Course:  INFO 531/ISTA 431: Data Warehousing and Analytics in the Cloud

Module/Week:  2 - Week of January 18, 2023

Topic:  SQL Basics; ETL Metadata Models

**Week 2 Assignment**: Metadata Models for Data Warehousing

PURPOSE:

This assignment introduces you to an ETL (Extract-Transform-Load) Metadata Model in a data warehouse. In enterprise data warehouses hundreds of batch cycles run to refresh the tables in different subject areas. The batch jobs run in different frequencies. To manage the batch cycles and jobs and refresh data warehouses in a timely manner and with accuracy, a metadata-driven batch cycle refresh is very important. The articles for this assignment provide details of how batch cycles and jobs run and how the load metrics are captured in the metadata tables.

**The assignment will consist of 50 points.**

Write a single-spaced, one-and-half-page summary of the learnings from one of the below two articles. The last paragraph of your write-up should list at least 5 key takeaways from the article. This article is uploaded to the Reading subfolder of the Week Module in D2L.

**1. Rahman, N, Marz, J., and Akhter, S. (2012). An ETL Metadata Model for Data Warehousing. *Journal of Computing and Information Technology (CIT)*, 20(2), 95–111.**

**2. Vetterli, T., Vaduva, A., and Staudt, M. (2000). Metadata Standards for Data Warehousing: Open Information Model vs. Common Warehouse Metadata, ACM SIGMOD Record, Vol. 29, Issue 3, September 2000.**

**Upload your submission to the Week 2 Assignment Dropbox in D2L.**