XML Q4

Compare and contrast your use of XML versus your use of MySQL for these data sets. Under what conditions does XML provide a better database approach? Under what conditions does MySQL provide a better database approach?

In question 1, the data set was not significant in size and not very complex. It wouldn’t have made sense, with such a small data set to create tables when we can get the results using a simple XML schema.

In question 3, the university dataset is comparatively larger and more complex than the baseball dataset used in question 1. It is more efficient to use MySQL for storing the data in a table structure format.

Under what conditions does XML provide a better database approach? Under what conditions does MySQL provide a better database approach?

* When the data is hierarchical in nature XML should be the way to go. An XML document contains information about the relationship of data items to each other in the form of the hierarchy. MySQL would be a better choice for inherently tabular data. (<https://www.ibm.com/support/knowledgecenter>)
* If the data design/schema requires frequent changes, XML makes for a natural choice. MySQL tables, like other relational databases, are quite rigid and not as flexible and making the changes may not be very cost and time effective and may even be more difficult to implement.
* When database performance is important, we are better off with MySQL. Unlike MySQL, serializing, parsing, and interpreting data in XML can make it a bit slower.
* Applicability of attributes to all/subset of data can also define if XML or MySQL should be used. For data that includes a very large list of attributes, many that are typically null (that is, sparse attributes), XML may be a better alternative than a long list of null columns or a normalized table. (<http://xml.coverpages.org/IBM-XML-GC34-2497.pdf>)
* When frequent updates to data are needed, MySQL is more beneficial. Updating a substantial number of rows can be costly with XML as XML data in an XML column is updated by replacing full documents. (<https://www.ibm.com/support/knowledgecenter>)
* Many situations involve highly structured information in very small quantities. Representation of that data with a MySQL model can involve complex star schemas in which each dimension table is joined to many more dimension tables, and most of the tables have only a few rows. A better way to represent this data is to use a single table with an XML column, and to create views on that table, where each view represents a dimension. (<https://www.ibm.com/support/knowledgecenter>)
* XML uses hierarchy and is very efficient with modeling parent/child relationships. MySQL should be used where data might be inherently hierarchical in nature, but the child components do not need the parents to provide value. For example, a purchase order might contain part numbers. The purchase orders with the part numbers might be best represented as XML documents. However, each part number has a part description associated with it. It might be better to include the part descriptions in a relational table, because the relationship between the part numbers and the part descriptions is logically independent of the purchase orders in which the part numbers are used. (<https://www.ibm.com/support/knowledgecenter>)