Marking Rubric for Assessment 4

	subtasks	Weight (w.r.t task total)	HD	D	C	P	F
Task 1(65%)	Suburb	15%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	train_station_id	15%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	distance_to_train_station	5%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	travel_min_to_CBD	15%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	over_priced?	4%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	crime_A_average	3%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	crime_B_average	3%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	crime_C_average	3%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	closest_primary_school	3%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	distance_to_closest_primary	3%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	primary_school_ranking	11%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	closest_secondary_school	3%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	distance_to_closest_secondary	3%	80-100% of the entries are correct.	70-79% of the entries are correct.	60-69% of the entries are correct..	50-59% of the entries are correct.	0-49% of the entries are correct.
	secondary_school_ranking	12%	80-100% of the	70-79% of the entries are	60-69% of the	50-59% of the	0-49% of the

			entries are correct.	correct.	entries are correct..	entries are correct.	entries are correct.
	Other Data Integration Issues	2%	Detect and fix 80% or more of other data integration issues if they present.	Detect and fix 70-79% of other data integration issues if they present.	Detect and fix 60-69% of other data integration issues if they present.	Detect and fix 50-59% of other data integration issues if they present.	Detect and fix 0-49% of other data integration issues if they present.
Task 2 (15%)		100%	Different methods for data reshaping are well analyzed. The effects of using different data reshaping methods are shown in details. The student uses different attributes to evaluate the performance of linear regression models and gives details explanation. The student has made a convincing recommendation to the readers based on the observation in the analysis.	Different methods for data reshaping are reasonably analyzed. The effects of using different data reshaping methods have been demonstrated. The student uses different attributes to evaluate the performance of linear regression models and gives a detailed explanation. The student has made a reasonable recommendation to the readers based on the observation in the analysis.	Different methods for data reshaping are fairly analyzed. The effects of using different data reshaping methods have been demonstrated. The student uses different attributes to evaluate the performance of linear regression models and gives little explanation. The student has made a recommendation to the readers based on the observation in the analysis.	Different methods for data reshaping are poorly analyzed. The effects of using different data reshaping methods have been demonstrated. The student builds a simple linear regression model without much consideration and explanation. The student then rushes to a recommendation to the readers.	Different methods for data reshaping are badly analyzed. The effects of using different data reshaping methods have not been demonstrated. The student builds a simple linear regression model without any explanation. The student then rushes to a recommendation to the readers.
Methodology and Documentation (20%)	Methodology	40%	The report has demonstrated a good solution in a proper way. No error or only minor errors are found in the methodology. The methodology consists of all required steps for the tasks, and produces output with HD scores.	The report has demonstrated a reasonable solution. One key step is missing or leads to the errors in the output.	The report has provided a fair solution. Two key steps are missing or lead to incorrect output.	The report has provided a poor solution. Three key steps are missing or lead to incorrect output.	The report has provided a bad solution. Only one or two steps are demonstrated, which clearly cannot achieve the correct output.
	Documentation	60%	The report has proper sections and subsections (e.g.	The report has proper sections and subsections (e.g. introduction,	The report has some sectioning but it is not well	The report has some sectioning but it is not well organized.	The report has poor/no sectioning and it is not well

			introduction, methodology, conclusion,...). The methodology is explained properly and the code is well commented.	methodology, conclusion,...). The methodology is explained fairly and the code is fairly commented. The report can be improved.	organized. The explanation of the report is limited and the code is well commented.	The explanation of the report is poor and the code is well commented.	organized. The explanation of the report is poor and the code is poorly commented.

Individual Contribution Score (ICS):

The above rubric will be prorated based on each student's ICS, which will be calculated based on the contributions sheet, reflective diaries, and the interview performance (optional). For example: if a student achieves 50% ICS score, he/she will get only half the mark resulting from above rubric. Hence, please document your individual contributions thoroughly and be very well prepared during the interview to explain the steps you have taken to obtain your results based on a true understanding of the underlying concepts.