

Bases de Dades

P1 FASE 2

Aleix Ollé Requena
Ferran Mateu Solans

aleix.olle
ferran.mateu
La Salle curs 19/20

1. MODEL CONCEPTUAL.....	4
2. MODEL RELACIONAL	5
3. MODEL FÍSIC	6
3.1 Descripció de les Taules.	6
-Property.	6
-Host.....	6
-Verification.....	6
-Verifies	6
-Reviewer	6
-ReviewerCommentProperty	7
-Amenity.....	7
-HasAmenity.....	7
-Street	7
-Neighbourhood.....	7
-City	7
-State	7
-Country	8
3.2 Sobre les Taules Intermitges.	8
3.3 Verificació del Model Físic.	8
-Procediment Importació.....	8
-State	9
-Country	9
-City	9
-Neighbourhood.....	10
-Street	10
-Property	11
-Amenity.....	11
-HasAmenity.....	12

-Host.....	13
-Reviewer	14
-ReviewerCommentProperty	14
-Verification.....	15
-Verifies	16

4. PROCÉS D'IMPORTACIÓ 18

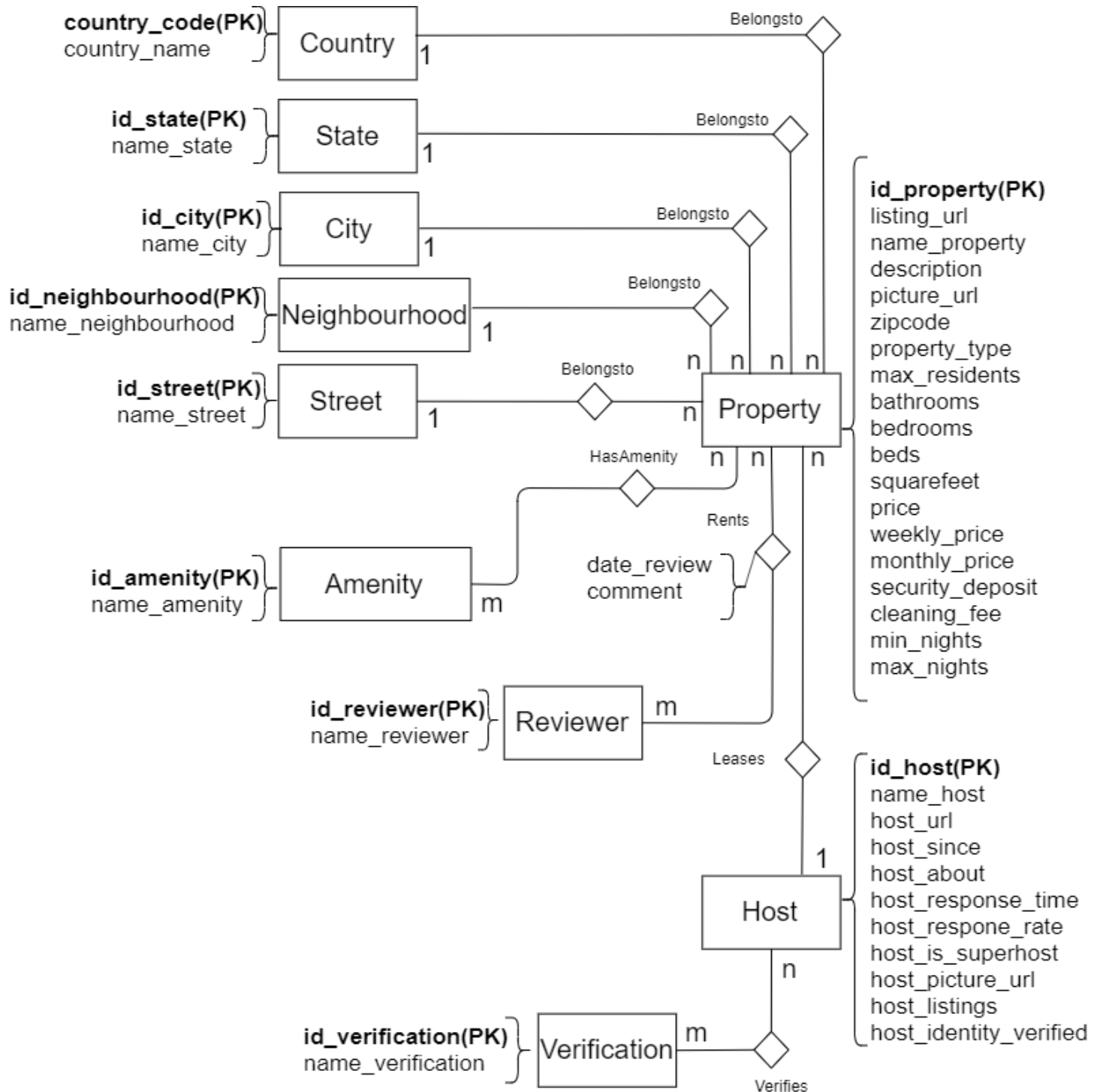
4.1Consideracions durant la importació.....	18
4.2 Verificació de la importació	19
-ImportApartment.....	19
-Property	20
-City	20
-Neighbourhood.....	20
-Street	21
-Host.....	21
-ImportHosts	21

5.FASE 2 22

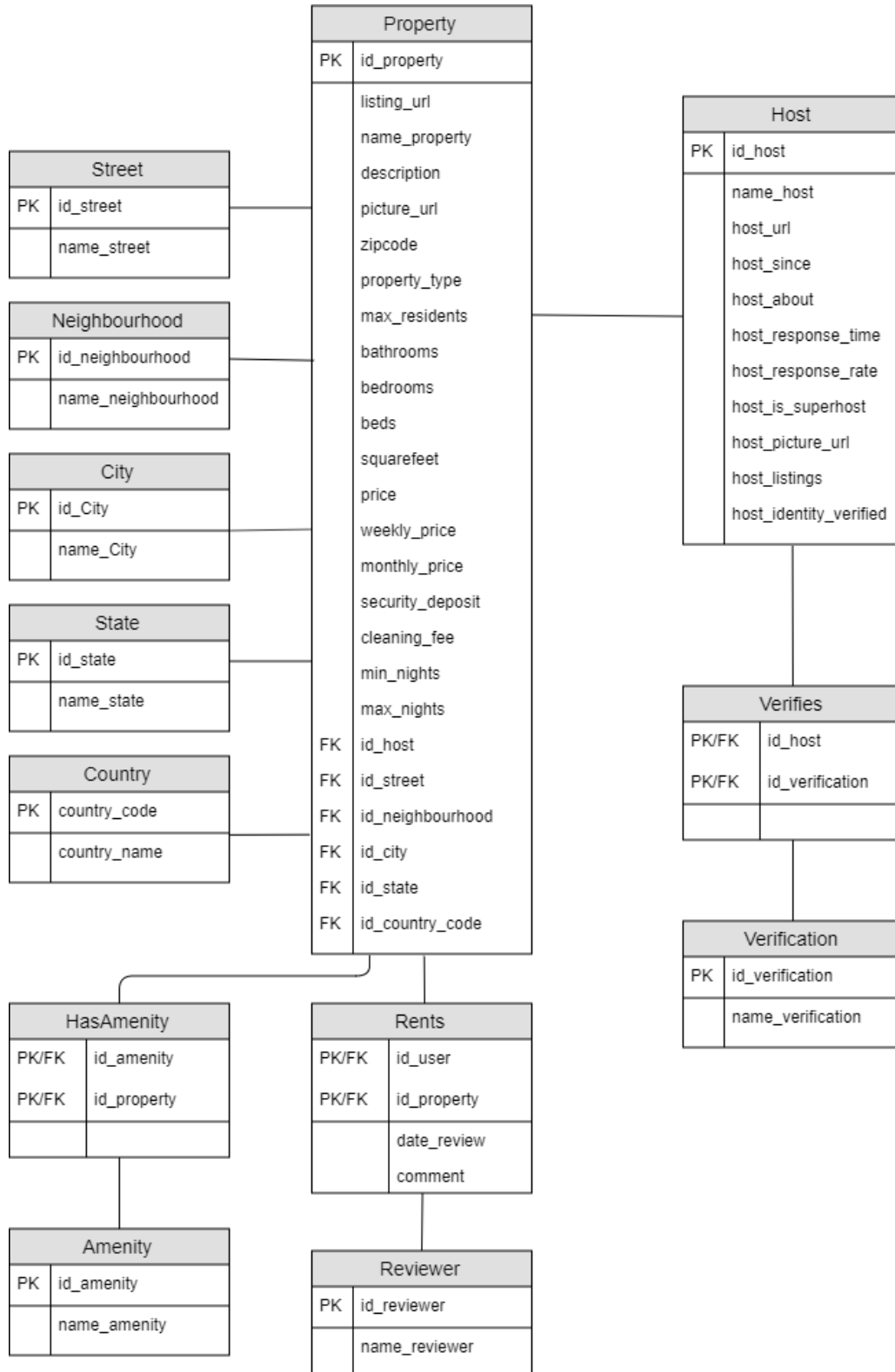
5.1 Query 1.....	22
Comprovació	22
5.2 Query 2.....	22
Comprovació	23
5.3 Query 3.....	23
Comprovació	23
5.4 Query 4.....	24
Comprovació	24
5.5 Query 5.....	24
Comprovació	24
5.6 Query 6.....	25
Comprovació	25

5.7 Query 7.....	26
Comprovació	26
5.8 Query 8.....	27
Comprovació	28
5.9 Query 9.....	29
Comprovació	30
5.10 Query 10.....	31
Comprovació	32
 6.CONCLUSIONS	 32

1. MODEL CONCEPTUAL



2. MODEL RELACIONAL



3. MODEL FÍSIC

3.1 Descripció de les Taules.

A continuació procedirem a fer un llistat de les taules necessàries per a implementar el nostre model físic juntament amb una petita explicació del contingut i necessitat de cada una d'elles.

-Property.

Aquesta taula conté tota la informació respectiva a cada un dels tipus d'habitatges. Per respectar les diverses normes formals s'ha separat la informació referent a ubicació física i aquelles columnes no atòmiques en taules diferents.

Tal com se'ns ha instat al enunciat, hem descartat el identificador que teníem i n'hem creat un de nou, per aquesta raó el camp "id_property" es de tipus "SERIAL"

Es important comentar que el tipus money ens dona un error ja que no accepta el caràcter '\$' per tant ens hem vist obligats a utilitzar "VARCHAR"

-Host

Aquí hi trobarem tota la informació relativa al host, es a dir a l'amo de cada una de les propietats.

Una vegada mes, amb l'objectiu de respectar les formes normals, hem separat el camp no atòmic "verification" creant així una nova taula.

-Verification

Tal com acabem de comentar, era necessari crear una nova taula per emmagatzemar els diferents mètodes de verificació d'un propietari. Al ser una taula creada per nosaltres i al no tenir un camp suficientment únic com per ser clau primària hem decidit afegir una nova columna "SERIAL" que farà aquestes funcions.

-Verifies

Aquesta es la taula sorgida de la relació "NxM" entre les taules "Host" i "Verification" no contindrà més informació que la necessària per relaciona-les adequadament.

-Reviewer

Taula amb la informació del usuari que lloga una propietat i deixa una ressenya sobre aquest últim. Una vegada més la normalització ens obliga a separar la informació que fa referencia a la ressenya en si en una taula diferent, o en aquets cas en atributs de la relació "NxM"

Cal destacar també que encara que el primer nom que vam donar a la taula va ser “User” aquesta és una paraula reservada per tant va ser necessary trobar una alternativa a aquest nom per la taula.

-ReviewerCommentProperty

Taula de relació entre “Property” i “Reviewer” conté també els atributs de la mateixa relació, els comentaris.

Aquesta taula és interessant ja que ens hem trobat amb un nou problema. Els dos PK/FK que caldria crear per fer una taula de relació NxM es restringeixen entre ells mateixos, per tant no els podem crear. Per solucionar-ho tal com se’ns indica al la diapositiva 22 del tema 3 dels apunts de l’assignatura, hem limitat els dos PK/FK a només FK i creat una nova clau primària.

-Amenity

Com s’ha esmentat anteriorment el camp “Amenities” de la taula original no era atòmic per aquesta raó ha estat necessari separar-lo en diverses entrades d’una nova taula.

-HasAmenity

Taula de relació entre “Amenity” i “Property”, com la relació no té atribut aquesta només contindrà les claus primàries de les taules d’origen com a claus primàries i forànees.

-Street

Tant aquesta taula com les que vindran a continuació són producte de la normalització de la base de dades.

D’altre costat, en els arxius originals el mateix nom pot estar escrit en formats diferents. No separar en taules diferents provocaria dificultats addicionals per unificar-los.

Per acabar en el cas particular d’aquesta taula i trobem el nom del carrer juntament amb un ID generat per nosaltres amb el tipus “SERIAL”

-Neighbourhood

Taula que conté nom dels barris i una clau primària generada per nosaltres.

-City

Taula que conté nom de les ciutats i una clau primària generada per nosaltres.

-State

Taula que conté nom dels estats i una clau primària generada per nosaltres.

Cal destacar que originalment el nom dels diferents estats podia estar complert o abreviat. Durant la fase d’importació hem unificat totes les entrades. Ampliarem aquest procés en la seva secció corresponent.

-Country

La taula country destaca per tenir un ID original i no creat per nosaltres. Hem considerat el codi de cada país com a suficientment únic i distintiu com per poder ser clau primària.

Aquesta taula contindrà a més el nom sencer del país.

3.2 Sobre les Taules Intermittents.

Considerem adient fer un petit comentari sobre certes taules auxiliars que s'han utilitzat durant el període d'importació de la informació però que no tenen l'objectiu d'emmagatzemar dades sinó només són una eina per tal d'introduir les dades a la seva destinació final.

Tant en el cas de les verificacions dels hosts com les amenitats de cada apartament hem utilitzat taules intermitges per, primer de tot separar tota la informació continguda a les cades i a continuació treure les comes i altres símbols que no aportaven informació per, finalment, unificar totes les tuples i passar-les a les taules finals.

3.3 Verificació del Model Físic.

A continuació afegim captures de pantalla que verifiquen la creació del model físic.

-Procediment Importació

```
COPY ImportApartments FROM 'D:\LASALLE\Bases de dades\1920\Practica\apartments.csv'
CSV HEADER DELIMITER ',';
COPY ImportHosts FROM 'D:\LASALLE\Bases de dades\1920\Practica\hosts.csv' CSV HEADER
DELIMITER ',';
COPY ImportReviews FROM 'D:\LASALLE\Bases de dades\1920\Practica\review.csv' CSV
HEADER DELIMITER ',';
```

```
[2019-11-30 13:42:57] completed in 25 ms
postgres.public> COPY ImportApartments FROM 'D:\LASALLE\Bases de dades\1920\Practica\apartments.csv' CSV HEADER DELIMITER ','
[2019-11-30 13:42:59] 22895 rows affected in 1 s 636 ms
postgres.public> COPY ImportHosts FROM 'D:\LASALLE\Bases de dades\1920\Practica\hosts.csv' CSV HEADER DELIMITER ','
[2019-11-30 13:43:01] 22895 rows affected in 1 s 356 ms
postgres.public> COPY ImportReviews FROM 'D:\LASALLE\Bases de dades\1920\Practica\review.csv' CSV HEADER DELIMITER ','
[2019-11-30 13:43:33] 486920 rows affected in 32 s 128 ms
```

-State

```
INSERT INTO State(name_state) SELECT DISTINCT state FROM ImportApartments;
```

id_state	name_state
1	NSW
2	QLD
3	VIC

-Country

```
INSERT INTO Country SELECT DISTINCT country_code, country FROM ImportApartments;
```

country_code	country_name
AU	Australia

-City

```
INSERT INTO City(name_city) SELECT DISTINCT city FROM ImportApartments;
```

id_city	name_city
1	Abbotsford
2	Aberfeldie
3	Airport West
4	Albanvale
5	Albert Park
6	Albion
7	Alphington
8	Altona
9	Altona Meadows
10	Altona North
11	Armadale
12	Armadale North
13	Ascot Vale
14	Ashburton
15	Ashwood
16	Aspendale
17	Aspendale Gardens
18	Attwood
19	Avondale Heights
20	Badger Creek
21	Balaclava
22	Balwyn
23	Balwyn North
24	Bangholme
25	Bayles
26	Bayswater
27	Bayswater North
28	Beaconsfield
29	Beaconsfield Upper
30	Beaumaris
31	Belgrave

-Neighbourhood

```
INSERT INTO Neighbourhood(name_neighbourhood) SELECT DISTINCT neighbourhood_cleansed
FROM ImportApartments;
```

id_neighbourhood	name_neighbourhood
1	Banyule
2	Bayside
3	Boroondara
4	Brimbank
5	Cardinia
6	Casey
7	Darebin
8	Frankston
9	Glen Eira
10	Greater Dandenong
11	Hobsons Bay
12	Hume
13	Kingston
14	Knox
15	Manningham
16	Maribyrnong
17	Maroondah
18	Melbourne
19	Melton
20	Monash
21	Moonee Valley
22	Moreland
23	Nillumbik
24	Port Phillip
25	Stonnington
26	Whitehorse
27	Whittlesea
28	Wyndham
29	Yarra
30	Yarra Ranges

-Street

```
INSERT INTO Street (name_street) SELECT DISTINCT street FROM ImportApartments ;
```

id_street	name_street
1	Abbotsford, Melbourne, VIC, Australia
2	Abbotsford, VIC, Australia
3	Aberfeldie, VIC, Australia
4	Airport West, VIC, Australia
5	Albanvale, VIC, Australia
6	Albert Park, VIC, Australia
7	Albert Park, Melbourne, VIC, Australia
8	Albion, VIC, Australia
9	Alphington, VIC, Australia
10	Altona Meadows, VIC, Australia
11	Altona North, VIC, Australia
12	Altona, VIC, Australia
13	Armada North, VIC, Australia
14	Armada North, VIC, Australia
15	Armada North, VIC, Australia
16	Ascot Vale, VIC, Australia
17	Ashburton, VIC, Australia
18	Ashwood, VIC, Australia
19	Aspendale Gardens, VIC, Australia
20	Aspendale, VIC, Australia
21	Attwood, Australia
22	AU, Australia
23	Avondale Heights, VIC, Australia
24	Badger Creek, VIC, Australia
25	Balaclava, VIC, Australia
26	Balwyn North, VIC, Australia
27	Balwyn North, VIC, Australia
28	Balwyn, VIC, Australia
29	Bangholme, VIC, Australia
30	Bayles, VIC, Australia
31	Bayswater North, VIC, Australia

-Property

```
INSERT INTO Property(listing_url, name_property, description, picture_url, zipcode,
property_type, max_residents, bathrooms, bedrooms, beds, squarefeet, price,
weekly_price, monthly_price, security_deposit, cleaning_fee, min_nights, max_nights
,id_host, country_code, id_state, id_city, id_neighbourhood, id_street)
SELECT DISTINCT ImportApartments.listing_url, ImportApartments.name,
ImportApartments.description, ImportApartments.picture_url, zipcode, property_type,
accommodates, bathrooms, bedrooms, beds, square_feet, price, weekly_price,
monthly_price, security_deposit, cleaning_fee, minimum_nights, maximum_nights,
host_id, ImportApartments.country_code,State.id_state, City.id_city,
Neighbourhood.id_neighbourhood, id_street
FROM ImportApartments, ImportHosts, state, city, Neighbourhood, street WHERE
ImportHosts.listing_url = ImportApartments.listing_url AND (ImportApartments.state =
name_state AND ImportApartments.city = name_city AND neighbourhood_cleansed =
name_neighbourhood AND name_street = ImportApartments.street) ORDER BY listing_url
ASC;
```

id_prop...	listing_url	name_property	description	picture_url
1	https://www.airbnb.com/rooms/10000147	Sanctuary Parkview Resort	A tranquil, light-filled sanctuary, away from the rush and bustle, where you can relax & re	
2	https://www.airbnb.com/rooms/10000335	Furnished Opposite Monash Uni Clayton (Room 5)	Registered Rooming house with 8 bedrooms and 3 bathrooms. . https://a0.muscache.com/im/pictu	
3	https://www.airbnb.com/rooms/10000999	Sanctuary Parkview Resort	The light-filled tranquility of our parkland & lake setting. https://a0.muscache.com/im/pictu	
4	https://www.airbnb.com/rooms/10001744	Tulloch Brae #1	this is the first bedroom at Tulloch Brae Queen size bed a. https://a0.muscache.com/im/pictu	
5	https://www.airbnb.com/rooms/10003237	Ideal location with Gym and Pool	This apartment is located just behind the Iconic Arts cent. https://a0.muscache.com/im/pictu	
6	https://www.airbnb.com/rooms/1001457	Docklands Apartment with River Views	A generous size apartment located in Docklands with River . https://a0.muscache.com/im/pictu	
7	https://www.airbnb.com/rooms/1001739	Private room 1, Box Hill, Melbourne	We have two comfortable bedrooms house available in easter. https://a0.muscache.com/im/pictu	
8	https://www.airbnb.com/rooms/10020122	Stunning Boutique Master Suite	Stay in our Master bedroom with ensuite in our recently re. https://a0.muscache.com/im/pictu	
9	https://www.airbnb.com/rooms/1002094	Heart of the CBD	Spacious apartment fully equipped for the short to medium . https://a0.muscache.com/im/pictu	
10	https://www.airbnb.com/rooms/10021017	Beautiful Terrace House + Courtyard	This renovated two bedroom home is easy walking distance t. https://a0.muscache.com/im/pictu	
11	https://www.airbnb.com/rooms/10021541	Luxury apartment in South Yarra. Walk to MCG	The apartment is located in the heart of south yarra just . https://a0.muscache.com/im/pictu	
12	https://www.airbnb.com/rooms/10022934	quiet stylish spacious inner north	Melbourne's inner north sunny 2 room spacious studio apart. https://a0.muscache.com/im/pictu	
13	https://www.airbnb.com/rooms/10024832	Right on the Yarra River	Chic and stylish 1 bedroom apartment in a picturesque sett. https://a0.muscache.com/im/pictu	
14	https://www.airbnb.com/rooms/10026095	Fiesty-Independent? (1r/1p*)	This is 1 of 4 rooms. (1r = 1 room + 1 set of keys 1p = 1. https://a0.muscache.com/im/pictu	
15	https://www.airbnb.com/rooms/10026682	Lucky in Love? (1r/1-2 pp)	This is 1 of 4 rooms hired to Guests. * (1r = 1 room + 1 s. https://a0.muscache.com/im/pictu	
16	https://www.airbnb.com/rooms/10028048	"Carlisle" - Pearl Apartments	Great spacious studio right in the heart of St Kilda. Fir. https://a0.muscache.com/im/pictu	
17	https://www.airbnb.com/rooms/10030726	A home away from home with carpark!	Filled with an abundance of natural light, this spacious 1. https://a0.muscache.com/im/pictu	
18	https://www.airbnb.com/rooms/10037624	Furnished, convenient, scenic room	You Se looking at one of the most convenient, scenic and s. https://a0.muscache.com/im/pictu	
19	https://www.airbnb.com/rooms/10037870	Single bed, sunset views, quiet Apt	Quiet 2 bedroom apartment with Gorgeous light filled Art D. https://a0.muscache.com/im/pictu	
20	https://www.airbnb.com/rooms/10039033	Cosy Comfort in Berwick	One bedroom with Queen size bed & shared bathroom or Maste. https://a0.muscache.com/im/pictu	
21	https://www.airbnb.com/rooms/10039270	Green Urban Living	Marvellous verdant location for all seasons including easy. https://a0.muscache.com/im/pictu	
22	https://www.airbnb.com/rooms/10039723	Two Storey Townhouse Richmond Hill	In an exceptional well-maintained townhouse complete with . https://a0.muscache.com/im/pictu	
23	https://www.airbnb.com/rooms/10040503	Restored 1910 Edwardian Home.	Conveniently located on the iconic #11 tram line. This bea. https://a0.muscache.com/im/pictu	
24	https://www.airbnb.com/rooms/10041233	Vance, eclectic Melb city APTMT	In the very centre of Melbourne, Vance is a cool and colour. https://a0.muscache.com/im/pictu	
25	https://www.airbnb.com/rooms/10041602	5* Luxurious Apartment - Perfect for business	Perfect for business trips or leisure. The complex offers. https://a0.muscache.com/im/pictu	
26	https://www.airbnb.com/rooms/10042345	Bay View Apartment The Icon 1202	Free Parking inside the building Free WIFI Fully access. https://a0.muscache.com/im/pictu	
27	https://www.airbnb.com/rooms/1004260	Apartment in Fitzroy Melbourne	Well kept modern first floor apartment. Moments from resta. https://a0.muscache.com/im/pictu	
28	https://www.airbnb.com/rooms/10042775	Family Friendly Bayside Residence	Bright freshly painted, carpeted and floor boarded street . https://a0.muscache.com/im/pictu	
29	https://www.airbnb.com/rooms/10045067	Wonderful Victorian spacious home	Well located near hawkesburn station and chapel st and too. https://a0.muscache.com/im/pictu	
30	https://www.airbnb.com/rooms/10047160	Cosy, sun filled apartment	1 bedroom apartment in a small block, this space is perfec. https://a0.muscache.com/im/pictu	

-Amenity

```
-- Insertem les diferents amenities dels apartaments
-- Insertem les amenities utilitzant la funció regexp_split_to_table() Per separar la
cadena de caracters en cada amenity i les funcions trim i replace per treure
caracters no desitjats
DROP TABLE IF EXISTS AmenityIntermig;
CREATE TABLE AmenityIntermig(
    name_Amenity VARCHAR(255)
);

INSERT INTO AmenityIntermig SELECT regexp_split_to_table(amenities, ',') FROM
ImportApartments;
UPDATE AmenityIntermig SET name_Amenity = replace(name_Amenity, ',', '');
UPDATE AmenityIntermig SET name_Amenity = ltrim(name_Amenity, '{');
UPDATE AmenityIntermig SET name_Amenity = rtrim(name_Amenity, '}');
```

```

UPDATE AmenityIntermig SET name_Amenity = trim(name_Amenity);

-- Afegim la informació obtinguda a la taula Amenity
INSERT INTO Amenity(name_amenity) SELECT DISTINCT * FROM AmenityIntermig;

-- Eliminem la verificació que és un espai en blanc
UPDATE Amenity SET name_amenity = 'No té amenities' WHERE name_amenity = '';

DROP TABLE IF EXISTS AmenityIntermig;

```

id_amenity	name_amenity
1	24-hour check-in
2	Bathtub with bath chair
3	Dog(s)
4	Hangers
5	Breakfast table
6	Full kitchen
7	Coffee maker
8	Roll-in shower
9	Wide clearance to bed
10	Safety card
11	Beach essentials
12	Body soap
13	TV
14	Other
15	Mini fridge
16	Fire pit
17	Gas oven
18	Pocket wifi
19	Projector and screen
20	Hot water
21	Iron
22	Children's dinnerware
23	Jetted tub
24	Cat(s)
25	Babysitter recommendations
26	Pets allowed
27	Warming drawer
28	Amazon Echo
29	Patio or balcony
30	Printer
31	Paid parking off premises

-HasAmenity

```

INSERT INTO HasAmenity SELECT DISTINCT id_property, id_amenity FROM Amenity,
Property, ImportApartments WHERE Property.listing_url = ImportApartments.listing_url
AND position(name_amenity IN amenities) != 0 ;

```

id_property	id_amenity
1	1
1	4
1	13
1	21
1	42
1	46
1	58
1	61
1	65
1	70
1	76
1	80
1	83
1	86
1	138
1	140
1	142
1	153
1	157
1	163
1	172
1	173
1	180
2	13
2	21
2	42
2	46
2	58
2	74
2	76
2	80

-Host

```
INSERT INTO Host(id_host, name_host, host_url, host_since, host_about,
host_response_time, host_response_rate, host_is_superhost, host_picture_url,
host_listings, host_identity_verified)
SELECT DISTINCT host_id, host_name, host_url, host_since, host_about,
host_response_time, host_response_rate, host_is_superhost, host_picture_url,
host_listings_count, host_identity_verified FROM ImportHosts;
```

id_host	name_host	host_url	host_since	host_about	host_response_time	host_response_r
9082	Dennis	https://www.airbnb.com/users/show/9082	2009-03-05	I am a complete geek, from the automated least distance pa. within a few hours	100%	
12740	Nicky	https://www.airbnb.com/users/show/12740	2009-04-09	30 something female photographer, easy going friendly, lov. N/A	N/A	
15565	Minnie	https://www.airbnb.com/users/show/15565	2009-05-02	I am an mid 30's professional living and working in Melbou. within an hour	100%	
17308	Romi	https://www.airbnb.com/users/show/17308	2009-05-16	Hello!@My first experience with Air BnB was as a guest ir within an hour	100%	
18785	Kate	https://www.airbnb.com/users/show/18785	2009-05-25	Friendly filmmaking professional who loves to travel, meet. within an hour	100%	
23632	Carmen	https://www.airbnb.com/users/show/23632	2009-06-29	Having lived between Melbourne, Paris and New York in the . N/A	N/A	
26089	Erin	https://www.airbnb.com/users/show/26089	2009-07-17	within a few hours	100%	
26687	Rachel	https://www.airbnb.com/users/show/26687	2009-07-21	I'm an editor and I work across a broad spectrum of projec. within an hour	100%	
28967	Adam And Jackie	https://www.airbnb.com/users/show/28967	2009-07-31	We're a married couple, living and enjoying life in Heales. N/A	N/A	
33057	Manju	https://www.airbnb.com/users/show/33057	2009-08-21	About us: We are an East Indian couple in our sixties who . a few days on more	8%	
34705	Rahul	https://www.airbnb.com/users/show/34705	2009-08-29	Hi. I am Rahul and I am a 44 yr old married man. My wife A. within a day	100%	
38901	Lindsay	https://www.airbnb.com/users/show/38901	2009-09-16	As an artist working in animation and video I am involved . N/A	N/A	
40864	Jane	https://www.airbnb.com/users/show/40864	2009-09-23	N/A	N/A	
41698	Trang	https://www.airbnb.com/users/show/41698	2009-09-27	Full time professional currently living in Melbourne, Aus.. N/A	N/A	
108825	Allissa	https://www.airbnb.com/users/show/108825	2010-04-15	Photographer from Melbourne. Live between Melbourne and Da. within a day	100%	
111551	Blake	https://www.airbnb.com/users/show/111551	2010-04-20	Melbourne-based actor-singer. within an hour	100%	
47366	Inger	https://www.airbnb.com/users/show/47366	2009-10-21	My partner and I have relocated to Melbourne from several . N/A	N/A	
48538	Michael	https://www.airbnb.com/users/show/48538	2009-10-27	Laid back Australian undergrad student. Enjoy having fun a. within an hour	100%	
50121	Frank & Vince	https://www.airbnb.com/users/show/50121	2009-10-31	Hi, we're Frank and Vince.@Thanks for reading our profile within an hour	100%	
57747	Pippa	https://www.airbnb.com/users/show/57747	2009-11-26	Having lived in London and Western Australia, we are now b. N/A	N/A	
59786	Eleni	https://www.airbnb.com/users/show/59786	2009-12-03	Artist, academic, community activist, traveller, meditator. N/A	N/A	
65090	Colin	https://www.airbnb.com/users/show/65090	2009-12-22	Retired Australian traveller, I run a small gay - friendly. within an hour	100%	
65179	Robyn	https://www.airbnb.com/users/show/65179	2009-12-22	I am a lover of plants, especially herbs, gardening, the o. within an hour	100%	
66318	Paul	https://www.airbnb.com/users/show/66318	2009-12-28	Relaxed, silly, adventurous and interested.... beer, wine,. within an hour	100%	
82422	Amy	https://www.airbnb.com/users/show/82422	2010-02-17	I'm a fun loving laid back Aussie girl who loves to travel. within an hour	100%	
89766	Georgia	https://www.airbnb.com/users/show/89766	2010-03-08	Hi, I am 35 and an architect and artist. @I am a true home within a day	100%	
94675	Ariel	https://www.airbnb.com/users/show/94675	2010-03-16	I am a 33 year old music producer from Australia. @I've tr within a day	100%	
96945	Kylie	https://www.airbnb.com/users/show/96945	2010-03-21	I am a mature and independent woman from Melbourne who lov. within an hour	100%	
98515	Aleksandra	https://www.airbnb.com/users/show/98515	2010-03-24	My housemate and I are both friendly, professional creativ. within a few hours	100%	
100972	Cass & Jeremy	https://www.airbnb.com/users/show/100972	2010-03-30	We are busy travelling and living life => @Please enjoy y within a few hours	100%	

-Reviewer

```
INSERT INTO Reviewer SELECT DISTINCT reviewer_id, reviewer_name FROM ImportReviews;
```

id_reviewer	name_reviewer
262	Jonathan
299	Tristan
1262	Rebecca
1510	John
1698	Luc
1872	Juan Pablo
3000	Andreas
3172	David
3274	Kristen
4168	Chloe
4392	Lee-Ann
5184	Bjorgvin
5598	Henok
5629	Irina
6467	Lee
6995	Stephen
7054	Mei Lin
7365	Elisa
7834	Friska
7859	Noam
8057	Chris And Nick
8222	Jessica
8340	Craig
8728	Rosie
9631	Phil
10347	Sam
10592	Susan
10869	Deborah
10886	Anja & JP
10957	Ben
11097	Hamish

-ReviewerCommentProperty

```
INSERT INTO ReviewerCommentProperty(id_property, id_reviewer, date_review, comment)
SELECT id_property, reviewer_id, date_review, comments FROM ImportReviews, Property
WHERE ImportReviews.listing_url = Property.listing_url;
```

id_reviewercomment	id_property	id_reviewer	date_review	comment
1	584	4421189	2013-01-12	It was very convenient to stay in Lindsay's apartment in.
2	584	1763045	2013-01-18	Perfect isnt enough! Lindsay was the best host we could ev.
3	584	4423532	2013-01-23	Living with Lindsay was very relaxed. The room is chosen .
4	584	4551787	2013-02-11	Beautiful home, great location very friendly and helpful t.
5	584	632036	2013-03-01	It was great staying at Lindsay's apartment. He's friendly.
6	584	2553303	2013-03-26	Lindsay was an excellent host, his apartment is in a great.
7	584	5610658	2013-04-05	Lovely room and home, and very gracious hosts. Lots of cha.
8	584	1836341	2013-06-02	A great room in a great location! The room is very large .
9	584	5630666	2013-06-04	The reservation was canceled 2 days before arrival. This i.
10	584	6553329	2013-06-19	Lindsay is a really kind and accommodating host! His place.
11	584	6578448	2013-06-27	Very spacious, comfy and quiet room, & lovely host. Super
12	584	6589030	2013-08-05	The reservation was canceled 28 days before arrival. This .
13	584	8378014	2013-12-13	Lindsay has a great living space over his yoga studio. The.
14	584	10719843	2014-01-09	I had a wonderful stay. Lindsay is very friendly, flexible.
15	584	10941298	2014-01-18	I really enjoyed my stay at Lindsay's apartment and I woul.
16	315	13438280	2015-06-18	Helen was really helpful and nice, box hill is a quiet are.
17	584	10174200	2014-02-05	Lindsay is relaxed and easy to get along with a pleasure t.
18	584	11360615	2014-02-24	Lindsay is a very nice host: respectful and helpful. The r
19	584	12407146	2014-03-02	Staying for 4 days in Lindsay's house was an amazing exper.
20	584	13599943	2014-07-13	Lindsay was great. Really a sweet and extremely kind guy. .
21	584	2350539	2014-07-22	Lindsay's home is easily accessible by public transport, c.
22	315	40904079	2015-08-18	房子非常干净,房东非常好
23	584	17961657	2014-08-04	The location is very good and there are nice places really.
24	4548	56566140	2016-09-27	Welcoming host. Fresh and clean. Warm and comfortable bed..
25	584	15035042	2014-08-11	Lindsay is a gentle soul and I felt very safe being in his.
26	584	2838033	2014-08-16	Conveniently located near the tram stop. Comfortable room
27	584	19715778	2014-08-20	Lindsay was a fantastic host. The room was very clean and
28	21280	1824481	2012-04-05	This was our first airbnb experience and certainly not the.
29	584	3242826	2014-09-12	All good. I loved looking at all your stuff and perusing y.
30	584	3242826	2014-09-14	Even though Lindsay wasn't there when I arrived, he left .
31	584	467452	2014-09-26	Great place, cool vibe, Lindsay's very easy-going and the..

-Verification

```

DROP TABLE IF EXISTS VerificationIntermig;
CREATE TABLE VerificationIntermig(
    name_verification VARCHAR(255)
);
INSERT INTO VerificationIntermig SELECT regexp_split_to_table(host_verifications,
',') FROM ImportHosts;

UPDATE VerificationIntermig SET name_verification = replace(name_verification, '',
'');
UPDATE VerificationIntermig SET name_verification = ltrim(name_verification, '[');
UPDATE VerificationIntermig SET name_verification = rtrim(name_verification, ']');
UPDATE VerificationIntermig SET name_verification = trim(name_verification);

-- Afegim la informació obtinguda a la taula Verification
INSERT INTO Verification(name_verification) SELECT DISTINCT * FROM
VerificationIntermig;

DROP TABLE IF EXISTS VerificationIntermig;

```


id_verification	name_verification
1	email
2	jumio
3	facebook
4	google
5	zhima_selfie
6	kba
7	manual_offline
8	None
9	government_id
10	manual_online
11	selfie
12	offline_government_id
13	work_email
14	phone
15	sent_id
16	sesame
17	reviews
18	identity_manual
19	sesame_offline
20	weibo

-Verifies

```
INSERT INTO Verifies SELECT DISTINCT id_verification, id_host FROM Verification,
Host, ImportHosts WHERE host_id = id_host AND position(name_verification IN
host_verifications) != 0 ;
```

id_verification	id_host
1	9082
1	12740
1	15565
1	17308
1	18785
1	23632
1	26089
1	26687
1	28967
1	33057
1	34705
1	38901
1	41698
1	47366
1	48538
1	50121
1	57747
1	59786
1	65090
1	65179
1	66318
1	82422
1	89766
1	94675
1	96945
1	98515
1	100972
1	108825
1	111551
1	112497
1	113346

4. PROCÉS D'IMPORTACIÓ

4.1 Consideracions durant la importació

Al llarg del procés durant el qual hem copiat les dades de les taules d'importació a les seves taules finals del model físic hem pres algunes decisions que cal comentar. Aquestes son les següents.

- Modificació camps State per tal de unificar-los sota un mateix criteri

- Omplir camps State buids

- Modificació Hanzl

- Modificació taula reviews/rents

- Al camp city de ImportApartments hem fet diverses modificacions per tal que tot estigués sota un mateix criteri. Hem retirat paraules que no aportaven informació i feien que hi haguessin tuples que un cop exportades es dupliquessin. En el cas que no hi hagués cap ciutat hem introduït el nom del barri ja que fent una cerca hem vist que coincidia amb el nom de la ciutat.

- A ImportHosts hem tret els camps Verification en els que no hi havia res escrit o simplement hi havia '[]' i els hem canviat per 'None' per tal que no donessin problemes a l'hora de l'exportació cap a les taules corresponents

- Per poder fer les quèries de la fase 2 i les seves comprovacions hem hagut de modificar alguns paràmetres de les taules *importapartments* i *importhosts* per tal de poder dur a terme dites comprovacions. Aquests canvis són bàsicament canvis de tipus en les columnes dels preus ja que nosaltres els havíem guardat com a Varchar i els passem a real per poder operar amb ells. En el fitxer Fase2.sql hi ha comentaris abans dels canvis on s'explica el perquè de cadascun d'ells.

4.2 Verificació de la importació

A continuació afegim captures per comprovar que la importació a les taules de la base de dades ha sigut correcte. Ho farem utilitzant un exemple de relació entre Property i les seves diferents característiques.

-ImportApartment

Aquí podem veure la informació proporcionada a través del fitxer .csv. Serà la informació que ens servirà de referència per la comprovació.

```
SELECT * FROM importapartments WHERE listing_url =  
'https://www.airbnb.com/rooms/3802605'
```

	1
id	3802605
listing_url	https://www.airbnb.com/rooms/3802605
name	Stunning town residence, Melbourne
description	My home is located in a great central location. It's in a beautiful and quiet tree lined
picture_url	https://a0.muscache.com/im/pictures/35485dd4-263b-40b1-868c-f3fbfa62fb06.jpg?aki_policy=l
street	Elsternwick, VIC, Australia
neighbourhood_cleansed	Glen Eira
city	Elsternwick
state	VIC
zipcode	3185
country_code	AU
country	Australia
property_type	House
accommodates	4
bathrooms	2.5
bedrooms	2
beds	3
amenities	{TV,Internet,Wifi,"Air conditioning","Wheelchair accessible",Kitchen,"Free parking on pre
square_feet	<null>
price	339
weekly_price	2373
monthly_price	<null>
security_deposit	0
cleaning_fee	90
minimum_nights	1
maximum_nights	1125

-Property

Taula que conté la informació dels apartaments, relacionada amb el host, state, country, city, neighbourhood i street

```
SELECT * FROM property WHERE listing_url = 'https://www.airbnb.com/rooms/3802605'
```

	1
id_property	19615
listing_url	https://www.airbnb.com/rooms/3802605
name_property	Stunning town residence, Melbourne
description	My home is located in a great central location. It's in a beautiful and qu:
picture_url	https://a0.muscache.com/im/pictures/35485dd4-263b-40b1-868c-f3fbfa62fb06.jpg
zipcode	3185
property_type	House
max_residents	4
bathrooms	2.5
bedrooms	2
beds	3
squarefeet	<null>
price	339
weekly_price	2373
monthly_price	10170
security_deposit	0
cleaning_fee	90
min_nights	1
max_nights	1125
id_host	7035315
country_code	AU
id_state	3
id_city	151
id_neighbourhood	9
id_street	184

-City

```
SELECT * FROM city WHERE id_city = 151;
```

	1
id_city	151
name_city	Elsternwick

-Neighbourhood

```
SELECT * FROM neighbourhood WHERE id_neighbourhood = 9;
```

	1
id_neighbourhood	9
name_neighbourhood	Glen Eira

-Street

```
SELECT * FROM street WHERE id_street = 184;
```

	1
id_street	184
name_street	Elsternwick, VIC, Australia

-Host

```
SELECT * FROM host WHERE id_host = 7035315;
```

	1
id_host	7035315
name_host	Jonathan
host_url	https://www.airbnb.com/users/show/7035315
host_since	2013-06-21
host_about	I am (currently) a single father to a beautiful 10 y/o, who I have
host_response_time	within an hour
host_response_rate	100
host_is_superhost	<input checked="" type="checkbox"/>
host_picture_url	https://a0.muscache.com/im/pictures/2aeaa052-5dd1-4660-b0a3-c37cb
host_listings	3
host_identity_verified	<input checked="" type="checkbox"/>

-ImportHosts

```
SELECT * FROM importhosts WHERE listing_url = 'https://www.airbnb.com/rooms/3802605';
```

	1
listing_url	https://www.airbnb.com/rooms/3802605
name	Stunning town residence, Melbourne
description	My home is located in a great central location. It's in a beautiful and quiet tree lined street that
picture_url	https://a0.muscache.com/im/pictures/35485dd4-263b-40b1-868c-f3fbfa62fb06.jpg?aki_policy=large
host_id	7035315
host_url	https://www.airbnb.com/users/show/7035315
host_name	Jonathan
host_since	2013-06-21
host_about	I am (currently) a single father to a beautiful 10 y/o, who I have part of the time only. Am the dire
host_response_time	within an hour
host_response_rate	100
host_is_superhost	<input checked="" type="checkbox"/>
host_picture_url	https://a0.muscache.com/im/pictures/2aeaa052-5dd1-4660-b0a3-c37cb2b763ab.jpg?aki_policy=profile_x_med
host_listings_count	3
host_verifications	['email', 'phone', 'reviews', 'jumio', 'offline_government_id', 'government_id']
host_identity_verified	<input checked="" type="checkbox"/>

Amb aquestes captures podem veure com la informació s'ha importat correctament a totes les taules i s'han relacionat correctament entre elles.

5.FASE 2

5.1 Query 1

```
SELECT name_city AS name, ((avg(price*7) - avg(weekly_price)) / avg(price*7)) * 100
AS saving_percentage
FROM city, property, host WHERE city.id_city = property.id_city AND host.id_host =
property.id_host
AND host_identity_verified IS TRUE AND price > 0 AND min_nights < 7 AND max_nights >=
7 AND price*7 > weekly_price
GROUP BY city.id_city
ORDER BY saving_percentage DESC LIMIT 3;
```

	name	saving_percentage
1	Cranbourne South	79.22077922077922
2	Wattle Glen	64.28571428571429
3	Doncaster East	58.38509316770186

Comprovació

```
SELECT city AS name, ((avg(price*7) - avg(weekly_price)) / avg(price*7)) * 100 AS
saving_percentage
FROM importapartments, importhosts WHERE importhosts.listing_url =
importapartments.listing_url
AND host_identity_verified IS TRUE AND price > 0 AND minimum_nights < 7 AND
maximum_nights >= 7 AND price*7 > weekly_price
GROUP BY city
ORDER BY saving_percentage DESC LIMIT 3;
```

	name	saving_percentage
1	Cranbourne South	79.22077922077922
2	Wattle Glen	64.28571428571429
3	Doncaster East	58.38509316770186

5.2 Query 2

```
SELECT name_property, round((price/squarefeet)::numeric,2):: money AS price_m2,
count(reviewercommentproperty.id_property) AS reviews
FROM property, reviewercommentproperty
WHERE property.id_property = reviewercommentproperty.id_property AND squarefeet IS
NOT NULL AND squarefeet > 0
AND property_type = 'Guesthouse'
```

```
GROUP BY name_property, (price/squarefeet)
HAVING count(id_reviewercomment) >=200
ORDER BY price_m2 DESC LIMIT 1;
```

	name_property	price_m2	reviews
1	The Stables, Island of Richmond	\$0.37	538

Comprovació

```
SELECT importapartments.name, round((price/square_feet)::numeric,2):: money AS
price_m2, count(importreviews.listing_url) AS reviews
FROM importapartments, importreviews
WHERE importreviews.listing_url = importapartments.listing_url AND square_feet IS NOT
NULL AND square_feet > 0
AND property_type = 'Guesthouse'
GROUP BY importapartments.name, round((price/square_feet)::numeric,2):: money
HAVING count(importreviews.listing_url) >=200
ORDER BY price_m2 DESC LIMIT 1;
```

	name	price_m2	reviews
1	The Stables, Island of Richmond	\$0.37	538

5.3 Query 3

```
SELECT name_property AS name, listing_url AS url, (((price*5)*6) + cleaning_fee) +
(security_deposit*0.1)::numeric:: money AS price FROM property, neighbourhood,
hasamenity, amenity, host
WHERE property.id_property = hasamenity.id_property AND property.id_neighbourhood =
neighbourhood.id_neighbourhood AND host.id_host = property.id_host
AND amenity.id_amenity = hasamenity.id_amenity
AND max_residents = 6 AND name_amenity = 'Balcony' AND bathrooms > 1.5 AND
host_response_rate > 90 AND name_neighbourhood = 'Port Phillip'
ORDER BY price ASC LIMIT 1;
```

	name	url	price
1	Spacious Designer Apartment near Albert Park	https://www.airbnb.com/rooms/12528535	\$10,050.00

Comprovació

```
SELECT importapartments.name AS name, importapartments.listing_url AS url,
(((price*5)*6) + cleaning_fee) + (security_deposit*0.1):: numeric) AS price FROM
importapartments, importhosts
WHERE importhosts.listing_url = importapartments.listing_url
AND accommodates = 6 AND amenities ILIKE '%,Balcony%' AND bathrooms > 1.5 AND
host_response_rate > 90 AND neighbourhood_cleansed = 'Port Phillip'
ORDER BY price ASC LIMIT 1 ;
```

	name	url	price
1	Spacious Designer Apartment near Albert Park	https://www.airbnb.com/rooms/12528535	10050

5.4 Query 4

```
UPDATE host SET host_is_superhost = TRUE WHERE host_since <= '2014-12-31';
UPDATE host SET host_is_superhost = FALSE WHERE host_since > '2014-12-31';
SELECT count(id_host) AS superhost FROM host WHERE host_is_superhost = TRUE;
SELECT count(id_host) AS normal_hosts FROM host WHERE host_is_superhost = FALSE;
```

	normal_hosts		superhost
1	10028	1	4957

Comprovació

```
UPDATE importhosts SET host_is_superhost = TRUE WHERE host_since <= '2014-12-31';
UPDATE importhosts SET host_is_superhost = FALSE WHERE host_since > '2014-12-31';
SELECT count(id_host) AS superhost FROM host WHERE host_is_superhost = TRUE;
SELECT count(id_host) AS normal_hosts FROM host WHERE host_is_superhost = FALSE;
```

	normal_hosts		superhost
1	10028	1	4957

5.5 Query 5

```
SELECT name_street AS street, COUNT(street.id_street) AS num,
round(avg(price)::numeric,2)::money AS price FROM street, property
WHERE street.id_street = property.id_street
GROUP BY street.id_street
HAVING AVG(price) < 100
ORDER BY num DESC LIMIT 3;
```

	street	num	price
1	Brunswick, VIC, Australia	437	\$94.80
2	Footscray, VIC, Australia	165	\$93.05
3	Preston, VIC, Australia	142	\$94.37

Comprovació

```
SELECT street AS street, COUNT(listing_url) AS num,
round(avg(price)::numeric,2)::money AS price FROM importapartments
GROUP BY street
HAVING AVG(price) < 100
ORDER BY num DESC LIMIT 3;
```

	street	num	price
1	Brunswick, VIC, Australia	437	\$94.80
2	Footscray, VIC, Australia	165	\$93.05
3	Preston, VIC, Australia	142	\$94.37

5.6 Query 6

```
SELECT name_reviewer, listing_url AS url, COUNT(id_reviewercomment) AS num_reviews
FROM reviewercommentproperty, reviewer, property
WHERE reviewercommentproperty.id_reviewer = reviewer.id_reviewer AND
property.id_property = reviewercommentproperty.id_property
GROUP BY reviewer.id_reviewer, name_reviewer, listing_url
ORDER BY num_reviews DESC, name_reviewer DESC LIMIT 3 ;
```

	name_reviewer	url	num_reviews
1	Cameron	https://www.airbnb.com/rooms/7842455	81
2	Therese	https://www.airbnb.com/rooms/1995746	24
3	Michael	https://www.airbnb.com/rooms/9263414	23

Veient els comentaris dels diferents reviewers hem decidit que el Cameron està fent reviews falses ja que té moltes reviews en dies molt seguits i tots els comentaris són molt semblants, fet que ens fa sospitar que és ell qui està escrivint reviews falses. Per això no el tindrem en compte en pròximes queries, ja que pot alterar els resultats desitjats.

Comprovació

```
SELECT reviewer_name, importreviews.listing_url AS url,
COUNT(importreviews.listing_url) AS num_reviews FROM importapartments,importreviews
WHERE importapartments.listing_url = importreviews.listing_url
GROUP BY reviewer_name,importreviews.listing_url, reviewer_id
ORDER BY num_reviews DESC, reviewer_name DESC LIMIT 3 ;
```

	reviewer_name	url	num_reviews
1	Cameron	https://www.airbnb.com/rooms/7842455	81
2	Therese	https://www.airbnb.com/rooms/1995746	24
3	Michael	https://www.airbnb.com/rooms/9263414	23

5.7 Query 7

```
SELECT property.id_property AS id, name_property AS name, (((price*2*2) +
cleaning_fee) + (security_deposit*0.1))::numeric:: money AS total_price FROM
city,property, amenity, hasamenity, host, verification, verifies
WHERE host.id_host = property.id_host AND city.id_city = property.id_city AND
amenity.id_amenity = hasamenity.id_amenity
AND property.id_property = hasamenity.id_property AND verifies.id_verification =
verification.id_verification
AND host.id_host = verifies.id_host
AND max_residents >= 2 AND beds >= 2 AND name_city = 'Saint Kilda' And name_amenity =
'Kitchen' AND name_verification = 'phone'
AND (((price*2*2) + cleaning_fee) + (security_deposit*0.1)) < 5000 AND min_nights <=
2
GROUP BY property.id_property
ORDER BY total_price DESC;
```

	id	name	total_price
1	19156	★ Beach at your Doorstep: Enjoy Summer in St Kilda	\$4,166.00
2	3370	Premium property in great location	\$2,000.00
3	16064	Luxury Vintage Melbourne Loft	\$1,807.00
4	6160	Gorgeous Stkilda Apartments Sleeps 14 parking	\$1,805.00
5	1408	Spacious Designer Apartment near Albert Park	\$1,600.00
6	19777	Design House near Saint Kilda	\$1,550.00
7	4147	Huge St Kilda Beach PentHouse Beauty	\$1,540.00
8	14558	Stylish and Spacious Contemporary Apartment in St Kilda	\$1,500.00
9	10585	Family and Large Group Parking WIFI Exc Location	\$1,436.00
10	5225	2 Br Apartment fully furnished near albert park	\$1,364.00
11	11498	Modern Apart Opposite Albert Park with Breakfast	\$1,360.00
12	10506	St Kilda Lux Penthouse	\$1,305.00
13	11678	**Totally Renovated in 2018. Beautiful Interiors**	\$1,300.00
14	14294	**Newly Renovated. Cityview. Stunning Interiors**	\$1,256.00
15	5777	ST KILDA - Grand Mansion Living.	\$1,245.00
16	11120	Mediterranean Townhouse in the heart of it all!	\$1,226.00
17	1025	Huge, stylish, central, 3 br apartment with pool.	\$1,212.00
18	7710	cosy townhouse with nature light	\$1,200.00
19	7737	Classy 3BR apt on iconic Chapel St!	\$1,196.00
20	11731	From St Kilda to Melbourne City the Events Capital	\$1,156.00
21	16113	Dream St Kilda Apartment	\$1,144.00
22	14891	Brand new Apartment 2km to Luna park beach	\$1,126.00
23	9760	Top-Floor Contemporary Designer Flat	\$1,119.00
24	10333	★Massive bright 3BR home by the bay★	\$1,113.00
25	17145	Best View in Melbourne (St Kilda), SunRise&SunSet	\$1,111.00
26	3571	Contemporary Apartment with Botanic Gardens view	\$1,099.00
27	4123	Modern apartment in the heart of St Kilda	\$1,060.00
28	11010	Large 2 BRM apt, close to restaurants St Kilda bch	\$1,059.00

Comprovació

```
SELECT id AS id, importapartments.name AS name, (((price*2*2) + cleaning_fee) +
(security_deposit*0.1))::numeric:: money AS total_price FROM importapartments,
importhosts
WHERE importhosts.listing_url = importapartments.listing_url
AND accommodates >= 2 AND beds >= 2 AND city = 'Saint Kilda' And amenities ILIKE
'%Kitchen%' AND host_verifications ILIKE '%phone%'
AND (((price*2*2) + cleaning_fee) + (security_deposit*0.1)) < 5000 AND minimum_nights
<= 2
GROUP BY id, importapartments.name, (((price*2*2) + cleaning_fee) +
(security_deposit*0.1))::numeric:: money
ORDER BY total_price DESC;
```

	id ↕ name	↕ total_price ↕
1	30479523 ★ Beach at your Doorstep: Enjoy Summer in St Kilda	\$4,166.00
2	15737004 Premium property in great location	\$2,000.00
3	28227216 Luxury Vintage Melbourne Loft	\$1,807.00
4	18634383 Gorgeous Stkilda Apartments Sleeps 14 parking	\$1,805.00
5	12528535 Spacious Designer Apartment near Albert Park	\$1,600.00
6	4238287 Design House near Saint Kilda	\$1,550.00
7	16416097 Huge St Kilda Beach PentHouse Beauty	\$1,540.00
8	26338530 Stylish and Spacious Contemporary Apartment in St Kilda	\$1,500.00
9	22392766 Family and Large Group Parking WIFI Exc Location	\$1,436.00
10	17573198 2 Br Apartment fully furnished near albert park	\$1,364.00
11	23125153 Modern Apart Opposite Albert Park with Breakfast	\$1,360.00
12	22349738 St Kilda Lux Penthouse	\$1,305.80
13	23280244 **Totally Renovated in 2018. Beautiful Interiors**	\$1,300.00
14	25990845 **Newly Renovated. Cityview. Stunning Interiors**	\$1,256.00
15	18196590 ST KILDA - Grand Mansion Living.	\$1,245.00
16	22780316 Mediterranean Townhouse in the heart of it all!	\$1,226.00
17	11802035 Huge, stylish, central, 3 br apartment with pool.	\$1,212.00
18	20446300 cosy townhouse with nature light	\$1,200.00
19	20486536 Classy 3BR apt on iconic Chapel St!	\$1,196.00
20	23330390 From St Kilda to Melbourne City the Events Capital	\$1,156.00
21	28267864 Dream St Kilda Apartment	\$1,144.00
22	26816876 Brand new Apartment 2km to Luna park beach	\$1,126.00
23	21951978 Top-Floor Contemporary Designer Flat	\$1,119.00
24	22283414 ★Massive bright 3BR home by the bay★	\$1,113.00
25	29201424 Best View in Melbourne (St Kilda), SunRise&SunSet	\$1,111.00
26	15955013 Contemporary Apartment with Botanic Gardens view	\$1,099.00
27	16398747 Modern apartment in the heart of St Kilda	\$1,060.00
28	22677674 Large 2 BRM apt, close to restaurants St Kilda bch	\$1,059.00

En aquest cas els Id no són els mateixos perquè nosaltres vam crear Id's propis usant el tipus *SERIAL*.

5.8 Query 8

```
-- Creem taules intermitges per guardar la quantitat de verificacions i la
d'apartaments, després les utilitzarem en la recerca
DROP TABLE IF EXISTS numeroVerifications;
CREATE TABLE numeroVerifications(
    id_host INT,
    numeroVerifications int
);

DROP TABLE IF EXISTS numeroApartments;
CREATE TABLE numeroApartments(
    id_host INT,
    numeroApartments int
);

INSERT INTO numeroVerifications SELECT host.id_host, count(verifyes.id_verification)
FROM host , verifyes , verification
WHERE verifyes.id_host = host.id_host AND verification.id_verification =
verifyes.id_verification GROUP BY host.id_host;

INSERT INTO numeroApartments SELECT host.id_host, count(property.id_host) FROM host ,
property
WHERE host.id_host = property.id_host GROUP BY host.id_host;

-- Fem la recerca final
SELECT name_host AS name, SUM((1/price)*(1 + host_is_superhost:: INT)*(SELECT
numeroVerifications.numeroVerifications
FROM numeroVerifications WHERE numeroVerifications.id_host = h2.id_host )*
```

```
(SELECT numeroApartments.numeroApartments FROM numeroApartments WHERE
numeroApartments.id_host = h2.id_host)) AS score, count(h2.id_host) -- Aquí s'acaba
el select
FROM host AS h2, property AS p2 WHERE h2.id_host = p2.id_host AND price > 0 AND
name_host IS NOT NULL
GROUP BY h2.id_host
ORDER BY score DESC LIMIT 10;
```

	name	score
1	Valeria	1146.059229792009
2	Jared SSP	1105.9727272741097
3	Adji	947.276430303675

Comprovació

```
/* PER AQUESTA COMPROVACIÓ HEM HAGUT DE CREAR UNA TAULA INTERMITJA PER TAL DE PODER
CONTAR LES VERIFICACIONS
DE CADA HOST, UN COP TENIEM A LA TAULA UNA FILA PER CADA VERIFICACIÓ DEL HOST HO
PASSEM A UNA ALTRA FENT UN COUNT I AIXÍ OBTENIM
QUANTES VERIFICACIONS TÉ CADA HOST
*/
DROP TABLE IF EXISTS VerificationsComprovacioIntermig;
CREATE TABLE VerificationsComprovacioIntermig(
    id_host INT,
    verificacio VARCHAR(50)
);

INSERT INTO VerificationsComprovacioIntermig SELECT DISTINCT host_id,
regexp_split_to_table(host_verifications, ',') FROM importhosts;

DROP TABLE IF EXISTS numeroApartmentsComprovacio;
CREATE TABLE numeroApartmentsComprovacio(
    id_host INT,
    numeroApartments int
);

INSERT INTO numeroApartmentsComprovacio SELECT host_id, COUNT(host_id) FROM
importhosts GROUP BY host_id;
DROP TABLE IF EXISTS VerificationsComprovacio;
CREATE TABLE VerificationsComprovacio(
    id_host INT,
    verificacio INT
);

INSERT INTO VerificationsComprovacio SELECT id_host, COUNT(id_host) FROM
VerificationsComprovacioIntermig GROUP BY id_host;

DROP TABLE IF EXISTS VerificationsComprovacioIntermig;

SELECT host_name AS name, SUM((1/price)*(1 + host_is_superhost:: INT)*(SELECT
VerificationsComprovacio.verificacio
FROM VerificationsComprovacio WHERE VerificationsComprovacio.id_host = h2.host_id )*
```

```
(SELECT numeroApartmentsComprovacio.numeroApartments FROM numeroApartmentsComprovacio
WHERE numeroApartmentsComprovacio.id_host = h2.host_id)) AS score -- Aquí s'acaba el
select
FROM importhosts AS h2, importapartments AS p2 WHERE h2.listing_url = p2.listing_url
AND price > 0 AND host_name IS NOT NULL
GROUP BY h2.host_id, h2.host_name
ORDER BY score DESC LIMIT 3;
```

	name	score
1	Valeria	1146.059229792009
2	Jared SSP	1105.9727272741097
3	Adji	947.2764303036748

5.9 Query 9

```
DROP TABLE IF EXISTS PointComments;
CREATE TABLE PointComments(
    name_reviewer VARCHAR(255),
    id_reviewer INT,
    points int
);

-- Inserim els punts dels usuaris per comentaris més llargs de 100 caràcters, NO
agafem el id_reviewer = 49929207 perquè és l'usuari que ha escrit reviews falses
INSERT INTO PointComments
SELECT name_reviewer, reviewer.id_reviewer, count(reviewer.id_reviewer)*15
FROM reviewer, reviewercommentproperty
WHERE reviewer.id_reviewer = reviewercommentproperty.id_reviewer AND length(comment)
>= 100 AND reviewer.id_reviewer != 49929207
GROUP BY reviewer.id_reviewer;

-- Inserim els punts dels usuaris per comentaris més curts de 100 caràcters
INSERT INTO PointComments
SELECT name_reviewer, reviewer.id_reviewer, count(reviewer.id_reviewer)*10
FROM reviewer, reviewercommentproperty
WHERE reviewer.id_reviewer = reviewercommentproperty.id_reviewer AND length(comment)
< 100 AND reviewer.id_reviewer != 49929207
GROUP BY reviewer.id_reviewer;

-- SELECT FINAL
SELECT name_reviewer AS name, SUM(PointComments.points) AS points
FROM PointComments
GROUP BY name_reviewer, PointComments.id_reviewer
ORDER BY points DESC LIMIT 10;
```

	name	points
1	Laurie	1520
2	Jessica	780
3	Michael	715
4	Thanh	540
5	Andrew	520
6	Suzanne	475
7	Anders	475
8	John	435
9	Jon	425
10	Andrew	425

Comprovació

```

DROP TABLE IF EXISTS PointCommentsComprovacio;
CREATE TABLE PointCommentsComprovacio(
    name_reviewer VARCHAR(255),
    id_reviewer INT,
    points int
);

-- Inserim els punts dels usuaris per comentaris més llargs de 100 caràcters, NO
-- agafem el id_reviewer = 49929207 perquè és l'usuari que ha escrit reviews falses
INSERT INTO PointCommentsComprovacio
SELECT reviewer_name, reviewer_id, count(reviewer_id)*15
FROM importreviews
WHERE length(comments) >= 100 AND reviewer_id != 49929207
GROUP BY reviewer_id, reviewer_name;

-- Inserim els punts dels usuaris per comentaris més curts de 100 caràcters
INSERT INTO PointCommentsComprovacio
SELECT reviewer_name, reviewer_id, count(reviewer_id)*10
FROM importreviews
WHERE length(comments) < 100 AND reviewer_id != 49929207
GROUP BY reviewer_id, reviewer_name;

-- SELECT FINAL
SELECT name_reviewer AS name, SUM(PointCommentsComprovacio.points) AS points
FROM PointCommentsComprovacio
GROUP BY name_reviewer, PointCommentsComprovacio.id_reviewer
ORDER BY points DESC LIMIT 10;

```


	name	points
1	Laurie	1520
2	Jessica	780
3	Michael	715
4	Thanh	540
5	Andrew	520
6	Suzanne	475
7	Anders	475
8	John	435
9	Jon	425
10	Andrew	425

5.10 Query 10

Per fer la nostra query hem triat mostrar totes les ciutats acompanyades de la quantitat d'apartaments i el preu mig d'aquests. d'aquesta forma els usuaris podran veure en quina ciutat tenen una oferta d'apartaments més gran amb un preu mig que s'ajusti més al seu preessupost.els usuaris podran triar si anar a una ciutat on la oferta és molt gran però el preu és molt alt o anar a un lloc amb poca oferta però preus més ajustats.

```
SELECT name_city, count(city.id_city) AS num_apartments, AVG(price)::numeric::money
AS preu_mig FROM city, property WHERE city.id_city = property.id_city
GROUP BY city.id_city ORDER BY num_apartments DESC, preu_mig ASC;
```

	name_city	num_apartments	preu_mig
1	Melbourne	4318	\$152.56
2	Southbank	1240	\$187.20
3	South Yarra	885	\$158.92
4	Saint Kilda	806	\$144.08
5	Docklands	612	\$177.07
6	Richmond	571	\$159.56
7	Carlton	538	\$111.84
8	Brunswick	442	\$95.28
9	Fitzroy	383	\$153.83
10	North Melbourne	341	\$111.66
11	Elwood	334	\$147.64
12	Collingwood	297	\$146.11
13	South Melbourne	289	\$205.31
14	Northcote	284	\$113.47
15	West Melbourne	264	\$129.83
16	Prahran	262	\$152.41
17	Abbotsford	259	\$130.37
18	Port Melbourne	258	\$200.32
19	Fitzroy North	238	\$118.08
20	St. Kilda	202	\$152.32
21	Windsor	195	\$142.02
22	East Melbourne	193	\$201.62
23	Point Cook	191	\$118.77
24	Hawthorn	184	\$187.79
25	Box Hill	182	\$117.73
26	Saint Kilda East	182	\$118.31
27	Brunswick East	181	\$149.43
28	Glen Waverley	174	\$126.60

Comprovació

```
SELECT city, count(city) AS num_apartments, AVG(price)::numeric::money AS preu_mig
FROM importapartments
GROUP BY city ORDER BY num_apartments DESC, preu_mig ASC;
```

	city	num_apartments	preu_mig
1	Melbourne	4318	\$152.56
2	Southbank	1240	\$187.20
3	South Yarra	885	\$158.92
4	Saint Kilda	806	\$144.08
5	Docklands	612	\$177.07
6	Richmond	571	\$159.56
7	Carlton	538	\$111.84
8	Brunswick	442	\$95.28
9	Fitzroy	383	\$153.83
10	North Melbourne	341	\$111.66
11	Elwood	334	\$147.64
12	Collingwood	297	\$146.11
13	South Melbourne	289	\$205.31
14	Northcote	284	\$113.47
15	West Melbourne	264	\$129.83
16	Prahran	262	\$152.41
17	Abbotsford	259	\$130.37
18	Port Melbourne	258	\$200.32
19	Fitzroy North	238	\$118.08
20	St. Kilda	202	\$152.32
21	Windsor	195	\$142.02
22	East Melbourne	193	\$201.62
23	Point Cook	191	\$118.77
24	Hawthorn	184	\$187.79
25	Box Hill	182	\$117.73
26	Saint Kilda East	182	\$118.31
27	Brunswick East	181	\$149.43
28	Glen Waverley	174	\$126.60

6.CONCLUSIONS

Per acabar ens agradaria fer un petit repàs del procés que ens ha portat des de la primera lectura del enunciat fins a la finalització de la practica així com comentar el que ens ha suposat cada un d'aquests passos.

Just després de la primera lectura del enunciat i d'una primera observació de les dades ens van aparèixer diversos dubtes. En primer lloc contràriament al que estem acostumats al enunciat hi havia poca informació que ens ajudes a construir el model relacional, tota aquesta s'havia d'obtenir observant i entenent els fitxers “.csv” adjuntats.

Destacaven, d'altre costat, en aquests fitxers, dos camps amb un format peculiar “Amenities” i “host_verification”, que no només no eren atòmics i per tant calia modificar-los, sinó que no enteníem ben be com fer-ho. Tot i això després d'una segona llegida del

enunciat i una mica d'investigació sobre l'eina "*egexp_split_to_table*" no vam tardar gestionar aquest format i entendre com emmagatzemar-lo adientment.

Un cop finalitzat el model conceptual, La implementació dels models relacional i físic no va tenir cap complicació més enllà de destacar algun petit error de disseny que ens va obligar a revisar el primer model. La dificultat real de la pràctica ha aparegut a l'hora de copiar les dades de les taules d'importació a les dades finals. Aquest procés ens ha portat a fer encara més redissenys de la bases de dades ja que el nostre primer model no era òptim.

A la vegada, ens hem donat compte de l'existència de camps amb formats o idiomes diferents però amb un mateix significat que provocaven tenir entrades duplicades. Un exemple d'això són els noms escrits en "Hàanzi" o els noms d'estat Abreviats ("VIC" en comptes de "Victoria").

Quant a la fase 2 de la pràctica, un cop llegides les diferents quèries que ens demanava, vam començar a resoldre les primeres quèries sense gaires complicacions, fet que ens confirmava que la importació de la base de dades havia estat la correcta. Quan ja portàvem unes quantes quèries resoltes ens van sorgir alguns problemes com que el nostre resultat s'assemblava molt al de l'enunciat però hi havia alguns camps que canviaven una mica, en aquest moment vam enviar un correu als becaris i amb la seva ajuda vam descobrir que ho teníem bé, però que en lloc de comparar els nostres resultats amb els de l'enunciat ho havíem de fer amb les taules d'importació.

Amb la query que vam tenir més problemes va ser la 8, la nostra primera solució va ser crear dues taules, una amb els punts dels missatges de menys de 100 caràcters i una altra amb els punts dels missatges de més de 100 caràcters per a continuació relacionar-les amb l'Id de l'usuari i sumar els punts. Quan la teníem feta ens vam adonar que aquesta solució no era correcta, ja que la query va arribar a estar 20 minuts executant-se sense donar resultat, en aquest moment ens vam adonar que el millor era posar-ho tot en una mateixa taula i a partir d'aquesta sumar els punts.

Finalment en perspectiva considerem que la nostra base de dades ha assolit com a mínim les exigències de qualitat que ens vam imposar en un primer moment, intentant que aquesta no només estigui normalitzada sinó que sigui també tant eficient com les nostres capacitats permetin.