

I first created each staging table, where stage_teams, stage_players, stage_tournaments and stage_dates have a source_db property to identify which database they are from. Each staging table also has an incrementing surrogate key.

Before inserting data into each staging table, I first checked to see whether the data is already in the table in order to prevent duplicates. Data inserted from database 2 into the stage_tournaments table had the tour_prize property modified to convert from dollars to euro. Dates are extracted from the tournament tables and inserted into the stage_dates table.

Certain data inserted into the stage_facts table that may not have a corresponding surrogate key in another dimensional table as it is duplicate data is updated. An example of this is Tiger Woods with a SK of 1 and 7, the SK of 1 is located in the dim_player table, so the SK of 7 in the stage_facts table needs to be updated to reflect this.

Example query:

```
/* Select all tournaments tiger woods participated in */
select dp.player_name, dt.tournament_desc
from dim_tournaments dt
join fact_results fr on fr.tournament_sk = dt.tournament_sk
join dim_players dp on dp.player_sk = fr.player_sk
where dp.player_name = 'Tiger Woods';
```