Project Rubric:

Criteria	Unsatisfactory (0-49%)	Developing (50-69%)	Proficient (70-89%)	Excellent (90-100%)	Weight
1. Data Collection & Cleaning	Dataset is missing, unusable, or from a single source. No cleaning process described.	Dataset is messy, very small (n<20), or cleaning description is vague.	Dataset is clean, from 2+ sources, and n>40. Cleaning process is described well. The submitted file is usable.	Dataset is exceptionally well-curated and clean. Cleaning challenges are thoughtfully discussed.	**30%**
2. EDA & Visualization	No descriptive stats or visualizations.	Basic descriptive stats are present but no visualizations.	Correct descriptive stats and a relevant visualization are included and labeled.	Insightful EDA with excellent visualizations that clearly support the analysis.	**20%**
3. Probability Calculations	Calculations are missing or fundamentally incorrect.	Some calculations are present but contain major conceptual errors	All required probabilities (Marginal, Joint, Conditional, Bayes') are calculated. May have minor arithmetic errors.	All calculations are 100% correct, clearly shown, and use proper notation.	**20%**
4. Next Steps: Fitting the variables in your dataset to an appropriate distribution (to be discussed with me)	To be determined	To be determined	To be determined	To be determined	**20%**
5. Organization (Rmd file/Jupyter notebook)	A single script file is submitted. No organization.	A notebook is used, but it's a "wall of code" that is hard to follow. Fails to knit/run.	The notebook is clean, well-commented, and runs/knits. Clear headings separate analysis sections.	The notebook is a model of clarity: elegant, efficient, and easy to read. Markdown text and code are perfectly balanced.	**10%**
Total					**100%**