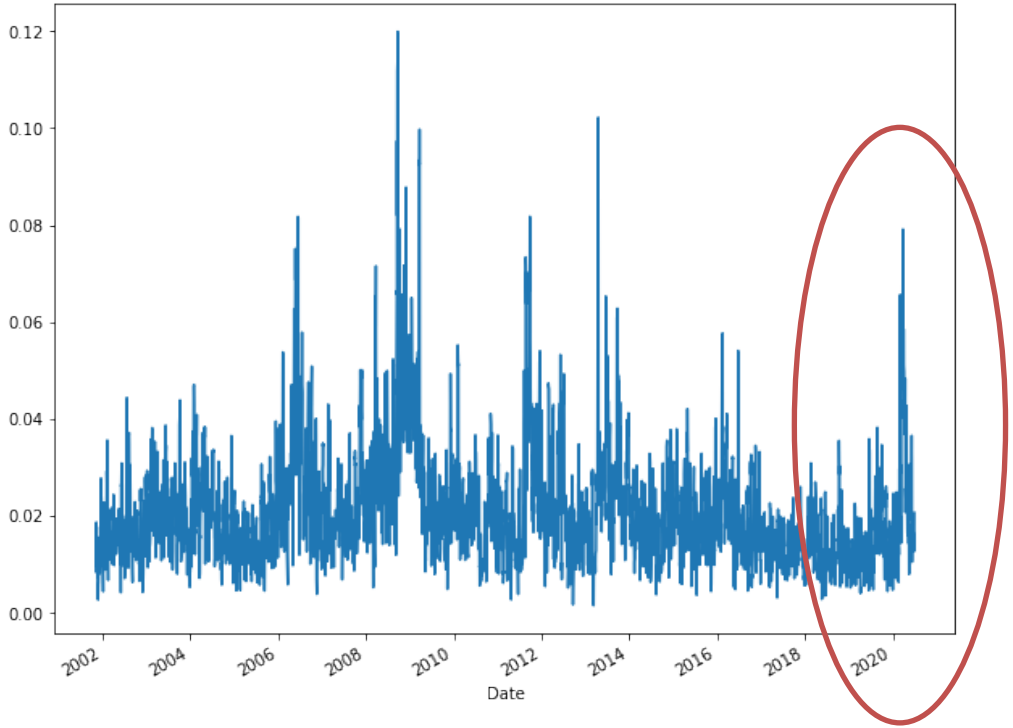
Financial Analysis

- Daily % Change



- This shows us the daily percentage change relationship and distribution of the assigned variables

Financial Analysis
-gold price volatility



- Gold violitility has been very significant in 2020, especially around Covid-19 timeline
- Also we can see the past financial crisis as per the spikes prior to 2020

Regression analysis
- OLS

OLS Regression Results

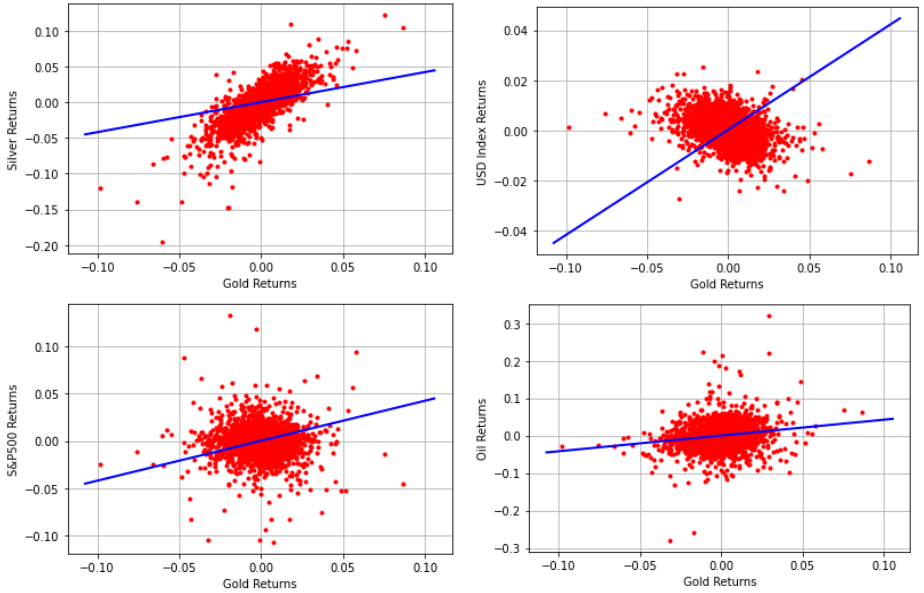
Dep. Variable:	GC=F	R-squared:	0.634
Model:	OLS	Adj. R-squared:	0.633
Method:	Least Squares	F-statistic:	4018.
Date:	Sat, 05 Dec 2020	Prob (F-statistic):	0.00
Time:	02:55:17	Log-Likelihood:	16592.
No. Observations:	4651	AIC:	−3.318e+04
Df Residuals:	4648	BIC:	−3.316e+04
Df Model:	2		
Covariance Type:	nonrobust		

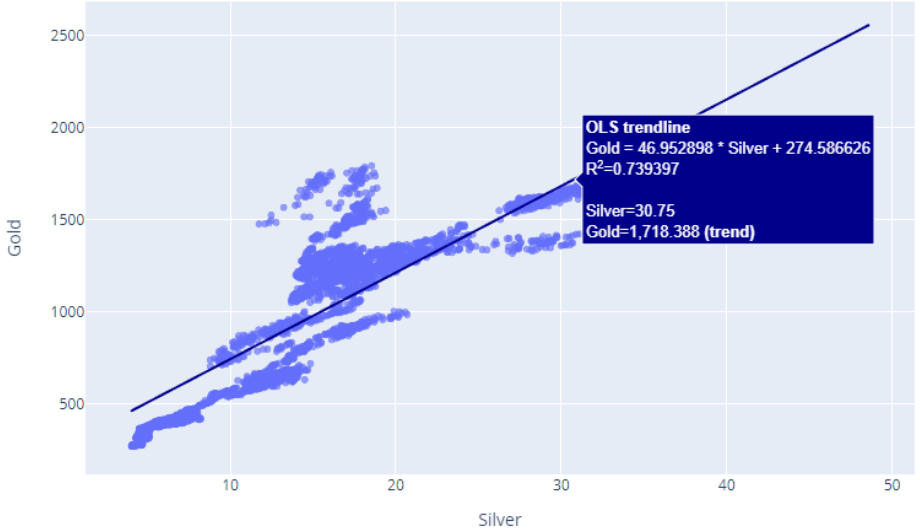
	coef	std err	t	P>|t|	[0.025	0.975]
Const	0.0002	0.000	2.220	0.026	2.6e-05	0.000
Silver	0.4179	0.005	77.291	0.000	0.407	0.429
USD Index	−0.2814	0.022	−12.884	0.000	−0.324	−0.239

Omnibus:	536.412	Durbin-Watson:	2.007
Prob(Omnibus):	0.000	Jarque-Bera (JB):	4670.808
Skew:	0.191	Prob(JB):	0.00
Kurtosis:	7.894	Cond. No.	219.

- Initial regression model included 4 variables: Silver, S&P500, Oil, US dollar Index
- We removed less significant variables in our second OLS regression model where silver and US dollar index are labelled as most ***influencing factors***

Regression analysis
- OLS





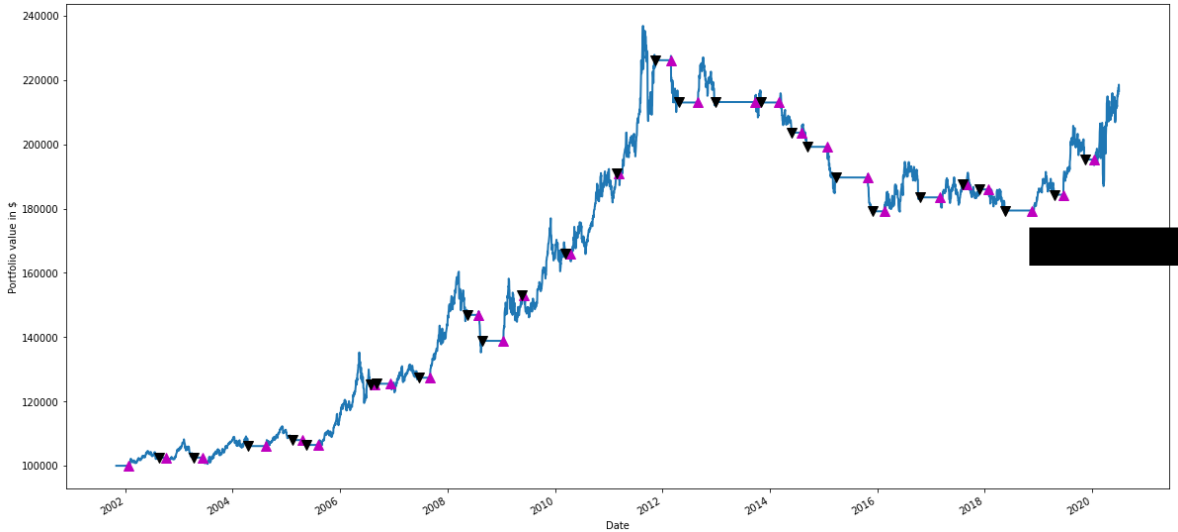
Gold trading strategy

– SMA 40 days VS SMA 100 days

Back-testing Results

Trading strategy: moving average crossover

- ❑ Buy Signal
- ❑ Sell Signal



Mock Portfolio

Date	GC=F	holdings	cash	total	returns
2020-06-26	177250.000000	177250.000000	39280.010986	216530.010986	0.004826
2020-06-28	178530.004883	178530.004883	39280.010986	217810.015869	0.005911
2020-06-29	177480.004883	177480.004883	39280.010986	216760.015869	-0.004821
2020-06-30	179300.000000	179300.000000	39280.010986	218580.010986	0.008396
2020-07-01	177319.995117	177319.995117	39280.010986	216600.006104	-0.009058

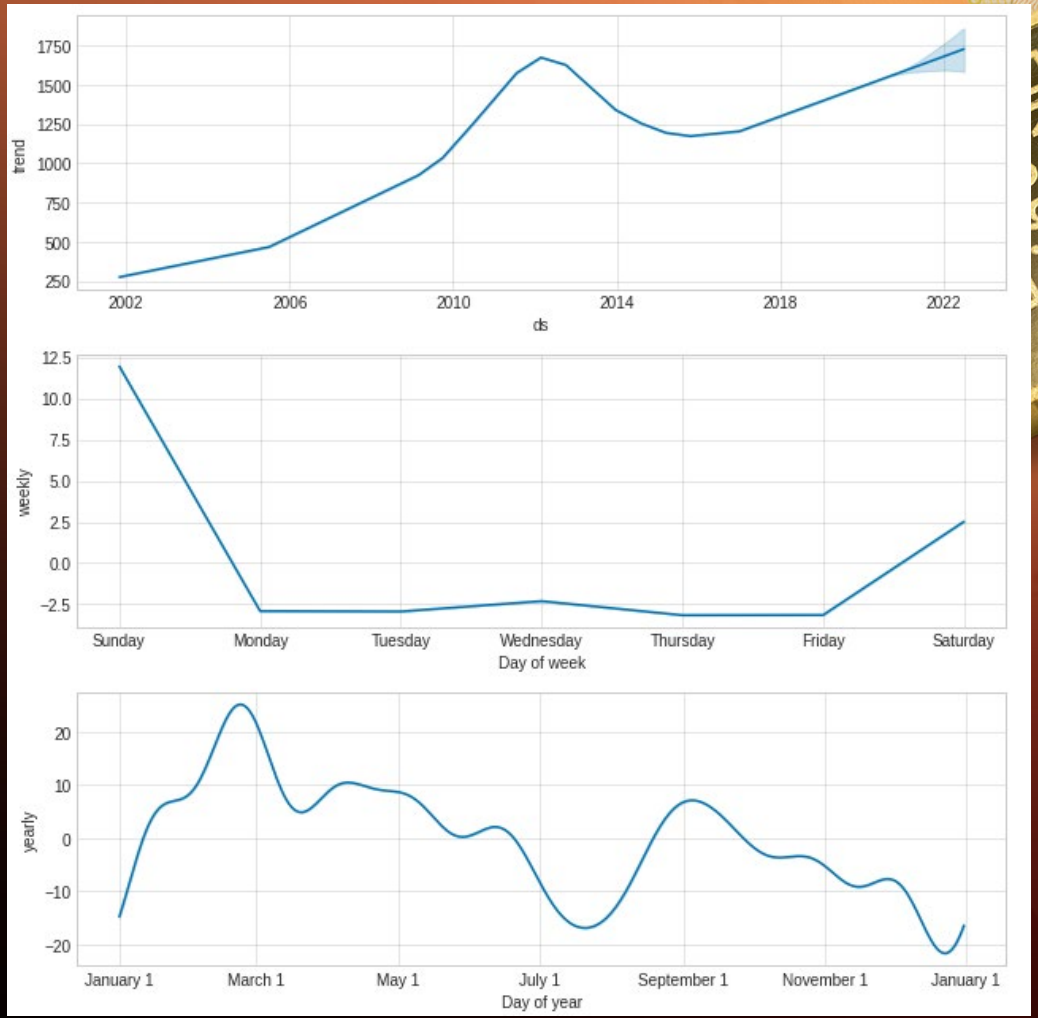
Compound Annual Growth Rate (CAGR) of 10.39%

Predicting gold prices via Facebook Prophet

- Prophet is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects
- It works best with time series that have strong seasonal effects and several seasons of historical data
- Prophet is robust to missing data and shifts in the trend, and typically handles outliers well

Gold price prediction - monthly upward trend!

Gold price prediction - seasonality / trends



Gold Price Checker

Gold: Analysis, Trading Strategy, Backtesting, Prediction and C...

File Edit View Insert Runtime Tools Help All changes saved

Comment Share Editing

+ Code + Text

RAM Disk

```
[ ] !pip install telepot
    import time
    import telepot
    import sys
    from telepot.loop import MessageLoop
    from telepot.namedtuple import ReplyKeyboardMarkup
    from telepot.namedtuple import InlineKeyboardMarkup, InlineKeyboardButton
    from telepot.delegate import pave_event_space, per_chat_id, create_open
    # Bot code here
    class PriceChecker(telepot.helper.ChatHandler):
    # The majority function for this bot is check the real gold price in RMB


        def __init__(self, *args, **kwargs):
            super(PriceChecker, self).__init__(*args, **kwargs)
            # give a indicator so bot knows where it is
            self.indicator = "price_check"
            # what it wanna send in the email
            self.wannasend = {}

        def on_chat_message(self, msg):
            content_type, chat_type, chat_id = telepot.glance(msg)
            print(content_type, chat_type, chat_id)
            # add the checking list, randomly select some, the most common use one is "AU99.99"
            p_list = [["AU99.99"], ["Au100g"], ["Au(T+N1"], ["Au(T+D)"], ["Au(T+N1)"],['MAUTD']]
            if self.indicator == "price_check":
              # for now this bot can check the price only
                if msg["text"] == "/price":
                    mark_up = ReplyKeyboardMarkup(keyboard=p_list, one_time_keyboard=True)
                    # To
                    bot.sendMessage(chat_id, text="What price you want to check?", reply_markup=mark_up)
                  # reply the price of selected product
                elif [msg["text"]] in p_list:
                  # Using the function in last chunk
                    data = get(response.json())
                    say0 = which_price(data, msg["text"])
                    say = take_this(say0)
                    # storage what we may send via email
                    self.wannasend["say"] = say
                    bot.sendMessage(chat_id, say)
                    self.indicator = "e-mail check"
                # Send an email
            if self.indicator == "e-mail check":
                mark_up = ReplyKeyboardMarkup(keyboard=[["Yes"], ["No"]], one_time_keyboard=True)
```

Communication Script

Archived chats

PC Price Checker 5:56 PM
What price you want to check?

Telegram Bot + Email Alerts

December 5

/start 5:56 PM

/price 5:56 PM

What price you want to check? 5:56 PM

Write a message...

AU99.99

Au100g

Au(T+N1

Au(T+D)

Conclusion

- Two most significant variables for gold price
  - Silver (+)
  - US Dollar index (-)
- The trading strategy – SMA40 VS SMA100
- The gold price would increase as Monthly Data Predictions (24 months ahead)