Module Code: TRAN 5032M

Module Title:
Transport Data Collection & Analysis

Institute for Transport Studies

January 2023

Calculator instructions:

- You are allowed to use a (computer) calculator in this assessment.

Dictionary instructions:

- You are allowed to use your own dictionary in this exam and the spell checker facility on your computer

Exam information:

- There are 4 pages to this exam.
- You will have 5 hours, plus your additional time allowance, if applicable, to complete this exam.
- You are recommended to take a maximum of 2 hours within the time available to complete the assessment.
- This assessment is worth 50% of the overall module mark.
- Answer **ALL** the questions.
- This paper has two sections A and B worth equal marks.
- Each section has two questions. In Section A, Question 1 is worth 35 marks and Question 2 15 marks. In Section B, the two questions are worth 25 marks each.
- The deadline for submission of your assessment is 2pm (UK Time) on Monday 16th January 2023.
- Please submit your assessment to the 'Submit Your Work' area on the module's Minerva page.
- Please include your Student Identification Number (SID) in the title of your submission.
- Please include an OTLA Cover Sheet with your submission.
- If there is anything that needs clarification or you have any problems, please email c.calastri@leeds.ac.uk, a.m.salasjones@leeds.ac.uk, d.p.watling@its.leeds.ac.uk and ITS-Support@leeds.ac.uk and we will respond to you as quickly as possible within normal working hours UK time (9:00-17:00 hours, Monday-Friday).
- You must not discuss or share the content of or answers to this assessment, with any fellow students, any staff or other contacts outside the school or the University's professional services. School contacts available to you will be detailed in the bullet point above.
- Your answers must be typed. Hand-written equations and hand-drawn sketches can be pasted as images into the documents.
- Sketches should be clearly labelled.
- You are required to submit one file, including the calculations, graphs, interpretations, and discussion. Marks will be awarded for the correctness of your answer, the relevance of the answer, the quality of discussion and critique, your synthesis, and (where appropriate) presentation.
- The level of referencing is the same as what is expected in a normal exam. You are expected to mention key authors/papers in the text, but you do not need to include a reference list.

- Please be reminded that you are expected to maintain the normal high standards of academic integrity that apply to all your assessments.

Aim:
The learning outcomes of this Online Time Limited Assessment are to assess your ability to:

- Evaluate the link between research needs and data requirements and collection
- Discuss the principles underlying statistical analysis and apply these on data
- Discuss and recommend data collection techniques relevant to transport issues
- Apply data acquisition skills
- Apply data handling, statistical and analytical skills

SECTION A

This section of the exam contains two questions.

You are provided with two data sets, one for each question in Section A. You are asked to conduct statistical analyses on these data sets, and to discuss the results. You can either conduct the analyses by hand or by using software. Whichever method you use you should present the main steps in your analysis. For example, do not just provide numerical estimates derived from software, without explaining the basis for how they were calculated (e.g. if using Excel, make it clear which Excel functions were used, how, and what those functions do).

Question 1:

The length of commuting trips for a sample of staff travelling to work at the Hospital has been computed from GPS data. The trip distance data are split into those commuting by car and those using other (non-car) modes. The data are provided in the "TDCA_Q1_Data_.xls" Excel file, along with histograms for both modes.

a) For both CAR and NOCAR data, compute the sample mean, median, standard deviation and skewness (stating which skewness definition or formula you used). Comment on the distribution of the data sets, referring both to the histogram and summary statistics. — 7 marks

b) Do there appear to be any outliers or erroneous data points? If yes, explain how you identified them and why such observations might have been recorded. Remove them from the data and re-compute the summary statistics before continuing to the remainder of the question. — 5 marks

c) Estimate a 95% confidence interval for the true mean commuting distance by CAR (you do <u>not</u> need to estimate for NOCAR). State any assumptions you make and explain the relevance of the central limit theorem. — 8 marks

d) Use a statistical test to decide whether these data give evidence that the mean commuting distance by car is different from the mean commuting distance by other modes. State any assumptions you — 15 marks

make. Interpret your findings to be understandable by an intelligent non-expert.

Question 2:

An attitudinal survey was conducted among peak-time rail commuters arriving in Leeds and in Manchester, to understand their views on levels of crowding on services following the lifting of Covid-19 restrictions. The resulting data are provided in the "TDCA OTLA Jan 2023 Attitudinal survey.xlsx" file.

a) Use a suitable test to determine whether the attitudes appear to be independent of the location (i.e. whether arriving in Leeds or Manchester). — 10 marks

b) Identify any potential problems with applying the test in 2(a) and suggest a resolution (you do not need to repeat the test). — 5 marks

SECTION B

Question 3

Results from the 2021 Leeds travel to work survey showed that over 35% of the respondents travel to work by car with a single occupant, while less than 10% use buses as their main transport mode. Leeds City Council (LCC) is considering investing to improve the performance of the bus network in the city and wants to know which improvements (e.g. in reliability, headway and/or network coverage) are more urgently needed. To fulfil this task, you are asked to:

a) Devise a data collection plan that details the data to be collected, and what would be the most appropriate method(s) to collect it. — 10 marks

b) How would you analyse the data, and how can the results from your analysis be used by LCC? Specify statistical techniques relevant to your proposed analysis. — 10 marks

c) Discuss how the limitations of the data collection techniques in your plan could affect the data quality and your analysis. — 5 marks

Question 4:

A large UK city is facing serious congestion and air quality issues and the council decides to run a data collection to better understand how people commute, with the aim of designing informed policies to encourage more sustainable mobility in the area. They have decided to run a survey and have formulated the following survey question:

What transport modes do you use for commuting?
- Car/Van
- Bike
- Bus
- Scooter
- Carpool

a) Is this a good survey question? If not, why? How would you improve it? Explain your answer. — 10 marks

b) What other information should the council collect as part of this survey? Explain your answer. — 5 marks

c) Propose a data collection protocol for this survey, including a target population, recruitment strategy, modality of data collection, bias limitation strategy, incentives, and timeline. — 10 marks