UNIVERSITY OF LONDON

GOLDSMITHS COLLEGE

B.Sc. Examination 2003

Mathematics et al

ST317 Time Series Analysis and Forecasting

Duration: 2 hours 15 minutes

Date and time:

- *Full marks will be awarded for complete answers to FOUR questions. Do not attempt more than FOUR questions on this paper.*
- *Electronic calculators may be used. The make and model should be specified on the script. The calculator must not be programmed prior to the examination. Calculators which display graphics, text or algebraic equations are not allowed.*

THIS EXAMINATION PAPER MUST NOT BE REMOVED FROM THE EXAMINATION ROOM

Question 1.

a) The plot below shows observations of half-hourly electricity demand data for the UK for a period of just over 5 days. Describe the pattern that you observe. **[5]**

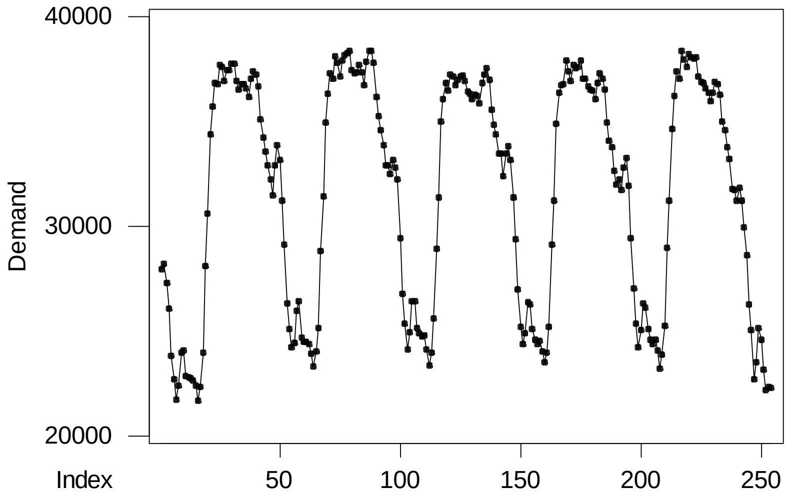


Figure 1: UK Demand Data (254 half hours)

b)- An analyst used 3360 observations from this time series and produced two forecasting models: a neural network and a Holt-Winters. For the latter he used Minitab 13 and part of his output is given below.

Winters' additive model

```
Data         Demand (MW)
Length       3360.00
NMissing     0

Smoothing Constants
Alpha (level):    0.20
Gamma (trend):    0.18
Delta (seasonal): 0.10

Accuracy Measures
MAPE:       4
MAD:     1101
MSD:  1822861
```

Summarise this output in your own words. **[5]**

c)- Why would the analyst choose the above type of model for this time series? What are the advantages and disadvantages of exponential smoothing methods? **[9]**

d) For the same period, the neural-network model gave the accuracy measures that are shown in the table below:

NeuralNet Fit Results (n=3360)

Mean Error	Mean Square Error	Mean Percentage Error	Mean Absolute Percentage Error
536.9995	1303063	0.94	3.16

Which model resulted in the best fit? Why? **[6]**

Question 2.

We consider the UK half-hourly electricity demand data, which is partly shown in figure 1, and decompose this time series (3360 observations or 70 days) using an additive model in Minitab. Part of the output is shown below.

Time Series Decomposition

```
Data        Demand (MW)
Length      3360.00
NMissing    0

Trend Line Equation

Yt = 36259.3 - 2.25225*t

Seasonal Indices
Period    Index
   1      -7593.50
   2      -6260.59
   3      -5003.91
   4      -4759.82
   5      -4759.03
   6      -5194.25
   7      -5502.84
   8      -5907.08
   9      -6664.67
  10      -6771.25
  11      -6501.22
  12      -5033.54
  13      -3077.11
  14       215.192
  15       2463.38
  16       3842.47
  17       3864.67
  18       4321.68
  19       4421.86
  20       4396.19
  21       4379.61
  22       4473.21
  23       4584.93
  24       4601.94
  25       4199.86
  26       3767.57
  27       3530.90
  28       3526.53
  29       3544.25
  30       3189.67
  31       3319.35
  32       3522.34
  33       3589.83
  34       3528.58
  35       2859.21
  36       2185.43
  37       1685.64
  38       1207.92
  39       973.442
  40       1055.78
  41       1534.62
  42       1836.83
  43       922.338
  44      -22.6832
  45      -1754.30
  46      -3807.08
  47      -5698.63
  48      -7233.74

Accuracy of Model
MAPE:            8
MAD:          2597
MSD:      10963591
```

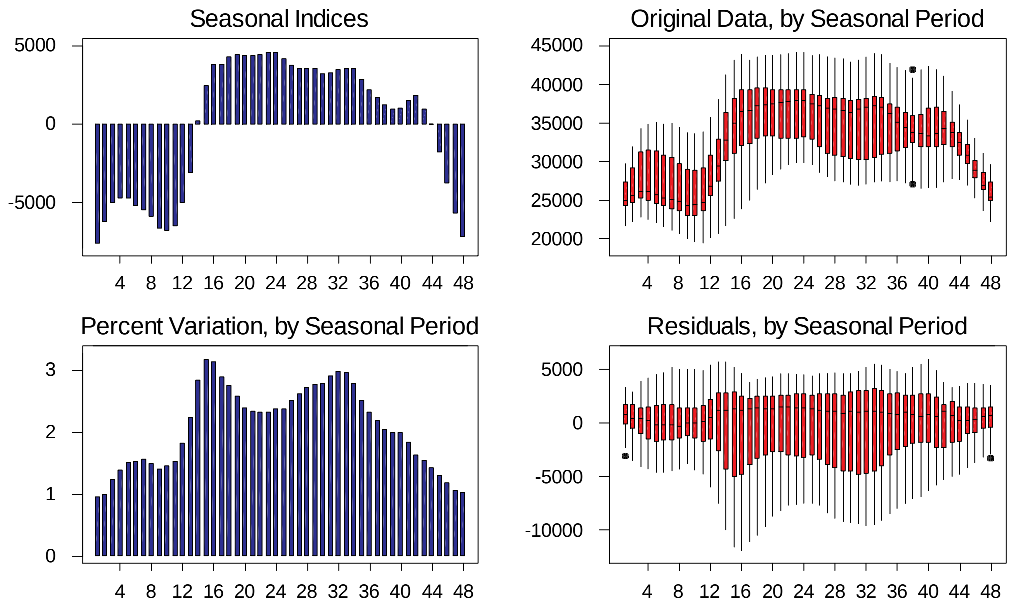
a)- Use the above information to produce forecasts for the next six half-hours (the first six half hours of the seventy first day). **[5]**

b)- The table below gives the actual demand and forecasts from the neural network model for the same period. Given these results, which of the two models (decomposition, neural network performed best? **[8]**

Time	Demand(MW)	Neural Net Forecast
3361	33529	33656
3362	35428	35227.7
3363	36034	36728.8
3364	36834	36699.7
3365	37296	37677.6
3366	37338	37817

c)- The graph below describes the seasonal pattern. Briefly describe what you observe. Is this an expected pattern for this type of data (energy demand, consumption)? **[5]**

Seasonal Analysis for Demand (MW)



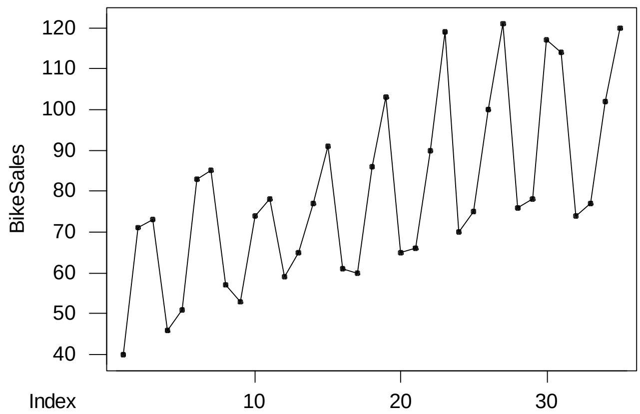
d)- A forecasting problem faced by those who are in charge of providing electricity demand forecasts is associated with especial days or events that may change the demand profile.

Briefly describe the role of judgement and expertise in such situations. Which kind of problems can one expect from judgemental forecasts? **[7]**

Question 3.

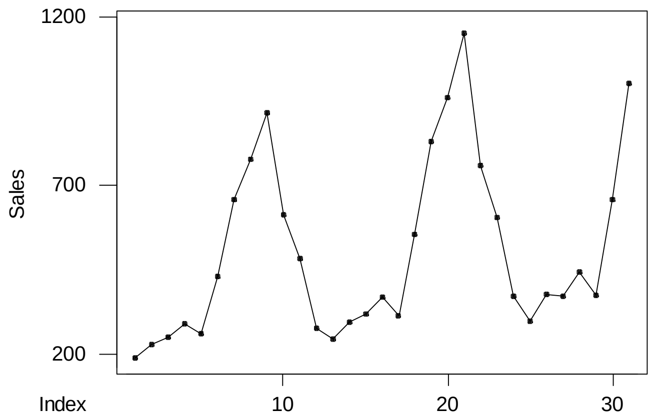
a)- Below are plots of different time series. Briefly describe how would you proceed in order to produce forecasts for these data: which forecasting method would you use? Justify your answers, assumptions and choices.

i)



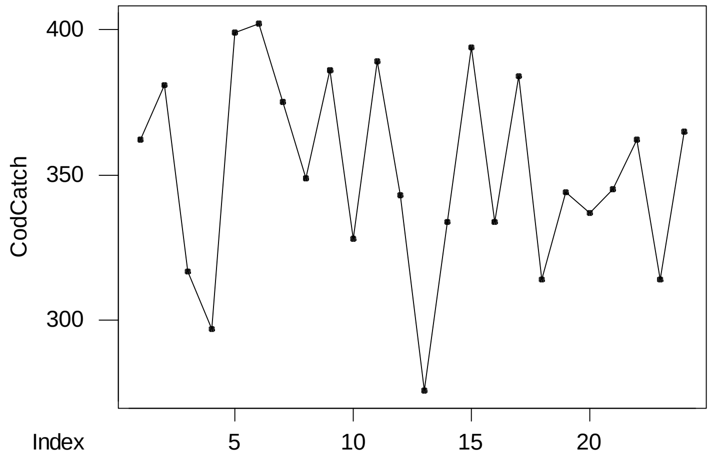
[5]

ii)



[5]

iii)

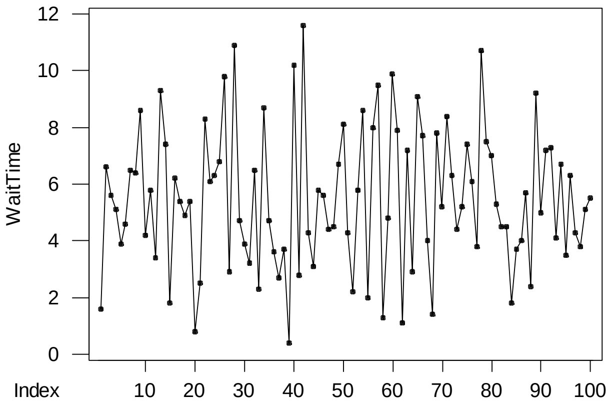


[5]

b) Why is cause and effect difficult to prove in business and economics? What is implied by the correlation coefficient and which values can it take? What can be verified by building a regression model, what are its standard assumptions in a regression model? Which of these are unlikely to hold in time series data? **[10]**

Question 4.

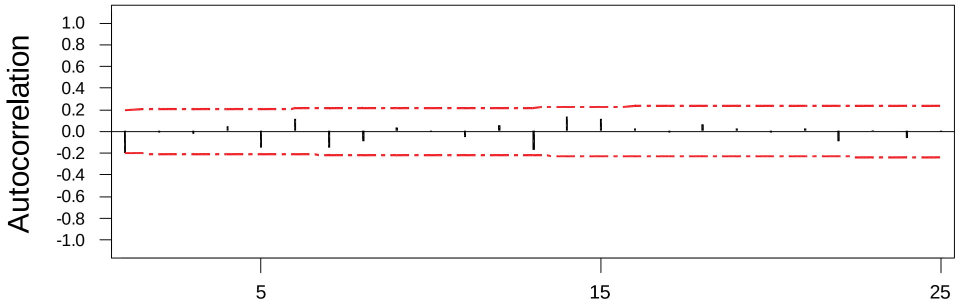
The data plotted below are the average daily waiting times in a 24-hour supermarket till. These are recorded in minutes.



a)- Describe the pattern that you observe. **[5]**

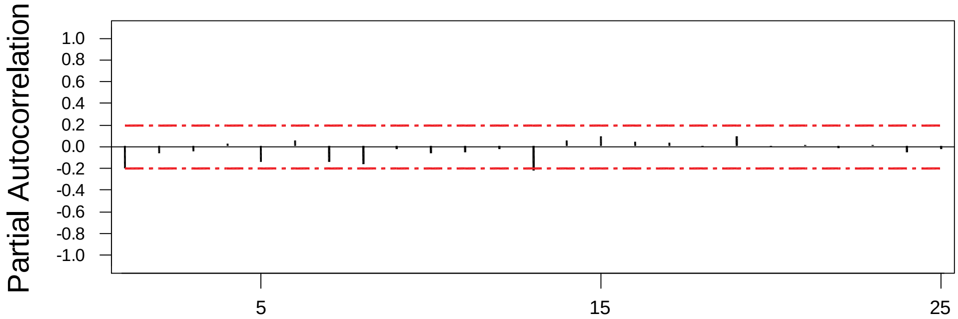
b)- According to the autocorrelation and partial autocorrelation functions that were computed and plotted using Minitab, as shown below,

Autocorrelation Waiting Time



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0.20	-2.02	4.22	8	-0.09	-0.81	11.93	15	0.12	1.00	19.74	22	-0.09	-0.78	21.56
2	-0.02	-0.16	4.25	9	0.03	0.30	12.05	16	0.02	0.20	19.81	23	0.00	0.00	21.56
3	-0.03	-0.25	4.32	10	-0.01	-0.09	12.06	17	-0.02	-0.14	19.84	24	-0.06	-0.52	22.05
4	0.04	0.42	4.52	11	-0.05	-0.47	12.37	18	0.06	0.52	20.29	25	-0.01	-0.05	22.06
5	-0.15	-1.48	7.07	12	0.05	0.45	12.67	19	0.02	0.19	20.35				
6	0.11	1.03	8.38	13	-0.17	-1.51	15.97	20	-0.02	-0.14	20.39				
7	-0.16	-1.45	11.04	14	0.13	1.19	18.13	21	0.02	0.20	20.46				

Partial Autocorrelation Function for WaitTime

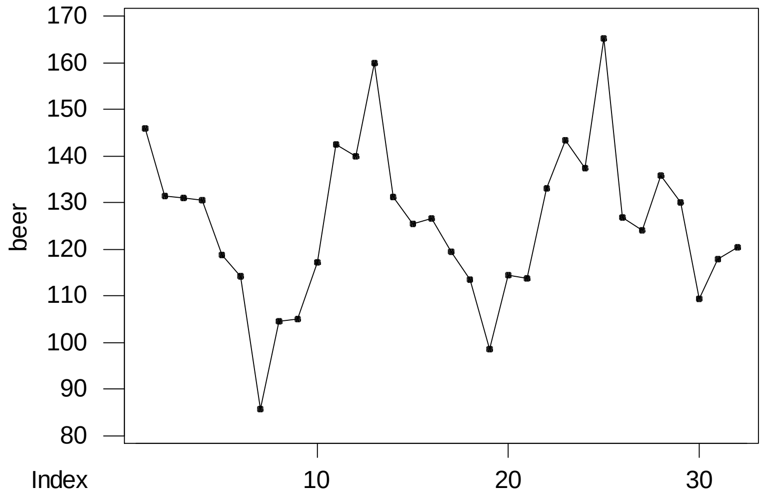


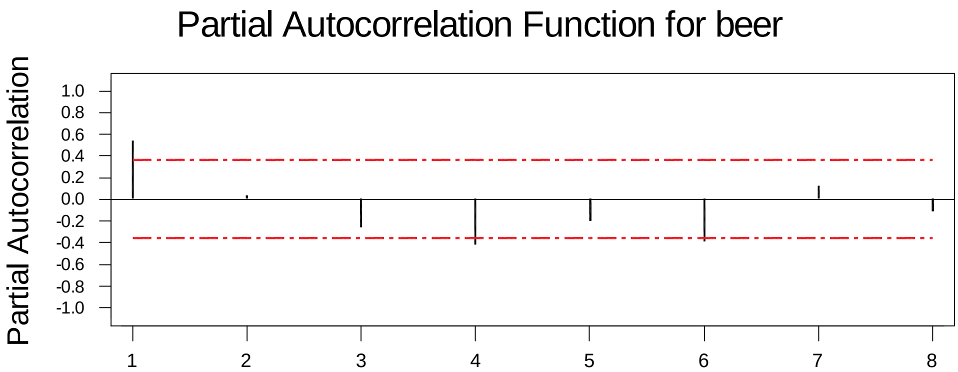
Lag	PAC	T	Lag	PAC	T	Lag	PAC	T	Lag	PAC	T
1	-0.20	-2.02	8	-0.16	-1.64	15	0.09	0.93	22	-0.02	-0.16
2	-0.06	-0.61	9	-0.03	-0.26	16	0.04	0.40	23	0.02	0.16
3	-0.04	-0.44	10	-0.06	-0.64	17	0.03	0.32	24	-0.05	-0.52
4	0.03	0.29	11	-0.06	-0.57	18	-0.00	-0.03	25	-0.02	-0.22
5	-0.15	-1.48	12	-0.02	-0.24	19	0.09	0.92			
6	0.05	0.53	13	-0.22	-2.17	20	-0.01	-0.09			
7	-0.14	-1.44	14	0.05	0.50	21	0.02	0.17			

i)- what can be said of this time series? **[5]**
ii)- how would you proceed if you were asked to produce forecasts of these data? **[15]**

Question 5.

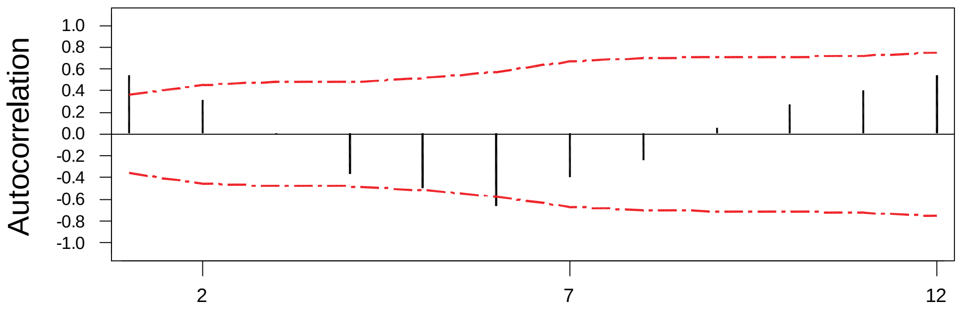
An analyst was studying the consumption of beer in his local pub during a period of just over a month. He collected the daily consumption (in l.) and ran an analysis using Minitab. He produced the following output.





Lag	PAC	T
1	0.53	3.03
2	0.04	0.20
3	-0.26	-1.48
4	-0.42	-2.37
5	-0.20	-1.14
6	-0.39	-2.20
7	0.12	0.69
8	-0.11	-0.62

Autocorrelation Function for beer



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.53	3.03	10.04	8	-0.24	-0.69	57.25
2	0.31	1.41	13.56	9	0.05	0.16	57.40
3	-0.01	-0.03	13.56	10	0.27	0.79	61.09
4	-0.37	-1.57	18.84	11	0.40	1.12	69.31
5	-0.50	-1.99	29.00	12	0.54	1.47	85.29
6	-0.67	-2.37	47.74				
7	-0.40	-1.22	54.68				

ARIMA Model: beer

ARIMA model for beer

Estimates at each iteration

Iteration	SSE	Parameters	
0	7774.73	0.100	112.961
1	6821.10	0.250	94.105
2	6235.27	0.400	75.278
3	6019.91	0.539	57.924
4	6013.90	0.559	55.460
5	6013.65	0.563	54.994
6	6013.64	0.564	54.897
7	6013.64	0.564	54.877

Relative change in each estimate less than 0.0010

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
AR 1	0.5644	0.1505	3.75	0.001
Constant	54.877	2.485	22.08	0.000
Mean	125.988	5.705		

Number of observations: 32

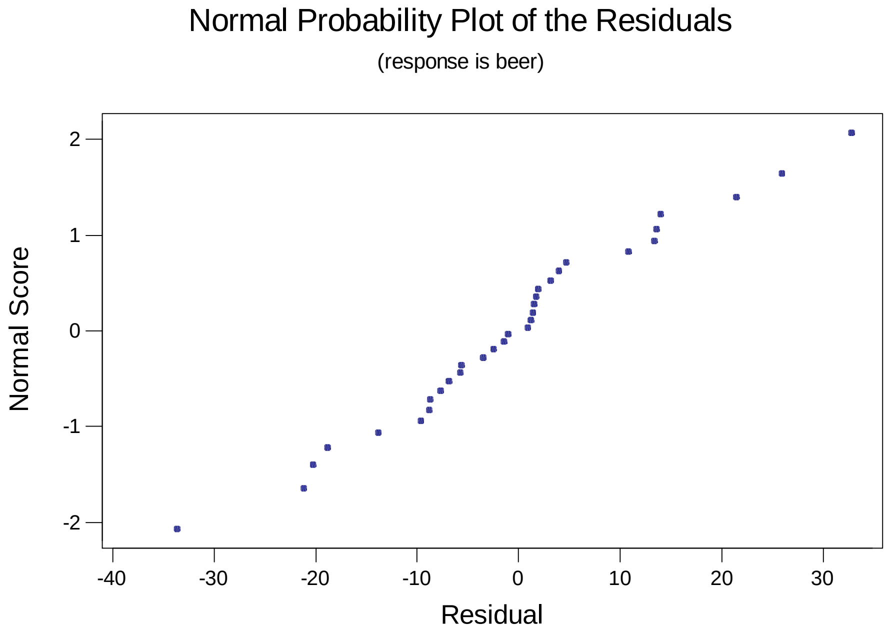
Residuals: SS = 5927.57 (backforecasts excluded)
MS = 197.59 DF = 30

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	31.3	41.7	*	*
DF	10	22	*	*
P-Value	0.001	0.007	*	*

Forecasts from period 32

		95 Percent Limits		
Period	Forecast	Lower	Upper	Actual
33	122.777	95.221	150.334	
34	124.176	92.533	155.819	
35	124.965	92.127	157.803	
36	125.411	92.201	158.620	
37	125.662	92.335	158.989	
38	125.804	92.439	159.169	
39	125.884	92.507	159.261	



a)- Describe the analysis that was undertaken, summarise the results that followed and comment on limitations. **[15]**

b)- The same analyst used an alternative model to produce a forecast for the next day, as shown in the output below. Comment on this analysis and on how this method differs from the used above. **[10]**

Single Exponential Smoothing

```
Data         beer
Length       32.0000
NMissing     0

Smoothing Constant
Alpha: 0.788346

Accuracy Measures
MAPE:    9.714
MAD:    11.920
MSD:   224.690


 Row  Period  Forecast     Lower     Upper

   1      33   119.597   90.3932   148.801
```

END OF EXAMINATION