



UNIVERSITAT OBERTA DE CATALUNYA (UOC)
MÀSTER UNIVERSITARI EN CIÈNCIA DE DADES (*DATA SCIENCE*)

TREBALL FINAL DE MÀSTER

ÀREA: DATA ANALYSIS I BIG DATA

Evolució de la gentrificació a la ciutat de Barcelona i anàlisi dels factors que la caracteritzen

Autor: Albert Salvador Yuste

Tutor: Anna Muñoz Bolas

Professor: Albert Solé Ribalta

Barcelona, 9 de juny de 2024

Crèdits/Copyright



Aquesta obra està subjecte 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: | Evolució de la gentrificació a la ciutat de Barcelona i anàlisi dels factors que la caracteritzen |
| Nom de l'autor: | Albert Salvador Yuste |
| Nom del/de la col·laborador/a docent: | Anna Muñoz Bolas |
| Nom del PRA: | Albert Solé Ribalta |
| Data de lliurament: | 06/2024 |
| Titulació o programa: | Màster Universitari en Ciència de Dades |
| Àrea del Treball Final: | Àrea 1 |
| Idioma del treball: | Català |
| Paraules clau | gentrificació, Barcelona, visualització |

Resum

La gentrificació s'ha convertit en un fenomen d'interès creixent en les principals ciutats del món, i Barcelona no n'és una excepció. A mesura que la ciutat experimenta transformacions socials, econòmiques i urbanístiques, la gentrificació apareix com un fenomen complex que afecta diverses àrees i comunitats, manifestant-se de manera diversa des dels barris històrics fins als perifèrics. El pes creixent del turisme a la ciutat juga un paper significatiu en aquest procés, potenciant canvis que influeixen en diversos aspectes, com l'increment del preu dels habitatges, canvis en la composició comercial dels barris, la pressió sobre els serveis públics i les infraestructures, així com el desplaçament forçat de residents que busquen alternatives en altres àrees de la ciutat. Aquest projecte té com a objectiu utilitzar dades obertes per identificar els factors que faciliten la determinació del nivell de gentrificació als diferents barris de Barcelona, creant una visualització interactiva d'aquestes variables juntament amb una representació de la seva evolució temporal. Posteriorment, s'avaluaran tècniques de classificació i de predicció del risc de gentrificació a curt termini. A través d'aquesta anàlisi, es busca entendre millor els processos de gentrificació a la ciutat i aportar un valor addicional per desenvolupar estratègies per afrontar els seus impactes socioeconòmics.

Paraules clau: gentrificació, Barcelona, visualització, indicadors, turisme, dades obertes, classificació, impactes socioeconòmics

Abstract

Gentrification has become a phenomenon of growing interest in major cities around the world, and Barcelona is no exception. As the city undergoes social, economic, and urban transformations, gentrification emerges as a complex phenomenon affecting various areas and communities, manifesting itself differently from historic neighborhoods to peripheral areas. The increasing weight of tourism in the city plays a significant role in this process, driving changes that influence various aspects, such as rising housing prices, changes in the commercial composition of neighborhoods, pressure on public services and infrastructure, as well as the forced displacement of residents seeking alternatives in other areas of the city. This project aims to use open data to identify the factors that facilitate determining the level of gentrification in different neighborhoods of Barcelona, creating an interactive visualization of these variables along with a representation of their temporal evolution. Subsequently, classification and short-term gentrification risk prediction techniques were evaluated. Through this analysis, the aim is to better understand the processes of gentrification in the city and provide added value to develop strategies to address its socioeconomic impacts.

Keywords: gentrification, Barcelona, visualization, indicators, tourism, open data, classification, socioeconomic impacts

Índex

| | |
|--|-----------|
| Resum | v |
| Abstract | vii |
| Índex | ix |
| Índex de Figures | xi |
| Índex de Taules | xv |
| 1 Introducció | 1 |
| 1 Context i justificació | 1 |
| 2 Motivació | 3 |
| 3 Objectius | 4 |
| 3.1 Objectiu principal | 4 |
| 3.2 Objectius parcials | 4 |
| 4 Impacte en sostenibilitat, ètic-social i de diversitat | 5 |
| 5 Enfoc i metodologia | 7 |
| 6 Planificació | 8 |
| 7 Resum dels productes del projecte | 11 |
| 8 Breu descripció de la resta de capítols de l'informe | 11 |
| 2 Estat de l'art | 13 |
| 1 El complex procés de la gentrificació | 13 |
| 2 El cas de Barcelona | 14 |
| 3 Detectar els processos de gentrificació | 20 |
| 4 Decisions tecnològiques | 22 |
| 3 Disseny i implementació del treball | 23 |
| 1 Obtenció i tractament de les dades | 24 |

| | | |
|----------|--|-----------|
| 1.1 | Nivell d'estudis | 24 |
| 1.2 | Estructura del domicili | 25 |
| 1.3 | Regió de naixement d'origen | 27 |
| 1.4 | Grups d'edat | 28 |
| 1.5 | Renda anual bruta | 29 |
| 1.6 | Preu del lloguer | 30 |
| 1.7 | Habitatges turístics | 31 |
| 1.8 | Unitats administratives | 32 |
| 2 | Unificació de les dades | 32 |
| 2.1 | Selecció de grups d'interès | 32 |
| 2.2 | Càlcul de les variacions interanuals | 33 |
| 3 | Càlcul del nivell de gentrificació | 34 |
| 3.1 | Determinació del nivell de gentrificació | 36 |
| 4 | Interfície visual interactiva | 37 |
| 4 | Resultats | 39 |
| 1 | Anàlisi exploratori de les dades | 39 |
| 1.1 | Nivell d'estudis | 39 |
| 1.2 | Estructura del domicili | 41 |
| 1.3 | Regió de naixement d'origen | 42 |
| 1.4 | Grups d'edat | 43 |
| 1.5 | Renda anual bruta | 45 |
| 1.6 | Preu del lloguer | 45 |
| 1.7 | Habitatges turístics | 46 |
| 2 | Nivells de gentrificació | 47 |
| 3 | Interfície visual interactiva | 53 |
| 5 | Conclusions i treball futur | 57 |
| | Bibliografia | 61 |

Índex de figures

| | | |
|-----|---|----|
| 1.1 | Fotografia aèria de Barcelona. Font: Casamitjana (2023) | 2 |
| 1.2 | Objectius de Desenvolupament Sostenible alineats amb el treball present (Nacions Unides, nd). | 6 |
| 1.3 | Diagrama de Gantt. Elaboració pròpia. | 10 |
| 1.4 | Captura de mostra de la interfície visual interactiva. S'hi observa informació referent a domicilis on hi viu una sola persona per districtes per a l'any 2023. . . | 12 |
| 2.1 | Plànol del terme municipal de Barcelona dividit per districtes (Ajuntament de Barcelona, nd). | 15 |
| 2.2 | Localització de la població arribada a Barcelona amb educació universitària segons nacionalitat. Pob. 25-49 anys, 233 AEB, 2011-2016 (López-Gay, 2018). . . | 17 |
| 2.3 | Evolució de la proporció de persones amb educació universitària entre totes les arribades a cada AEB de Barcelona. Pob. 25-49 anys, 2011-13 i 2014-16 (López-Gay, 2018). | 18 |
| 2.4 | Preus d'oferta dels locals comercials. Estimació del preu de lloguer mensual per districtes ($\text{€}/\text{m}^2/\text{mes}$)(Fresnillo, 2018). | 19 |
| 2.5 | Evolució dels preus mitjans d'oferta d'habitatges de lloguer. Catalunya. Primer trimestre de 2019 – Tercer trimestre de 2023 (Observatori Metropolità de l'Habitatge de Barcelona, 2024). | 20 |
| 2.6 | Ofertes de lloguer de temporada, en termes relatius sobre el total d'ofertes de lloguer. Barris de Barcelona. Juny de 2023 (Observatori Metropolità de l'Habitatge de Barcelona, 2023). | 21 |
| 2.7 | Resultats de l'indicador de gentrificació a Barcelona i Madrid, 2011-2017 (López-Gay et al., 2019). | 22 |
| 3.1 | Procés d'obtenció de les dades de pisos turístics mitjançant Power Query. | 31 |
| 3.2 | Selecció de les dades de treball una vegada unificades. | 35 |

| | | |
|------|---|----|
| 3.3 | Captura de mostra de la interfície visual interactiva. S'hi observa informació sobre la variació anual de les rendes brutes per districtes per a l'any 2023. | 38 |
| 4.1 | Distribució del nivell d'estudis universitaris per barris per a l'any 2023 i l'evolució anual per a tota la ciutat de Barcelona. | 40 |
| 4.2 | Variació absoluta (esquerra) i percentual (dreta) de població amb estudis universitaris per a l'any 2023 a la ciutat de Barcelona. | 40 |
| 4.3 | Distribució dels domicilis amb una persona de 18 a 64 anys per a l'any 2023 i l'evolució anual per a tota la ciutat de Barcelona. | 41 |
| 4.4 | Variació absoluta (esquerra) i percentual (dreta) de domicilis habitats per una persona de 18 a 64 anys per a l'any 2023 a la ciutat de Barcelona. | 42 |
| 4.5 | Distribució de la població europea per a l'any 2023 i l'evolució anual per a tota la ciutat de Barcelona. | 43 |
| 4.6 | Variació absoluta (esquerra) i percentual (dreta) de població estrangera d'origen europeu per a l'any 2023 a la ciutat de Barcelona. | 43 |
| 4.7 | Distribució de grups d'edat majoritaris per a l'any 2023 i l'evolució anual per a tota la ciutat de Barcelona. | 44 |
| 4.8 | Distribució de la variació percentual positiva dominant per grups d'edat (esquerra) i l'evolució d'aquestes variacions (dreta) per a l'any 2023 a la ciutat de Barcelona. | 44 |
| 4.9 | Distribució de la renda anual bruta per persona per a l'any 2023 a la ciutat de Barcelona. | 45 |
| 4.10 | Variació percentual de les rendes anuals brutes per a l'any 2023 a la ciutat de Barcelona. | 46 |
| 4.11 | Distribució del preu del lloguer per a l'any 2023 a la ciutat de Barcelona. | 46 |
| 4.12 | Variació percentual dels preus del lloguer per a l'any 2023 a la ciutat de Barcelona. | 47 |
| 4.13 | Distribució dels habitatges turístics per a l'any 2023 a la ciutat de Barcelona. | 47 |
| 4.14 | Determinació del nombre de <i>clusters</i> a utilitzar a partir del mètode de la silueta (esquerra) i el del colze (dreta) per a les dades del 2023. Elaboració pròpia. | 48 |
| 4.15 | Distribució dels nivells de gentrificació per barris pel mètode de <i>k-means</i> (esquerra) i per puntuació de gentrificació (dreta). | 49 |
| 4.16 | Distribució de les puntuacions segons el nivell de gentrificació calculat pel mètode <i>k-means</i> dels barris de Barcelona per al 2023. | 51 |
| 4.17 | Distribució de les puntuacions segons el nivell de gentrificació calculat pel mètode de puntuació de gentrificació dels barris de Barcelona per al 2023. | 52 |
| 4.18 | Captura de la pestanya 'presentació' de la pàgina web. | 54 |

| | | |
|------|--|----|
| 4.19 | Captura de la pestanya 'anàlisi anual' de la pàgina web, amb informació dels nivells d'estudi per districtes per a l'any 2014. | 54 |
| 4.20 | Captura de la pestanya 'evolució anual' de la pàgina web, amb informació dels preus del lloguer per districtes per a l'any 2016. | 55 |
| 4.21 | Captura de la pestanya 'càlcul gentrificació' de la pàgina web, amb informació dels nivells de gentrificació per a l'any 2020. | 56 |

Índex de taules

| | | |
|-----|--|----|
| 3.1 | Valors extrems i mediana de les variables utilitzades pel càlcul del nivell de gentrificació | 36 |
| 4.1 | Classificacions dels nivells de gentrificació per barris pel mètode de <i>k-means</i> i per puntuació de gentrificació per a l'any 2023. S'ordenen segons la puntuació de gentrificació. | 50 |

Capítol 1

Introducció

1 Context i justificació

La gentrificació s'ha convertit en una característica intrínseca i mantinguda de moltes ciutats del món, especialment en els mercats immobiliaris metropolitans i urbans. Això està provocant un canvi notori tant en els serveis locals com en les infraestructures que es construeixen, les quals anteriorment estaven enfocades als perfils de poblacions d'ingressos baixos ([Atkinson and Wulff, 2009](#); [Martin and Beck, 2018](#)), virant la seva orientació per ser més atractives pels nous visitants.

L'origen del mot prové de l'anglès *gentry*, una classe social històrica formada per l'alta burgesia i baixa noblesa britànica, i se li atribueix a la sociòloga Ruth Glass on en un estudi de l'any 1964 l'introduí per analitzar el procés de reapropiació del centre de Londres per part de la classe social esmentada ([Glass, 1964](#)). No obstant això, el terme gentrificació no és només un terme relacionat amb la variació de la població, sinó que inclou aquells altres canvis relacionats amb l'espai on es produeix. Alhora, l'arribada d'una població més benestant acostuma a provocar un encariment generalitzat en altres productes de servei o necessitat, no necessàriament vinculats directament amb el turisme, on el més àmpliament afectat és el preu dels immobles.

Aquests canvis sovint provoquen que els antics residents es sentin silenciosament forçats a desplaçar-se a d'altres barris o municipis on puguin assumir una vida quotidiana més adaptada al seu poder adquisitiu. A més, també provoquen augments de temps de desplaçament al treball i els llogaters insegurs s'enfronten a desnonaments o a desplaçaments del mercat ([Atkinson and Wulff, 2009](#)). Alhora, els nous requisits econòmics per poder establir-se a les zones gentrificades limiten el perfil del nou resident del barri, el qual acostuma a ocupar-lo un nou perfil que disposa d'una capacitat financera per assumir-los.

Entre els períodes clau que han influït en la gentrificació a la ciutat de Barcelona, el desenvolupament urbà del s. XIX i principis del XX, on un creixement demogràfic important vinculat



Figura 1.1: Fotografia aèria de Barcelona. Font: [Casamitjana \(2023\)](#).

amb la Revolució Industrial ([Maluquer de Motes, 2019](#)), va transformar la fisionomia de la ciutat amb l'aparició de barris industrials i obrers al costat del nucli històric. Aquest canvi va provocar una divisió territorial entre els barris de classe treballadora i les zones residencials de la burgesia emergent.

Fixant-nos en èpoques més recents, Barcelona està de moda. És una frase que s'ha sentit recurrentment, especialment si hom es fixa en la plana internacional, fent que s'hagi transformat en una ciutat turística en els últims 30 anys, causada principalment per dues referències clau: els Jocs Olímpics de Barcelona i la creació del consorci Turisme de Barcelona ([López Palomeque, 2015](#)). No és quelcom diferent del que s'observa en el nombre de vols i passatgers de l'aeroport Josep Tarradellas Barcelona-El Prat, el segon a l'Estat en nombre de passatgers després de l'aeroport de la capital espanyola el 2023 ([Aena, 2024](#)), els quals no només estan recuperant els nombres pre pandèmics, si no que es mantenen en un ascens continu. Aquesta última etapa de transformació de la ciutat, amb la promoció del turisme i la revitalització de barris degradats, han sigut un precursor molt important de la gentrificació.

L'impacte del turisme a Barcelona és quelcom que es fa evident quan passeges pels seus barris i carrers. Tant pel perfil de població que s'hi pot observar com per la tipologia de comerços que s'estableixen al múltiples locals comercials, on es percep un increment d'aquells

enfocats a satisfer les necessitats i desitjos del visitant de la ciutat comtal. Barris com el Raval, Poblenou i el Born han presentat una revitalització significativa, fet que també ha provocat una gentrificació accelerada, desplaçant els residents locals i transformant la identitat dels barris.

Arrel de la gentrificació present a la ciutat, en aquest projecte es presenta un estudi on s'analitzen aquells barris que estan actualment immersos en aquest procés, identificant-ne les causes i extraient-ne indicadors que aportin una caracterització i predicció evolutiva a curt termini. Per arribar a aquesta fi, es crearà una visualització interactiva amb diferents característiques escollides a partir de la lectura de l'estat de l'art, i s'aplicaran tècniques de classificació i predicció per avaluar la seva potencial evolució.

2 Motivació

Durant els meus darrers anys de formació acadèmica i parcialment laboral, m'he vist envoltat de mapes que representen tot tipus de variables, la majoria d'elles d'origen físic, geogràfic i meteorològic. Abans no me'n adonés em vaig acostumar a intentar establir relacions potencials entre elles amb les eines que en cada moment coneixia i disposava. Sovint encara veig mapes de temàtiques molts diverses i m'acostumen a despertar cert interès en les distribucions de valors, icones, colors, etc., que s'hi observen i, encara que sigui de manera molt preliminar i sense arribar a elaborades conclusions, intento establir certes relacions ràpides entre aspectes socials, zones geogràfiques, característiques físiques, evolucions temporals, entre d'altres.

Durant el curs del Màster en Ciència de Dades he après noves eines, tant de representació de dades geoespacionals en l'assignatura d'Anàlisi de Dades Geoespacionals, com d'altres de caire estadístic, de classificació, modelització, etc., repartides en tantes altres assignatures, així com evidentment de programació, ampliant substancialment i fonamentant els coneixements que fins ara disposava, la majoria apresos segons les necessitats de cada moment.

La combinació tant dels nous coneixements assolits com de les inquietuds d'interpretació de mapes i gràfics em despertaren un interès addicional en elaborar nous mapes per mi mateix, que resultin interessants, agradables d'interpretar i, a poder ser, que aportin algun valor afegit.

La temàtica de la gentrificació, en aquest cas enfocada a Barcelona, ha sigut quelcom que em va cridar l'atenció doncs, tot i que coneixia superficialment la problemàtica, no n'era conscient, fins que he anat llegint la literatura, de com de complexa era, el nombre elevat de factors que potencialment hi poden influir, la dificultat d'obtenir unes conclusions clares i definitives, possiblement exposades a una subjectivitat en la interpretació, etc., a més de la afectació que té a molts altres indrets del món, principalment en les grans ciutats. Així doncs, poder realitzar un treball d'aquesta especialitat em resulta interessant per volcar els nous coneixements apresos els darrer anys, per finalment presentar un producte interessant i enriquidor.

3 Objectius

Aquest treball té com a objectiu principal la identificació i avaluació d'aquells factors econòmics, municipals i de població que permeten analitzar el procés complex de la gentrificació, en aquest cas a la ciutat de Barcelona. Alhora, no només s'aspira a establir aquestes variables a una distribució espacial, sinó també temporal. D'aquesta manera s'espera poder comprovar quines són les zones i/o barris amb més grau de gentrificació i, conjuntament, realitzar una previsió tant de la seva evolució com de detectar quins altres barris són potencialment susceptibles d'estar exposats a un procés de gentrificació.

Les dades que s'utilitzaran per a l'elaboració d'aquest treball seran eminentment de caràcter públiques i obertes durant el moment de la seva obtenció, consegüentment qualsevol lector interessat pot obtenir-les de manera lliure i pública.

Per tal de representar els resultats obtinguts en aquest treball d'investigació, es crearà un producte de visualització interactiva que faciliti la seva interpretació i posterior explicació dels resultats observats. Aquesta serà també d'accés públic i obert de manera que hom pugui observar les imatges proporcionades en aquest treball, així com d'altres que potser no s'hi han inclòs amb les variables que allà s'hi disposin.

3.1 Objectiu principal

L'objectiu principal d'aquest treball és la determinació dels factors més influents en la classificació del nivell de gentrificació en els diferents barris de Barcelona, així com la seva evolució en el temps, representant-ne els resultats en una visualització interactiva. Posteriorment, es desenvoluparà una classificació dels barris de la ciutat segons el nivell de gentrificació, seguit de l'aplicació de tècniques de predicció per observar la variació al llarg del temps.

3.2 Objectius parcials

- Cerca de l'estat de l'art referent a la gentrificació.
- Identificació de bancs de dades públics que continguin la informació necessària per a elaborar el treball present.
- Classificació, neteja i anàlisi de les dades disponibles.
- Creació d'una visualització interactiva que faciliti l'anàlisi i explicació de les dades, així com dels resultats obtinguts.
- Desenvolupament d'una classificació per barris segons el nivell de gentrificació i posterior aplicació de tècniques de predicció.

- Sintetització de la feina realitzada al llarg de l'elaboració d'aquest projecte en la redacció d'una memòria.
- Realització d'una presentació i posterior defensa del projecte.

4 Impacte en sostenibilitat, ètic-social i de diversitat

La Universitat Oberta de Catalunya, en compromís amb la Competència de Compromís Ètic i Global, cerca actuar de manera honesta, ètica, sostenible, socialment responsable i respectuosa amb els drets humans i la diversitat, tant en la pràctica acadèmica com en la professional, i promou dissenyar solucions per millorar aquestes pràctiques. Per aconseguir-ho s'aborden tres grans dimensions: la sostenibilitat, el comportament ètic i responsabilitat social, i la diversitat i drets humans.

Aquests objectius s'alineen amb els ODS (Objectius de Desenvolupament Sostenible) per tal de ser utilitzats com a guia tant en general com en les investigacions que es desenvolupin. Així, el treball present cerca enquadrar-se en els següents ODS:

- **ODS01:** Fi de la pobresa. Extesa en moltes dimensions, algunes de les causes freqüents que les causen són la falta de feina, l'exclusió social i l'alta vulnerabilitat en front a múltiples aspectes.
- **ODS05:** Igualtat de gènere. Fonamental per crear una societat sana, pròsper i sostenible en tots el àmbits, es requereixen unes inversions i reformes polítiques integrals per aconseguir aquest objectiu transversal.
- **ODS08:** Treball decent i creixement econòmic. Promou un creixement econòmic inclusiu i sostenible, i treball decent per a tothom, reduint el treball informal, la desigualtat en el mercat laboral, promoció d'entorns segurs i protegits, així com millorar l'accés a serveis financers igualitaris.
- **ODS9:** Indústria, innovació i infraestructura. Persegueix l'assoliment d'infraestructures resilents, una industrialització sostenible i el foment de la innovació, presentant una igualtat d'oportunitats d'accessibilitat per a totes les persones.
- **ODS10:** Reducció de les desigualtats. Promovent una vida digna per a tots, es motiven polítiques econòmiques i socials universals i a parar atenció a les necessitats de les comunitats més desafavorides i marginades.
- **ODS11:** Ciutats i comunitats sostenibles. Entre d'altres, cerca aconseguir que les ciutats i els assentaments humans siguin inclusius, segurs, resilents i sostenibles.



Figura 1.2: Objectius de Desenvolupament Sostenible alineats amb el treball present ([Naciones Unidas, nd](#)).

El procés de la gentrificació es defineix de manera que múltiples aspectes i factors hi formen part, on els que presenten una vessant sociodemogràfica hi tenen un pes molt important. Com s'ha observat fins ara, una de les seves conseqüències immediates és el del desplaçament de la població econòmicament més desfavorida dels barris que la pateixen, al ser substituïda per una població més benestant.

Partint d'aquesta premissa, aquest treball enfoca algun dels seus objectius finals de manera que es posi especial atenció en el perfil de població més vulnerable, on entrant amb el ODS01, la població més pobre és potencialment la més afectada i, malauradament encara a dia d'avui, les desigualtats de gènere hi són presents, agreujant la situació de les dones en aquest procés, posant el ODS05 en el punt de mira. En conjunt de les dues anteriors, així com englobant altres perfils de població que també en puguin sortir perjudicats, el ODS10 cerca que aquestes desigualtats es redueixin.

Alhora, la identificació dels factors de gentrificació, així com possibles projeccions a curt termini, persegueixen la possibilitat de promoure polítiques adequades perquè les conseqüències negatives per a la població es redueixin al màxim. Així, en consonància també amb el ODS08 (on es promou un creixement econòmic i laboral igualitari, permetent una capacitat monetària justa per a tothom), el ODS09 (amb una industrialització i aplicació d'infraestructures igualitària i sostenible, amb igualtat d'oportunitats per a la població) i, de manera global, el ODS11 (promovent una ciutat amb creixement sostenible per als seus habitants), aquest treball cerca ser una eina per obrir noves línies de treball que finalitzin en aplicacions d'eines polítiques i socials que redueixin les desigualtats econòmiques i d'oportunitat entre totes les part involucrades.

5 Enfoc i metodologia

Seguint amb els objectius descrits a l'apartat d'[objectius](#), el destí final d'aquest treball és la representació en forma de visualització interactiva d'aquells factors que caracteritzen i defineixen el procés de gentrificació a la ciutat de Barcelona. S'analitzaran variables com ara els preus dels lloguers, les condicions socioeconòmiques, el perfil i origen de la població, estructura per edats, localització i característiques de la zona, entre d'altres. Són elements que inicialment poden ser sospitosos de tenir alguna influència en la determinació de l'objectiu final, però òbviament s'aniran modulant, corregint i/o ampliant a mida que la investigació de l'estat de l'art es vagi ampliant.

L'estat de l'art no només assenyalava algunes de les variables amb més pes a l'hora d'enfocar els esforços de la investigació, també serà important valorar les eines que s'han utilitzat en diferents estudis, quins són els que han aportat resultats més fidedignes a la realitat, o tanmateix si orienten cap a possibles altres tècniques d'anàlisi que no s'han aplicat fins el moment en aquesta línia d'investigació.

Un cop o en paral·lel s'han determinat quines són aquelles variables que s'utilitzaran en aquest projecte, és important conèixer un o múltiples orígens de dades que proporcionin les dades necessàries per poder avançar adequadament. Un exemple molt clar és el del projecte de Open Data BCN, el servei de dades obertes de l'Ajuntament de Barcelona, el qual disposa conjunts de dades de diferent tipologia, entre les que es destaquen les de territori, població, ciutat i serveis, administració i economia i empresa. Òbviament cal destacar que no és la única font de dades que pot ser d'interès a mida que el projecte vagi creixent, però fins al moment actual aquesta serà una de les fonts principals del treball en qüestió.

Alhora, és utòpic creure que les dades que s'obtinguin ja estiguin perfectament adequades i preparades a les necessitats del projecte, per lo que serà important realitzar tasques d'anàlisi d'aquestes, com observar les tipologies dels diferents camps, descartar potencials valors anòmals o clarament erronis, homogeneïtzar valors quan es necessiti, entre d'altres processos de tractament de dades. També serà d'utilitat realitzar anàlisis estadístics preliminars per conèixer els comportaments dels diferents camps amb els que es treballarà, per així agilitzar la feina posterior amb aquests. Tota aquesta feina s'implementarà amb *scripts* de R, utilitzant múltiples funcions que allà hi són disponibles.

Un dels altres passos serà la classificació d'aquells barris de la ciutat segons el seu nivell de gentrificació. Per realitzar-ho, s'aplicaran tècniques d'aprenentatge automàtic com *k-means*, de manera que faciliti la identificació d'aquelles variables significatives en la determinació de l'objectiu de caracteritzar la gentrificació. Alhora, s'aplicaran tècniques de predicció de manera que proporcioni una visió de com aquestes variables poden variar al llarg del temps, assenyalant quins són els altres barris que poden estar en risc de gentrificació a curt termini.

Una vegada les dades ja s'han adaptat als requeriments establerts, i es comencin a obtenir el primers resultats, es procedirà a realitzar representacions visuals interactives amb una eina encara pendent de determinar, on les eines com CARTO, ArcGis Online i la llibreria *Shiny* de R són potencials candidates de ser utilitzades per a aquesta fi. En elles, no només es representaran aquelles variables espacials, si no també la seva evolució temporal.

6 Planificació

Per a la realització d'aquest projecte s'han utilitzat els següents recursos:

- Programari:
 - **R** versió 4.3.2 ([R Core Team, 2023](#)): Llenguatge de programació
 - **RStudio** versió 2023.6.0.421 ([Posit team, 2023](#)): Entorn de desenvolupament integrat per al llenguatge de programació R
 - **Overleaf**: Editor \LaTeX
 - **Opera GX**: Navegador web
 - **Microsoft PowerPoint**: Programa de realització de presentacions
 - **Github**: Servei d'allotjament de repositoris git
 - **GanttProject** versió 3.3.3305: Programari d'administració de projectes
- Maquinari:
 - **Placa mare**: Gigabyte Technology Co. Ltd. Z97X-Gaming 3
 - **Processador**: Intel® Core™ i7-4790K @4.00GHz
 - **Memòria RAM**: 16.0 GB
 - **Targeta gràfica**: NVIDIA GeForce GTX 970 i Intel® HD Graphics 4600
 - **Sistema operatiu**: Windows 10 Pro

La planificació d'aquest treball es divideix en 5 mòduls diferents, on cadascun d'ells finalitza amb l'entrega de les respectives activitats avaluable, o bé amb la finalització i defensa del projecte en el cas de l'últim mòdul. Per a cadascun d'aquests mòduls s'indica la data de finalització fixada en el projecte, alhora que es contemplen un seguit de tasques associades:

- **Mòdul 1 - Definició i planificació del treball final** (12 de març del 2024)
 - Investigació general del projecte

-
- Descripció i justificació de la proposta
 - Explicació de la motivació personal
 - Definició dels objectius (principals i secundaris)
 - Descripció de la metodologia emprada
 - Planificació del projecte
 - Proposta del títol
 - Resum de la proposta (*abstract*) i definició de les paraules clau (*keywords*)
 - Bibliografia (s'anirà editant al llarg del projecte)
- **Mòdul 2 - Estat de l'art o anàlisi de mercat del projecte** (26 de març del 2024)
 - Cerca de bibliografia ad hoc al projecte
 - Anàlisi els estudis més propers a l'objectiu principal
 - Replantejament dels objectius parcials si s'escau
 - Identificació de les dades a obtenir
 - Definició de les eines a utilitzar
- **Mòdul 3 - Disseny i implementació del treball** (21 de maig del 2024)
 - Preparació de l'entorn de desenvolupament
 - Obtenció de dades obertes necessàries pel projecte
 - Anàlisi de les dades i tractament estadístic
 - Incorporació de dades en entorns de visualització interactiva
 - Explotació de la visualització per extreure'n coneixement
 - Disseny d'un model de classificació dels barris
- **Mòdul 4 - Redacció de la documentació del TFM** (18 de juny del 2024)
 - Redacció de la memòria al llarg de tot el projecte
 - Extreure conclusions del projecte realitzat
 - Bibliografia final
 - Identificació dels punts claus del projecte
 - Elaboració de la presentació i el vídeo corresponent
- **Mòdul 5 - Defensa del projecte** (7 de juliol del 2024)

– Exposició del projecte realitzat davant un tribunal

A la figura 1.3 es mostra el diagrama de Gantt amb la relació entre les fases i les tasques que es preveuen durant el projecte.

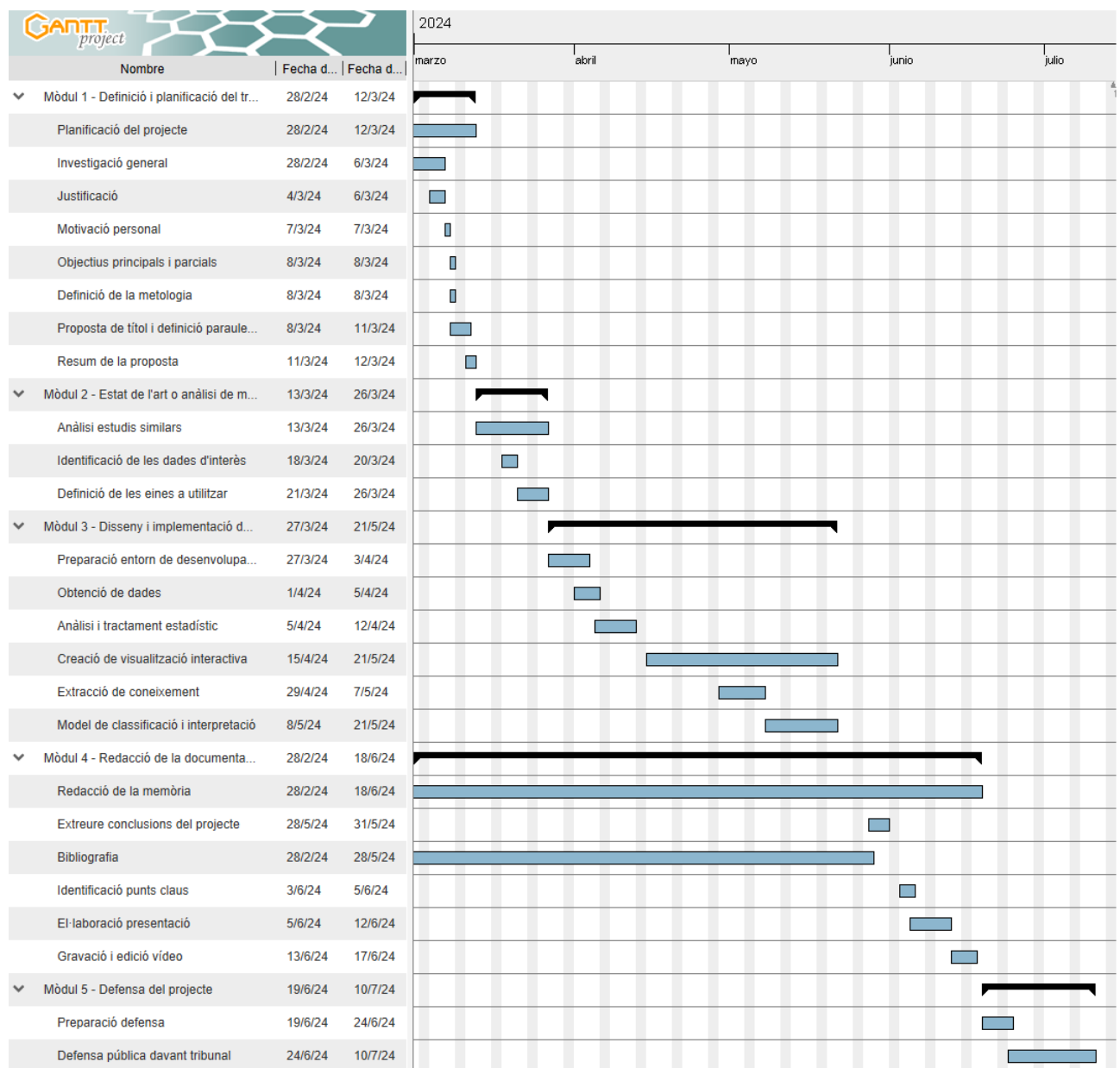


Figura 1.3: Diagrama de Gantt. Elaboració pròpia.

7 Resum dels productes del projecte

Com a resultat d'aquest projecte, s'han obtingut un seguit de productes, els quals es llisten a continuació:

- Un conjunt de fitxers en format ZIP, CSV i XLSX, obtinguts del portal de Open Data BCN i de la Oficina Municipal de Dades del Departament d'Estadística i Difusió de Dades (ambdúes de l'Ajuntament de Barcelona), i del Departament d'Habitatge de la Generalitat de Catalunya. En ells s'hi conté tota la informació de les diferents variables que s'utilitzen en aquest projecte.
- Dos repositoris de Github amb tots els elements de programació utilitzats durant el treball:
 - Enllaç a Github amb tots els *scripts* utilitzats per a la obtenció de les dades, tractament i anàlisi (https://github.com/ASYuste/TFM_UOC_Gentrificacio). El contingut es separa en:
 - * Carpeta `data`, on s'emmagatzemen les dades origen i del resultats obtinguts en les diferents etapes del treball.
 - * Carpeta `bin`, on es recullen els diferents *scripts* per a la obtenció de la majoria de dades utilitzades.
 - * Arxius `*.md`, on es detallen tots el processos de tractament i transformació de les dades per assolir els resultats finals. Al llarg del treball es mencionen cadascun d'aquests arxius per obtenir-ne informació addicional.
 - Enllaç a Github amb tots els *scripts* que gestionen la creació de la interfície visual interactiva (https://github.com/ASYuste/TFM_UOC_Gentrificacio_Visualitzador). La carpeta `data` conté la mateixa informació que la del repositori anterior.
- Interfície visual interactiva, publicada al portal *shinyapps.io* i disponible al següent enllaç (https://asyuste.shinyapps.io/uoc_tfm_shiny/), on es representen les dades utilitzades i els resultats obtinguts d'aquest treball. A la figura 1.4 es pot observar una captura d'exemple d'aquest visualitzador.

8 Breu descripció de la resta de capítols de l'informe

El següent capítol de la memòria contempla l'estudi de l'estat de l'art de la matèria, on es descriuen alguns estudis referents als processos i dinàmiques de la gentrificació, descripció de l'entorn de treball, i estudis relacionats amb la classificació del nivell de gentrificació.

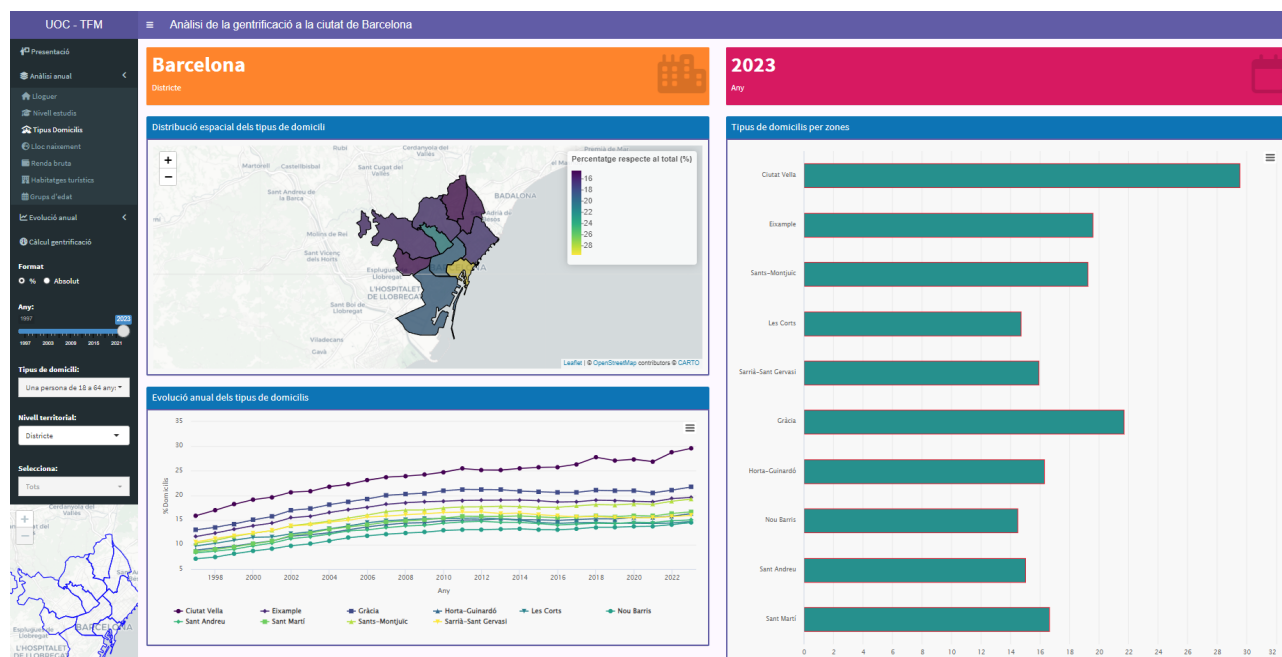


Figura 1.4: Captura de mostra de la interfície visual interactiva. S'hi observa informació referent a domicilis on hi viu una sola persona per districtes per a l'any 2023.

En la següent part es detalla el [disseny i la implementació del treball](#), on es defineixen totes les decisions preses per a un correcte desenvolupament del treball en quant a tractament i anàlisi de les dades, així com de les tècniques de classificació aplicades i del procediment de creació de la interfície visual interactiva.

El segueix el capítol dels [resultats](#), on s'analitza el comportament de les diferents variables estudiades geogràfica i temporalment, els resultats de l'aplicació de les tècniques de classificació i l'anàlisi pertinent. A més, es presenta la visualització interactiva amb els diferents elements de consulta i resultats que s'hi representen.

Finalment, al capítol de [conclusions i treball futur](#) es realitza un balanç de la feina realitzada al llarg del projecte, amb l'avaluació dels resultats obtinguts i dels objectius assolits, alhora que es detallen potencials punts de millora i d'ampliació a aplicar a potencials treballs futurs.

Capítol 2

Estat de l'art

Aquest treball d'investigació consisteix en la elaboració d'una visualització interactiva d'aquells factors que es determinin més influents en la identificació de la gentrificació a la ciutat de Barcelona. Aquesta visualització permetria ressaltar i facilitar la identificació de les variables significatives que, potencialment, s'utilitzarien per aplicar tècniques d'aprenentatge automàtic que permetin categoritzar els diferents barris de Barcelona segons nivells de gentrificació similars. Finalment, es valora la realització d'una predicció de com poden variar aquests nivells a curt termini, o inclús si poden aparèixer nous barris potencialment afectats.

Amb aquest objectiu clar, esdevé imprescindible conèixer la situació actual a la ciutat comtal enfront a aquest procés, obtenint-ne informació diversa i plural. Degut als múltiples factors que poden influir en la situació de la gentrificació actual, en aquest capítol es procurarà proporcionar al lector un anàlisi dels estudis, discussions i memòries que han abordat aquesta problemàtica de maneres diverses, tant analitzant la qüestió com conèixer què la causa i/o identifica.

1 El complex procés de la gentrificació

A partir de la introducció d'aquest treball, es pot arribar a la conclusió que la gentrificació és un procés complex i multifactorial, on a més també pot tenir una forta dependència segons la causa que la provoqui, especialment degut a la complexitat dels canvis al llarg del temps.

Conèixer els factors o indicadors per categoritzar la gentrificació és quelcom que s'ha analitzat des de diferents punts de vista. Per començar a realitzar algun enfoc i prenent una referència inicial en alguns estudis dels Estats Units, Canadà i la ciutat de Londres, en ells s'aborden diferents aspectes que es presenten com a rellevants per a la identificació de la gentrificació. Entre d'altres, s'analitzen aspectes com ara la naturalesa de la població (Atkinson, 2000), les condicions socioeconòmiques (e.g., Ding et al., 2016; Martin and Beck, 2018), la tipologia dels habitatges (e.g., Atkinson, 2000; Atkinson et al., 2011), la distribució per edats

(e.g., [Atkinson, 2000](#); [Atkinson et al., 2011](#); [Moos et al., 2019](#)), les mobilitats i rotacions dins de les mateixes ciutats ([Atkinson et al., 2011](#)), així com les característiques de les vivendes i de la zona afectada ([Freeman, 2005](#)).

No obstant, més enllà de caracteritzar la gentrificació amb paràmetres quantitius, es planteja una altra problemàtica, i és la qüestió temporal. Recordant que la gentrificació no és una situació puntual, sinó un procés de canvi dels barris segons la composició social i estructural, realitzar una instantània de la situació actual sovint no serà suficient per obtenir la informació completa.

Segons [Antón-Alonso et al. \(2018\)](#), aquests processos de substitució impliquen entendre que tant poden ser intensos i concentrats en períodes relativament curts de temps, com d'altres en què el procés es realitzi en un període més dilatat en el temps i, en conseqüència, de menys intensitat. Alhora, diferents casuístiques temporals poden afectar diferent segons les característiques dels barris, doncs no impactarà de la mateixa manera en un barri degradat de classe treballadora com en un barri amb millors infraestructures on les classes mitjanes hi són més dominants, on a més també poden existir múltiples fases intermèdies en què aquests canvis socials avancen, es paralitzen o inclús s'acceleren. Paral·lelament, dos barris poden estar en respectius processos de gentrificació, però segons el seu estadi evolutiu (fase incipient, intermèdia o avançada), la localització (centre o zona perifèrica), característiques dels habitatges i de l'entorn, la composició social del barri, agents implicats (govern local i/o regional, empreses d'inversió o promotores immobiliàries, etc.), entre d'altres, poden presentar característiques, identificadors i conseqüències diferents.

Adicionalment, també és clau la identificació del motiu principal sobre el què s'està produint el procés de gentrificació, doncs es presenta un ventall de possibilitats. Entre les més destacades, sovint s'aglutinen entre les causades per una regeneració urbana residencial, una reconversió industrial cap a una més residencial, compra d'habitatges per particulars, substitució de la població llogatera, gentrificació comercial, gentrificació turística, entre d'altres.

2 El cas de Barcelona

Barcelona, la capital de Catalunya, és una ciutat formada per un conjunt de 10 districtes, cadascun amb diversos barris que tenen un perfil sociocultural divers. Per ordre alfabètic es llisten tot seguit i es poden observar a la figura 2.1:

- **Ciutat Vella:** Comprès pels barris de la Barceloneta, el Gòtic, el Raval i Sant Pere, Santa Caterina i la Ribera, és el districte més antic de la ciutat amb una gran riquesa històrica i cultural, que el condueix a ser un lloc d'interès turístic.

de Palomar, la Sagrera i la Trinitat Vella.

- **Sant Martí:** Inclou els barris del Clot, el Besòs i el Maresme, el Camp de l'Arpa del Clot, Poblenou, Diagonal Mar i el Front Marítim del Poblenou, el Parc i la Llacuna del Poblenou, Provençals del Poblenou, Sant Martí de Provençals, la Verneda i la Pau i la Vila Olímpica del Poblenou, caracteritzat per àrees residencials, industrials i de negocis.
- **Sants-Montjuïc:** Format pels barris de la Bordeta, la Font de la Guatlla, Hostafrancs, la Marina del Port, la Marina del Prat Vermell, Sants, Poble-sec i Sants-Badal, és un districte amb una eminenment combinació de zones residencials i d'entreteniment.
- **Sarrià-Sant Gervasi:** Considerada una de les àrees més benestants de la ciutat, el formen els barris de Sarrià, el Putxet i el Farró, Sant Gervasi - la Bonanova, Sant Gervasi - Galvany, les Tres Torres i Vallvidrera, el Tibidabo i les Planes. És tradicionalment una zona residencial de luxe amb nombroses botigues exclusives.

Un cop definida la distribució territorial de la zona d'estudi d'aquest treball, es segueix amb la seva conjuntura en front de la gentrificació. Així, la situació a la ciutat de Barcelona no és, com ja s'ha vist, aliena a aquesta problemàtica que ja s'ha arribat a classificar com a planetària ([Merrifield, 2013](#)). Nombrosos estudis han procurat abordar la problemàtica a la ciutat barcelonina des de diferents punts de vista, de la mateixa manera que s'ha analitzat arreu d'altres regions del món.

Alguns exemples d'estudi són els de [Antón-Alonso et al. \(2018\)](#), on prenent de referència altres estudis com els de [Freeman \(2005\)](#) i [Martin and Beck \(2018\)](#) en què descriuen els barris potencialment gentrificables a partir de les variables d'antiguitat del parc residencial i de ingressos mitjans de les llars (relacionats amb l'aspecte residencial i social de la zona), presenta una evolució, en el període 1991-2011 i en la regió de Barcelona i la seva àrea metropolitana, de les ràtios entre classes mitjanes i treballadores, així com una classificació dels barris gentrificables segons les trajectòries de grups de població. Entre les seves conclusions, destaca les característiques i tinença d'habitatge, la centralitat del barri, la presència d'equipaments culturals, una intervenció urbanística intensa, entre d'altres, així com una elevada presència de població gran que faciliti el reemplaçament poblacional, com a elements clau per a que els processos de gentrificació acabin prosperant.

Un altre factor que s'ha observat en diversos estudis és la matèria del nivell educatiu dels nou vinguts a la ciutat. El perfil de població amb estudis superiors ha tingut un cert pes en la població barcelonina ([López-Gay, 2008](#)), però s'ha accentuat notòriament degut a la reducció de la immigració internacional de perfil socioeconòmic baix, provocant que en el període 2011-2016 la població passés d'un 51% amb estudis superiors fins a un 58% al final del període (partint d'un valor al voltant del 38% el 2001).

A l'estudi de López-Gay (2018) es destaca en el mateix període una evolució dels fluxes d'entrada i sortida dels barris de Barcelona, tot classificant entre edat, sexe, país de procedència i nivell educatiu. Separant en aquesta ocasió per la nacionalitat de la població entrant, a la figura 2.2 es destaquen les localitzacions d'arribada del perfil poblacional amb educació universitària entre els 25 i 49 anys, on s'observa una preferència general pels barris més cèntrics. Amb la salvetat que els que posseeixen la nacionalitat espanyola es reparteixen entre tots el barris de l'Eixample i cap a Gràcia, Sants-Montjuïc, Sarrià-Sant Gervasi i Sant Martí, els de nacionalitat estrangera destaquen per la seva presència dominant en els barris del districte de Ciutat Vella, els barris de la Dreta del Eixample, Sagrada Família i Poble Sec.

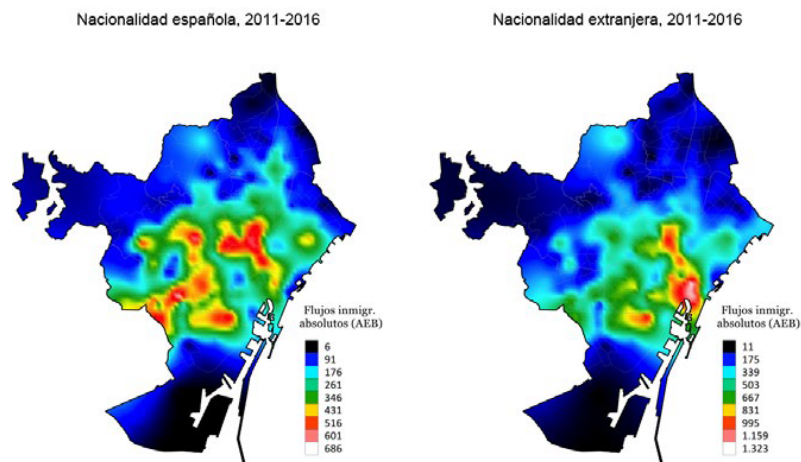


Figura 2.2: Localització de la població arribada a Barcelona amb educació universitària segons nacionalitat. Pob. 25-49 anys, 233 AEB, 2011-2016 (López-Gay, 2018).

A més, a la figura 2.3 també es pot destacar com ha evolucionat la proporció de persones amb educació superior dividint el període d'estudi en dos (2011-2013 i 2014-2016), destacant-ne les diferències o evolució entre elles. D'aquesta manera, els barris del Raval, la Barceloneta, Poble Sec, Hostafrancs, Sants, Sant Antoni, la Nova Esquerra de l'Eixample, Montbau, Sant Andreu, la Sagrega, Sant Martí de Provençals i Poblenou, són els exemples més clars de zones amb canvis més intensos del perfil de nous nadius amb estudis superiors. Es destaquen a banda zones com la Dreta de l'Eixample, el barri Gòtic, Sant Pere, Santa Caterina i la Ribera, Gràcia i Sarrià-Sant Gervasi i les Corts on el nivell d'estudis superiors ja era inicialment més elevat, i els barris més vulnerables al nord i sud de la ciutat on s'hi observa un increment més lleuger respecte a d'altres zones. Finalment, acaba conclouent que en més del 80% de la ciutat de Barcelona la població que arriba està més formada que la que abandona, fet que provoca que aquest fluxes migratoris i residencials tenen una alta capacitat transformadora sobre la composició de la nova població, destacant especialment els barris del Raval, la Barceloneta i el Poble Sec, delegant a la resta una transformació social més lenta. Finalment remarca que a més s'ha de tenir present

l'estructura de població actual, la qual afavoreix un canvi a favor dels més joves i educats a tota la ciutat.

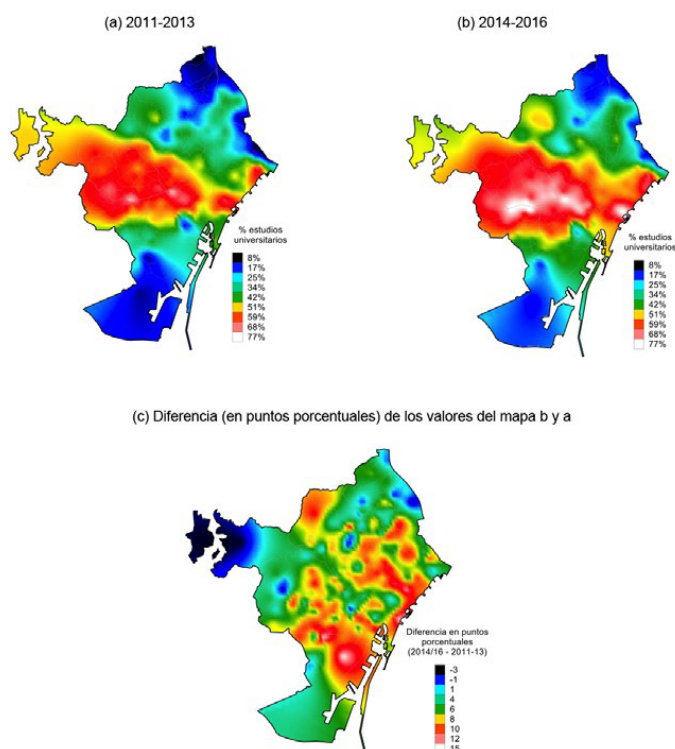


Figura 2.3: Evolució de la proporció de persones amb educació universitària entre totes les arribades a cada AEB de Barcelona. Pob. 25-49 anys, 2011-13 i 2014-16 (López-Gay, 2018).

Un altre aspecte que s'ha discutit a la introducció d'aquest treball és el pes que té el turisme en el complex procés de la gentrificació. Estudis com [Fresnillo \(2018\)](#) planteja el terme de gentrificació comercial, especificant-se en aquells processos de desplaçament en els barris gentrificats dels establiments amb productes assequibles per a la població habitual, per uns altres més enfocats a consumidors de classe mitjana i alta. Entre d'altres, destaca el cas dels mercats tradicionals, els quals estan modificant usos, preus i productes per adreçar-se a un altre perfil de població de més alt nivell adquisitiu, així com el destinat al turisme ([Cocola-Gant, 2015](#)). A més, aquest procés repercuteix en una espiral especulativa per als comerços de proximitat, els quals sovint es troben amb la impossibilitat de fer front als increments de preus dels lloguers al final dels contractes vigents. Tot i que amb un descens derivat de la crisi financera del 2008 a Espanya, a la figura 2.4 es pot observar com, a partir del inici de mil·lenni, els preus per m²/mes no han parat de pujar i, tot i que la sèrie s'interromp a l'any 2011, els valors del 2016 indiquen que el lloguer mitjà dels locals comercials a la ciutat de Barcelona s'ha incrementat un 170.23% en només 5 anys, on els districtes de Nou Barris, l'Eixample i Ciutat Vella en són els més perjudicats.

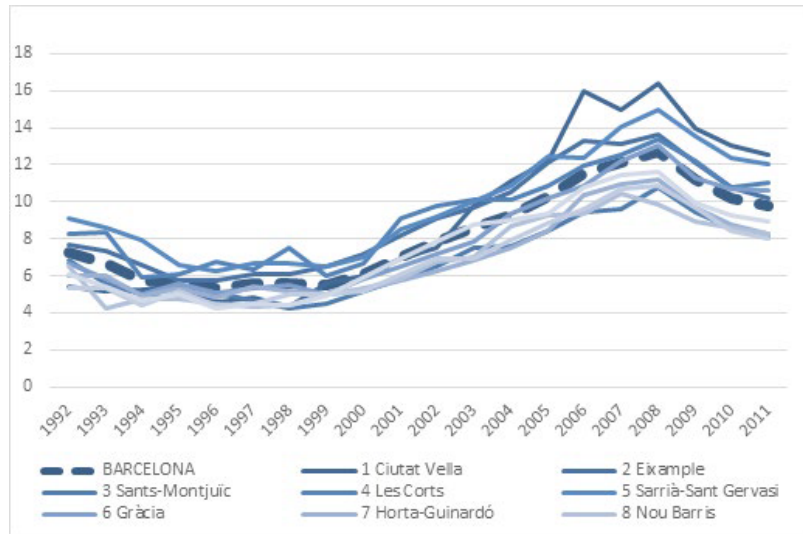


Figura 2.4: Preus d'oferta dels locals comercials. Estimació del preu de lloguer mensual per districtes (€/m²/mes) (Fresnillo, 2018).

La problemàtica no obstant no és tan sols per als locals comercials, ja que l'afectació als actuals residents no li és indiferent. El turisme i la gentrificació es presenten en entorns similars i es retroalimenten entre ells (Cocola-Gant, 2015), provocant que l'arribada del turisme n'augmenta el valor immobiliari, tant el comercial anteriorment explicitat com el residencial, remarcant l'expulsió d'un perfil de població amb un poder adquisitiu més reduït. Estudis com el de Cocola-Gant (2018) identifiquen tres punts d'enfoc vinculades al turisme a la ciutat comtal. El primer relaciona la pressió que exerceixen els allotjaments turístics en el mercat residencial. L'allotjament turístic està provocant un augment clar dels preus dels habitatges, provocant alhora una substitució directa de la població resident per una de flotant. A partir de dades recents, tal i com s'observa a Observatori Metropolità de l'Habitatge de Barcelona (2024), els preus mitjans de l'oferta de lloguers a Catalunya als portals creixen un 20.8% interanual, on el pes de la ciutat de Barcelona impacta amb un 23.7%. Aquesta afectació també provoca un desequilibri entre l'oferta i la demanda de lloguer, on la demanda de pisos per sota dels 1.000€ està al 58.3%, per una oferta de tan sols el 10.9%. En la figura 2.5 es pot veure l'evolució d'aquests preus tant a la capital catalana com a la resta de Catalunya. En paral·lel, a la figura 2.6 s'hi representa el pes del lloguer de temporada per al mes de juny de 2023.

El segon dels punts que destaca Cocola-Gant (2018) és el de la substitució del teixit comercial abans esmentada. Finalment, el tercer punt està associat a la pressió continuada de les activitats turístiques que comporta canvis en els serveis, mobilitat, activitats en l'espai urbà, etc. A partir de tots aquests factors, estudis com Arias (2018) conclou que les desigualtats provocades pel turisme s'haurien d'abordar a través d'instruments i polítiques urbanes redistribuïdes que garanteixin els drets de vivenda i permanència de la població local.

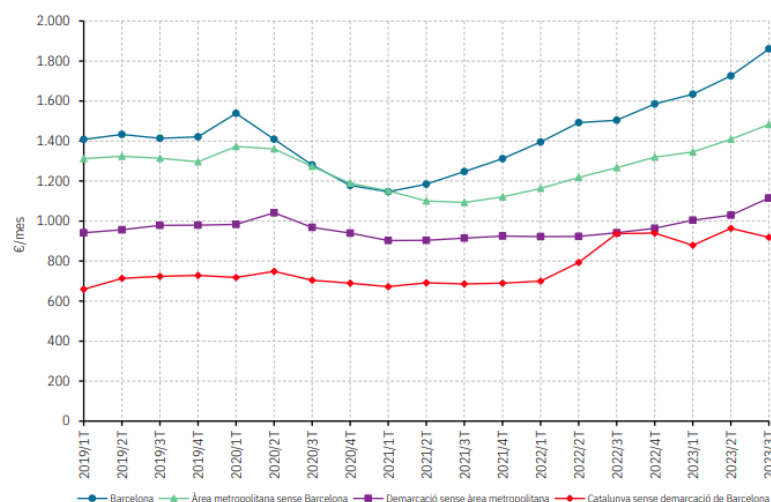


Figura 2.5: Evolució dels preus mitjans d'oferta d'habitatges de lloguer. Catalunya. Primer trimestre de 2019 – Tercer trimestre de 2023 (Observatori Metropolità de l'Habitatge de Barcelona, 2024).

3 Detectar els processos de gentrificació

Fins al moment s'han destacat algunes de les característiques i factors que poden facilitar la caracterització de la gentrificació, així com una descripció de la situació a Barcelona a partir d'alguns estudis que s'hi han destacat. Arrel d'aquesta informació, es planteja la conjuntura de plantejar un procediment per al que, mitjançant mesures socioeconòmiques com les anteriorment descrites, es pugui determinar el nivell de gentrificació d'una zona, així com el potencial risc d'esdevenir-ho.

A partir d'aquest plantejament, sorgeixen estudis com els de [Abarca Álvarez et al. \(2018\)](#), [López-Gay et al. \(2019\)](#) i [Díaz Parra and Apaolaza \(2020\)](#), en que es plantegen una metodologia per mesurar aquests processos de gentrificació. És d'especial interès en l'elaboració d'aquest treball l'estudi de [López-Gay et al. \(2019\)](#) doncs, en aquest cas, està enfocat a les ciutats de Barcelona i Madrid, proporcionant algunes conclusions sobre les quals es podran comparar els resultats del treball present.

Així, l'esmentat estudi presenta un indicador consistent en la utilització de set dimensions teòriques associades típicament a les transformacions socioeconòmiques dels barris. Entre elles es troben factors com el rejuveniment de la població, canvis en la naturalesa de la població, canvis en les estructures dels habitatges, atracció de la població amb estudis universitaris, migració i canvis de residència, velocitat del canvi, i finalment les transformacions del mercat de la vivenda. Els indicadors derivats de cadascun d'aquests factors (els quals posteriorment unifica en un de sol) consisteixen principalment en la diferència de dos quocients de dos períodes temporals diferents (en el cas de l'estudi en qüestió, entre els anys 2011 i 2017), on cadascun

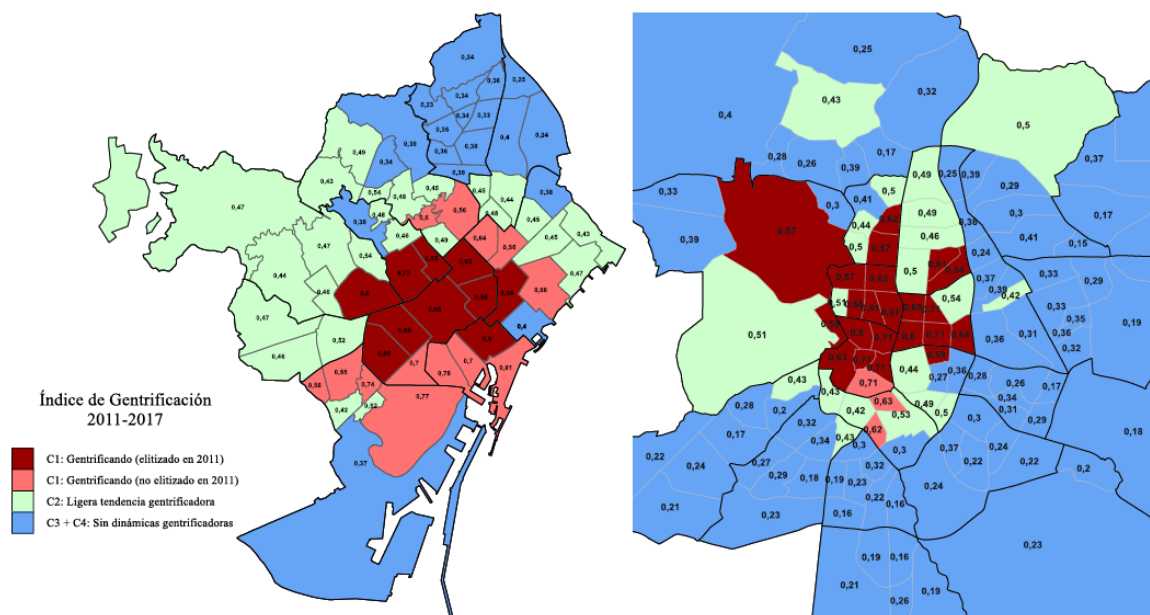


Figura 2.7: Resultats de l'indicador de gentrificació a Barcelona i Madrid, 2011-2017 (López-Gay et al., 2019).

4 Decisions tecnològiques

Per tal d'elaborar el present treball final de màster s'utilitzaran aquells coneixements adquirits durant el Màster en Ciència de Dades que puguin resultar d'utilitat per al seu desenvolupament. Entre d'altres, tècniques pròpies d'obtenció de dades, així com de neteja, estudi preliminar, elaboració d'estadístics, entre d'altres, seran imprescindibles en les primeres fases del procés.

Per a la facilitat de l'autor, s'ha escollit el llenguatge R com a llenguatge de programació estadístic majoritari d'aquest treball, gràcies a les eines estadístiques de que disposa, així com altres llibreries adaptades a múltiples i diverses necessitats interdisciplinars. Entre elles, es destaquen les que proporcionen informació i tractament georeferenciat, doncs resultarà imprescindible per traslladar la informació obtinguda en els posteriors mapes interactius. Alhora, l'entorn sobre el que es treballarà serà RStudio, permetent una interfície gràfica que facilita i agilitza els diferents passos a realitzar durant tot el procés. L'extensió R *Markdown* s'utilitzarà alhora per la elaboració de informes HTML on es plasmin els diferents passos, resultats i conclusions obtingudes a cada etapa.

Finalment, per a l'elaboració de les visualitzacions interactives s'ha escollit el llenguatge R, tot utilitzant la llibreria *Shiny*, així com algunes altres d'auxiliars d'aquesta per augmentar-ne les possibilitats tant estètiques com d'interacció de la interfície final. La conjunció de totes elles permetran plasmar els resultats obtinguts al llarg del present treball.

Capítol 3

Disseny i implementació del treball

En aquest apartat es pretén documentar tota la fase de desenvolupament del present projecte final de màster. Així, es detallaran tots els passos realitzats, començant per l'obtenció de les dades obertes sobre les que es fonamentarà la resta del treball, passant per la neteja i adequació d'aquestes, l'anàlisi i tractament, així com els càlculs de nous paràmetres realitzats per tal d'assolir l'objectiu del càlcul d'un nivell de gentrificació per a la ciutat de Barcelona. Finalment també, es descriurà el procediment realitzat per a la creació d'una visualització interactiva amb molts dels resultats que aquí es detallaran.

Els períodes de disponibilitat de dades varien de variable en variable, tal i com s'anirà especificant en cadascuna. Avançant però en el què el lector es trobarà en aquest treball, tot i que en el visualitzador interactiu es representaran tots els anys per els quals hi hagi dades disponibles, degut a la disponibilitat anual de les dades s'ha decidit que el càlcul del nivell de gentrificació es realitzarà de manera anual entre els anys 2020 i 2023. D'aquesta manera, definint variacions de les variables per barris de fins a un màxim de 5 anys de diferència, finalment s'acabarà treballant amb dades compreses entre el 2015 i 2023. Sota aquesta decisió de treball, les dades finals s'adequaran per a que compleixin aquest requisit.

Finalment, i com s'ha mencionat anteriorment, tot el programari creat per realitzar les diverses tasques s'ha fonamentat en el llenguatge R, tot incloent totes aquelles llibreries que hagin sigut menester per a la correcta elaboració del present treball. Totes les dades i *scripts* de R creats per a tal objectiu es poden trobar al següent enllaç de [Github](#) per al tractament i càlcul de les dades, i en aquest altre enllaç de [Github](#) per al desenvolupament de la visualització interactiva. L'accés a la visualització interactiva es pot trobar al següent [enllaç](#).

1 Obtenció i tractament de les dades

Tal i com s'ha fet esment anteriorment, les dades utilitzades han sigut totes de caràcter públic i obertes durant el moment de la seva obtenció. Així, la majoria de les dades s'han extret principalment del portal de Open Data BCN, el servei de dades obertes de l'Ajuntament de Barcelona, tot i que també s'han utilitzat dades de la Oficina Municipal de Dades del Departament d'Estadística i Difusió de Dades, també de l'Ajuntament de Barcelona, i del Departament d'Habitatge de la Generalitat de Catalunya. A mida que es detallin quines dades s'han utilitzat, s'anirà detallant la font. Al respecte de les dades obtingudes del portal Open Data BCN, aquestes s'han descarregat utilitzant els *scripts* que es poden trobar al següent [enllaç](#), la resta s'han descarregat manualment.

Tots els conjunts de dades, a mida que s'han anat descarregant, s'han anat emmagatzemant, dins de l'arrel del repositori, a la carpeta `data/L0/nom_del_conjunt/*`. Després del posterior anàlisi de les dades, tractament i adequació (processos que es detallaran tot seguit), s'han emmagatzemat cadascun d'ells a la carpeta `data/L1/*`. Per obtenir més detalls del procés, es recomana acompanyar la següent secció juntament amb el codi executat que es pot trobar al repositori de [Github](#).

1.1 Nivell d'estudis

Un dels conceptes que s'ha observat en l'estat de l'art referent a la gentrificació de la ciutat ha sigut el concepte del volum de població amb estudis universitaris, així com la seva distribució geogràfica pels diferents barris. Per obtenir informació referent a aquesta característica s'han obtingut dades del padró a través d'aquest [enllaç](#). En el moment de la consulta de les dades disponibles, aquestes comprenien el període entre 1997 i 2023, de les quals se n'han obtingut tots els anys amb els pertinents arxius en format CSV.

Aquests arxius comprenen, per a cada any indicat, informació agrupada del volum de població separada per:

- Districte (Codi del districte i nom)
- Barri (Codi del barri i nom)
- Nivell educatiu (Índex del nivell educatiu)
- Edat quinquenal (Índex d'agrupament d'edats en blocs de 5 anys)
- Sexe (Índex del sexe, corresponent al femení i masculí)

Després d'un anàlisi preliminar dels arxius, s'han identificat dos potencials casuístiques que tenen aquests conjunt de dades. Tot i que aquestes no són úniques d'aquest conjunt, es detallaran seguidament i se'n farà menció als següents conjunts que n'hagin sigut adaptats.

La primera de les casuístiques consisteix en l'aparició del valor "." quan el valor real és inferior a 5, tal i com ja s'indica a la pàgina web d'origen. Degut al previ coneixement d'aquest problema, i al obvi desconeixement de quin és el possible valor real, en el moment d'importar els múltiples *datasets* s'ha tractat automàticament aquest text "." com a un valor de NA.

La segona casuística que s'ha trobat en alguns dels conjunts de dades és la de falta de registres per a totes les possibles combinacions de any, districte, barri, i les múltiples combinacions d'altres variables que presenti les dades en concret. Per resoldre-ho (i a mode preventiu s'ha aplicat a tots els següents conjunts de dades que així s'indiqui), es crea una variable auxiliar que, amb una funció de combinació de taules, s'ha combinat amb les dades originals per tal de forçar la creació de registres que prèviament no existien. Òbviament els valors corresponents de població s'han adaptat com a un valor desconegut NA, però s'ha realitzat d'aquesta manera doncs, al combinar posteriorment els conjunts de dades de múltiples dimensions, no aparegués la possibilitat que hi haguessin diferents nombres de registres i dificultés la feina posterior de tractament.

En quant a possibles valors de nivell educatiu, estan llistats els següents (amb una codificació que posteriorment serà creuada amb la informació disponible en aquest [enllaç](#)):

- Sense estudis
- Batxillerat superior, BUP, COU, FPPI, CFGM grau mitjà
- Estudis primaris, certificat d'escolaritat, EGB
- Estudis universitaris, CFGS grau superior
- Batxillerat elemental, graduat escolar, ESO, FPI
- No consta

Com per a aquesta variable no és d'interès ni l'edat ni el sexe, s'han realitzat càlculs d'agrupament, de manera que s'han reduït les dades al total de població a nivell de barri.

Els noms dels barris i districtes també s'han eliminat de les dades, doncs serà posteriorment en l'anàlisi final quan ja se li atorgui novament el nom del barri o districte a partir dels corresponents codis.

1.2 Estructura del domicili

Un altra aspecte que s'ha considerat per calcular el nivell de gentrificació, és el volum de domicilis on hi resideixen persones soles entre 18 i 64 anys, les quals acostumen a ser nouvingudes a la ciutat per diversos motius, com pot ser per qüestions laborals, interès propi en la ciutat, etc., i que a més tenen la capacitat econòmica d'afrontar una vivenda de manera individual.

Per obtenir informació del volum de domicilis que compleixen aquesta condició, es remet al següent [enllaç](#) per obtenir-ne les dades del padró. Igual que en l'anterior cas, les dades

disponibles tornen a ser en el període comprès entre el 1997 i el 2023 en el moment de la consulta i descàrrega, també en arxius en format CSV. Aquests arxius comprenen, per a cada any indicat, informació agrupada del volum de població separada per:

- Districte (Codi del districte i nom)
- Barri (Codi del barri i nom)
- AEB (Àrea Estadística Bàsica)
- Secció Censal
- Tipus de domicili (Índex del tipus de domicili)

D'igual manera que el cas anterior, novament s'han tractat els valors inferiors a 5 (expressats en els arxius com a ". . ."), readaptant-los com a un valor desconegut NA, i també s'ha assegurat que existissin tots els possibles registres a nivell de barri i tipus de domicili, de tal manera que no es tinguessin registres buits en el *dataset* final.

A partir del índex dels tipus de domicili, i novament creuant la informació amb la informació disponible en el següent [enllaç](#), es tenen llistats els següents tipus de domicili:

- Una dona de 18 a 64 anys
- Un home de 18 a 64 anys
- Una dona de 65 anys i més
- Un home de 65 anys i més
- Dues persones de 18 a 64 anys
- Dues persones de 65 anys i més
- Dues persones: una de 18-64 anys i l'altra de 65 i més
- Dues persones o més totes majors de 18 anys
- Dues persones o més: una dona de 18 anys i més amb altres menors de 18 anys
- Dues persones o més: un home de 18 anys i més amb altres menors de 18 anys
- Tres persones o més: dues de 18 anys i més i la resta menors de 18 anys
- Altres domicilis amb una o més persones menors de 18 anys

Novament igual que l'anterior cas, es té informació ampliada per àrea estadística bàsica i per secció censal, però com al nivell que es treballarà tan sols s'arribarà fins a valors per barri, s'han realitzat altra vegada càlculs d'agrupament per reduir-ne el detall. Alhora, els noms dels barris i districtes s'han eliminat novament del conjunt.

1.3 Regió de naixement d'origen

La immigració, però sobretot destacant aquella que té un poder adquisitiu potencialment superior al de la mitjana de la ciutat de Barcelona, és un component molt important a l'hora de destacar els processos de gentrificació a la ciutat comtal, doncs són potencialment uns indicadors bastant poderosos alhora de valorar el nivell de gentrificació.

Les dades obtingudes novament provenen del padró, en concret del següent [enllaç](#). Novament hi ha dades anuals disponibles entre el 1997 i 2023, tots obtingut d'arxius en format CSV. Per a cada any indicat, es disposa d'informació de les següents variables:

- Districte (Codi del districte i nom)
- Barri (Codi del barri i nom)
- AEB (Àrea Estadística Bàsica)
- Secció Censal
- Regió del lloc de naixement (Índex de la regió de naixement)
- Sexe (Índex del sexe, corresponent al femení i masculí)

Igual que en els dos casos anteriors, s'han carregat els pertinents arxius tenint en compte els valors com a NA, i s'han contemplat tots aquells potencials registres que poguessin faltar arrel de la combinació de les múltiples variables. Així també com els casos anteriors, s'han agrupat les dades de manera que es tingui informació fins a només a nivell de barri i de regió del lloc de naixement.

Les regions de naixement que estan aquí codificades són les següents:

- Àfrica oriental
- Àfrica central
- Àfrica septentrional
- Àfrica meridional
- Àfrica occidental
- Carib
- Amèrica central
- Amèrica del sud
- Amèrica del nord
- Àsia central
- Àsia oriental
- Àsia meridional
- Àsia sud-oriental
- Àsia occidental
- Europa oriental
- Europa septentrional
- Europa meridional
- Europa occidental
- Austràlia i Nova Zelanda
- Melanèsia
- Micronèsia
- Polinèsia
- No consta

Com s'observa es tenen moltes possibles regions de naixement, però no es necessita tanta precisió, doncs al final tant sols es posarà especial enfoc en aquelles regions que tenen un potencial més elevat de posseir una capacitat econòmica més elevada que el de la població

local, les quals corresponen als països d'Europa i d'Amèrica del nord. Així, s'ha readaptat i agrupat els elements d'aquesta llista, de manera que ha quedat de la següent manera:

- Àfrica
- Amèrica central/sud
- Amèrica del nord
- Àsia
- Europa
- Oceania
- No consta

No obstant, s'ha de tenir present que, en el cas de les dades d'Europa, aquí també s'hi estan incloent a aquelles persones nascudes a Espanya, les quals no interessa considerar en aquesta variable. Així, s'ha accedit a un altre conjunt de dades equivalent, per al qual hi ha informació país per país (disponible en el següent [enllaç](#)).

Les característiques d'aquestes dades són equivalents a les anteriors (amb la diferència que en aquest cas es té separada la regió de naixement en dues noves variables: el país i el continent de naixement). El tractament d'aquestes dades és equivalent al que s'ha aplicat amb anterioritat, contemplant possibles registres buits i reduïnt la informació a nivell de barri, però amb la salvedat de que finalment només es tindrà en compte la informació corresponent a la d'Espanya.

Un cop s'ha tractat adequadament, es combina amb el *dataset* original per regions de naixement, de manera que al valor de població de la regió europea se li resta aquella població obtinguda del segon *dataset*, obtenint finalment tan sols aquella població europea estrangera. Abans però de restar aquest valor, s'emmagatzema en nou camp el total de població per barris, doncs en un posterior càlcul del percentatge de població, aquest es realitzarà sobre el total de població, no només sobre la immigrant, i per tant requerim guardar aquesta variable. Una vegada més, els noms del barris i dels districtes s'eliminen de la informació emmagatzemada.

1.4 Grups d'edat

En conjunció amb les dues variables anteriors, el perfil de població el qual se li atribueix un cert potencial econòmic que pugui influir en els nivells de gentrificació és el de les persones amb capacitat de treballar i que han pogut acumular uns estalvis adequats pels nous requisits de la ciutat. Per tant, tenir una aproximació de les edats dels habitants dels diferents barris i districtes de Barcelona segur que aporta una informació valuosa per a la determinació de la gentrificació.

Les dades d'aquest conjunt són les últimes de la selecció de variables que s'obtenen del padró, disponibles en el següent [enllaç](#). Una vegada més, es disposa informació anual entre el període 1997 i 2023, obtenint un arxiu per any en format CSV. Les variables d'aquest *dataset* són les següents:

- Districte (Codi del districte i nom)
- Barri (Codi del barri i nom)
- AEB (Àrea Estadística Bàsica)
- Secció Censal
- EDAT_1 (Índex de l'edat)

Els valors de l'índex de l'edat (els valors de les dades comprenen entre el 0-100), equivalen al grup de l'edat equivalent, amb l'excepció del valor 100, el qual contempla els casos en que es tinguin 100 anys o més.

Novament, també com els altres casos fins ara, les dades al importar-se han sigut corregides amb els valors de NA corresponents, i han sigut també completades perquè no es tinguessin registres buits, seguint el mètode anteriorment descrit.

No obstant, s'ha decidit crear una nova variable per realitzar agrupaments d'edat més grans que els presentats, en aquest cas en blocs de 10 anys, que serien els següents:

- 0-9 anys
- 10-19 anys
- 20-29 anys
- 30-39 anys
- 40-49 anys
- 50-59 anys
- 60-69 anys
- >70 anys

Un cop creat el nou índex (que s'ha anomenat com a EDAT_10), s'han agrupat els valors de població a nivell d'aquest nou índex i de barri. No obstant però, coneixent la casuística del següent *dataset* de la renda anual bruta que es detallarà tot seguit, s'ha creat a partir d'aquest d'edats un *dataset* paral·lel en què registrem el total de població per secció censal que s'utilitzarà tot seguit.

1.5 Renda anual bruta

S'ha mencionat ja en altres variables, però pot aixecar sospites de processos de gentrificació que alguns barris puguin tenir augments dels sous dels seus habitants de manera destacada respecte a la resta, on també el poder adquisitiu previ n'és també un factor important.

El *dataset* que aquí s'utilitza s'ha obtingut a través del següent [enllaç](#). A diferència de la resta de *datasets* fins ara, el període que comprèn les dades disponibles està entre els anys 2015 i 2021, separats amb arxius anuals en format CSV. Entre les variables disponibles es té:

- Districte (Codi del districte i nom)
- Barri (Codi del barri i nom)
- Secció Censal
- Import renda mitjana bruta per persona (en €)

Amb aquestes dades es realitza el mateix que fins ara, procurant que no quedi cap possible registre per contemplar, a més de corregir els valors de NA. Com s'ha mencionat, aquest *dataset*

només contempla fins l'any 2021, i l'objectiu d'aquest treball és arribar a calcular el nivell de gentrificació fins al 2023. Així, en el llistat de tots els possibles registres, també es contempla una extensió de fins el 2023, òbviament amb valors de NA, els quals es tractaran seguidament.

Un cop s'han assegurat que les dades són correctes, el lector s'adona que la variable de l'import de renda mitjana és el primer camp calculat que tenim, és a dir, la traslació a valors de barri o districte no és immediata, ja que es necessita conèixer la població per secció censal de tal manera que es pugui calcular una variable comodí que es consideraria com la suma de totes les rendes de tota la població per secció censal.

Per realitzar-ho, es recupera el *dataset* que s'ha creat al final de l'estudi de les edats amb la població per secció censal, la qual multiplicada per l'import de renda mitjana obtenim aquest valor total de renda per secció censal i, ara sí, reduir la granularitat de la informació fins a nivell de barri resulta més senzill. S'ha de tenir en compte també que s'ha de mantenir present en aquest *dataset* la variable del total de població heretada de l'anterior, doncs al final la variable que es considerarà serà novament un import de renda mitjana per persona.

No obstant, es posa de relleu novament el fet de que només es tenien dades disponibles fins l'any 2021. Per solucionar-ho, s'ha decidit calcular imports de renda bruta totals per secció mitjançant tècniques d'interpolació, calculades secció a secció individualment. D'aquesta manera es pot estimar les rendes mitjanes d'aquestes zones per aquests dos últims anys amb absència de dades.

Finalment, es vol destacar també que no s'ha tingut en compte la població treballadora per zones, i es té present que s'està sobreestimant la població que té una renda ja que s'està considerant el total de població, per lo que els imports de renda mitjans recalculats són potencialment inferiors als reals.

1.6 Preu del lloguer

El preu del lloguer és una variable també molt important a l'hora de valorar el nivell de gentrificació d'un barri, doncs alhora que els sous, increments sobtats d'aquests seran sospitosos de ser indicadors de que un barri està patint processos gentrificadors.

Per obtenir informació sobre el preu del lloguer, en aquesta ocasió es realitza la consulta al Departament d'Habitatge de la Generalitat de Catalunya, disponible al següent [enllaç](#). Aquí, entre els múltiples arxius disponibles, s'escull treballar amb la informació anual del lloguer mitjà per metre quadrat, el qual està contingut en un arxiu en format XLSX.

A diferència dels anteriors conjunts de dades fins ara, es tenen períodes d'informació variable segons si són per ciutat o districte (2000-2023) o per barri (2013-2023), fet que caldrà tenir en compte. Així, de manera anual, ja es té el valor del lloguer mitjà per àrea territorial. Alhora, tot i que el càlcul de gentrificació es realitzarà a nivell de barri, es vol conservar la informació

dels lloguers per ciutat i districte per al visualitzador interactiu. Per això, es realitzen un seguit de processos sobre l'arxiu original, de tal manera que es creen tres variables noves (per ciutat, districte i barri), les quals s'agrupen posteriorment en una de sola.

A més a més, després d'un anàlisi exploratori de les dades, s'ha observat com alguns barris tenen alguns valors de import de lloguer buits. Per corregir-ho, d'igual manera que s'ha realitzat pels imports de les rendes, s'han calculat els valors buits mitjançant tècniques d'interpolació, calculats barri a barri.

1.7 Habitatges turístics

Finalment, s'ha volgut posar un enfoc en la variable turística en el càlcul del nivell de gentrificació a la ciutat. Així, d'entre les dades disponibles que s'han trobat amb granularitat per barris, s'ha decidit utilitzar el nombre d'habitatges turístics com a representant d'aquesta variable.

Aquesta informació s'ha obtingut a través del següent [enllaç](#) del Departament d'Estadística i Difusió de Dades de l'Ajuntament de Barcelona. El mètode d'importació d'aquestes dades ha sigut el de la opció 2 detallada en aquest altre [enllaç](#), on s'ha utilitzat el programari ofimàtic Excel, mitjançant tècniques de Power Query per adaptar les dades a un format més adequat, tal i com s'observa a la figura 3.1, el qual s'emmagatzemarà per seguir-lo tractant posteriorment en el llenguatge R.

| | Codi_Districte | Codi_Barri | Nom_Barri | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
|----|----------------|------------|--|------|------|------|------|------|------|
| 1 | 1 | 1 | el Raval | 180 | 180 | 177 | 177 | 177 | 177 |
| 2 | 1 | 2 | el Barri Gòtic | 184 | 184 | 184 | 188 | 188 | 188 |
| 3 | 1 | 3 | la Barceloneta | 69 | 69 | 68 | 68 | 68 | 68 |
| 4 | 1 | 4 | Sant Pere, Santa Caterina i la Ribera | 171 | 171 | 169 | 169 | 169 | 169 |
| 5 | 2 | 5 | el Fort Pienc | 343 | 343 | 331 | 331 | 331 | 331 |
| 6 | 2 | 6 | la Sagrada Família | 794 | 794 | 772 | 772 | 772 | 772 |
| 7 | 2 | 7 | la Dreta de l'Eixample | 1747 | 1747 | 1653 | 1653 | 1653 | 1653 |
| 8 | 2 | 8 | l'Antiga Esquerra de l'Eixample | 727 | 727 | 709 | 709 | 709 | 709 |
| 9 | 2 | 9 | la Nova Esquerra de l'Eixample | 428 | 428 | 418 | 418 | 418 | 418 |
| 10 | 2 | 10 | Sant Antoni | 568 | 568 | 590 | 590 | 590 | 590 |
| 11 | 3 | 11 | el Poble Sec - AEI Parc Montjuïc | 546 | 546 | 518 | 518 | 518 | 518 |
| 12 | 3 | 12 | la Marina del Prat Vermell - AEI Zona Franca | 3 | 3 | 3 | 3 | 3 | 3 |
| 13 | 3 | 13 | la Marina de Port | 6 | 6 | 7 | 7 | 7 | 7 |
| 14 | 3 | 14 | la Font de la Guatlla | 74 | 74 | 73 | 73 | 73 | 73 |
| 15 | 3 | 15 | Hostafrancs | 187 | 187 | 168 | 168 | 168 | 168 |
| 16 | 3 | 16 | la Bordeta | 42 | 42 | 44 | 44 | 44 | 44 |
| 17 | 3 | 17 | Sants - Badal | 90 | 90 | 106 | 106 | 106 | 106 |
| 18 | 3 | 18 | Sants | 189 | 189 | 245 | 245 | 245 | 245 |
| 19 | 4 | 19 | les Corts | 171 | 171 | 210 | 210 | 210 | 210 |
| 20 | 4 | 20 | la Maternitat i Sant Ramon | 36 | 36 | 46 | 46 | 46 | 46 |
| 21 | 4 | 21 | Pedrabes | 8 | 8 | 11 | 11 | 11 | 11 |
| 22 | 5 | 22 | Valldorera, el Tibidabo i les Planes | 3 | 3 | 3 | 3 | 3 | 3 |
| 23 | 6 | 23 | Sarrià | 74 | 74 | 77 | 77 | 77 | 77 |

Figura 3.1: Procés d'obtenció de les dades de pisos turístics mitjançant Power Query.

Les dades disponibles de pisos turístics contempnen informació per barris en el període entre

2014 i 2022, per lo que s'interpolaran els valors per barri per l'últim any 2023 tot i que, en general, s'han mantingut pràcticament constants any a any.

1.8 Unitats administratives

Tot i que no és una variable en sí, no es deixa de banda l'obtenció del detall de les diferents unitats administratives de la ciutat, concretament amb l'objectiu d'obtenir en format SHP els polígons fronterers a nivell de ciutat, districtes i barris.

Per obtenir-los, s'ha accedit al següent [enllaç](#), descarregant en aquest cas l'arxiu d'informació geogràfica més actualitzat (`Unitats_Administratives_BCN.shp`) en format ZIP, de manera que s'hi continguin tot un conjunt d'arxius que, entre ells, estan els dels nivells administratius desitjats. Cadascun d'aquests serà carregat a l'hora de reproduir mapes, segons convingui.

2 Unificació de les dades

Un cop ja es disposen de totes les dades correctament tractades, el següent pas immediat consisteix a unificar-les totes en un únic *dataset*.

Segons la disponibilitat d'informació completa de totes les variables, com ja s'ha fet esment anteriorment en el moment en que s'han interpolat valors faltants per barris (destacant novament els dels preus del lloguer d'alguns barris i anys, així com el de les rendes anuals de manera generalitzada per a tots els barris per als dos últims anys, i els valors de pisos turístics per a l'últim any), s'ha limitat el conjunt de dades finals a el període comprès entre 2015 i 2023.

Es recorda que el càlcul del nivell de gentrificació es realitzarà a nivell de barri, per tant tota la informació referent a ciutat i/o districte no serà considerada en aquest *dataset* final.

Per obtenir més detalls del procés, es recomana acompanyar la següent secció juntament amb el codi executat que es pot trobar al repositori de [Github](#).

2.1 Selecció de grups d'interès

Abans però de la unificació de totes les variables, es necessita aplicar filtratges addicionals segons la variable en cada cas, doncs no es vol conservar tota la informació disponible fins ara, només aquella que es considera important per la determinació del nivell de gentrificació.

A excepció dels imports de les rendes brutes mitjanes per barris, els preus dels lloguers, i el nombre d'habitatges turístics (dels quals només se'n conserva els respectius valors absoluts), de la resta de variables es seleccionaran alguns conjunts de població/domicili en concret.

Així, per al *dataset* de nivell d'estudis, es mantindrà el conjunt de població que té un nivell equivalent o superior al de estudis universitaris, mantenint per a cada barri i any el total de

població universitària i el percentatge corresponent respecte al volum de població total.

En el *dataset* d'estructura de domicilis es realitzarà un tractament similar, on en aquest cas es consideren conjuntament els domicilis habitats per un home o una dona majors a 18 anys que viuen sols. Igual que l'anterior cas, es mantindrà el total de domicilis d'aquest conjunt, alhora que el percentatge corresponent respecte al total de domicilis.

El *dataset* de regió de naixement d'origen se li aplica un filtratge equivalent, on en aquest cas es consideren d'especial atenció el perfil de població amb origen europeu i d'Amèrica del nord. Així, novament es manté el total de població estrangera d'aquestes dues regions, i el percentatge corresponent respecte al total de la població (no només estrangera si no també la nascuda a Espanya).

Finalment, es seleccionen els grups d'edat entre els 20-50 anys del *dataset* de grup d'edat i, de manera equivalent als anteriors, se'n conserva el total de població d'aquest interval d'edat, així com el percentatge corresponent respecte al total.

2.2 Càlcul de les variacions interanuals

Els processos de gentrificació són, tal i com diu la paraula, una conseqüència que es deriva de l'evolució d'un seguit de característiques al llarg del temps. Per tant, quedar-se amb una instantània anual de la situació d'aquestes variables comportarà una pèrdua d'informació que no interessa descartar.

Per tal de contemplar finestres temporals suficientment àmplies, i també limitats amb la disponibilitat de les dades de treball, s'han considerat finestres temporals de 5 anys, és a dir, que en el període entre 2015 i 2023 es poden establir quatre finestres temporals diferents:

- 2015 - 2020
- 2016 - 2021
- 2017 - 2022
- 2018 - 2023

D'aquesta manera, es requereix un càlcul addicional per obtenir els valors de les variacions entre el període 2020 i 2023 respecte a com eren 5 anys enrere (sempre tenint present que es calcula barri per barri de manera individual). El mètode de càlcul, inspirat en el treball de [López-Gay et al. \(2019\)](#), es divideix en dos grups, els quals seran els següents.

Per als *datasets* del preu del lloguer i la renda mitjana per persona, el càlcul de la variació serà tal com segueix:

$$VAR_f^j = \frac{X_f^j - X_i^j}{X_i^j}$$

On VAR_f^j correspon a la variació de la variable j (preu del lloguer / renda mitjana) a l'any final (f), X_f^j correspon al valor de la variable j a l'any final f i X_i^j correspon al valor de la variable j a l'any inicial i .

Un cop calculades aquestes variacions, per a aquests *datasets* llistats es conservarà, per barri i any final (2020, 2021, 2022 i 2023), el valor de la variable (preu del lloguer / import de renda mitjana) i el valor de la variació corresponent.

Per als *datasets* del nivell d'estudis, l'estructura del domicili, la regió de naixement d'origen i el grup d'edat, el càlcul serà el corresponent a la diferència entre els valors percentuals de les proporcions entre el conjunt de població o domicili seleccionat (segons correspongui), respecte als respectius totals. En aquest punt, es recorda que a l'apartat anterior ja s'havia detallat el càlcul d'aquests percentatges. Així, el càlcul de la variació d'aquestes variables segueix la següent expressió:

$$VAR_f^k[\%] = X_f^k[\%] - X_i^k[\%]$$

On, en aquesta ocasió, VAR_f^k correspon a la variació percentual de la variable k (proporció de població / domicili respecte al total, per a cadascuna de les variables indicades) a l'any final (f), X_f^k correspon al valor de la variable k a l'any final f i X_i^k correspon al valor de la variable k a l'any inicial i .

Novament, i una vegada s'ha calculat aquestes variacions, per a aquests altres *datasets* llistats es conservarà, per barri i any final (2020, 2021, 2022 i 2023), el valor percentual de la proporció de la població o domicili (segons pertoqui) respecte als respectius totals i el valor de la variació percentual corresponent.

A banda queda la variable dels habitatges turístics. Això és degut a que s'ha observat que les variacions d'aquesta al llarg dels anys, i per a tots els barris, són menyspreables, per lo que excepcionalment per a aquesta variable només es tindrà en compte la variable absoluta del total d'habitatges turístics per barri i any.

Un cop s'han calculat totes aquestes variacions interanuals, es procedeix, finalment, a la unificació de les dades en un únic *dataset* mitjançant tècniques de combinació on l'any i l'identificador del barri en siguin el nexa comú. A la figura 3.2 es pot veure una selecció d'aquestes dades, on en total s'emmagatzemen 292 registres (els 73 barris per 4 anys) amb 15 atributs diferents (dos per a cada variable - exceptuant la dels pisos turístics que n'és un sol atribut -, l'any i el codi del barri). Finalment, aquestes dades s'han emmagatzemat en un únic arxiu `dades_L2.csv`, dins la carpeta `data/L2/`.

3 Càlcul del nivell de gentrificació

Una vegada s'han unificat tots els *datasets* en un d'únic de més fàcil manipulació, es procedeix a descriure la metodologia descrita per a la determinació del nivell de gentrificació, calculat any a any i a nivell de barri. Per obtenir més detalls del procés, es recomana acompanyar la

| ▲ | Data_Referencia | Codi_Barrri | Renda | Var_Renda | Pisos | Perc_Domicilis | Var_Domicilis | Perc_Edat | Var_Edat | Perc_Regio | Var_Regio | Perc_Uni | Var_Uni | Lloguer | Var_Lloguer |
|----|-----------------|-------------|-----------|-------------|-------|----------------|---------------|-----------|---------------|-------------|---------------|------------|--------------|-----------|--------------|
| 1 | 2020 | 01 | 10971.809 | 0.129893070 | 191 | 0.24861719 | 1.803408e-02 | 0.5683862 | -2.122929e-03 | 0.114427237 | 1.996428e-02 | 0.27200192 | 0.058486026 | 833.8682 | 0.332438021 |
| 2 | 2020 | 02 | 16221.377 | 0.165311540 | 193 | 0.28498036 | -2.848647e-03 | 0.6698634 | 7.059869e-02 | 0.179037274 | -2.233036e-02 | 0.32766954 | -0.039368733 | 1035.7777 | 0.317709752 |
| 3 | 2020 | 03 | 14798.554 | 0.153096560 | 64 | 0.28565317 | 3.080060e-02 | 0.5816918 | 4.555252e-02 | 0.184837303 | 3.148957e-02 | 0.30585605 | 0.068627890 | 867.9499 | 0.440296258 |
| 4 | 2020 | 04 | 16963.035 | 0.173374320 | 159 | 0.29960823 | 1.704371e-02 | 0.5856155 | 7.516646e-03 | 0.210227273 | 2.905939e-02 | 0.42162110 | 0.059443055 | 954.4768 | 0.319342446 |
| 5 | 2020 | 05 | 21968.756 | 0.139534339 | 332 | 0.16774143 | -7.296136e-03 | 0.4673381 | 2.215853e-02 | 0.089384629 | 2.042133e-02 | 0.40345459 | 0.044554555 | 1003.1163 | 0.305930958 |
| 6 | 2020 | 06 | 21403.563 | 0.125532526 | 765 | 0.18167342 | 3.239229e-03 | 0.4538223 | 9.866840e-03 | 0.069061146 | 1.547339e-02 | 0.39476226 | 0.057636039 | 942.0899 | 0.321029655 |
| 7 | 2020 | 07 | 28957.790 | 0.107266911 | 1662 | 0.20356660 | -5.210792e-03 | 0.4552970 | 1.190388e-02 | 0.126514995 | 2.808318e-02 | 0.52816322 | 0.044901941 | 1275.5032 | 0.257768310 |
| 8 | 2020 | 08 | 26912.775 | 0.153399593 | 709 | 0.20808548 | -7.742214e-04 | 0.4764926 | 2.324991e-02 | 0.099286619 | 1.462127e-02 | 0.49733094 | 0.053841560 | 1155.7804 | 0.268493959 |
| 9 | 2020 | 09 | 24456.657 | 0.139888672 | 413 | 0.17895120 | -1.631166e-03 | 0.4490224 | 9.197874e-03 | 0.064217324 | 1.132788e-02 | 0.44949117 | 0.059932627 | 1034.1529 | 0.319349224 |
| 10 | 2020 | 10 | 21165.285 | 0.133298573 | 586 | 0.18718639 | 3.739819e-03 | 0.4631697 | 7.988642e-03 | 0.084719375 | 1.702117e-02 | 0.38919652 | 0.059363962 | 971.8817 | 0.269423007 |
| 11 | 2020 | 11 | 15111.450 | 0.171458878 | 515 | 0.21650388 | 2.218419e-02 | 0.5104678 | 3.001396e-03 | 0.103514104 | 2.566573e-02 | 0.29591202 | 0.066540161 | 827.4298 | 0.351688077 |
| 12 | 2020 | 12 | 12952.000 | 0.272798742 | 3 | 0.23983740 | 2.522553e-02 | 0.4406355 | -1.100688e-02 | 0.020798669 | 1.776930e-03 | 0.04850361 | 0.022468109 | 395.2389 | -0.023406786 |
| 13 | 2020 | 13 | 15465.161 | 0.136008206 | 6 | 0.14154152 | 3.935574e-03 | 0.4044987 | -1.776737e-02 | 0.028830340 | 3.726577e-03 | 0.16780644 | 0.032085750 | 787.0947 | 0.451492424 |
| 14 | 2020 | 14 | 19422.976 | 0.146006320 | 73 | 0.17540964 | -7.282472e-03 | 0.4420311 | -1.190917e-02 | 0.066992187 | 2.076139e-03 | 0.31910986 | 0.042373369 | 859.2797 | 0.236020658 |
| 15 | 2020 | 15 | 18272.371 | 0.137300722 | 177 | 0.20827113 | 1.602829e-02 | 0.4896154 | 9.102573e-03 | 0.076750000 | 1.873238e-02 | 0.33630430 | 0.068069516 | 852.2671 | 0.364394895 |
| 16 | 2020 | 16 | 18801.929 | 0.168421879 | 44 | 0.15365112 | -8.501009e-04 | 0.4169948 | -3.069263e-03 | 0.034267913 | 4.732515e-03 | 0.26647514 | 0.054003147 | 861.7976 | 0.349810565 |
| 17 | 2020 | 17 | 18074.201 | 0.111388417 | 106 | 0.16840433 | -2.859690e-03 | 0.4413245 | 6.356892e-03 | 0.042688871 | 9.497705e-03 | 0.28219714 | 0.053247909 | 845.3940 | 0.349248053 |
| 18 | 2020 | 18 | 20518.652 | 0.146934370 | 245 | 0.19274100 | 5.461431e-03 | 0.4680861 | 1.671201e-02 | 0.055358344 | 9.940117e-03 | 0.33592720 | 0.043546101 | 874.4672 | 0.338531686 |
| 19 | 2020 | 19 | 28360.328 | 0.141145569 | 206 | 0.15363333 | -3.261934e-03 | 0.3921146 | -3.396543e-03 | 0.044999672 | 1.011104e-02 | 0.45006150 | 0.048857491 | 1060.7121 | 0.238652604 |
| 20 | 2020 | 20 | 25007.906 | 0.120133416 | 44 | 0.13507109 | 4.435979e-04 | 0.3938678 | -2.131835e-03 | 0.038898657 | 7.956486e-03 | 0.40466833 | 0.043357801 | 992.6913 | 0.312014631 |
| 21 | 2020 | 21 | 41190.154 | 0.133961841 | 11 | 0.11598967 | 1.426562e-03 | 0.3584075 | -7.576364e-03 | 0.103653863 | 1.054603e-02 | 0.52795112 | 0.021527027 | 1718.2406 | -0.006769037 |
| 22 | 2020 | 22 | 27499.135 | 0.130892527 | 3 | 0.13878563 | -1.466041e-02 | 0.3871595 | -2.117013e-02 | 0.112505409 | 1.954891e-02 | 0.43957219 | 0.025509693 | 1233.0746 | 0.225070871 |
| 23 | 2020 | 23 | 36904.854 | 0.159252090 | 27 | 0.14354930 | -5.578512e-03 | 0.3693701 | -1.357017e-02 | 0.092454578 | 1.265439e-02 | 0.51900835 | 0.028304460 | 1366.2417 | 0.231222807 |
| 24 | 2020 | 24 | 43132.991 | 0.140873134 | 35 | 0.11615892 | -8.590673e-03 | 0.3565138 | -2.471353e-02 | 0.061409417 | -6.616609e-04 | 0.53882422 | 0.028115821 | 1696.1170 | 0.224602106 |
| 25 | 2020 | 25 | 37045.028 | 0.155551362 | 32 | 0.14472436 | -7.196415e-03 | 0.3784118 | -4.763010e-03 | 0.063297541 | 1.354530e-02 | 0.51779384 | 0.042986716 | 1301.1298 | 0.243838536 |

Figura 3.2: Selecció de les dades de treball una vegada unificades.

següent secció juntament amb el codi executat que es pot trobar al repositori de [Github](#).

Per tal de tenir unificats els diferents atributs segons el seu origen, es proposa seguir amb la metodologia proposada per [López-Gay et al. \(2019\)](#), on per a cada parella d'atributs per variable (a excepció, en el cas d'aquest treball, de la variable dels pisos turístics que té un únic atribut) se li atribueix una puntuació única on la variació temporal pondera $\frac{2}{3}$, i el valor de cada variable a l'any final pondera $\frac{1}{3}$. Per calcular-ne el valor, s'ha realitzat, atribut a atribut (sempre realitzant el càlcul per anys de manera diferenciada), una interpolació lineal entre el cinquè valor més petit i el cinquè més gran (tant de la variació temporal com del valor a l'any final per a cada interval d'estudi), de manera que els cinc valors més grans puntuen 1 i els cinc més petits puntuen 0. Aquests valors de referència, any a any i per atribut, es llisten a la taula [3.1](#), juntament amb les respectives medianes.

Un cop realitzada la interpolació lineal entre els respectius valors extrems, per parelles d'atributs es realitza una suma d'aquests tenint en compte les respectives ponderacions (es recorda, $\frac{2}{3}$ per la variació temporal i $\frac{1}{3}$ pel valor anual) i, un cop calculades, se'ls hi torna a aplicar una interpolació lineal, aquesta vegada entre el valor màxim i mínim d'aquesta nova variable suma, on el valor màxim se li atribueix un valor de 1, i el mínim un valor de 0. En quant a la variable dels habitatges turístics, en aquest punt no se li ha aplicat cap transformació. Així, aquesta puntuació final s'interpreta com, any per any, aquell barri que lideri cada variable tindrà una puntuació de 1, i a l'últim se li haurà atorgat una puntuació de 0.

Un cop calculades les puntuacions per variable, es descarten els atributs que s'han utilitzat

| | | 2020 | | | 2021 | | | 2022 | | | 2023 | | |
|----------------------|---------------------------------|----------|---------|-----------|----------|---------|-----------|----------|---------|-----------|----------|---------|-----------|
| | | 5è + alt | mediana | 5è + baix | 5è + alt | mediana | 5è + baix | 5è + alt | mediana | 5è + baix | 5è + alt | mediana | 5è + baix |
| Estudis | Diferència (%) | 6.29 | 4.14 | 1.61 | 6.07 | 3.84 | 1.95 | 5.36 | 3.61 | 1.39 | 5.75 | 3.40 | 1.43 |
| | Valor anual (%) | 52.76 | 29.81 | 8.79 | 53.51 | 30.42 | 8.82 | 54.34 | 31.24 | 9.25 | 54.83 | 31.89 | 9.92 |
| Estructura Domicili | Diferència (%) | 2.04 | 0.37 | -1.47 | 1.71 | 0.34 | -1.17 | 2.38 | 0.65 | -0.49 | 2.13 | 0.79 | -0.35 |
| | Valor anual (%) | 24.86 | 16.33 | 11.59 | 24.58 | 16.03 | 11.70 | 25.62 | 16.22 | 11.30 | 26.34 | 16.71 | 11.94 |
| Regió Naixement | Diferència (%) | 2.44 | 0.95 | -0.07 | 2.63 | 0.99 | 0.06 | 2.38 | 0.85 | -0.23 | 2.52 | 0.77 | -0.05 |
| | Valor anual (%) | 12.32 | 4.64 | 2.08 | 12.98 | 4.74 | 2.01 | 13.39 | 4.99 | 2.25 | 13.53 | 5.17 | 2.20 |
| Grup d'Edat | Diferència (%) | 2.22 | -0.39 | -2.47 | 2.67 | 0.02 | -1.83 | 2.17 | -0.31 | -2.06 | 2.56 | 0.08 | -2.46 |
| | Valor anual (%) | 56.84 | 42.83 | 36.24 | 55.99 | 42.84 | 36.42 | 54.39 | 42.36 | 35.58 | 50.97 | 42.40 | 35.48 |
| Renda anual bruta | Diferència (%) | 18.96 | 14.09 | 10.46 | 21.85 | 17.06 | 12.23 | 18.32 | 13.47 | 9.38 | 12.58 | 9.25 | 2.66 |
| | Valor anual (€) | 36.905 | 18.802 | 10.781 | 38.398 | 19.779 | 11.620 | 38.408 | 19.785 | 11.620 | 38.361 | 19.811 | 11.620 |
| Preu Lloguer | Diferència (%) | 42.09 | 27.38 | 13.46 | 28.05 | 10.92 | -1.79 | 20.88 | 10.31 | -5.01 | 45.64 | 19.81 | 11.17 |
| | Valor anual (€/m ²) | 16.52 | 12.98 | 9.25 | 15.66 | 12.60 | 8.93 | 17.24 | 13.56 | 9.72 | 18.98 | 15.46 | 10.76 |
| Habitatges turístics | Valor anual | 586 | 32 | 0 | 576 | 31 | 0 | 554 | 30 | 0 | 554 | 30 | 0 |

Taula 3.1: Valors extrems i mediana de les variables utilitzades pel càlcul del nivell de gentrificació

per calcular-les (és a dir, sense tenir en compte l'any i el codi del barri, es passa de 13 variables a un total de 7). Finalment, s'atribueix any per any i barri per barri una puntuació de gentrificació consistent en la suma de la resta de puntuacions amb un pes equivalent per a cadascuna d'elles (és a dir, $\frac{1}{7}$). Així, si per un any en concret, un barri ha liderat cadascuna de les set variables, se li haurà atorgat una puntuació de gentrificació de 1 i, en el cas totalment contrari, se li haurà atorgat una puntuació de gentrificació de 0.

Aquests últims resultats amb els valors de cadascuna de les puntuacions de les variables, juntament amb el de la puntuació del nivell de gentrificació, s'emmagatzema en un arxiu `gentrificacio.csv` a la carpeta `data/results/`.

3.1 Determinació del nivell de gentrificació

Arribats a aquest punt, s'han decidit aplicar dues metodologies simultàniament per, finalment, atribuir anualment i per barri un nivell de gentrificació.

La primera d'aquestes metodologies és la de l'aplicació d'un algoritme d'aprenentatge automàtic no supervisat com és *k-means*. En aquest algoritme s'utilitzen totes les puntuacions obtingudes al final de l'últim apartat (descartant no obstant la de la puntuació final de gentrificació la qual, tal i com s'ha obtingut, és una combinació lineal de la resta de puntuacions).

A partir d'aquí, mitjançant tècniques com són la interpretació de l'anàlisi de la silueta i la del mètode del colze, es determina el nombre de *clusters* que finalment es seleccionen.

Adicionalment, degut a que es realitza un aprenentatge automàtic per anys sobre una mateixa població, es tindran en compte els centres calculats dels diferents *clusters* per a cada any per a calcular els nous *clusters* dels anys següents, de manera que es permeti una traçabilitat entre *clusters* més senzilla d'identificar i classificar.

A més a més, també s'aplica un anàlisi de components principals (PCA, de l'anglès *Principal Component Analysis*) amb variació ortogonal varimax, de manera que es conegui la contribució de cada variable en el resultat final. Alhora, s'aplica el test de Esfericitat de Barlett i el test de Kaiser-Meyer-Olkin (KMO) per avaluar l'aplicabilitat del PCA. Per obtenir més detalls del procés, es recomana consultar el codi executat que es pot trobar al repositori de [Github](#).

La segona metodologia aplicada és la proposada novament per [López-Gay et al. \(2019\)](#) on, a partir de les puntuacions de la gentrificació, s'agrupen els barris segons el quartil al qual pertanyen les respectives puntuacions de gentrificació. Al treball de [López-Gay et al. \(2019\)](#) a més, s'aplica una diferència en el primer quartil segons la proporció de persones amb estudis universitaris en l'any d'inici. No obstant, en aquest treball, i per homogeneïtat amb el nombre de *clusters* que s'han obtingut amb l'aplicació primer mètode (tal i com es detalla posteriorment al corresponent [apartat](#) dins del capítol de resultats), s'han escollit només tres grups segons si pertanyen al primer quartil, al segon, o al tercer i quart conjuntament.

Un cop aplicats tots dos mètodes, es comparen els valors de les puntuacions de les variables individuals segons el mètode aplicat, i se'n detallen els aspectes comuns i diferències.

4 Interfície visual interactiva

La metodologia d'aquest treball finalitza amb el procés de realització amb el llenguatge R d'una taula de comandament o *dashboard* interactiu on l'usuari final pot consultar la representació de les diferents variables que s'han anat mencionant al llarg del present treball, de manera que tot modificant els paràmetres d'entrada, es puguin generar noves representacions les quals no s'hagin explicitat en el present document.

Aquest *dashboard* s'ha creat a partir de les característiques i eines que permeten les llibreries *Shiny* i *ShinyDashboard*, aportant de manera senzilla selectors interactius de diferent modalitat, pestanyes per estructurar la informació, espais de representació de gràfics, textos i mapes, entre d'altres. En quant a la representació gràfica, s'han utilitzat fonamentalment les llibreries *Highchart* i *Leaflet* ja que permeten una fàcil interacció amb l'usuari final, tot aportant elements interactius de més senzilla interpretació i interacció. En quants els gràfics creats amb la llibreria *Highchart*, aquests se'ls hi ha afegit la característica de que l'usuari final pot descarregar-los des del mateix visor, amb les dades que en aquell moment s'hi representin.

El *dashboard* final utilitzat en aquest treball obté la informació sobre la qual es fonamenta

dels *datasets* finals que s'han obtingut en cadascun dels apartats anteriors, és a dir, aquella informació continguda a les carpetes *data/L1/**, *data/L2/** i *data/results/**. Tot i que s'amplia la informació més endavant, s'ha decidit que aquest *dashboard* aplegui el contingut de cadascuna d'aquestes carpetes en pestanyes diferents, per tal d'aportar un nivell organitzatiu de la informació disponible a l'usuari final.

Finalment, aprofitant la disponibilitat per compartir aquest *dashboard* de manera que sigui consultable de manera pública i gratuïta sota certes condicions com, entre d'altres, d'un límit d'hores mensuals acumulades de visualització per compte, s'ha escollit el portal *shinyapps.io* com a punt de compartició del *dashboard* realitzat. Aquest es pot consultar a través del següent [enllaç](#) i es mostra una captura d'exemple en la figura 3.3.

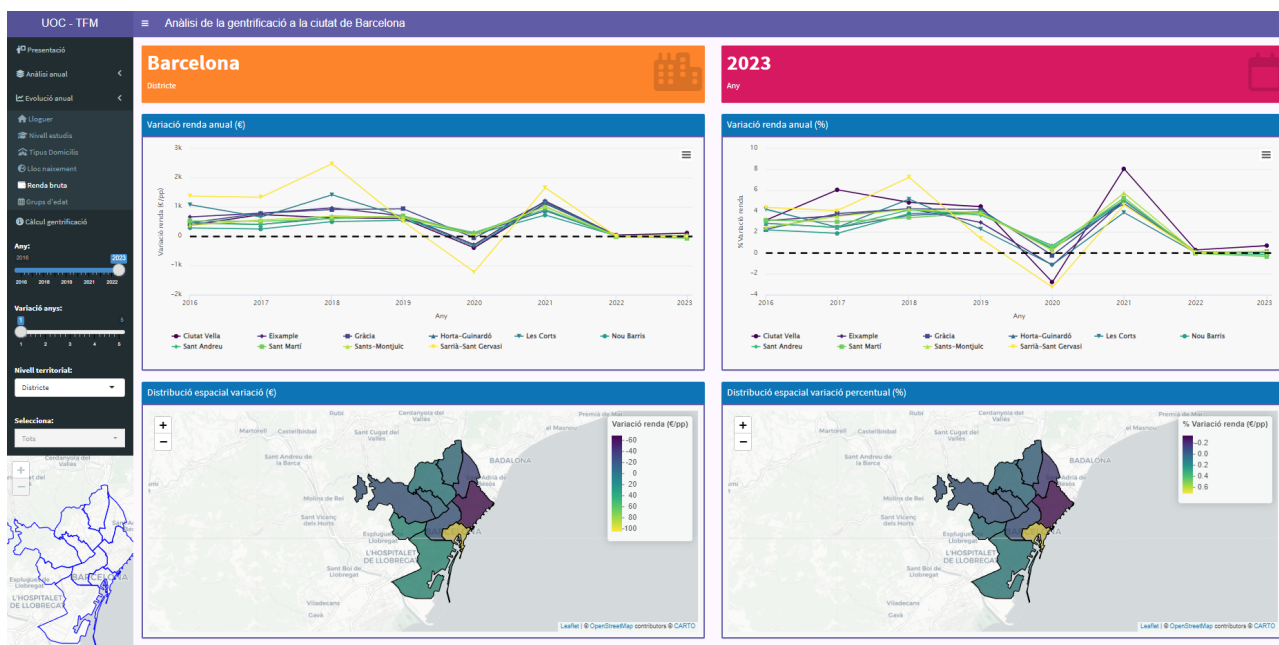


Figura 3.3: Captura de mostra de la interfície visual interactiva. S'hi observa informació sobre la variació anual de les rendes brutes per districtes per a l'any 2023.

Capítol 4

Resultats

Un cop definida la metodologia de treball, en aquest capítol es realitza tant un anàlisi exploratori de les diferents variables, des d'un punt de vista de variabilitat espacial i temporal, que han aportat informació que ha acabat derivant en la determinació d'un nivell de gentrificació a nivell anual i per barri, així com dels càlculs dels nivells de gentrificació pròpiament, i una presentació de la visualització interactiva.

1 Anàlisi exploratori de les dades

Tot i la importància d'analitzar com la situació ha anat canviant al llarg dels anys, en aquesta secció d'anàlisi exploratòria es posa un especial enfoc en la situació de l'any 2023 i les variacions de les seves variables respecte als últims cinc anys, per no estendre en excés la longitud d'aquest document. Alhora, es presenta una imatge general a nivell de ciutat del comportament de cada variable i, si es presenta algun comportament anòmal d'especial interès en algun barri al llarg dels anys, també se'n fa esment.

Tots els gràfics i mapes d'aquest capítol s'han extret de la interfície visual interactiva. Per a altres consultes fora de les contemplades en aquest treball, es convida el lector a accedir a la visualització a través del següent [enllaç](#).

1.1 Nivell d'estudis

Tal i com s'observa a la figura 4.1, per a l'any 2023 el domini dels estudis universitaris o de CFGS de grau superior és majoria a una gran part dels barris de Barcelona. De fet, es pot extreure com aquest domini comença a aparèixer aproximadament a partir de l'any 2010, ja que en termes generals es veu com és l'únic grup de nivell d'estudis que ha tingut un augment sostingut en el percentatge de la població fins a les últimes dades disponibles.

També es remarca una variació espacial clara en quant a la proporció d'aquest nivell d'estudis respecte al total de la població, on en destaquen especialment la Vila Olímpica del Poblenou (56.83%), les Tres Torres (56.25%), la Dreta de l'Eixample (56.09%), Sant Gervasi - Galvany (55.74%), Pedralbes (54.83%) i el Putxet i el Farró (54.06%), entre d'altres.

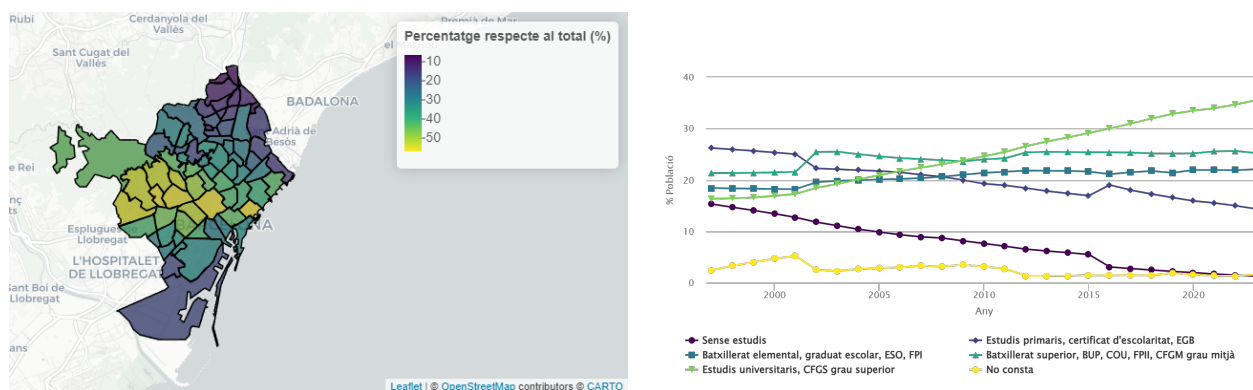


Figura 4.1: Distribució del nivell d'estudis universitaris per barris per a l'any 2023 i l'evolució anual per a tota la ciutat de Barcelona.

En quant a la variació interanual de la població amb estudis universitaris del 2023 respecte als valors de cinc anys enrere (figura 4.2), s'observa com els valors de població més elevats es concentren predominantment en els barris més cèntrics. No obstant, si s'observa la variació en quant a porcentage de població, destaca especialment en els últims dos anys el cas del barri de la Marina del Prat Vermell, on en l'últim any la proporció de la seva població universitària respecte cinc anys enrere ha augmentat un 14.01%, quan comparativament la resta de barris presenten un augment entre el 2% i el 6% aproximadament. A banda queda el barri Gòtic, l'únic que presenta els últims anys una tendència negativa en quant a la proporció d'estudis universitaris (sent pel 2023 una variació del -7.93%).

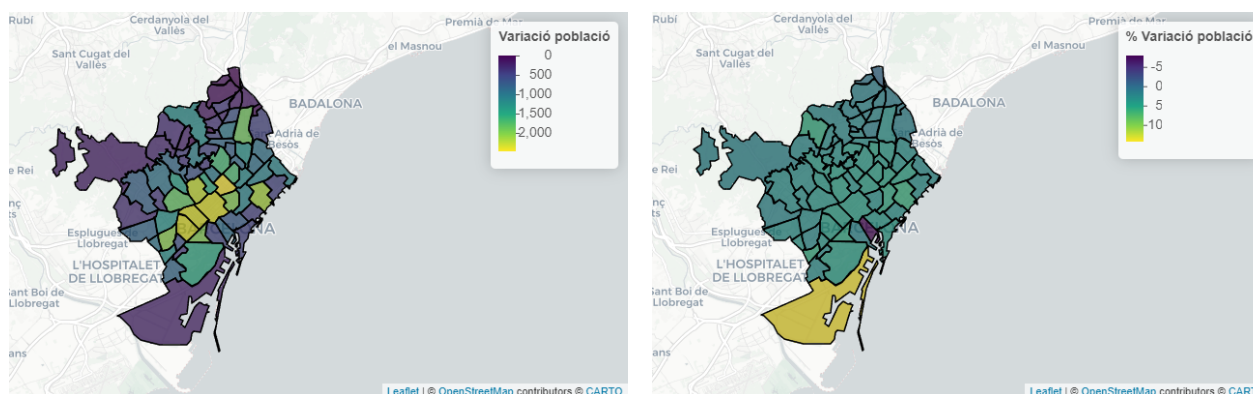


Figura 4.2: Variació absoluta (esquerra) i percentual (dreta) de població amb estudis universitaris per a l'any 2023 a la ciutat de Barcelona.

1.2 Estructura del domicili

En aquest apartat s'analitza la distribució de domicilis on hi habita una sola persona entre els 18 i 64 anys. A la figura 4.3 es pot observar la distribució a nivell de barris per a l'any 2023, així com el pes que té a nivell de ciutat en els últims anys. En general s'observa com el percentatge de domicilis habitats per una persona ha anat augmentant durant la primera dècada del segle XXI, mantenint-se relativament constant, just per sota dels domicilis habitats per dues persones o més on totes són majors de 18 anys. Just són els dos últims anys on aquest percentatge ha augmentat, equiparant-se al de la primera posició.

En la distribució per barris destaquen especialment els barris de la Barceloneta (31.85%), Sant Pere, Santa Caterina i la Ribera (31.66%), el Barri Gòtic (31.42%) el Raval (26.62%) i especialment els dos últims anys la Marina del Prat Vermell (24.63%) els més proporcionalment habitat per persones soles entre els 18 i 64 anys.

L'evolució respecte a cinc anys enrere es pot observar a la figura 4.4. En quant a increments en el nombre de domicilis destaquen especialment els barris del Poble-sec i el Raval (el qual destaca el seu augment en els últims anys). Tenint però en compte els valors percentuals, la Barceloneta (3.82%), Can Peguera (3.49%), Montbau (2.89%) i la Clota (2.13%) encapçalen el rànking dels barris amb més variació. Al respecte de la Clota, els dos últims anys ha passat de tenir una variació negativa del percentatge de domicilis a ocupar les primeres posicions de variació positiva, tot i que la casuística d'aquest barri, amb un volum aproximat de 300 domicilis, petites variacions poden provocar grans dispersions en els resultats obtinguts.

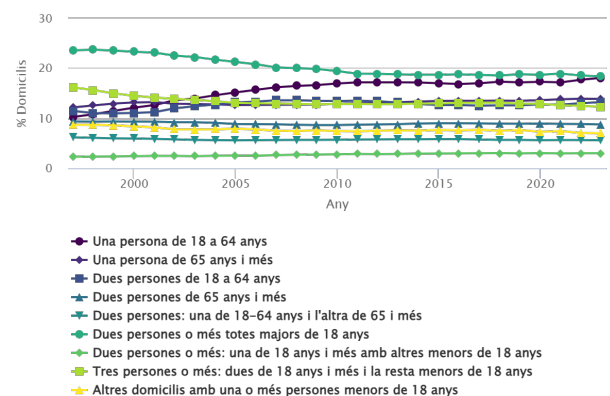
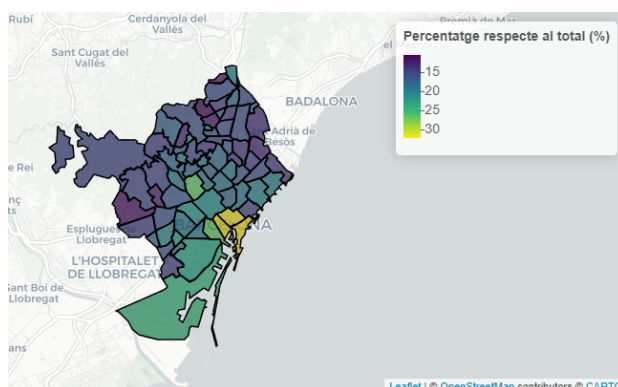


Figura 4.3: Distribució dels domicilis amb una persona de 18 a 64 anys per a l'any 2023 i l'evolució anual per a tota la ciutat de Barcelona.

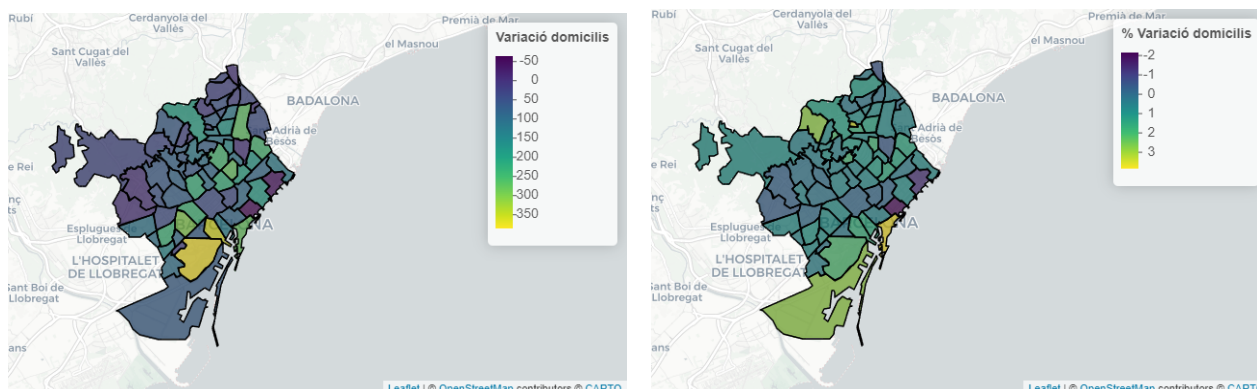


Figura 4.4: Variació absoluta (esquerra) i percentual (dreta) de domicilis habitats per una persona de 18 a 64 anys per a l'any 2023 a la ciutat de Barcelona.

1.3 Regió de naixement d'origen

Les regions de naixement susceptibles de causar processos de gentrificació són aquelles que potencialment tenen un poder adquisitiu més elevat i, en el cas d'aquest treball, s'ha considerat com a tal aquell conjunt de població provinent de la resta d'Europa i d'Amèrica del nord. No obstant això, i tot i que en els càlculs de gentrificació s'han utilitzat els valors de totes dues poblacions conjuntament, en aquest punt es proporciona informació únicament de la població europea (108.014 a data de 2023) degut a la seva predominància sobre la nord-americana (6.736 a data de 2023). Per observar el comportament de la població nord-americana s'anima al lector a consultar la visualització interactiva.

En termes generals la població estrangera dominant a tota la ciutat de Barcelona prové d'Amèrica central/sud, amb les salvetat a l'any 2023 dels barris de Vallvidrera, el Tibidabo i les Planes, Sarrià, Pedralbes, Diagonal Mar i el Front Marítim del Poblenou i la Vila Olímpica del Poblenou, on n'és majoritària la població europea, i els barris del Raval i el Barri Gòtic on hi predomina la població asiàtica.

En quant a la població europea es pot observar la seva distribució en la figura 4.5, on en els barris de Sant Pere, Santa Caterina i la Ribera (20.19%), la Barceloneta (18.5%), el Barri Gòtic (14.43%), la Dreta de l'Eixample (12.75%), Diagonal Mar i el Front Marítim del Poblenou (12.65%) i la Vila Olímpica del Poblenou (12.33%) encapçalen els barris amb més proporció de població europea estrangera.

En quant a les variacions respecte a cinc anys enrere (figura 4.6), el total de població estrangera europea augmenta especialment als barris de la Dreta de l'Eixample, el Poblenou, Sant Gervasi - Galvaní i la Sagrada Família, però en la variació percentual destaca especialment la Marina del Prat Vermell els últims dos anys (3.21%), seguint amb Diagonal Mar i el Front Marítim del Poblenou (2.94%), el Poblenou (2.49%) i la Dreta de l'Eixample (2.41%), éssent els

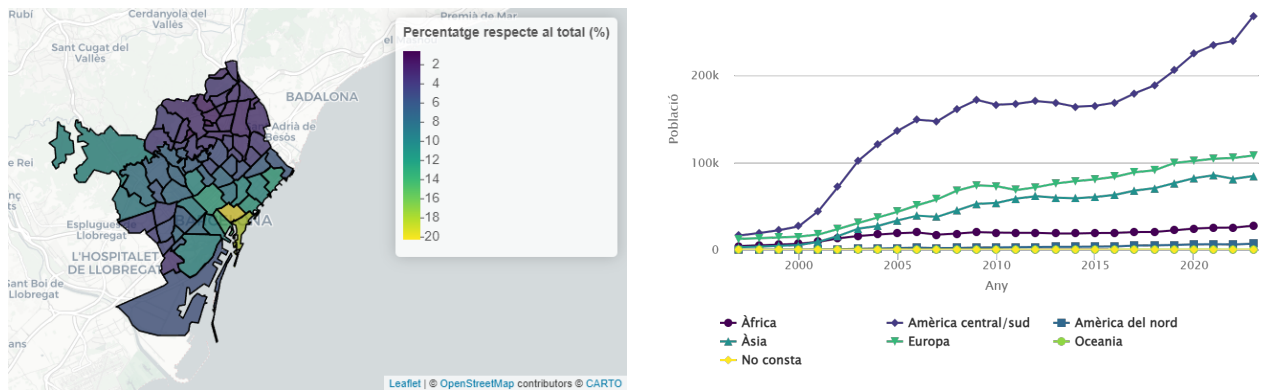


Figura 4.5: Distribució de la població europea per a l'any 2023 i l'evolució anual per a tota la ciutat de Barcelona.

màxims representants d'aquesta variació. Novament, igual que ens les variacions de la població universitària, el Barri Gòtic destaca com el barri amb un descens més sobtat de població estrangera europea (-4.33%), molt lluny de la resta de barris, els quals presenten augments o es mantenen relativament estables.

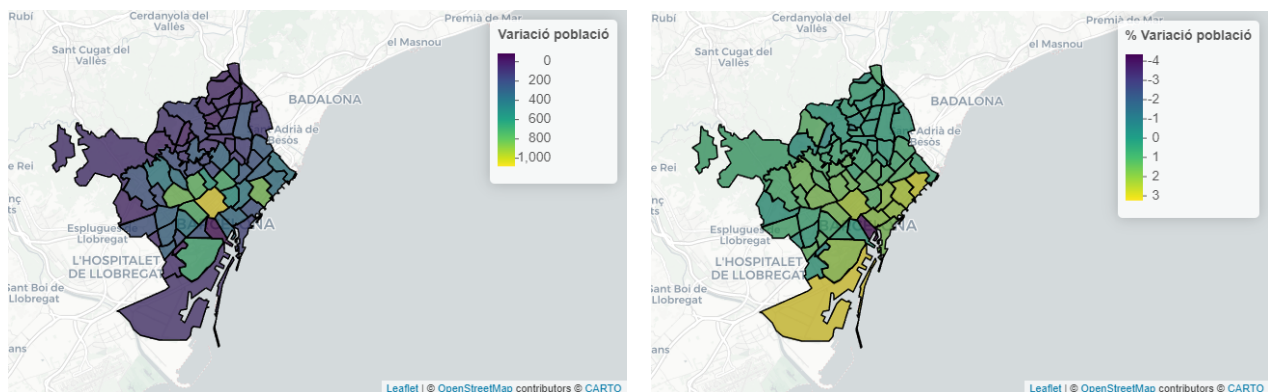


Figura 4.6: Variació absoluta (esquerra) i percentual (dreta) de població estrangera d'origen europeu per a l'any 2023 a la ciutat de Barcelona.

1.4 Grups d'edat

En el global de la ciutat de Barcelona, la població dominant en quant a grups d'edat segueix la distribució de la figura 4.7. Es pot concloure ràpidament com, en termes generals, a la part alta de la ciutat predomina la població de més de 70 anys, a la part més cèntrica i sud predomina la població entre els 30-39 anys, sent la resta de barris generalment dominats per població entre els 40-59 anys. A nivell de ciutat, s'observa com entre l'inici de segle i el 2017 predominava la població entre els 30-39 anys, posició compartida amb les poblacions de 40-49 anys i majors de

70 anys en els últims anys. La població entre els 20-29 anys queda en cinquena posició després d'un fort descens entre el 2003 i 2015, just darrera de la població entre els 50-59 anys.

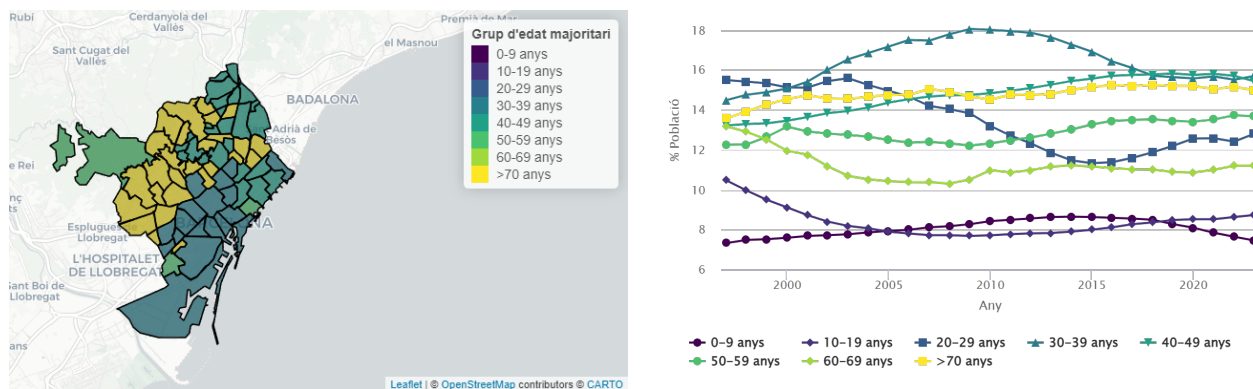


Figura 4.7: Distribució de grups d'edat majoritaris per a l'any 2023 i l'evolució anual per a tota la ciutat de Barcelona.

Les variacions percentuals a nivell de barri es presenten difícils de recollir en un text, doncs són molt heterogènies, però sí que se'n destaca especialment l'augment percentual de la població entre els 20-29 anys ens els barris més centrals de la ciutat (figura 4.8). En termes generals de ciutat, l'augment percentual el lidera la població entre els 20-29 anys (0.93%), seguit de 10-19 anys (0.37%) i 60-69 anys (0.19%) i 50-59 anys (0.16%). La última posició la encapçala la població d'entre 0-9 anys, amb un descens del -1.04%.

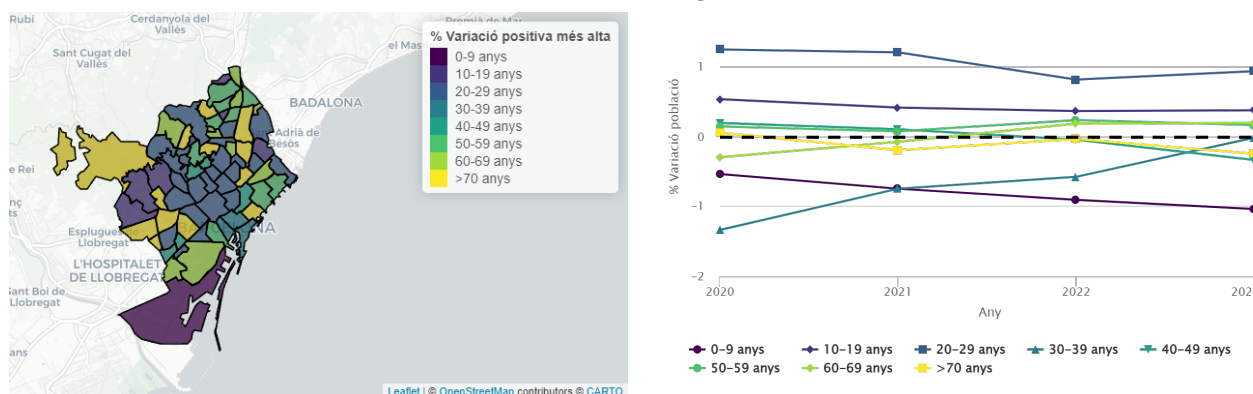


Figura 4.8: Distribució de la variació percentual positiva dominant per grups d'edat (esquerra) i l'evolució d'aquestes variacions (dreta) per a l'any 2023 a la ciutat de Barcelona.

1.5 Renda anual bruta

La distribució de la renda anual bruta per persona a la ciutat de Barcelona per a l'any 2023 es pot observar a la figura 4.9. En ella es destaquen els barris típicament coneguts per les rendes més altes, on barris com les Tres Torres (44.370€), Pedralbes (42.488€), Sant Gervasi - Galvany (39.258€), Sant Gervasi - la Bonanova (38.761€) i Sarrià (38.361€) n'encapçalen la classificació.

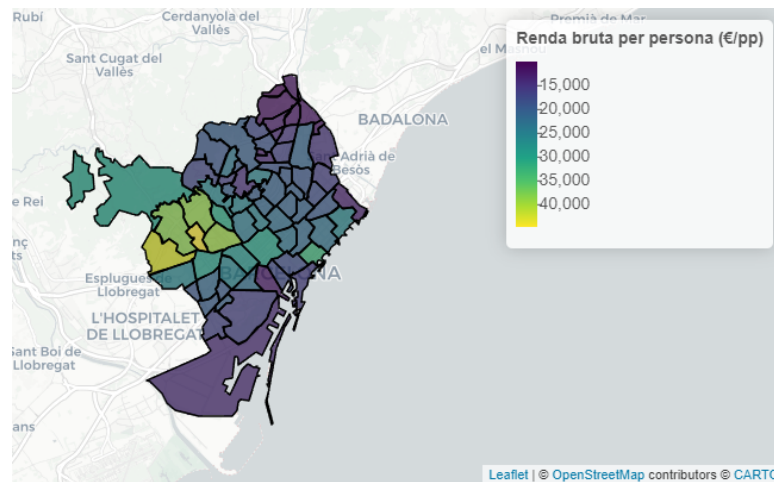


Figura 4.9: Distribució de la renda anual bruta per persona per a l'any 2023 a la ciutat de Barcelona.

En quan a les seves variacions temporals en canvi, són els barris amb rendes més altes els que presenten variacions percentuals més baixes. Són altres barris com la Marina del Port Vermell (19.37%), Torre Baró (16.21%), el Poblenou (13.7%) i la Trinitat Nova (13.49%) els que presenten variacions percentuals més altes (figura 4.10).

1.6 Preu del lloguer

El preu del lloguer d'habitatges és un dels factors que sovint es veu afectat per aquests processos de gentrificació, i analitzar-ne l'evolució és un factor clau per entendre la seva casuística. Analitzant la situació a 2023 del preu del lloguer per metre quadrat (figura 4.11), s'observa com els barris que més destaquen són els de la Bordeta, la Vila Olímpica del Poblenou, Sarrià i les Tres Torres.

Per altra banda s'observen també les variacions percentuals del preu del lloguer a la figura 4.12, on especialment aquest 2023 alguns barris han destacat en aquest increment, on la Marina del Prat Vermell (96.05%), la Clota (93.96%), la Bordeta (73.31%), Can Peguera (47.38%, el qual, a excepció del 2022, destaca des de 2020), i la Verneda i la Pau (45.67%) en són els casos més rellevants.

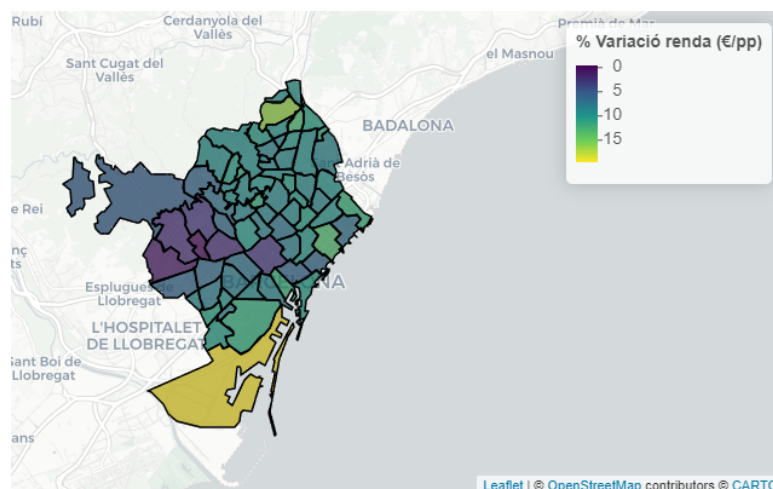


Figura 4.10: Variació percentual de les rendes anuals brutes per a l'any 2023 a la ciutat de Barcelona.

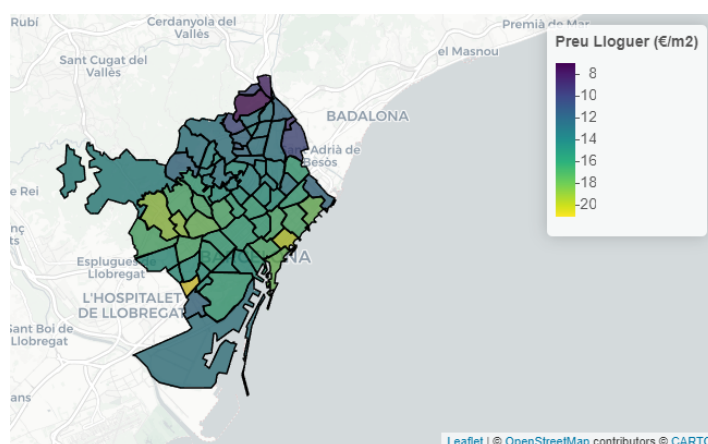


Figura 4.11: Distribució del preu del lloguer per a l'any 2023 a la ciutat de Barcelona.

1.7 Habitatges turístics

La presència d'habitatges turístics a la ciutat té una potencial afectació en els preus dels lloguers de llarga durada, no només pels canvis de preu d'aquests últims, si no per la tipologia de perfil de població que atrau i per l'impacte que això comporta en els processos de gentrificació dels barris, aportant una variable de pressió turística.

La seva distribució al llarg de tot Barcelona és clarament centrada, on el barri de la Dreta de l'Eixample destaca de manera inequívoca respecte a la resta (1.631), seguits pels barris de la Sagrada Família (750), la Vila de Gràcia (725), l'Antiga Esquerra de l'Eixample (670), Sant Antoni (554) i la Nova Esquerra de l'Eixample (520), els quals tenen més de mig miler d'habitatges turístics cadascun. Per l'altra banda, 19 dels 73 barris de Barcelona no tenen cap habitatge turístic, els quals bàsicament són els més allunyats del centre.

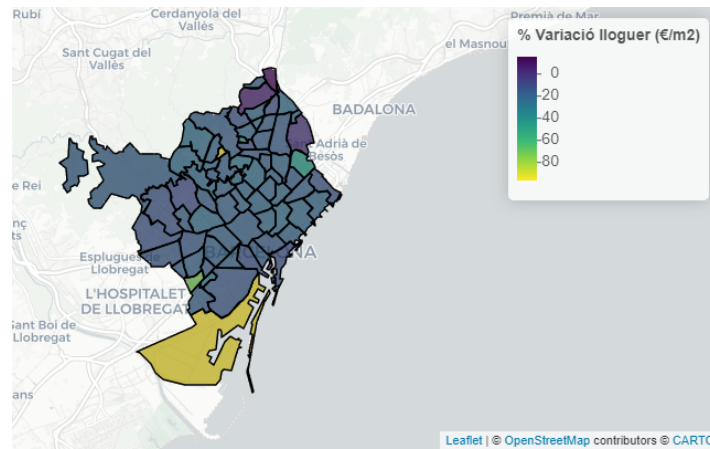


Figura 4.12: Variació percentual dels preus del lloguer per a l'any 2023 a la ciutat de Barcelona.

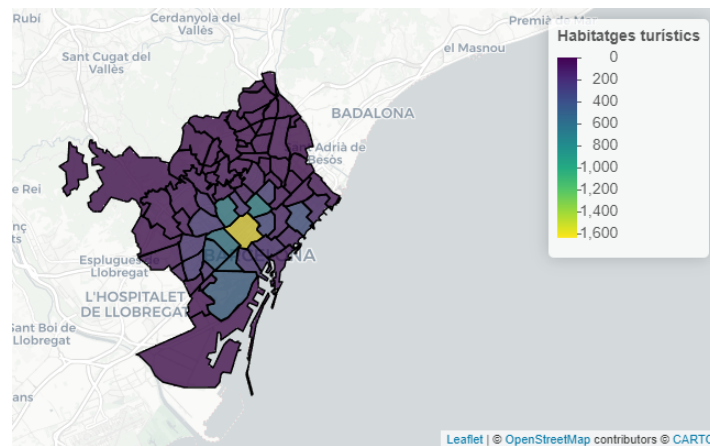


Figura 4.13: Distribució dels habitatges turístics per a l'any 2023 a la ciutat de Barcelona.

2 Nivells de gentrificació

Un cop s'han analitzat cadascuna de les variables, el següent pas consisteix en la classificació dels barris segons els diferents mètodes proposats, és a dir, mitjançant l'aplicació d'un algoritme d'aprenentatge automàtic no supervisat com és *k-means*, i el segon seguint l'exemple del treball de López-Gay et al. (2019) on, a partir de les puntuacions atorgades a cadascuna de les variables estudiades per cada any, es calcula una nova puntuació de la gentrificació i, tot aplicant grups per quantils, s'identifiquen els diferents nivells de gentrificació.

Seguint el procés de *k-means*, s'han determinat un total de 3 *clusters* a partir del mètode de la silueta i del colze, els quals es poden observar a la figura 4.14. En aquesta figura només es mostren els resultats d'aquests dos mètodes utilitzant les dades del 2023, però els resultats que s'han realitzat de manera equivalent per a la resta d'anys retornen un valor de *clusters* similars, i per homogeneïtzar tots els anys s'han escollit el nombre indicat.

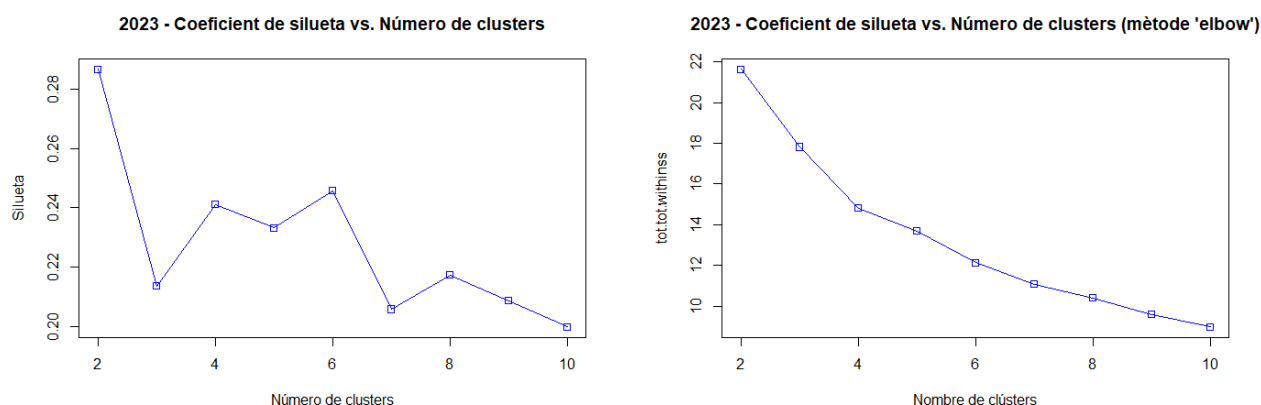


Figura 4.14: Determinació del nombre de *clusters* a utilitzar a partir del mètode de la silueta (esquerra) i el del colze (dreta) per a les dades del 2023. Elaboració pròpia.

En el cas de *k-means*, degut a que s'estan utilitzant les mateixes dades però per anys consecutius, s'utilitzen els centres resultants dels *clusters* dels primer any com a centres del següent any, i així consecutivament, de manera que un etiquetatge d'un grup de dades s'utilitzi en els anys posterior per a dades similars. Es pot obtenir més informació consultant el repositori de [Github](#).

Degut als tres agrupament de barris que s'utilitzen segons el mètode de *k-means*, per homogeneïtzar els resultats s'escull el mateix nombre de *clusters* per al mètode de la puntuació de gentrificació, escollint de més gentrificat a menys gentrificat les puntuacions de gentrificació d'aquells barris que estiguin en el primer quartil, el segon, i el tercer i el quart quartil conjuntament, respectivament.

Els resultats d'aquesta classificació per als dos mètodes es poden observar a la figura 4.15. En general s'observen moltes similituds entre els dos mètodes, tot i que analitzant amb més detall per anys, la classificació per puntuació de gentrificació accentua més les diferències entre el primer grup gentrificat i els barris sense dinàmiques gentrificants, causant que el nombre de barris amb una lleugera tendència gentrificadora sigui reduït. Per altra banda, el mètode de *k-means* aparentment classifica de manera similar els barris més gentrificats al del primer mètode per a tots els anys, però respecte al primer mètode augmenta el nombre de barris amb classificació intermitja, reduint el nombre de barris sense processos de gentrificació. Però en termes globals, s'observen moltes similituds i són pocs els barris que per un mateix any són classificats en categories oposades, com a molt només presentant un nivell de diferència cap amunt o cap avall segons cada mètode.

Els resultats d'aquestes dues classificacions es poden observar, pels valors calculats de l'any 2023, a la taula 4.1.

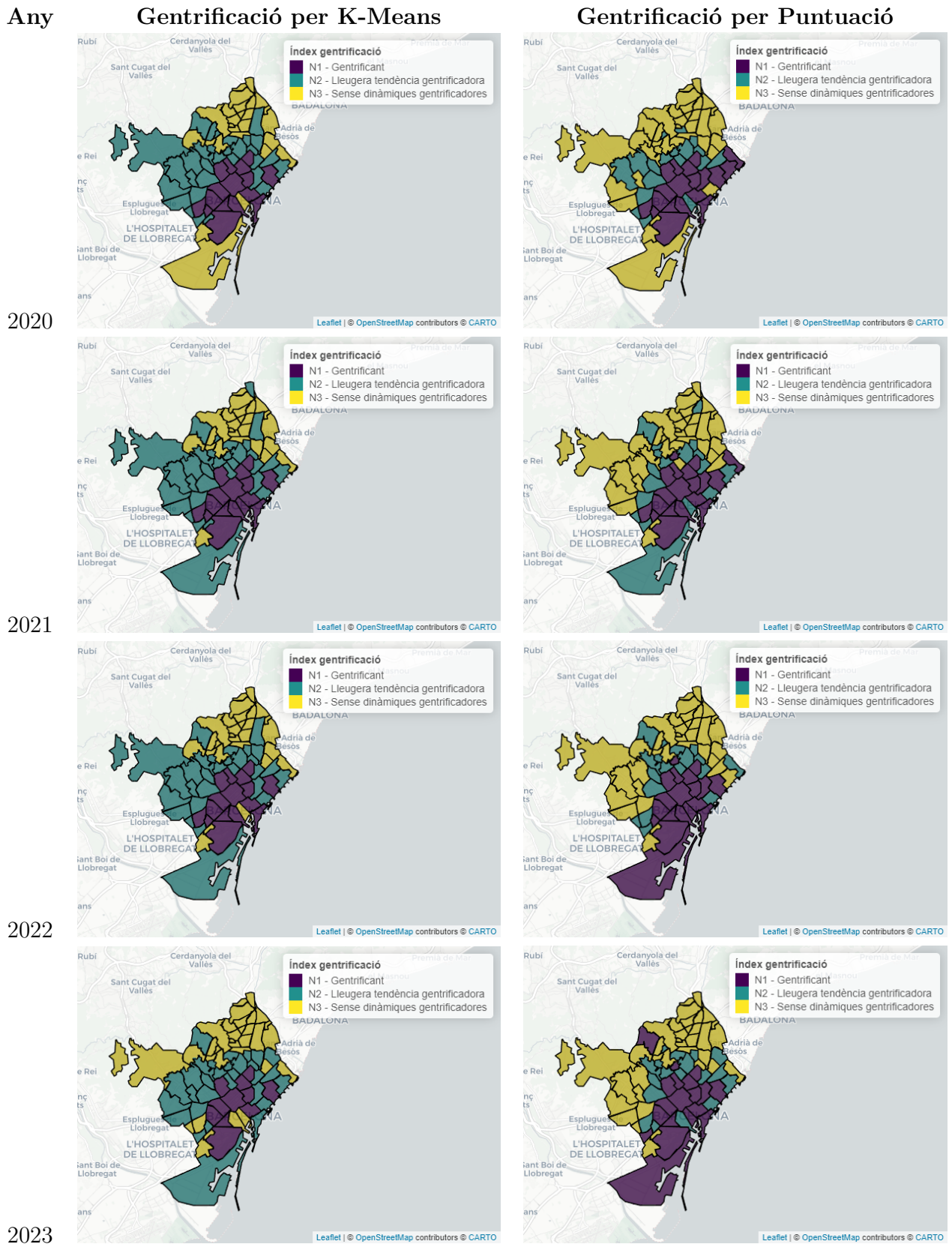


Figura 4.15: Distribució dels nivells de gentrificació per barris pel mètode de *k-means* (esquerra) i per puntuació de gentrificació (dreta).

| Posició | Barri | Classificació K-Means | Puntuació Gentrificació | Posició | Barri | Classificació K-Means | Puntuació Gentrificació |
|---------|--|-----------------------|-------------------------|---------|---------------------------------------|-----------------------|-------------------------|
| 1 | la Vila de Gràcia | N1 | 0.76 (N1) | 38 | el Besòs i el Maresme | N3 | 0.41 (N3) |
| 2 | el Poble-sec | N1 | 0.75 (N1) | 39 | Sants - Badal | N3 | 0.39 (N3) |
| 3 | l'Antiga Esquerra de l'Eixample | N1 | 0.73 (N1) | 40 | Navas | N2 | 0.39 (N3) |
| 4 | el Poblenou | N1 | 0.71 (N1) | 41 | les Corts | N2 | 0.37 (N3) |
| 5 | la Marina del Prat Vermell | N2 | 0.7 (N1) | 42 | la Font de la Guatlla | N3 | 0.36 (N3) |
| 6 | la Barceloneta | N2 | 0.65 (N1) | 43 | Horta | N3 | 0.36 (N3) |
| 7 | la Sagrada Família | N1 | 0.65 (N1) | 44 | Sant Gervasi - la Bonanova | N2 | 0.34 (N3) |
| 8 | la Dreta de l'Eixample | N1 | 0.65 (N1) | 45 | Sant Martí de Provençals | N3 | 0.34 (N3) |
| 9 | Sant Antoni | N1 | 0.65 (N1) | 46 | la Maternitat i Sant Ramon | N2 | 0.33 (N3) |
| 10 | la Nova Esquerra de l'Eixample | N1 | 0.61 (N1) | 47 | la Teixonera | N3 | 0.33 (N3) |
| 11 | la Bordeta | N2 | 0.6 (N1) | 48 | Porta | N3 | 0.33 (N3) |
| 12 | el Barri Gòtic | N3 | 0.57 (N1) | 49 | el Turó de la Peira | N3 | 0.33 (N3) |
| 13 | el Camp de l'Arpa del Clot | N1 | 0.57 (N1) | 50 | el Carmel | N3 | 0.32 (N3) |
| 14 | el Parc i la Llacuna del Poblenou | N2 | 0.57 (N1) | 51 | Can Peguera | N3 | 0.32 (N3) |
| 15 | Sant Pere, Santa Caterina i la Ribera | N1 | 0.56 (N1) | 52 | la Sagrera | N2 | 0.32 (N3) |
| 16 | Can Baró | N2 | 0.56 (N1) | 53 | les Tres Torres | N2 | 0.31 (N3) |
| 17 | el Fort Pienc | N1 | 0.54 (N1) | 54 | Sarrià | N2 | 0.3 (N3) |
| 18 | el Camp d'en Grassot i Gràcia Nova | N2 | 0.54 (N1) | 55 | la Vall d'Hebron | N3 | 0.3 (N3) |
| 19 | Montbau | N2 | 0.53 (N1) | 56 | Vilapicina i la Torre Llobeta | N3 | 0.3 (N3) |
| 20 | Hostafrancs | N2 | 0.52 (N2) | 57 | Sant Andreu | N3 | 0.3 (N3) |
| 21 | el Clot | N2 | 0.51 (N2) | 58 | Verdun | N3 | 0.29 (N3) |
| 22 | la Clota | N3 | 0.5 (N2) | 59 | la Marina de Port | N3 | 0.27 (N3) |
| 23 | la Vila Olímpica del Poblenou | N2 | 0.5 (N2) | 60 | Vallvidrera, el Tibidabo i les Planes | N3 | 0.27 (N3) |
| 24 | Sants | N3 | 0.48 (N2) | 61 | la Prosperitat | N3 | 0.27 (N3) |
| 25 | Sant Gervasi - Galvany | N2 | 0.48 (N2) | 62 | la Trinitat Nova | N3 | 0.27 (N3) |
| 26 | el Putxet i el Farró | N2 | 0.47 (N2) | 63 | la Trinitat Vella | N3 | 0.25 (N3) |
| 27 | la Salut | N2 | 0.46 (N2) | 64 | Sant Genís dels Agudells | N3 | 0.24 (N3) |
| 28 | el Congrés i els Indians | N2 | 0.46 (N2) | 65 | les Roquetes | N3 | 0.24 (N3) |
| 29 | el Guinardó | N2 | 0.45 (N2) | 66 | Torre Baró | N3 | 0.24 (N3) |
| 30 | Vallcarca i els Penitents | N2 | 0.44 (N2) | 67 | el Bon Pastor | N3 | 0.24 (N3) |
| 31 | la Font d'en Fargues | N2 | 0.44 (N2) | 68 | Pedralbes | N2 | 0.22 (N3) |
| 32 | Diagonal Mar i el Front Marítim del Poblenou | N2 | 0.44 (N2) | 69 | la Guineueta | N3 | 0.22 (N3) |
| 33 | el Coll | N2 | 0.43 (N2) | 70 | Baró de Viver | N3 | 0.21 (N3) |
| 34 | el Baix Guinardó | N2 | 0.43 (N2) | 71 | Canyelles | N3 | 0.19 (N3) |
| 35 | la Verneda i la Pau | N2 | 0.43 (N2) | 72 | Ciutat Meridiana | N3 | 0.16 (N3) |
| 36 | el Raval | N3 | 0.42 (N2) | 73 | Vallbona | N3 | 0.15 (N3) |
| 37 | Provençals del Poblenou | N2 | 0.42 (N2) | | | | |

Taula 4.1: Classificacions dels nivells de gentrificació per barris pel mètode de *k-means* i per puntuació de gentrificació per a l'any 2023. S'ordenen segons la puntuació de gentrificació.

Seguint com a referència de la puntuació de gentrificació, els barris més gentrificats a la ciutat de Barcelona per a l'any 2023 són la Vila de Gràcia, el Poble-sec, l'Antiga Esquerra de l'Eixample, el Poblenou, la Marina del Prat Vermell, la Barceloneta, la Sagrada Família, la Dreta de l'Eixample, Sant Antoni, la Nova Esquerra de l'Eixample, la Bordeta i el Barri Gòtic, entre d'altres, la majoria d'ells classificats per tots dos mètodes com a barris gentrificats.

En quant a evolucions destacades en el nivell de gentrificació, destaca especialment el cas de la Marina del Prat Vermell, el qual ha passat d'un barri per ambdós mètodes classificat com que no tenia dinàmiques gentrificadores al 2020, a ser un alt candidat per a ser un barri gentrificat, causat principalment tant per motius d'augment de població universitària, augment sobtat els dos últims anys del preu del lloguer i una incorporació de població europea estrangera. La resta de barris es mantenen relativament estables en quant a nivell de gentrificació, cadascun respecte als respectius mètodes de classificació.

El següent punt a analitzar és la distribució de les puntuacions de les diferents variables (es recorda, renda, pisos turístics, domicilis unifamiliars, edat entre els 20-50 anys, lloc naixement Europa - Amèrica del nord, població universitària i lloguer) afegint la puntuació de gentrificació calculada anteriorment.

A les figures 4.16 i 4.17 es poden observar aquestes distribucions en forma de diagrama de caixa per a cada nivell de gentrificació calculat segons el mètode dels *k-means* o de la puntuació de gentrificació, respectivament. Es detallen les següents característiques:



Figura 4.16: Distribució de les puntuacions segons el nivell de gentrificació calculat pel mètode *k-means* dels barris de Barcelona per al 2023.

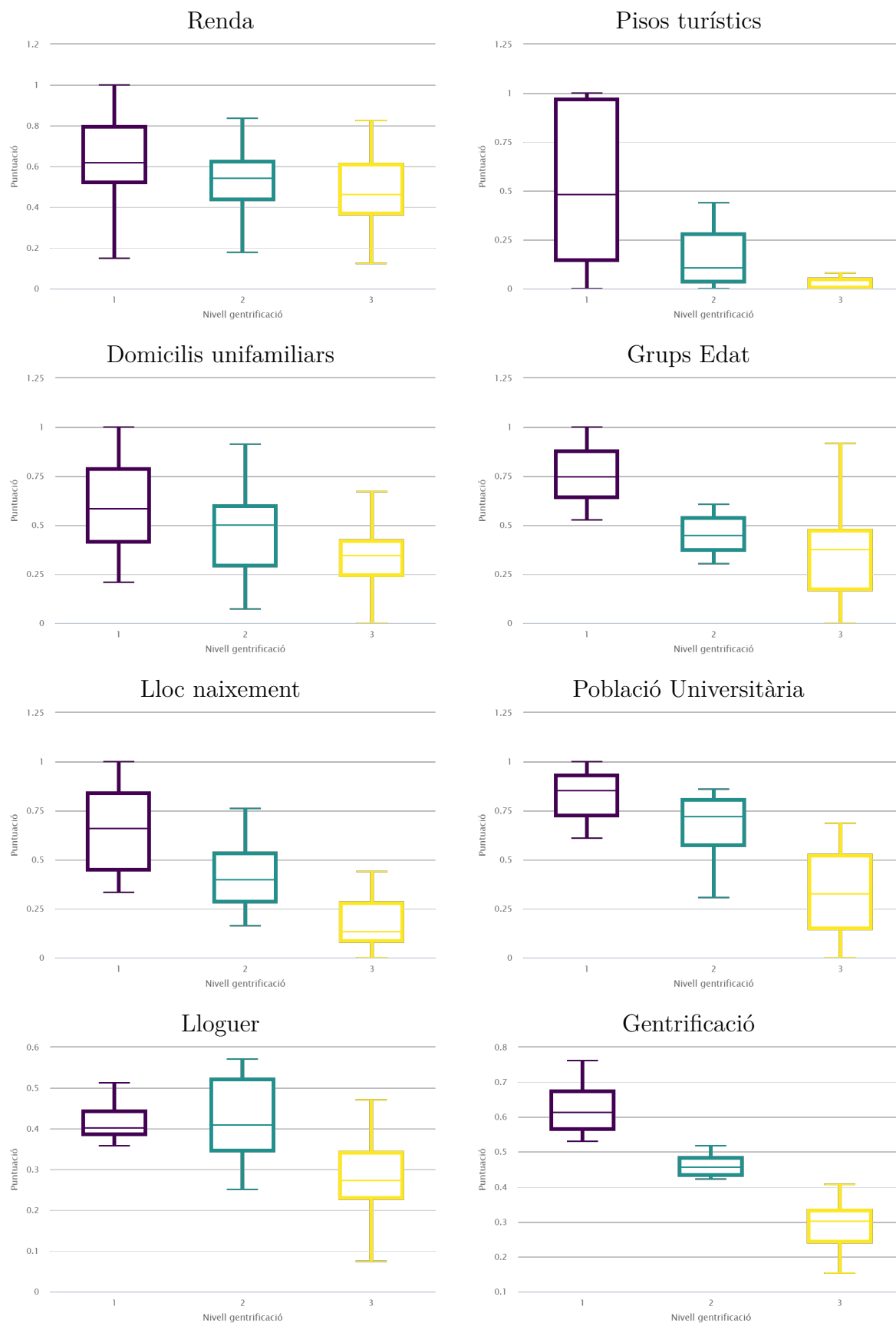


Figura 4.17: Distribució de les puntuacions segons el nivell de gentrificació calculat pel mètode de puntuació de gentrificació dels barris de Barcelona per al 2023.

- **Renda:** Cap dels dos mètodes veu gaire diferència en quant a la puntuació de la renda bruta segons els barris classificats de maneres diferents. El mètode de puntuació de gentrificació té una lleugera tendència en l'augment de la puntuació de la renda com més gentrificat està, però la diferència és molt petita.
- **Pisos turístics:** Tots dos mètodes veuen una clara diferència entre els barris, on com més gentrificat està un barri més pisos turístics conté.
- **Domicilis unifamiliars:** Comportament similar a la renda, on per *k-means* té un comportament similar entre classificacions, però pel mètode de puntuació de gentrificació, la puntuació dels domicilis augmenta lleugerament amb el nivell de gentrificació, tot i que aquí les diferències es veuen més clares que en el cas anterior.
- **Grups d'edat:** Comportament similar en tots dos mètodes, on la puntuació dels grups d'edat augmenta segons el nivell de gentrificació.
- **Lloc de naixement:** Una altra característica de comportament compartit entre mètodes, amb una puntuació que augmenta segons el nivell de gentrificació.
- **Població universitària:** La puntuació d'aquesta variable també és creixent segons el nivell de gentrificació, tot i que en tots dos mètodes els dos primers nivells presenten diferències més petites respecte al tercer nivell.
- **Lloguer:** Tots dos mètodes tenen un comportament similar, en el què la puntuació pràcticament no canvia entre el primer i segon nivell de gentrificació, on de fet els barris amb una lleugera tendència gentrificadora arriben a tenir una puntuació més elevada.
- **Puntuació de gentrificació:** Aquesta puntuació final, com és d'esperar del segon mètode per la seva definició, però també pel mètode de *k-means*, la seva puntuació augmenta segons el seu nivell de gentrificació.

3 Interfície visual interactiva

L'últim apartat d'aquest capítol consisteix en descriure la visualització interactiva creada amb les llibreries *Shiny* i *ShinyDashboard* del llenguatge R, el qual es recorda es pot visitar a través del següent [enllaç](#). La pàgina web que aquí s'hi pot consultar es divideix en quatre parts.

La primera part i pestanya que es trobarà el visitant és la pàgina de presentació (figura 4.18), on se l'introduirà en la pàgina, se li presenten els objectius que aquesta persegueix, i li detallarà alguns consells bàsics per moure's dins de les diferents pestanyes. Entre d'altres,

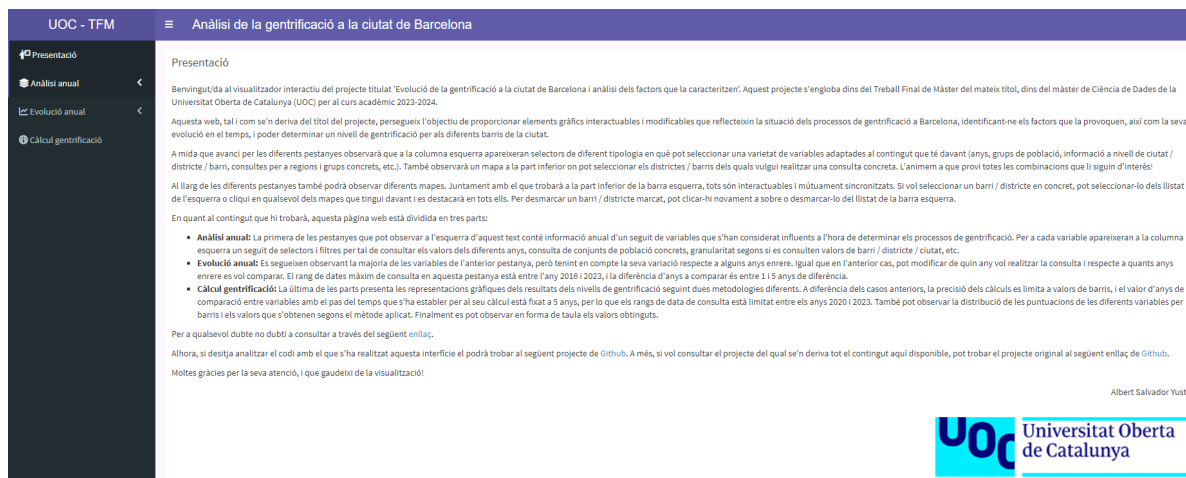


Figura 4.18: Captura de la pestanya 'presentació' de la pàgina web.

es destaca que a totes les següents seccions d'aquesta pàgina web hi apareixeran mapes amb els nivells territorials seleccionats (ciutat / districte / barri) en forma de polígons (així com també a la part inferior de la columna de l'esquerra). Tots aquestes mapes són interactuables i, a més d'obtenir-ne informació seleccionada a mida que es passa el cursor per sobre, es poden seleccionar i deseleccionar barris perquè aquests siguin destacats i, en cas que s'hagi contemplat així, els diferents gràfics s'adaptaran a les regions seleccionades. La selecció dels diferents barris i districtes s'hereta entre consultes, provocant que tots els mapes estàn sempre sincronitzats.

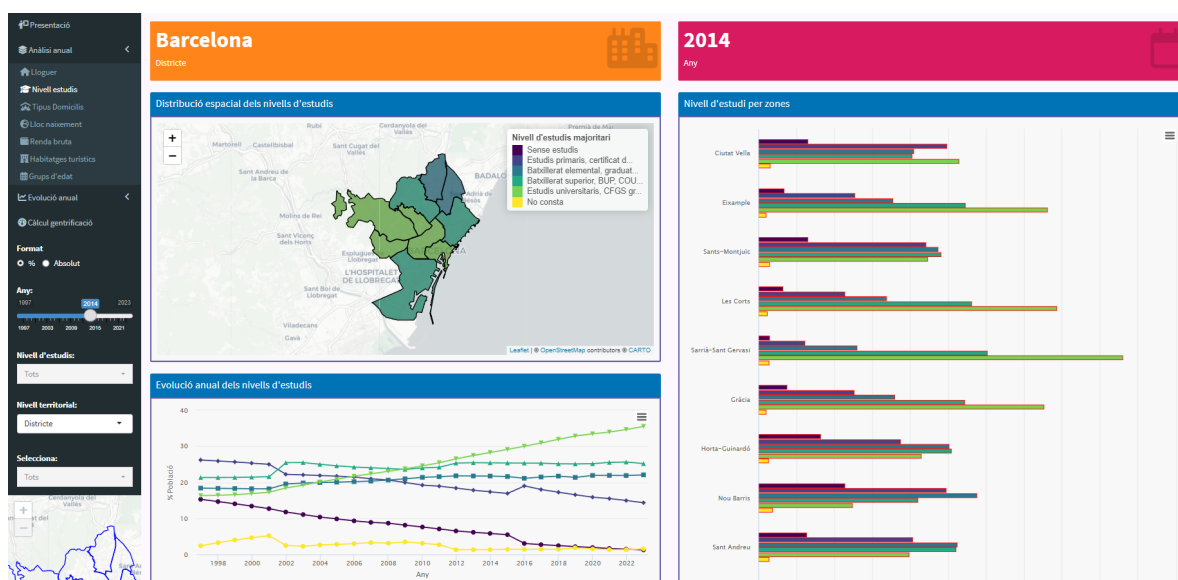


Figura 4.19: Captura de la pestanya 'anàlisi anual' de la pàgina web, amb informació dels nivells d'estudi per districtes per a l'any 2014.

La segona part consisteix en l'anàlisi anual de les diferents variables. Es pot consultar

una captura d'exemple a la figura 4.19. Un cop entra en aquesta pestanya, al usuari se li obre un selector de les diferents variables estudiades en aquest treball. Cadascuna d'elles obre una pestanya diferent amb informació relacionada on, entre d'altres, s'hi conté un mapa amb la distribució de les diferents variables, així com un gràfic vertical de barres on, per el nivell territorial seleccionat, es poden veure els diferents grups que s'inclouen en cadascuna de les variables. Es pot observar també a la barra de l'esquerra un seguit de selectors i filtres que l'usuari pot modificar de manera que els gràfics s'adaptin als paràmetres seleccionats. Exceptuant la variable dels habitatges turístics, la resta de variables presenta a més una evolució al llarg dels anys dels quals es disposen dades.

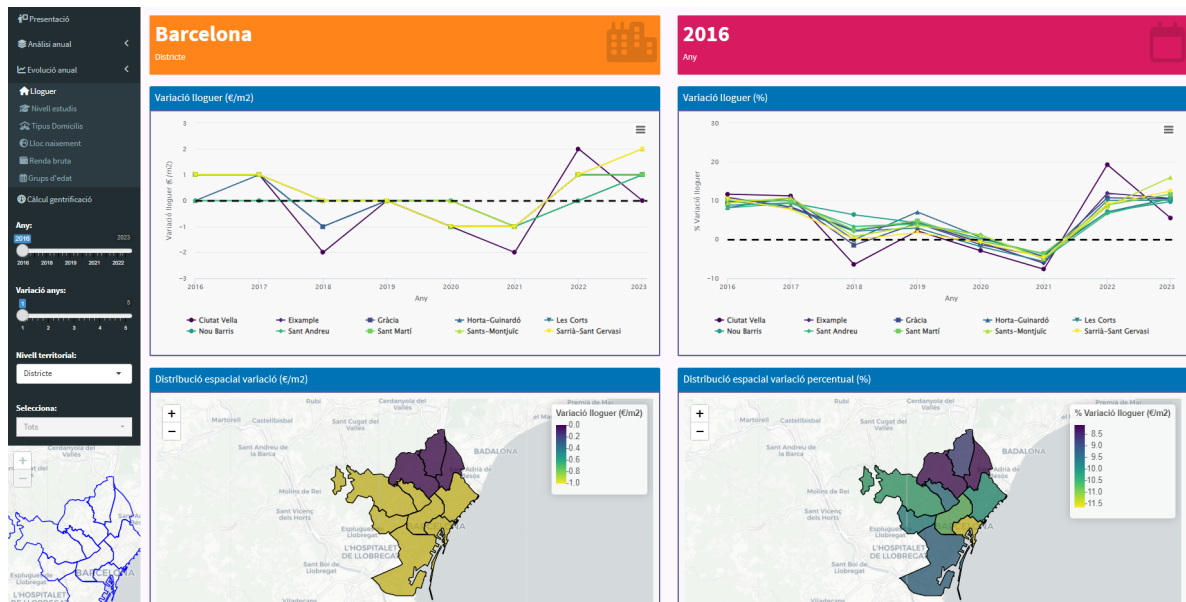


Figura 4.20: Captura de la pestanya 'evolució anual' de la pàgina web, amb informació dels preus del lloguer per districtes per a l'any 2016.

La tercera part contempla l'evolució anual de les diferents variables (figura 4.20). Degut a la seva pràcticament nul·la variació, els habitatges turístics no s'inclouen en aquest apartat. En ella s'estudien les variacions de les diferents variables, on l'usuari pot modificar, addicionalment als selectors que pot trobar a l'anterior part, un selector de variació d'anys, el qual implica observar la variació d'una variable respecte al nombre d'anys seleccionats en el indicat selector. Degut a que per aquesta part només es disposa d'informació des del 2015 fins al 2023 i la variació d'anys contemplada està entre 1 i 5 anys, el selector de l'any de consulta pot variar segons el nombre de variació d'anys des del 2016-2023 (per a una variació d'anys de 1) fins al període 2020-2023 (per a una variació d'anys de 5). En aquesta pestanya es pot observar a la part superior les variacions absolutes (esquerra) i percentuals (dreta) de les diferents variables contemplades, i a la part inferior les respectives representacions gràfiques en forma de mapa.

Finalment a l'última part l'usuari pot consultar els resultats d'aplicar els diferents mètodes de càlcul de gentrificació (figura 4.21). Els controls de la barra esquerra es limiten a seleccionar l'any de consulta entre el 2020 i el 2023, i poder destacar els barris. A la part superior es pot veure la classificació segons els dos mètodes per a l'any seleccionat, i just a sota la distribució de les diferents puntuacions anteriorment calculades segons els nou selector que apareix just a sobre. Més enllà de que tots els gràfics s'actualitzen segons l'any de consulta, en aquesta segona fila de gràfics, el mapa de l'esquerra només varia segons la variable seleccionada, mentre que el gràfic de diagrama de caixes varia tant amb el de variable com el de mètode seleccionat. Finalment sota d'aquests, l'usuari pot consultar una taula amb els diferents resultats de puntuació obtinguts.

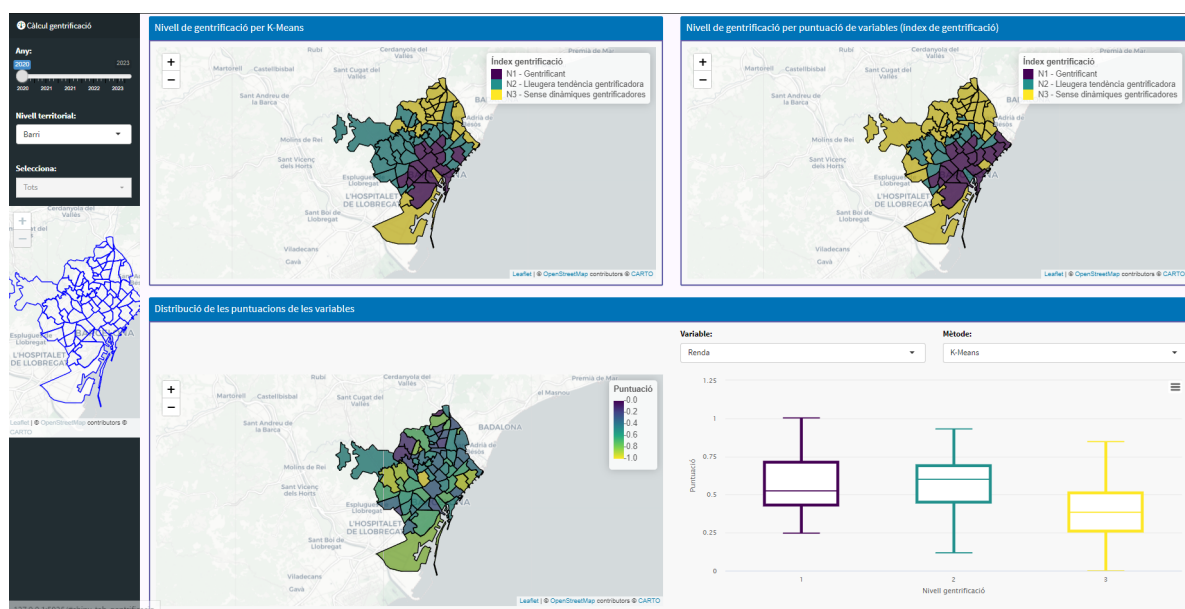


Figura 4.21: Captura de la pestanya 'càlcul gentrificació' de la pàgina web, amb informació dels nivells de gentrificació per a l'any 2020.

Les diferents pestanyes utilitzen els resultats obtinguts durant el desenvolupament del projecte. D'aquesta manera, les pestanyes de 'Anàlisi anual' s'alimenten dels resultats emmagatzemats a la carpeta `data/L1/*`, les pestanyes de 'Evolució anual' dels de la carpeta `data/L2/*` i, finalment la pestanya 'Càlcul gentrificació' utilitza les de la carpeta `data/results/*`. Al respecte d'aquesta última, la classificació per *k-means* es realitza en temps real.

Finalment destacar que, exceptuant els mapes, tots els gràfics són exportables en els formats PNG, JPEG, PDF i SVG, seleccionant les tres línies horitzontals de cada gràfic i seleccionant-ne la opció corresponent en cada cas. Per consultar tot el codi relacionat amb l'elaboració d'aquesta visualització, es pot consultar a través del següent enllaç de [Github](#).

Capítol 5

Conclusions i treball futur

El present projecte representa la culminació de múltiples coneixements adquirits al llarg del màster en Ciència de Dades de la Universitat Oberta de Catalunya. Entre d'altres, es destaquen els processos de captura, neteja i anàlisi de les dades utilitzades, representació gràfica, aplicació d'algoritmes de classificació no supervisada, així com el desenvolupament d'una interfície gràfica on exposar tot el material utilitzat i resultats obtinguts.

Sota aquesta premisa, aquest treball ha perseguit l'objectiu de realitzar una classificació del nivell de gentrificació en els diferents barris de Barcelona a partir d'un seguit de variables determinades segons la documentació detallada a l'estat de l'art d'aquest projecte. Alhora, tota la informació de les dades utilitzades, juntament amb els resultats que se n'han obtingut, han derivat en la creació d'una interfície visual interactiva en què el lector interessat pot consultar tots els resultats aquí exposats, així com realitzar consultes noves no plantejades en aquest treball. Es destaca la disponibilitat pública de totes les dades utilitzades per a una reproductibilitat accessible per al lector interessat.

La classificació dels nivells de gentrificació s'ha realitzat seguint dues metodologies de manera simultània. L'inici és comú, on s'avalua anualment entre el període 2020-2023 la situació actual de cada variable i la variació respecte als últims anys, donant més pes a aquesta última degut al procés transformatiu de la gentrificació. A partir dels resultats obtinguts, s'agrupen els barris segons la tècnica de classificació *k-means*, i alhora segons una puntuació de gentrificació calculada per combinació lineal de la resta de variables utilitzades, i classificant els diferents barris segons el quartil al qual pertanyen. Un cop classificats els barris segons els dos mètodes, s'analitzen les seves característiques comunes i divergents. En quant als resultats obtinguts, es destaquen els barris més centrals de la ciutat com als més afectats pels processos de gentrificació per a tots els anys d'estudi.

Totes les tasques de programació realitzades en aquest treball han estat majoritàriament a través del llenguatge de programació R, tant des de l'obtenció de la majoria de les dades, el

seu tractament i anàlisi, així com la implementació de les diferents tècniques de classificació aplicades. Alhora, tota la visualització interactiva s'ha realitzat utilitzant les llibreries *Shiny* i *ShinyDashboard*, el qual conjuntament amb els gràfics de la llibreria *Highchart* i els mapes de la llibreria *Leaflet* han permès la creació d'una eina amb una alta interactivitat, facilitat de consulta i extracció de resultats.

Revisitant els [objectius](#) d'aquest treball, es pot observar com la majoria d'ells, tant principals com parcials, s'han aconseguit assolir durant el desenvolupament del projecte. No obstant, tot i que s'ha analitzat el comportament al llarg del temps de les diferents variables utilitzades, així com de la classificació dels barris segons el nivell de gentrificació per a cada any d'estudi, ha restat pendent l'aplicació de tècniques de predicció per tal de poder avançar els respectius nivells de gentrificació segons l'evolució potencialment prevista de la resta de variables. Aquest punt sense dubte seria un aspecte a analitzar en un treball futur. Alhora, ha mancat un anàlisi més acurat a nivell de barri de les diferents variables, doncs la descripció d'aquestes s'ha realitzat més per agrupació de barris segons el nivell de gentrificació.

La planificació prevista durant l'inici del treball s'ha seguit de la manera més curosa que ha estat possible. Es destaca no obstant la part de disseny i implementació del treball, en la qual van anar apareixent imprevistos que van reduir els temps contemplats per a altres tasques. Així, la selecció final de les variables, l'obtenció d'algunes d'elles, i finalment el tractament de les dades i la presa de decisió de quin o quins mètodes de classificació es seguirien, van demorar l'inici de la creació de la visualització interactiva, la qual ha ocupat gran part del temps disponible. L'assoliment final d'aquesta, creada i dissenyada des de zero, ha sigut possible gràcies a la incorporació de *sprints* en els processos finals, incorporant, descartant i procurant millorar elements interactius segons s'anava observant i analitzant l'"aspecte i tacte" (traducció de l'expressió anglesa *look and feel*) de la pàgina web.

En quant als ODS plantejats al [capítol](#) inicial, tot i que les conclusions descrites no es vinculen de manera directa amb ells, es considera que el rerefons del treball realitzat continua estant alineat amb aquests objectius. Conservant l'objectiu de la identificació d'aquells factors que causen els processos de gentrificació a la ciutat de Barcelona, es desitja que la feina realitzada porti informació addicional que pugui ser considerada a l'hora de reduir els efectes perjudicials dels processos de gentrificació. Així, comportaria de manera general a una reducció de les desigualtats des de diferents punts de vista, la implementació de mesures per crear comunitats sostenibles, així com la distribució de feina, infraestructures i creixement econòmic de manera més homogènia i equitativa.

Finalment, en quants als temes per a un treball futur potencial, es destaca novament la implementació de tècniques de predicció dels nivells de gentrificació als diferents barris de la ciutat de Barcelona, així com un anàlisi més profund de les diferents variables per a cada barri

de manera individual. Alhora, es considera important tenir en compte altres variables que es puguin incorporar a l'estudi, així com descartar algunes de les utilitzades que no han presentat una significància molt elevada als resultats obtinguts. A més, es destaca la reproductibilitat dels diferents mètodes aplicats d'aquest treball en altres llocs d'estudi, adaptant-se a la disponibilitat de dades i a les particularitats de cada indret, amb l'objectiu novament de trobar elements en comú o diferències segons la zona d'estudi que permetin un anàlisi més acurat dels processos de gentrificació.

Bibliografia

- Abarca Álvarez, F. J., Campos Sánchez, F. S., and Reinoso Bellido, R. (2018). Señales de gentrificación a través de la Inteligencia Artificial: identificación mediante el censo de vivienda.
- Aena (2024). Los aeropuertos de Aena en España cierran 2023 con más de 283 millones de pasajeros. <https://www.aena.es/es/prensa/los-aeropuertos-de-aena-en-espana-cierran-2023-con-mas-de-283-millones---de-pasajeros.html> [Data de consulta: 2 de març de 2024].
- Ajuntament de Barcelona (n.d.). Mapes imprimibles de Barcelona. <https://ajuntament.barcelona.cat/mapes-imprimibles/ca/> [Data de consulta: 16 de març de 2024].
- Antón-Alonso, F., Sergio, P. L., and Irene, C. (2018). Factors contextuais associats als processos de gentrificació de l'Àrea Metropolitana de Barcelona. *Papers: Regió Metropolitana de Barcelona: Territori, estratègies, planejament*, (60):64–79.
- Arias, A. (2018). Turisme i gentrificació: apunts des de Barcelona. *Papers: Regió Metropolitana de Barcelona: Territori, estratègies, planejament*, (60):130–139.
- Atkinson, R. (2000). Measuring Gentrification and Displacement in Greater London. *Urban Studies*, 37(1):149–165.
- Atkinson, R. and Wulff, M. (2009). Gentrification and displacement: a review of approaches and findings in the literature. *AHURI Positioning Paper*, 115:1 – 29.
- Atkinson, R., Wulff, M., Reynolds, M., and Spinney, A. (2011). Gentrification and displacement: The household impacts of neighbourhood change. *AHURI Final Report*, pages 1–89.
- Casamitjana, L. (2023). La gentrificació verde (o como se pasó de los ultramarinos al 'brunch'). <https://ethic.es/2023/06/la-gentrificacion-verde-o-como-se-paso-de-los-ultramarinos-al-brunch/> [Data de consulta: 10 de març de 2024].
- Cocola-Gant, A. (2015). Tourism and commercial gentrification.

- Cocola-Gant, A. (2018). *Tourism gentrification*, pages 281–293.
- Ding, L., Hwang, J., and Divringi, E. (2016). Gentrification and Residential Mobility in Philadelphia. *Regional science and urban economics*, 61:38–51.
- Díaz Parra, I. and Apaolaza, R. (2020). Una propuesta metodológica para identificar gentrificación a partir de los censos de población. *Estudios Demográficos y Urbanos*, 35(3):629–661.
- Freeman, L. (2005). Displacement or Succession?: Residential Mobility in Gentrifying Neighborhoods. *Urban Affairs Review*, 40(4):463–491.
- Fresnillo, I. (2018). La transformació del comerç de proximitat als barris. *Papers: Regió Metropolitana de Barcelona: Territori, estratègies, planejament*, (60):140–150.
- Glass, R. (1964). *London: Aspects of Change*. Centre for Urban Studies report. MacGibbon & Kee.
- López-Gay, A. (2008). *Canvis residencials i moviments migratoris en la renovació poblacional de Barcelona*. PhD thesis, Universitat Autònoma de Barcelona. Departament de Geografia.
- López-Gay, A. (2018). Cambio en la composición social y gentrificación en Barcelona: una mirada a través de los flujos migratorios y residenciales. *Papers: Regió Metropolitana de Barcelona: Territori, estratègies, planejament*, (60):80–93.
- López-Gay, A., Sales i Favà, J., Solana Solana, M., Fernández, A., and Peralta, A. (2019). *Midiendo los procesos de gentrificación en Barcelona y Madrid: una propuesta metodológica*. Centre de Política de Sol i Valoracions, CPSV / Universitat Politècnica de Catalunya, UPC.
- López Palomeque, F. (2015). Barcelona, de ciutat amb turisme a ciutat turística. Notes sobre un procés complex i inacabat. *Documents d'Anàlisi Geogràfica*, 61(3):483–506.
- Maluquer de Motes, J. (2019). La Revolució Industrial a Catalunya (1832-1935). *Butlletí de la Societat Catalana d'Estudis Històrics*, 30(3):125–179.
- Martin, I. W. and Beck, K. (2018). Gentrification, Property Tax Limitation, and Displacement. *Urban Affairs Review*, 54(1):33–73.
- Merrifield, A. (2013). The Urban Question under Planetary Urbanization. *International Journal of Urban and Regional Research*, 37(3):909–922.
- Moos, M., Revington, N., Wilkin, T., and Andrey, J. (2019). The knowledge economy city: Gentrification, studentification and youthification, and their connections to universities. *Urban Studies*, 56(6):1075–1092.

- Naciones Unidas (n.d.). Objetivos y metas de desarrollo sostenible - Desarrollo Sostenible. <https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/> [Data de consulta: 25 de març de 2024].
- Observatori Metropolità de l'Habitatge de Barcelona (2023). *L'oferta de lloguer de temporada a Barcelona. Informe preliminar.*
- Observatori Metropolità de l'Habitatge de Barcelona (2024). *L'oferta i la demanda d'habitatges de lloguer. Dades de portals immobiliaris. Tercer trimestre de 2023.*
- Posit team (2023). *RStudio: Integrated Development Environment for R.* Posit Software, PBC, Boston, MA.
- R Core Team (2023). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

