

	PROJECT 2	Qingyi Chen
1	<b>Problem Statement</b>	
	Is it clear what the student plans to do?	3.0
	What type of model will be developed?	3.0
	How will success be evaluated?	3.0
	Is the scope of the project appropriate?	3.0
	Is it clear who cares about this or why this is important to investigate?	3.0
	Does the student consider the audience and the primary and secondary stakeholders?	3.0
	<b>Problem Statement - Averaged Marks (out of 3)</b>	3.0
2	<b>Data Cleaning and EDA</b>	
	Are missing values imputed appropriately?	2.5
	Are distributions examined and described?	2.0 [1]
	Are outliers identified and addressed?	3.0
	Are appropriate summary statistics provided?	3.0
	Are steps taken during data cleaning and EDA framed appropriately?	3.0
	Does the student address whether or not they are likely to be able to answer their problem statement with the provided data given what they've discovered during EDA?	3.0
	<b>Data Cleaning and EDA - Averaged Marks (out of 3)</b>	2.8
3	<b>Preprocessing and Modeling</b>	
	Are categorical variables one-hot encoded?	3.0
	Does the student investigate or manufacture features with linear relationships to the target?	3.0
	Have the data been scaled appropriately?	2.5
	Does the student properly split and/or sample the data for validation/training purposes?	3.0
	Does the student utilize feature selection to remove noisy or multi-collinear features?	3.0
	Does the student test and evaluate a variety of models to identify a production algorithm (AT MINIMUM: linear regression, lasso, and ridge)?	3.0
	Does the student defend their choice of production model relevant to the data at hand and the problem?	2.0 [2]
	Does the student explain how the model works and evaluate its performance successes/downfalls?	2.0 [3]
	<b>Preprocessing and Modeling - Averaged Marks (out of 3)</b>	2.7
4	<b>Evaluation and Conceptual Understanding</b>	
	Does the student accurately identify and explain the baseline score?	1.5 [4]
	Does the student select and use metrics relevant to the problem objective?	3
	Is more than one metric utilized in order to better assess performance?	3
	Does the student interpret the results of their model for purposes of inference?	3
	Is domain knowledge demonstrated when interpreting results?	1.5 [5]
	Does the student provide appropriate interpretation with regards to descriptive and inferential statistics?	3
	<b>Evaluation and Conceptual Understanding - Averaged Marks (out of 3)</b>	2.5
5	<b>Conclusion and Recommendations</b>	
	Does the student provide appropriate context to connect individual steps back to the overall project?	3.0
	Is it clear how the final recommendations were reached?	3.0
	Are the conclusions/recommendations clearly stated?	3.0
	Does the conclusion answer the original problem statement?	3.0
	Does the student address how findings of this research can be applied for the benefit of stakeholders?	3.0
	Are future steps to move the project forward identified?	2.0
	<b>Conclusion and Recommendations - Averaged Marks (out of 3)</b>	2.8

<b>6</b>	<b>Project Organization</b>	
	Are modules imported correctly (using appropriate aliases)?	3
	Are data imported/saved using relative paths?	3
	Does the README provide a good executive summary of the project?	2.0 [6]
	Is markdown formatting used appropriately to structure notebooks?	3
	Are there an appropriate amount of comments to support the code?	3
	Are files & directories organized correctly?	3
	Are there unnecessary files included?	3
	Do files and directories have well-structured, appropriate, consistent names?	3
	<b>Project Organization - Averaged Marks (out of 3)</b>	<b>2.9</b>
<b>7</b>	<b>Visualizations</b>	
	Are sufficient visualizations provided?	3.0
	Do plots accurately demonstrate valid relationships?	3.0
	Are plots labeled properly?	3.0
	Are plots interpreted appropriately?	3.0
	Are plots formatted and scaled appropriately for inclusion in a notebook-based technical report?	3.0
	<b>Visualizations - Averaged Marks (out of 3)</b>	<b>3.0</b>
<b>8</b>	<b>Python Syntax and Control Flow</b>	
	Is care taken to write human readable code?	3.0
	Is the code syntactically correct (no runtime errors)?	3.0
	Does the code generate desired results (logically correct)?	3.0
	Does the code follows general best practices and style guidelines?	2.5 [7]
	Are Pandas functions used appropriately?	3.0
	Are sklearn methods used appropriately?	3.0
	<b>Python Syntax and Control Flow - Averaged Marks (out of 3)</b>	<b>2.9</b>
<b>9</b>	<b>Presentation</b>	
	Is the problem statement clearly presented?	2.5
	Does a strong narrative run through the presentation building toward a final conclusion?	2.5
	Are the conclusions/recommendations clearly stated?	2.5
	Is the level of technicality appropriate for the intended audience?	2.5
	Is the student substantially over or under time?	2.5
	Does the student appropriately pace their presentation?	2.5
	Does the student deliver their message with clarity and volume?	2.5
	Are appropriate visualizations generated for the intended audience?	2.5
	Are visualizations necessary and useful for supporting conclusions/explaining findings?	2.5
	<b>Presentation - Averaged Marks (out of 3)</b>	<b>2.5</b>
<b>Total:</b>	<b>Overall Marks (out of 27)</b>	<b>25.1</b>
	<b>Average Marks (out of 3)</b>	<b>2.79</b>
<b>27</b>	<b>Percentage</b>	<b>93.0%</b>

Comments	<p>[1] Well done in this section, but missing target variable distribution.</p> <p>[2] Stated Lasso can zero-out insignificant features, but could improving reasoning with model metric evaluation</p> <p>[3] Could improve by including general knowledge of each model, their pro/con</p> <p>[4] Explained concept of baseline model, but done multiple modelling - should be 1 baseline model</p> <p>[5] Support with external evidence of domain knowledge for higher score (not required)</p> <p>[6] Well done on README, clear and concise. Could include overview of pre-processing step (not detailed), i.e. Null value, outlier, etc treatment - provide better understand if your prediction are reliably derived</p> <p>[7] Generally good practice to lower &amp; snake case column name to prevent error when calling columns"</p>
----------	--