

Data Mining & Visualisation – Dataset Presentation			
Student	Shahriar Khan		
Dataset	Road Traffic Dataset		
Marking			
Presentation/Slides (Clarity, readability, flow, accuracy, completeness, performance) Good relaxed presentation style. Be careful with your use of colour – pale cream can be difficult to read against mid grey.			(10 marks) 6
Dataset Description (Where did the data come from? What does the data describe? Is the sample large enough? What data formats and types are involved?) More detail could be provided about the attributes, for instance by providing example records rather than simply a list of the attributes and their meanings. Three columns is a subset of the data structure rather than the whole data structure.			(50 marks) 28
Dataset Suitability (What features of the dataset lend themselves to data mining? Has it been mined previously? Why does the student feel that they have the contextual background to work with this data?) Some good points made about the scale of the dataset and its capacity for supporting time series analysis. There are, however, a lot of external influences on this that you don't mention or have data for – weather being the most obvious. How will that impact the patterns found in the data? How do you know where the roadworks are, or what other external events may impact the traffic?			(20 marks) 11
Objectives (What does the student plan to investigate/expect to find? Are these SMART?) The research questions are quite simplistic, possibly because you don't have the full context for this dataset. What kind of patterns are you expecting to find beyond the obvious ones? You have a lot of information about different types of vehicle – do you intend to use that and, if so, how?			(20 marks) 11
General comments Understanding techniques is not the same as understanding the context of the data. You should be able to discover some patterns in the dataset but it will be harder for you to interpret them well without the necessary context.			
Date		Final Grade (%)	56