

Homework 9
 Ali Lacey
 Thomas Pattara
 Stat 560

1. Analyze the data ex1. See attached.
 - a.) Plot the production data in ex1 and its ACF and PACF. (5 points)

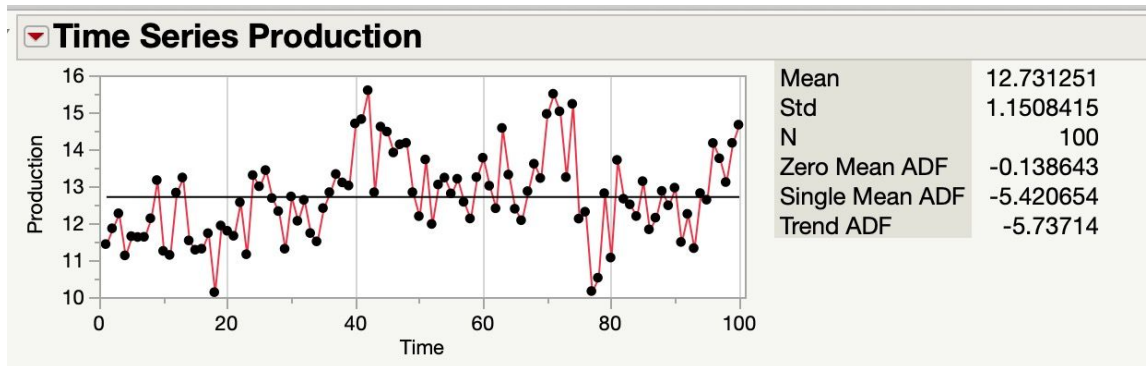


Figure 1

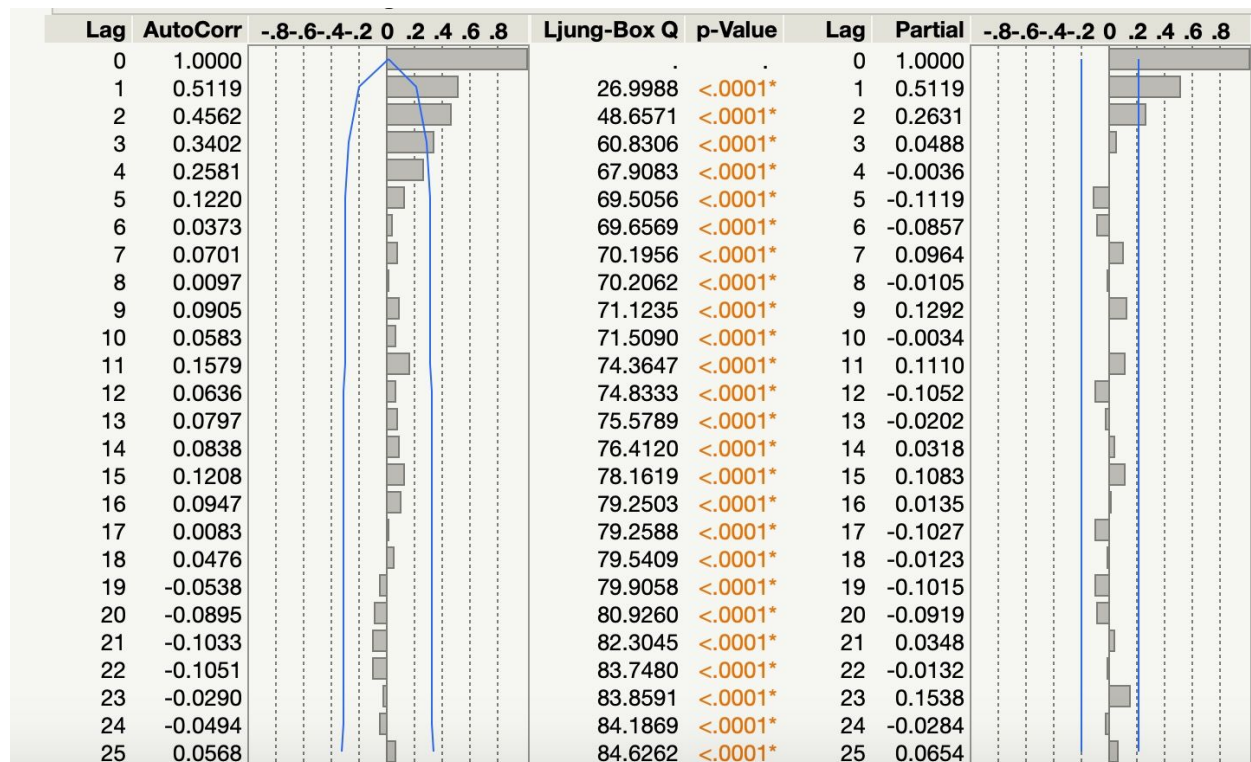


Figure 2

b.) Fit an ARIMA(p,d,q) model to the production by ML estimate methods. (Hint: choose the model which has the minimum aic.) Specify the model, i.e. don't forget the coefficients and intercept. (10 points)

Table 1

Model Comparison															
Report	Graph	Model	DF	Variance	AIC	SBC	RSquare	-2LogLH	Weights	.2	.4	.6	.8	MAPE	MAE
<div><div></div><div></div></div>	<input type="checkbox"/>	AR(2)	97	0.914854	278.34377	286.15929	0.325	272.34377	0.278507	<div><div></div><div></div><div></div><div></div></div>				6.234054	0.788601
<div><div></div><div></div></div>	<input type="checkbox"/>	ARMA(1, 1)	97	0.9234249	279.25024	287.06575	0.319	273.25024	0.177011	<div><div></div><div></div><div></div><div></div></div>				6.188474	0.782409
<div><div></div><div></div></div>	<input type="checkbox"/>	AR(3)	96	0.9223851	280.13498	290.55566	0.327	272.13498	0.113732	<div><div></div><div></div><div></div><div></div></div>				6.212497	0.785602
<div><div></div><div></div></div>	<input type="checkbox"/>	ARMA(2, 1)	96	0.9225321	280.15040	290.57108	0.327	272.1504	0.112858	<div><div></div><div></div><div></div><div></div></div>				6.216114	0.786062
<div><div></div><div></div></div>	<input type="checkbox"/>	ARIMA(2, 1, 2)	94	0.9166869	280.17607	293.15167	0.325	270.17607	0.111419	<div><div></div><div></div><div></div><div></div></div>				6.236817	0.789325
<div><div></div><div></div></div>	<input type="checkbox"/>	ARMA(2, 2)	95	0.921483	281.05349	294.07934	0.334	271.05349	0.071850	<div><div></div><div></div><div></div><div></div></div>				6.260308	0.791883
<div><div></div><div></div></div>	<input type="checkbox"/>	IMA(1, 1)	97	0.9790051	281.17802	286.36826	0.273	277.17802	0.067513	<div><div></div><div></div><div></div><div></div></div>				6.274145	0.790651
<div><div></div><div></div></div>	<input type="checkbox"/>	ARIMA(1, 1, 1)	96	0.9890568	283.14727	290.93263	0.273	277.14727	0.025221	<div><div></div><div></div><div></div><div></div></div>				6.293290	0.793089
<div><div></div><div></div></div>	<input type="checkbox"/>	AR(1)	98	0.9822435	284.30294	289.51328	0.270	280.30294	0.014152	<div><div></div><div></div><div></div><div></div></div>				6.303655	0.797538
<div><div></div><div></div></div>	<input type="checkbox"/>	AR(2, 1)	96	1.0016498	284.34006	292.12542	0.264	278.34006	0.013892	<div><div></div><div></div><div></div><div></div></div>				6.353501	0.800693
<div><div></div><div></div></div>	<input type="checkbox"/>	ARI(1, 1)	97	1.0155273	284.67872	289.86896	0.246	280.67872	0.011728	<div><div></div><div></div><div></div><div></div></div>				6.539766	0.826014
<div><div></div><div></div></div>	<input type="checkbox"/>	MA(2)	97	1.0111499	288.12373	295.93924	0.257	282.12373	0.002095	<div><div></div><div></div><div></div><div></div></div>				6.361182	0.805958
<div><div></div><div></div></div>	<input type="checkbox"/>	MA(1)	98	1.1210418	297.32076	302.53110	0.169	293.32076	0.000021	<div><div></div><div></div><div></div><div></div></div>				6.533314	0.827013
<div><div></div><div></div></div>	<input type="checkbox"/>	I(1)	98	1.262934	305.05506	307.65018	0.053	303.05506	0.000000	<div><div></div><div></div><div></div><div></div></div>				6.972124	0.881097
<div><div></div><div></div></div>	<input type="checkbox"/>	ARIMA(0, 0, 0)	99	1.3378144	313.88639	316.49156	-0.00	311.88639	0.000000	<div><div></div><div></div><div></div><div></div></div>				7.143400	0.905056

Table 2

Model: AR(2)							
Model Summary							
DF		97	Stable	Yes			
Sum of Squared Errors		88.7408339	Invertible	Yes			
Variance Estimate		0.91485396					
Standard Deviation		0.95647998					
Akaike's 'A' Information Criterion		278.343775					
Schwarz's Bayesian Criterion		286.159285					
RSquare		0.32548413					
RSquare Adj		0.31157659					
MAPE		6.23405442					
MAE		0.78860093					
-2LogLikelihood		272.343775					
Parameter Estimates							
Term	Lag	Estimate	Std Error	t Ratio	Prob> t	Constant	Estimate
AR1	1	0.385786	0.0961087	4.01	0.0001*	4.26387237	
AR2	2	0.279732	0.0969573	2.89	0.0048*		
Intercept	0	12.747696	0.2740860	46.51	<.0001*		

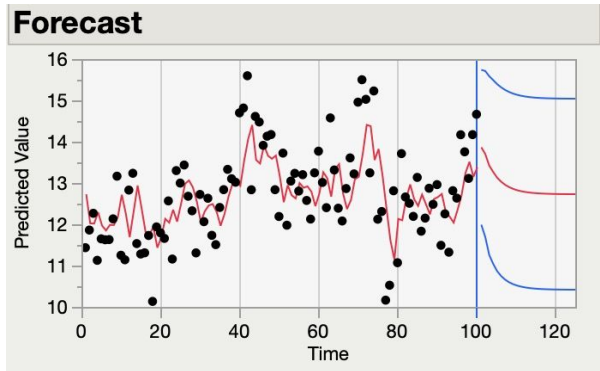


Figure 3

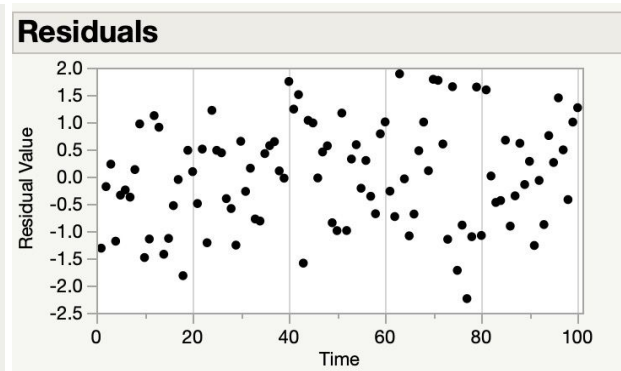


Figure 4

Table 3

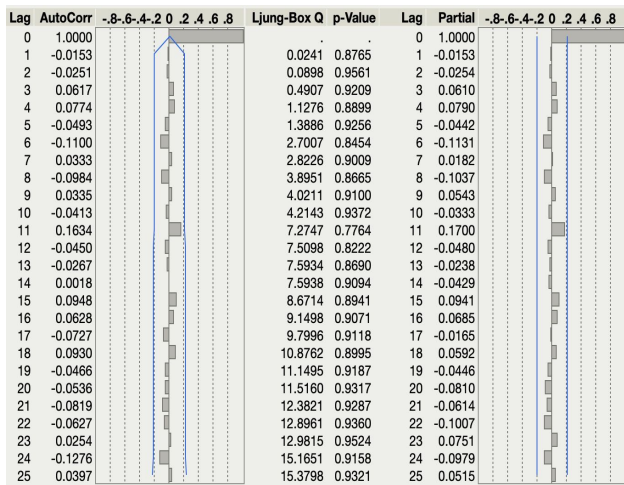


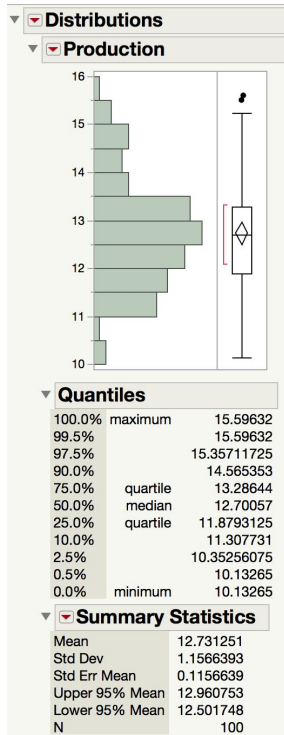
Figure 5

Iteration History				
Iter	Iteration History	Step	Delta-Criterion	Obj-Criterion
0	244.34692124	Initial	.	1.35205381
0	244.33597112	BFGS	.	1.27483219
1	244.32607771	BFGS	.	1.20249666
2	244.24604115	BFGS	.	0.11164698
3	244.245808	BFGS	.	0.09361527
4	244.24554343	BFGS	.	0.07560436
5	244.24525394	BFGS	.	0.05162916
6	244.2451319	BFGS	.	0.03042582
7	244.24511067	BFGS	.	0.0222385
8	244.24510738	BFGS	.	0.02065212
9	244.24510302	BFGS	.	0.01821158
10	244.24509605	BFGS	.	0.01319606
11	244.24509006	BFGS	.	0.00674992
12	244.24508786	BFGS	.	0.00214985
13	244.2450876	BFGS	.	0.00041853
14	244.24508759	BFGS	.	0.00004214
15	244.24508759	BFGS	.	2.39696e-6
16	244.24508759	BFGS	.	6.44554e-7
17	244.24508759	BFGS	.	3.80882e-7
18	244.24508759	BFGS	.	5.77425e-7

We used the AR(2) model because it had the lowest minimum AIC. The model is:

$$y_t = 4.263872 + 0.38578 y_{t-1} + 0.27973 y_{t-2}$$

c.) Use MOM (Method of Moments) to estimate the coefficients of the model that you selected in part b). (Hint: Calculate the coefficients by hand. Type your results. You can use the acf and pacf function in R to obtain the corresponding estimators.) (10 points)



```
ex1 <- read.csv("/Users/AliLacey/Desktop/TimeSeries/ex1.csv")
head(ex1)
var(ex1$production)
mean(ex1$production)
```

Table 4

```
> var(ex1$production)
[1] 1.337814
> mean(ex1$production)
[1] 12.73125
>
```

Figure 6

$$\gamma_{(0)} = 1.337$$

$$\hat{\rho}_{(1)} = 0.5119$$

$$\hat{\rho}_{(2)} = 0.4562$$

$$0.5119 = \Phi_{(1)} + \Phi_{(2)}(0.44562)$$

$$0.4562 = \Phi_{(2)}(0.5119) + \Phi_{(2)}$$

$$\Phi_{(1)} = 0.3772$$

$$\Phi_{(2)} = 0.2631$$

$$\bar{y} = 12.7313$$

$$\delta = 12.7313(1 - 0.3772 - 0.2631)$$

$$\delta = 4.579$$

$$y_{(t)} = 4.579 + 0.377 y_{(t-1)} + 0.263 y_{(t-2)} + \varepsilon_t$$

d.) Use the model in part b) to obtain the fitted values. (10 points)

Table 5

Predicted Production
12.747695854
12.043793423
12.037704854
12.31363359
11.987934602
11.870118103
12.006873385
12.002374849
12.197257926
12.736224438

e.) Calculate the residuals. But don't print all values, list only the first 10 residuals. (5 points)

Table 6

Residual Production
-1.314195854
-0.183553423
0.2283451458
-1.18608359
-0.340254602
-0.245098103
-0.377083385
0.1291151512
0.9675120738
-1.486014438

f.) Analyze the residuals. Your analysis should include: Box-Ljung test and interpretation (5 points), ACF, PACF, and their interpretation (5 points) and 4-in-1 plot of residuals and interpretation (5 points).

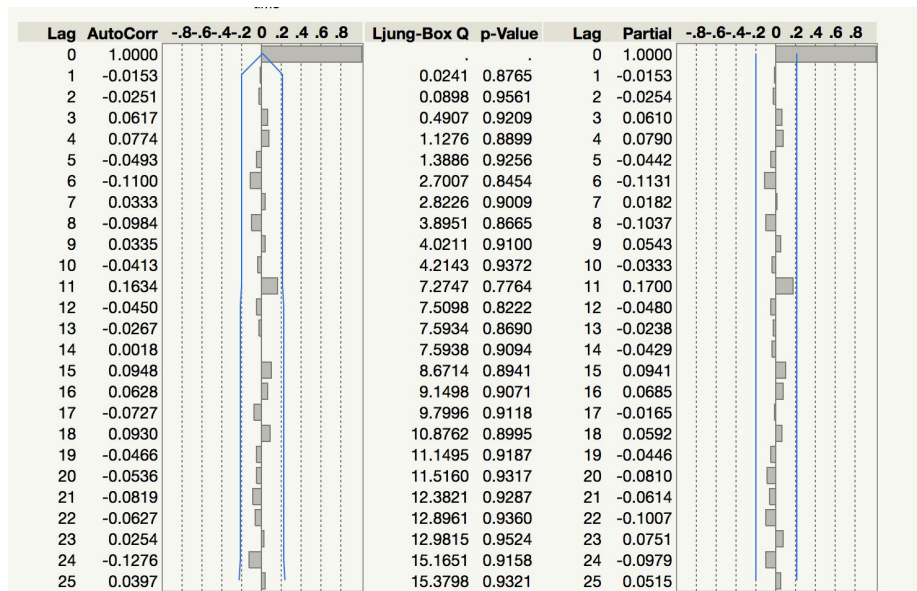


Figure 7

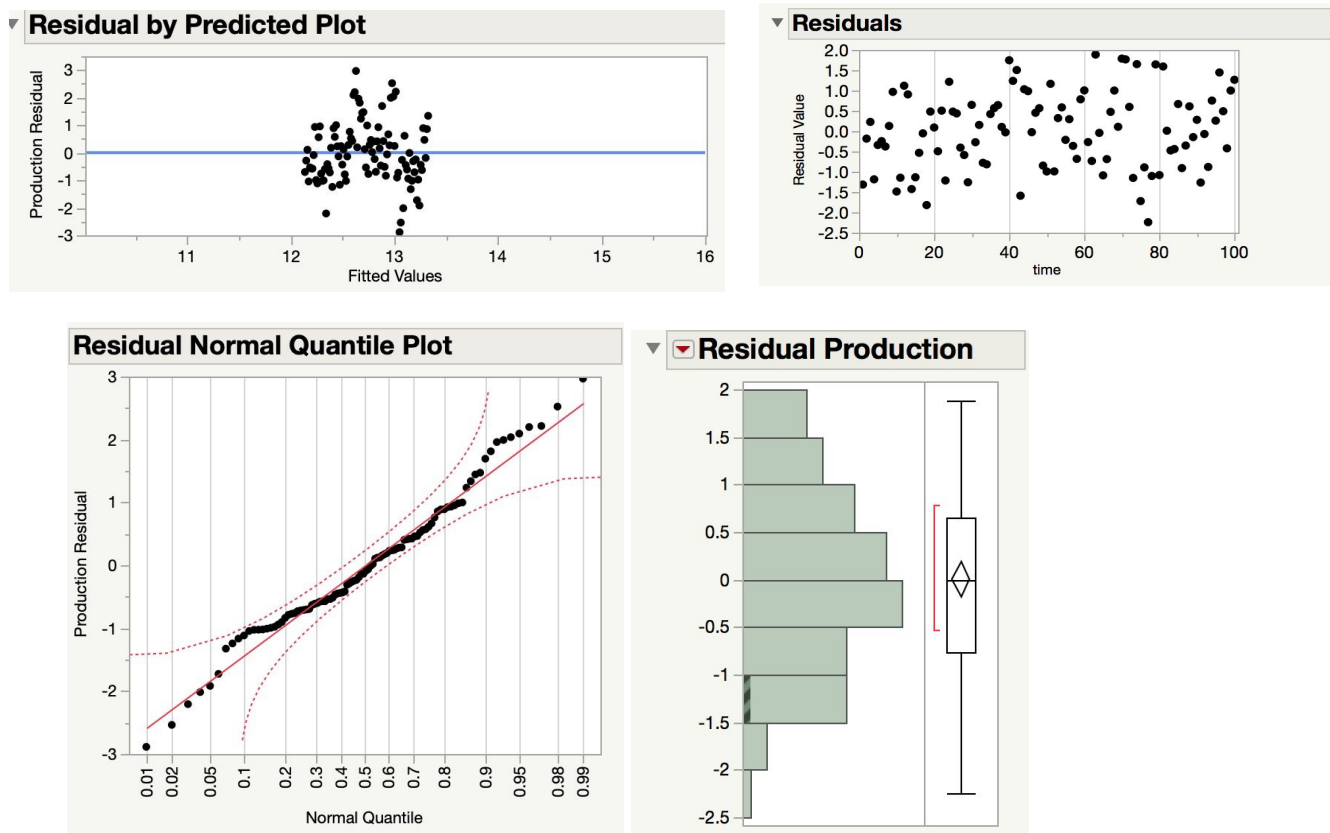


Figure 8

In Figure 8, we can see a normal distribution of residuals in both the histogram and the normal quantile plot. The ACF and PACF for each lag are within the confidence interval and fluctuates around 0, between -0.2 and 0.2. Also, with the Residual by row plot we can finally conclude the time series is stationary.

The Ljung–Box test is uncommonly used in (ARIMA) modeling. Note that it is applied to the residuals of a fitted ARIMA model, not the original series, and in such applications the hypothesis actually being tested is that the residuals from the ARIMA model have no autocorrelation. Hence from Figure 7, none of the Ljung–Box Q values have significant p-values. Therefore we accept the hypothesis that residuals from the ARIMA model have no autocorrelation.

g.) Plot the observations and fitted values in the same graph. Are you satisfactory with your model? (5 points)

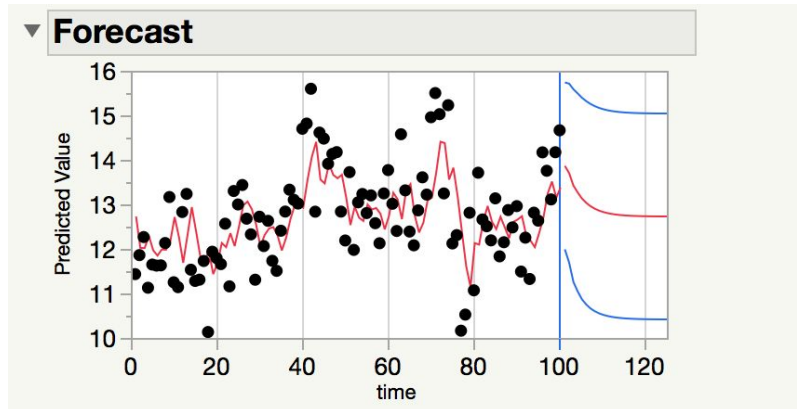


Figure 9

Yes, we are satisfied with our model. The observations seem to follow the fitted values.

2. Analyze the data ex3. See attached.

a.) Plot the data y in ex3, and its ACF and PACF. (10 points)

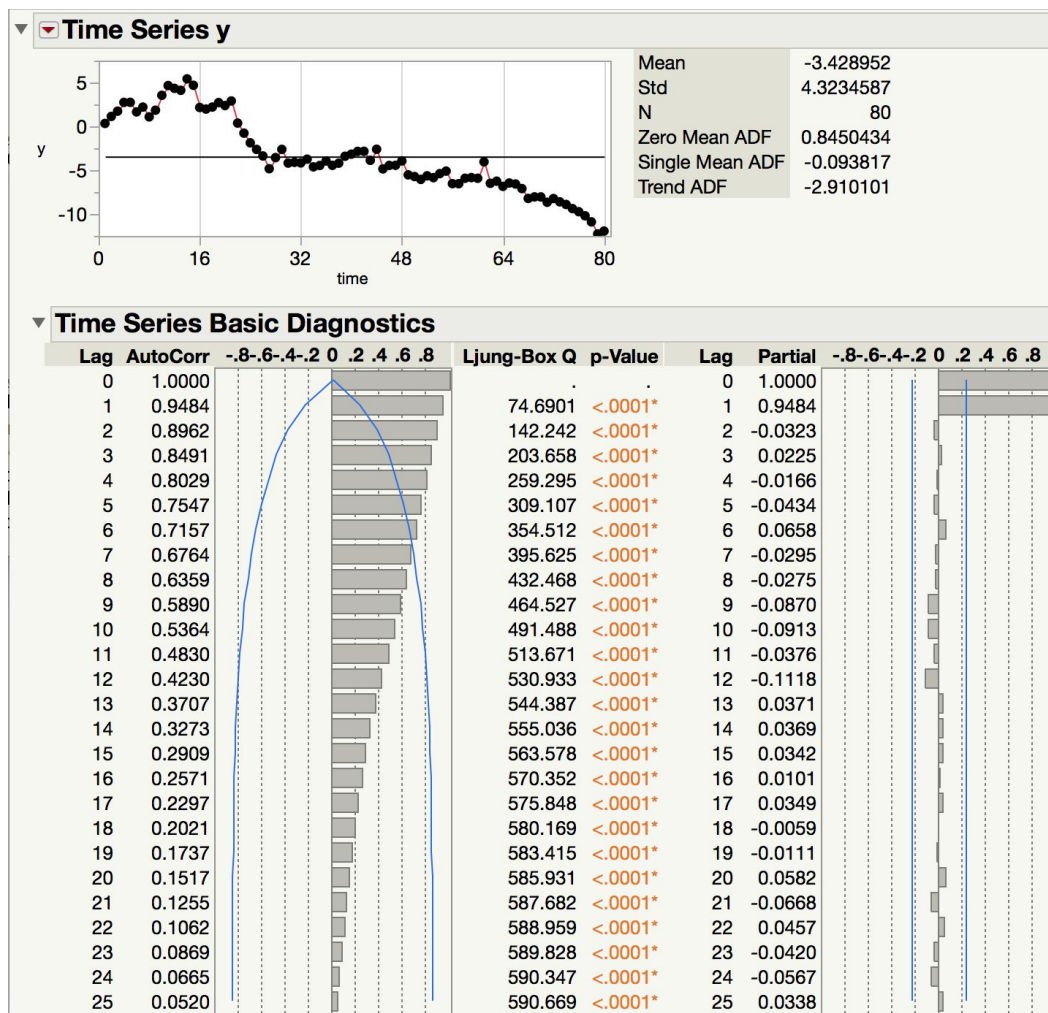


Figure 10

b.) Plot the first difference of y and its ACF and PACF. Interpret your outputs. (10 points)

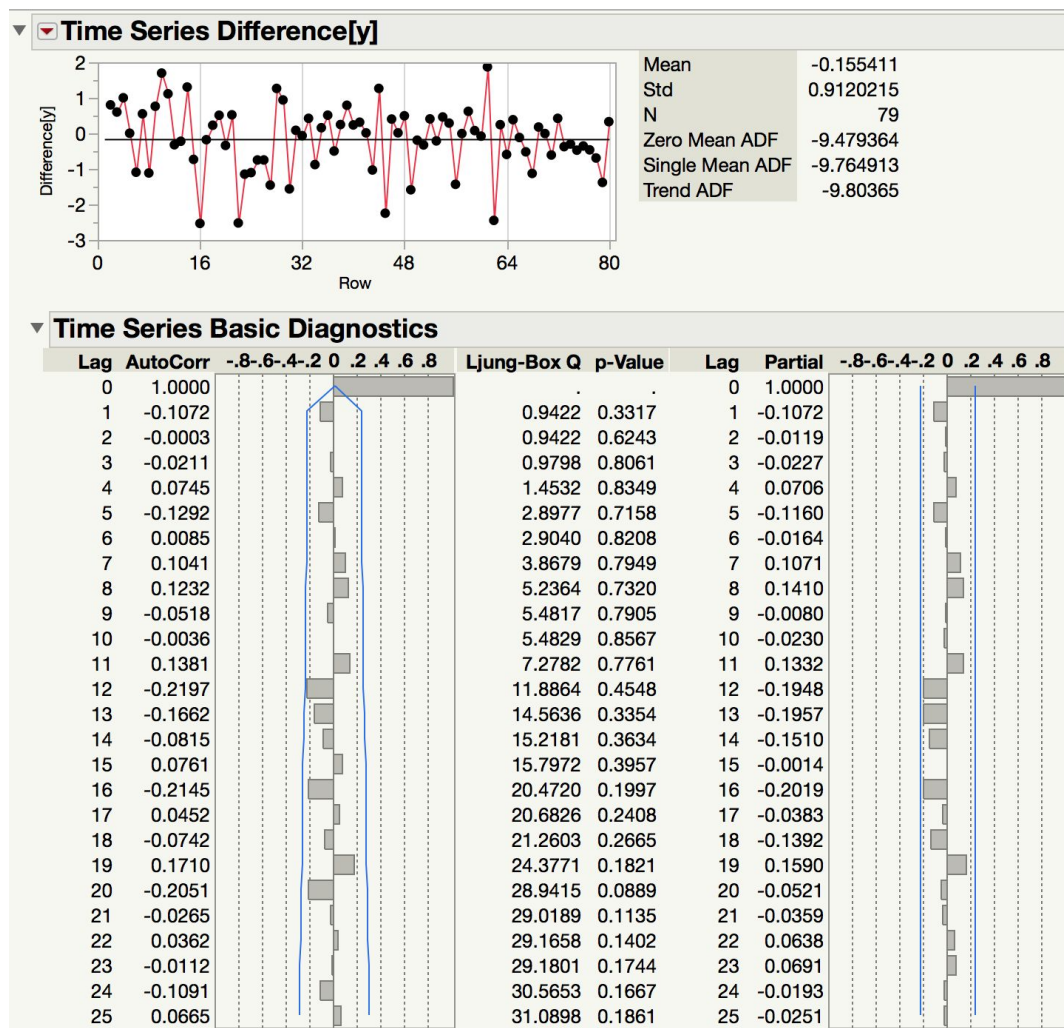


Figure 11

The first difference transforms the initial non stationary times series into a stationary time series, all the ACF and PACF values lie between the upper and lower limits.

c.) Fit an ARIMA(p,d,q) model to the y by ML estimate methods. Specify the model and its coefficients. (10 points)

Model Comparison

Report	Graph	Model	DF	Variance	AIC	SBC	RSquare	-2LogLH	Weights	.2	.4	.6	.8	MAPE	MAE
		I(1)	78	0.8424471	211.64180	214.01125	0.956	209.6418	0.321567					28.859894	0.708050
		IMA(1, 1)	77	0.843235	212.70841	217.44731	0.956	208.70841	0.188651					29.078591	0.703738
		ARI(1, 1)	77	0.84341	212.72434	217.46324	0.956	208.72434	0.187154					29.067227	0.703497
		ARIMA(2, 1, 2)	74	0.7820896	213.04459	224.89183	0.959	203.04459	0.159463					27.487774	0.693176
		ARIMA(1, 1, 1)	76	0.8543134	214.70693	221.81528	0.956	208.70693	0.069452					29.082631	0.704060
		ARI(2, 1)	76	0.8543702	214.71198	221.82033	0.956	208.71198	0.069277					29.082800	0.703833
		AR(1)	78	0.8732183	221.68141	226.44546	0.937	217.68141	0.002124					48.107917	0.761075
		AR(2)	77	0.8806664	223.37076	230.51684	0.936	217.37076	0.000913					48.607339	0.759368
		ARMA(1, 1)	77	0.8809226	223.39233	230.53841	0.936	217.39233	0.000903					48.590004	0.759742
		ARMA(2, 2)	75	0.8552723	224.58811	236.49824	0.937	214.58811	0.000497					49.130474	0.772073
		I(2)	77	1.874702	271.36702	273.72373	0.901	269.36702	0.000000					38.500909	1.029355
		MA(2)	77	3.3125841	328.02072	335.16680	0.820	322.02072	0.000000					54.869477	1.529641
		MA(1)	78	6.654127	381.97421	386.73826	0.644	377.97421	0.000000					72.169920	2.127034

Figure 12

Model: ARIMA(2, 1, 2)

Model Summary

DF	74	Stable	Yes
Sum of Squared Errors	57.874629	Invertible	Yes
Variance Estimate	0.78208958		
Standard Deviation	0.88435829		
Akaike's 'A' Information Criterion	213.044586		
Schwarz's Bayesian Criterion	224.891826		
RSquare	0.95896877		
RSquare Adj	0.95675087		
MAPE	27.4877742		
MAE	0.69317615		
-2LogLikelihood	203.044586		

Parameter Estimates

Term	Lag	Estimate	Std Error	t Ratio	Prob> t	Constant Estimate
AR1	1	1.536987	0.0712308	21.58	<.0001*	-0.0452739
AR2	2	-0.820221	0.0853553	-9.61	<.0001*	
MA1	1	1.739716	0.0566906	30.69	<.0001*	
MA2	2	-0.999960	0.0551029	-18.15	<.0001*	
Intercept	0	-0.159846	0.0879117	-1.82	0.0731	

Model: IMA(1, 1)

Model Summary

DF	77	Stable	Yes
Sum of Squared Errors	64.9290947	Invertible	Yes
Variance Estimate	0.843235		
Standard Deviation	0.91827828		
Akaike's 'A' Information Criterion	212.708415		
Schwarz's Bayesian Criterion	217.44731		
RSquare	0.95615336		
RSquare Adj	0.95558392		
MAPE	29.0785906		
MAE	0.70373751		
-2LogLikelihood	208.708415		

Parameter Estimates

Term	Lag	Estimate	Std Error	t Ratio	Prob> t	Constant Estimate
MA1	1	0.1098417	0.1128434	0.97	0.3334	-0.1573595
Intercept	0	-0.1573595	0.0909428	-1.73	0.0876	

Model: ARI(1, 1)

Model Summary

DF	77	Stable	Yes
Sum of Squared Errors	64.9425688	Invertible	Yes
Variance Estimate	0.84340998		
Standard Deviation	0.91837355		
Akaike's 'A' Information Criterion	212.724345		
Schwarz's Bayesian Criterion	217.46324		
RSquare	0.95614453		
RSquare Adj	0.95557498		
MAPE	29.0672265		
MAE	0.70349659		
-2LogLikelihood	208.724345		

Parameter Estimates

Term	Lag	Estimate	Std Error	t Ratio	Prob> t	Constant Estimate
AR1	1	-0.1077329	0.1121228	-0.96	0.3396	-0.1741245
Intercept	0	-0.1571899	0.0921966	-1.70	0.0922	

Model: I(1)

Model Summary

DF	78	Stable	Yes
Sum of Squared Errors	65.7108765	Invertible	Yes
Variance Estimate	0.84244713		
Standard Deviation	0.91784919		
Akaike's 'A' Information Criterion	211.6418		
Schwarz's Bayesian Criterion	214.011248		
RSquare	0.95563299		
RSquare Adj	0.95563299		
MAPE	28.8598943		
MAE	0.70805012		
-2LogLikelihood	209.6418		

Parameter Estimates

Term	Lag	Estimate	Std Error	t Ratio	Prob> t	Constant Estimate
Intercept	0	-0.1554110	0.1025851	-1.51	0.1338	-0.155411

Figure 13

The model chosen is ARIMA(2,1,2) since this model has the best fit under the maximum likelihood function (minimum -2LogLikelihood). All the parameter estimates can be seen significant for ARIMA(2,1,2), whereas in Figure 12, there are three models which have lower AIC than ARIMA(2,1,2), but from Figure 13 we see that these models do not have significant parameters. The ARIMA(2,1,2) model equation can be written as

$$y_t = -0.045273 + 1.53698 y_{t-1} - 0.8202 y_{t-2} + 1.73971 \varepsilon_{t-1} - 0.9999 \varepsilon_{t-2}$$

d.) Plot a realization of your fitted model (you will need the outputs in part c)), also show its ACF and PACF. Compare them with the results in part a). (10 points)

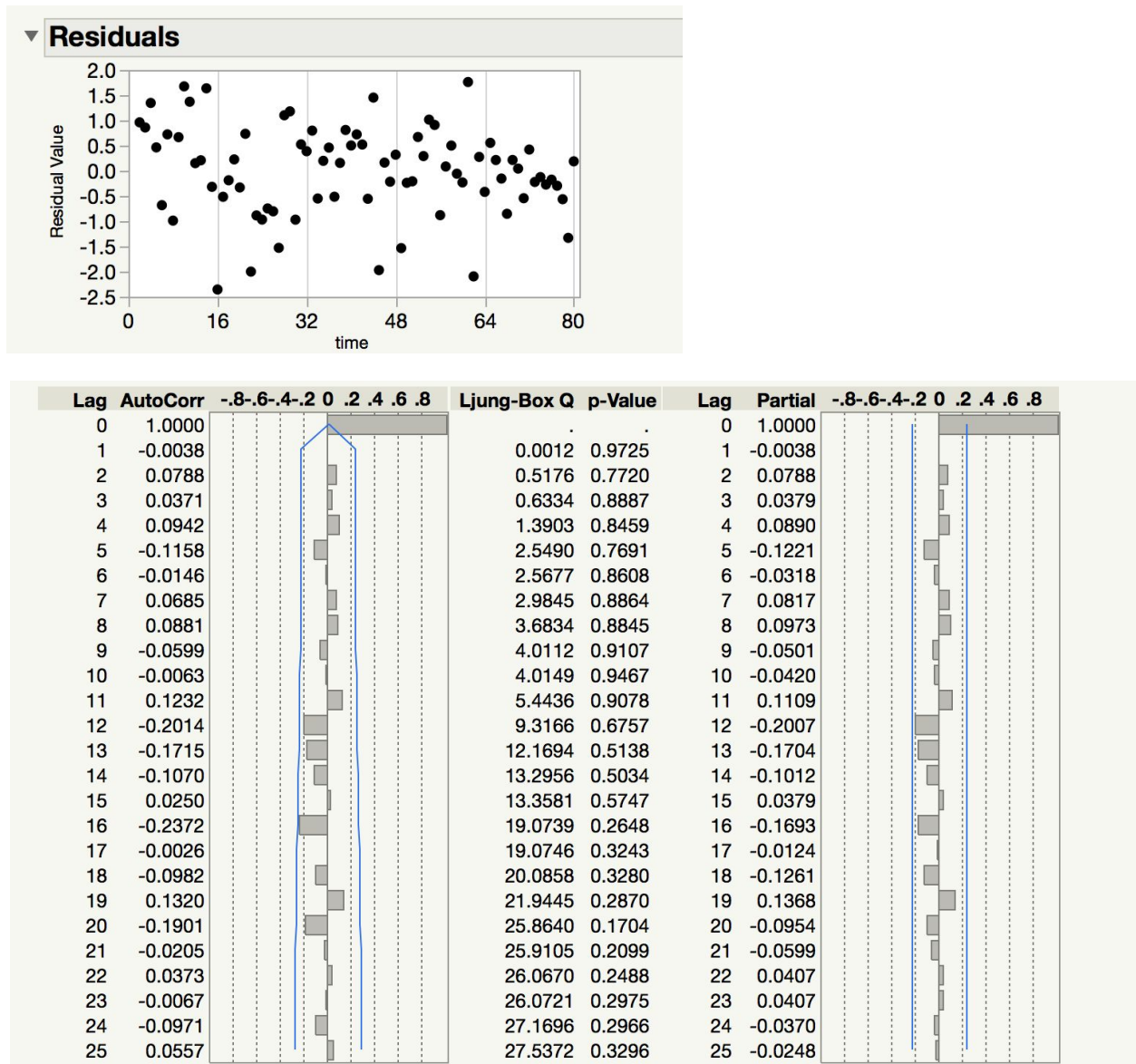


Figure 14

The non stationary time series in part a is transformed to a stationary time series. The model is stable with all ACF and PACF values in between the upper and lower limits.