Scenario I:

Option A: Modifying the Instructor dimension by adding special rows representing instructor teams. This allows the Instructor dimension to match the intended grain and be included in the fact table. Strengths of this option include maintaining the integrity of the fact table by including all relevant dimensions. However, a weakness is that it requires additional processing and handling of the special rows representing instructor teams. It may also complicate reporting and analysis, as queries need to account for the special rows.

Option B: Changing the grain of the fact table to be one row per student enrollment per course per instructor. This option addresses the issue of multiple instructors by allocating enrollments equally among them. The strength of this option is its simplicity, as it aligns with the desired grain without requiring modifications to the dimensions. However, a weakness is that it may lead to fractional values in the EnrollmentCount field, which could be confusing for users and affect certain aggregations or calculations.

Option C: Creating two fact tables, one without the Instructor dimension and one with the Instructor dimension. This option allows for separate fact tables optimized for different types of queries. The strength is that it provides flexibility and performance optimization, as users can choose the appropriate fact table based on their query requirements. However, a weakness is the increased complexity in managing and maintaining two separate fact tables.

Question 2: I would choose Option C, creating two fact tables. This option provides the most flexibility and performance optimization. Users can select the appropriate fact table based on their specific query needs, allowing for efficient retrieval of data. Additionally, it avoids the complexities introduced by fractional values in Option B and the additional processing of special rows in Option A.

Question 3: If the majority of classes had multiple instructors, Option B might become more appealing. Having a significant number of classes with multiple instructors could make it impractical to allocate enrollments equally among them, leading to fractional values that may be difficult to interpret or use in calculations. However, if only one or two classes had multiple instructors, Option A or Option C could still be viable, as the impact on the overall system complexity would be minimal.

Scenario II:

Option A: The scores are attributes of the Customer dimension, and when scores change, the old score is overwritten with the new score (Type 1 Slowly Changing Dimension). Strengths of this option include simplicity and efficiency in terms of storage and maintenance. However, a weakness is that it does not preserve the historical information of customer scores, making it difficult to analyze changes in activity and profitability levels over time.

Option B: The scores are attributes of the Customer dimension, and when scores change, new Customer dimension rows are created using the updated scores (Type 2 Slowly Changing Dimension). The strength of this option is that it preserves the historical information of customer scores, allowing for analysis of changes in activity and profitability levels over time. However, a weakness is that it increases storage requirements and potentially complicates queries that need to consider different versions of customer scores.

Option C: The scores are stored in a separate CustomerScores dimension, which contains 45 rows representing different combinations of activity and profitability scores. The Trades fact table includes a foreign key to the CustomerScores dimension. The strength of this option is that it provides a direct and efficient way to join the fact table with the scores. However, a weakness is that it may not easily support future changes or additions to the scoring system, as it relies on the predefined rows in the CustomerScores dimension.

Question 6: I would choose Option B, which utilizes a Type 2 Slowly Changing Dimension. This option preserves the historical information of customer scores, allowing for analysis of changes in activity and profitability levels over time. It provides the necessary flexibility for understanding how and why customer scores change, which is of particular interest to the users. Although it increases storage requirements, the benefits of historical analysis outweigh this concern.

Question 8: In this design, we introduce a separate bridge table, called the Score History Bridge Table, to capture and track changes in customer scores over time. The main fact table remains the Trades fact table with the grain of one row per stock trade. The dimensions, such as Date, Customer, Account, and Security, remain the same. However, the Customer dimension would include a foreign key to the Score History Bridge Table.

The Score History Bridge Table would have the following structure:

CustomerID (foreign key to the Customer dimension)

ScoreType (indicating whether it is an activity score or profitability score)

Score

EffectiveStartDate

EffectiveEndDate