ABHINAV VIJAYAKUMAR

19BCE1311

CSE3506 – ESSENTIALS OF DATA ANALYTICS LAB-2

DR. LAKSHMI PATHI JAKKAMPUTI (L21 + L22)

Tasks for Week-2: Forecasting

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

Aim: To develop a forecasting model that forecasts the value 24 units ahead of time

Algorithm:

1. Set the working directory
2. Convert the dataset into time-series data using the 'ts' function.
3. Using acf and pacf check if the data is stationary visually, if the readings are below the blue line then the data is stationary
4. Using adf.test check if the data is stationary using p-value.
5. Using auto ARIMA, find out which model is the best.
6. Pass the best model to the forecast function so that the forecasting is done with 95% confidence for the next 24 units of data.

Inference:

i) For Gold Prices:

The given data is not stationary. We can see this using p value and acf and pacf graph.
p-value = 0.4359

Series goldts

Series goldts
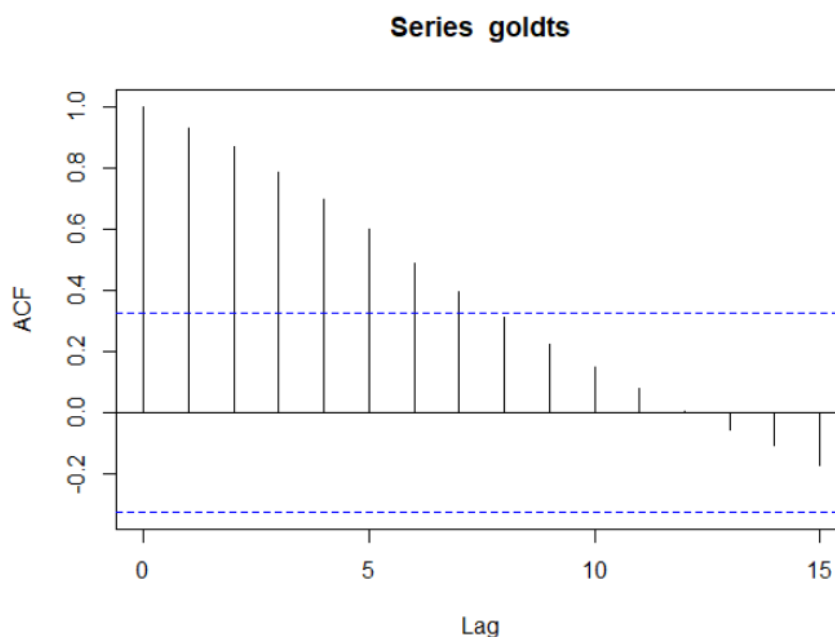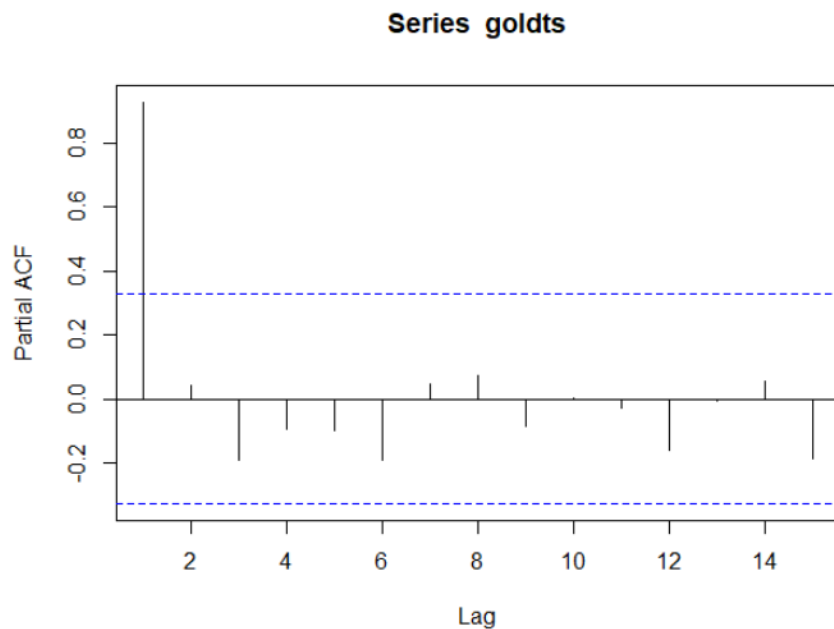
Best ARIMA Model:

Auto ARIMA

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

Accuracy of the Model:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	0.08218409	155.5098	116.6965	-0.1799051	2.960037	0.9286895	-0.07882193

ii) For gdp.csv:

The given data is stationary. We can see this using p value and acf and pacf graph.
p-value = 0.01

Series gdpts

Series gdpts
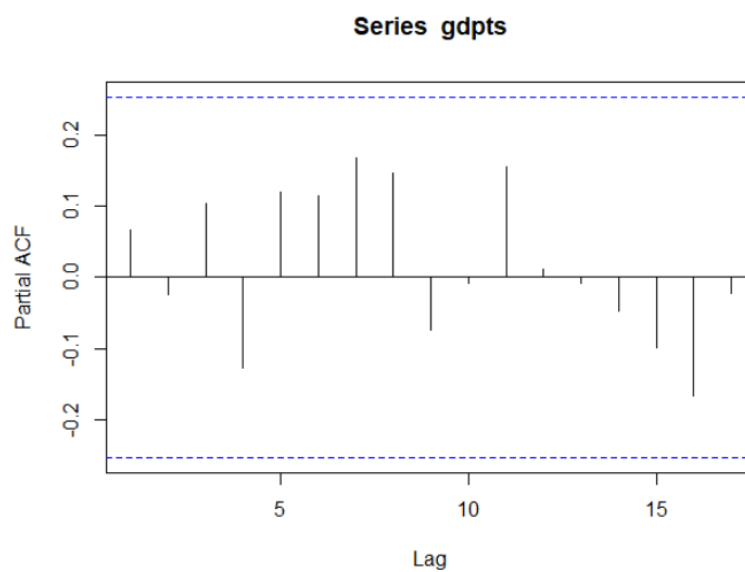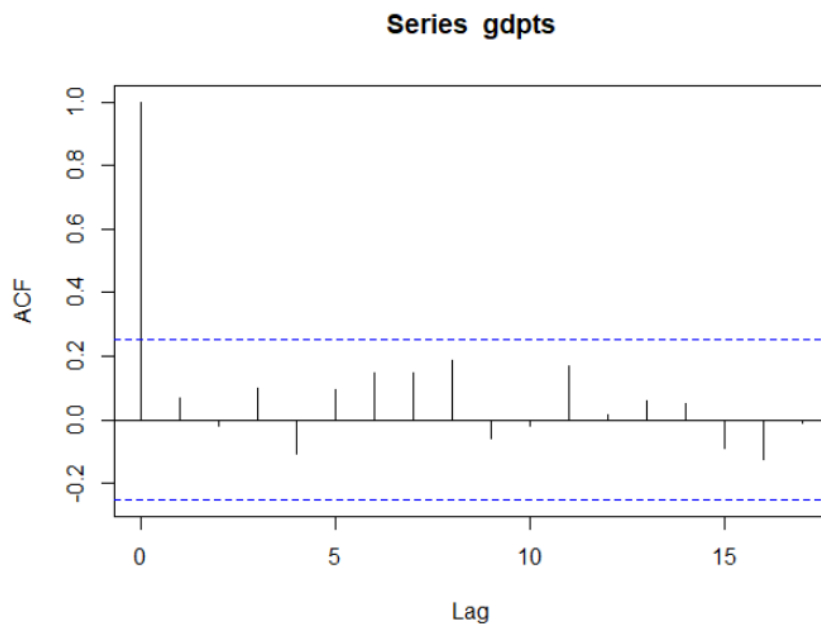
Best ARIMA Model:

Auto ARIMA

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)

Accuracy of the Model:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	0.2704179	3.287709	2.345416	121.6616	161.0542	0.7720211	-0.02667223

Program:

i) For gold.csv:

```
setwd("C:/Users/Abhinav Vijayakumar/Desktop/VIT Academics/Sem 6/Essentials of
Data Analytics/LAB/LAB 2")


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

library(forecast)

library(tseries)

View(gold)

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

class(goldts)
```

```
plot(goldts)

acf(goldts)

pacf(goldts)

adf.test(goldts) # stationary only if p value <0.05


# To make it stationary, differentiate

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

goldf=forecast(goldmodel, level=c(95), h=24)

goldf

plot(goldf)

accuracy(goldmodel)
```

ii) For gdp.csv

```
setwd("C:/Users/Abhinav Vijayakumar/Desktop/VIT Academics/Sem 6/Essentials of
Data Analytics/LAB/LAB 2")


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

library(forecast)

library(tseries)

View(gdp)

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

class(gdpts)

plot(gdpts)

acf(gdpts)

pacf(gdpts)

adf.test(gdpts) # stationary only if p value <0.05


# To make it stationary, differentiate

gdpmodel=auto.arima(gdpts, ic='aic', trace = TRUE)

gdpf=forecast(gdpmodel, level=c(95), h=24)

gdpf

plot(gdpf)

accuracy(gdpmodel)
```

Result:

Forecast for gold prices:

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

Forecasts from ARIMA(0,1,0) with drift
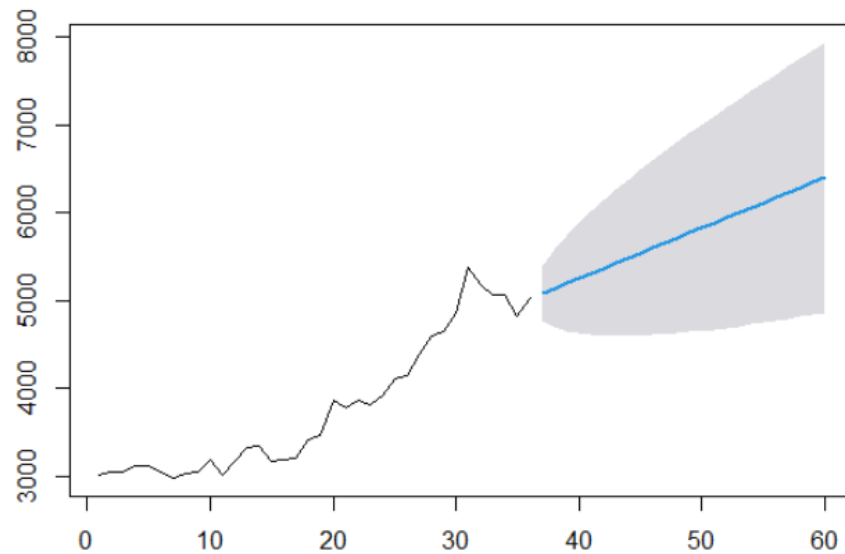
Forecast for gdp:

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
2031	5.177274	-1.625500	11.98005
2032	5.177274	-1.649882	12.00443
2033	5.177274	-1.674178	12.02873
2034	5.177274	-1.698389	12.05294
2035	5.177274	-1.722514	12.07706
2036	5.177274	-1.746555	12.10110

2037	5.177274	-1.770513	12.12506
2038	5.177274	-1.794389	12.14894
2039	5.177274	-1.818183	12.17273
2040	5.177274	-1.841896	12.19644
2041	5.177274	-1.865530	12.22008
2042	5.177274	-1.889085	12.24363
2043	5.177274	-1.912561	12.26711
2044	5.177274	-1.935960	12.29051

Forecasts from ARIMA(0,1,1)