ETL Assignment

Aqib Hameed

1   Aircraft Dimention

From the excel file we load the data to the Aircraft diemention table, to improve the data quaility we remove the dublicates values, first we sort the rows by registration code then with the help of unique rows template we will eliminate identical enteries then we load the data into the aircraft dimention table.

2   People Dimention

The perposon who report about aircraft record exist in both places like xml file and in the TLB table.To improve the consistency and data completeness is crossing data from one source wih other source.We first filter the data from TLB table on the basis of role, if the role is 'M' then we join the record with MAREP xml file, and remove the dublicate value of the reporteurid. As we know pilots independent of the Airport and the Airport value will be null and then we agian join the record and remove the dublicates, then for recoverability we add blocking steps until to wait for intermediate steps, in order to start filling target tables when all the steps are finished. Like that, we avoid FK and PK constraint problems and deleting each time data from those tables in case of errors.

3   LogBookReporting Fact

From the TLB table we fetch the record of person who report the aircraft and then we group by monthid, personid and aircraftid and count the value and then load the record in the LogBookReporting dimention table.
To improve the performance the same task run in Parallelization, by adding mutiple copies(threads). So we add parallelization in the memory group by tempelate.

4 MaintenanceEvents

In the Maintenance event business rule dealing with the respect to duration
aircraft on ground less than an hour
Maintenance greater than one hour and less or equal than one day
Revision greater than one day and less or equal than one month.
Filter the record of the basis of the Maintenance kind and Revision detecting
the overlaping record.
Two Slots of the same aircraft cannot overlap also dealing in the Maint-
EventsInSlots.

5 AircraftUtilization Fact

As In the Aircraftdimention table records come from four different sources
like Maintenance event, Flights without delays and canceled, flight with de-
lays and OperationInterruption delays and at the end we group by aircraftId
and timeId and calculated the measurements of this fact table and load into
the AircraftUtilization fact table.

6 Month

In the month table record comes from different sources like TLB, Flights
without delays and canceled, maintenance events. Remove the dublicates
value and then load the record into the Month table.

7 TemporalDimension

In this table record comes from different sources like light with delays, Oper-
ationInterruption delays, Flights without delays and canceled, maintenance
events. Remove the dublicates value.Wait untill when in the month table
record load.
After that load the record into the TemporalDimension table.