



# PRÀCTICA 1

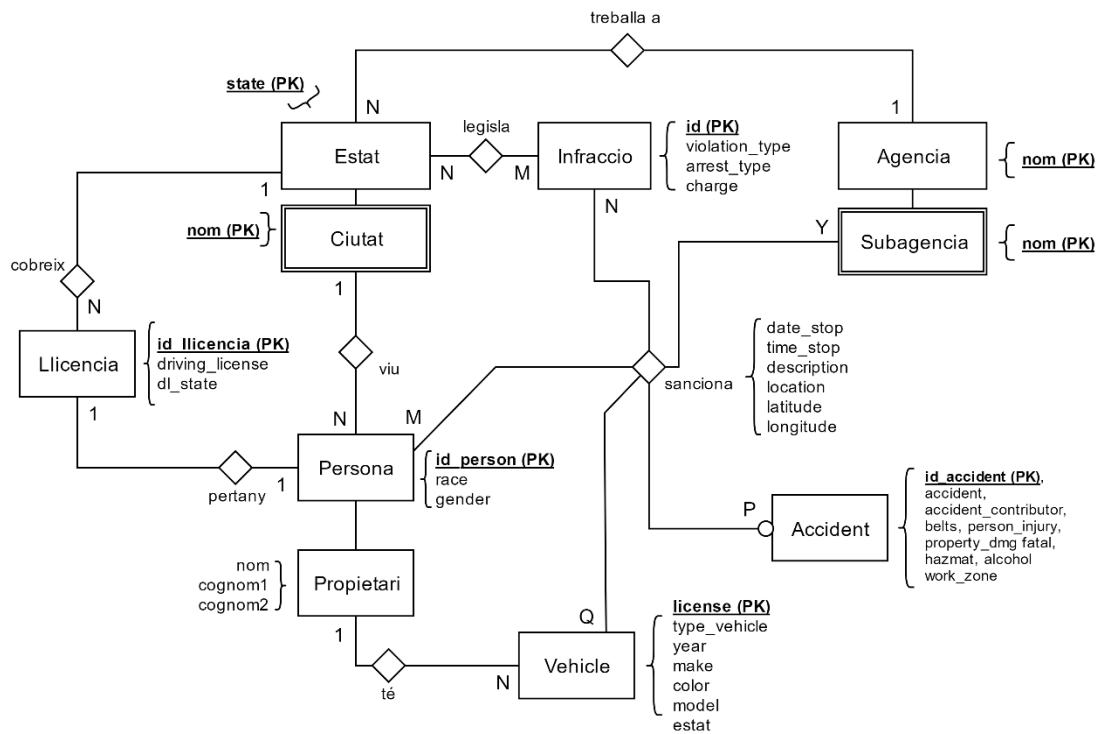
Arxius i Bases de Dades

Joel López – joel.lopez.2015  
Alex Pons – ls31155

## Índex

1. Model Conceptual (diagrama Entitat-Relació) .....	2
2. Model Lògic (diagrama Relacional).....	4
3. Estudi de mercat dels SGBD .....	5
3.1 Han fet bé els policies en sol·licitar l'estructuració de les seves dades en el programari emprat? .....	7
3.2. Com s'han solucionat les incoherències de les dades? .....	7
4. Model Físic .....	8
5. Test del procés d'importació. ....	14
6. Conclusions .....	20

# 1. Model Conceptual (diagrama Entitat-Relació)

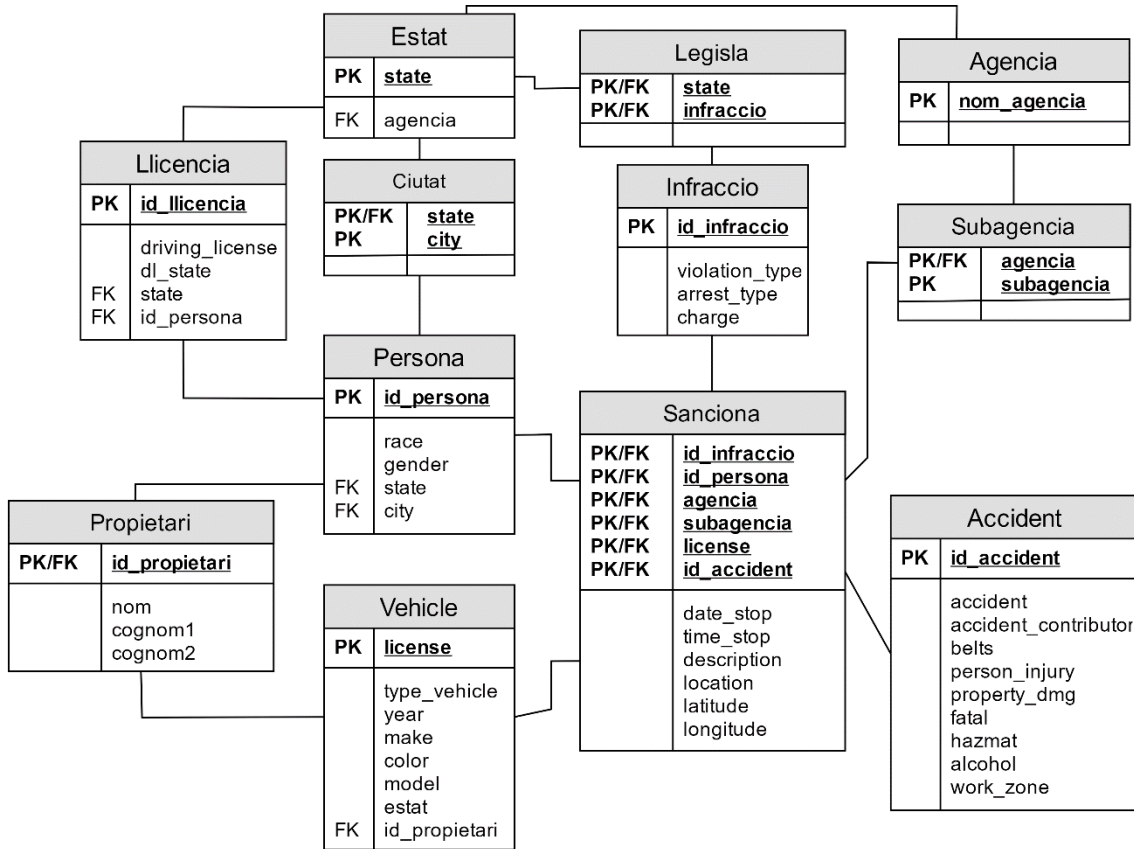


Aquest es el nostre model conceptual de la pràctica 1 de bases. Hem estructurat la informació en diferents entitats: Agència, Subagència, Accident, Vehicle, Infracció, Propietari, Persona, Estat, Ciutat i Llicència. A continuació expliquem les seves relacions i el perquè:

- Agència té una relació 1: N amb Estat, perquè hi ha una sola Agència (l'estació policial del Comtat de Montgomery) per tots els Estats. En cas d'haver més d'una, hauríem de canviar la relació per una N-M.
- Subagència és una subentitat d'Agència, és a dir, que si no hi ha Agència no pot existir una Subagència.
- Persona té diverses relacions amb diferents entitats:
  - o La primera relació és la de multiplicitat 1:1 amb Llicència; una persona només pot tenir una llicència.
  - o La segona és 1:N amb Ciutat; una persona només viu a una ciutat, mentre que a una ciutat hi viuen N persones.
  - o La tercera és Sancio, una relació de 5 entitats (Vehicle, Accident, Persona, Subagència, Infracció). Utilitzem aquesta relació perquè en una sanció intervenen totes aquestes entitats.

- Propietari es una herència de Persona, es a dir, Propietari i Persona tenen els mateixos atributs, i de Propietari també ens guardem el nom i els cognoms. Té una relació 1:N amb Vehicle, perquè una persona pot tindre un o més vehicles, però un vehicle només un propietari.
- Accident forma part de la relació de 5 entitats, però en el seu cas es opcional perquè pot ser que la sanció tingui un accident o no.
- Infraccio forma una relació N:M amb Estat, perquè una infracció es pot cometre en diferents estats i tots els estats poden tindre alguna infracció.
- Llicencia manté dos relacions amb dos Entitats.
  - Una és la relació 1:N amb Estat: en un estat hi ha moltes llicencies però una llicencia no pot ser de més d'un estat.
  - L'altra relació 1:1 és la que hem dit anteriorment amb Persona.
- Ciutat es una subentitat d'Estat, és a dir, que un estat conté moltes ciutats, algunes de les quals poden tenir el mateix nom (fet que considerem aquest tipus de relació).

## 2. Model Lògic (diagrama Relacional)



En aquest model podem veure les taules que es formen degut a les relacions entre les entitats, i els atributs PK de algunes entitats que passen a ser FK de altres.

Observations:

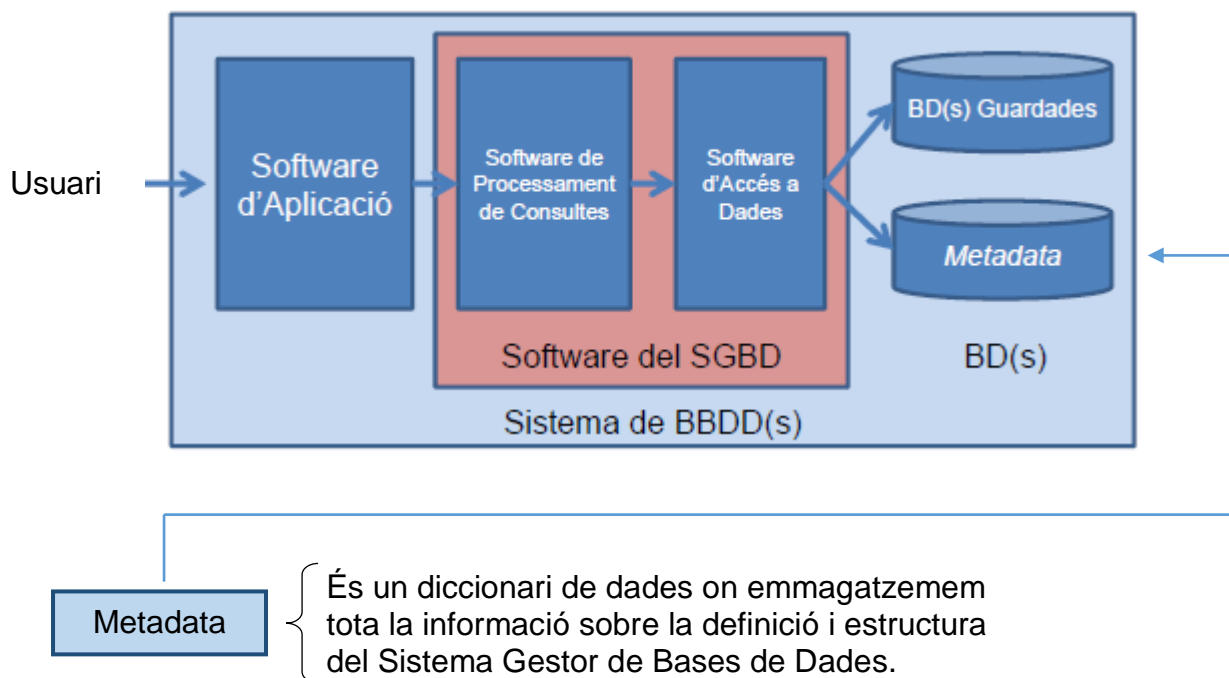
- La multiplicitat entre Estat i Infraccio la considerem N:M, fet que resulta en una taula de relació intermitja anomenada “Legisla”. Aquesta també simbolitza el fet que, com es comentava a l’enunciat, als Estats Units cada territori té la seva pròpia legislació (afectant doncs a la taula Infraccio).

### 3. Estudi de mercat dels SGBD

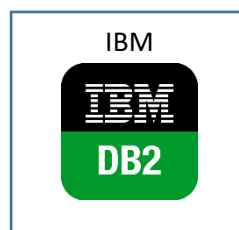
Abans de fer l'estudi de mercat dels SGBD, definim que és un **SGBD**.

Un **SGBD** és un sistema de *Software* de propòsit general que ajuda als usuaris en el procés de creació i manteniment d'una o més bases de dades i que proporciona eines per aquesta tasca.

Arquitectura General:



Un cop explicat que es un SGBD, tenim diferents SGBD en el mercat. Els organitzarem per empreses propietàries dels SGBD:



Aquí tenim una taula on comparem els diferents SGBD per veure les diferències, avantatges i desavantatges que tenen entre elles.

	Oracle	MySQL	DB2	SQLServer	PostgreSQL
<b>Llicència</b>	Propietària	Propietària Gratuïta	Propietària	Propietària	Gratuïta
<b>Seguretat</b>	Alta	Mitjana-Baixa	Alta	Alta	Mitjana-Baixa
<b>Escalabilitat</b>	Molt Alta	Mitjana-Baixa	Molt Alta	Alta-Mitjana	Baixa
<b>Rendiment</b>	Molt Alt	Baix	Alt	Alt	Baix
<b>Sistemes operatius</b>	Windows MAC OS X Linus & Unix	Windows MAC OS X Linus & Unix Etc...	Windows MAC OS X Linus Unix	Windows	Windows MAC OS X Linus & Unix BSD
<b>Interacció</b>	API GUI SQL	SQL	GUI SQL	GUI SQL	GUI SQL
<b>Prestacions de gestió d'índexs</b>	Elevades	Elevada-Mitjana	Mitjanes	Mitjanes	Molt Altes

Després de observar la taula podem arribar a una conclusió. Oracle i DB2 són dos dels líders del mercat dels SGBD (i que potser aniria millor a la policia fer servir per aquests encàrrecs):

## ORACLE

- Conjunt més complet de característiques/funcionalitats per BBDDs.
- Recomanat per arquitectures escalables i que precisin alt rendiment.
- Permet gestionar *terabytes* de informació i milers d'usuaris.
- L'únic problema és el cost de la llicència.
- Oracle Corp. domina el mercat de les grans empreses, i el seu propòsit actual és abordar el segment de les petites/mitjanes empreses.

## DB2

- Producte molt sòlid i en un molt bon moment (MAC, Linux, Unix, Windows).
- Recomanat per desenvolupament d'aplicacions i si es precisa un SGBD amb una elevada robustesa i rendiment.
- És un SGBD d'alta velocitat i el millor processant transaccions.
- Elevat rendiment (*availability*) i sorprenents característiques, només superades pel paquet de funcionalitats d'Oracle.

### 3.1 Han fet bé els policies en sol·licitar l'estructuració de les seves dades en el programari emprat?

Estructurar les dades que tenen en un sistema de bases de dades facilita la gestió i estructuració alhora de buscar, modificar o cercar informació sobre alguna recerca desitjada. A més si volem canviar el format de com volem organitzar o guardar les dades es molt més eficaç i ràpid si esta estructurat en un SGBD. Pel que fa a la pregunta si es bo estructurar les dades en PostgreSQL, no sabem del pressupost que tenen els policies, per tant, l'utilitzem ja que es gratuït i es pot utilitzar en qualsevol sistema operatiu. Nosaltres recomanariem fer servir un altre pel que fa referencia a la seguretat (un punt important, a més, tenint en compte la informació que estem emmagatzemant), ja que no és molt bona i hi ha de millors.

### 3.2. Com s'han solucionat les incoherències de les dades?

Gracies a un PDF que ens proporcionaven, on ens donava informació i ens explicava que significa cada atribut, hem pogut muntar un model on estructuràvem la informació. Això si, algunes de les relacions entre les entitats hem tingut que informar-nos de com funcionava com per exemple els estats i subestats.

També hem tingut alguns problemes a l'hora de introduir informació referent a accident, ja que establir que només s'inserís un atribut definit, en comptes de tots, si es complia una condició no va ser gaire productiu, però vam descobrir una combinació per garantir la informació.

Per últim, respecte als excels, vam haver de reemplaçar totes les cel·les sense informació (nul·les) per 0, fet que no va complicar l'execució del codi al ser tractats com a cadenes de caràcters en comptes de nombres. La majoria d'índexs van haver de ser de tipus *bigint*, degut a la longitud de les xifres en qüestió. I algunes lletres no permeses per la codificació de lectura d'aquests (en especial la 'ç'), van haver de ser igualment canviades per 'c'. A més vam haver de separar el nom dels propietaris, ficant dues columnes més (*surname1* i *surname2*), des de l'excels, amb l'assistent del programa, com s'explica aquí:

<https://support.office.com/es-es/article/Dividir-texto-en-diferentes-columnas-con-el-Asistente-para-convertir-texto-en-columnas-30b14928-5550-41f5-97ca-7a3e9c363ed7>



## 4. Model Físic

**Diferència entre *charset* i *collate*:** *Charset* fa referència a com guarda internament una dada PostgreSQL, mentre que *Collate* es una manera de dir-li com ha de comparar, llegir, ordenar... el text.

- **CHARSET:** utf-8, perquè permet caràcters com ñ
- **COLLATE:** utf8-general-ci (perquè considera sinònimes algunes lletres com la ñ i la n)

**¿Perquè no hem escollit un altre, com Charset ASCII o Latin?** Perquè ASCII no permet ingressar altra dada que no aparegui en la tabla Ascii, així que tant els accents com algunes lletres més complicades o no pròpies del llatí podrien donar problemes.

### ARA EL MODEL FÍSIC:

Aquest és el conjunt de sentències empleat en l'execució del model físic. Els comentaris necessaris es troben al codi mateix:

---

```
-- EL NOSTRE DIAGRAMA:
-- Primer les entitats que no depenen de ningú:
drop table if exists agencia cascade;
create table agencia(
    nom          varchar(255),

    primary key(nom)
);
-- select * from agencia;

drop table if exists subagencia cascade;
create table subagencia(
    agencia      varchar(255),
    subagencia   varchar(255),

    foreign key(agencia) references agencia(nom),
    primary key(agencia, subagencia)
);
-- select * from subagencia;

-- ESQUERRA DEL DIAGRAMA:
drop table if exists estat cascade;
create table estat(
    state        varchar(255),
    agencia      varchar(255),

    foreign key(agencia) references agencia(nom),
    primary key (state)
);
-- select * from estat;

drop table if exists ciutat cascade;
create table ciutat(
    city         varchar(255),
    state        varchar(255),

    foreign key(state) references estat(state),
    primary key(city, state)
);
-- select * from ciutat;

drop table if exists persona cascade;
```

```

create table persona(
    id_persona    bigint,
    race          varchar(255),
    gender        char,
    state         varchar(255),
    city          varchar(255),

    foreign key(state, city) references ciutat(state, city),
    -- al definir una pk composta (tabla ciutat), la fk deu ser també composta
    primary key(id_persona)
);
-- select * from persona;

drop table if exists llicencia cascade;
create table llicencia(
    id_llicencia  serial,
    dl_state      varchar(255),
    driving_license varchar(255),
    state         varchar(255),
    id_persona    bigint,

    foreign key(state) references estat(state),
    foreign key(id_persona) references persona(id_persona),
    primary key(id_llicencia)
);
-- select * from llicencia;

drop table if exists propietari cascade;
create table propietari(
    id_propietaribigint,
    nom           varchar(255),
    cognom1       varchar(255),
    cognom2       varchar(255),

    foreign key(id_propietari) references persona(id_persona),
    primary key(id_propietari)
);
-- select * from propietari;

drop table if exists vehicle cascade;
create table vehicle(
    license        varchar(255), -- matricula
    type_vehicle  varchar(255),
    year          int,
    make          varchar(255),
    color         varchar(255),
    model         varchar(255),
    id_propietaribigint,
    estat         varchar(255),

    foreign key(id_propietari) references propietari(id_propietari),
    primary key(license)
);
-- select * from vehicle;

-- DRETA DEL DIAGRAMA
drop table if exists accident cascade;
create table accident(
    id_accident    bigint,
    accident       boolean,
    belts          boolean,
    person_injury  boolean,
    property_dmg   boolean,
    fatal          boolean,
    hazmat         boolean,
    alcohol        boolean,
    work_zone      boolean,

```

```

        primary key(id_accident)
);    -- select * from accident;

-- CENTRE DEL DIAGRAMA:
drop table if exists infraccio cascade;
create table infraccio(
    id_infraccio bigint,
    violation_type    varchar(255),
    arrest_type    varchar(255),
    charge          varchar(255),

    primary key(id_infraccio)
);    -- select * from infraccio;

drop table if exists legisla cascade;
create table legisla(
    state            varchar(255),
    infraccio        bigint,

    foreign key(state) references estat(state),
    foreign key(infraccio) references infraccio(id_infraccio),
    primary key(state, infraccio)
);    -- select * from legisla;

drop table if exists sanciona cascade;
create table sanciona(
    id_infraccio bigint,
    id_persona    bigint,
    agencia        varchar(255),
    subagencia    varchar(255),
    license        varchar(255),
    id_accident    bigint,    -- perque la necessitem per fer la fk composta
    (de la pk composta de la taula Subagencia)
    date_stop     varchar(255), -- no pot ser de tipus DATE, perque és DD-MM-
YYYY, y DATE: YYYY-MM-DD
    time_stop     time,
    description    varchar(255),
    location       varchar(255),
    latitude       varchar(255),
    longitude      varchar(255),

    foreign key(id_infraccio) references infraccio(id_infraccio),
    foreign key(id_persona) references persona(id_persona),
    foreign key(agencia, subagencia) references subagencia(agencia,
subagencia),
    foreign key(license) references vehicle(license),
    foreign key(id_accident) references accident(id_accident),
    primary key(id_infraccio, id_persona, agencia, subagencia, license,
id_accident)
);    -- select * from sanciona;

-- SOBRE LA TAULA IMPORT_TRAFFIC_VIOLATIONS:
-- També agreguem l'atribut id_itv a la taula d'importacio de traffic_violations,
que ens serà útil per l'entitat Accident (id_accident) i Infraccio:
alter table import_traffic_violations add id_itv serial; -- (30 segons)

/*
Ara inserim la informació a les taules del nostre diagrama des de les de
importacio,
seguint el mateix ordre que el de la seva creació, per evitar problemes amb les
fks:

```

```

*/

-- PRIMER LES QUE NO DEPENEN DE NINGU:
-- Agencia: (1 secs)
insert into agencia(nom)
select distinct itv.agency
from import_traffic_violations as itv;
-- select * from agencia;

-- Subagencia: (5 segons)
insert into subagencia(agencia, subagencia)
select distinct a.nom, itv.subagency
from agencia as a, import_traffic_violations as itv
where a.nom = itv.agency;
-- select * from subagencia order by agencia;

-- ESQUERRA DEL DIAGRAMA:
-- farem servir les variables driver_state i driver_city per a l'index d'estats
i ciutats.
-- Estat: (3 segons, 67 estats)
insert into estat
select distinct itv.driver_state, a.nom
from import_traffic_violations as itv, agencia as a
where itv.agency like a.nom;
-- select * from estat;

-- Ciutat: (15 segons, 8K ciutats)
insert into ciutat(city, state)
select distinct itv.driver_city, estat.state
from estat, import_traffic_violations as itv
where estat.state like itv.driver_state;
-- select * from ciutat;

-- Persona: (ho farem per passos, porque si no la PK de la taula peta amb els
distincts) (50 segons, 510K persones)
-- 1) INSERT per a les ID_PERSONAS (5 segundos)
insert into persona(id_persona)
select distinct itv.id_persona
from import_traffic_violations as itv
order by id_persona;
-- 2) UPDATES de la TAULA ITV: (17 segons)
update persona
set
-- para los generos y razas
race = itv.race,
gender = itv.gender,
-- para los estados y ciudades
state = itv.driver_state,
city = itv.driver_city
from import_traffic_violations as itv
where itv.id_persona = persona.id_persona
and persona.id_persona = itv.id_persona;
-- select * from persona; -- (20 segons): 510K persones.
-- 3) UPDATES de la TAULA IP: (2 segons)
insert into persona(id_persona)
select ip.id_propietari
from import_propietaris as ip
where ip.id_propietari not in(select p.id_persona
from persona as p
where p.id_persona = ip.id_propietari);

-- Llicencia: (1 minut)
-- 1) Primer les ID_PERSONAS, ja que si no surtiras repetides algunes llicències
(517K en vez de 498K), i només hi ha 1 llicència per persona.

```

```

-- (1 minut, 498K llicencies)
insert into llicencia(id_persona)
select distinct      p.id_persona
from import_traffic_violations as itv, persona as p
where itv.id_persona = p.id_persona
and not(itv.driving_license like '%XX%'); -- perquè XX vol dir que no tenen
llicencia.
-- select * from llicencia;
-- 2) Després, quan no tenim llicències repetides i les ID_PERSONAS siguin
introduïdes, agreguem l'altra informació. (30 segons)
update llicencia
set      driving_license = itv.driving_license,
        dl_state = itv.dl_state,
        state = itv.driver_state -- utilitzem aquesta per estalviar-nos una
taula.
from import_traffic_violations as itv
where llicencia.id_persona = itv.id_persona;
-- select * from llicencia; -- (30 segons, 498K llicencies)

-- Propietari: (15 segons)
insert into propietari(id_propietari)
select distinct p.id_persona
from persona as p;

update propietari
set      nom = ip.name,
        cognom1 = ip.surname1,
        cognom2 = ip.surname2
from import_propietaris as ip
where propietari.id_propietari = ip.id_propietari;
-- select * from propietari order by id_propietari;

-- Vehicle: (10 segons)
-- 1) Ficar les llicències de els vehicles a la taula (2 segons, 125K vehicles)
insert into vehicle(license)
select distinct      itv.license
from import_traffic_violations as itv;
-- 2) Ficar les altres dades: (7 segons, 125K vehicles (=propietaris))
update vehicle
set      type_vehicle = itv.vehicle_type,
        year = itv.year,
        make = itv.make,
        color = itv.color,
        model = itv.model,
        estat = itv.state,
        id_propietari = ip.id_propietari
from import_traffic_violations as itv, import_propietaris as ip, propietari as
p
where itv.id_persona = p.id_propietari
and ip.id_propietari = p.id_propietari
and vehicle.license = itv.license;
-- IMPORTANT: hi ha dos propietaris que no es veuen perquè no apareixen a la
"Import_Traffic_Violations" però sí a la "Import_Propietaris".
-- select * from vehicle; -- (12 segons)

-- DRETA DEL DIAGRAMA:
-- Accident: (1 sec, 28K accidents)
-- 1) Primer si ha hagut accident
insert into accident(id_accident, accident, belts, person_injury, property_dmg,
fatal, hazmat, alcohol, work_zone)
select id_itv, accident, belts, personal_injury, property_dmg, fatal, hazmat,
alcohol, work_zone
from import_traffic_violations

```

```

where contributed_accident = true;

-- 2) Després, si no hi ha hagut accident, ja que llavors tots els altres
atributs son null.
insert into accident(id_accident)
select itv.id_itv
from import_traffic_violations as itv
where itv.id_itv not in(select a.id_accident
                        from accident as a
                        where itv.id_itv = a.id_accident);
-- select * from accident order by id_accident;

-- CENTRE DEL DIAGRAMA:
-- Infraccio: (15 segons)
insert into infraccio
select id_itv, violation_type, arrest_type, charge
from import_traffic_violations;
-- select * from infraccio; -- (+1 minut)

-- Legisla: (2 minut)
insert into legisla(state, infraccio)
select e.state, i.id_infraccio
from estat as e, infraccio as i, import_traffic_violations as itv
where e.state like itv.driver_state
and itv.id_itv = i.id_infraccio;
-- select * from legisla; -- (20 segons)

-- Sanciona 1: (2 minuts, 1204K sancions)
-- tot i que tot això podríem treure-ho de ITV, farem servir les altres taules
per tal de comprovar que la informació introduïda és correcta.
insert into sanciona(id_infraccio, id_persona, agencia, subagencia, license,
id_accident,
date_stop, time_stop, description, location, latitude, longitude)
select i.id_infraccio, itv.id_persona, ag.nom, sag.subagencia, v.license,
a.id_accident, -- perque les pks no poden valer null, i es el seu
valor si contributed_accident = false.
itv.date_stop, itv.time_stop, itv.description, itv.location,
itv.latitude, itv.longitude -- atributs normals
from infraccio as i,
agencia as ag,
accident as a,
subagencia as sag,
vehicle as v,
import_traffic_violations as itv
where i.id_infraccio = itv.id_itv
and itv.agency = ag.nom
and sag.agencia = ag.nom
and sag.subagencia = itv.subagency
and a.id_accident = itv.id_itv
and v.license = itv.license;

-- select * from sanciona order by id_accident; -- 3 minuts.

```

---

Hauria de trigar entre 6 i 8 minuts en fer l'execució completa del codi.

## 5. Test del procés d'importació.

Aquest és el codi d'importació de tots els excels cap a les taules d'importació "import\_propietaris" e "import\_traffic\_violations":

---

```
-- TAULES D'IMPORTACIO:

drop table if exists import_propietaris;
create table import_propietaris(
    id_propietaribigint,
    license          varchar(255),
    name             varchar(255),
    surname1         varchar(255),
    surname2         varchar(255)
); -- select * from import_propietaris (triga uns 7 segons en carregar la
taula sencera);

drop table if exists import_traffic_violations;
create table import_traffic_violations(
    date_stop        varchar(255),
    time_stop        time,
    agency            varchar(255),
    subagency         varchar(255),
    description       varchar(255),
    location          varchar(255),
    latitude          varchar(255),
    longitude         varchar(255),
    accident          boolean,
    belts            boolean,
    personal_injury   boolean,
    property_dmg      boolean,
    fatal             boolean,
    hazmat            boolean,
    alcohol           boolean,
    work_zone         boolean,
    state             varchar(255), -- aquesta la farem servir per a la classe
vehicle
    license           varchar(255),
    vehicle_type      varchar(255),
    driving_license   char(255),
    year             int,
    make             varchar(255),
    model            varchar(255),
    color            varchar(255),
    violation_type    varchar(255),
    charge            varchar(255),
    contributed_accident boolean,
    race             varchar(255),
    gender           char,
    driver_city       varchar(255),
    driver_state      varchar(255), -- aquesta es per a la classe estat
    dl_state          varchar(255), -- aquesta es per a la classe llicencia
    arrest_type       varchar(255),
    id_persona        bigint
); -- select * from import_traffic_violations; -- (triga uns 9 minuts aprox
en mostrar la taula sencera)

-- Fiquem la informació dels propietaris del csv: (aprox 4 segons en copiar
les dades)
```

```

copy import_propietaris from
'G:\Universidad\3o\BBDD\p1\Arxius\propietaris.csv' delimiter ';';

-- Fiquem la info de les violacions de trafic dels 23 csv: (aprox 35 segons en
copiar les dades)
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_1.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_2.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_3.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_4.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_5.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_6.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_7.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_8.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_9.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_10.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_11.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_12.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_13.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_14.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_15.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_16.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_17.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_18.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_19.csv' delimiter ';' csv
header;

```



```

copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_20.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_21.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_22.csv' delimiter ';' csv
header;
copy import_traffic_violations from
'G:\Universidad\3o\BBDD\p1\Arxius\traffic_violations_23.csv' delimiter ';' csv
header;

-- trigarà uns 30 segons en carregar totes les dades d'importació.

```

---

Pas a pas els resultats serien els següents (*descomentant* el select que hi ha rere cada creació de la taula):

### 1) Al crear la taula **Import\_Propietaris**:

Query - importacio\_p1 en postgres@localhost:5432 - [G:\Universidad\3o\BBDD\p1\Entrega\archivos ...]

Archivo Editar Consulta Favoritos Macros Vista Ayuda

importacio\_p1 en postgres@localhost:5432

Editor SQL &Constructor Gráfico de Consultas

Consultas anteriores Borrarr Eliminar Todos

```

-- Database: "BD18_p1"
/*
-- DROP DATABASE "BD18_p1";

CREATE DATABASE "BD18_p1"
  WITH OWNER = postgres
       ENCODING = 'UTF8'
       TABLESPACE = pg_default
       LC_COLLATE = 'Spanish_Spain.1252'
       LC_CTYPE = 'Spanish_Spain.1252'
       CONNECTION LIMIT = -1;
*/

-- TAULES D'IMPORTACIO:

drop table if exists import_propietaris;
create table import_propietaris(
  id_propietari    bigint,
  license          varchar(255),
  name             varchar(255),
  surname1         varchar(255),
  surname2         varchar(255)
);
select * from import_propietaris;

drop table if exists import_traffic_violat
create table import_traffic_violations(

```

Panel de Salida

Salida de datos Comentar Mensajes Historial

id_propietari	license	name	surname1	surname2
bigint	character	character	character v	character v

OK. DOS Lín 16 Col 1 Car 316 233 caracteres 0 filas. 186 msec

## 2) Al crear la taula **Import\_Traffic\_Violations**:

The screenshot shows the SQL Editor interface with the following SQL code:

```
drop table if exists import_traffic_violations;
create table import_traffic_violations(
    date_stop varchar(255),
    time_stop time,
    agency varchar(255),
    subagency varchar(255),
    description varchar(255),
    location varchar(255),
    latitude varchar(255),
    longitude varchar(255),
    accident boolean,
    belts boolean,
    personal_injury boolean,
    property_dmg boolean,
    fatal boolean,
    hazmat boolean,
    alcohol boolean,
    work_zone boolean,
    state varchar(255),
    license varchar(255),
    vehicle_type varchar(255),
    driving_license char(255),
    year int,
    make varchar(255),
    model varchar(255),
    color varchar(255),
    violation_type varchar(255),
    charge varchar(255),
    contributed_accident boolean,
    race varchar(255),
    gender char,
    driver_city varchar(255),
    driver_state varchar(255),
    dl_state varchar(255),
    arrest_type varchar(255),
    id_persona bigint
);
select * from import_traffic_violations;

-- Fiquem la informació dels propietaris
copy import_propietaris from 'G:\Universi';
-- Fiquem la info de les violacions de t
```

The right panel shows the table structure for **import\_traffic\_violations** with columns: date\_stop, time\_stop, agency, subagency, description, location, latitude, longitude, accident, belts, personal\_injury, property\_dmg, fatal, hazmat, alcohol, work\_zone, state, license, vehicle\_type, driving\_license, year, make, model, color, violation\_type, charge, contributed\_accident, race, gender, driver\_city, driver\_state, dl\_state, arrest\_type, id\_persona.

## 3) Al ficar la informació a **Import\_Propietaris**:

The screenshot shows the SQL Editor interface with the following SQL code:

```
license varchar(255),
vehicle_type varchar(255),
driving_license char(255),
year int,
make varchar(255),
model varchar(255),
color varchar(255),
violation_type varchar(255),
charge varchar(255),
contributed_accident boolean,
race varchar(255),
gender char,
driver_city varchar(255),
driver_state varchar(255),
dl_state varchar(255),
arrest_type varchar(255),
id_persona bigint
);
select * from import_traffic_violations; -- (tarda 9 minutos aprox en cargar la tabla entera)

-- Fiquem la informació dels propietaris del csv: (4 segons en copiar les dades)
copy import_propietaris from 'G:\Universiad\3o\BDD\p1\Arxius\propietaris.csv' delimiter ';';
-- Fiquem la info de les violacions de trafic dels 23 csv: (35 segons en copiar les dades)
copy import_traffic_violations from 'G:\Universiad\3o\BDD\p1\Arxius\traffic_violations_1.csv';
copy import_traffic_violations from 'G:\Universiad\3o\BDD\p1\Arxius\traffic_violations_2.csv';
copy import_traffic_violations from 'G:\Universiad\3o\BDD\p1\Arxius\traffic_violations_3.csv';
```

The right panel shows the message: "Query returned successfully: 125756 rows affected, 352 msec execution time."

[si fem un select:]

Panel de Salida					
Salida de datos					
Comentar					
Mensajes					
Historial					
	id_propietari bigint	license character	name character	surname1 character v	surname2 character v
1	30774427	AB-1158	Si	Bucknell	
2	14389340	AK-1295	Wilson	Harly	
3	32234465	AL-1142	Aleci	Mattaser	
4	62096627	AL-1159	Micke	Crocker	
5	5424320	AL-1167	Conwa	Hathaway	
6	9467090	AL-1094	Lynne	McAlinder	
7	3206578	AL-1097	Moses	Blesing	
8	29961594	AL-1214	Gaven	Pennetta	
9	33692977	AL-1143	Jo	ann	Hartnup
10	13142275	AL-1229	Zache	Bernath	
11	59619475	AL-1156	Karlo	Tite	
12	1736643	AL-1266	Linnie	Carradice	
13	27981858	AL-1190	Norbe	O'Leahy	
14	19339120	AR-9524	Lynet	O'Scollee	
15	23485048	AR-1262	Sebas	Dyott	
16	34895387	AR-1221	Yvor	Sigg	
17	17040176	AZ-9866	Jacki	Edeson	
18	8010537	AZ-1177	Denne	Gynn	
19	7765490	AZ-1064	Nanet	Gillford	
20	16067466	AZ-1105	Allys	Hackeltor	
21	32824381	AZ-1229	Gilbe	Lahy	
22	13181734	AZ-1219	Grego	Petticrev	
23	11020790	AZ-1229	Drusy	Hardwicke	
24	5202238	AZ-1162	Ruthe	Barfield	
25	10614203	AZ-1234	Sheff	Sesons	
26	58384458	AZ-1202	Cheri	Kenningha	
27	17297635	AZ-1210	Lissi	Bletsoe	
28	16736589	AZ-1278	Addy	Kingsnode	
29	5633673	AZ-1317	Marne	Chesterfi	
30	22277527	AZ-1339	Skyla	Barcroft	
31	27620501	AZ-1289	Grace	Bacher	
32	2507319	CA-1060	Inger	Jaher	
33	2913068	CA-1058	Pauli	Greystock	
34	1461829	CA-1077	Iorgo	Gammack	
35	3444093	CA-1026	Tipha	Standbroc	
36	14897772	CA-1148	Garre	Christol	
37	7749291	CA-1156	Gunth	Rosenster	

Podem veure les diferents columnes de què vam parlar abans:

- name
- surname1
- surname2.

4) Al **ficar** la informació a **Import\_Traffic\_Violations**:

Query - importacio\_p1 en postgres@localhost:5432 - [G:\Universidad\3o\BDD\p1\Entrega\archivos nuevo\importacio.sql] \*

Archivo Editor Consulta Favoritos Macros Vista Ayuda

Importacio\_p1 en postgres@localhost:5432

Editor SQL &Constructor Gráfico de Consultas

Consultas anteriores

```

copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_1.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_2.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_3.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_4.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_5.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_6.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_7.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_8.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_9.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_10.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_11.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_12.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_13.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_14.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_15.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_16.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_17.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_18.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_19.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_20.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_21.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_22.csv' delimiter ';' csv header;
copy import_traffic_violations from 'G:\Universidad\3o\BDD\p1\Arxius\traffic_violations_23.csv' delimiter ';' csv header;

-- triga 30 segons en carregar totes les dades d'importació.

```

Panel de Salida

Salida de datos Comentar Mensajes Historial

Query returned successfully: 105598 rows affected, 01:07 minutes execution time.

105598 filas afectadas. DOS Lin 68 Col 1 Car: 1859 2935 caracteres 01:07 minutes

[si fem un select:]

Panel de Salida																						
Salida de datos Comentar Mensajes Historial																						
	date_stop	time_stop	agency	subagency	description	location	latitude	longitude	accident	belts	personal_in	property_d	fatal	hazmat	alcohol	work_zone	state	license	vehicle_type	driving_lice	year	
	character	time without	character	character	character var	character	character	character	boolean	boolean	boolean	boolean	boolean	boolean	boolean	boolean	charac	character	character var	character(2	integ	
1	15/12/20	15:08:00	MCP	4th dist	PERSON DF	GEORGIA	3.900	-7.710	0	f	f	f	f	f	f	f	AB	AB-118	02 - Autom	B	2016	
2	26/09/20	15:08:00	MCP	4th dist	PERSON DF	GEORGIA	3.900	-7.710	0	f	f	f	f	f	f	f	AB	AB-118	02 - Autom	B	2016	
3	21/02/20	15:08:00	MCP	4th dist	PERSON DF	GEORGIA	3.900	-7.710	0	f	f	f	f	f	f	f	AB	AB-115	02 - Autom	B	2013	
4	05/06/20	15:08:00	MCP	4th dist	PERSON DF	GEORGIA	3.900	-7.710	0	f	f	f	f	f	f	f	AB	AB-115	02 - Autom	B	2013	
5	07/03/20	13:15:00	MCP	2nd dist	DRIVER FR	WARREN	3.901	-7.704	8	f	f	f	f	f	f	f	AK	AK-129	02 - Autom	B	2009	
6	03/05/20	23:43:00	MCP	4th dist	DRIVER FR	UNIVERS	3.900	-7.705	4	f	f	f	f	f	f	f	AL	AL-114	02 - Autom	B	2012	
7	11/04/20	23:43:00	MCP	4th dist	DRIVING V	UNIVERS	3.900	-7.705	4	f	f	f	f	f	f	f	AL	AL-114	02 - Autom	B	2012	
8	11/09/20	23:43:00	MCP	4th dist	DRIVER FR	UNIVERS	3.900	-7.705	4	f	f	f	f	f	f	f	AL	AL-114	02 - Autom	B	2012	
9	06/09/20	00:02:00	MCP	3rd dist	FAILURE T	SB NEW	3.903	-7.700	0	f	f	f	f	f	f	f	AL	AL-115	05 - Light	E	2012	
10	06/03/20	14:07:00	MCP	2nd dist	EXCEEDING	N/B 35	3.899	-7.710	0	f	f	f	f	f	f	f	AL	AL-116	02 - Autom	B	2016	
11	27/08/20	11:11:00	MCP	4th dist	FAILURE C	WIMBLEY	3.910	-7.704	7	f	f	f	f	f	f	f	AL	AL-109	02 - Autom	B	2004	
12	10/11/20	11:11:00	MCP	4th dist	FAILURE T	WIMBLEY	3.910	-7.704	7	f	f	f	f	f	f	f	AL	AL-109	02 - Autom	B	2004	
13	13/01/20	11:11:00	MCP	4th dist	OPERATING	WIMBLEY	3.910	-7.704	7	f	f	f	f	f	f	f	AL	AL-109	02 - Autom	B	2004	
14	18/07/20	20:26:00	MCP	2nd dist	DISPLAYIN	EDGEMO	3.900	-7.710	0	f	f	f	f	f	f	f	AL	AL-109	02 - Autom	B	2004	
15	11/11/20	22:02:00	MCP	3rd dist	DRIVING M	N/B UN	3.900	-7.700	2	f	f	f	f	f	f	f	AL	AL-121	02 - Autom	B	2014	
16	09/08/20	18:49:00	MCP	6th dist	FAILURE Y	N. FREI	3.920	-7.730	0	f	t	f	f	f	f	f	AL	AL-114	02 - Autom	B	2003	
17	08/10/20	00:30:00	MCP	5th dist	DISPLAYIN	CRYSTAL	3.918	-7.730	0	f	f	f	f	f	f	f	AL	AL-122	02 - Autom	B	2011	
18	06/09/20	00:32:00	MCP	2nd dist	FAILURE C	CONNEC	3.899	-7.707	6	f	f	f	f	f	f	f	AL	AL-115	02 - Autom	B	2003	
19	16/05/20	13:54:00	MCP	4th dist	FAILURE T	NEW HAM	3.908	-7.700	0	f	f	f	f	f	f	f	AL	AL-126	02 - Autom	B	2016	
20	12/11/20	22:01:00	MCP	4th dist	PERSON DF	N/B GE	3.910	-7.710	0	f	f	f	f	f	f	f	AL	AL-119	02 - Autom	B	2006	
21	26/04/20	22:01:00	MCP	4th dist	PERSON DF	N/B GE	3.910	-7.710	0	f	f	f	f	f	f	f	AL	AL-119	02 - Autom	B	2006	
22	10/06/20	22:01:00	MCP	4th dist	DRIVER FR	N/B GE	3.910	-7.710	0	f	f	f	f	f	f	f	AL	AL-119	02 - Autom	B	2006	
23	28/03/20	22:01:00	MCP	4th dist	DRIVING M	N/B GE	3.910	-7.710	0	f	f	f	f	f	f	f	AL	AL-119	02 - Autom	B	2006	
24	04/07/20	22:01:00	MCP	4th dist	PERSON DF	N/B GE	3.910	-7.710	0	f	f	f	f	f	f	f	AL	AL-119	02 - Autom	B	2006	
25	05/02/20	02:26:00	MCP	3rd dist	EXCEEDING	PINEY I	3.900	-7.700	0	f	f	f	f	f	f	f	AR	AR-952	02 - Autom	B	1989	
26	11/04/20	02:26:00	MCP	3rd dist	DRIVINGTC	PINEY I	3.900	-7.700	0	f	f	f	f	f	f	f	AR	AR-952	02 - Autom	B	1989	
27	09/07/20	03:57:00	MCP	4th dist	DRIVING V	FOREST	3.905	-7.700	0	f	f	t	f	f	f	f	AR	AR-126	02 - Autom	B	2007	
28	05/12/20	03:57:00	MCP	4th dist	DRIVING V	FOREST	3.905	-7.700	0	f	f	t	f	f	f	f	AR	AR-126	02 - Autom	B	2007	
29	19/10/20	03:57:00	MCP	4th dist	DRIVING V	FOREST	3.905	-7.700	0	f	f	t	f	f	f	f	AR	AR-126	02 - Autom	B	2007	
30	11/08/20	03:57:00	MCP	4th dist	NEGLIGENT	FOREST	3.905	-7.700	0	f	f	t	f	f	f	f	AR	AR-126	02 - Autom	B	2007	
31	14/01/20	03:57:00	MCP	4th dist	RECKLESS	FOREST	3.905	-7.700	0	f	f	t	f	f	f	f	AR	AR-126	02 - Autom	B	2007	
32	02/11/20	03:57:00	MCP	4th dist	EXCEEDING	FOREST	3.905	-7.700	0	f	f	t	f	f	f	f	AR	AR-126	02 - Autom	B	2007	
33	09/03/20	03:57:00	MCP	4th dist	DRIVING V	FOREST	3.905	-7.700	0	f	f	t	f	f	f	f	AR	AR-126	02 - Autom	B	2007	
34	28/05/20	03:57:00	MCP	4th dist	FAILURE T	FOREST	3.905	-7.700	0	f	f	t	f	f	f	f	AR	AR-126	02 - Autom	B	2007	
35	18/10/20	03:57:00	MCP	4th dist	FAILURE C	FOREST	3.905	-7.700	0	f	f	t	f	f	f	f	AR	AR-126	02 - Autom	B	2007	
36	23/06/20	03:57:00	MCP	4th dist	FAILURE T	FOREST	3.905	-7.700	0	f	f	t	f	f	f	f	AR	AR-126	02 - Autom	B	2007	

## 6. Conclusions

Per acabar les conclusions que podem treure en fer aquesta pràctica son principalment dues: primerament, hem pogut ampliar els nostres coneixements sobre l'ús de PostgreSQL i el llenguatge en sí mateix, especialment a l'hora de provar d'inserir la informació i després "refrescar" la taula (procés que hem repetit un parell de cops, especialment per evitar problemes de repetició de *primary keys*) i les seves eines. La segona, però no per això menys important, consisteix al disseny de diagrames. Hem hagut de refer tantes vegades els nostres diagrames abans de començar a implementar els models definitius, que fins i tot vam poder millorar en quant a la teoria: el treball no era difícil (potser una mica ambigu en quant a les dades, generalment), però havíem de dedicar-li temps, i ha estat una bona pràctica trobar gairebé tots els elements de teoria o de les evaluacions continues en un mateix exercici (selects, inserts, updates, deletes, import -evidentment-... fins i tot un alter!).

Per últim, també hem après bastant sobre el funcionament de les taules més enllà de Postgre, com seria als excels, tant a l'hora d'agafar i extreure dades, com a llegir els csv's i a fer servir eines d'utilitat, com l'Asistent (quina meravella a l'hora de separar els noms en nom, cognom u i cognom dos).

Ha estat una mica estressant, però una bona pràctica al cap i a la fi.