Assessment Method Used	Characteristics of Incorrectly Answered Questions
BERTscore	• Listing Elements • Repeating Incorrect Elements
sacreBLEU	• Summarization Tasks
Manual Error Analysis	• Answering multiple questions • Mismatch of Nomenclature • Inferring Numerical Operations from Attribute Names • Ambiguous Questions