**Principles of Data Science**

**Justification For Question A**

**Justify your action for this task:**

It’s the process of handling missing values in a dataset. It emphasizes identifying which columns have missing values and the severity of these gaps to determine appropriate imputation methods. For categorical data, using the mode is recommended to preserve data distribution, while the median is suggested for the numerical ‘Seats’ column to minimize the impact of outliers. If the ‘New\_Price’ column has a significant percentage of missing values, dropping it may be better to maintain data integrity. The overall goal is to ensure that the dataset is clean and ready for analysis, minimizing bias. The code described uses straightforward imputation methods, filling missing values in bulk for efficiency and maintaining clarity for easy maintenance and adaptation.