# **TASCA 5: ESQUEMA MODEL I CÀRREGA DA UN MAGATZEM DE DADES (DATAWAREHOUSE)**

### Objectiu

### Executar la càrrega de dades transformades a un magatzem de dades real en PostgreSQL, estructurant-les segons un **model dimensional** dissenyat per l’alumne (estrella, cop de neu o galàxia), tenint en compte criteris d’escalabilitat, eficiència i facilitat.

### Requisits previs

* Entorn Python preparat (Docker ETL)
* Dispondre d’un conjunt de dades netes,normalitzades i enriquides( Tasca4)
* Coneixements bàsics de Python y Pandas

### Procediment pas a pas

1. Decidir el tipus de model de dades amb el que vol contruir el Datawarehouse: **estrella** o **cop de neu**.
2. Definir l’esquema i la relació entre fets i dimensions.
3. Definir la granularitat de la taula de fets
4. Crear l’estructura de la BD i carregar les dades mitjançant Pandas.



*Exemple: Ús del mètodeto\_sql de Pandas*

1. Validar les dades introduïdes.

### Notes Importants

* Defineix clarament les claus: Les taules de dimensions han de tindre claus primàries, i la taula de fets ha de contenir claus foranes que enllacen amb aquestes.
* Escull el granulat adequat: El nivell de detall de la taula de fets ha de ser coherent amb les preguntes d’anàlisi que es volen respondre. Un granulat massa fi pot generar dades innecessàriament pesades; massa gros pot limitar les anàlisis

# ENTREGA

La prova d’entrega de la Tasca 5 inclourà:

* Codi font .py, amb totes les funcions de creació i càrrega de dades amb Pandas.
* Diagrama de l’esquema dimensional o explicació escrita clara.
* Captures o consultes que mostren les dades carregades