

Name: Vikram Surya R

Reg No: 19BCE1490

Essentials of Data Analytics

Tasks for Week-2: Time-series Forecasting

Understand time-series operations/functions and forecast the annual gdp growth rate of India based on given instructions.

Aim:

To understand time-series operations/functions and forecast the annual gdp growth rate of India based on given instructions.

gold.csv CODE:

```
setwd("C:/Users/VIKRAM SURYA/Desktop/EDA_LAB")
```

```
gold<-read.csv("gold.csv")
```

```
library(dplyr)
```

```
library(forecast)
```

```
library(tseries)
```

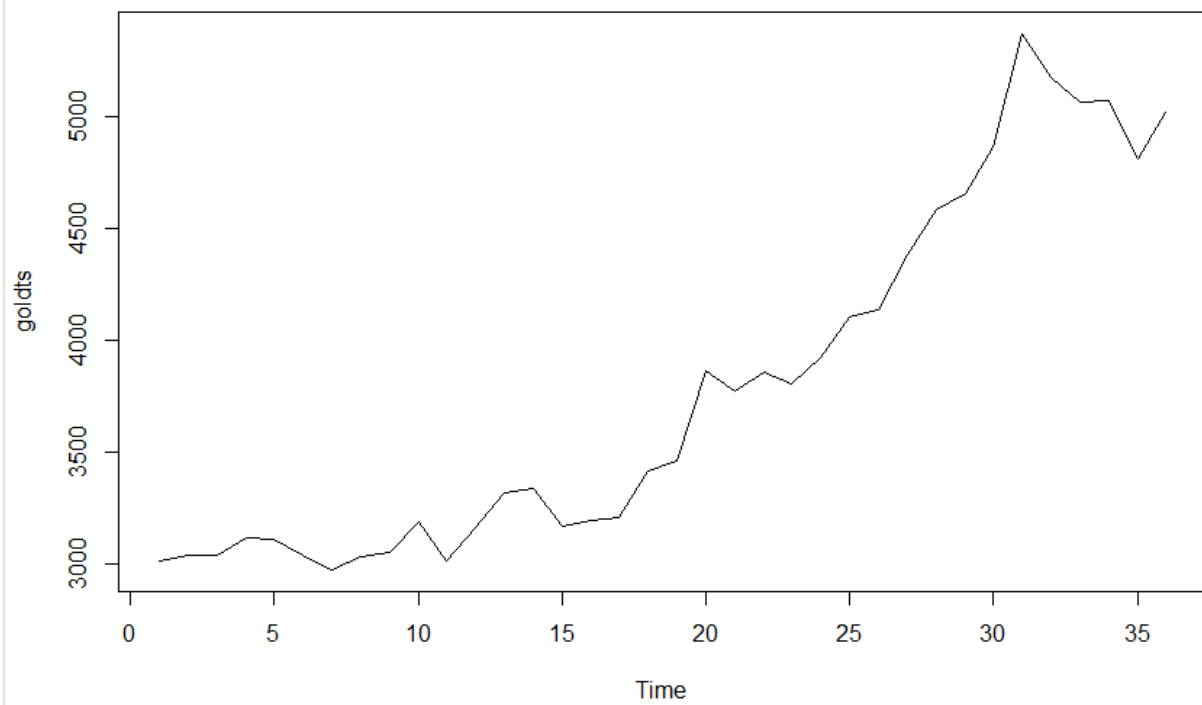
```
View(gold)
```

	Month	Price
1	1	3016
2	2	3044
3	3	3041
4	4	3121
5	5	3111
6	6	3043
7	7	2977
8	8	3036
9	9	3051
10	10	3191
11	11	3016
12	12	3164
13	13	3321
14	14	3338
15	15	3170
16	16	3194
17	17	3212
18	18	3420
19	19	3465

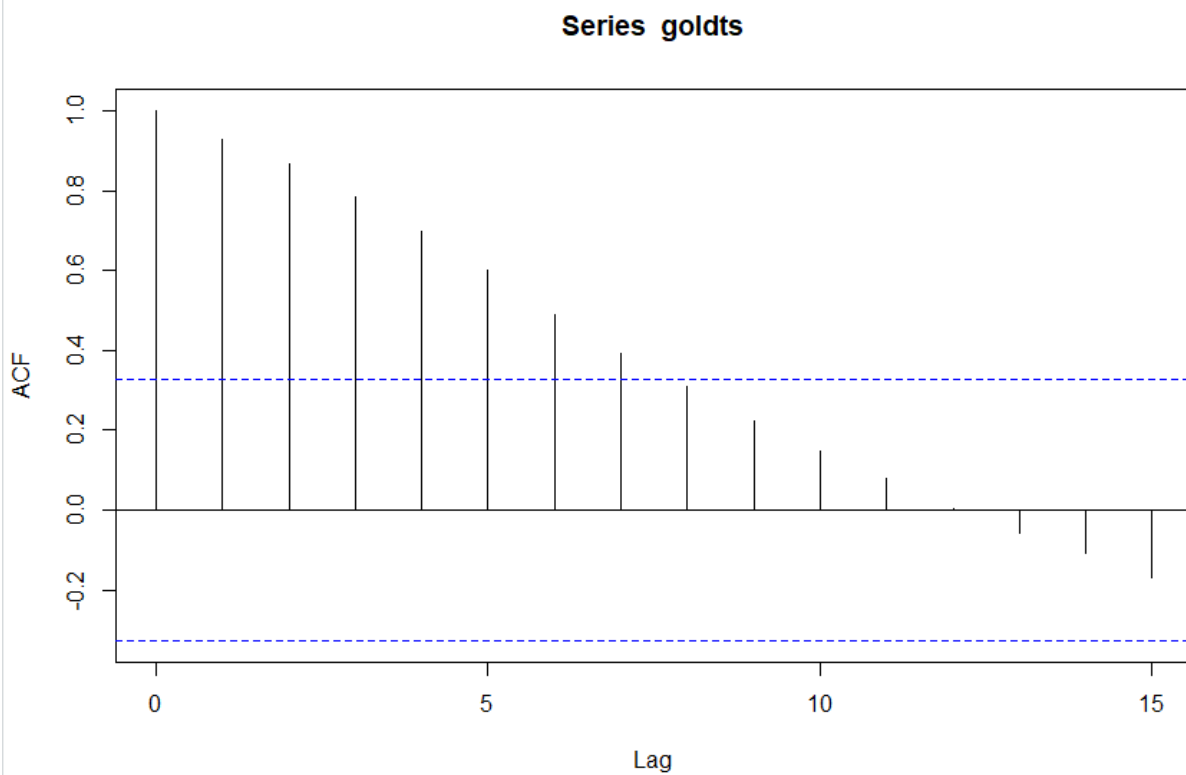
```
goldts<-ts(gold$Price,start = min(gold$Month),end = max(gold$Month),frequency=1)
```

```
class(goldts)
```

```
plot(goldts)
```



acf(goldts)



```

> adf.test(goldts)

        Augmented Dickey-Fuller Test

data:  goldts
Dickey-Fuller = -2.3526, Lag order = 3,
p-value = 0.4359
alternative hypothesis: stationary

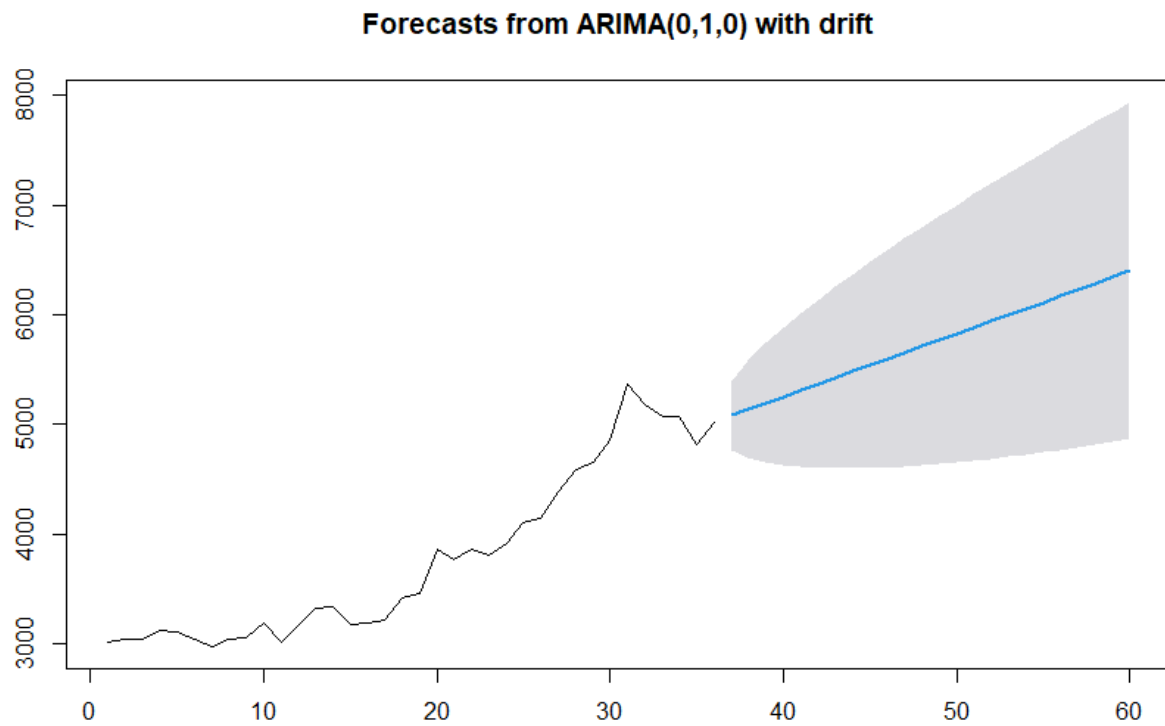
> goldmodel=auto.arima(goldts,ic="aic",trace = TRUE)

ARIMA(2,1,2) with drift      : Inf
ARIMA(0,1,0) with drift     : 457.5809
ARIMA(1,1,0) with drift     : 459.3633
ARIMA(0,1,1) with drift     : 459.385
ARIMA(0,1,0)                : 459.9305
ARIMA(1,1,1) with drift     : 461.3121

Best model: ARIMA(0,1,0) with drift

> goldf=forecast(goldmodel,level=c(95),h=24)
> goldf
   Point Forecast    Lo 95    Hi 95
37      5081.371 4767.741 5395.001
38      5138.743 4695.203 5582.283
39      5196.114 4652.891 5739.338
40      5253.486 4626.226 5880.746
41      5310.857 4609.559 6012.155
42      5368.229 4599.995 6136.462
43      5425.600 4595.813 6255.387
44      5482.971 4595.892 6370.051
45      5540.343 4599.453 6481.233
46      5597.714 4605.929 6589.500
47      5655.086 4614.892 6695.279
48      5712.457 4626.011 6798.904
49      5769.829 4639.019 6900.638
50      5827.200 4653.704 7000.696
51      5884.571 4669.887 7099.255
52      5941.943 4687.423 7196.463
53      5999.314 4706.184 7292.444
54      6056.686 4726.066 7387.305
55      6114.057 4746.975 7481.139
56      6171.429 4768.832 7574.025
57      6228.800 4791.566 7666.034
58      6286.171 4815.116 7757.227
59      6343.543 4839.426 7847.660
60      6400.914 4864.447 7937.382
> plot(goldf)

```



STATISTICS:

```
> accuracy(goldmodel)
               ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 0.08218409 155.5098 116.6965 -0.1799051 2.960037 0.9286895 -0.07882193
> |
```

gdp.csv CODE:

```
setwd("C:/Users/VIKRAM SURYA/Desktop/EDA_LAB")

gdp<-read.csv("gdp.csv")

library(dplyr)

library(forecast)

library(tseries)

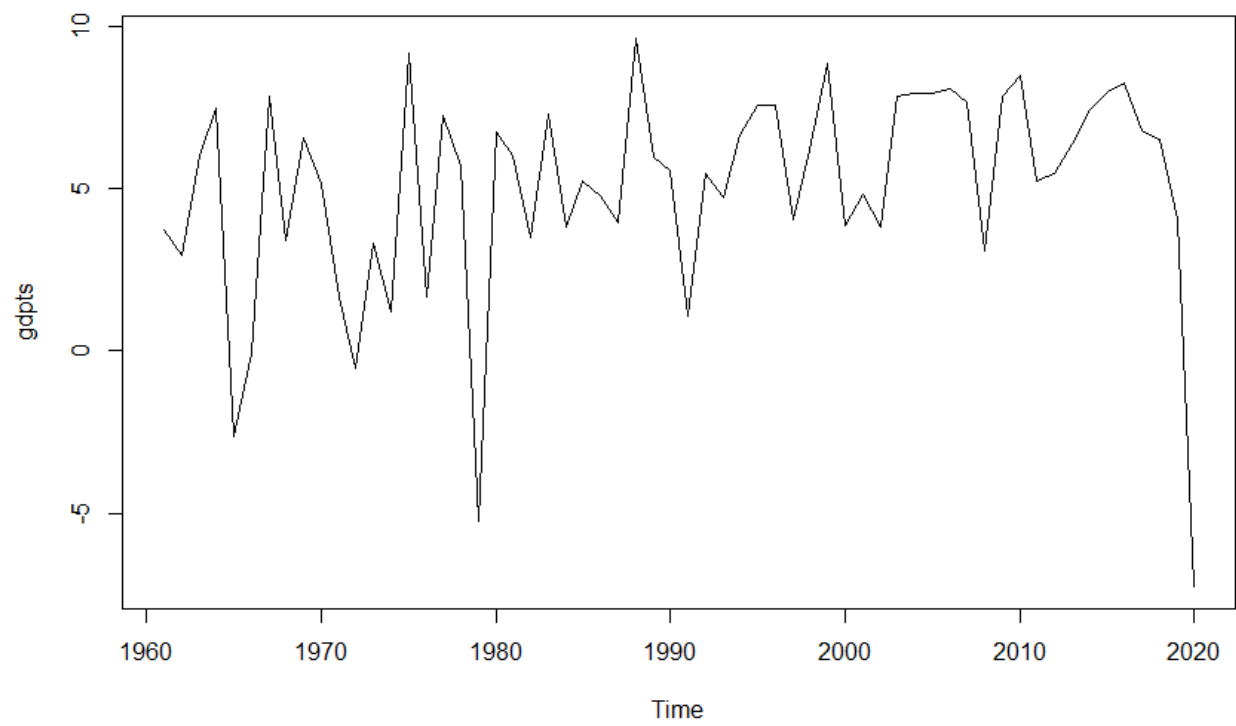
View(gdp)
```

	Year	GDP_gr
1	1961	3.72274253
2	1962	2.93112774
3	1963	5.99435326
4	1964	7.45295012
5	1965	-2.63577011
6	1966	-0.05532877
7	1967	7.82596303
8	1968	3.38792918
9	1969	6.53970030
10	1970	5.15722974
11	1971	1.64293038
12	1972	-0.55330131
13	1973	3.29552113
14	1974	1.18533626
15	1975	9.14991201
16	1976	1.66310364
17	1977	7.25476459
18	1978	5.71253209
19	1979	-5.23818270

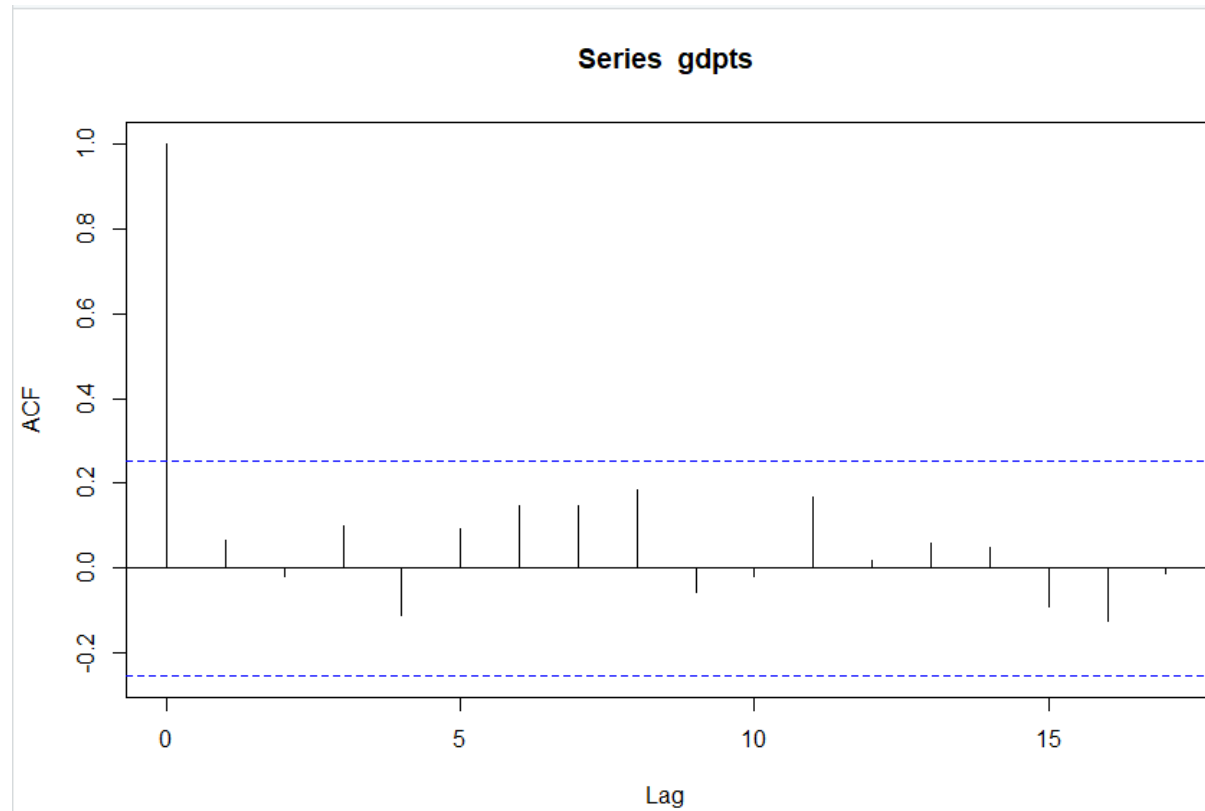
```
gdpts<-ts(gdp$GDP_gr,start = min(gdp$Year),end = max(gdp$Year),frequency=1)
```

```
class(gdpts)
```

```
plot(gdpts)
```



acf(gdpts)



```
> adf.test(gdpts)
```

Augmented Dickey-Fuller Test

```
data: gdpts
Dickey-Fuller = -4.7448, Lag order = 3,
p-value = 0.01
alternative hypothesis: stationary
```

warning message:

In adf.test(gdpts) : p-value smaller than printed p-value

```
> gdpmodel=auto.arima(gdpts,ic="aic",trace = TRUE)
```

```
ARIMA(2,1,2) with drift      : Inf
ARIMA(0,1,0) with drift     : 341.4397
ARIMA(1,1,0) with drift     : 332.4653
ARIMA(0,1,1) with drift     : Inf
ARIMA(0,1,0)                : 339.554
ARIMA(2,1,0) with drift     : 326.0715
ARIMA(3,1,0) with drift     : 327.9755
ARIMA(2,1,1) with drift     : Inf
ARIMA(1,1,1) with drift     : Inf
ARIMA(3,1,1) with drift     : Inf
ARIMA(2,1,0)                : 324.2097
ARIMA(1,1,0)                : 330.5929
ARIMA(3,1,0)                : 326.1139
ARIMA(2,1,1)                : 317.8228
ARIMA(1,1,1)                : 316.651
ARIMA(0,1,1)                : 314.6516
ARIMA(0,1,2)                : 316.6508
ARIMA(1,1,2)                : 316.6275
```

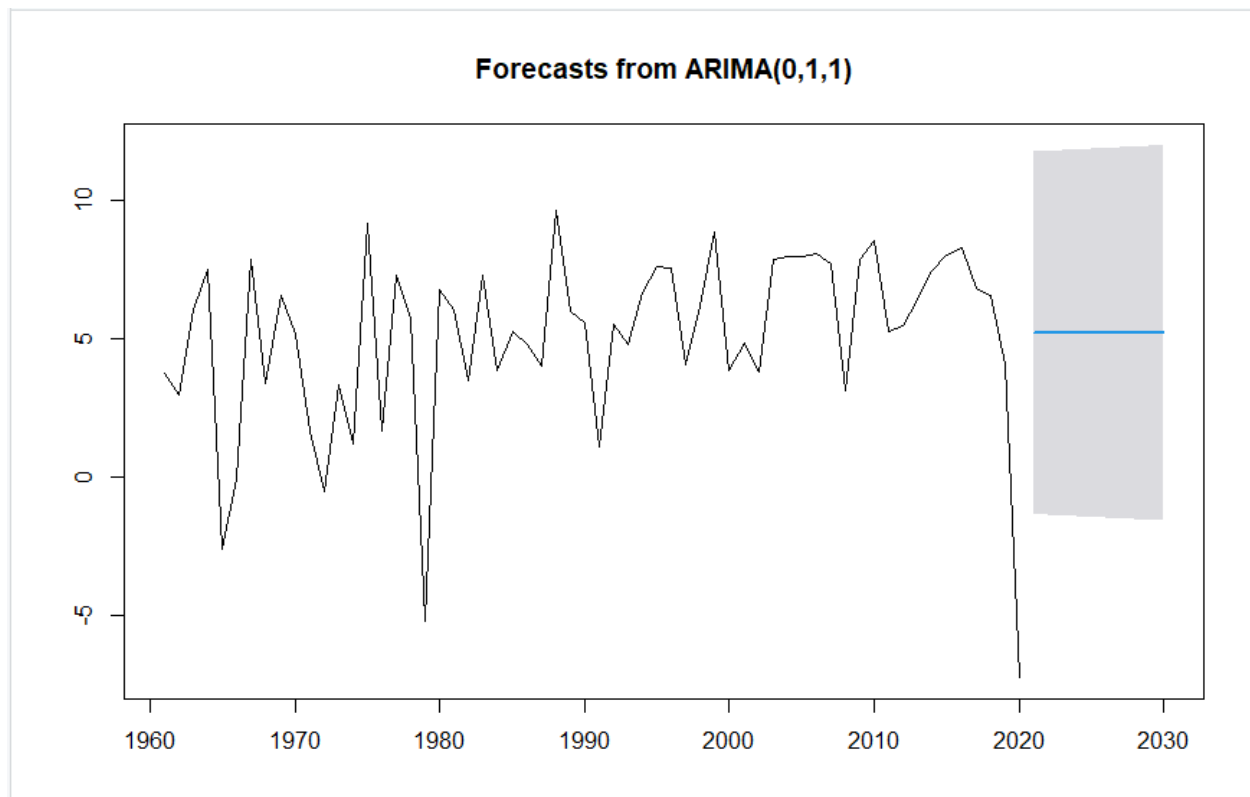
Best model: ARIMA(0,1,1)

```
> gdpf=forecast(gdpmodel,level=c(95),h=10)
```

```
> gdpf
```

	Point	Forecast	Lo 95	Hi 95
2021		5.177274	-1.376684	11.73123
2022		5.177274	-1.401989	11.75654
2023		5.177274	-1.427197	11.78174
2024		5.177274	-1.452309	11.80686
2025		5.177274	-1.477327	11.83187
2026		5.177274	-1.502250	11.85680
2027		5.177274	-1.527082	11.88163
2028		5.177274	-1.551821	11.90637
2029		5.177274	-1.576470	11.93102
2030		5.177274	-1.601029	11.95558

```
> plot(gdpf)
```

STATISTICS:

```
> accuracy(gdpmodel)
              ME      RMSE      MAE      MPE
Training set 0.2704179 3.287709 2.345416 121.6616
              MAPE      MASE      ACF1
Training set 161.0542 0.7720211 -0.02667223
> |
```

INFERENCE:

1. gold.csv

The expected values for the coming 24 months have been plotted into a graph using time-series operations/functions and forecast command.

2. gdp.csv

The expected values for the coming 10 years have been plotted into a graph using time-series operations/functions and forecast command.