FIT5149 Assignment 1 - Marking Rubric

Marking Rubric - Assignment 1 contributes 25% to your final unit mark

Indicating the level the student is work at*:	**(N) Fail, (P) Pass, (C) Credit, (D) Distinction, (HD) High Distinction**

This assessment meets **Unit Learning Outcomes** 1,2,4,5

1. Analyse data sets with a range of statistical, graphical and machine-learning tools;

2. Evaluate the limitations, appropriateness and benefits of data analytics methods for given tasks;

3. Design solutions to real-world problems with data analysis techniques;

4. Assess the results of analysis;

5. Communicate the results of an analysis for both specific and broad audiences.

Total: 25 marks	Did not attempt	Poor	Limited	Some	Most	All

Data exploration	No or incomplete data exploration potentially without supporting plots or statistical methods	Incomplete descriptive exploration of variables with inappropriate plots and statistical methods. No justification for the association or interaction between variables.	Limited exploration of variables or analysis of findings with suitable plots and statistical methods. The use of plots and statistical methods is poorly justified (if at all).	Some exploration of variables and analysis of findings with suitable plots and statistical methods, however the use of these plots and methods is not well justified.	Clear and succinctly explores the variables (and analyses findings) with suitable plots and statistical methods. The use of the plots/statistical methods are well justified however the exploration and analysis is not comprehensive.	A comprehensive and descriptive exploration of variables and analysis with suitable plots and statistical methods. The use of the plots/statistical methods is well justified. To be comprehensive the exploration and analysis should: 1. Justify the characteristics of variables, potential association or interaction between variables, and

						variable transformation using proper plots and diagnostics. 2. Provide a comprehensive analysis of important findings, which demonstrates excellent data insight and high level of understanding of exploratory data analysis.
3 marks	0	0.5	1.5	2	2.5	3

Model description and justification	No attempt	No, or limited explanation of how the models were developed with no evidence and proper justification.	Limited explanation of how the models were developed with insufficient evidence and proper justification. The exploratory analysis does not logically lead to the development of the modes	A fair explanation of how the models were developed with some relevant supporting information and statistical diagnostics. The exploratory analysis shows some connection to the development of the models	A logical explanation of the model development process with supporting information derived from exploratory data analysis. The exploratory analysis and diagnostics logically lead to the development of the models however some points are not fully justified	Clearly explained how the models were developed with sufficient supporting information derived from exploratory data analysis and proper statistical diagnostics. The exploratory analysis and diagnostics logically lead to the development of the models (including attribute or feature selection).
3 marks	0	0.5	1.5	2	2.5	3

The quality of model comparison	No comparison of the models	A limited comparison of the models with poor analysis and lack of details.	Comparison of the models is limited. The description is without sufficient analytics and details. Demonstrates a limited understanding of the difference between the compared models	Some logical comparison of the models with sufficient statistical analysis and an appropriate level of details. Demonstrates some understanding of the difference between the compared models.	Comparison of the models with solid analysis and appropriate details to support the argument. Demonstrates a solid understanding of the difference between the compared models.	Critically assess the accuracy of the models with indepth statistical analysis. The comparison is logical and solid with an appropriate level of details. Demonstrates a high-level of understanding of the differences (i.e., advantages and disadvantages) between the models in regard to the task.
3 marks	0	0.5	1.5	2	2.5	3

Explanation of why the chosen subset of attributes have a significant impact on the prediction of the response variable	No explanation	Limited justification of why a subset of attributes have a significant impact on the prediction and there is no statistical analysis to support their argument	Limited justification of why a subset of attributes have a significant impact on the prediction with insufficient statistical analysis and accompanying interpretation	Some Explanation of why a subset of attributes have a significant impact on the prediction with proper statistical analysis and accompanying interpretation, in line with the findings from the model development.	Good justification of why a subset of attributes have a significant impact on the prediction with sufficient and proper statistical analysis and interpretation, in line with the findings from the model development.	The subset of attributes leading to the prediction are identified and clearly justified with in-depth statistical analysis. The description and accompanying interpretation are comprehensible and convincing, in line with the findings from the model development.
3 marks	0	0.5	1.5	2	2.5	3

The quality of the Jupiter notebook/ the R Markdown file, quality and readability of the R implementation	The Jupyter Notebook is messy and incomplete. No commenting of code	The Jupyter Notebook is unorganised with no evidence to support the development of the answers The R code is incomplete and the code readability is poor.	The Jupyter Notebook is poorly organised with no or insufficient evidence to support the development of the answers The R code is complete but poorly structured and the code readability is poor with no or insufficient comments to explain how the code is intended to work.	The Jupyter Notebook is acceptably structured with some evidence to support the development of the answers The R code is structured but the logic is not clear. Code readability is good with adequate comments to explain how the code is intended to work	The Jupyter Notebook is well structured with sufficient evidence to support the development of the answers. The R code is logically structured and easy to read. Comments clearly explain how the code is intended to work. However, there are some redundant or unnecessary code and/or comments.	The Jupyter Notebook is well structured in a logical order that clearly supports the development of the answers. The R code is logically structured and easy to read. Concise but precise code comments clearly explain how the code is intended to work.
3 marks	0	0.5	1.5	2	2.5	3

Leadership, teamwork, and project management	No evidence of leadership or contribution. Student was a free rider. No documentation of tasks. Consistently failed to meet deadlines. Had a negative impact on the project.	Minimal leadership and contribution. Rarely participated or communicated. Poorly documented team contributions. Often missed deadlines and was unreliable. Had little to no positive impact.	Limited leadership and inconsistent contribution. Occasionally participated but often relied on others. Limited task documentation. Sometimes missed deadlines. Had a limited positive impact.	Some leadership shown and regular contributions made. Adequate task documentation. Generally met deadlines and fulfilled most responsibilities. Had a moderate positive impact.	Good leadership skills and consistent valuable contributions. Tasks were clearly assigned and managed. Consistently met deadlines and fulfilled responsibilities. Had a significant positive impact.	Excellent leadership and consistently substantial contributions. Comprehensive documentation of all tasks. Always met or exceeded deadlines and expectations. Had an exceptional impact, driving the project's progress and outcomes.
5 marks	0	1	2	3	4	5
Model Performance	No valid predictions made	RMSE > 2.5 or extreme overfitting/underfitting evident	$2.0 \leq$ RMSE < 2.5 with some evidence of proper model fitting	$1.7 \leq$ RMSE < 2.0 with good evidence of proper model fitting	$1.3 \leq$ RMSE < 1.7 with strong evidence of proper model fitting	RMSE < 1.3 with excellent model fitting and robustness across cross-validation
5 marks	0	1	2	3	4	5