

Coursework 2: Group analysis case study (worth 40% of your final module grade)

The objective is to assess your ability to analyse data and communicate your findings. You need to do this coursework in the groups that are defined on Minerva. A group comprises a maximum of four people.

Your task is to analyse some road safety data (called STATS19) for England for the 12 months to 31/12/2019, to provide answers these questions:

- a) What patterns are there in the demographics of casualties?
- b) Looking at accidents which included a killed or seriously injured (KSI) casualty, what patterns are there between the local authorities?
- c) What patterns are there in pedestrians who were KSI casualties?

The data you need is available from the UK Government at

<https://www.data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data>

The accident and casualty data is in two separate files, and the metadata is provided by the Road Safety Open Dataset Data Guide.

Mark breakdown

The report has a limit of eight pages and should be written single-spaced in an 11 pt font (or larger). The report should have the following section headings and content

- 1. Introduction [4 marks]:** State the coursework aim, the data files that were used and provide links to them.
- 2. Data quality [20 marks]:** Describe the data quality issues that are present in the data.
- 3. Data characterisation [20 marks]:** After you have cleaned the data, describe and characterise every variable.
- 4. Detailed analysis [30 marks]:** Provide answers to questions (a), (b) and (c), with appropriate qualitative and quantitative evidence.
- 5. Conclusions [6 marks]:** Restate your main findings.

To gain full marks your report needs to adhere to the specified structure, be well presented and easy to understand.

Do not include references in your report.

Deadline

One person should make the submission on behalf of the group. The submission deadline is 10am on Wed 7th Dec 2022. Late submissions will be penalised in accordance with departmental guidelines.

Your submission must be one PDF file that contains: (a) a coursework header sheet that states the names and student IDs of all the group members, and (b) your report.

Optionally: Before the header sheet, insert a signed copy of the COMP5122M Data Science Group Work Form for your group. You only need to do that if the group members did not contribute equally to the coursework.

Roy Ruddle and Layik Hama, November 2022