

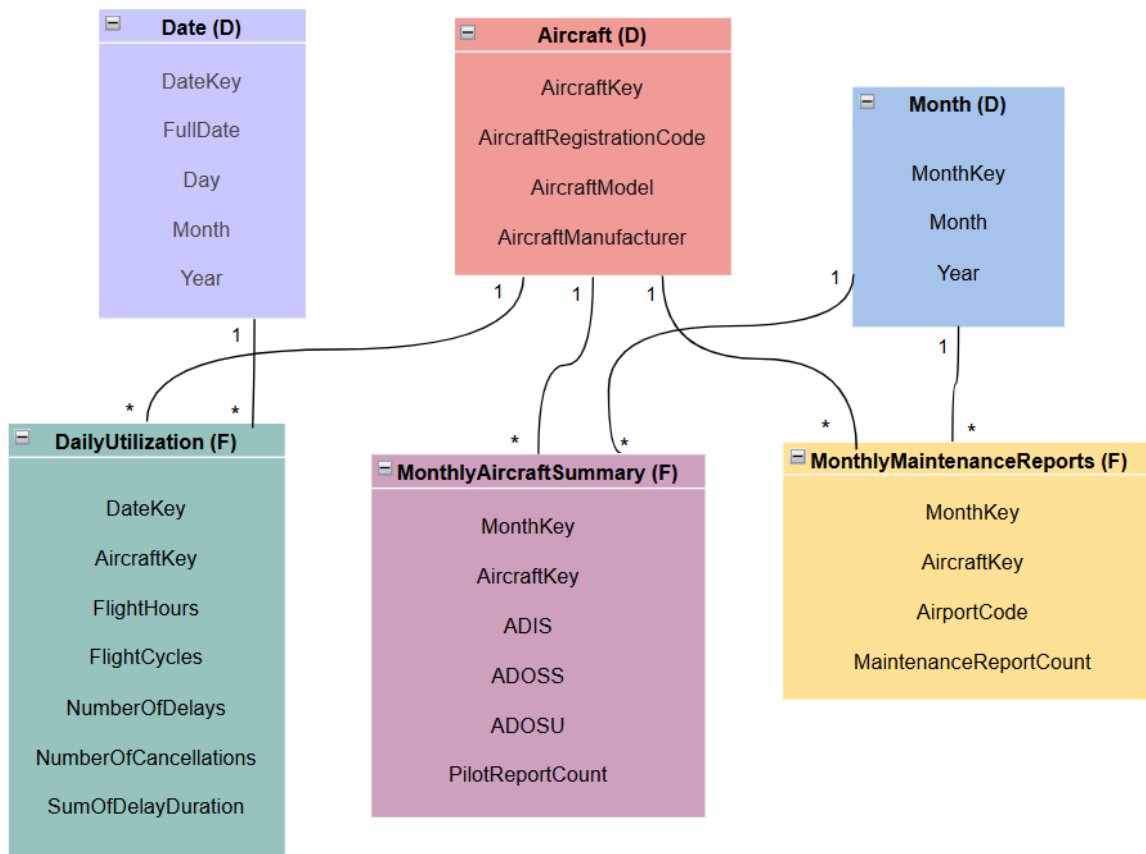
Universitat Politècnica de Catalunya
Grau en Ciència i Enginyeria de Dades

Projecte Data Warehousing

Grup 12A:

Rebeca Torrecilla Domínguez
Joel Márquez Álvarez

Bases de Dades Avançades
2025-2026

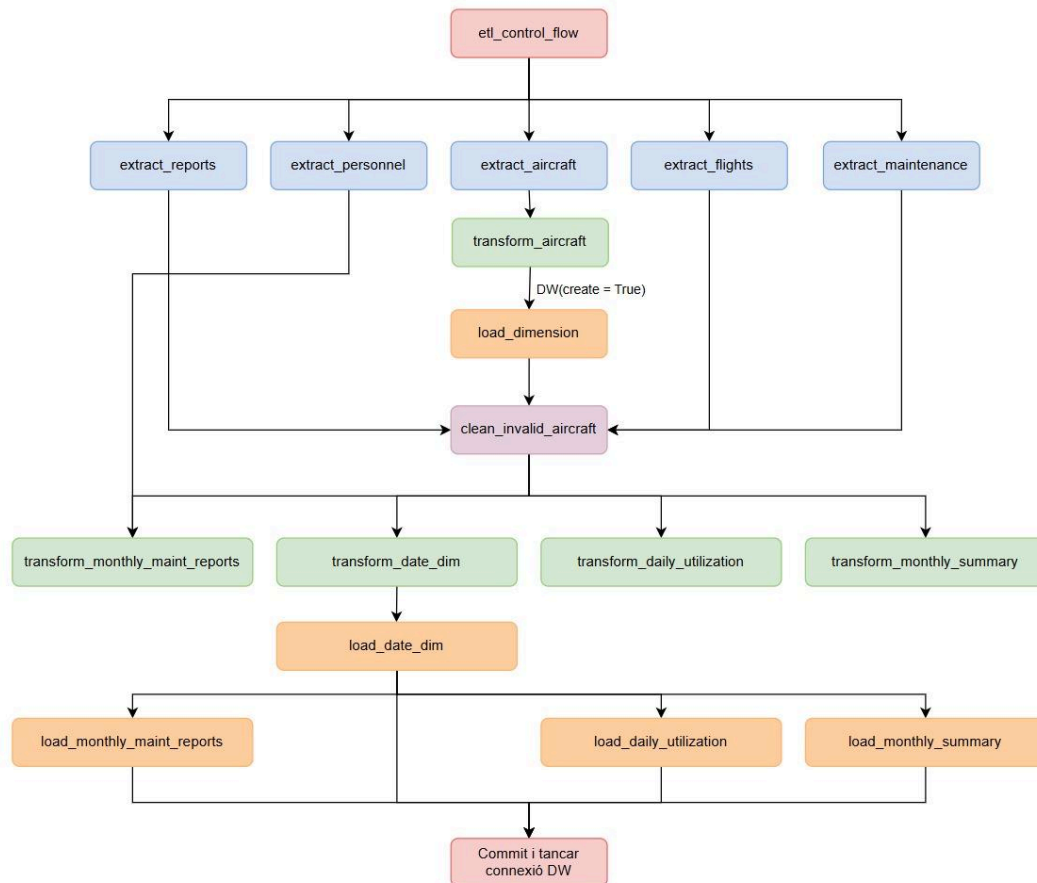


El disseny final es basa en un esquema en constel·lació amb tres taules de fets, una arquitectura escollida per gestionar de manera eficient les diferents granularitats dels KPIs: DailyUtilization per a les dades diàries per aeronau, MonthlyAircraftSummary a nivell mensual per aeronau, i MonthlyMaintenanceReports per als reports de manteniment a nivell mensual per aeronau i aeroport.

Una decisió d'optimització va ser la creació de la dimensió Month, separada de Date, per dues raons: primer, per claredat, evitant la convenció d'usar el primer dia del mes com a clau, i segon, per rendiment, ja que les consultes mensuals fan JOIN contra una dimensió molt més petita.

També es va considerar un model alternatiu amb una única taula de fets mensual per reduir joins, però es va optar pel disseny actual per evitar haver d'incloure un AirportCode que seria irrellevant per a la majoria de mètriques com l'ADIS, resultant en una taula de fets molt més gran (i redundant). En especial, es va voler evitar el problema de doble comptabilització quan un avió tingués informes de diferents aeroports en un mateix mes, duplicant mètriques com l'ADOS en fer les consultes si no es tracta amb cura.

Per tot això, s'ha optat per aquest model, més intuïtiu i també flexible en disposar ja d'una taula de fets on qualsevol mètrica futura de manteniment dependent de l'aeroport s'integri fàcilment.



El control flow (aquesta imatge) i data flow (següent pàgina) son força estàndard, l'únic punt destacable és que per fer la neteja de dades, hem de carregar Aircraft abans per permetre els lookups dels registres vàlids. Fer la neteja al principi assegura que les dimensions Date i Month només incloguin dies amb activitat vàlida, reduint la seva mida. Respecte a les consultes:

- **Query Aircraft Utilization:** resultats idèntics, amb només petites diferències decimals pels càlculs intermitjos. DW va trigar 0.0419 segons, 19.6 vegades més ràpid que Baseline amb 0.8230 segons.
- **Query Reporting:** resultats idèntics, i DW va ser més ràpid, completant la query en 0.0394 segons (9.6 vegades més ràpid, 0.3781 segons del Baseline)
- **Query Reporting per Role:** resultats idèntics, DW amb 0.0371 segons, aproximadament 10.5 vegades més ràpid que Baseline (0.3900 segons)

Tot això ha sigut possible gràcies al disseny exposat, pensat per pre-calcular i simplificar les queries per fer anàlisi de dades.

