Problem 1 (60 marks)

The London Commissioning Support Unit (LCSU) has collected data concerning continuing care (CC) services across 26 of the 31 Clinical Commissioning Groups (CCGs) in London. CC is long-term care (whole or in part NHS funded) provided either at home or in residential/nursing homes (placement) for those patients that require any of the following categories of care (care groups):

Functional mental health (FMH) Learning disability (LD)  
Organic mental health (OMH)

Palliative care (PC)  
Physically disabled adult (PDA) Physically frail (PF)

Aware of your expertise in data analysis, LCSU has contacted you to undertake the analysis of the data [file LCSU\_CC\_Data.csv]. In particular they would like you to answer the questions below. For each question, indicate what assumptions you have made. For questions 2-10, justify why you choose a particular approach (graph, hypothesis test, both, etc.), and summarise your result in plain English (no more than 2-3 lines

**a)**Describe how you prepared and cleaned the data. What assumptions have you made to do this? In particular, you will notice that there are quite a few missing values. Describe and suggest ways to deal with the problem (with references), even if you end up deleting records.

**b)**Produce suitable charts to compare:

* 1. Split between provision types (home vs Placement) across CCGs
  2. Variations in care groups across CCGs

**c)**Derive age at admission and calculate mean, min, max and median age at admission overall, overall per care group, overall per provision type. Also calculate this for each CCG.

**d)**Produce box plots to compare age at admission across CCGs for each care group

**e)**Produce box plots to compare weekly rate across CCGs for each care group

**f)**Plot histograms of age at admission and weekly rate and comment on the distributions of these variables

**g)**Determine whether in each care group, the age at admission is higher in CCG W than in CCG X.

**h)**Determine whether the weekly rate of LD is higher in CCG W than in CCG X.

**i)**Are females more (or less) likely than males to be cared for at home?

**j)**Is age at admission correlated with weekly rate?

Problem 2 (35 marks)

page3image37974016

Please answer the following questions after carefully reading the paper:  
P. Falcoz, C. Binquet, F. Clement, D. Kaili, C. Quantin, S. Chocron and J. Etievent (2003) “Management of the second episode of spontaneous Pneumothorax: a decision analysis,” *Annals of Thoracic Surgery*, Vol. 76, pp. 1843–8

**a)**What kind of model was developed? What was the aim(s) of the model?

**b)**Describe briefly the possible decision alternatives and outcomes.

**c)**Using a suitable package such as Precision Tree, draw, for this problem, a decision tree such as the one in Figure 1, including all parameters (baseline values).

**d)**Determine the expected cost and effectiveness of each decision alternative. Use the tree drawn in c) and show all working. Compare your results with those in Table 2.

**e)**Study:

* + The effect of mean length for Conservative Management (CM) on expected cost
  + The two-way effects of mean length of stay for CM and VATS on expected cost
  + The two-way effects of mean length of stay for CM and VATS on ICER (Incremental Cost Effectiveness Ratio)

Comment on your results, in particular on:

-how they compare with those in the paper  
-the usefulness of sensitivity analysis in decision analysis.

**Problem 3**

The London Metropolitan (Met) Police has been informed of your expertise in Analytics and has commissioned you to undertake an analysis of reported offences between April 2010 and March 2019 (nine financial year), with view to develop forecasts for the following year (Apr 2019 – Mar 2020). For that purpose, they have recorded in a separate worksheet offence data for that year. Given the limited time available they would like you to focus on four crime categories: Burglary, Sexual Offences, Theft, and Violence Against the Person (11 subcategories – it is recommended that you search for more information about the meaning of these categories). You will find the data in the Excel file “Met Police Offence Data 2010-2020 CW.xlsx”. Make sure to always justify your answers.

**a)**Plot the data (Apr 2010 – Mar 2019) as time series (11 of them). Use the graphs, summarise which series exhibit trends (linear / non-linear) and/or seasonality.

**b)**Analyse for seasonality the data for “Burglary - Business and Community”, “Burglary – Residential”, “Rape”, “Bicycle Theft”, “Violence with Injury”, “Violence without Injury”. Include the seasonal indices for each month, and comment on the high seasonal and low seasonal sales months. Do the seasonal indices make intuitive sense? Discuss.

**c)** For each of the three offence types exhibiting highest seasonality in b), use two appropriate forecasting techniques to forecast offences for Apr 2019 to Mar 2020.  
For each of the three time-series, compare the results obtained with the two methods using the mean error (bias), the mean absolute deviation (MAD), the mean square error (MSE), and the mean absolute percentage error (MAPE) of the forecasts on the data between Apr 2010 and Mar 2019.

**d)**Check and discuss autocorrelation and error randomness for each series.

**e)**Produce forecasts for the three offences in c) for Apr 2019 to Mar 2020 using your two chosen

methods and compare your results with actual values for Apr 2019 – Mar 2020 using again the mean absolute deviation (MAD), the mean square error (MSE), and the mean absolute percentage error (MAPE) for that period. Are the best performing forecasting methods still the same as in c)?

**f)**Discuss correlations between the 11 offence subcategories, potential issues with calculating these, and how to alleviate them.

**Problem 4**

FarmProd produces two types of cattle feed, both consisting totally of wheat and barley. Feed 1 must contain at least 80% wheat, and feed 2 must contain at least 60% barley. Feed 1 sells for £1.5 per pound, and feed 2 sells for £1.30 per pound. FarmProd can purchase up to 1,000 pounds of wheat at £0.50 per pound and up to 800 pounds of barley at £0.40 per pound. demand for each type of feed is unlimited.

**a)** Formulate algebraically a linear programme to maximise FarmProd’s profit.

To ensure the loyalty of its main customer, FarmProd has decided to give him a quantity discount. If the customer purchases more than 300 pounds of feed 1, each pound over the first 300 pounds will sell for only £1.25 per pound. similarly, if the customer purchases more than 300 pounds of feed 2, each pound over the first 300 pounds will sell for £1.00 per pound.

**b)**Re-formulate algebraically the LP from the previous question to account for the presence of quantity discount.

**c)**Develop a spreadsheet model for the problem and solve. Describe your solution.

**d)**Discuss how sensitive the optimal solution is to key inputs.