

Lab 3

Data Acquisition and ETL

- A short discussion of your experience connecting data in Domo and creating ETL processes.
 - Discuss any issues you ran into while going through the exercise.
 - List each of the ETL processes you completed and what they are supposed to accomplish.
-
- The review of the data import and ETL process in Domo was positive. The importing of datasets was smooth and efficient, with manual update setting providing necessary control for the practice exercise. The creation of the ETL process and integration of multiple datasets generated a comprehensive overview of the data and an efficient approach to data processing
 - During the process of combining multiple data sets, I inadvertently assigned an incorrect data type to one of the columns. This resulted in an error message indicating that all data types within the columns must be the same. I immediately rectified the issue by assigning the correct data type other than this small known mistake rest of things were good and no issues were faced
 - Throughout the ETL process, I have executed a sequence of data join operations on 7 input datasets, creating 6 intermediate join datasets and 1 final output dataset. The first step was to merge the rental and customer datasets based on the customer_id column, resulting in the rental data. Then, I joined the rental data with the inventory dataset using the inventory_id column, yielding the inventory data. The film_id column was used to integrate the film, inventory data at first and combined the film data and film_category datasets, producing film category data. Finally, I combined the film category data with the category dataset using the category_id column, resulting in the final output dataset which is DVDPurchases. The utilization of join data objects enabled efficient integration of the imported tables, ensuring the accuracy of the combined data.
 - I have accomplished it by combining tables in the process of merging the data from multiple tables into a single table. In the context of the mentioned tables (Customer, Rental, Payment, Film_category, Inventory, Film, and Category), the data from each of these tables is combined to create a single, comprehensive table that contains all relevant information related to the categories of DVD's which helps in better analysis of the datasets.

