IST722: Unit 08 Participation Questions

**This is an individual assignment.**

**Before you begin, please make sure you’ve read and understand 1) our class honor code, 2) course policies on late work and 3) participation policies as posted on the syllabus. “I didn’t know” is not an excuse.**

**You should cite your sources in a standard format like MPA or APA and include a list of works cited.**

|  |  |
| --- | --- |
| Your Name: | Harsh Rajesh Darji |
| Your Email: | hrdarji@syr.edu |

# Instructions

Answer each of the following questions as concisely as possible. More is not necessarily better. Please justify your answer by citing your sources from the assigned readings from our textbooks, our class lectures, or online if directed to do so. Be sure to cite in text and include a list of works cited. Place your answer below each question. When you’re finished, print out this document and bring it to class as part of your participation grade.

# Questions

1. What are the goals of an ETL solution?

* **Extract** the data from its source
* **Cleanse and Conform** to improve data accuracy and quality (transform)
* **Deliver** the data into the presentation server (load)
* **Managing** the ETL process itself.

1. Is redundant data acceptable in a ETL solution? Explain.

Yes, redundant data is acceptable in a ETL solution. It should only be done when database is unable to perform

It’s operations. This is called De normalization:

It can make implementation more complex.

It hampers flexibility.

It might slow downs updates.

So, even if it allows redundant data, we can get unsatisfactory performance and relation has low update rate and high query rate.

1. Why is ETL the most time-consuming step in a DW/BI project?

Business intelligence projects are more time consuming and they require a successful methodology to employ all the business-related operations. According to Kimball’s approach, the business intelligence model suggests:

* Understanding the requirements and delivering the valuable business output
* Act in accordance with the proven DW Lifecycle.
* Building and delivering progressively within the organization’s data framework
* Designing flexible, useful and high-performance datasets.
* Providing a complete business solution via reports, query tools, documentation.

WORKS CITED: