

Student: HAN LUPercentage: 70%

QUESTION	EXCEPTIONAL 5	MERITORIOUS 4	COMPETENT 3	PASS 2	FAIL 1
Dataset description (15%)	Infers multiple pre-processing AND analysis techniques from descriptive statistics	Infers pre-processing OR analysis techniques from descriptive statistics	Provides descriptive statistics for the dataset	Correctly defines the data types; describes the original data source	Misinterprets data types and meanings
Data pre-processing (10%+10%)	Performs multiple complex primary pre-processing operations	Codes data sensibly and discusses any data removal or manipulation	Performs obvious coding (Yes/No) OR justifies data manipulation	Describes the data used but does not perform any significant pre-processing	Performs no primary pre-processing
	Performs multiple secondary pre-processing operations	Recodes data to allow further analyses AND adds new calculated fields	Recodes data to allow further analyses OR adds new calculated fields	Performs appropriate secondary pre-processing	Performs no secondary pre-processing
Data analysis - data exploration (15%+10%)	Uses multiple methods other than decision trees Full and detailed discussion of the selected model	Uses multiple types of algorithm AND multiple options with each of these.	Uses multiple types of algorithm appropriately and shows a logical progression within an analysis	Uses the same type of algorithm in each analysis	Incorrect, misleading or missing
	Fully tests each model in each analysis	Uses MVT methodology correctly	Includes prediction and comparison of algorithms	Only creates models on the total dataset	Incorrect, misleading or missing
Reporting of results (20%)	Four complete analyses with discussion of the implications of the results	Three complete analyses with discussion of the implications of the results	Two complete analyses with discussion of the implications of the results	One complete analysis with discussion of the implications of the results	Incorrect, misleading or missing
Reflection (10%)	Full and detailed reflection with discussion of how the analysis might be repeated and extended	Discussion of how the research might be repeated OR extended.	Discussion of satisfying analyses and problems encountered.	List of skills learned.	Incorrect, misleading or missing
Report writing (10%)	Reads like a technical review article	Good use of signposting	Clear sectioning of content	Basic use of introduction, body and conclusion; correctly referenced	No clear structure; incorrectly referenced

Good introduction and clear data description

I disagree that there was nothing interesting in the initial data visualisation. The shapes for casual, registered, and total are clearly exponential - why is this?

Pulling in two datasets from different locations is a great idea - how did you join the 2 datasets? This is also part of the pre-processing and should be documented.

Excellent Tableau analysis, with strong, clearly-stated conclusions.

Very good and thoughtful analysis in Weka. Using the Tableau results to guide your initial experiments with Weka is an excellent idea.

Excellent secondary pre-processing, particularly in dealing with the time series analysis.

Thank you for your honesty about how the work sometimes made ~~you~~ you feel depressed. Data mining is often like that, when you have what seems a brilliant idea that ~~it~~ just doesn't work. But when the ideas do work, you feel great!

Overall a very good piece of work. Well done!

Data Mining – Final Poster			
Student	LO HAN		
Dataset	Shared Bike Dataset		
Marking			
Abstract (Clarity, readability, flow, accuracy, completeness) Good including the questions in the abstract.			(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? What were the objectives of the project?) Good description of the dataset and its preprocessing.			(10 marks) 8
Data Processing (What primary and secondary preprocessing was needed? How was data integrity maintained?) Good coverage of both primary and secondary preprocessing.			(20 marks) 15
Analysis (What analyses were carried out? Are they appropriate?) Each question clearly answered, with some technical details given.			(20 marks) 15
Results (Are the results clearly stated? Do they meet the objectives?) Results are clear, and are linked to each question.			(20 marks) 15
Graphic Design (Is it balanced and readable? Do the colours work? Are the graphics appropriate?) Layout and colour are good, but the text is a bit too small. A graphic would also help. You could put the question being addressed in each data box.			(20 marks) 13
General comments Well-structured.			
Date	20/5/19	Agreed Final Grade (%)	73
Tutor Signature	Nik Whithead		
Second Marker Signature	 		

$$15^{07} - 15^{25} + 2$$

Data Mining – Final Presentation			
Student	LU HAN		
Dataset	SHARED BIKE DATASET.		
Marking			
Presentation/Slides (Clarity, readability, flow, accuracy, completeness, performance) Good thematic slide design. Text is a little small at times			(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? What were the objectives of the project?) Good description. Sensible idea to split into 4 general types.			(10 marks) 7
Preprocessing (What primary and secondary preprocessing was needed? How was data integrity maintained?) Sensible preprocessing steps.			(25 marks) 19
Analysis (What analyses were carried out? Are they appropriate?) Good descriptions of highly relevant analyses. Excellent explanation of multiple possibilities for each question			(20 marks) 18
Results (Are the results clearly stated? Do they meet the objectives?) Good display of similarities and differences in use numbers over time. Good strong cluster analysis. Good identification of periodicity. Excellent summary slide.			(25 marks) 22
Results (What further analyses could be carried out if you had more time?) Good ideas for further analysis. Identifies the known is			(20 marks) 14
General comments Try to look at the audience occasionally. Very good description of not only what you did but also why you did things.			
Date	24/5/19	Agreed Final Grade (%)	85
Tutor Signature	Nik Whitehead		
Second Marker Signature	 		

86.