Question 1 (20 Marks)

a) A time series can be decomposed into a number of components. List three such components and describe each one. (6 Marks)

b) Discuss the difference between additive and multiplicative seasonality. Explain how this difference affects the choice of model transformation when seasonal adjustment of the series involves a regression stage.

(7 Marks)

c) Explain how you would identify and remove the seasonal component for an additive model? (7 Marks)

Question 2 (20 Marks)

Suppose you have fitted an ARIMA(1,0,1) model to a time series and you intend to implement a residual analysis in order to assess the model's fit.

a) Write out the ARIMA(1,0,1) model for series X_t in terms of autoregressive parameters, differencing (if any), and moving average parameters. Hence obtain an expression for the residuals. (8 Marks)

b) Why and how would you use the correlogram of the residuals in order to assess the quality of the model? (6 Marks)

c) Describe other ways in which you could use residuals to assess model fit.

(6 Marks)

Question 3 (20 Marks)

From a time series x_t of 156 daily values of electricity consumption, an analyst has calculated the series of first differences $y_t = x_t - x_{t-1}$

i) Suggest a reason why the analyst, having examined a time series plot of the data, decided to analyze the first differences. (1 Mark)

ii) The analyst fitted the following model to the series y_t:

$$y_t = 0.4y_{t-1} + \varepsilon_t + 0.7\varepsilon_{t-1}$$

What assumptions are usually made about the errors, ε_t? (3 Marks)

iii) Write out the model equation for x_t in terms of past x_t and past and present ε_t only. The model for x_t can be classified as a member of the $ARIMA(p, d, q)$ family having $p = 1, d = 1$ and $q = 1$. Explain these values of p, d and q in the light of your model equation. (5 Marks)

iv) Describe why and how the analyst could use a plot of the residual autocorrelation function (i.e. the correlogram of the residuals) in order to assess the quality of the model. (4 Marks)

v) Why might the analyst also examine

a) a time series plot of the residuals,

b) a $Q - Q$ plot of the residuals? (4 Marks)

vi) As part of the computer output from the fitting routine, the analyst received the following information on portmanteau lack-of-fit tests for the residuals.

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

Lag	12	24	36	48
Chi-Square	18.7	40.2	49.5	56.9
DF	10	22	34	46
P-Value	0.045	0.010	0.041	0.130

Briefly explain what the chi-square statistic here is measuring. What conclusion should the analyst reach about the adequacy of the model that has been fitted? (3 Marks)

Question 4 (20 Marks)

i) For a time series y_t write down

a) a formula for smoothing the series using a symmetric $(2k + 1) -$ point moving average with weights $a_r \, (r = -k, \ldots, k)$ (2 Marks)

b) the equation, in simple exponential smoothing with smoothing parameter α, for calculating the smoothed estimate m_t of the series in terms of y_t and m_{t-1}. (2 Marks)

ii) The table below shows the half-yearly sales figures y_t of a certain product for the years 2011 to 2014.

Date		Sales (Kshs. 10,000)	Smoothed A	Smoothed B
2011	June	51		
	December	25		
2012	June	49		
	December	31		
2013	June	53		
	December	29		
2014	June	55		
	December	25		

Copy the table and complete the extra columns, giving (correct to 2 decimal places)

- Smoothed A: the 3-point simple symmetric moving average of y_t
- Smoothed B: the exponentially weighted moving average of y_t with smoothing parameter 0.3 and initial smoothed value set equal to the actual value for June 2011. (6 Marks)

Describe the strengths and weaknesses of each of these methods from the point of view of

a) forecasting,

b) smoothing seasonal time series. (6 Marks)

iii) Explain briefly how moving averages can be used to extract the seasonal and trend components from a time series. (4 Marks)

Question 5 (20 Marks)

a) An additive trend and seasonal decomposition for a quarterly time series $y(t)$ may be represented as

$$y(t) = T(t) + C(t) + R(t)$$

where $T(t)$ and $R(t)$ are respectively the trend component and the random or irregular component. Explain why the constraint $C(t) = C(t-4)$ should be applied to $C(t)$ and state the usual 'standardizing' condition imposed on the sum $C(1) + C(2) + C(3) + C(4)$ (3 Marks)

Analyst A applies an additive trend and seasonal decomposition to some sales data for a cake manufacturing company. The table below gives the data plus partial results of using a simple 3-point moving average to smooth the sales series along with the corresponding errors for each quarter, Q1 to Q4.

Period	**Sales($K)** $y(t)$	**Smoothed value,** $y^*(t)$	**Error**
Q1 2012	241		
Q2 2012	283	268.33	14.67
Q3 2012	281	296.33	-15.33
Q4 2012	325	297.67	27.33
Q1 2013	287	296.00	-9.00
Q2 2013	276	288.00	-12.00
Q3 2013	301	309.67	-8.67
Q4 2013	352	325.00	27.00
Q1 2014	322	335.67	-13.67
Q2 2014	333		
Q3 2014	351		
Q4 2014	410		

b) Calculate the missing smoothed values and errors for Q2 and Q3 of 2014.

(3 Marks)

[Note: Do not attempt to calculate smoothed values for the 'end-points', Q1 2012 or Q4 2014.]

c) What weakness does the choice of a 3-point moving average have when smoothing quarterly data? (1 Mark)

d) Calculate estimates of the seasonal components for Q1 to Q4. (4 Marks)

e) A linear regression model fitted to the smoothed values yields

$$y^{*}(t) = 254.99 + 8.72t$$

where t is the number of quarters from Q1 2012 with $t = 1$ for Q1 2012. Use the additive trend and seasonal model from part (a) to find A's forecasts of sales for Q3 and Q4 of 2015. (3 Marks)

f) Analyst B fits a linear regression model of $y(t)$ on t and an indicator variable for Q4, obtaining the following output.

Coefficients:

	Estimate	Std. Error
(Intercept)	242.282	8.006
t	9.157	1.122
Q4	46.798	8.947

Discuss whether the explanatory variables, t and Q4, should be retained in the model. (4 Marks)

g) Calculate forecasts of sales for Q3 and Q4 of 2015 from B's model and compare them with those from A's model as calculated in part (e).

(2 Marks)