**Dataware house assignment-2**

Option A: The strength of this option is that it keeps the instructor dimension true to the grain of the fact table, allowing for easy querying of instructor-related attributes. The weakness is that it can be confusing for users, as a single instructor row represents multiple instructors.

Option B: The strength of this option is that it allows for proper aggregation of enrollment counts when multiple instructors are present. The weakness is that it requires changing the grain of the fact table and can be confusing for users.

Option C: The strength of this option is that it keeps the instructor dimension true to the grain and allows for proper aggregation of enrollment counts. The weakness is that it requires creating and maintaining two separate fact tables, which can be confusing for users and increases complexity.

Question 2:

I would choose option C. Even though it increases complexity and requires maintaining two fact tables, it keeps the instructor dimension true to the grain and allows for proper aggregation of enrollment counts which is important for querying and reporting.

Question 3:

If the majority of classes had multiple instructors, I would still choose option C. This design allows for proper aggregation of enrollment counts, which is important for querying and reporting. If only one or two classes had multiple instructors, it would be better to consider using the alternative design to keep the complexity of the data model minimal.

Question 4:

Another alternative design would be to use a bridge table to connect the student-course enrollment fact table with the instructor dimension. This would allow for proper aggregation of enrollment counts when multiple instructors are present while keeping the instructor dimension true to the grain. However, this will increase complexity of the data model as well as the need for maintaining the bridge table.

Question 5:

Option A: The strength of this option is that it is simple to implement and maintain, as it only requires one version of the customer dimension. The weakness is that it does not allow for tracking of historical changes in customer scores.

Option B: The strength of this option is that it allows for tracking of historical changes in customer scores. The weakness is that it increases the complexity of the data warehouse, as it requires maintaining multiple versions of the customer dimension.

Option C: The strength of this option is that it separates the customer scores from the customer dimension, making it easy to update and track changes in scores. The weakness is that it increases the complexity of the data warehouse, as it requires a separate dimension and a join between the fact table and the dimension.

Option D: Similar to option C, this option separates the customer scores from the customer dimension and allows for easy tracking of changes in scores. The weakness is that it increases the complexity of the data warehouse, as it requires an outrigger table and a join between the customer dimension and the outrigger table.

Question 6:

I would choose Option B because it allows for tracking of historical changes in customer scores, which is important for the data analysts' analyses. It also separates the customer scores from the customer dimension, making it easy to update and track changes in scores.

Question 7:

If the number of customers was much larger, option B may not be the best choice as it will cause the dimension table to grow quickly. In this case, option C or D may be a better choice as they separate the customer scores from the customer dimension. If the time interval between score recalculations was much smaller, option B may also not be the best choice as it will cause the dimension table to grow quickly. In this case, option C or D may be a better choice as they separate the customer scores from the customer dimension.

Question 8:

Another reasonable alternative design would be to store the customer scores in a separate table, with a foreign key to the customer dimension. The table would have one row per customer per score recalculation, with columns for the activity and profitability scores. The advantage of this design is that it allows for easy tracking of historical changes in customer scores, while also keeping the customer dimension simple. The disadvantage is that it increases the complexity of the data warehouse, as it requires a separate table and a join between the fact table and the table.