

Data Mining – Dataset Presentation			
Student	Paige Harris		
Dataset	Online Retail Set		
Marking			
Presentation/Slides (Clarity, readability, flow, accuracy, completeness, performance) Good slide design. Well presented. Clear summary. Smooth transition between topics and a nice pace to the presentation.			(10 marks) 8
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?) Interesting data set – 1 year. Good statistical description – gives a really good feel for the content of the dataset. Nice idea to open up the UCI site to show the origins of the data. Good understanding of the limitations of the dataset, and that some cleaning may be required.			(50 marks) 42
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?) Plenty of records, although a bit low on the number of attributes. Good consideration and amelioration of limitations – e.g. ‘unspecified’ country values. Good background for context.			(20 marks) 16
Objectives (What does the student plan to investigate/expect to find? Are these SMART?) Questions lead on from each other – logical progression. Lots of potential. Looking at seasonal trends is interesting. Looking at low stock, how does that affect the business? Consider using classification and clustering. Only having one year of data may limit the data mining options a little.			(20 marks) 16
General comments Try not to read from the slides. Nice idea to use a UCI dataset – this means you can use the referred papers for hints. Look at trends like days of the week vs. weekend buying, end of the month payday increase in sales, etc. Good audience interaction, confident attitude.			
Date	21/1/2019	Agreed Final Grade (%)	82
Tutor Signature			
Second Marker Signature			