Ramsay Data Insights

Take-home Technical Test

Instructions:

- Submit your answers in a well-organised document or set of files, including any code (preferrable in a Github Repo) and visualisations.
- Ensure your work is clearly documented and explained.

Part	Questions
Part 1: Data Exploration and Preparation	1. Identify and describe 2 data quality issues present in the dataset. Briefly propose strategies to address these issues. Document the steps taken and provide a summary of the data quality improvements.
	2. Using the data provided create a feature that could be valuable for analysis or modelling. Explain the rationale behind the feature you created and how they might be useful for analysis.
Part 2: Data Analysis and Visualisation	3. Using the data provided produce a piece of analysis that describes to Ramsay which DRGs accrue the largest charges and your hypotheses for the drivers of these charge. Visualise these trends using appropriate charts or graphs and describe the results.

	4. Write an SQL query to calculate the total and average admissions for each month over the last two years. Include the month and year in the results. 5. Write an SQL query to analyse the distribution of `TotalCharges` by `PrincipalDiagnosis` and `Sex`. Use percentiles to describe the distribution.
Part 4: Strategic Insights and Recommendations	6. Based on your analysis, identify two strategic insights that could help Ramsay improve hospital operations or patient care. Justify your insights with evidence from your data analysis.
Part 5: Model Development	7. Describe a predictive model that could help Ramsay act upon one of the insights you found in Part 4. Describe your model purpose, model choice, preprocessing steps, and evaluation metrics.
Part 6: MLOps and Deployment	8. Describe how you would deploy the trained model into a production environment. Discuss how you would monitor its performance and handle any potential issues.