

Data Mining – Dataset Presentation			
Student	Yan Xue 9:57		
Dataset	PM2.5 Data in 5 Chinese Cities		
Marking			
Presentation/Slides (Clarity, readability, flow, accuracy, completeness, performance) Excellent use of initial slides to set the scene for the project. Clear slides.			(10 marks) 7
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?) Good description of the background to the problem and the source of the data. Clear explanation of how the data has been combined into one table as initial data cleaning.			(50 marks) 36
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?) Good justification for the dataset’s use, including sustainability issues.			(20 marks) 14
Objectives (What does the student plan to investigate/expect to find? Are these SMART?) Clear objectives, and an understanding of the necessary contextual knowledge required.			(20 marks) 14
General comments Looks like a really interesting dataset.			
Date	11/3/21	Final Grade (%)	71
Tutor Signature			