

Week 11 Homework

Please refer to the Week 11 video as further details about this homework have been addressed during the class.

Instructions

Modify your individual project to de-normalize the customer information (Refer to **01-high-level-technical-architecture.pdf** for overall architecture). Here are the general steps to follow:

1. Modify **extract-customer.ktr** so to use the **oltp_hr** as the source in extracting the tables into **source_db**. You are to use the same DDLs in the **oltp_hr** to create corresponding tables in **source_db** except that all the foreign key constraints must be removed in **source_db** (That is, do not include the **CONSTRAINT** clause of the DDLs).

Here is the connection information for the **oltp_hr** :

- **Host:** rdsinstance1.cg057s6hdp7p.us-west-2.rds.amazonaws.com
- **User:** etl_read
- **Pass:** etl_read_user
- **DB:** oltp_hr

(You can hard-code the DB user/pass of **oltp_hr** in your ETL)

2. Run the **05-ddl-script-customer-stg.sql** provided to you to create the **stage_db.customer_stg** table. Create a new transformation, **trans-denorm-customer.ktr**, in the **02-transfer** folder, to de-normalize the extracted tables into **stage_db.customer_stg**. Refer to **02-physical_data_model_oltp_hr.pdf** and **08-customer-denormalization.sql** for the entity-relationship of the customer related tables.
3. Modify the **main-datamart-sales.kjb** so that the transformation created in the previous step is placed between the **extract-all** and the **dim-load-all**.



4. Run the **06-ddl-script-dim-customer.sql** provided to you to re-create the **datamart_db.dim_customer** as two fields in the table have been renamed as below:
 - address_line1 → **home_address_line1**
 - address_line2 → **home_address_line2**

Modify **load-dim-customer.ktr** so that it uses the **stage_db.customer_stg** as the source in loading the **datamart_db.dim_customer**.

5. Make sure your ETL main job, **main-datamart-sales.kjb**, completes successfully and all the tables were populated correctly.
6. Add any additional DDLs you created into the **etl-datamart-sales/ddl-scripts.sql**

Extra Credits (10 points maximum)

Modify your individual project to de-normalize the product information. Here are the overall steps:

1. Modify **extract-product.ktr** so to use the **oltp_product** as the source in extracting the tables into **source_db**. You are to use the same DDLs in the **oltp_product** to create corresponding tables in **source_db** except that all the foreign key constraints must be removed in **source_db** (That is, do not include the **CONSTRAINT** clause of the DDLs).

Here is the connection information is the same as that for the **oltp_product** :

- **Host:** rdsinstance1.cg057s6hdp7p.us-west-2.rds.amazonaws.com
- **User:** etl_read
- **Pass:** etl_read_user
- **DB:** oltp_product

(You can hard-code the DB user/pass of **oltp_product** in your ETL)

2. Create a new transformation, **trans-denorm-product.ktr** in the **02-transfer** folder, to de-normalize the extracted tables into **stage_db.product_stg** (The DDL script for this table must be created on your own). Refer to **03-physical_data_model_oltp_product** and **09-product-denormalization.sql** for entity-relationship.
3. Modify the **main-datamart-sales.kjb** so that the transformation created in the previous step is placed between the **trans-denorm-customer** and the **dim-load-all**.
4. Modify **load-dim-product.ktr** so that it uses the **stage_db.product_stg** as the source in loading the **datamart_db.dim_customer**.
5. Make sure your ETL main job, **main-datamart-sales.kjb**, completes successfully and all the tables were populated correctly.
6. Add any additional DDLs you created into the **etl-datamart-sales/ddl-scripts.sql**

Deliverables

Submit the entire folder, **etl-datamart-sales**, in a single zip file with the naming convention below:

yourID_week11.zip
(e.g. dlee52_week11.zip)