DS 1.1 Final Project Rubric

Term: _____ Student: _____

Date: _____ Instructor: _____

Project goal: Demonstrate mastery of data cleaning, exploration, analysis, and visualization

	At Risk (1)	Initial Mastery (2)	Approaching Mastery (3)	Mastery (4)
Problem Identification Identify the problem that needs data analysis and visualization	❑ Misidentified problem as data analysis	❑ Correctly identifies the problem as data analysis project	❑ Correctly identifies the problem as data analysis problem with interesting motivation and state why chose this project	
Dataset Preparation and Analysis Provide data cleaning Provide data manipulation Provide data visualization	❑ Does not contain any data cleaning or does not check if it needs cleaning ❑ Does not contain any data manipulation ❑ Does not report any analysis visually	❑ Mention that data needs or does not need cleaning ❑ Basic data manipulation was done ❑ Only one visualization was plotted	❑ Clean the dataset or check in the code that it is a clean dataset ❑ Acceptable level of data manipulation ❑ Couple of visualizations were in the presentation	❑ Provide data cleaning and join different datasets ❑ Provide many efforts on data manipulation ❑ Provide different, novel and meaningful visualizations
Statistical Analysis Report statistics of dataset/Outlier detection Comparison of samples among different groups Find correlation/Scatter plot	❑ Did not show statistics of dataset ❑ Did not provide comparison among groups ❑ Does not apply scatter plot or does represent the correlation among features	❑ Did show the statistics of dataset ❑ Talk about different classes/ groups ❑ Mentioned that correlation can be computed for the features	❑ Did show statistics about dataset and report outliers in the dataset ❑ Compared classes/groups in terms of mean, variance ❑ Compute correlation among features (no visualization)	❑ Did show statistics about dataset and report outliers in the dataset with box-plot (IQR) visualization ❑ Compared groups by t-test ❑ Compute the correlation and show it in scatter plot or heatmap
Communicating Results Comments and explanations of process Analysis and interpretation of results Write-up and presentation	❑ No comments or explanation of process or decisions made ❑ Very little or no analysis or interpretation of results ❑ Write-up or presentation unclear or communicated ineffectively to audience	❑ Some comments and explanation of process and decisions made ❑ Basic analysis and interpretation of results ❑ Written or presented to communicate results to technical audience	❑ Frequent comments and explanation of process and decisions made ❑ Thorough analysis and interpretation of results ❑ Written or presented to communicate results to technical audience	❑ Thorough analysis and interpretation of results including discussion of trade-offs (e.g., inductive bias of model choice, limitations of dataset, etc) ❑ Written or presented to communicate results to non-technical audience