DW Project Week 2: Multidimensional Design

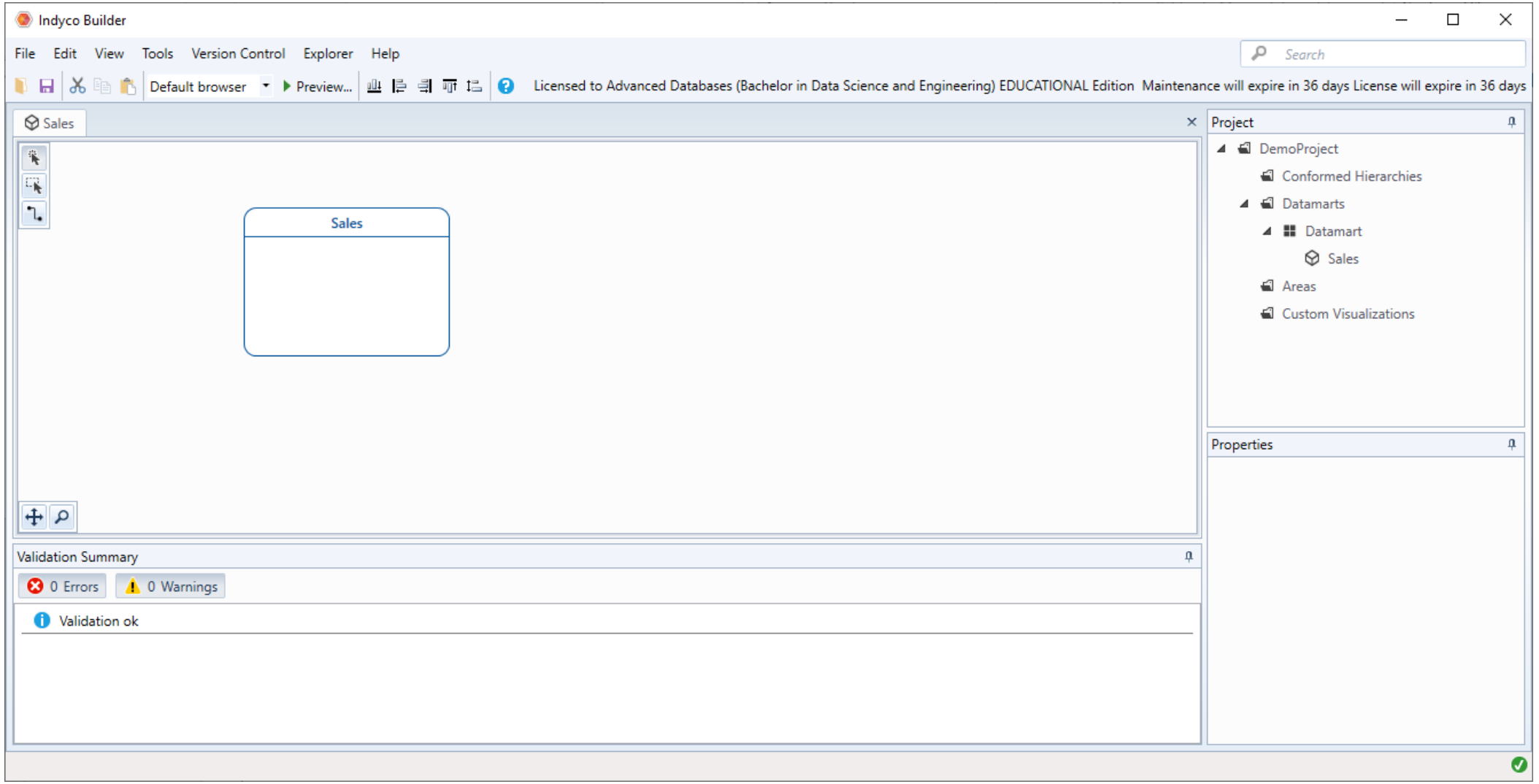
Design a multidimensional model in Indyco Builder

Week 2: Multidimensional modeling

Prerequisites

- Connect to PostgreSQL using DBeaver (Week 1)
  - You will find two databases: AMOS and AIMS
  - Understand the domain
- Install Indyco Builder (Week 2)
  - Only for Windows
  - Follow the steps provided in LearnSQL
- Tutorial on Indyco Builder (Week 2)
  - Learn how to model with Indyco Builder using the tutorial provided in LearnSQL

Demo: Indyco Builder



Week 2: Multidimensional modeling

Check the slides of the ACME use case to see how KPIs are calculated:
RRh – Report Rate per hour
RRc – Report Rate per cycle
PRRh – PIREP rate hours
PRRc – PIREP rate cycles
MRRh – MIREP rate hours
MRRc – MIREP rate cycles

Check the slides of the ACME use case to see how KPIs are calculated:
FH – Flight hours
TO – Flight cycles
ADOS – Aircraft Days Out of Service
DU – Daily utilization
DC – Daily cycles
DYR – Delay rate
CNR - Cancellation rate
TDR – Technical Dispatch Reliability
ADD – Average Daily Duration

LAB ON MULTIDIMENSIONAL MODELING FOR THE ACME-FLYING USE CASE

You must create a multidimensional model, potentially consisting of different stars, that allow to easily retrieve the KPIs about aircraft utilization (namely FH, TO, ADIS, ADOS, ADOSS, ADOSU, DU, DC, DYR, CNR, TDR and ADD) and logbook reporting (namely RRh, RRc, PRRh, PRRc, MRRh and MRRc). All these metrics would be obtained from the available data in AIMS and AMOS, that can be complemented with the following sources:

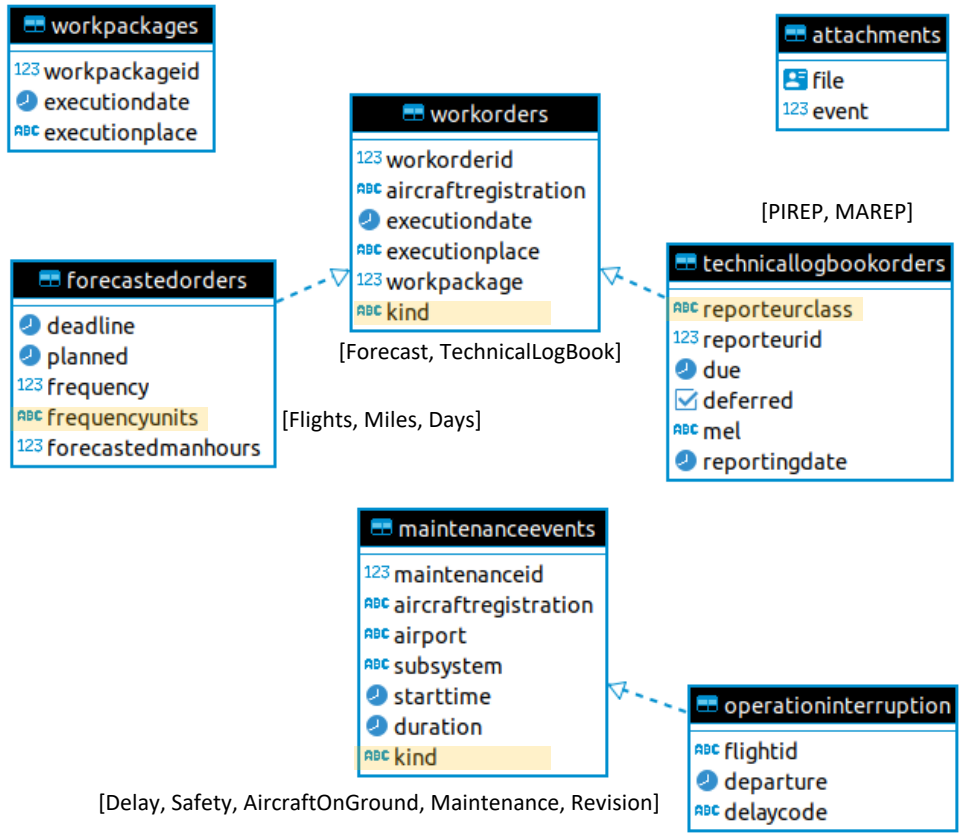
- A file containing for every aircraft registration code, its manufacturer registration code, the aircraft model and manufacturer. aircraft-manufacturerinfo-lookup.csv
- Another file containing for the maintenance personnel, their identifier and the code of the airport where they work. maintenance-personnel-airport-lookup.csv

Finest temporal granule for flight hours and cycles is the day, while the ADIS and ADOS are calculated per month or year, like the DYR, CNR, TDR, ADD, and all Report Rates. The latest can also be analysed per person and airport of the reporting person (just in case of MAREP), as well. Thus, required queries are:

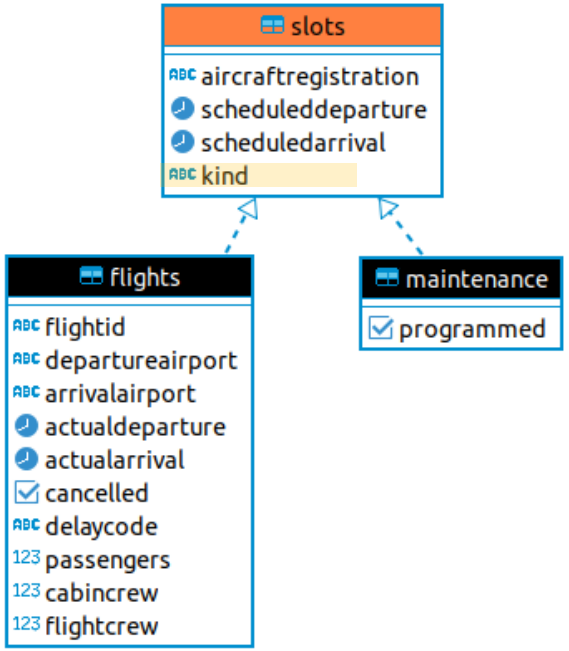
a) Give me FH and TO per aircraft (also per model) per day (also per month and per year).
b) Give me ADIS, ADOS, ADOSS, ADOSU, DYR, CNR, TDR, ADD per aircraft (also per model) per month (also per year).
c) Give me the RRh, RRc, PRRh, PRRc, MRRh and MRRc per aircraft (also per model and manufacturer) per month (also per year).
d) Give me the MRRh and MRRc per airport of the reporting person per aircraft (also per model).

Data sources

AMOS
Aircraft Maintenance Operation System



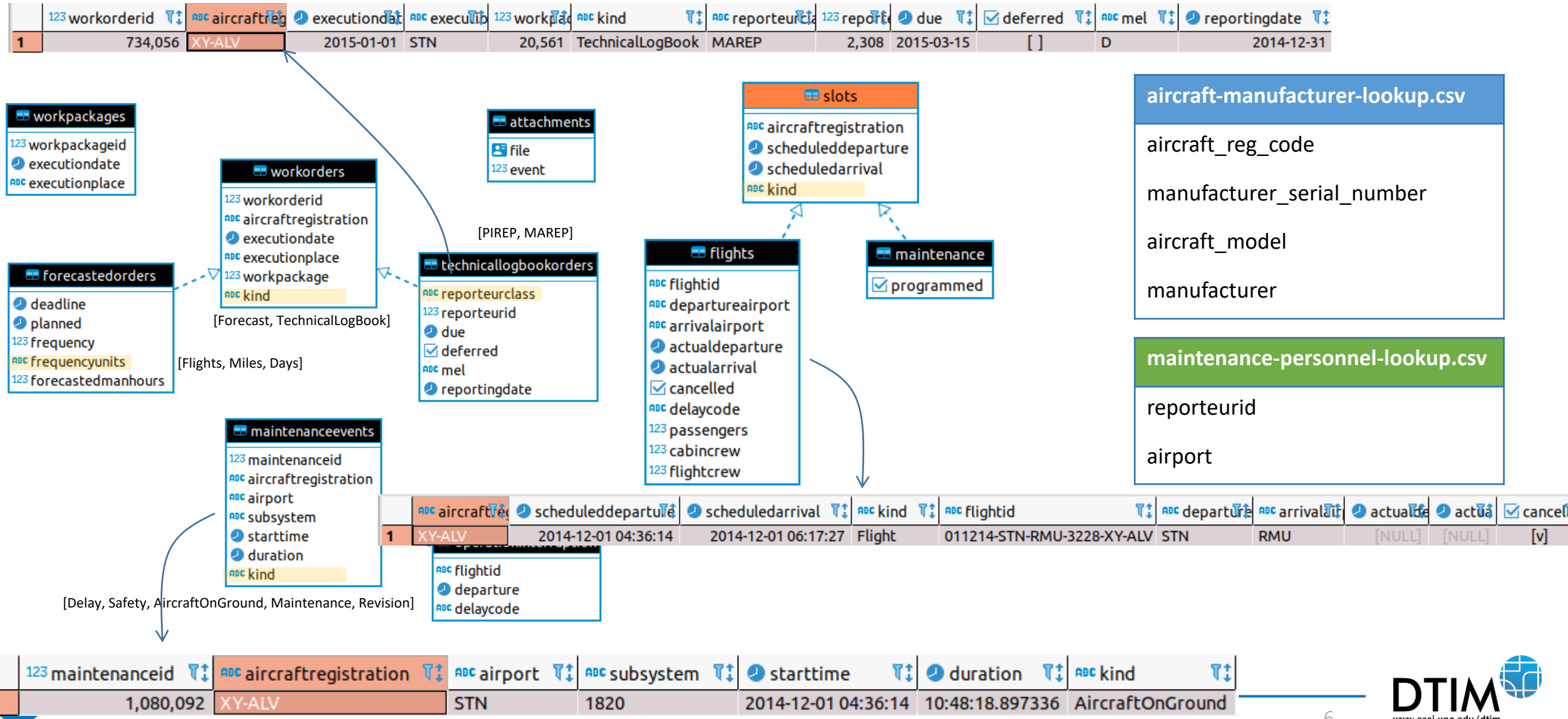
AIMS
Air Information Management System



aircraft-manufacturer-lookup.csv
aircraft_reg_code
manufacturer_serial_number
aircraft_model
manufacturer

maintenance-personnel-lookup.csv
reporteurid
airport

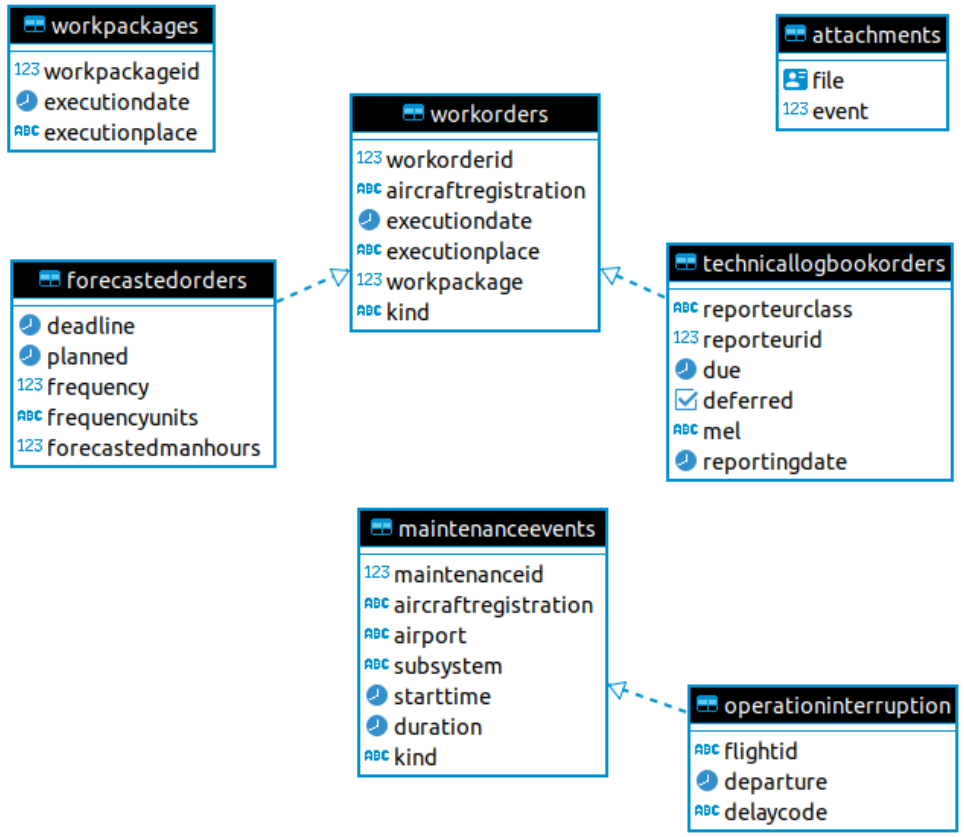
Data sources



DTIM
www.essi.upc.edu/dtim

Data sources

AMOS
Aircraft Maintenance Operation System



KPIs (see slides)

Flight hours (FH)
Flight cycles (TO) – number of take offs
Aircraft days out of service
- ADOSS – scheduled maintenance [days]
- ADOSU – unscheduled maintenance [days]

…

Requirements / Queries (see Statement)

1. Give me FH and TO per aircraft (also per model) per day (also per month and per year)
2. Give me ADIS, ADOS, ADOSS, ADOSU, DYR, CNR, TDR, ADR per aircraft (also per model) per month (also per year)

…

Week 2 | Activity 1.1

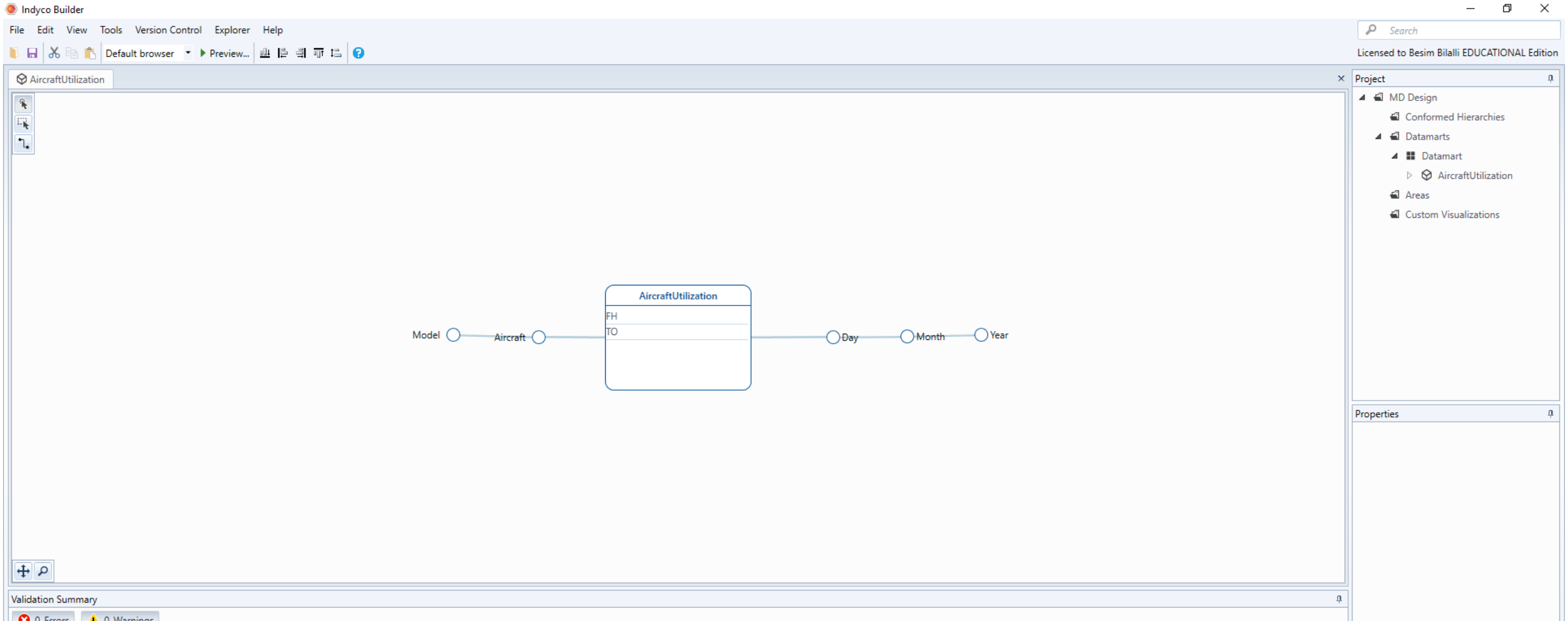
- Given the required KPIs and queries, **create independent MD schemas, in the Indyco Builder tool**, that satisfy (i.e., that can answer) these information requirements.

For instance,

- KPIs: **FH and TO**
- Req1: **Give me FH and TO per aircraft (also per model) per day (also per month and per year)**
  - Can I compute these KPIs using the data sources I have?
  - Do I have the required dimensions and do they satisfy the multidimensional modeling constraints?

Week 2 | Activity 1.1 | Independent MD schema

- Req1: Give me FH and TO per aircraft (also per model) per day (also per month and per year)



Week 2 | Activity 1.2

a)	Give me FH and TO per aircraft (also per model) per day (also per month and per year).
b)	Give me ADIS, ADOS, ADOSS, ADOSU, DYR, CNR, TDR, ADD per aircraft (also per model) per month (also per year).
c)	Give me the RRh, RRc, PRRh, PRRc, MRRh and MRRc per aircraft (also per model and manufacturer) per month (also per year).
d)	Give me the MRRh and MRRc per airport of the reporting person per aircraft (also per model).

- Given the required MD schemas, **consider integrating them**:
  i. you obtain a unified MD schema that satisfies all information requirements at once, and
  ii. the unified MD schema has the minimal structural complexity (regarding the number of classes, attributes, and relationships).

  - Do fact tables share any dimensions?
  - Can I compute a galaxy schema out of the independent MDs?

Week 2 | Activity 2

- Given a simple multidimensional (i.e., the previously created galaxy) schema:
  i. Propose a logical database schema (i.e., a set of CREATE TABLE statements) corresponding to that multidimensional schema.

  - How do I physically store the designed model?
  - Do I use only Tables or maybe also Views?

Week 2 | Activity 2 | Logical Schema

- Propose a logical database schema (i.e., a set of CREATE TABLE statements) corresponding to that multidimensional schema.

```
CREATE TABLE AircraftUtilization (
aircraftID CHAR(6),
timeID DATE,
flightHours NUMBER(2),
flightCycles NUMBER(2),
PRIMARY KEY (aircraftID, timeID),
FOREIGN KEY (aircraftID) REFERENCES AircraftDimension(ID),
FOREIGN KEY (timeID) REFERENCES TemporalDimension(ID)
);
```

```
CREATE TABLE AircraftDimension (
ID CHAR(6),
model VARCHAR2(100) NOT NULL,
manufacturer VARCHAR2(100) NOT NULL,
PRIMARY KEY (ID)
);
```

```
CREATE TABLE TemporalDimension (
ID DATE,
month CHAR(7) NOT NULL,
year NUMBER(4) NOT NULL,
PRIMARY KEY (ID),
);
```

Deliverables

Deliverables:

1) Indyco folder (in a single zip file) with all the files of the conceptual design
2) Commented SQL file with the CREATE TABLE statements and auxiliary views (if any) in Postgres
3) PDF file (**one single A4 page, 2.5cm margins, font size 12, inline space 1.15**) with all assumptions made and justifying the decisions you made (if any)

Assessment criteria:

i) Conciseness of explanations (only first page will be considered in the evaluation)
ii) Understandability
iii) Coherence
iv) Soundness

Evaluation:

- 60% Deliverables
- 40% Exercises related to the project done individually in the day of the partial exam