Project	**Exemplary**	**Acceptable**	**Inadequate**
Question Framing	**3**	**2**	**1**
10%	Question(s) framed are clear, interesting, and thoughtful. They involve one of the provided datasets and do not have an obvious answer.	Question framed is relevant to provided datasets	No clear or relevant question framed.
Data Cleaning			
15%	Data cleaning methods are sound and clearly utilized. Categorical values and Null values are clearly addressed in a rational way, and numerical variables are standardized. Data is in an easy ready-to-use format.	There are some problems with the resulting data, but not enough to materially affect the conclusion.. Some columns could have been addressed more suitably.	There is only minimal data cleaning. Techniques taught in class are used wrongly or not used.
Data Visualization	**3**	**2**	**1**
20%	Visualization is the appropriate kind for the kind of data (i.e. doesn't use histogram for qualitative data) Captions are informative, axes are properly labeled, plotting symbols and colors are judiciously selected. All graphs serve a clear purpose and are relevant to the framed question. The graphs have variety.	Several small issues are present in the graphs, or one major error, which are relevant to the data and the framed question.	There are no good visualizations, or the visualizations are irrelevant to the framed question. More than one major error.
Method and Experiments	**3**	**2**	**1**
20%	Methods are suitable for answering the proposed question(s). Many approaches are correctly tried on the dataset to address the question(s), even if they didn't eventually work out. The results of each method are clearly reported.	Some flaws in methods used or the execution of the experiments. The flaws are noticeable but not significant.	Serious flaws in methods used or experiments performed. Showed very limited work.
Analysis and Conclusion	**3**	**2**	**1**
25%	Made correct analyses based on the experimental results and answered soundly the 7 questions outlined in "Report Format and Submission" section in the guideline accurately. Explored opportunities for further research.	Some flaws in interpreting experimental results or errors in answering the 7 questions outlined in "Report Format and Submission" section in the guideline accurately. The flaws are noticeable but not significant.	Major flaws in interpreting experimental results. Either did not answer or answer incorrectly the majority of the 7 questions outlined.
Composition	**3**	**2**	**1**
10%	The report is clear and the text flows well. There is a clear structure in the writeup. Very few grammatical errors and spelling mistakes.	Parts of the writeup are confusing or not explained clearly. There isn't a clear structure in the write-up. There are some grammatical errors and spelling mistakes.	The write-up is very hard to understand and follow. There are many grammatical errors and spelling mistakes.

Any group that does not attempt any of the above sections will receive a 0 for that section, not a 1. Along with the above rubric, we expect the code to follow the above guidelines: uniform naming convention, clear and informative comments, follows the data science life cycle, easy to read and understand. **Failure to adhere to these guidelines may result in point deductions in the corresponding sections.**