

## Data Warehousing Assignment

Submitted by: Aditya Singh Rajput  
Batch: BigData Batch 1

### SCENARIO (I):

#### Answer 1:

Option A:

**Strengths:** It allows to include the Instructor dimension in the fact table, while preserving the grain of one row per student per course enrollment. It also makes it easy to understand and interpret the results of the analysis.

**Weaknesses:** It can create confusion for users, as the instructor team may not always be the same for all courses.

Option B:

**Strengths:** It allows to include the Instructor dimension in the fact table, and provides a more accurate representation of the data, as it takes into account the multiple instructors for a course.

**Weaknesses:** It increases the complexity of the fact table, and may make it harder for users to understand and interpret the results of the analysis. It also requires additional calculations to aggregate the data properly.

Option C:

**Strengths:** It allows to include the Instructor dimension in the fact table, while preserving the grain of one row per student per course enrollment, and it allows to use the first fact table for other queries. It also allows to maintain the accuracy of data.

**Weaknesses:** It increases the complexity of the data model and may be more difficult for users to navigate and understand.

#### Answer 2:

Option A would be the best option in this scenario as it allows to include the Instructor dimension in the fact table while preserving the grain of one row per student per course enrollment, and it makes it easy to understand and interpret the results of the analysis.

#### Answer 3:

If the majority of classes had multiple instructors, option B or C would be a better choice as it allows to include the Instructor dimension in the fact table, and provides a more accurate representation of the data, as it takes into account the multiple instructors for a course.

If only one or two classes had multiple instructors, option A would be a better choice as it allows to include the Instructor dimension in the fact table while preserving the grain of one row per student per course enrollment, and it makes it easy to understand and interpret the results of the analysis.

#### Answer 4:

Another alternative design would be to create a separate table for multiple instructors, where we can store the instructor details, and instead of having the instructor dimension in the fact table, we can have the foreign key to this table. This will allow the fact table to maintain its grain, and will also provide the details of multiple instructors. However, this approach could increase the complexity of the data model and may be more difficult for users to navigate and understand.

#### SCENARIO (II):

#### Answer 5:

##### Option A:

**Strengths:** It is simple to implement and maintain, as the old score is overwritten with the new score. It doesn't require creating new dimension rows when scores change.

**Weaknesses:** It doesn't allow to track how and why the customer scores change over time.

##### Option B:

**Strengths:** It allows to track how and why the customer scores change over time. It also allows to maintain a history of customer scores.

**Weaknesses:** It requires creating new dimension rows when scores change, which can increase complexity and maintenance efforts.

##### Option C:

**Strengths:** It allows to track how and why the customer scores change over time. It also allows to maintain a history of customer scores.

**Weaknesses:** It requires creating a separate dimension table which can increase complexity and maintenance efforts.

##### Option D:

**Strengths:** It allows to track how and why the customer scores change over time. It also allows to maintain a history of customer scores.

**Weaknesses:** It requires creating a separate outrigger table which can increase complexity and maintenance efforts.

#### Answer 6:

Option B would be the best option in this scenario as it allows to track how and why the customer scores change over time. It also allows to maintain a history of customer scores.

#### Answer 7:

If the number of customers and/or the time interval between score recalculations was much larger, the option B would still be the best choice, as it allows to track how and why the customer scores change over time, and maintain a history of customer scores.

If the number of customers and/or the time interval between score recalculations was much smaller, option A could be considered as it is simple to implement and maintain, as the old score is overwritten with the new score.

#### Answer 8:

Another reasonable alternative design would be to create a separate table that stores the history of customer