

Individual Assignment Question and Guidance

The Structure

The assignment should be completed in RMarkdown and knitted into a html file. Your submission should be a zip file containing the RMarkdown file and the html output it creates. When the RMarkdown file is re-run, it must produce the same html output that you submit.

This should have two parts. The first part should include the code that you produce to perform all stages of the data analysis. It should read in the datafile provided, check the integrity of data, calculate all of the statistics requested, and produce all of the figures/visualisations requested.

The purpose of this first part is that it could be shared with someone else who is not familiar with the data or project, but is expert in R, tidyverse, and statistics. They should be able to see how you achieved your results, and if necessary, easily modify the code to work with new or updated data. Therefore, the appropriateness of your comments and data documentation will be assessed, as well as clarity and consistency of approach such as consistent use of compatible packages.

The second part should be a short, polished and professional report presenting and interpreting the findings for an audience of managers or policy makers. You should assume that the managers/policy makers have a basic knowledge of the data and where it comes from, but are not experts in statistics or R. Therefore, in addition to reporting the results of statistical tests using standard reporting conventions, you should also state the results in plain English and say what they mean or what they tell us about the real scenario – in this case what they tell us about the number of pedestrians in different locations in York. The visualisations and figures in this section should be as clear as possible and of publication standard.

The official word count for this assignment is 1500 words, but **the written elements of the report should be much shorter than this**. Since this is a code and analysis based assignment, the word count is not directly applicable in the way that it is for an essay. The word count is provided as a guide: that this assignment is assumed to be approximately the same amount of work as a 1500 word essay.

The comments (in code and between code blocks) in the first part should be short, brief, and sufficient for another data analyst to quickly determine what you are doing and why. The conclusions and writing in your second part should be no more than a few hundred words. If you include more than this then you are likely including irrelevant information. This could have a negative impact upon your grade.

The Request

York Footfall Data

We are taking the role of a promoter in York who is deciding what is the best location to place their stall so that the most people walk by it. To do this we are going to use daily footfall data. This data comes

from cameras that automatically measure the number of people who walk past a particular location: the “footfall”.

Load in the data and check it for integrity/quality.

Provide a summary table that shows the following for each location where footfall was measured:

- The date of the first and last day when footfall was measured at this location
- The mean daily footfall
- The standard deviation of the daily footfall
- The highest daily footfall
- The lowest daily footfall

For all tasks from this point onwards, you should only use data from the year 2019 (the last full year before the COVID pandemic).

Plot the distribution of footfall across all days in the dataset. This should show the distributions separated by the locations where footfall is measured (i.e. for each LocationName).

Perform a t-test to compare the footfall on Coney Street to Stonegate.

Perform an additional t-test to compare the footfall on Coney Street to Stonegate only on weekends (i.e. Saturdays/Sundays).

What would you advise the promoter about whether they should place their stall on Coney Street or Stonegate to maximise the number of passers by:

- if they were doing so every day of the week
- if they were doing so only at weekends

Maximum Length: 1500 words (*this is a strict limit, not a guideline*)

Assessment Weighting: 20%

Submission Deadline: Thursday 31st October 2024 before 12:00:00 (Midday, UK time)

Marks Released by: Thursday 28 September 2024
(*we will aim to release marks by this date, but in the event of an unavoidable delay a message will be posted on the module page*)

Artificial Intelligence: Permitted

Use of Artificial Intelligence

The University recognises an increasing number of technologies such as Artificial Intelligence and that they may be applicable in your completing this assessment. The assessment brief sets out specific requirements or restrictions, and your student handbook has further guidance and advice.

You are reminded that the inappropriate use of such a technology may constitute a breach of University policy, such as the Proofreading Policy or [Regulation 11 \(Academic Integrity\)](#). If you breach these policies, it may have significant consequences for your studies. Please make sure you read and understand the assessment brief and how AI may or may not be used.

When you submit the assessment, you will be required to explain the use of any AI. Failure to disclose at the point of submission may be prejudicial in any later investigations should they arise.

For this assessment, AI is:

PERMITTED (neither prohibited or required)

If you use a generative Artificial Intelligence (AI) in the process of completing this assessment you **MUST** set out clearly the following:

- WHY you used a generative AI
- WHAT it was used for
- WHICH AI was used; and
- If any generated content has been used directly in this submission, if so where.

Note that this declaration does NOT contribute towards the word count for the assessment.

You will also have to confirm in your declaration that the work remains yours and you have intellectual ownership of it. You may be called for viva or other interview to demonstrate such intellectual ownership. A failure to disclose the use of AI, or the use of a misleading description of its use may have significant consequences for your studies. As a result, keeping good records of your interactions is strongly advised.

Submitting your work

Before you submit your assessment, you should ensure you are familiar with the guidance and rules in the “Your Assessments” section of your Student Handbook, paying particular attention to:

- Academic Integrity (including Plagiarism)
- Referencing
- Word Count Policy and Formatting
- Confidentiality of your work
- Extensions and late submission of work
- Guidelines for Online Submission

Mitigating Circumstances

Mitigating circumstances MUST be submitted **within 20 working days following the submission deadline**. Mitigating circumstances not submitted by the relevant deadline are not required to be considered by the School/Department and may have to be considered by an Academic Appeals Committee as part of an academic appeal – for further information, please see:
<https://warwick.ac.uk/services/gov/calendar/section2/regulations/reg42academicappeals>

Please see your Student Handbook for more guidance on mitigating circumstances
