



UNIVERSITAT OBERTA DE CATALUNYA (UOC)  
MÀSTER UNIVERSITARI EN CIÈNCIA DE DADES (*Data Science*)

## TREBALL FINAL DE MÀSTER

ÀREA: BIG DATA EN SISTEMA D'INFORMACIÓ GEOGRÀFICA

**Visibilitzant la gentrificació a Barcelona. Evolució,  
anàlisi i detecció de zones en risc.**

---

Autor: Oriol Martínez García

Tutora: Anna Muñoz Bolas

Professor: Albert Solé Ribalta

---

Barcelona, 11 de juny de 2024



# Crèdits



Aquesta obra està subjecta a una llicència de Reconeixement - NoComercial - SenseObra-Derivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/).



# FITXA DEL TREBALL FINAL

Títol del treball:	Visibilitzant la gentrificació a Barcelona. Evolució, anàlisi i detecció de zones en risc
Nom de l'autor:	Oriol Martínez García
Nom del col·laborador/a docent:	Anna Muñoz Bolas
Nom del PRA:	Albert Solé Ribalta
Data d'entrega (mm/aaaa):	06/2024
Titulació o programa:	Màster Universitari en Ciència de Dades
Àrea del Treball Final:	Data Analysis i Big Data
Idioma del treball:	Català
Paraules clau	Gentrificació, Big data, Barcelona



# Dedicatòria

A la meva parella i als meus pares pel suport brindat durant tot el camí acadèmic.





# Agraïments

Vull agrair a la meva tutora Anna Muñoz Bolas per la seva labor i ajuda.



# Abstract

The effect of gentrification in cities is increasingly present and worrying for society. Therefore, it is important to identify which are the factors that favor this effect and to be able to study them. Good identification can allow regulations to be applied or actions to reduce the negative effects that are being observed of higher prices and population displacement. The study aims to do all this in a large city like Barcelona where there are already areas that have suffered gentrification. The study finally hopes to be able to see the evolution and factors associated with visualizations.

**Keywords:** Gentrification, Barcelona, R, Clusterization, Master Thesis



# Resum

L'efecte de la gentrificació a les ciutats és cada cop més present i preocupant per a la societat. Per això es troba important identificar quins són els factors que afavoreixen aquest efecte i poder-los estudiar. Una bona identificació pot permetre aplicar regulacions o fer accions que disminueixin els efectes negatius que s'estan observant d'encariment de preus i de desplaçament de població. L'estudi pretèn fer tot això en una gran ciutat com Barcelona on ja hi ha zones que han patit gentrificació. L'estudi finalment espera poder veure l'evolució i els factors associats amb visualitzacions.

**Paraules clau:** Gentrificació, Barcelona, R, Clusterització, Treball Final Màster



# Índex

<b>Abstract</b>	<b>ix</b>
<b>Resum</b>	<b>xi</b>
<b>Índex</b>	<b>xiii</b>
<b>Llistat de figures</b>	<b>xv</b>
<b>Llistat de taules</b>	<b>1</b>
<b>1 Introducció</b>	<b>3</b>
1 Context i justificació del treball . . . . .	4
2 Objectius . . . . .	5
2.1 Objectius generals . . . . .	5
2.2 Objectius específics . . . . .	5
3 Impacte en sostenibilitat, diversitat i desafiament ètic i social . . . . .	5
4 Enfocament i mètode seguit . . . . .	6
5 Planificació . . . . .	6
5.1 Resum dels productes del projecte . . . . .	8
5.2 Breu descripció dels altres capítols de la memòria . . . . .	8
<b>2 Estat de l'art</b>	<b>11</b>
1 Indicadors estudiats per a la identificació de la gentrificació . . . . .	11
2 Mètodes d'anàlisi de les variables . . . . .	14
3 Plantejament i desenvolupament tecnològic . . . . .	15
<b>3 Materials i mètodes</b>	<b>17</b>
1 Recollida de dades . . . . .	17
2 Neteja de dades . . . . .	18
2.1 Tractament de valors NA . . . . .	19

---

3	Taules finals . . . . .	19
3.1	Contingut de les taules . . . . .	19
4	Anàlisi de les dades . . . . .	20
5	Representació gràfica . . . . .	21
6	Clusterització . . . . .	21
7	Model de decisió . . . . .	23
<b>4</b>	<b>Resultats</b>	<b>25</b>
1	Les dades . . . . .	25
2	Anàlisi gràfic . . . . .	25
3	Arbre de decisió . . . . .	32
4	Creació d'una visualització del procés de gentrificació . . . . .	39
<b>5</b>	<b>Conclusions i treballs futurs</b>	<b>43</b>
1	Conclusions . . . . .	43
2	Assoliment objectius generals . . . . .	44
3	Assoliment objectius específics . . . . .	44
4	Planificació i metodologia . . . . .	44
5	Treballs futurs . . . . .	45
	<b>Glossari</b>	<b>47</b>
	<b>Bibliografia</b>	<b>48</b>
	<b>Annexos</b>	<b>53</b>



# Índex de figures

1.1	Encariment del lloguer. Font: Incasol - <a href="https://beteve.cat/economia/pujada-lloguer-barcelona-triplica-ipc-10-anys-mapa/">https://beteve.cat/economia/pujada-lloguer-barcelona-triplica-ipc-10-anys-mapa/</a> . . . . .	4
1.2	Planificació del projecte del TFM. Elaboració pròpia . . . . .	8
2.1	Característiques socio espacials de Nanjing. Font: [23] . . . . .	12
2.2	Taula de variables de l'estudi de Brisbane. Font: Bottero, 2022 [8] . . . . .	13
2.3	Percepció de la gentrificació per zones. Font: Antunes, 2020 [6] . . . . .	14
2.4	Comparativa nombre Airbnb 2016 vs 2020. (A) Mapa de codis postals (B) Airbnb gener 2016 (C) Airbnb febrer 2020. Font: [19] . . . . .	15
3.1	Mètode del colze per a l'elecció del valor de k. . . . .	22
3.2	Resultats de l'aplicació del pairwiseAdonis als resultats de la clusterització. . . . .	22
4.1	Representació de totes les variables en El Gòtic. Font: Elaboració pròpia . . . . .	26
4.2	Evolució del preu del lloguer en El Gòtic. Font: Elaboració pròpia . . . . .	26
4.3	Evolució preu del lloguer i nombre d'habitatges d'ús turístic en El Gòtic. Font: Elaboració pròpia. . . . .	28
4.4	Evolució preu del lloguer i nombre d'habitatges d'ús turístic a la Sagrada Família. Font: Elaboració pròpia. . . . .	28
4.5	Comparació entre l'evolució de la població espanyola i la població europea. Font: Elaboració pròpia. . . . .	29
4.6	Comparativa evolució de la població espanyola segons el nivell d'estudis. Font: Elaboració pròpia. . . . .	29
4.7	Comparativa evolució de la població europea segons el nivell d'estudis. Font: Elaboració pròpia. . . . .	31
4.8	Clusterització mitjançant mètode kmeans (k=3). Font: Elaboració pròpia. . . . .	31
4.9	Gentrificació de Barcelona segons l'estudi realitzat per la Direcció de Serveis de Gènere i Polítiques del Temps. Font: Direcció de Serveis de Gènere i Polítiques del Temps [9] . . . . .	33

4.10	Fragment dels resultats de la clusterització per kmeans. Font: Elaboració pròpia.	34
4.11	UMAP sobre les dades amb assignació de colors per cluster. Font: Elaboració pròpia.	35
4.12	Pes de les variables en la clusterització. Font: Elaboració pròpia.	36
4.13	Visualització del procés de gentrificació a Barcelona en 4 anys diferents a nivell de barri (1997, 2005, 2014, 2023). Font: Elaboració pròpia	37
4.14	Visualització del procés de gentrificació a Barcelona en 4 anys diferents a nivell de districte (1997, 2005, 2014, 2023). Font: Elaboració pròpia	38
4.15	Mapa de calor de la correlació entre les variables utilitzades en l'estudi. Font: Elaboració pròpia.	39
4.16	Arbre de decisió elaborat a partir de les dades de clusterització. Font: Elaboració pròpia	40
4.17	Resultats de l'arbre de decisió.	40
4.18	Visualització de la interfície de l'aplicació shiny per veure l'evolució de les variables i del procés de gentrificació a Barcelona. Font: Elaboració pròpia	41
5.1	Taula amb dades de l'edat de la població de Barcelona diferenciat per lloc de procedència del 1997 al 2023	54
5.2	Taula amb dades de la formació de la població de Barcelona diferenciat per lloc de procedència del 1997 al 2023.	55
5.3	Taula amb dades del nombre d'habitatges d'ús turístic a Barcelona del 2014 al 2023.	56
5.4	Taula amb dades de la població de Barcelona del 2014 al 2023.	57
5.5	Taula amb dades del preu mitjà del lloguer (en euros) a Barcelona del 2000 al 2023.	58

# Índex de taules

4.1	Dades del districte de Ciutat Vella per a la taula de preus de lloguer. . . . .	27
4.2	Dades del barri del Raval per a la taula de preus de lloguer. . . . .	30



# Capítol 1

## Introducció

Hi ha una paraula que ha agafat força aquests darrers anys, la gentrificació. No és una paraula innovadora perquè aquesta fou encunyada el 1964 per la sociòloga Ruth Glass en la introducció del llibre “London: Aspects of change”. I és que, com bé indica l’autora, un cop començada la gentrificació aquesta s’expandirà i ocuparà tot o gairebé tot l’espai de la classe treballadora fins a canviar l’aspecte social del barri [13].

A partir d’aquí diversos grups han volgut donar una explicació al procés i trobar quines variables poden ser identificadors d’un procés de gentrificació. Així alguns estudis han pres com variables el sexe, l’edat, nivell d’estudis, nacionalitat i composició familiar, com seria el cas de l’estudi de López-Gay al barri de Sant Antoni [16], o variables com el percentatge de població, la mitjana d’ingressos, el valor de la propietat i l’ocupació de les persones en l’estudi de Voorhees en la ciutat de Chicago [21].

Per tant, la situació de gentrificació no s’ha quedat a l’estudi de Londres, des d’on es va forjar la paraula, sinó que els seus efectes són visibles en altres ciutats del món. I de manera més propera en punts importants de la geografia del país, com ara Madrid o Barcelona. I és que a Barcelona l’efecte de la gentrificació es va poder notar l’any 1991. Aquest any 78 dels 380 barris de l’àrea metropolitana tenien condicions idònies per a la gentrificació, i 18 d’ells a Barcelona. Un exemple la Vila Olímpica (Jocs Olímpics, 1992) i Diagonal Mar i Front Marítim (Fòrum de les cultures, 2004). Aquestes zones van passar a ser zones residencials d’alt nivell de renda. I s’hi uneixen altres barris com el Gòtic, Santa Caterina, la Ribera, la Barceloneta o el Bon Pastor, els quals s’identifica que han patit o estan patint gentrificació, en l’anàlisi que realitza Milagros Pérez. [18].

En el Gòtic, per exemple, es troba que només un 15% de la població d’entre 25-34 anys són de Barcelona. Aquest percentatge fa entreveure que una part de la població (la nascuda a Barcelona) ha sigut desplaçada per una altra de fora, o que aquesta zona ha deixat de ser interessant per a la gent local. Aquest desplaçament de població, que ja descrivia Ruth Glass,

porta associat un augment dels lloguers de la zona. En l'article de Marcos Arizmendi, publicat a Betevé, s'indica que Barcelona ha patit un augment de lloguers (Figura 1.1) d'un 63% des del 2014 (lloguer mitjà de 688.23 €) el 2023 (lloguer mitjà de 1123.56 €). [7]

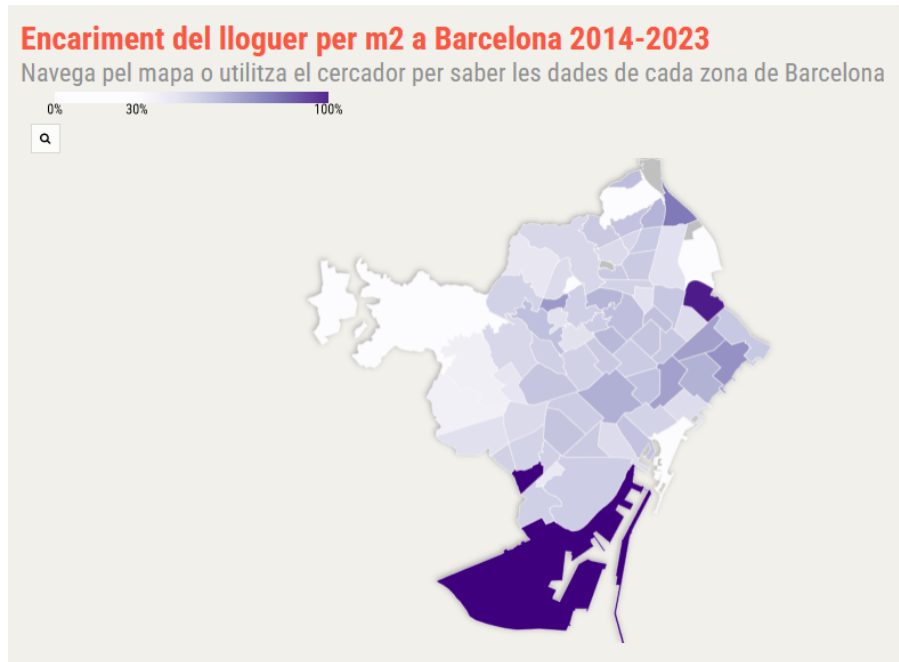


Figura 1.1: Encariment del lloguer. Font: Incasol - <https://beteve.cat/economia/pujada-lloguer-barcelona-triplica-ipc-10-anys-mapa/>

És a partir del conjunt d'efectes negatius de la gentrificació que sorgeix l'interès per l'estudi i trobar les raons i variables que hi tenen més pes.

## 1 Context i justificació del treball

El treball plantejat neix de l'interès a intentar anar un pas més enllà o poder ampliar els coneixements que hi ha actualment sobre la gentrificació. Per un ciutadà en el seu dia a dia és important saber que pot viure en una zona, que pot arribar a pagar un lloguer assequible, que pot trobar un habitatge digne, que la zona disposa dels serveis adequats, que hi ha una certa activitat social, entre d'altres. Tot això es veu afectat de forma negativa amb la gentrificació i el desplaçament de la població. Per tant, es vol abordar aquest problema d'una manera quantitativa i poder trobar els factors que caracteritzen les zones ja gentrificades. I veure quines zones estan en risc de patir-la.

En el treball es tractarà l'àrea de Barcelona, per ser una zona propera i, per tant, generar un interès de partida més gran. S'espera que al llarg del treball es puguin obtenir visualitzacions diverses que permetin d'una manera senzilla i clara entendre i observar els factors estudiats.

## 2 Objectius

En aquest treball es voldran assolir els següents objectius generals i específics.

### 2.1 Objectius generals

- Capacitat de recerca de l'estat de l'art sobre la gentrificació i els factors associats.
- Extreure i analitzar informació de les dades a partir de mètodes de big data.
- Elaborar una memòria del treball fet.
- Defensar els resultats obtinguts.

### 2.2 Objectius específics

- Obtenir, transformar i integrar les dades de diferents fonts.
- Fer una anàlisi estadística de les dades.
- Extreure informació de l'anàlisi estadística.
- Crear visualitzacions interactives dels factors relacionats amb la gentrificació.
- Desenvolupar un model capaç d'identificar barris ja gentrificats o en procés de gentrificació.

## 3 Impacte en sostenibilitat, diversitat i desafiament ètic i social

**Sostenibilitat** En el desenvolupament del treball no es preveu un tractament en específic de l'àmbit de la sostenibilitat. En el cas concret del treball sobre gentrificació no es preveu fer un estudi sobre consum d'energia, contaminació, esgotament de matèria primera i com es pot evitar, reduir o millorar.

**Comportament ètic i responsabilitat social** El projecte espera aconseguir un resultat final amb visualitzacions i conclusions que puguin tenir cert impacte positiu en l'àmbit social. Amb el projecte no es posa en perill cap lloc de treball sinó ser una eina més d'ajuda.

**Diversitat, gènere i drets humans** El projecte a desenvolupar pot tenir un petit impacte a l'hora de plantejar noves regulacions a partir de conèixer els factors que afecten la gentrificació. Es pot emmarcar dins de l'ODS 17 - Ciutats i comunitats sostenibles. L'ODS 17 busca que les ciutats estiguin preparades per una ràpida urbanització i reduir el creixement de zones marginals. L'augment del preu de l'habitatge, afavorit per la gentrificació, i el consegüent desplaçament de gent amb menys recursos alimenta el creixement de barris marginals.

## 4 Enfocament i mètode seguit

Per dur a terme el projecte caldrà cobrir diferents àrees. Una d'elles serà la consulta sobre informació en el tema. Donat que l'objectiu és trobar factors de gentrificació cal conèixer prèviament quins han sigut proposats per altres investigadors en la matèria i com han sigut tractats. Així com entendre algunes de les conclusions a les quals hagin pogut arribar i evitar caure en errors o anàlisis poc productives que ja s'hagin fet.

La segona àrea que caldrà abordar és la de les dades. L'obtenció de les dades es farà en gran part de bases de dades obertes. Donat que l'estudi es vol realitzar en la ciutat de Barcelona gran part de la informació la voldrem obtenir de llocs com INE (Instituto Nacional de Estadística), el portal de dades obertes de Barcelona o fins i tot del registre cadastral. De les dades que s'hagin obtingut caldrà fer una neteja i identificar quines són necessàries, quines no aporten valor o quines poden esdevenir interessants si es combinen entre elles. Quan les dades estiguin correctament treballades s'analitzaran de manera que permetin extreure índexs o factors interessants per fer-ne estudis estadístics. Que a la vegada seran útils per analitzar la informació obtinguda.

Finalment, es voldrà crear una o diverses visualitzacions que permetin veure l'evolució d'aquests factors al llarg del temps.

## 5 Planificació

El desenvolupament del projecte es dividirà en 5 parts coincidint amb les cinc entregues del projecte.

### Etapa 1: Definició del TFM

- Introducció i context del projecte: Es defineix l'abast del projecte i l'interès del seu desenvolupament.



- Definir objectius: Es defineixen els objectius a assolir durant i al final del projecte.
- Enfocament i metodologia: Es defineix la manera com es vol desenvolupar el projecte i que s'espera obtenir.

### **Etapa 2: Estat de l'art**

- Buscar articles: Recerca d'articles relacionats amb el tema treballat que ajudin a entendre com està l'estat de l'art actualment i quins progressos s'han fet fins ara.
- Buscar repositoris de dades: Recerca de dades adients a les necessitats de l'estudi i dels factors que es preveuen treballar.

### **Etapa 3: Disseny i implementació**

- Preparació entorn de treball: Preparació del programari necessari (R/RStudio).
- Tractament dades: Es tractaran les dades de manera que siguin coherents i estiguin totes integrades. S'agafaran les dades necessàries i es farà neteja de les que no siguin útils.
- Anàlisi de dades: Es durà a terme l'anàlisi estadística de les dades.
- Creació visualitzacions: Es durà a terme la creació de diverses visualitzacions a partir dels resultats obtinguts amb el treball previ de les dades.

### **Etapa 4: Redacció memòria**

- Redacció preliminar: Entrega de la redacció prèvia del projecte.
- Redacció final: Entrega de la redacció final del projecte.
- Presentació audiovisual: Preparació i entrega de la presentació audiovisual del projecte.

### **Etapa 5: Defensa TFM**

- Lliurament projecte: Entrega de tota la documentació del projecte.
- Defensa TFM: Defensa del projecte davant un tribunal.

Finalment, es mostra la planificació a partir del diagrama de Gantt (figura 1.2) elaborat.

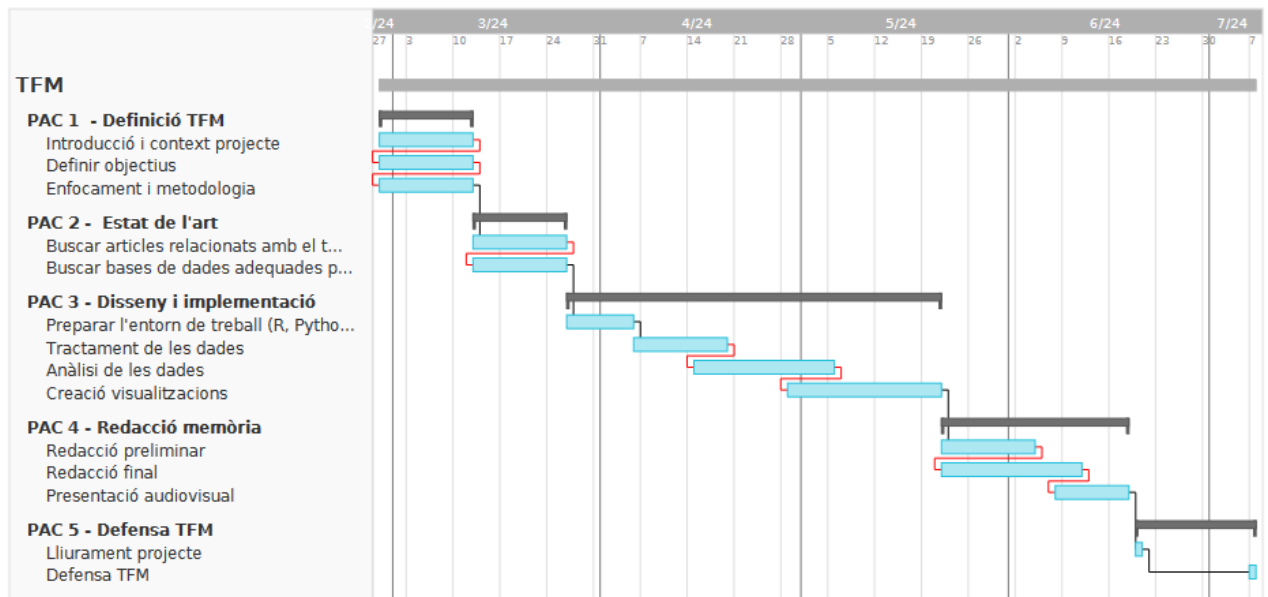


Figura 1.2: Planificació del projecte del TFM. Elaboració pròpia

## 5.1 Resum dels productes del projecte

En aquest projecte s'han creat els següents productes:

- Visualització interactiva gràfica de línies per barri (fitxer html i aplicació Shiny)
- Visualització interactiva mapa de Barcelona (aplicació Shiny)

## 5.2 Breu descripció dels altres capítols de la memòria

### Capítol 2: Estat de l'art

En el capítol 2 es duu a terme una anàlisi d'estudis relacionats amb la temàtica tractada. Permet amb la resta de capítols de disposar d'un punt de partida inicial sobre variables que poden ser d'utilitat al llarg del treball.

### Capítol 3: Materials i mètodes

En el capítol 3 s'indiquen quins materials s'han utilitzat per poder dur a terme el desenvolupament de les diferents visualitzacions. Es pot trobar el tractament de les dades i les seves casuístiques. I finalment l'elecció de les representacions gràfiques escollides i el model de clustertització escollit per a l'estudi.

### Capítol 4: Resultats

En el capítol 4 s'analitza amb més detall el format de les dades i el canvi que pateixen per poder ser utilitzades. Es fa una anàlisi amb més profunditat de les visualitzacions creades per poder extreure informació rellevant. I finalment es mostren els resultats de la clusterització i l'elaboració d'un model classificador.

**Capítol 5: Conclusions i treballs futurs**

En el capítol 5 es presenten les conclusions extretes de tot el treball. S'associen els resultats amb els objectius establerts i es plantegen diverses propostes de continuïtat del treball.



# Capítol 2

## Estat de l'art

### 1 Indicadors estudiats per a la identificació de la gentrificació

Al llarg dels anys s'han anat desenvolupant diferents estudis sobre la gentrificació. A partir d'estudis qualitatius o quantitatius els investigadors han provat de trobar indicadors o índexs que els permetin descriure i identificar quan, on i com es produeix aquest efecte gentrificador. Alguns identificadors utilitzats per Wu (2014) en el seu estudi de la variació socio espacial de Nanjing són els trets comuns en la demografia, les característiques d'ocupació de la població i l'habitatge [23]. L'estudi aprofita les dades del cens de la població de Nanjing amb informació espacial i temporal. D'entre les dades del cens se'n seleccionen 36 d'interès com el percentatge de població en un cert rang d'edat, el percentatge de transport, el percentatge d'estat socioeconòmic, la mitjana de preus, entre d'altres. Per reduir la dimensionalitat es fa ús d'una PCA que permet trobar 8 components entre les quals algunes es poden utilitzar per interpretar l'estructura socio espacial de les ciutats estudiades (Guangzhou i Nanjing). Finalment, aplicant una clusterització a les seves dades els permet crear una visualització de l'estructura social, com es pot veure en la figura 2.1.

En la mateixa línia de l'estudi de la gentrificació com la combinació de diverses variables hi ha un estudi sobre una eina per a les àrees metropolitanes del Canadà. L'estudi de Firth (2021) [12] incorpora en diversos models variables d'estudi com l'estat de salut, la salut mental, el crim, la puntuació sobre el crèdit i característiques de l'espai, com poden ser les línies de tren i els camins de bicicletes. És interessant destacar com els diferents models a partir de les dades poden identificar més o menys zones *gentrificades*. Amb aquests resultats han desenvolupat una eina que permet l'observació dels models per una millor comparació i facilitar d'aquesta manera possibles actuacions per part de l'administració.

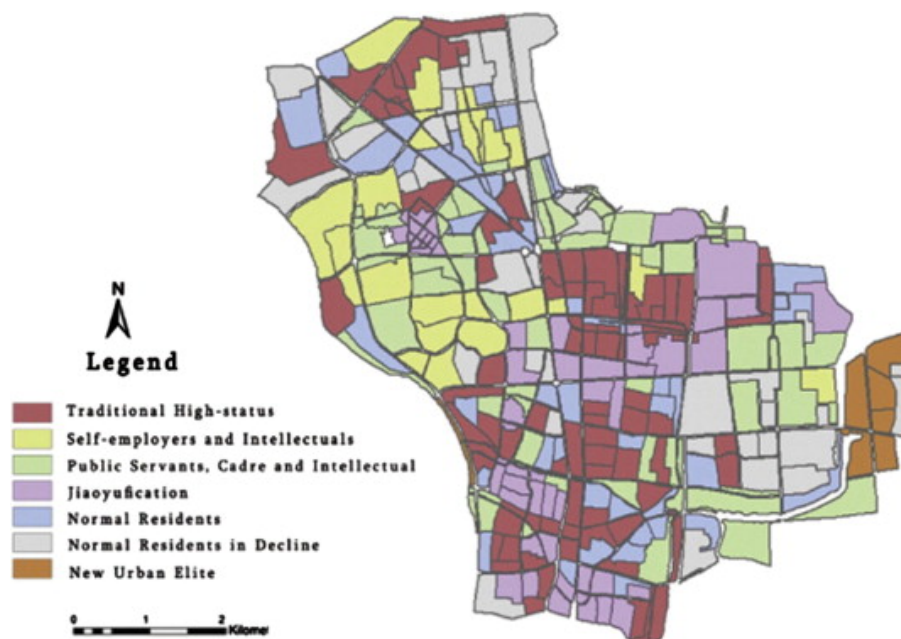


Figura 2.1: Característiques socio espacials de Nanjing. Font: [23]

Articles com el de Bottero (2022) [8] i Wolch (2014) [22] han estudiat el procés de gentrificació a partir de l'anàlisi dels espais verds. Per una banda, en l'estudi de Bottero s'ha pogut trobar en l'estudi de la ciutat de Brisbane que el preu augmenta en zones amb parcs recreatius mentre hi ha un petit descens (1 %) dels habitatges propers a parcs esportius. En aquest cas s'ha tingut en compte el preu i les característiques dels habitatges i la distància als parcs, com es pot veure en la figura 2.2. Per la seva banda, l'estudi de Wolch se centra en com la millora dels espais verds d'una ciutat té repercussió en l'augment dels preus de l'habitatge i, per tant, es produeix un desplaçament de gent de baix nivell econòmic. Aquest efecte es podria definir com a gentrificació verda (green gentrification). I s'ha observat que certs grans projectes de parcs tenen com a objectiu l'increment del valor del terreny buscant així un efecte gentrificador.

No només la millora dels espais verds és un aspecte analitzat per estudiar la gentrificació. L'article d'Ellen (2019) [10] busca trobar si la reducció de la criminalitat en una zona facilita o propicia un efecte gentrificador. S'arriba a observar que les zones que mostren un descens més marcat de la delinqüència tenen un augment important de compradors amb nivell econòmic alt. Tot i que el model que proposen pot ser sensible al tipus de dades que s'afegeixin a l'estudi sí que es constata que la reducció de la criminalitat pot tenir a llarg termini implicació en el procés gentrificador.

Altres indicadors que també s'han estudiat per identificar processos de gentrificació són la bretxa de lloguer (Antunes, 2020) [6], l'aparició d'Airbnb (Rabieidastjerdi, 2022) [19] o la variació del lloguer i el desplaçament i formació de la gent (López-Gay, 2022) [16].

Variables		Unit	Min	Max	Mean	Standard Deviation
<b>Apartment characteristics</b>	Price	[AUS \$]	1099	2365,0000	840,893,355	740,052,483
	Bedroom	[number]	0	20	2643	1370
	Bathroom	[number]	0	16	1711	0896
	Car	[number]	0	17	1462	1019
<b>Distance</b>	Urban park	[meter]	3262	5789,337	282,086	168,731
	Recreational urban park	[meter]	3262	5789,337	310,515	186,042
	Sport urban park	[meter]	8841	6627,113	636,947	344,772

Figura 2.2: Taula de variables de l'estudi de Brisbane. Font: Bottero, 2022 [8]

La bretxa de lloguer que planteja l'estudi d'Antunes [6] indicaria que el procés gentrificador ve produït per un efecte de devaluació de capital, per exemple una zona en decadència durant molt de temps i seguida llavors d'una reurbanització i ampliació urbana. L'article [6] pren com a exemple Vallcarca on a causa d'aquesta bretxa el preu del lloguer va pujar un 32.7% entre els anys 2014-2017. L'estudi de forma qualitativa, a partir d'enquestes, acaba representant la percepció del veïnat sobre la gentrificació que pateix Vallcarca, i en la figura 2.3 es pot veure les zones que es perceben com a gentrificades per part de llogaters, propietaris, ocupes i govern.

Les rendes més baixes es veuen obligades a desplaçar-se en les zones que comencen a gentrificarse amb una pujada dels preus d'accés a l'habitatge. Però a més si pot sumar la falta d'habitatge en haver-hi cada cop més habitatges que passen a ser Airbnb. L'article de Rabieidastjerdi (2022) [19] busca trobar patrons espacials, a partir de dades d'Airbnb, fent ús de l'anàlisi de punts calents que els permetin trobar llocs amb un gran potencial de gentrificació per a la ciutat de Dublín. En el seu estudi es poden observar l'aplicació de diverses metodologies com el tractament qualitatiu i quantitativ de les dades, amb visualitzacions i una clusterització mitjançant DBSCAN. Tot i l'observació d'una gran aparició d'habitatges utilitzades per a estances de curta durada (figura 2.4) no arriben a la conclusió de què Airbnb sigui l'únic procés que propiciï la gentrificació en solitari. Per tant s'han de tenir en compte altres variables.

López-Gay (2022) [16] realitza un estudi de la gentrificació a partir de la variació del preu del lloguer i del desplaçament de la població segons la seva formació en el cas del barri de Sant Antoni. En aquest cas en particular troben que el preu del lloguer pateix un gran augment del preu del lloguer. A més el percentatge de població en règim en lloguer també augmenta,

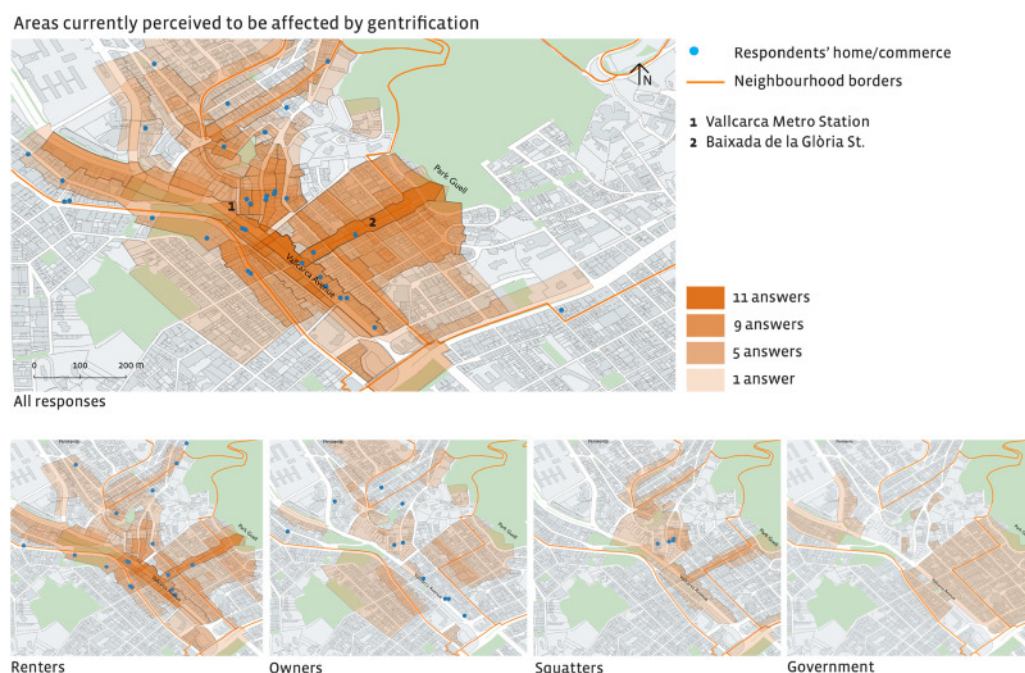


Figura 2.3: Percepció de la gentrificació per zones. Font: Antunes, 2020 [6]

passant d'un 36.5 % el 2011 al 58.3 % el 2017. El barri a més també experimenta en aquests anys un augment de la gent que arriba amb un nivell d'estudis superiors. Per donar més completitud a l'estudi s'afegeix una part d'anàlisi a partir d'entrevistes a les persones de la zona. És interessant veure com la percepció de què s'està produint un canvi pel canvi de teixit comercial i l'aparició d'habitatges turístics.

De l'anàlisi de diferents articles permet tenir una idea de l'estat de l'art i observar que s'ha intentat l'estudi del procés de gentrificació a partir de diferents tipus d'indicadors. Ara bé, no tots els indicadors com s'ha vist anteriorment tenen perquè tenir una incidència directa o ser el factor que defineix clarament si una zona pateix efecte gentrificador.

## 2 Mètodes d'anàlisi de les variables

Per extreure totes aquestes dades de manera general els investigadors han fet ús de repositoris de dades obertes, moltes d'elles del propi cens de la zona estudiada ([23], [12], [8], [22], [6], [16]), o extreures de pàgines web ([19]) sigui perquè estan accessibles o mitjançant processos de web crawling, com seria el cas de l'obtenció d'un llistat de preus i propietats de les estructures per l'estudi de Shanghai fet per LI (2019) [14].

Un cop en disposició de les dades alguns estudis (WU, 2014)[23] (Rabieidastjerdi, 2022) [19]) apliquen mètodes de PCA per reduir la gran quantitat de variables amb les que es treballa. I a partir d'aquí crear una clusterització dels resultats. Altres mètodes aplicats són la creació



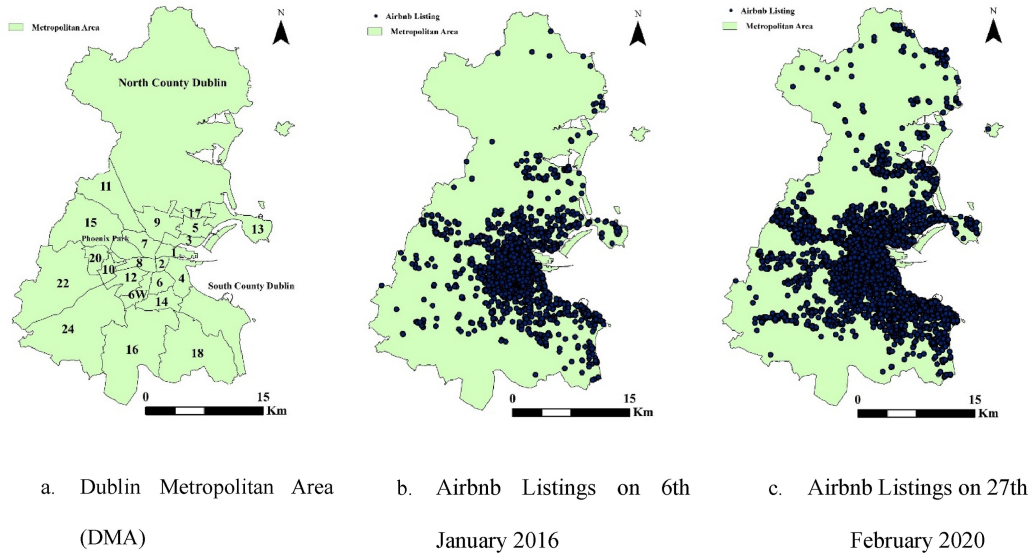


Figura 2.4: Comparativa nombre Airbnb 2016 vs 2020. (A) Mapa de codis postals (B) Airbnb gener 2016 (C) Airbnb febrer 2020. Font: [19]

d'índexs a partir de la combinació de dades (López-Gay, 2019) [17], també es fa ús de la clusterització mitjançant la tècnica de K-means (LIU, 2019) [15], altres apliquen mètodes de Random forests (Alejandro, 2019) [5] i també tècniques de machine learning (ML) (Reades, 2019) [20]. De forma general, segons l'anàlisi de diferents publicacions feta per Finio (2022) [11], la gentrificació s'estudia a partir d'una única variable juntament amb la seva variació en el temps.

### 3 Plantejament i desenvolupament tecnològic

A partir de l'anàlisi dels diferents estudis revisats es conclou que una part ([22], [10], [6]) acaben centrant-se en la variació temporal d'alguna variable determinada i en la representació individual de cada una d'elles sobre un mapa. La resta mostren variacions temporals o visualitzacions en mapes on la variable representada és algun índex o factor obtingut de la combinació de dades.

És interessant la proposta de creació d'aquest tipus d'estudi, però afegint una part interactiva que permeti una observació de canvis de forma més senzilla. La mateixa visualització podria permetre la navegació per diferents anys sigui de forma manual o automàtica. D'entre els mètodes d'estudi de les dades per extreure clústers es planteja l'ús del k-means, com del random forest o de dendogrames. La naturalesa pròpia de les dades i les variables disponibles afectarà en l'elecció del mètode més adequat.

L'estudi requerirà diferents eines tecnològiques. Per això tot el tractament de dades es

planteja fer en R Studio. Per a la visualització interactiva es proposa fer ús de la plataforma shiny i la seva implementació dins de R Studio.

# Capítol 3

## Materials i mètodes

Tot el procés de l'estudi es pot dividir en tres grans parts (obtenció i neteja de les dades, representació gràfica i clusterització i anàlisi). Les tres parts es duran a terme en l'IDE RStudio. El codi desenvolupat es realitza en llenguatge R i es pot trobar en el repositori de GitHub [https://github.com/Gamaor4/gentrification\\_bcn.TFM](https://github.com/Gamaor4/gentrification_bcn.TFM)

### 1 Recollida de dades

Les dades utilitzades han sigut extretes de la pàgina web de l'Ajuntament de Barcelona, **Barcelona Dades** [1].

Per la selecció de dades s'ha tingut en compte en primera instància els indicadors comentats en l'estat de l'art (demografia, preu lloguer, criminalitat, espais verds, habitatges turístics, nivell de formació, edat de la població...). Això ha servit de guia per fer la recerca de dades dins el portal de dades obertes de l'Ajuntament de Barcelona. El portal disposa d'una gran quantitat de taules dels indicadors comentats i que en un primer pas han sigut candidats per a ser analitzats. Un cop seleccionades les taules s'han filtrat per aquelles en què hi ha un seguiment anual i, com a mínim, informació a escala de districte. Un cop feta la recerca i seleccionades les taules candidates es fa una altra passada revisant en profunditat la informació continguda. En aquest pas s'acaben de filtrar les taules i s'afavoreixen les dades que contenen informació a escala de barri, districte i que són de seguiment anual.

La selecció final es queda en cinc taules. Aquestes són:

- Preu mitjà (€) del lloguer d'habitatges - <https://portaldades.ajuntament.barcelona.cat/ca/estad%C3%ADstiques/b37xv8wcjh>
- Edat mitjana de la població segons nacionalitat (Espanya - UE - Resta del món) - <https://portaldades.ajuntament.barcelona.cat/ca/estad%C3%ADstiques/te8dlihpxo>

- Població per nacionalitat (Espanya - UE - Resta del món) - <https://portaldades.ajuntament.barcelona.cat/ca/estad%C3%ADstiques/zkai27y5wu>
- Població de 16 anys i més per titulació acadèmica i nacionalitat (Espanya - UE - Resta del món) - <https://portaldades.ajuntament.barcelona.cat/ca/estad%C3%ADstiques/xnbrptprct>
- Nombre d'habitatges d'ús turístic - <https://portaldades.ajuntament.barcelona.cat/ca/estad%C3%ADstiques/z1wuyvykvf>)

Indicadors com els crims a Barcelona han quedat descartats al no trobar una taula adequada. Informació sobre el nombre d'habitatges que estan a la plataforma Airbnb s'ha substituït per una taula amb dades de tots els habitatges turístics de la ciutat. La selecció de taules ha buscat combinar diferents aspectes com la població i l'edat [23], el nombre d'habitatges d'ús turístic [19] i la formació i preu del lloguer [6]. A més a més, afegint un extra com és que moltes de les taules disposen de les dades separades per identificadors com Espanya, EU (Europa) i resta del món.

A més a més d'aquestes taules ha interessat també disposar d'un fitxer amb informació administrativa a escala de districtes i barris de Barcelona. La informació s'ha extret del portal open data de l'Ajuntament de Barcelona [2].

## 2 Neteja de dades

Tot i que les dades contenen informació a escala de barri, districte i anys aquesta es presenta de forma diferent en les diferents taules essent necessari una neteja i una aplicació de format unificat per a totes.

Actuacions que s'han dut a terme:

- S'han revisat i estandarditzat els noms utilitzats per a barris i districtes
- S'han modificat les columnes amb anys perquè només contingui l'any.
- S'han transformat les paraules a format ASCII per eliminar símbols problemàtics (accents)
- S'han transformat totes les paraules a minúscules per estandarditzar el format.
- S'han canviat les classes de tipus "char" a "num" o "int" en variables de tipus numèric.
- S'han modificat el nom d'algunes columnes per facilitar el seu ús.
- S'ha modificat l'ordre i presentació d'algunes columnes.

Per dur a terme totes aquestes accions en la neteja de les dades s’han utilitzat les llibreries *readr*, *tidyr*, *dplyr*, *stringr*, *stringi*.

## 2.1 Tractament de valors NA

Algunes de les taules presenten valors NA. En alguns molts casos per la transformació de classe “character” a “numeric”. L’aparició de valors NA no és uniforme en les diferents taules, de manera que l’eliminació d’aquestes pot produir una falta d’informació que afecti els càlculs posteriors.

La solució per la qual s’opta és introduir el valor de la mitjana, corresponent a cada variable estudiada, en els valors NA. S’assumeix que l’error que pot afegir en els processos posteriors serà menor que si es treballa sense valors i, per tant, amb una pèrdua important d’informació de diverses variables.

# 3 Taules finals

El tractament de les dades ha permès crear noves taules que faciliten la seva manipulació. Aquestes noves dades es guarden en fitxers nous amb el format adequat per treballar-les.

El nou format de les taules permet tractar amb les taules de forma individual i també facilita la seva combinació en una sola taula. Aquestes taules disposen de columnes per identificar el barri o districte a partir del nom propi del lloc, una columna per poder identificar el lloc a escala de barri o districte, una columna amb els anys i després la columna o columnes de les variables a estudiar.

## 3.1 Contingut de les taules

S’exposa de forma resumida el contingut de cada taula i el seu format.

- **Taula - Preu del lloguer:**

- territori [char]: noms propis de barris, districtes, municipis
- tipus de territori [char]: noms dels diferents nivells de territori
- anys [int]: any de quan s’ha pres la mesura
- preu mitjà lloguer [num]: valor del preu de lloguer

- **Taula - Edat mitjana per nacionalitat:**

- territori [char]: noms propis de barris, districtes, municipis

- tipus de territori [char]: noms dels diferents nivells de territori
  - anys [int]: any de quan s’ha pres la mesura
  - edat mitjana [num]: valor de l’edat mitjana. Hi ha una columna d’edat mitjana per cada nacionalitat (espanya - ue - resta del món)
- **Taula - Població de Barcelona:**
    - territori [char]: noms propis de barris, districtes, municipis
    - tipus de territori [char]: noms dels diferents nivells de territori
    - anys [int]: any de quan s’ha pres la mesura
    - població [num]: valor del nombre de població. Hi ha una columna de població per cada nacionalitat (espanya - ue - resta del món)
- **Taula - Estudis per nacionalitat:**
    - territori [char]: noms propis de barris, districtes, municipis
    - tipus de territori [char]: noms dels diferents nivells de territori
    - anys [int]: any de quan s’ha pres la mesura
    - nivell d’estudis [num]: valor del nombre de persones per nivell d’estudis. Hi ha una columna per cada nivell d’estudis amb cada nacionalitat (espanya - ue - resta del món)
- **Taula - Habitatges d’ús turístic:**
    - territori [char]: noms propis de barris, districtes, municipis
    - tipus de territori [char]: noms dels diferents nivells de territori
    - anys [int]: any de quan s’ha pres la mesura
    - habitatges us turístic [num]: valor del nombre d’habitatges d’us turístic.

## 4 Anàlisi de les dades

De les dades seleccionades es vol saber si hi ha relacions que poden ser interessants per investigar. Per això s’han unit totes les dades en una sola taula. En aquesta gran taula apareixen valors NA, ja que moltes taules no comparteixen els mateixos anys en els quals s’han anat obtenint dades. Per aquests valors, com bé s’ha comentat, se’ls hi assignarà el valor de la seva mitjana.

Com es volen trobar possibles correlacions s'ha fet ús de la funció de R *corr*. Abans s'han escalat tots els valors entre 0 i 1 per evitar que la discrepància entre dades molt grans i molt petites afecti els càlculs posteriors. Els valors de la correlació (figura 4.15) mostra que les variables seleccionades presenten interaccions interessants, com per exemple lloc de procedència i nivells d'estudis.

## 5 Representació gràfica

Mitjançant gràfiques es vol observar si les dades presenten tendències o canvis interessants. Per a la realització de les gràfiques s'ha fet ús de la llibreria *ggplot2* i per a fer-les interactives s'ha fet ús de la llibreria *plotly*. Per dotar de color s'ha fet ús de *RColorBrewer*. I altres llibreries útils per a la representació i que s'utilitzaran són *gridExtra* i *gplots*.

El tipus de visualització amb més ús al llarg de l'estudi és la gràfica de línies i el mapa. En aquest cas la gràfica de punts i línies en el seu format interactiu permet seleccionar les dades i comparar-les d'una forma senzilla. Amb la representació en mapa es vol fer més evident els canvis que pateixen els barris al llarg dels anys per cada variable.

## 6 Clusterització

Tot el tractament de dades i anàlisi previ porta a la part important del treball que és la clusterització. Es realitza també una anàlisi de les components principals mitjançant una PCA per veure si les clusteritzacions i la diferenciació creada és significativa. S'ha fet ús de les llibreries *cluster*, *factoextra*, *zoo*, *tibble*, *umap*, *nnet*, *vegan*, *pairwiseAdonis*, *stats*. I de les dades s'entrenarà un model d'arbre de decisió que permetrà poder identificar a partir de noves dades si una zona pot patir o pateix gentrificació. Per això s'ha fet ús de les llibreries *rpart*, *rpart.plot*, *caret*.

El mètode utilitzat per a la clusterització és el kmeans. Per poder aplicar el kmeans primer s'han seleccionat les dades en funció de si es treballa amb dades de districtes o barris. Un cop escollida la zona de territori a estudiar s'han escalat les dades entre 0 i 1. El kmeans necessita que s'especifiquin quants grups es volen en la variable "k". Per poder escollir un valor adequat de "k" s'ha fet ús del model del colze. A partir de la gràfica (figura 3.1) s'ha decidit que aquest valor sigui  $k = 3$ .

El valor seleccionat de k funciona bastant bé perquè permet separar les agrupacions per **no gentrificats, en procés de gentrificació i gentrificats**.

Per veure que el mètode ha funcionat s'aplica un UMAP sobre les dades. El UMAP també permet veure les dades agrupades en 2D. Per poder dotar de color el UMAP es faciliten els

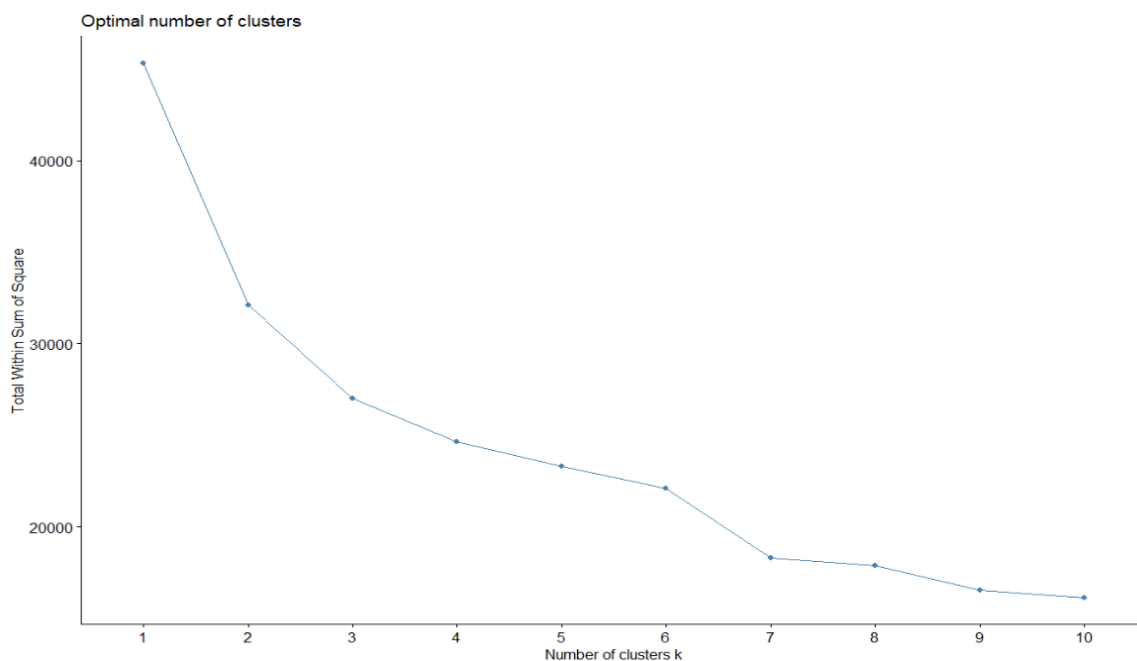


Figura 3.1: Mètode del colze per a l'elecció del valor de k.

pairs <chr>	Df <dbl>	SumsOfSqs <dbl>	F.Model <dbl>	R2 <dbl>	p.value <dbl>	p.adjusted <dbl>	sig <chr>
2 vs 1	1	14.49219	531.9078	0.3616187	0.001	0.003	*
2 vs 3	1	28.66814	731.0874	0.3046086	0.001	0.003	*
1 vs 3	1	40.78100	972.9639	0.4228506	0.001	0.003	*

3 rows

Figura 3.2: Resultats de l'aplicació del pairwiseAdonis als resultats de la clusterització.

valors del clúster generat en el kmeans. Això permet veure si el UMAP mostra una agrupació similar (figura 4.11).

De les agrupacions interessa saber quines variables han sigut les que han tingut més pes a l'hora de fer les agrupacions. A més, és interessant veure possibles canvis per cada component, per això s'aplica una PCA. La PCA o anàlisi de components principals permetrà discernir com han anat actuant cada una de les variables.

Finalment, s'aplica un mètode d'anàlisi per veure si hi ha diferències significatives entre els grups creats. Amb la idea de l'ANOVA es fa ús d'un mètode anomenat PERMANOVA. Per això s'utilitzen les llibreries *vegan* i *pairwiseAdonis*. Amb l'aplicació d'adonis es pot saber si els clústers són significativament diferents entre ells. Per acabar de completar que ho són entre tots ells s'aplica el pairwiseAdonis que indica com cada grup és significativament diferent l'un de l'altre. En el cas d'estudi es troba que els tres grups creats (1,2,3) tots són significament diferent entre ells (figura 3.2).



## 7 Model de decisió

Del resultat de saber que els grups creats en la clusterització són diferents entre ells es pot plantejar que la combinació de certs valors de les dades permet una classificació per a cadascun dels grups. A partir d'aquí s'ha elaborat un model d'arbre de decisió. Per aquest arbre s'ha pres la taula amb els valors originals de les dades, amb els valors NA corregits amb el valor de la mitjana, i s'ha creat un conjunt d'entrenament i un de test. Es crea el model predictiu del conjunt d'entrenament i s'aplica el conjunt de test. S'observa que el model arriba a un 95% de precisió.



# Capítol 4

## Resultats

### 1 Les dades

El tractament de les dades inicials té com a resultat una taula final com la del preu del lloguer que es pot veure en la taula 4.1 a escala de districte i en la taula 4.2 a escala de barri. La resta de taules es poden consultar en l'annex.

Es pot observar, en l'exemple de taula 4.1, com a escala de districte es disposen de valors des de l'any 2000 mentre que per als barris, taula 4.2, només es comença a registrar a partir de 2014. Els valors NA són els que després es corregiran per a les anàlisis.

### 2 Anàlisi gràfic

Tot i disposar de dades a escala de municipi i districte l'interès més gran està en veure com evolucionen i canvien les variables en cada barri. En la figura 4.1 es pot veure un exemple de representació de punts i línies del barri del Gòtic, on les variables s'han escalat de 0 a 1 per evitar l'efecte de valors molt grans respecte a la resta.

Amb totes les variables representades no es pot extreure res, però el fet de ser interactiva ajudarà a veure coses interessants. Tot i que només es mostra com a exemple el cas del Gòtic es poden trobar la resta de gràfiques en format HTML per descarregar en el repositori de Github [3], i també en format d'aplicació shiny interactiva [4]. Un dels efectes que s'associen amb la gentrificació és una pujada del preu en la zona [6] i això, per tant, es veu reflectit en el preu del lloguer (veure figura 4.2). Si se selecciona aquesta variable es pot observar com aquesta mostra un ascens destacat. Però també s'observa un descens que coincideix amb els anys 2019-2021 en el qual es va passar el període de COVID-19. Per tant, es pot veure el que es pot considerar com a efecte COVID en les dades.

Però a banda de l'augment de preus del lloguer és interessant observar com evoluciona el

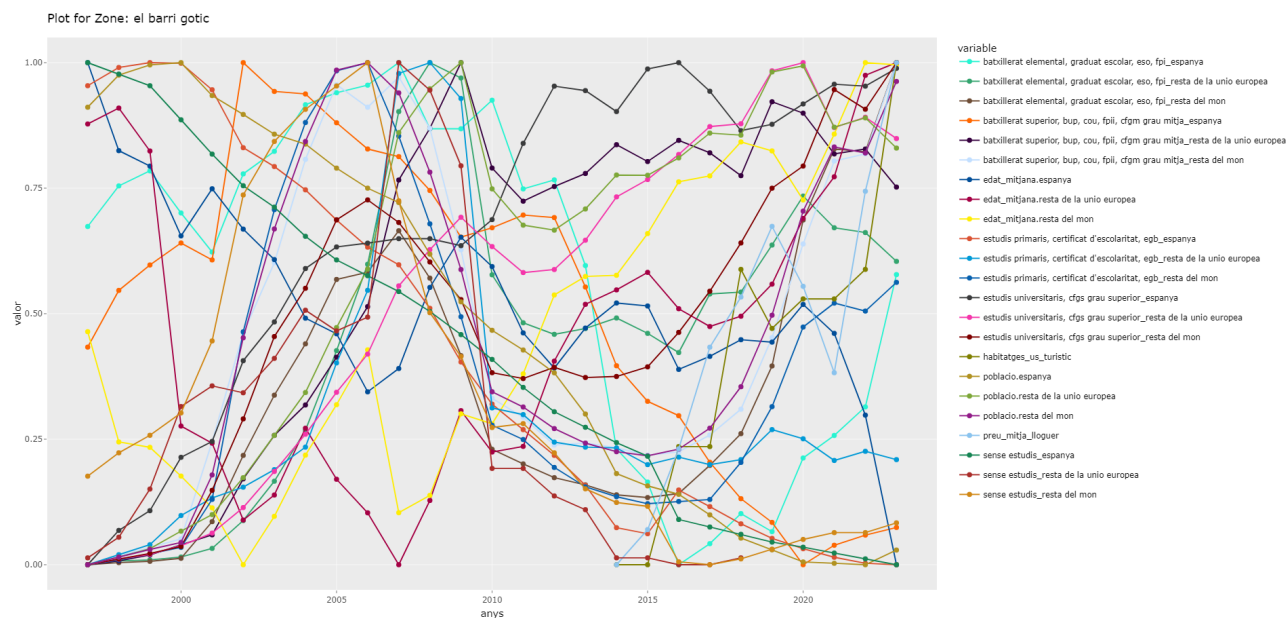


Figura 4.1: Representació de totes les variables en El Gòtic. Font: Elaboració pròpia

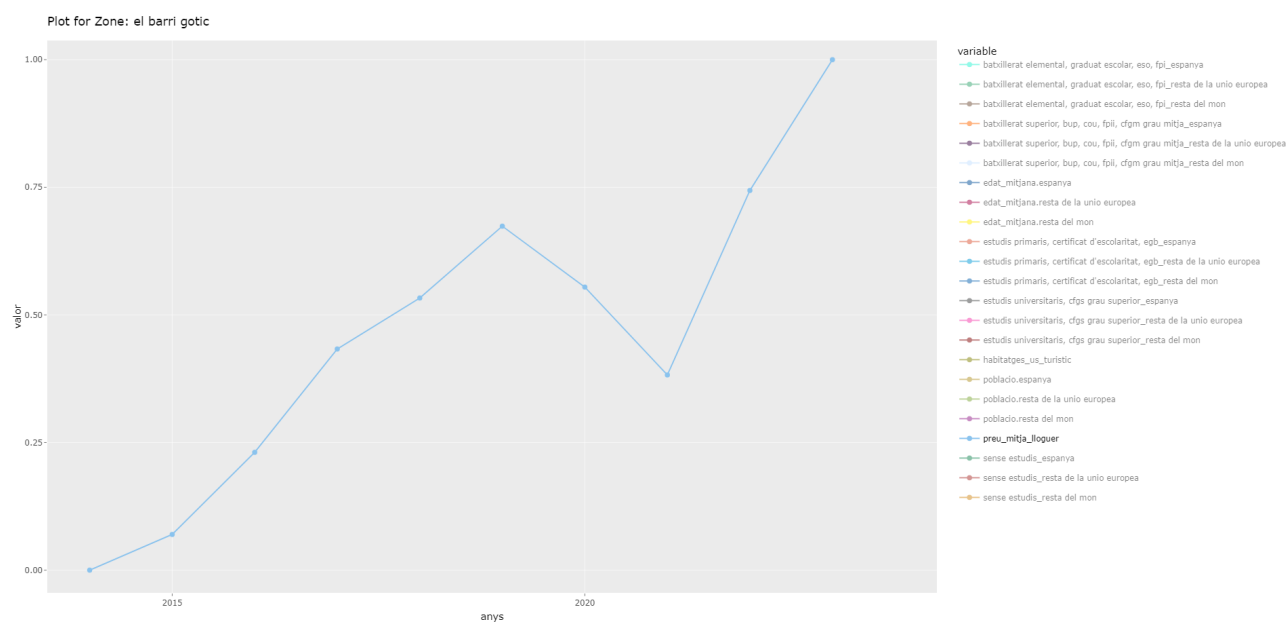


Figura 4.2: Evolució del preu del lloguer en El Gòtic. Font: Elaboració pròpia

territori	tipus de territori	anys	preu mitja lloguer
ciutat vella	districte	2000	325.191895
ciutat vella	districte	2001	350.9249161
ciutat vella	districte	2002	412.0531833
ciutat vella	districte	2003	443.1504516
ciutat vella	districte	2004	469.8563702
ciutat vella	districte	2005	539.2881261
ciutat vella	districte	2006	597.5975604
ciutat vella	districte	2007	660.9569614
ciutat vella	districte	2008	729.5869728
ciutat vella	districte	2009	715.6755064
ciutat vella	districte	2010	678.7898241
ciutat vella	districte	2011	663.0184923
ciutat vella	districte	2012	644.1452034
ciutat vella	districte	2013	618.6695989
ciutat vella	districte	2014	639.1097631
ciutat vella	districte	2015	675.7432
ciutat vella	districte	2016	739.2322591
ciutat vella	districte	2017	828.7861342
ciutat vella	districte	2018	914.041285
ciutat vella	districte	2019	945.5301408
ciutat vella	districte	2020	913.664714
ciutat vella	districte	2021	846.4039719
ciutat vella	districte	2022	1008.350958
ciutat vella	districte	2023	1079.006385

Taula 4.1: Dades del districte de Ciutat Vella per a la taula de preus de lloguer.

nombre d'habitatges turístics. Crida l'atenció com la corba de variació dels habitatges turístics fa un cert seguiment de la corba dels preus del lloguer (vegeu figura 4.3).

Tot i l'interès que desperta que la corba del nombre d'habitatges turístics sembla que segueixi la del preu de lloguer, aquesta tendència no es repeteix en la resta de barris. En aquest cas es compara amb un altre barri (figura 4.4), la Sagrada Família, també classificat com a gentrificat que no presenta aquesta dinàmica.

Un altre efecte comentat sobre la gentrificació és el desplaçament de la gent del territori per gent amb un nivell educatiu superior [6] i amb un poder adquisitiu més gran. Es pot observar, en la figura 4.5, com la població d'El Gòtic disminueix al llarg dels anys en pros de la gent que prové de la Unió Europea.

Però quin tipus de gent d'Espanya és la que es veu desplaçada. Es pot veure en la figura 4.6 com la població espanyola que disminueix és la que presenta que no té estudis o estudis primaris, i, en canvi, es pot veure com augmenta la gent amb formació universitària. I així i tot el total és que la població espanyola disminueix de la zona.

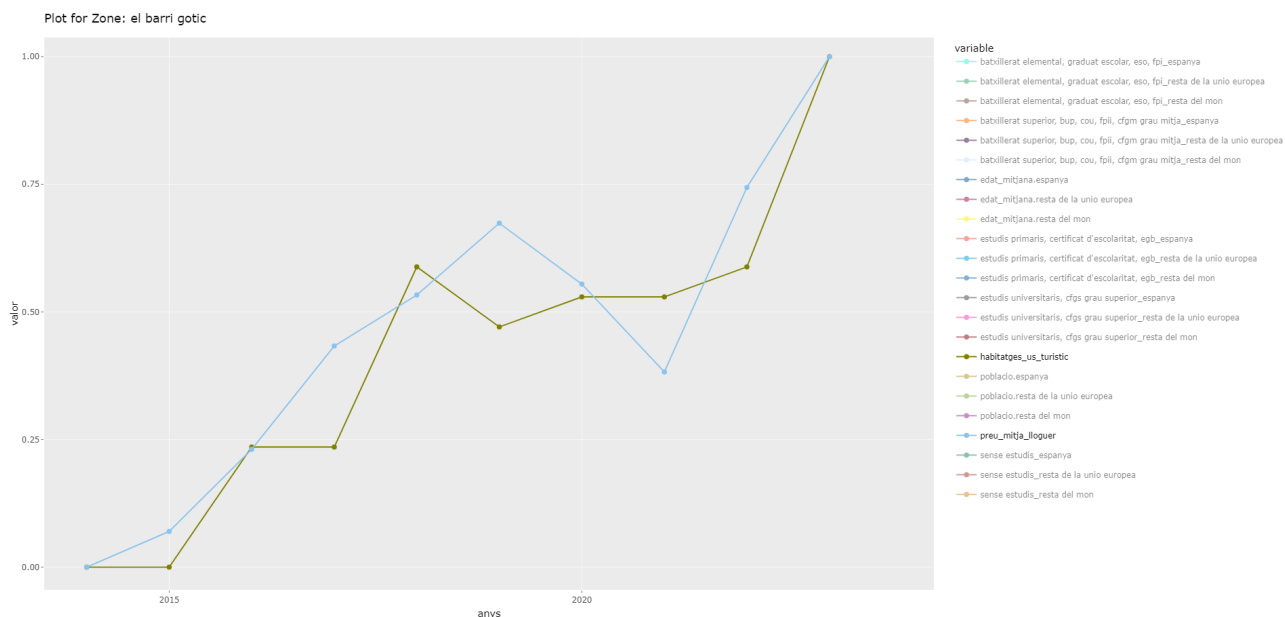


Figura 4.3: Evolució preu del lloguer i nombre d’habitatges d’ús turístic en El Gòtic. Font: Elaboració pròpia.

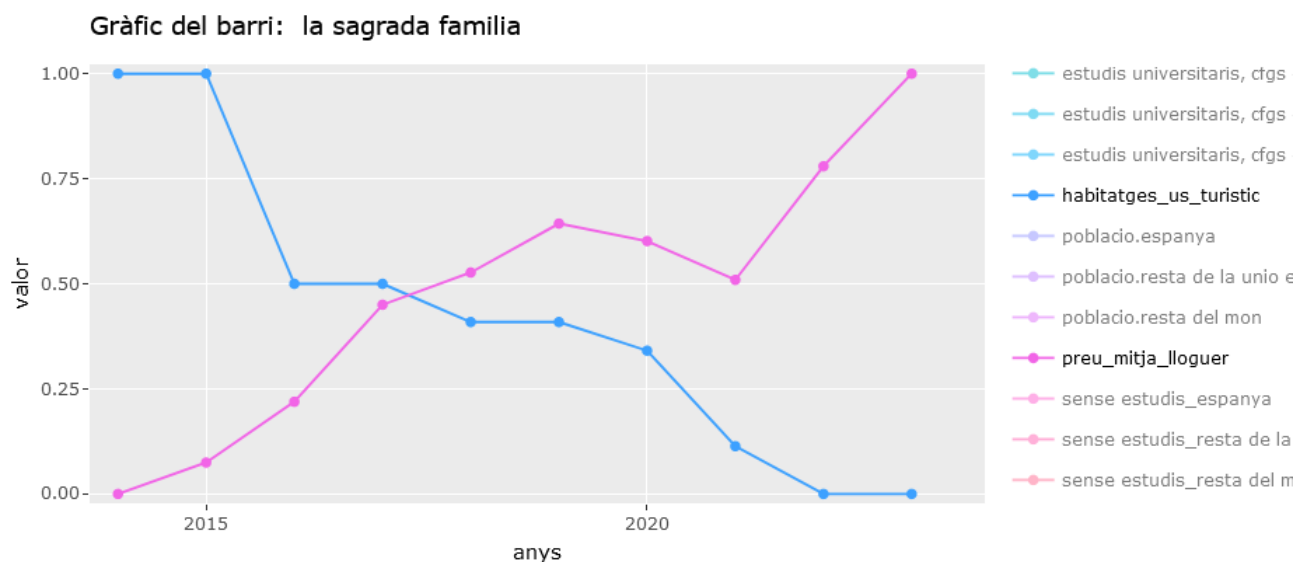


Figura 4.4: Evolució preu del lloguer i nombre d’habitatges d’ús turístic a la Sagrada Família. Font: Elaboració pròpia.

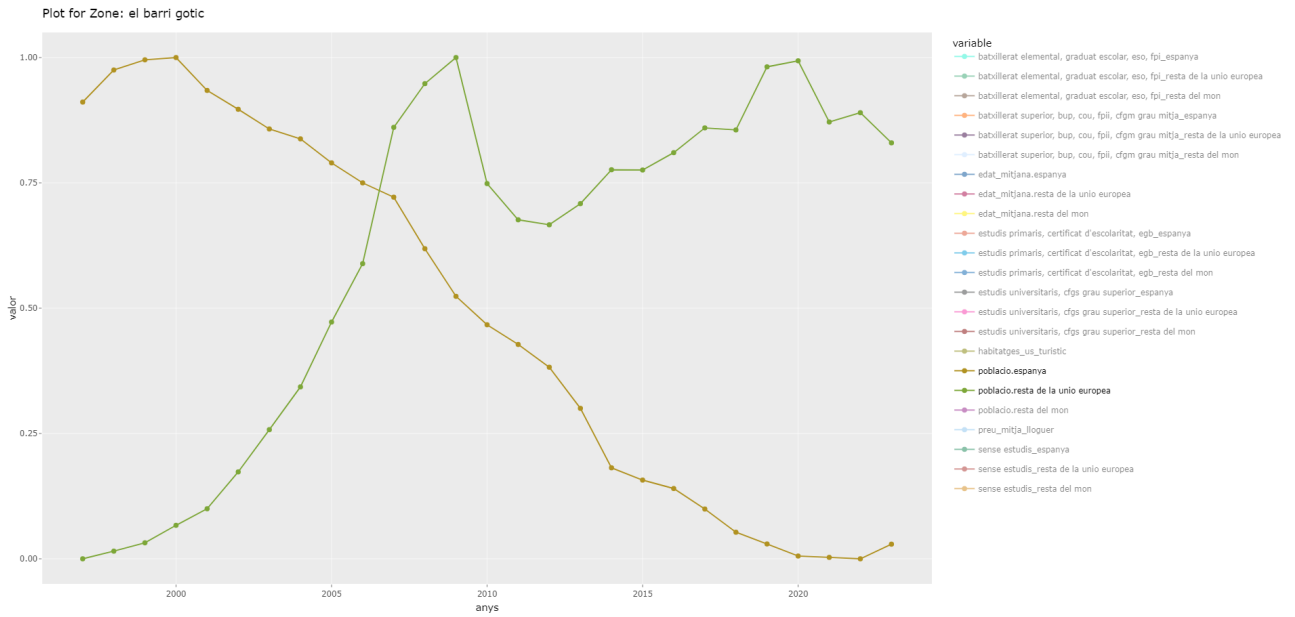


Figura 4.5: Comparació entre l'evolució de la població espanyola i la població europea. Font: Elaboració pròpia.

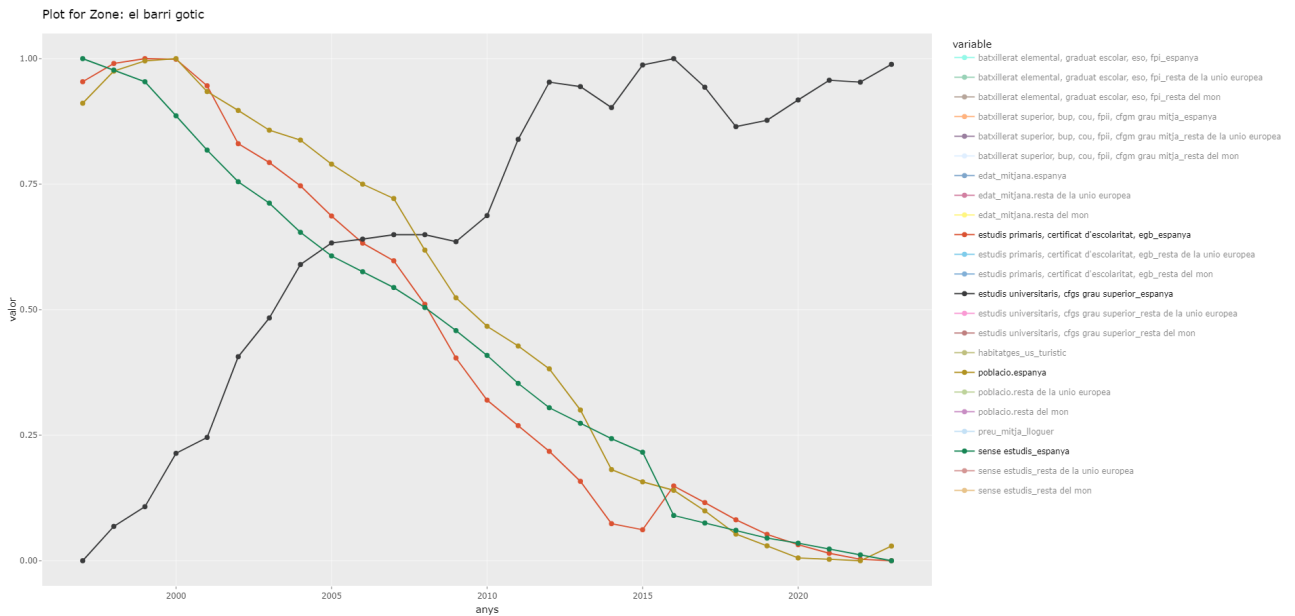


Figura 4.6: Comparativa evolució de la població espanyola segons el nivell d'estudis. Font: Elaboració pròpia.

territori	tipus de territori	anys	preu mitja lloguer
el raval	barri	2000	NA
el raval	barri	2001	NA
el raval	barri	2002	NA
el raval	barri	2003	NA
el raval	barri	2004	NA
el raval	barri	2005	NA
el raval	barri	2006	NA
el raval	barri	2007	NA
el raval	barri	2008	NA
el raval	barri	2009	NA
el raval	barri	2010	NA
el raval	barri	2011	NA
el raval	barri	2012	NA
el raval	barri	2013	NA
el raval	barri	2014	577.8528364
el raval	barri	2015	625.8214
el raval	barri	2016	667.0037911
el raval	barri	2017	748.5794858
el raval	barri	2018	821.8404327
el raval	barri	2019	850.1123432
el raval	barri	2020	833.8682277
el raval	barri	2021	768.7829456
el raval	barri	2022	911.974139
el raval	barri	2023	959.5023488

Taula 4.2: Dades del barri del Raval per a la taula de preus de lloguer.

En canvi, la població europea augmenta juntament amb aquella que té una formació superior o universitària, com es pot veure en la figura 4.7.

De les representacions observades es pot veure doncs que aquest desplaçament de gent del territori, del canvi a una població més formada i de l'augment del preu de lloguer al llarg del temps hi ha un efecte.

Per tant, la clusterització és necessària per observar en quin moment es produeixen aquests canvis. Aplicant el mètode kmeans el que s'obté és una clusterització en 3 grups que queden bastant ben marcats (veure figura 4.8)

Revisant els clústers associats (veure figura 4.10) amb els anys en què es produeixen es pot extreure que el clúster 3 representa “no gentrificat”, ja que es considera que el 1997 aquest efecte ha de ser menor en molts barris. El clúster 2 indicaria barris en un “procés de gentrificació” i el clúster 1 indica un barri totalment “gentrificat”. També per poder fer aquestes associacions s'han comparat els resultats obtinguts amb els que presenta l'estudi fet per l'Ajuntament de Barcelona per la Direcció de Serveis de Gènere i Polítiques del Temps [9] on s'aplica l'estudi



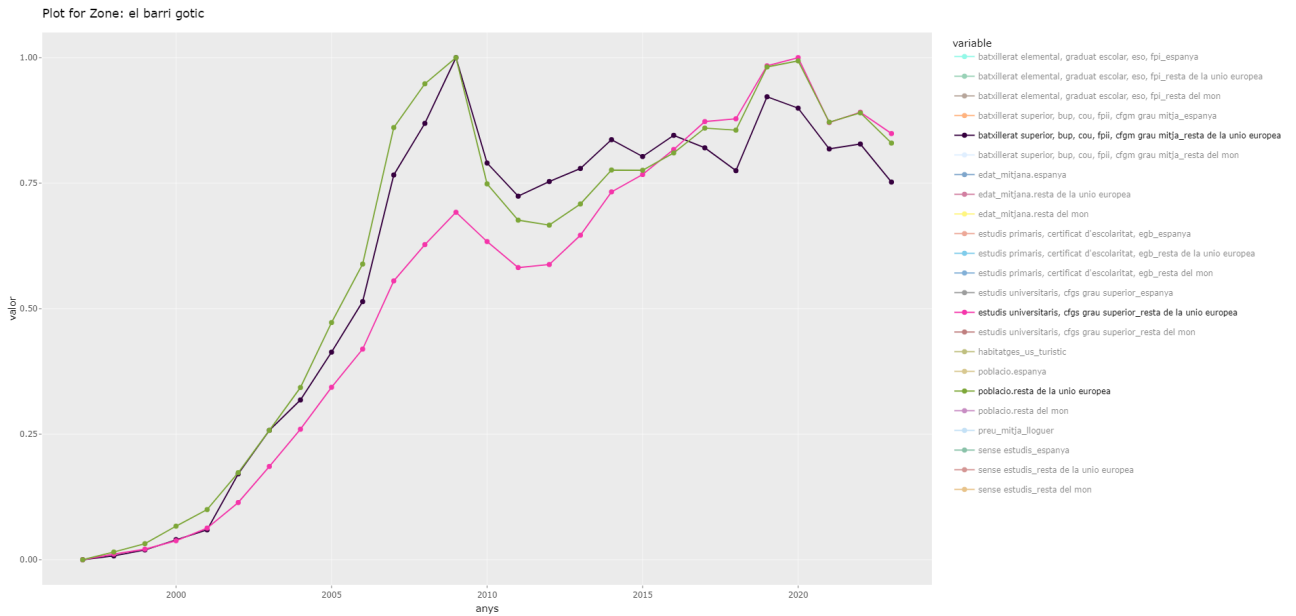


Figura 4.7: Comparativa evolució de la població europea segons el nivell d'estudis. Font: Elaboració pròpia.

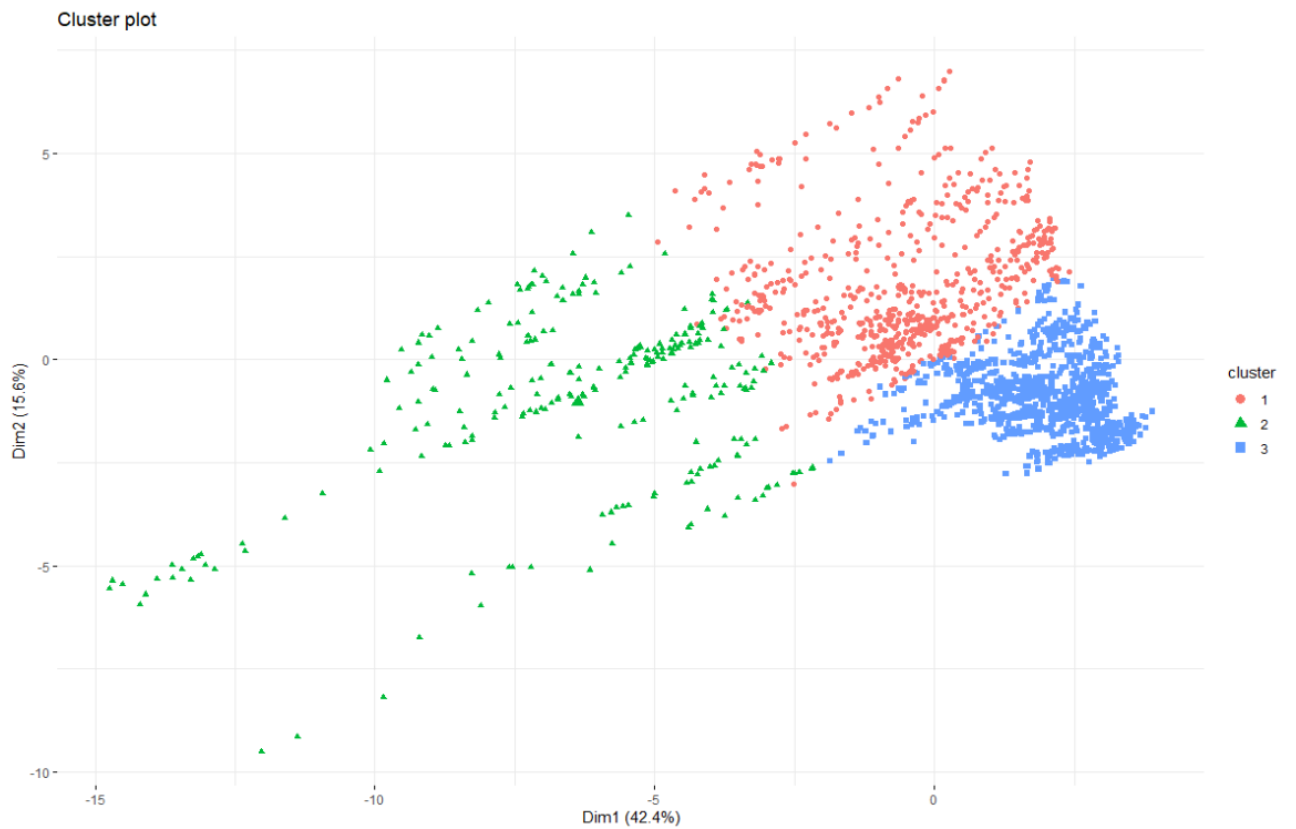


Figura 4.8: Clusterització mitjançant mètode kmeans (k=3). Font: Elaboració pròpia.

de López-Gay (2019). I es pot veure que en aquell any hi ha barris que coincideixen en el fet d'estar gentrificats (figura 4.9). A més a més, les dimensions que conformen el kmeans expliquen més del 50 % de la variació en les dades.

Per poder confirmar els resultats també s'ha realitzat un UMAP. Es pot veure el resultat en la figura 4.11. Aplicant el color dels clústers i el resultat del UMAP s'observa que la gran part dels punts queden ben agrupats. Es pot dubtar d'alguns resultats que pot ser que vinguin influenciats pels valors introduïts per substituir els valors NA.

Un cop obtinguts els clústers és possible identificar aquesta transició i efecte que s'ha definit que és la gentrificació. Però cal saber com s'estan agrupant aquestes zones. A partir dels centroides es pot obtenir el resultat de quines variables han tingut més pes a l'hora d'agrupar. Es pot observar que el que més afecta és la població de la resta del món. I de forma molt interessant és que totes les variables que agafen pes són les relacionades amb el nivell d'estudi de la població, i més concretament de la població de la Unió Europea i de la resta del món. S'observa doncs que té molt de pes la població i el nivell d'estudis de la gent que és de fora d'Espanya, com es pot veure en la figura 4.12.

El que es podria haver pensat que podria haver tingut molt d'efecte com seria el preu del lloguer o la presència d'habitatge turístic queden en posicions molts baixes respecte a la resta, per sota fins i tot de l'edat de la població.

Amb aquests resultats de la clusterització es fa una representació gràfica en un mapa, figura 4.13, que permet veure com varia al llarg del temps. Es pot observar com el procés gentrificador comença al centre de Barcelona i aquesta es va estenent cap enfora. Fet que es pot relacionar amb la sensació que comparteix la gent de desplaçament i d'expulsió de la ciutat.

El mateix procés s'ha realitzat per a veure aquest efecte a escala de districte. En la figura 4.14 es pot observar com és l'evolució de l'efecte gentrificador en els diferents districtes a partir de l'estat dels districtes en 4 anys diferents (1997, 2005, 2014, 2023).

Es pot veure en la figura 4.15 del mapa de calor com les variables de població d'Europa i resta del món estan molt relacionades amb les variables de nivell d'estudis superiors. Mentre que la població espanyola està molt relacionada amb el nivell d'estudis inferiors o sense estudis.

També es pot observar que el nombre d'habitatges turístics i el preu de lloguer en una zona tenen una relació molt petita (0.2) amb les dades que s'estan manipulant.

### 3 Arbres de decisió

El kmeans ha permès crear tres grups de classificació per a les dades estudiades. Donat que aquestes dades són significativament diferents entre elles es planteja elaborar un model classificador. Aquest classificador ha de permetre que a partir de la introducció de nous valors per

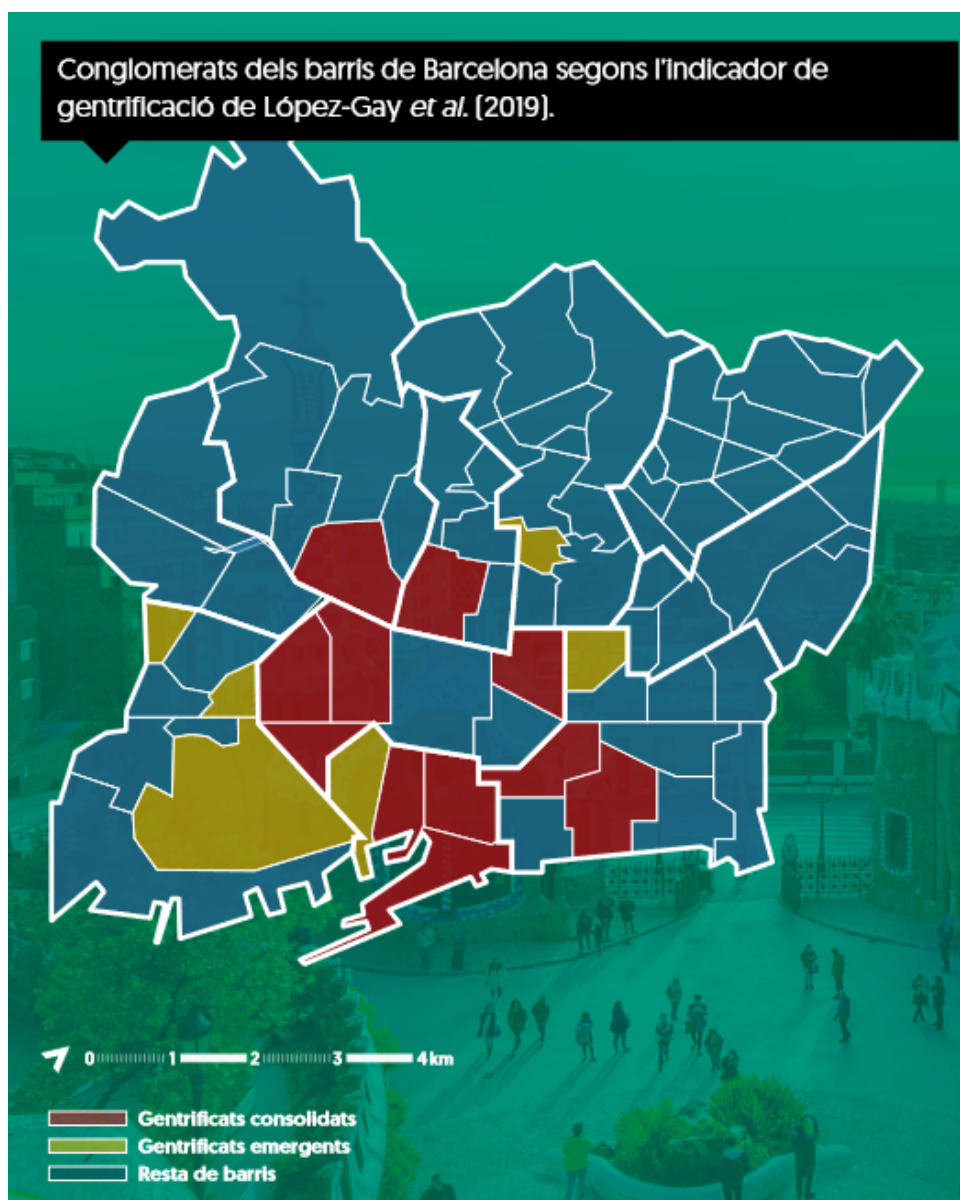


Figura 4.9: Gentrificació de Barcelona segons l'estudi realitzat per la Direcció de Serveis de Gènere i Polítiques del Temps. Font: Direcció de Serveis de Gènere i Polítiques del Temps [9]

	<b>cluster</b> <int>	<b>territori</b> <chr>
el raval_1997	2	el raval_1997
el raval_1998	2	el raval_1998
el raval_1999	2	el raval_1999
el raval_2000	2	el raval_2000
el raval_2001	2	el raval_2001
el raval_2002	2	el raval_2002
el raval_2003	1	el raval_2003
el raval_2004	1	el raval_2004
el raval_2005	1	el raval_2005
el raval_2006	1	el raval_2006
el raval_2007	1	el raval_2007
el raval_2008	1	el raval_2008
el raval_2009	1	el raval_2009
el raval_2010	1	el raval_2010
el raval_2011	1	el raval_2011
el raval_2012	1	el raval_2012
el raval_2013	1	el raval_2013
el raval_2014	1	el raval_2014
el raval_2015	1	el raval_2015
el raval_2016	1	el raval_2016
el raval_2017	1	el raval_2017
el raval_2018	1	el raval_2018
el raval_2019	1	el raval_2019
el raval_2020	1	el raval_2020
el raval_2021	1	el raval_2021
el raval_2022	1	el raval_2022
el raval_2023	1	el raval_2023
el barri gotic_1997	3	el barri gotic_1997
el barri gotic_1998	3	el barri gotic_1998
el barri gotic_1999	3	el barri gotic_1999
el barri gotic_2000	3	el barri gotic_2000
el barri gotic_2001	3	el barri gotic_2001
el barri gotic_2002	3	el barri gotic_2002
el barri gotic_2003	3	el barri gotic_2003

Figura 4.10: Fragment dels resultats de la clusterització per kmeans. Font: Elaboració pròpia.

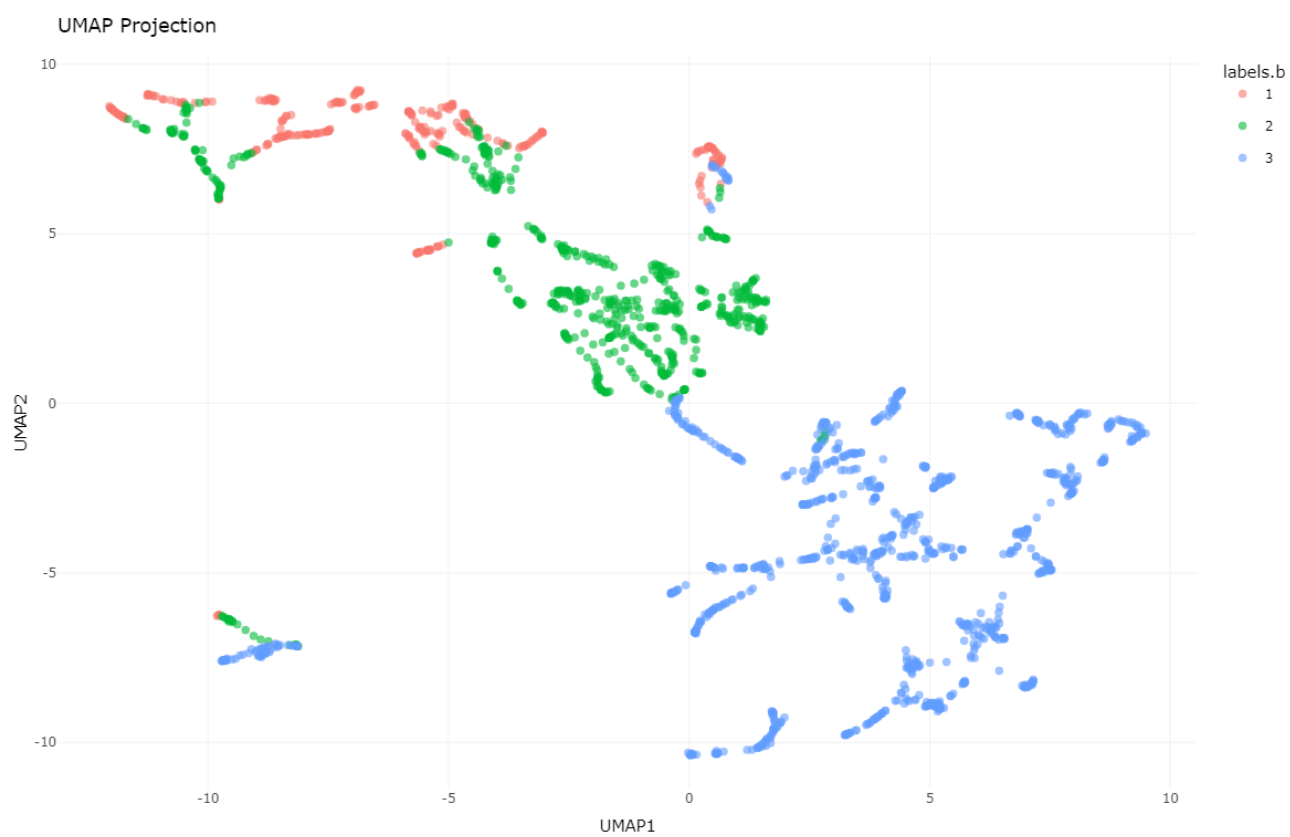


Figura 4.11: UMAP sobre les dades amb assignació de colors per cluster. Font: Elaboració pròpia.

```

poblacio.resta del mon
1.9559022
batxillerat superior, bup, cou, fpii, cfgm grau mitja_rest de la unio europea
1.8997525
batxillerat elemental, graduat escolar, eso, fpi_rest de la unio europea
1.8993410
estudis primaris, certificat d'escolaritat, egb_rest del mon
1.8712640
poblacio.rest de la unio europea
1.8589438
estudis primaris, certificat d'escolaritat, egb_espanya
1.8567195
batxillerat elemental, graduat escolar, eso, fpi_rest del mon
1.8509365
batxillerat superior, bup, cou, fpii, cfgm grau mitja_rest del mon
1.8130528
estudis primaris, certificat d'escolaritat, egb_rest de la unio europea
1.7958981
sense estudis_espanya
1.7898470
estudis universitaris, cfigs grau superior_rest de la unio europea
1.7593879
estudis universitaris, cfigs grau superior_rest del mon
1.6974410
edat_mitjana.rest del mon
1.6057270
batxillerat elemental, graduat escolar, eso, fpi_espanya
1.3990781
preu_mitja_lloguer
1.3841117
edat_mitjana.espanya
1.0641174
edat_mitjana.rest de la unio europea
1.0441781
sense estudis_rest del mon
1.0279259
poblacio.espanya
1.0126166
estudis universitaris, cfigs grau superior_espanya
0.9306404
habitatges_us_turistic
0.7414366
batxillerat superior, bup, cou, fpii, cfgm grau mitja_espanya
0.6932268
sense estudis_rest de la unio europea
0.6468568

```

Figura 4.12: Pes de les variables en la clusterització. Font: Elaboració pròpia.

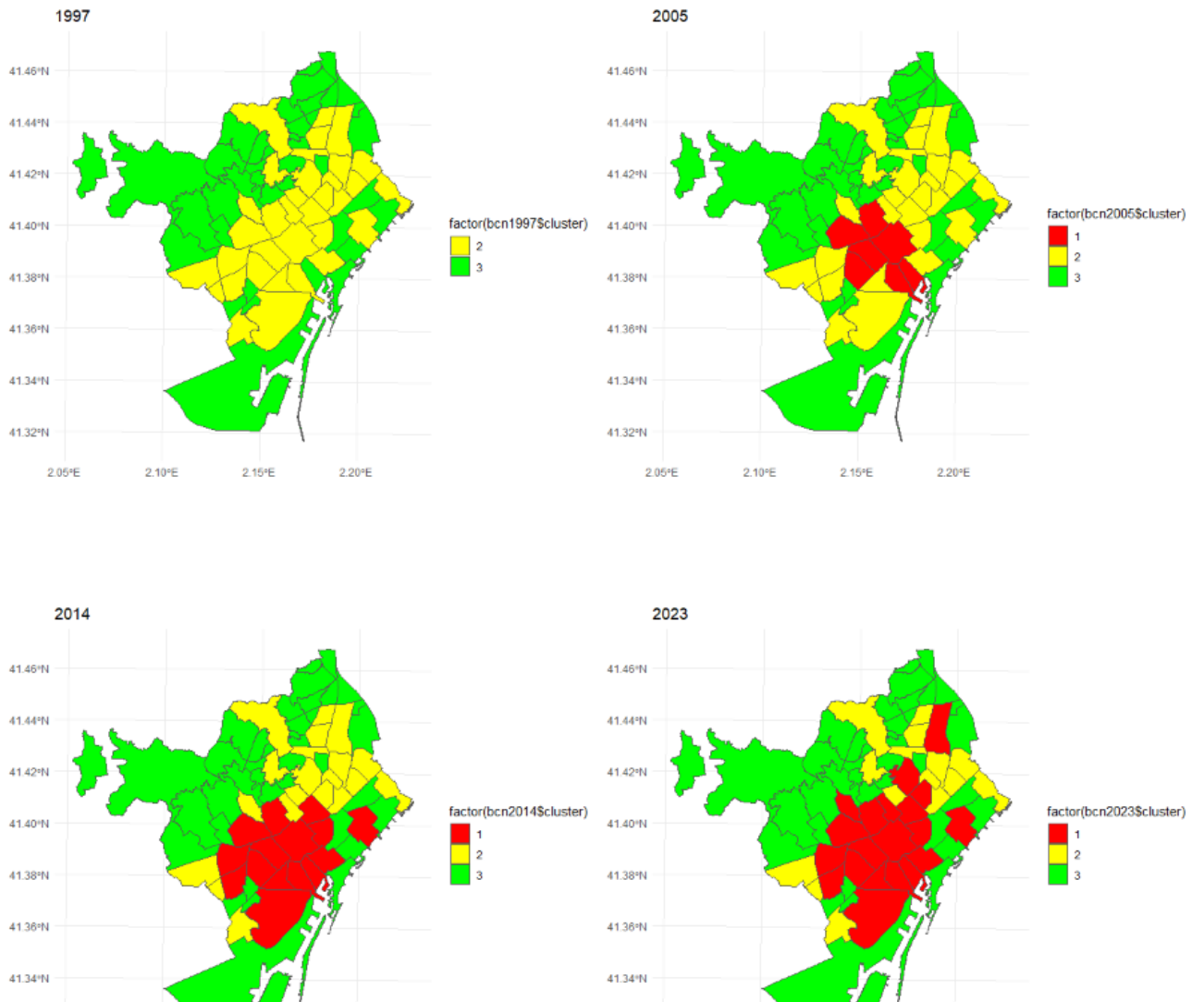


Figura 4.13: Visualització del procés de gentrificació a Barcelona en 4 anys diferents a nivell de barri (1997, 2005, 2014, 2023). Font: Elaboració pròpia

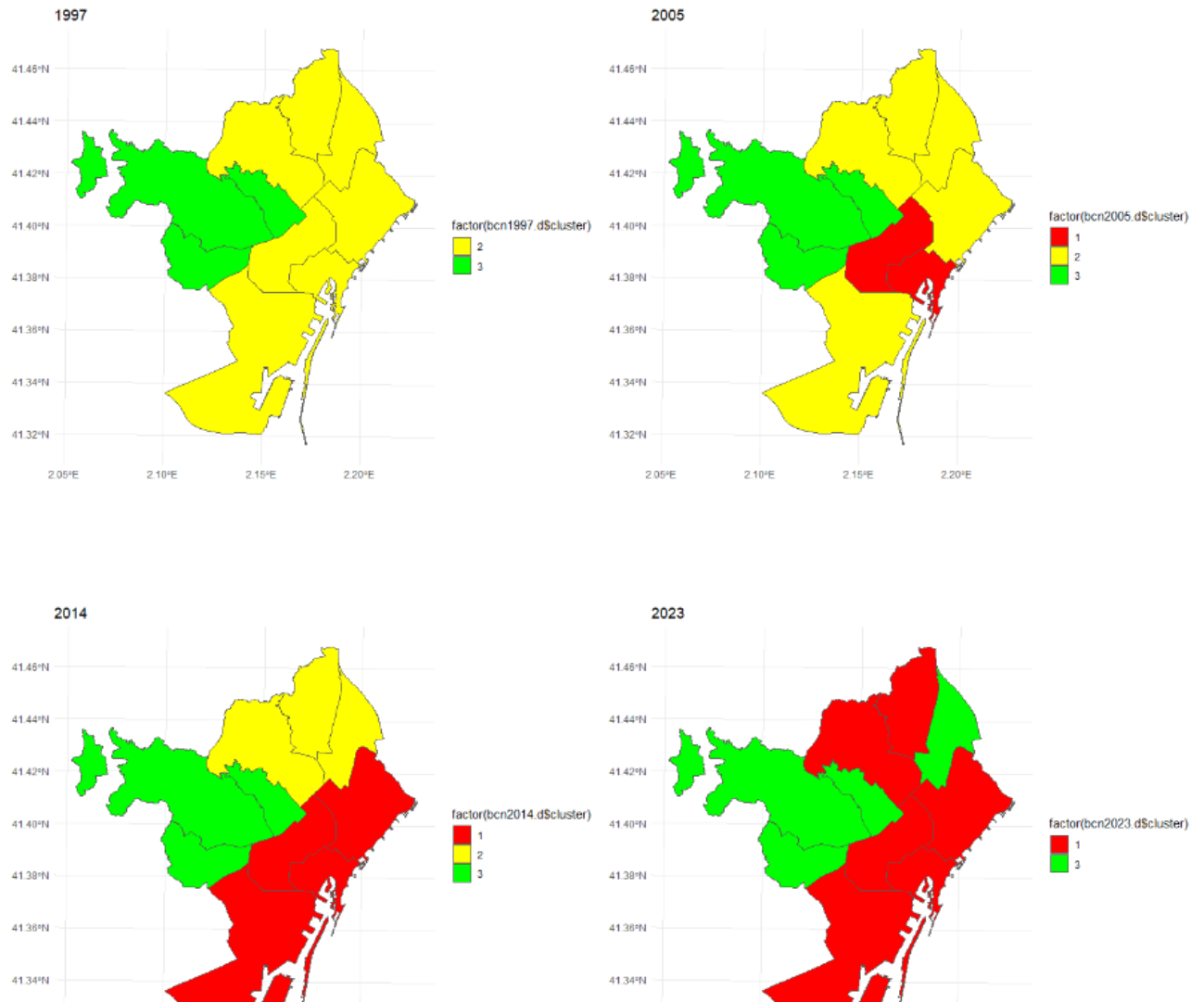


Figura 4.14: Visualització del procés de gentrificació a Barcelona en 4 anys diferents a nivell de districte (1997, 2005, 2014, 2023). Font: Elaboració pròpia



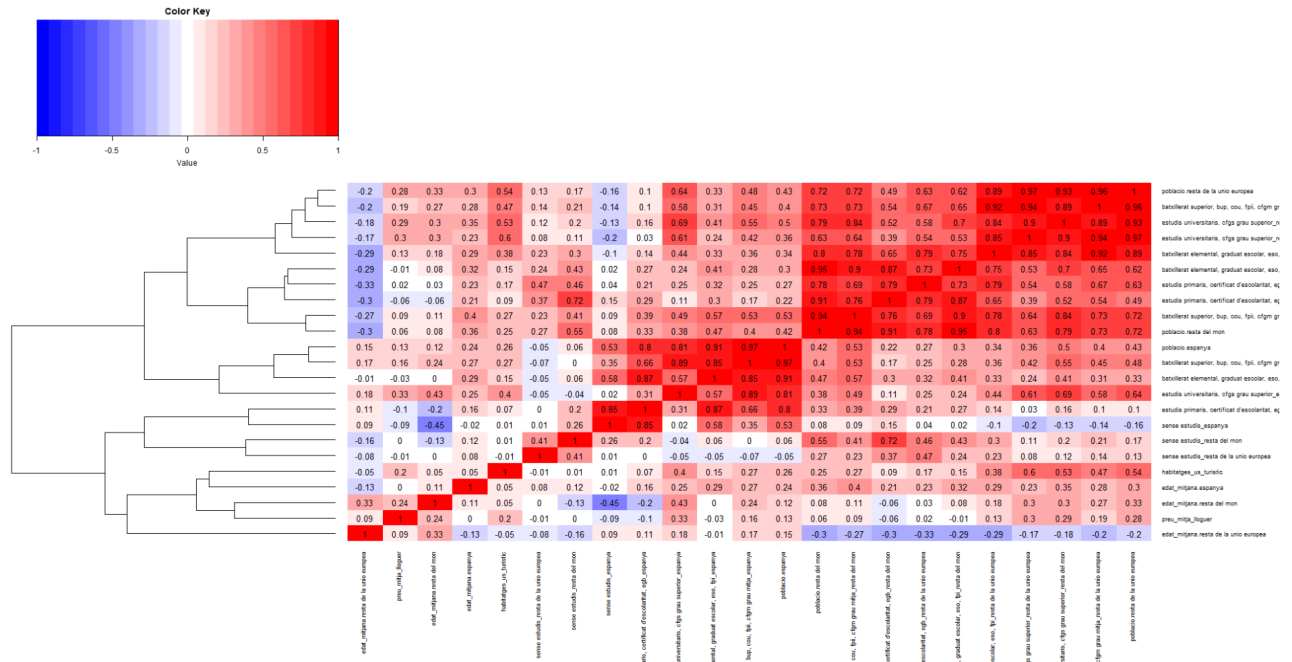


Figura 4.15: Mapa de calor de la correlació entre les variables utilitzades en l'estudi. Font: Elaboració pròpia.

cada una de les variables indiqui si hi ha un procés gentrificador o no en la zona d'interès. En el cas estudiat les dades provenen de les dades originals. Les quals han sigut dividides en un grup d'entrenament i un de test. Amb les dades d'entrenament es crea el model predictiu de classificació de l'arbre de decisió i un cop obtingut el model s'aplica sobre les dades de test. El resultat de l'arbre es pot observar en la figura 4.16. Es pot veure en l'arbre de decisió que les decisions es basen en el nivell d'estudis i per tipus de població. L'arbre de decisió ha detectat que el nivell d'estudis a tenir en compte és el batxillerat elemental/graduat escolar però no només això, sinó els que estan relacionats amb la població de la Unió Europea o espanyola que té aquests títols. Pel que fa a nombre de població l'arbre de decisió ha tingut en compte totes les variables possibles amb la població espanyola com el primer nivell de divisió i després la població de la Unió Europea i de la resta del món.

El model creat presenta una precisió del 95-96% amb les dades utilitzades. Es pot veure en la matriu de confusió, figura 4.17, que on més errors es comet error és en la classificació en "gentrificat".

## 4 Creació d'una visualització del procés de gentrificació

Per finalitzar, un dels objectius del treball era crear una visualització del procés de gentrificació. El que s'ha desenvolupat és una visualització interactiva, aplicació shiny, en la qual es pot veure

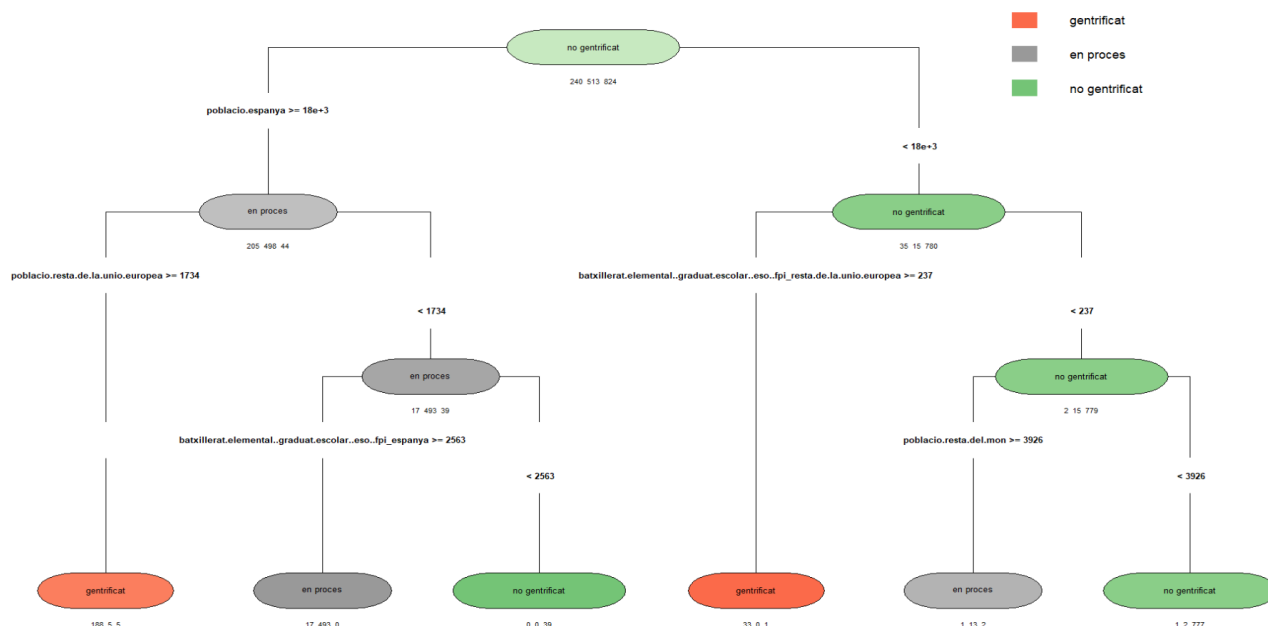


Figura 4.16: Arbre de decisió elaborat a partir de les dades de clusterització. Font: Elaboració pròpia

#### Confusion Matrix and Statistics

Prediction	Reference		
	gentrificat	en proces	no gentrificat
gentrificat	49	0	0
en proces	11	128	1
no gentrificat	0	0	205

#### Overall Statistics

Accuracy : 0.9695  
 95% CI : (0.9474, 0.9842)  
 No Information Rate : 0.5228  
 P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.9487

Mcnemar's Test P-Value : NA

#### Statistics by Class:

	Class: gentrificat	Class: en proces	Class: no gentrificat
Sensitivity	0.8167	1.0000	0.9951
Specificity	1.0000	0.9549	1.0000
Pos Pred Value	1.0000	0.9143	1.0000
Neg Pred Value	0.9681	1.0000	0.9947
Prevalence	0.1523	0.3249	0.5228
Detection Rate	0.1244	0.3249	0.5203
Detection Prevalence	0.1244	0.3553	0.5203
Balanced Accuracy	0.9083	0.9774	0.9976

[1] "Accuracy: 0.969543147208122"

Figura 4.17: Resultats de l'arbre de decisió.

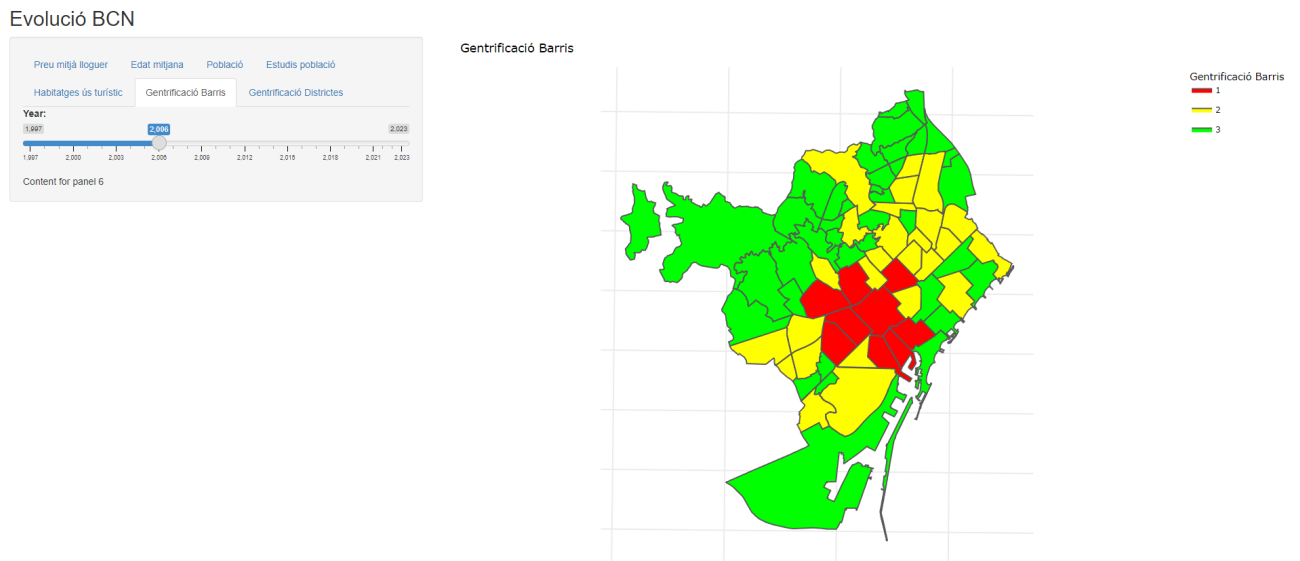


Figura 4.18: Visualització de la interfície de l'aplicació shiny per veure l'evolució de les variables i del procés de gentrificació a Barcelona. Font: Elaboració pròpia

la variació al llarg dels anys de cada una de les variables en un mapa. A més a més, es permet, si és que existeix, la selecció de la variable per tipus de població (espanya, eu o resta del món). En l'aplicació també s'ha afegit l'evolució del procés de gentrificació obtingut de la clusterització tant a escala de districte com de barri al llarg dels anys.

L'aplicació es pot consultar en el següent enllaç <https://omartinezgarciauoc.shinyapps.io/gentrifications/>

Un exemple de la interfície es pot veure en la figura 4.18.



# Capítol 5

## Conclusions i treballs futurs

### 1 Conclusions

L'elaboració d'aquest treball ha permès veure que en la clusterització ha tingut molt de pes el lloc d'origen de la població. L'augment de la població de fora d'Espanya en una zona és un primer marcador per identificar si una zona té un cert poder atractiu i, per tant, pot començar a patir canvis que propiciïn la gentrificació. A més, aquesta població que arriba al llarg del temps ho fa amb un nivell d'estudis superior. La població autòctona amb un baix nivell d'estudis disminueix, per tant, s'observa l'efecte desplaçament que s'associa a la gentrificació, tant a escala de barri com de districte. Pel que fa al preu de lloguer de la zona s'ha pogut observar que aquest tot i mostrar certa tendència a anar incrementant al llarg dels anys ho fa de forma més pronunciada on hi ha una major població europea o de la resta del món. Aquests preus es van veure afectats en el període de la COVID-19. La presència dels habitatges d'ús turístic en el treball no ha sigut un factor que el model hagi tingut en compte a l'hora d'agrupar els diferents districtes o barris. Per tant, podem concloure que no són un efecte directament gentrificador però sí un producte a causa d'aquest procés gentrificador. Ha sigut interessant observar també com aquest nombre d'habitatges en zones classificades com gentrificades mostrava una corba d'evolució similar a l'evolució dels preus de lloguer. Tot i suposar que l'edat podria ser un element que tindria efecte en la gentrificació s'ha trobat que aquest no era dels elements principals, per a la classificació proposada. Tot i això, es veu afectada per la població que arriba. En el cas de la població de la Unió Europea i la resta del món aquesta edat augmenta, mentre que per la gent d'Espanya aquesta edat disminueix degut, molt possiblement, pel desplaçament de la població.

La visualització en mapa ha permès veure com amb les dades proposades la ciutat de Barcelona ja parteix d'un estat de "en procés de gentrificació". Es podria relacionar aquest efecte amb els Jocs Olímpics del 92 i això va portar grans canvis a la ciutat, i les dades comencen

des de 1997. El procés s'ha pogut veure que comença en barris propers al mar i després s'estén cap a barris centrals. En la seva evolució aquest procés es va estenent de dins cap enfora.

## 2 Assoliment objectius generals

Al principi del treball es van plantejar diferents objectius a assolir. Un cop finalitzat l'estudi els objectius generals s'han assolit amb cert èxit. S'han trobat i presentat diferents estudis que analitzen els efectes gentrificadors en diverses ciutats i des de diferents perspectives que han permès agafar idees i combinar algunes d'elles per poder dur a terme el treball actual. S'ha assolit amb èxit l'extracció de dades i anàlisi d'aquestes a partir del portal de dades obertes de l'Ajuntament de Barcelona. Amb el IDE de R s'ha pogut posar en pràctica diferents tècniques de neteja i tractament de les dades per tal de fer-les adequades per a la seva manipulació. A més d'aplicar diverses funcions de big data per a poder extreure informació i coneixement d'elles. S'ha assolit l'elaboració de tota una memòria per mostrar els resultats del treball fet.

## 3 Assoliment objectius específics

Es van plantejar cinc objectius específics i de forma general s'han assolit tots. S'ha pogut obtenir, transformar i integrar les dades a partir d'una font. En aquest cas el plantejament era l'obtenció de diverses fonts, però el portal de dades ha mostrat disposar d'una gran quantitat de dades que han reduït molt la recerca d'aquestes. En certa manera s'ha fet una anàlisi de les dades. Tot i centrar el treball en gràfiques de línies i mapes les dades permeten treballar i mostrar-les en diferents formats. El plantejament seguit ha sigut reduir el nombre de formats possibles per a no sobrecarregar la dedicació en un sol apartat. Tot i això, la representació de les dades i la part interactiva aporten prou interès i informació que se'n pot extreure i fer una anàlisi de diferents factors, sigui individualment o en conjunt. S'assoleix amb èxit l'objectiu de la creació de visualitzacions interactives, algunes en format fitxer HTML i altres en format aplicació shiny. Finalment, un objectiu que s'ha assolit amb cert èxit de desenvolupar ha sigut el model de classificació. Aquest model classificador ha sigut possible pel fet que els models de clusterització hagin sigut prou adequats per agrupar i diferenciar les diferents zones d'estudi a partir de les variables disponibles.

## 4 Planificació i metodologia

La planificació proposada al principi d'aquest treball era conservadora quant als períodes de temps que s'assignaven a cada tasca o mòdul. Tot i això, el mateix desenvolupament del

treball ha mostrat que la recerca tant d'articles sobre l'estat de l'art com de taules amb dades interessants per ser treballades ha dut més temps del plantejat. No només la recerca sinó també el temps dedicat per entendre i plantejar una estratègia adequada per treballar cada element adequadament. Les variables seleccionades han sigut valorades per sobre de la resta per poder ser prou diferenciades entre elles i obrint així el ventall d'opcions de clusteritzacions possibles que poguessin ocórrer. Una part millorable en la metodologia aplicada en el treball és el tractament de l'absència de dades o NA. En el treball s'ha justificat l'ús del valor de la mitjana d'una variable per substituir els valors NA i evitar la pèrdua d'informació. Ara bé, aquesta decisió pot haver tingut efecte en la manera com després s'han classificat les diferents zones en la clusterització. Per tant, és possible que una altra manera de tractar les dades pugui ser més adient o complementària. És millorable també l'ús de les gràfiques i les diferents representacions que permeten les dades. Un major nombre de representacions també implica més temps dedicat en aquesta part i en aquest treball s'ha volgut trobar un equilibri en el temps disponible per dur a terme totes les parts.

En conjunt, s'ha pogut complir amb la planificació establerta i amb la metodologia aplicada s'han pogut arribar a resultats i complir els objectius establerts.

## 5 Treballs futurs

El tema de la gentrificació en una ciutat és molt interessant i pot ser treballat des de moltes branques. En el cas concret, on es fa ús del big data com a eina per extreure informació de les dades té recorregut per ser estudiat amb la gran quantitat de dades que hi ha disponibles sobre barris i districtes, en el cas de Barcelona. Es podria investigar aquest procés a partir d'altres variables i veure si els patrons observats es repeteixen, és a dir, si tornen a aparèixer els mateixos barris gentrificats i si ho fan els mateixos anys. Seria interessant potser fer un estudi similar, però en comptes d'en tota la ciutat fer-ho directament a escala de barri. I això combinar-ho amb coneixements més sociològics sobre quins esdeveniments o actuacions han pogut tenir lloc en aquell barri en concret i intentar veure si hi ha una relació entre variable i procés gentrificador. Amb les dades també es poden fer combinacions de variables en el cas de Barcelona i provar de clusteritzar amb aquests factors. Es podria treballar a identificar si hi ha alguna variació que indiqui encara amb més precisió que actua en el procés gentrificador. Ara bé, treballar amb tantes dades també pot arribar a produir un efecte de donar voltes sobre un punt central (gentrificació) però sense donar en l'efecte clau.

A banda de l'estudi de les taules de dades, una investigació que es podria fer seria veure com han canviat els locals que hi són presents en una zona. Trobar dades d'ubicació geogràfica d'establiments que puguin estar associats a anys i veure com ha variat el teixit comercial, les

activitats que es realitzen al barri.

Un altre treball possible seria intentar trobar alguna variable que segueixi alguna tendència, intentar trobar el seu patró i extrapolar-ho en el temps. Aquest treball hauria de permetre fer prediccions a llarg termini i amb els valors i la combinació d'un model ja de gentrificació, com l'elaborat en aquest treball per exemple, veure en quins anys els barris fan o no un canvi en l'estat en el qual es troben actualment.

En conclusió, la quantitat de dades disponibles permet fer l'estudi del procés gentrificador des de diverses perspectives, i la combinació amb altres coneixements de caràcter més social poden donar un punt extra de valor al que s'observa en les dades.



# Glossari

**Airbnb:** Plataforma que ofereix l'opció de llogar l'espai que tinguin disponible, sigui propietat completa o part d'ella.

**ANOVA:** Tècnica estadística que s'utilitza per comparar la mitjana de tres o més grups de determinar si existeixen diferències significatives entre elles.

**ASCII:** Codi de caràcters basat en l'alfabet llatí.

**char:** Tipus de dades que emmagatzema dades de caràcters en un camp de longitud fixa.

**clúster:** Grup d'elements similars o propers, agrupats en funció d'alguna característica o variable.

**DBSCAN:** Algorisme que defineix els clústers mitjançant l'estimació de la densitat local.

**dendograma:** Diagrama que mostra les distàncies d'atributs entre cada parell de classes fusionades seqüencialment.

**diagrama de Gantt:** Eina de gestió de projectes que il·lustra el treball fet durant un període de temps en relació amb el temps previst per al treball.

**gentrificació:** Procés de renovació d'una zona urbana, generalment popular o deteriorada, mitjançant un procés que implica el desplaçament de la seva població original per part d'una de major poder adquisitiu.

**github:** Plataforma en línia de desenvolupament de programari que s'utilitza per emmagatzemar, supervisar i treballar amb projectes de programari.

**HTML:** Codi que s'utilitza per estructurar i desplegar una pàgina web i els seus continguts.

**int:** Tipus de dades que emmagatzema nombres enters.

**K-means:** Algorisme de classificació no supervisat que agrupa objectes en k grups basant-se en les seves característiques.

**machine learning:** Ciència de desenvolupament d'algorismes i models estadístics.

**NA:** Abreviatura per indicar l'omissió d'informació existent en una taula.

**num:** Tipus de dades que emmagatzema números de separador flotant.

**PCA:** Mètode estadístic que permet reduir la dimensionalitat de la base de dades.

**PERMANOVA:** Model utilitzat en ecologia i biologia per explicar comunitats amb variables ambientals.

**R Studio:** Entorn de desenvolupament integrat (IDE) per al llenguatge de programació R, dedicat a la computació estadística i els gràfics.

**Random forests:** Algoritme d'aprenentatge automàtic que combina la sortida de múltiples arbres de decisió per arribar a un sol resultat.

**shiny:** Paquet de R que crea aplicacions web interactives directament des de R.

**UMAP:** Mètode de reducció de dimensionalitat i clústering de dades.

**web crawling:** Programa informàtic que inspecciona les pàgines d'internet de forma automatitzada.

# Bibliografia

- [1] Portal Barcelona Dades. (s.f.). <https://portaldades.ajuntament.barcelona.cat/ca/>.
- [2] Unidades administrativas de la ciudad de Barcelona - Open Data Barcelona. (2024, 18 enero). <https://opendata-ajuntament.barcelona.cat/data/es/dataset/20170706-districtes-barris>.
- [3] Gràfiques per barris (html). [https://github.com/Gamaor4/gentrification\\_bcn.TFM/tree/main/grafiques](https://github.com/Gamaor4/gentrification_bcn.TFM/tree/main/grafiques).
- [4] Anàlisi variables per barri. [https://omartinezgarciauoc.shinyapps.io/Interactive\\_graphs\\_bcn/](https://omartinezgarciauoc.shinyapps.io/Interactive_graphs_bcn/).
- [5] Yesenia Alejandro and Leon Palafox. Gentrification prediction using machine learning. In Lourdes Martínez-Villaseñor, Ildar Batyrshin, and Antonio Marín-Hernández, editors, *Advances in Soft Computing*, pages 187–199, Cham, 2019. Springer International Publishing.
- [6] Bianca Antunes, Hug March, and James J.T. Connolly. Spatializing gentrification in situ: A critical cartography of resident perceptions of neighbourhood change in vallcarca, barcelona. *Cities*, 97:102521, 2020.
- [7] Marcos Arizmendi. La pujada del preu del lloguer triplica l'ípc dels últims 10 anys a 28 barris de barcelona, 2023. Accessed: 2023-10-13.
- [8] Marta Bottero, Caterina Caprioli, Marcus Foth, Peta Mitchell, Markus Rittenbruch, and Marco Santangelo. Urban parks, value uplift and green gentrification: An application of the spatial hedonic model in the city of brisbane. *Urban Forestry & Urban Greening*, 74:127618, 2022.
- [9] Direcció de Serveis de Gènere i Polítiques del Temps, Gerència Municipal, Centre d'Estudis Demogràfics, Universitat Autònoma de Barcelona, Antonio López Gay, Joan Sales Favà, and Anna Ortiz Guitart. La gentrificació a la ciutat de barcelona des d'una perspectiva de gènere: aproximació quantitativa. 2020.

- 
- [10] Ingrid Gould Ellen, Keren Mertens Horn, and Davin Reed. Has falling crime invited gentrification? *Journal of Housing Economics*, 46:101636, 2019.
- [11] Nicholas Finio. Measurement and definition of gentrification in urban studies and planning. *Journal of Planning Literature*, 37(2):249–264, 2022.
- [12] Caislin L. Firth, Benoit Thierry, Daniel Fuller, Meghan Winters, and Yan Kestens. Gentrification, urban interventions and equity (genuine): A map-based gentrification tool for canadian metropolitan areas. *Health Reports*, 32, 2021.
- [13] Ruth Glass. London: Aspects of change. In *London: Aspects of Change*, pages 13–42. MacGibbon & Kee, London, 1964.
- [14] Han Li, Yehua Dennis Wei, Yangyi Wu, and Guang Tian. Analyzing housing prices in shanghai with open data: Amenity, accessibility and urban structure. *Cities*, 91:165–179, 2019.
- [15] Cheng Liu, Yu Deng, Weixuan Song, Qiyan Wu, and Jian Gong. A comparison of the approaches for gentrification identification. *Cities*, 95:102482, 2019.
- [16] A. Lopez-Gay, A. Ortiz-Guitart, and M. Solana-Solana. Vivienda, cambio poblacional y desplazamiento en un barrio en proceso de gentrificación. el caso de sant antoni (barcelona). *Dialnet*, 48, 2022.
- [17] Antonio López-Gay, Joan Sales-Favà, Miguel Solana, Ana Fernández, and Andrés Peralta. Midiendo los procesos de gentrificación en barcelona y madrid: una propuesta metodológica. 2019.
- [18] Milagros Perez Oliva. Gentrificació, l’aleta del tauró. Technical report, Arquitectura Tècnica de Barcelona, 2022.
- [19] Hamidreza Rabiei-Dastjerdi, Gavin McArdle, and William Hynes. Which came first, the gentrification or the airbnb? identifying spatial patterns of neighbourhood change using airbnb data. *Habitat International*, 125:102582, 2022.
- [20] Jonathan Reades, Jordan De Souza, and Phil Hubbard. Understanding urban gentrification through machine learning. *Urban Studies*, 56(5):922–942, 2019.
- [21] Nathalie Voorhees. The socioeconomic change of chicago’s community areas (1970-2010). Technical report, University of Illinois - College of Urban Planning and Public Affairs, 2014.

- 
- [22] Jennifer R. Wolch, Jason Byrne, and Joshua P. Newell. Urban green space, public health, and environmental justice: The challenge of making cities ‘just green enough’. *Landscape and Urban Planning*, 125:234–244, 2014.
- [23] Qiyang Wu, Jianquan Cheng, Guo Chen, Daniel J. Hammel, and Xiaohui Wu. Socio-spatial differentiation and residential segregation in the chinese city based on the 2000 community-level census data: A case study of the inner city of nanjing. *Cities*, 39:109–119, 2014.



# Annexos

## Taules de dades

Les taules de dades netes completes es poden consultar en l'enllaç [https://github.com/Gamaor4/gentrification\\_bcn.TFM/tree/main/dades](https://github.com/Gamaor4/gentrification_bcn.TFM/tree/main/dades)

Les següent figures mostren fragments de les taules:

territori	tipus_de_territori	anys	edat_mitjana.espanya	edat_mitjana.resta de la unio europea	edat_mitjana.resta del mon
barcelona	municipi	1997	43.32376794	41.9414097	32.99814867
barcelona	municipi	1998	43.50189689	41.25545255	32.67634371
barcelona	municipi	1999	43.77756644	40.81533231	32.48814122
barcelona	municipi	2000	44.0269762	40.33433997	32.14361146
barcelona	municipi	2001	44.31718089	39.24046458	31.45333697
barcelona	municipi	2002	44.51043911	37.19833908	31.06700371
barcelona	municipi	2003	44.80827157	36.10706371	30.93472077
barcelona	municipi	2004	45.09822827	35.60944664	31.0107081
barcelona	municipi	2005	45.39967719	34.82311043	31.32640468
barcelona	municipi	2006	45.59865281	34.3216967	31.59828746
barcelona	municipi	2007	45.78006455	33.58148753	31.55772193
barcelona	municipi	2008	45.93560773	33.45169018	31.64172918
barcelona	municipi	2009	46.06181731	33.63536953	31.67924368
barcelona	municipi	2010	46.12009807	33.73076521	31.92148861
barcelona	municipi	2011	46.20886281	33.95216982	32.1475168
barcelona	municipi	2012	46.31852917	34.23960098	32.40795591
barcelona	municipi	2013	46.44355269	34.56192024	32.74271641

Figura 5.1: Taula amb dades de l'edat de la població de Barcelona diferenciat per lloc de procedència del 1997 al 2023



territori	tipus_de_territori	anys	sense estudis_espanya	sense estudis_rest_a de la unio europea	sense estudis_rest_a del mon	estudis primaris, certificat d'escolaritat, egb_espanya	estudis primaris, certificat d'escolaritat, egb_rest_a de la unio europea
barcelona	municipi	1997	197608	223	1422	336150	790
barcelona	municipi	1998	191008	248	1763	334619	901
barcelona	municipi	1999	182658	252	2009	329548	1024
barcelona	municipi	2000	173543	249	2219	322495	1118
barcelona	municipi	2001	164884	242	2782	314519	1209
barcelona	municipi	2002	156463	287	4035	278937	1252
barcelona	municipi	2003	149178	285	4412	268955	1485
barcelona	municipi	2004	140291	284	4404	260172	1853
barcelona	municipi	2005	133815	273	4689	251959	2409
barcelona	municipi	2006	127031	255	4869	243675	2875
barcelona	municipi	2007	120485	389	3994	235478	5292
barcelona	municipi	2008	118910	427	3942	228562	6388
barcelona	municipi	2009	112169	415	4173	218484	6418
barcelona	municipi	2010	104108	332	3988	209733	5560
barcelona	municipi	2011	96094	312	4189	204442	4956
barcelona	municipi	2012	88608	292	4163	199553	4775

Figura 5.2: Taula amb dades de la formació de la població de Barcelona diferenciat per lloc de procedència del 1997 al 2023.

territori	tipus_de_territori	anys	habitatges_us_turistic
barcelona	municipi	2014	9606
barcelona	municipi	2015	9606
barcelona	municipi	2016	9606
barcelona	municipi	2017	9606
barcelona	municipi	2018	9581
barcelona	municipi	2019	9547
barcelona	municipi	2020	9530
barcelona	municipi	2021	9391
barcelona	municipi	2022	9402
barcelona	municipi	2023	9818
ciutat vella	districte	2014	604
ciutat vella	districte	2015	604
ciutat vella	districte	2016	602
ciutat vella	districte	2017	602
ciutat vella	districte	2018	607

Figura 5.3: Taula amb dades del nombre d'habitatges d'ús turístic a Barcelona del 2014 al 2023.

territori	tipus_de_territori	anys	poblacio.espanya	poblacio.resta de la unio europea	poblacio.resta del mon
barcelona	municipi	1997	1467063	7409	19108
barcelona	municipi	1998	1474740	8465	23713
barcelona	municipi	1999	1465518	9476	28721
barcelona	municipi	2000	1451247	10614	35401
barcelona	municipi	2001	1433986	12593	60317
barcelona	municipi	2002	1435676	18076	100725
barcelona	municipi	2003	1414749	23510	142611
barcelona	municipi	2004	1391575	29262	167345
barcelona	municipi	2005	1372461	36312	190065
barcelona	municipi	2006	1359040	43095	207381
barcelona	municipi	2007	1348056	58463	187430
barcelona	municipi	2008	1341107	69375	204228
barcelona	municipi	2009	1336490	75738	219375
barcelona	municipi	2010	1335882	75065	208498
barcelona	municipi	2011	1335652	69538	209227
barcelona	municipi	2012	1336800	72667	210194
barcelona	municipi	2013	1332044	77349	203258

Figura 5.4: Taula amb dades de la població de Barcelona del 2014 al 2023.

territori	tipus_de_territori	anys	preu_mitja_lloguer
catalunya	comunitat autonoma	2000	NA
catalunya	comunitat autonoma	2001	NA
catalunya	comunitat autonoma	2002	NA
catalunya	comunitat autonoma	2003	NA
catalunya	comunitat autonoma	2004	NA
catalunya	comunitat autonoma	2005	552.0147952
catalunya	comunitat autonoma	2006	602.7587591
catalunya	comunitat autonoma	2007	645.3627677
catalunya	comunitat autonoma	2008	677.9943908
catalunya	comunitat autonoma	2009	650.3302911
catalunya	comunitat autonoma	2010	614.54642
catalunya	comunitat autonoma	2011	599.1649943
catalunya	comunitat autonoma	2012	571.218656
catalunya	comunitat autonoma	2013	542.1825304
catalunya	comunitat autonoma	2014	539.5361585
catalunya	comunitat autonoma	2015	557.4956
catalunya	comunitat autonoma	2016	595.6554984

Figura 5.5: Taula amb dades del preu mitjà del lloguer (en euros) a Barcelona del 2000 al 2023.