DWH Asisignment 2:

Scenario 1

Question 1. What are the strengths and weaknesses of each option?

Option 1 strengths:

This is an effective solution as there will be teams formed for instructors so there will not be any repetition of rows. The number of rows in table will be lesser and hence analysis will be easier.

Option 1 weakness:

As there are teams formed, more processing should be done to get data on individual instructors

Option 2 Strengths:

Here since the question says less than 5% instructors are into multiple courses , this option would be because there will be more clarity while processing despite increased number of rows.

Option 2 weakness:

There will be multiple similar kind of rows(except change in instructor) , which takes more memory.

Option 3 strengths:

Since there are 2 fact tables there will be more clarity while wanting details particularly only about teacher or student

Option 3 weakness:

Difficult to draw information requiring both data together.

Wastage of memory.

Question 2. Which option would you choose and why?

I would choose option 2.

Here since the question says less than 5% instructors are into multiple courses , this option would be because there will be more clarity while processing despite increased number of rows.

Question 3. Would your answer to Question 2 be different if the majority of classes had multiple instructors? How

about if only one or two classes had multiple instructors? (Explain your answer.)

In this case I would choose option 1.

This is an effective solution as there will be teams formed for instructors so there will not be any repetition of rows. The number of rows in table will be lesser and hence analysis will be easier. Option 2 will not be suitable here because it will take too much memory.

Scenario 2:

Question 5. What are the strengths and weaknesses of each option?

Option 1 strength:

We can easily get the latest score

Option 1 weakness:

We lose the score history

Option 2 strength:

It keeps track of all data including past and present

Option 2 weakness:

Higher usage of memory

Option 3 strength:

Effective memory usage having all data

Option 3 weakness:

Analysts need to use complex sql queries

Option 4 strength:

none

Option 4 weakness:

Unrelated tables are in use and wanted data is missing

Question 6. Which option would you choose and why?

I would choose option 3 as Effective memory usage and has all data both present and past.

Question 7. Would your answer to Question 6 be different if the number of customers and/or the time interval

between score recalculations was much larger or much smaller? (Explain your answer.)

No it would still be the same because I feel in stocks irrespective of interval all data is important.