



UNIVERSITAT OBERTA DE CATALUNYA (UOC)

MÁSTER UNIVERSITARIO EN CIENCIA DE DATOS (*Data Science*)

TRABAJO FINAL DE MÁSTER

ÁREA: DATA ANALYSIS Y BIG DATA

Estudio de los factores de gentrificación de los barrios de las grandes ciudades

Autor: Alberto Giménez Aragón

Tutor: Anna Muñoz Bollas

Profesor: Albert Solé Ribalta

Barcelona, 2 de enero de 2022

Esta obra está sujeta a una licencia de Reconocimiento - NoComercial - SinObraDerivada
[3.0 España de CreativeCommons](#).



FICHA DEL TRABAJO FINAL

Título del trabajo:	Estudio de los factores de gentrificación de los barrios de las grandes ciudades
Nombre del autor:	Alberto Giménez Aragón
Nombre del colaborador/a docente:	Anna Muñoz Bollas
Nombre del PRA:	Albert Solé Ribalta
Fecha de entrega (mm/aaaa):	01/2022
Titulación o programa:	Máster Universitario en Ciencia de Datos
Área del Trabajo Final:	Data Analysis y Big Data
Idioma del trabajo:	Español
Palabras clave	Gentrificación, Geoespacial, Visualización

Abstract

Gentrification is a socio-economic phenomenon inherent of large cities where certain lower-class neighborhoods become fashionable due to a set of factors, which are urbanistically renewed and attract new businesses. These facts cause an increase in prices, making citizens with higher purchasing power move to the neighborhood, causing a displacement of the original citizens and traditional businesses. Currently, despite the fact that some studies have been carried out on gentrification, all the factors inherent to this phenomenon are not known with certainty, so it is not easy to predict the neighborhoods that will undergo this transformation in the short term. This study proposes the elaboration of an interactive map of the neighborhoods of Barcelona that allows to visualize different social and economic factors throughout the years to discover the factors most related to gentrification. In a second part, we will try to classify neighborhoods according to their level of gentrification using unsupervised algorithms in order to predict which neighborhoods are more likely to be gentrified in the short term.

Key words: Gentrification, Data Science, Geospatial data, Visualization, Map, GIS, Machine Learning, Clustering, Society, Barcelona, Master's Final Project

Resumen

La gentrificación es un fenómeno socio-económico propio de las grandes ciudades en donde ciertos barrios de clase baja se ponen de moda gracias a un conjunto de factores, se renuevan urbanísticamente y atraen nuevos negocios, provocando un aumento de precios y que ciudadanos con un poder adquisitivo superior se muevan al barrio, causando un desplazamiento de los ciudadanos originales y negocios tradicionales. Actualmente, pese a que se han realizado algunos estudios sobre la gentrificación, no se sabe con certeza todos los factores propios de este fenómeno, por lo que no es sencillo predecir los barrios que sufrirán esta transformación a corto plazo. Este estudio propone la elaboración de un mapa interactivo de los barrios de Barcelona que permita visualizar distintos factores sociales y económicos a lo largo de los años para descubrir los factores más relacionados con la gentrificación. En una segunda parte, intentaremos clasificar los barrios según su nivel de gentrificación mediante algoritmos no supervisados con el objetivo de predecir qué barrios son más propensos a ser gentrificados a corto plazo.

Palabras clave: Gentrificación, Ciencia de Datos, Datos geoespaciales, Visualización, Mapa, GIS, Machine Learning, Clustering, Sociedad, Barcelona, Trabajo de Final de Máster

Índice general

Abstract	III
Resumen	IV
Índice	IV
Llistado de Figuras	VII
Listado de Tablas	XII
1. Introducción	2
1.1. Contexto y justificación	2
1.2. Motivación	5
1.3. Objetivos	5
1.3.1. Objetivos generales	6
1.3.2. Objetivos específicos básicos	6
1.4. Enfoque y metodología	7
1.5. Planificación del trabajo	9
1.6. Breve sumario de productos obtenidos	14
1.7. Breve descripción de los otros capítulos de la memoria	15
2. Estado del arte	16

2.1. Estudio sobre los factores sociales y económicos relacionados con la gentrificación	16
2.2. Predicción de riesgo de gentrificación mediante modelos de <i>Machine Learning</i>	19
2.3. Planteamiento a partir del estado del arte	24
2.4. Decisiones tecnológicas	25
3. Diseño e implementación del trabajo	26
3.1. Extracción, limpieza y análisis	26
3.1.1. Lugares y ubicaciones de interés de los barrios (por año)	26
3.1.2. Lugares y ubicaciones de interés de los barrios (año 2019)	31
3.1.3. Indicadores de gentrificación de los barrios	34
3.2. Generación de las visualizaciones en Carto	41
3.2.1. Visualización de ubicaciones relevantes	42
3.2.2. Visualización de indicadores relevantes	45
3.3. Análisis y resultados del número de ubicaciones relevantes y evolución temporal de indicadores	46
3.3.1. Análisis de ubicaciones por barrio	46
3.3.2. Análisis de la evolución de los indicadores por barrio	50
3.3.3. Conclusiones del análisis y barrios gentrificados	59
3.4. <i>k-means</i> para clasificar el grado de gentrificación	63
3.5. Análisis de los resultados del <i>k-means</i>	65
4. Conclusiones	70
A. Gráficos y mapas	73
A.1. Gráficos de radar de ubicaciones	74
A.2. Mapas de coropletas de número de ubicaciones de cada tipo	78
A.3. Evolución de indicadores de 2015 a 2018	81

A.3.1. Renta	81
A.3.2. Precio de compra-venta por metro cuadrado	83
A.3.3. Precio del alquiler al mes por metro cuadrado	85
A.3.4. Número de personas por domicilio	87
A.3.5. Número de incidentes	89
A.3.6. Tasa de inmigración	91
A.3.7. Tasa de natalidad	93
A.4. Mapa de barrios de Barcelona	95
B. Código	96
Bibliografía	97

Índice de figuras

1.1. Ilustración del proceso de gentrificación por Grayson Perry. Una antigua industria se transforma en estudios, luego en un centro de creatividad y en café antes de convertirse en apartamentos lujosos. Fuente: [5]	4
1.2. Gráfico de la metodología en cascada. Fuente: [11]	8
1.3. Diagrama de Gantt. Fuente: Elaboración propia	13
2.1. Ejemplo de como funciona el financiamiento de desplazamiento. Fuente: [14]	18
2.2. Causas de la gentrificación. Fuente: [7]	20
2.3. Diagrama radial con el número de locales de cada tipo en los barrios de Madrid. Fuente: [7]	21
2.4. Resultados del <i>clustering</i> de los barrios de Nueva York. Fuente: [7]	22
2.5. Tipologías de transformaciones de los barrios en el condado de Davidson (Carolina del Norte). Fuente: [19]	23
2.6. Clustering de transformación de los barrios del condado de Davidson ($K = 4$). Fuente: [19]	24
3.1. Diagrama del proceso de extracción, limpieza y transformación de los lugares de interés.	27
3.2. Muestra del <i>dataset</i> de lugares.	31
3.3. Diagrama del proceso de extracción, limpieza y transformación de los lugares de interés del año 2019.	32
3.4. Muestra del <i>dataset</i> completo de lugares de 2019.	34

3.5. Diagrama del proceso de extracción, limpieza y transformación de los indicadores de cada barrio potencialmente relacionados con la gentrificación.	35
3.6. Muestra del <i>dataset</i> de indicadores.	40
3.7. Datasets en Carto en los que se basarán las visualizaciones. Fuente: elaboración propia.	41
3.8. Visualización de los lugares relevantes anualmente por barrio con filtro por categoría gastronomía y año 2018. Fuente: elaboración propia. Accesible en Carto	42
3.9. Visualización completa de los lugares relevantes por barrio en 2019 con el filtro de restaurantes. Fuente: elaboración propia. Accesible en Carto	44
3.10. Visualización de los indicadores relevantes por barrio, filtrando por renta del año 2015. Fuente: elaboración propia. Accesible en Carto	45
3.11. Diagramas de radar con los barrios que parecen estar más gentrificados según el número de ubicaciones de distintos tipos del año 2019. Fuente: Elaboración propia.	48
3.12. Diagramas de radar con los barrios que parecen estar en una fase inicial de gentrificación según el número de ubicaciones de distintos tipos del año 2019. Fuente: Elaboración propia.	49
3.13. Diagramas de radar con los barrios que parecen estar en situación estable según el número de ubicaciones de distintos tipos del año 2019. Fuente: Elaboración propia.	49
3.14. Diagramas de radar con los barrios que no tienen ningún riesgo de ser gentrificados según el número de ubicaciones de distintos tipos del año 2019. Fuente: Elaboración propia.	50
3.15. Gráficos de líneas de los barrios con más evolución de la renta. Fuente: Elaboración propia.	51
3.16. Gráficos de líneas de los barrios con más crecimiento del precio de compra-venta de viviendas de 2015 a 2018. Fuente: Elaboración propia.	52
3.17. Gráfico de líneas de la evolución del precio del alquiler del barrio de Baró de Viver de 2015 a 2018. Fuente: Elaboración propia.	53

3.18. Mapa de coropletas con el número de personas por domicilio en 2018. Fuente: Elaboración propia.	54
3.19. Gráficos de líneas de la evolución del número de incidentes de los barrios sospechosos de estar gentrificados según análisis previos. Fuente: Elaboración propia.	56
3.20. Gráficos de líneas con los barrios con mayor evolución en la tasa de inmigra- ción. Fuente: Elaboración propia.	58
3.21. Gráficos de líneas de la evolución de la tasa de natalidad de los barrios sos- pechosos de estar gentrificados según análisis previos. Fuente: Elaboración propia.	59
3.22. Mapa con los barrios gentrificados según el grado de gentrificación a partir de los análisis de ubicaciones y el resto de indicadores. De color amarillo se muestran los barrios en una fase inicial de gentrificación (el Raval y l'Antiga Esquerra de l'Eixample), en naranja los que están en una fase intermedia de gentrificación (la Dreta de l'Eixample y el Poble Sec) y en rojo los que están en una fase avanzada (Sant Pere y el Barri Gòtic - muy gentrificado) Fuente: Elaboración propia.	62
3.23. Selección de la k óptima con el análisis de la silueta y método del codo. Fuente: Elaboración propia.	64
3.24. Resultado del <i>clustering</i> utilizando <i>k-means</i> con $k = 3$ (Carto) y $k = 4$ (Carto). .	65
3.25. Evolución temporal media de las variables por cluster para $k = 3$. Versión interactiva en Flourish . Fuente: Elaboración propia.	66
3.26. Evolución temporal media de las variables por cluster para $k = 4$. Versión interactiva en Flourish . Fuente: Elaboración propia.	69
A.1. Diagramas de radar que muestran la distribución de las distintas categorías de negocios o ubicaciones por barrio. La línea marrón de cada gráfico muestra la media de Barcelona, mientras que la de color distinto muestra la propia del barrio. Fuente: Elaboración propia.	77
A.2. Mapa de coropletas para el número de alojamientos, lugares de culto, instal- ciones deportivas y restaurantes. Fuente: Elaboración propia.	78

A.3. Mapa de coropletas para el número de ubicaciones de ocio y cultura, tiendas de ropa y centros comerciales, centros sanitarios y farmacias y centros educativos. Fuente: Elaboración propia.	79
A.4. Mapa de coropletas para el número de ubicaciones de peluquerías y centros de estética, negocios tradicionales y alimentación y supermercados. Fuente: Elaboración propia.	80
A.5. Evolución de la renta de los años 2015 a 2018. Fuente: Elaboración propia. .	81
A.6. Evolución de la renta de los años 2015 a 2018. Fuente: Elaboración propia. .	82
A.7. Evolución de los precios de compra-venta por metro cuadrado de los años 2015 a 2018. Fuente: Elaboración propia.	83
A.8. Evolución de los precios de compra-venta por metro cuadrado de los años 2015 a 2018. Fuente: Elaboración propia.	84
A.9. Evolución de los precios mensuales de alquiler por metro cuadrado de los años 2015 a 2018. Fuente: Elaboración propia.	85
A.10. Evolución de los precios mensuales de alquiler por metro cuadrado de los años 2015 a 2018. Fuente: Elaboración propia.	86
A.11. Evolución del número de personas por domicilio de los años 2015 a 2018. Fuente: Elaboración propia.	87
A.12. Evolución del número de personas por domicilio de los años 2015 a 2018. Fuente: Elaboración propia.	88
A.13. Evolución del número de incidentes de los años 2015 a 2018. Fuente: Elaboración propia.	89
A.14. Evolución del número de incidentes de los años 2015 a 2018. Fuente: Elaboración propia.	90
A.15. Evolución de la tasa de inmigración de los años 2015 a 2018. Fuente: Elaboración propia.	91
A.16. Evolución de la tasa de inmigración de los años 2015 a 2018. Fuente: Elaboración propia.	92
A.17. Evolución de la tasa de natalidad de los años 2015 a 2018. Fuente: Elaboración propia.	93

A.18.Evolución de la tasa de natalidad de los años 2015 a 2018. Fuente: Elaboración propia.	94
A.19.Mapa de los barrios de Barcelona. Fuente: [21]	95
B.1. Captura del repositorio en Github. Fuente: Elaboración propia	96

Índice de cuadros

3.1. Ejemplo de un restaurante del dataset original con una selección de columnas.	29
3.2. Ejemplo de la transformación del mismo restaurante en el que se obtiene una fila por cada año del intervalo definido por los años de las columnas originales <code>created</code> y <code>modified</code> .	30
3.3. Barrios con más crecimiento del precio de compra-venta de viviendas de 2015 a 2018.	52
3.4. Barrios con más crecimiento del precio del alquiler de viviendas de 2015 a 2018.	53
3.5. Barrios con más evolución del número de incidentes de 2015 a 2018.	55
3.6. Barrios con más evolución de tasa de inmigración de 2015 a 2018.	57

Capítulo 1

Introducción

1.1. Contexto y justificación

Aquellas personas que vivimos en grandes ciudades seguro que somos conscientes de las transformaciones que viven los barrios por los que transitamos cada día a lo largo de los años. Dependiendo de la zona, puede que incluso nos llame la atención ciertos contrastes que nos pueden parecer algo extraños a primera vista. Por ejemplo, pequeñas tiendas de alimentación que conviven con restaurantes de moda, ferreterías, bazares y locutorios con hoteles, gimnasios, joyerías y tiendas de ropa con renombre. Además, se impulsa la construcción de nuevas viviendas, se llevan a cabo reformas para mejorar el barrio, como la peatonalización de calles, construcción de parques y jardines, etc.

Al final, lo que percibimos es que los negocios tradicionales del barrio se van sustituyendo poco a poco por locales de moda y que las mejoras urbanísticas hacen que la calidad del barrio mejore, atrayendo a gente distinta de la que solía acoger el barrio.

Todas estas mejoras hacen que los precios de las viviendas y servicios de la zona suban, provocando que las familias de toda la vida no puedan hacer frente a esta subida de precios y se vean obligadas a desplazarse poco a poco a otro barrio, mientras que otras personas con un mayor poder adquisitivo y un perfil distinto lleguen a este barrio transformado que ahora les parece atractivo. [1]

Esta situación que se acaba de explicar es realmente el significado del concepto gentrificación, palabra que proviene del inglés *gentry*, que se puede traducir como "alta nobleza". Según el NCRC (National Community Reinvestment Coalition) de Estados Unidos [2], se define la gentrificación como el proceso mediante el cual vecindarios de ingresos bajos reci-

ben niveles masivos de nuevas inversiones, agregando comodidades, aumentando el valor de las viviendas y atrayendo nuevos residentes con un poder adquisitivo mayor, hecho que puede provocar un desplazamiento cultural de ciertos miembros de la población que han vivido desde siempre en el barrio.

Pese a que el concepto de gentrificación parece surgido recientemente, la realidad es que los historiadores afirman que este concepto nace en el siglo III d.C. en la antigua Roma, donde grandes villas empezaron a sustituir a pequeños comercios. Sin embargo, la palabra "gentrificación" fue utilizada por primera vez con el sentido que le damos hoy en día por la socióloga Ruth Glass en 1964 para describir la afluencia de personas de clase media que desplazan a los trabajadores de clase baja residentes en los barrios urbanos, como por ejemplo el distrito londinense de Islington [3].

Si estudiamos los casos de gentrificación en España, el origen se sitúa en Madrid, concretamente en Malasaña. Hace unos 50 años, esta zona estaba bastante degradada y apenas tenía presencia institucional. Sin embargo, fue de los primeros barrios donde se pudo observar claramente la gentrificación a finales de los años 90 y inicios de los 2000, convirtiéndose en una zona alternativa en la que se reúnen habitantes de diferentes nacionalidades y especialmente joven, con más del 20 % de la población entre los 25 y 35 años que se han visto atraídos por la modernización del barrio con sus plazas, librerías, restaurantes, vinotecas y barberías. Como consecuencia de esta transformación se produjo una clara subida de los precios. Por ejemplo, el precio de compra de vivienda de segunda mano en el 2001 era de 2317€ por metro cuadrado, mientras que en 2018 este valor ascendía hasta 5068€ [1], [4].

No obstante, Malasaña no es el único caso de gentrificación en España. Por ejemplo, el barrio Gòtic en Barcelona es ya un barrio gentrificado. Después de diversas reformas que ha experimentado este barrio y una gran rehabilitación urbanística, hoy en día el precio medio de los alquileres es de los más altos de Barcelona, con excepción de las zonas de clase alta de Pedralbes y Sant Gervasi. Para exemplificar este incremento del alquiler, en 2013 el precio medio se situaba alrededor de 715€, mientras que en 2018 subió hasta los 1025€, teniendo en cuenta también que el 76 % de la población eran solteros y además el 12 % eran jóvenes. Pero para realmente tener en cuenta la importancia de esta subida y los posibles problemas económicos de sus habitantes que podrían verse obligados a desplazarse a zonas más baratas, es importante mencionar que la renta media anual de sus ciudadanos era de 26.885€ en 2018, siendo mucho menor que la del barrio de Vila Olímpica, donde este valor rondaba los 58.000€ y sus alquileres eran solo un poco más caros que en el Gòtic.

Otro concepto bastante relacionado con la gentrificación es la turistificación, especialmente en aquellos barrios donde el turismo ocupa gran parte de la actividad, en los que se produ-

cen fenómenos como la museificación, privatización, apropiaciones de sus espacios cotidianos o aperturas de tiendas de souvenires. Es decir, el barrio se transforma para atraer aún más a los turistas. Aquí es donde surge la principal diferencia con la gentrificación. Mientras que en esta última se produce una sustitución de la población de clase baja por la de media o media-alta, en la turistificación no se produce ninguna sustitución de la población. Lo que realmente sucede en los lugares turistificados es una progresiva despoblación, como consecuencia de cambios urbanísticos enfocados a las necesidades de los turistas y no para obedecer las necesidades de consumo más exclusivas que son propias de la gentrificación [6].

No obstante, ya que la gentrificación es un concepto más complejo y a mi juicio, más interesante, este trabajo se centrará en el estudio de la gentrificación de Barcelona, ya que es la gran ciudad que suelo transitar frecuentemente y que tiene barrios que ya han sido gentrificados. Aunque la existencia de la gentrificación es clara una vez se ha producido, hoy en día no hay un consenso claro de cuáles son sus causas, hecho que abre una línea clara de estudio sobre el tema.

Por lo tanto, este proyecto se centrará en la visualización interactiva de diferentes datos que puedan ser causa de la gentrificación de los barrios de Barcelona, con la que se podrá evaluar su evolución en los últimos años y detectar cuáles son los factores que caracterizan barrios con un índice alto de gentrificación y cuáles están en riesgo de ser gentrificados a corto plazo.

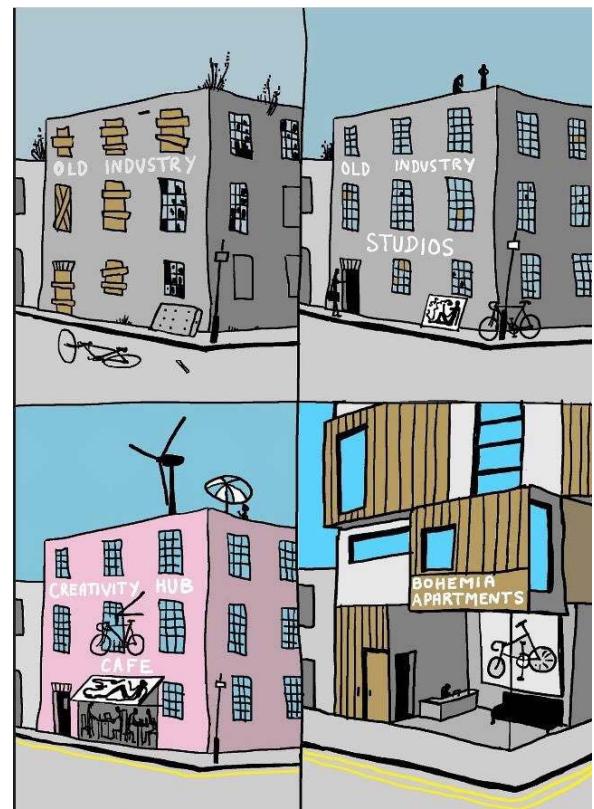


Figura 1.1: Ilustración del proceso de gentrificación por Grayson Perry. Una antigua industria se transforma en estudios, luego en un centro de creatividad y en café antes de convertirse en apartamentos lujosos. Fuente: [5]

1.2. Motivación

Pese a que actualmente vivo en una de las ciudades colindantes a la ciudad condal, en un futuro tendré que plantearme en qué zona querré residir y asentarme. Barcelona, debido a su proximidad, será uno de los lugares candidatos. Con el análisis de la gentrificación de los distintos barrios de la ciudad, y sabiendo mis necesidades y mis capacidades económicas, podré ser capaz de elegir la zona que se pueda adaptar mejor a estos factores. Imaginemos por ejemplo que me gusta la zona de la Barceloneta y que actualmente con mis condiciones económicas me lo podría permitir. Sin embargo, este barrio es uno de los que están en proceso de gentrificación, por lo que los precios de la zona aumentarán progresivamente y probablemente el coste de residir en la zona en unos pocos años estaría por encima de mis posibilidades, obligándome a desplazarme a otra zona con precios más asequibles. Este es un ejemplo concreto, pero podría aplicarse a cualquier barrio o incluso a cualquier ciudad, pese a que no haya tenido aún ningún caso claro de gentrificación, convirtiendo así los resultados de este estudio en una herramienta muy interesante para valorar diferentes ubicaciones a la hora de alquilar o comprar un piso teniendo en cuenta el poder adquisitivo de cada uno y predecir como podría evolucionar esa zona con los años dependiendo de si está o no en proceso de ser gentrificada.

Por otro lado, gracias a haber cursado en el máster de Ciencia de Datos la asignatura de *Análisis de datos geoespaciales*, he podido incrementar el interés que ya tenía previamente por este tipo de datos y podido descubrir un conjunto de herramientas SIG que me serán realmente útiles para llevar a cabo este proyecto y poder aplicarlas en el estudio de la gentrificación de los barrios de Barcelona.

1.3. Objetivos

Este proyecto tiene como objetivo el desarrollo de una herramienta de visualización interactiva de los distintos factores económicos y sociales a nivel temporal de los barrios de Barcelona que permita analizar los factores que caracterizan un alto grado de gentrificación y permita detectar los barrios en riesgo de ser gentrificados a corto plazo.

No obstante, para conseguir esta meta, es necesario definir unos objetivos más concretos, los cuales puedes ser divididos en dos grupos.

1.3.1. Objetivos generales

El primer grupo es el de los objetivos generales, que son los propios de un proyecto de final de máster y por tanto, con un objetivo más académico y de aprendizaje. Se determinan los siguientes:

1. Planificar y desarrollar un proyecto SIG
2. Investigar y analizar el estado del arte de los factores característicos de la gentrificación y su detección
3. Extraer e integrar datos de diferentes fuentes y tipología
4. Documentar el trabajo a través de una memoria
5. Realizar una presentación que resuma los principales puntos del proyecto
6. Defender el proyecto ante un tribunal

1.3.2. Objetivos específicos básicos

En este segundo bloque de objetivos, se encuentran aquellos más específicos del proyecto, por lo que son más técnicos y característicos del producto que se obtiene con el trabajo realizado.

1. Obtener, extraer e integrar las diferentes fuentes de datos (pueden ser o no geoespaciales) necesarias para realizar el proyecto
2. Realizar un análisis estadístico de los diferentes datos recogidos
3. Visualizar los resultados de forma interactiva
4. Extraer conocimiento sobre los factores de la gentrificación a través de esta herramienta de visualización
5. Identificar los factores que caracterizan un alto índice de gentrificación
6. Implementar una inteligencia capaz de identificar los barrios gentrificados y aquellos que estén en riesgo de serlo a corto plazo

1.4. Enfoque y metodología

Para garantizar el éxito del proyecto, es importante definir la manera en la que se va a enfocar el proyecto y qué metodología se seguirá durante todo el desarrollo. En primer lugar, para partir con cierto conocimiento del área de investigación, se buscarán otros proyectos que hayan analizado el proceso de gentrificación en alguna gran ciudad, con el objetivo de conocer el estado del arte del área y poder aplicar técnicas similares o más complejas. Esta tarea es fundamental para el desarrollo del proyecto, puesto que sirve como punto de partida, consigue proporcionar un conocimiento inicial, y permite que enfocar y guiar este proyecto mucho mejor, evitando dedicar esfuerzos en métodos probados anteriormente que no han resultado satisfactorios para dedicarlo en otros con más probabilidad de éxito. Por tanto, será importante dedicar una parte notable del tiempo a investigar el estado del arte.

Para poder empezar el estudio, es fundamental disponer de datos. En este caso, se necesitarán datos tanto espaciales como estadísticos de la ciudad de Barcelona. Estos datos los podremos encontrar con cierta facilidad en páginas *open data*, como por ejemplo el servicio de datos abiertos del Ayuntamiento de Barcelona o la Iniciativa open data Barcelona. No obstante, con ayuda de la investigación realizada del estado del arte, podremos tener una idea más clara de qué datos necesitamos exactamente, ya que deberíamos tener en cuenta aquellos que tengan una relación con el proceso de gentrificación. Sin embargo, como ya se ha mencionado en la introducción de este proyecto, no se puede afirmar con certeza que ciertos factores puedan afectar o no en este proceso de gentrificación, por lo que debaremos ser precavidos y analizar determinadas variables hasta llegar a la conclusión de que no tienen relevancia.

Una vez se han obtenido los datos, es importante realizar un preprocesado de estos para garantizar que son usables. Por ejemplo, se realizará un conteo de valor nulos, se comprobarán los tipos y formatos de las variables, y se realizará un análisis estadístico simple para tener algo más de conocimiento sobre los datos.

Posteriormente, se procederá a la creación del mapa interactivo con Carto. En este, se deberán visualizar las distintas variables por cada barrio de Barcelona, pudiendo ser seleccionadas desde un *widget*. Además, es fundamental que el mapa permita ver la evolución de estas variables a lo largo de los años, consiguiendo así relacionar los distintos factores con el proceso de gentrificación. Esta funcionalidad podría ser conseguida con otro *widget* que permitiese seleccionar un año o una fecha concreta.

Cuando se haya finalizado el desarrollo de este mapa interactivo, se deberá explotarlo. Con

ello, conseguiremos identificar los factores que son propios de un barrio gentrificado, así como poder detectar los barrios que pueden sufrir este proceso a corto plazo o bien están ya en una fase inicial de gentrificación.

Después, una vez adquirido conocimiento sobre los factores en el proceso de gentrificación, se aplicarán diferentes técnicas de aprendizaje automático para poder clasificar los barrios según su nivel de gentrificación. Para ello, se probarán distintos tipos de algoritmos no supervisados, como por ejemplo *k-means*, que es el algoritmo utilizado en el estudio de gentrificación de Nueva York, explicado en un *webinar* por Luis Nadal, Senior Business Consultant en LUCA [7].

Hoy en día existen diferentes metodologías de trabajo a seguir. No hay una opción mejor que las demás, ya que la más adecuada depende de muchos factores, como el tamaño del equipo, los objetivos y otros factores [8], [9]. A pesar que las metodologías ágiles se están convirtiendo en la corriente principal en los proyectos empresariales [10], utilizaremos el método tradicional de metodología cascada (ver Figura 1.2), ya que pensamos que se adapta mejor a las características del proyecto, el cual cuenta con tareas bastante acotadas y con objetivos claros, además de ser desarrollado por una sola persona.

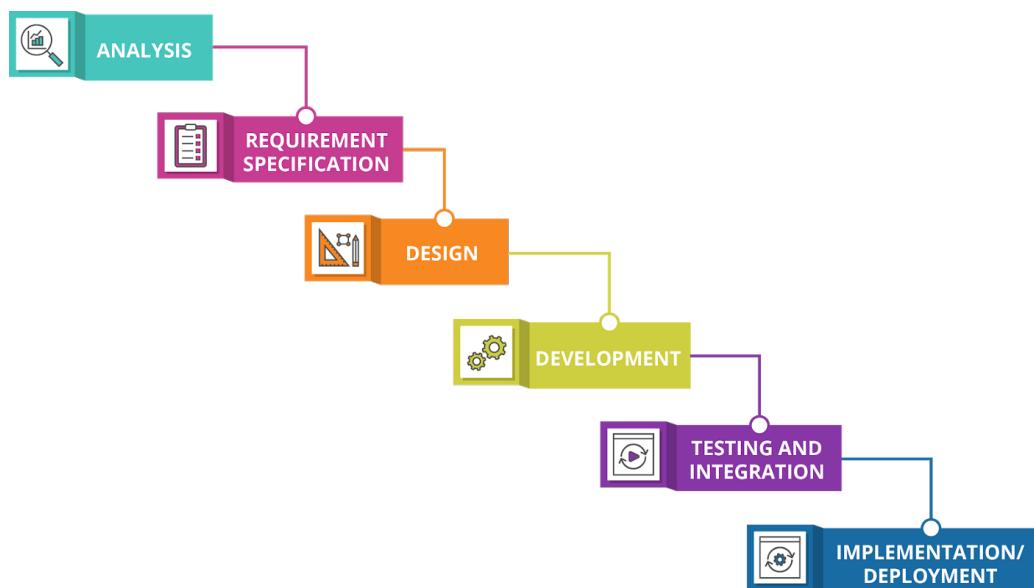


Figura 1.2: Gráfico de la metodología en cascada. Fuente: [11]

1.5. Planificación del trabajo

Para llevar a cabo el proyecto se necesitan distintos recursos, tanto se *software* como de *hardware*.

■ Software

- **PyCharm y Visual Studio Code:** IDE's de desarrollo de Python para análisis estadísticos, limpieza de datos y programación de modelos de *Machine Learning*.
- **Carto:** Herramienta SIG para el desarrollo de mapas interactivos.
- **Overleaf:** Editor de LaTeX basado en la nube.
- **Power Point:** Herramienta para realizar la presentación.
- **Gantt Project:** Aplicación para la programación y gestión de proyectos con diagrama de Gantt.
- **Google Chrome:** Navegador web.
- **Git:** Sistema de control de versiones.
- **Github:** Plataforma de desarrollo colaborativo para alojar proyectos utilizando el sistema de control de versiones Git.

■ Hardware

- PC de sobremesa con Intel Core i7-4770K @3.5Ghz, GPU NVIDIA GeForce RTX 3070 Ti, 16 GB RAM DDR3 y Windows 10.

La realización de este trabajo se divide en 6 fases diferentes, que se corresponden a las 6 PEC's en las que está organizado este curso, las cuales se explican a continuación.

1. **PEC1 - Definición y planificación del trabajo final.** Esta primera fase tiene como principal objetivo definir claramente cuál es la temática del trabajo, justificar su interés y/o relevancia, qué se quiere conseguir al finalizar el TFM, definir los objetivos principales y una planificación temporal del proyecto. Además, incluye la motivación personal del estudiante para hacer este trabajo final.
2. **PEC2 - Estado del arte o análisis de mercado del proyecto.** El objetivo consiste en explicar qué investigación (o soluciones) se han desarrollado hasta el momento para resolver la problemática en la que el estudiante trabajará. Esto es lo que se denomina "Estado del arte" o "Estado de la cuestión".

3. **PEC3 - Diseño e implementación del trabajo.** El objetivo, a grandes rasgos, es realizar las tareas propias para el desarrollo de un proyecto y/o investigación
4. **PEC4 - Redacción de la memoria.** Aunque el documento de la memoria se irá trabajando durante todo el curso, la PEC4 destina un tiempo específico a la redacción y configuración final de la memoria dada su importancia,
5. **PEC5 - Presentación y defensa del proyecto.** Se debe realizar una presentación con voz que resuma el proyecto realizado y que junto a la memoria se entrega a los miembros del tribunal para su evaluación.
6. **Defensa pública.** Finalmente, una vez presentadas y evaluadas todas la PEC sólo queda la evaluación del TFM por parte de un tribunal. La defensa consiste en una meet síncrono de unos 30 minutos durante los cuales los miembros del tribunal que evalúan tu trabajo te harán preguntas relativas al mismo. Normalmente son preguntas para que justifiques algún aspecto metodológico o para que lo aclares o lo completes.

A continuación, se definen las tareas asignadas a cada una de estas 6 fases. Junto a cada tarea se especifica su identificador, que nos será útil para hacer referencias concretas a tareas o para nombrar las distintas ramas de Git a la hora del desarrollo para saber rápidamente qué funcionalidad incluyen.

PEC1 - Definición y planificación del trabajo final

- [DEF-01] Contextualizar y justificar del trabajo.
- [DEF-02] Explicar la motivación del proyecto.
- [DEF-03] Definir de los objetivos del proyecto.
- [DEF-04] Enfocar el proyecto y determinar la metodología a seguir.
- [DEF-05] Identificar los recursos necesarios para la elaboración del proyecto.
- [DEF-06] Definir las principales fases del proyecto y las tareas necesarias de cada una de ellas.
- [DEF-07] Planificar temporalmente el desarrollo del trabajo.

PEC2 - Estado del arte o análisis de mercado del proyecto

- [ART-01] Buscar y explicar diferentes investigaciones desarrolladas hasta el momento que tratan o intentan resolver el mismo problema que se quiere solventar en este trabajo.
- [ART-02] Decidir si se quieren aplicar y evolucionar los mismos procedimientos efectuados por otras investigaciones o pensar e implementar una solución desde cero.
- [ART-03] Investigar herramientas que nos puedan ser útiles para el desarrollo, como fuentes de datos abiertos, librerías de Python para datos geoespaciales, ejemplos de código de aplicados en problemas similares, etc.

PEC3 - Diseño e implementación del trabajo

- [IMP-01] Definir qué datos de los barrios de Barcelona se necesitan obtener.
- [IMP-02] Buscar y seleccionar las fuentes de datos abiertos de Barcelona necesarias.
- [IMP-03] Preprocesar los datos (tratamiento de nulos, tipo y formato de columnas, filtros a aplicar, etc.).
- [IMP-04] Realizar un análisis estadístico para conocer los datos.
- [IMP-05] Validar el acceso a la herramienta GIS (Carto).
- [IMP-06] Cargar los datos en Carto (shapefiles y los propios datos)
- [IMP-07] Implementar las *queries* para mostrar las distintas variables sobre el mapa por barrios.
- [IMP-08] Añadir elementos de interacción en el mapa para permitir al usuario interactuar con él (selección de variable a mostrar, línea temporal, etc.)
- [IMP-09] Investigar sobre cómo implementar la componente temporal de los datos, permitiendo al usuario visualizar las variables en momentos temporales distintos.
- [IMP-10] Explotar la visualización para extraer conocimiento de las variables y su relación con la gentrificación.
- [IMP-11] Diseñar un modelo no supervisado capaz de clasificar los barrios según su riesgo de gentrificación.

PEC4 - Redacción de la memoria

- [MEM-01] Redactar la memoria de forma transversal durante todo el desarrollo e implementación del proyecto.

PEC5 - Presentación y defensa del proyecto

- [PRE-01] Identificar los puntos claves del proyecto que se querrán focalizar en la presentación.
- [PRE-02] Elaborar la presentación con Power Point con los puntos identificados en el apartado anterior.
- [PRE-03] Preparar y practicar la presentación asegurando que los distintos mensajes se transmiten con claridad y son fácilmente entendibles.
- [PRE-04] Grabar el vídeo de la presentación.

Defensa pública

- [DPU-01] Preparar la defensa.
- [DPU-02] Defender el proyecto respondiendo las preguntas del tribunal.

A continuación, en la Figura 1.3 se muestra el diagrama de Gantt correspondiente a las tareas descritas previamente.

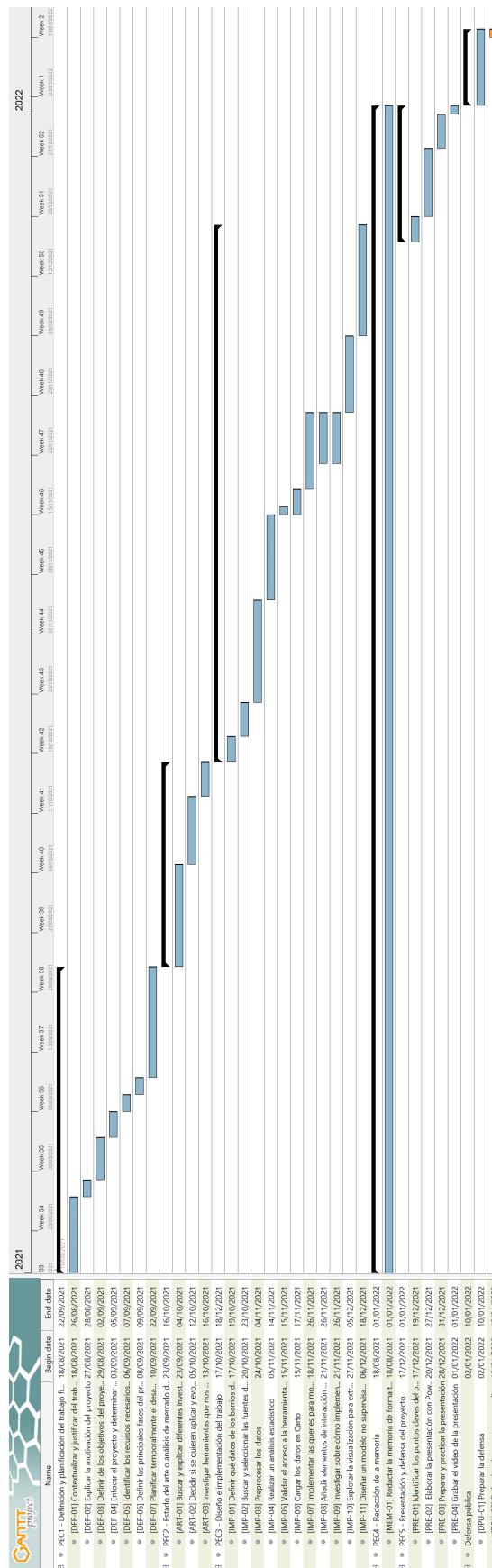


Figura 1.3: Diagrama de Gantt. Fuente: Elaboración propia

1.6. Breve sumario de productos obtenidos

Como resultado de este proyecto, han sido múltiples los productos que se han obtenido, los cuales se listan a continuación.

- Un conjunto de ficheros CSV y GEOJSON descargados de la web de Open Data BCN [12] y los correspondientes generados a partir de los scripts en Python en los que se transforman estos datos, dando lugar principalmente a *datasets* con lugares o negocios ubicados en cada barrio y la evolución anual de indicadores relevantes para la gentrificación.
- Un repositorio en [Github](#) con el código Python, así como los ficheros de configuración necesarios y una explicación de cómo ejecutarlo.
- Un conjunto de visualizaciones de mapas de coropletas de los barrios de Barcelona utilizando Carto:
 - [Ubicación](#) y evolución de lugares por barrio de los años 2015 a 2018 (datos parciales de todos los negocios):
 - [Ubicación](#) de lugares por barrio de 2019 (dataset completo):
 - [Evolución](#) de indicadores por barrio de los años 2015 a 2018:
 - [Resultado](#) de la ejecución de *k-means* con $k = 3$:
 - [Resultado](#) de la ejecución de *k-means* con $k = 4$:
- Un conjunto de visualizaciones realizadas con Flourish:
 - [Gráficos de líneas](#) con la evolución de los indicadores por barrio de 2015 a 2018.
 - [Gráficos de radar](#) con la distribución de ubicaciones por barrio.
 - [Gráficos de líneas](#) con la evolución de los indicadores por cluster de barrios de 2015 a 2018 obtenidos con $k = 3$.
 - [Gráficos de líneas](#) con la evolución de los indicadores por cluster de barrios de 2015 a 2018 obtenidos con $k = 4$.

1.7. Breve descripción de los otros capítulos de la memoria

La memoria del proyecto comienza explicando en primer lugar el estado del arte de los factores sociales y económicos relacionados con la gentrificación, así como de su predicción mediante modelos de *Machine Learning*.

En el siguiente capítulo de Diseño e implementación del trabajo se explica primeramente el proceso de extracción y transformación de los datos y cómo se generan las visualizaciones en Carto a partir de todos estos datos. Seguidamente, se analizan detalladamente todas las visualizaciones generadas para explicar la gentrificación de los barrios de Barcelona y descubrir el grado de gentrificación de cada uno de ellos. Por último, se implementa un modelo de *clustering (k-means)* para agrupar los barrios según su nivel de gentrificación y se analiza el resultado.

En el último capítulo se exponen las conclusiones del proyecto, tanto de su realización como de los resultados obtenidos, y se explica las líneas de trabajo futuro que no se han podido explorar.

Finalmente, en el apéndice de la memoria se pueden encontrar todos los gráficos que se han generado y analizado, tanto en Carto como en Flourish.

Capítulo 2

Estado del arte

Tal como se ha explicado en capítulos anteriores, este trabajo se centra primeramente en descubrir cuáles son los factores que están relacionados con el proceso de gentrificación (centrándonos en la ciudad de Barcelona a través de la visualización de distintas variables) y luego en implementar una inteligencia capaz de predecir los distintos grados de gentrificación en los barrios para detectar cuáles son los que están en riesgo de sufrir esta transformación a corto plazo. Por lo tanto, se van a seguir principalmente dos líneas de investigación:

- Estudio sobre los factores sociales y económicos relacionados con la gentrificación.
- Predicción de riesgo de gentrificación mediante modelos de *Machine Learning*

A continuación, se entra en detalle al estado del arte de cada uno de los puntos, que nos servirá como punto de partida y nos aportarán cierto conocimiento inicial que nos permitirá dirigir el proyecto de una forma más óptima y con más posibilidades de éxito.

2.1. Estudio sobre los factores sociales y económicos relacionados con la gentrificación

La primera parte del proyecto consiste en encontrar los factores sociales y económicos que están relacionados con la gentrificación de los barrios en la ciudad de Barcelona utilizando una visualización interactiva en la que el usuario podrá inspeccionar la evolución de las distintas variables y ver cuáles tienen más relación con este fenómeno.

2.1. Estudio sobre los factores sociales y económicos relacionados con la gentrificación 17

No obstante, estas variables que se probarán no pueden estar escogidas al azar. Deben ser variables que intuitivamente puedan tener relación. Sin embargo, no es una tarea trivial escogerlas, ya que hay infinitas variables posibles y solo unas pocas serán las que realmente sean útiles para detectar la gentrificación. Es por este motivo que se deberá realizar un ejercicio de investigación sobre qué factores se han considerado relevantes y cuales no en estudios previos, pudiendo de esta forma enfocar los esfuerzos de la recogida de datos en aquellos datos que aportarán más valor a partir de las conclusiones extraídas en los múltiples análisis anteriores.

En la página web de City Observatory [13], la cual está formada por grupo de expertos dedicado al análisis basado en datos de las ciudades y las políticas que les dan forma, Joe Cortright, economista y presidente de una empresa consultora en análisis económicos, nos explica una serie de factores que distintas personas con perfiles muy variados acusan de ser los causantes de la gentrificación. En un primer ejemplo, recupera el caso de un artista de Brooklyn que en un TED Talk afirma que el arte es una causa de la gentrificación. También se hace referencia a un artículo del Forbes [14] que culpa a los bancos como otro de los causantes de la gentrificación, especialmente debido a que prestan dinero a especuladores que compran multitudes de inmuebles a un precio mínimo que luego alquilan por un precio superior, haciendo que los inquilinos no puedan permitirse dicha subida. En la figura 2.1 se puede ver este proceso en más detalle.

También menciona el cambio climático como otro posible factor [15], aunque en este caso sería más propio de ciudades costeras, ya que el ejemplo que se describe, habla sobre como el crecimiento del nivel del mar pone en riesgo las viviendas a pie de playa en la zona de Florida, haciendo que los promotores inmobiliarios compren viviendas en las zonas más elevadas, las que originalmente eran las menos deseadas por su lejanía al mar.

Otra publicación que referencia [16], afirma que otro factor de la gentrificación es la disminución de la delincuencia en un barrio, ya provoca un aumento del precio de las viviendas al hacer el barrio más atractivo.

Seguidamente, también explica el tipo de perfil de gente que suele ser la gentrificadora. Es muy habitual prototiparla como solteros, aunque demográficamente, muchos de ellos tienen o están planificando tener su primer hijo.

Otro de los causantes que menciona es el espectáculo deportivo. Concretamente, explica el caso de cómo la construcción de un estadio de fútbol en Orlando ha desencadenado la revitalización del distrito de Parramore, el cual ha sido uno de los distritos más pobres en los últimos 50 años y con altas tasas de desempleo y criminalidad. [17]. Sin embargo, con la

Displacement Financing

An example of how Displacement Financing works:

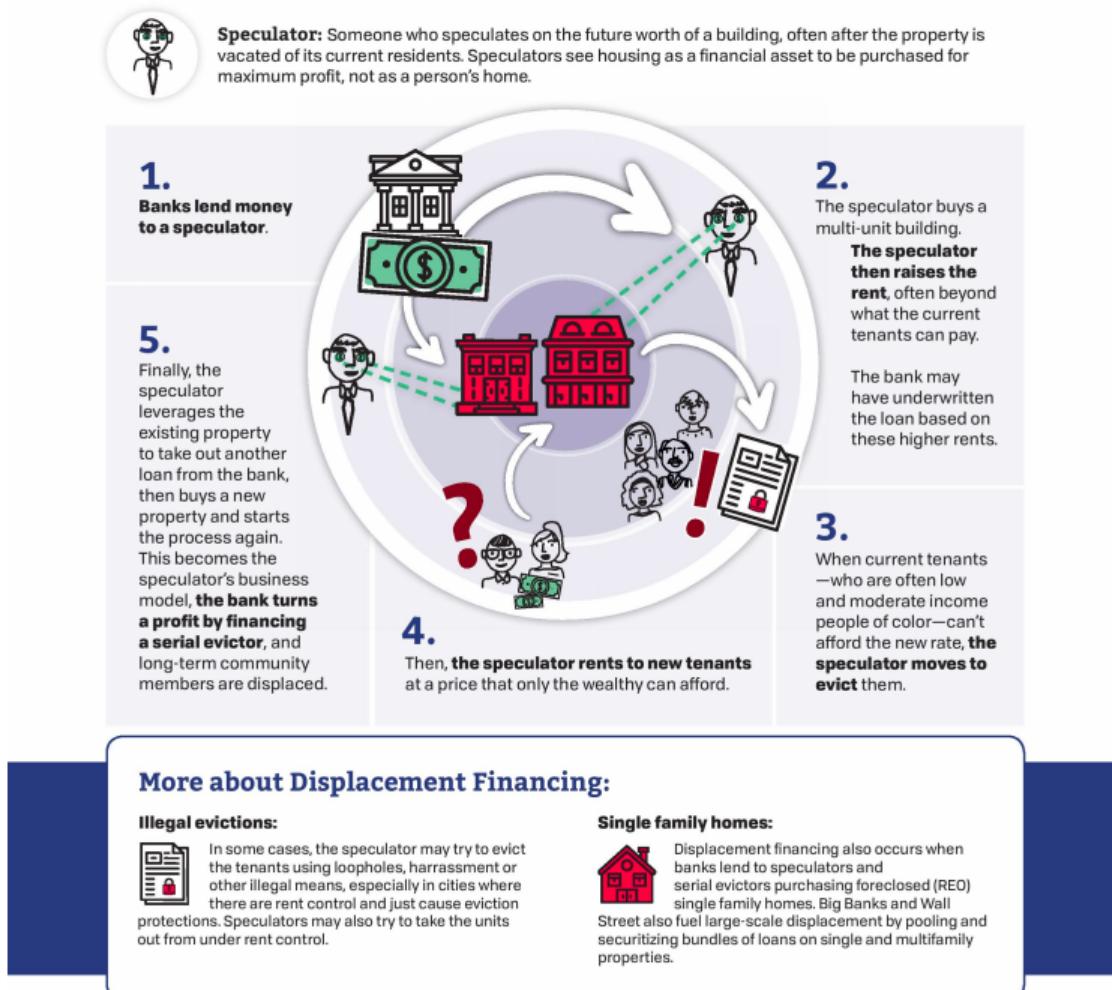


Figura 2.1: Ejemplo de como funciona el financiamiento de desplazamiento. Fuente: [14]

llegada de este estadio, se ha mejorado la movilidad de la zona, ha entrado dinero nuevo y se ha producido un aumento del valor de las propiedades y, lo que es más importante, el aumento de los alquileres.

A parte de estos factores que presenta Joe Cortright, hay muchas más investigaciones que analizan y argumentan más variables, puesto que se trata de un tema con el que los investigadores no se ponen completamente de acuerdo.

Por ejemplo, en un *paper* del departamento de investigación del Federal Reserve Bank de Philadelphia [18], Jackelyn Hwang y Jeffrey Lin explican los factores que según estudios recientes, son las causas de la gentrificación.

Uno de estos factores es el trabajo, ya que presentan estudios que afirman que los cambios en el acceso al trabajo han afectado a la gentrificación actual. Se muestra el ejemplo de cómo los trabajadores más cualificados se mueven a los grandes centros urbanos, siendo capaces de seguir y aceptar los cambios de precios de las viviendas, aunque realmente estos cambios se podrían ver más como una consecuencia de la gentrificación en lugar de una causa.

Otro de los factores que todavía no se había mencionado es el de los cambios en la composición racial y étnica de los barrios, hecho que puede también atraer gentrificadores. Cada vez más, según distintos estudios realizados, es más común encontrar barrios multiétnicos, ayudados de una mejora en las actitudes raciales de la población en general, siendo una llamada a la diversidad y autenticidad cultural.

Por último, volviendo al *webinar* LUCA Talk: Open Data para entender la transformación de los barrios [7] que se ha mencionado en el primer capítulo (y al que más tarde se volverá), aquí también se realiza una pequeña introducción sobre los factores causantes de la gentrificación. A alto nivel, habla sobre el precio de la vivienda, el cual sufre una evolución ascendente con el paso de los meses, el cambio en los ingresos de los habitantes, que también crece debido a los cambios de la población. Además, la actividad económica cambia de acuerdo a las necesidades de los nuevos habitantes del barrio. También se hace mención a otros factores que ya han sido tratados previamente, como la inversión pública que se realiza en los barrios para mejorarlo, las infraestructuras que se construyen con el mismo objetivo y los servicios que se ofrecen, todo para acabar atrayendo a una población con un mayor poder adquisitivo, la cual acabará provocando una migración de los habitantes que ya no puedan permitirse el lujo de vivir en su barrio gentrificado. En la Figura 2.2, se puede ver un diagrama con las causas de la gentrificación, extraído del *webinar*. Es importante remarcar el énfasis que hace el locutor en que muchos de estos motivos no se sabe con certeza si son causa o consecuencia de la gentrificación debido a las interrelaciones que existen entre todos estos factores.

2.2. Predicción de riesgo de gentrificación mediante modelos de *Machine Learning*

Una vez realizado el estudio de los factores relacionados con la gentrificación, se procederá a analizar investigaciones previas realizadas sobre la predicción del riesgo de gentrificación de los barrios mediante *Machine Learning*.

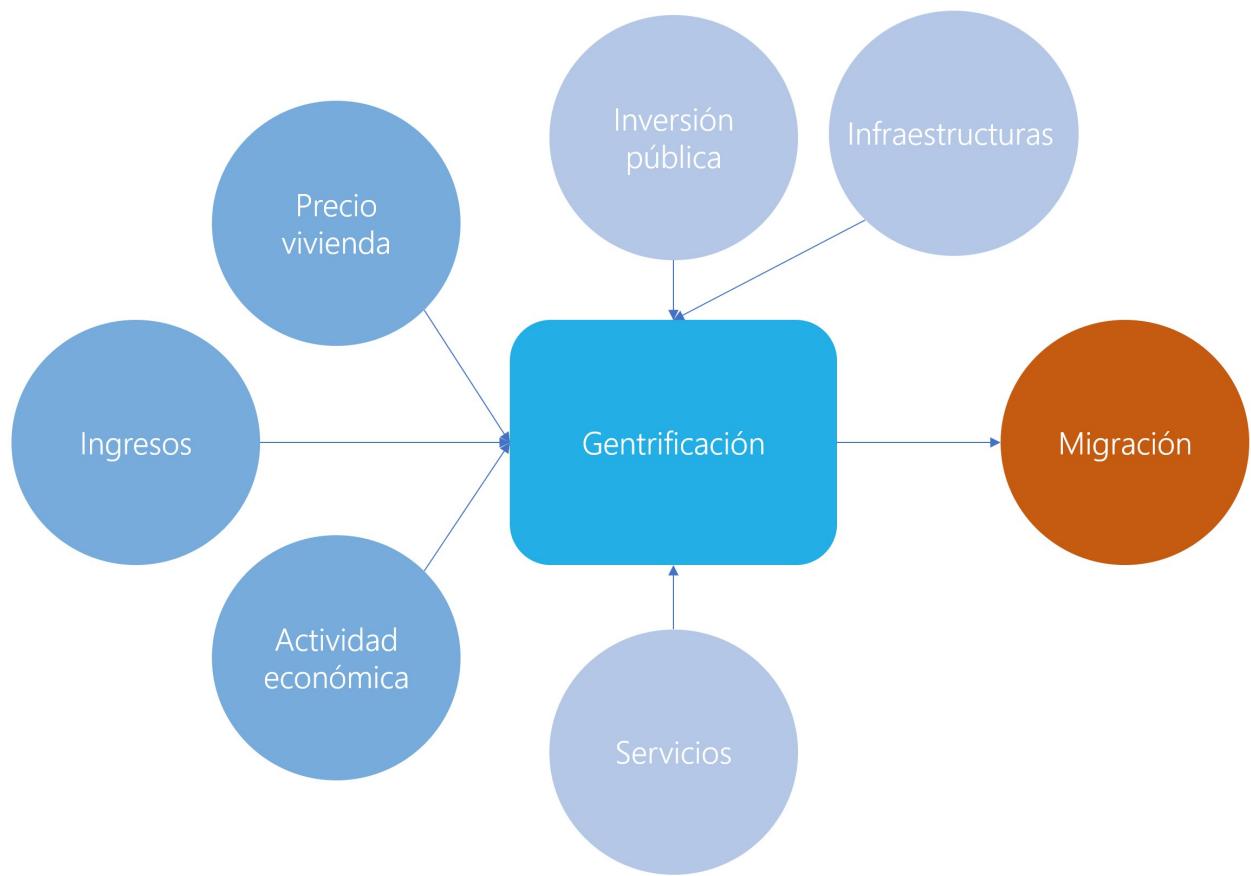


Figura 2.2: Causas de la gentrificación. Fuente: [7]

Primero de todo, se continuará analizando el mismo *webinar* del que se ha hablado anteriormente [7], que después de hablar de los factores de la gentrificación, introduce en primer lugar un análisis de la gentrificación en Madrid. Partiendo del hecho de que el tipo de servicios o locales que se hay en un barrio es un factor relevante para determinar si un barrio está gentrificado o no, comienzan mostrando distintas visualizaciones que enseñan el número de locales de cada tipología que hay en cada barrio madrileño. De una extensa lista inicial con 27 clases de locales, los acaban agrupando en 6 categorías, las cuales son las más útiles para detectar a simple vista si un barrio está gentrificado o no. Estas categorías son las siguientes:

- Minoristas de frutas, verduras, etc.
- Cafeterías
- Tiendas de ropa
- Peluquería y estética
- Restaurantes
- Educación

De entre estas categorías, las que serían más propias de un barrio gentrificado son las tien-

das de ropa, cafeterías y restaurantes. Si uno se fija en la Figura 2.3, la cual se ha extraído del *webinar*, para cada barrio se muestra la cantidad de locales de cada tipo en un gráfico radial. En cada gráfico está marcado con una área roja la forma que tendría un barrio con un alto grado de gentrificación (número elevado de restaurantes, cafeterías y tiendas de ropa). Por lo tanto, en aquellos barrios donde el polígono se parezca más a esta zona roja, más gentrificado estará. Por ejemplo, de los cuatro barrios que se muestran en la imagen, vemos que los que más gentrificados están son Malasaña y Sol, mientras que Canillejas y Puerta Bonita no lo están.

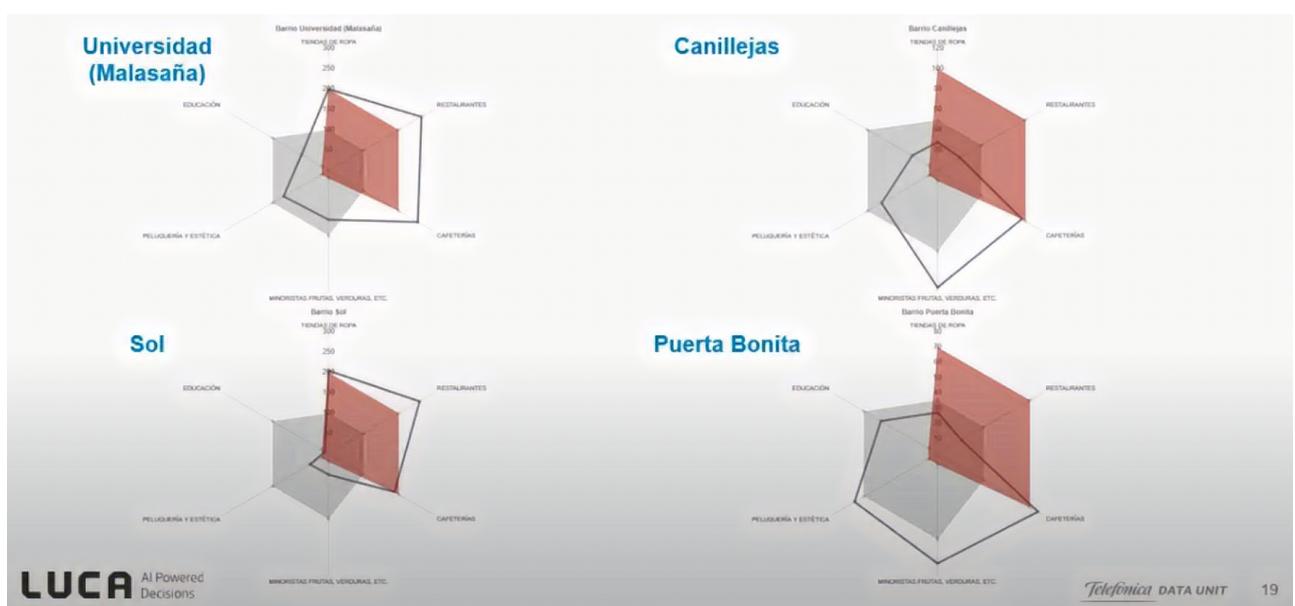


Figura 2.3: Diagrama radial con el número de locales de cada tipo en los barrios de Madrid. Fuente: [7]

Seguidamente, introducen al estudio las variables de renta media y precio del alquiler por metro cuadrado, llegando a la conclusión de que los barrios más sospechosos de estar gentrificados cumplen las siguientes condiciones.

1. Las cafeterías, restaurantes y tiendas de ropa suman más del 25 % de los locales del barrio.
2. La renta per cápita media del barrio está cerca de la media global de la ciudad (estando entre el 70 % y el 130 % de este valor).
3. El precio medio del alquiler por metro cuadrado se encuentra en el 50 % más caro.

Sin embargo, llevaron el estudio todavía más lejos, haciendo una colaboración con la universidad de Columbia para estudiar el caso de la gentrificación en los barrios de Nueva York.

No obstante, de cara a este estudio, se añadieron más variables a considerar, como cifras de crimen, de transporte, datos de salud, nivel de estudios, datos demográficos (sexo, edad y raza), migración entre distritos y nuevas aperturas de negocios, a parte de los ingresos y precios de las propiedades. El objetivo, a parte de realizar un primer análisis exploratorio, era construir un modelo de *Machine Learning* para detectar los barrios que se puedan estar gentrificando. Para ello, utilizan el algoritmo no supervisado *k-means*, usando las variables anteriores normalizadas y teniendo en cuenta su variación a lo largo de los últimos cinco años considerados, además de determinar el número de *clusters* en $k = 3$. Una vez obtuvieron las etiquetas de cada barrio, pudieron relacionar visualmente cada etiqueta con un nivel de gentrificación (bajo, medio y alto), ya que al tratarse de un algoritmo no supervisado, los resultados no están asignados a una clase concreta.

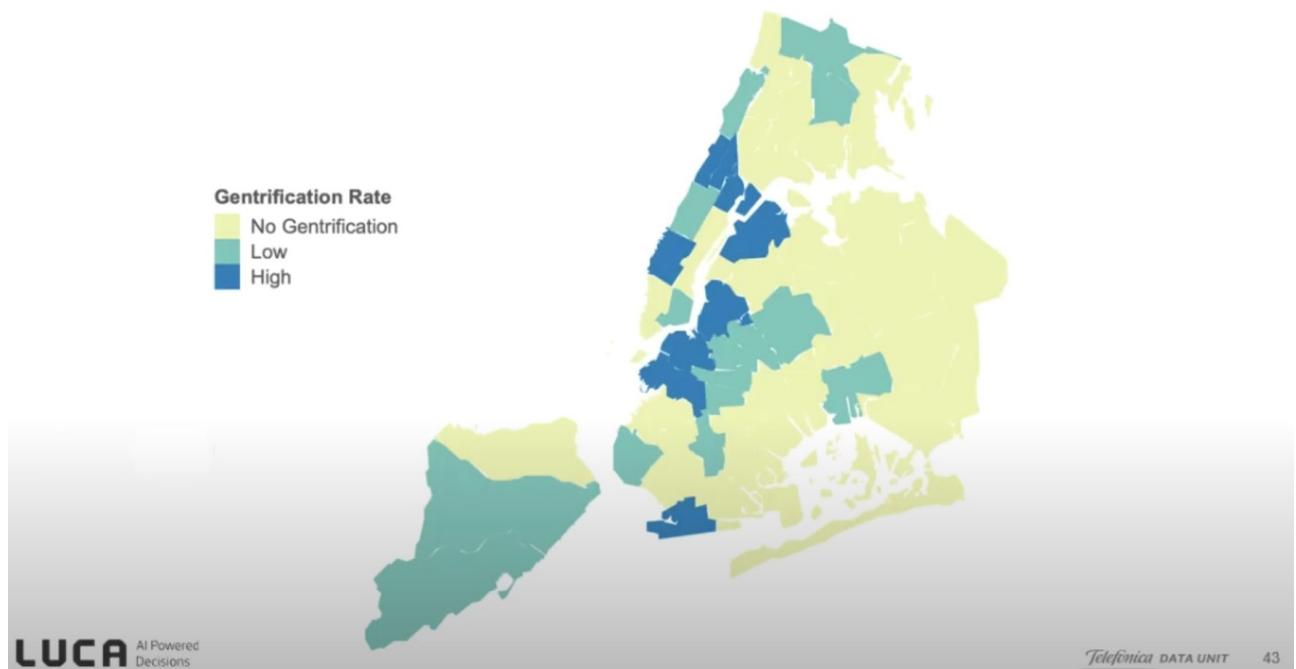


Figura 2.4: Resultados del *clustering* de los barrios de Nueva York. Fuente: [7]

Además, según el modelo, las variables que resultaron ser las más significativas fueron la renta, el nivel educativo y la raza de los habitantes. No obstante, como conclusión mencionan que es muy complicado generar un modelo que sea aplicable a cualquier ciudad del mundo, puesto que cada lugar tiene unas particularidades concretas que hacen que sea necesario individualizar el modelo para cada caso.

Otro trabajo muy interesante que intenta predecir la gentrificación tanto de forma no supervisada como supervisada es la tesis de David Knorr, *Using Machine Learning to Identify and Predict Gentrification in Nashville, Tennessee* [19]. Para el caso no supervisado, usaron de

nuevo el algoritmo *k-means* con las siguientes variables.

- Mediana del precio de compra-venta de las viviendas
- Mediana del precio del alquiler
- mediana de los ingresos en el hogar
- Porcentaje de personas mayores de 25 años con estudios universitarios
- Porcentaje de personas no blancas
- Porcentaje de viviendas multifamiliares

Sin embargo, en este caso no se escogieron directamente el número de clusters, sino que se tomó mediante el método del codo y el coeficiente de la silueta [20]. El método del codo indicaba una $k = 4$ como óptima, mientras que el coeficiente de la silueta más alto era 2. Sin embargo, como no querían una clasificación binaria, tomaron $k = 4$ como valor final, ya que además era el segundo coeficiente más alto con este último método.



Figura 2.5: Tipologías de transformaciones de los barrios en el condado de Davidson (Carolina del Norte). Fuente: [19]

Una vez obtenidos los grupos, pudieron etiquetarlos según sus características (ver figura 2.5 y su representación en el mapa en la figura 2.6), cuyas etiquetas serían asequible - cluster K1 (indicativos de desinversión y diversidad racial), gentrificado - cluster K2 (exhibe los patrones distinguibles de gentrificación y evidencia de desplazamiento), estable - cluster K3 (valores muy cercanos a la media en todos los aspectos), y mezclado - cluster K4 (altos aumentos en el alquiler, pero disminuciones generales en el valor de la vivienda).

Aprovechando los resultados del *clustering*, etiquetaron los barrios binariamente en gentrificados y no gentrificados para poder aplicar un modelo supervisado. Para ello, añadieron bastantes más variables relacionadas con la demografía, los hogares, los trabajos, el transporte y comodidades del barrio. Entrenando con un *Random Forest*, consiguieron identificar el 83 % de las zonas gentrificadas, a costa de obtener un 45 % de falsos positivos. De este modelo, la variable de distancia al centro resultó ser con diferencia la más importante, seguida de la pobreza, porcentaje de terreno sin edificar y desempleo.

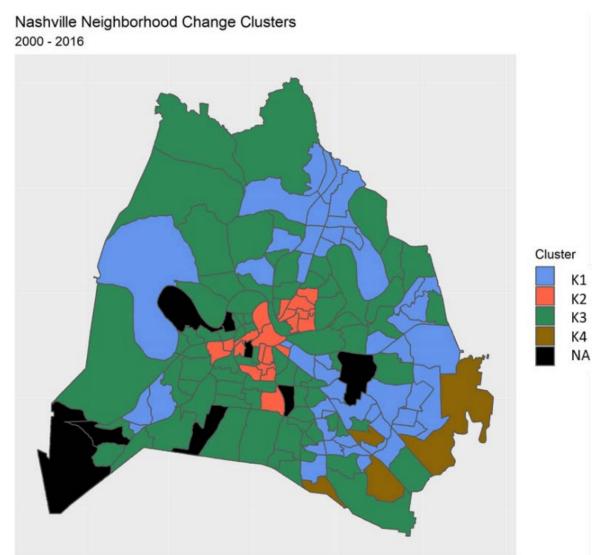


Figura 2.6: Clustering de transformación de los barrios del condado de Davidson ($K = 4$). Fuente: [19]

2.3. Planteamiento a partir del estado del arte

Después de haber investigado el estado del arte, se ha aprendido cuáles son los factores que han sido considerados como relevantes a la hora de detectar y predecir la gentrificación, por lo que se podrán centrar los esfuerzos en obtener los datos de la ciudad de Barcelona que realmente puedan ser útiles. También se ha visto que la parte de exploración y visualización de datos será muy importante, pues se podrán detectar las variables que realmente estén relacionadas con la gentrificación en el caso particular de la ciudad de Barcelona, ya que como se ha visto, los factores varían mucho dependiendo de la ciudad y por lo tanto no se pueden generalizar.

Posteriormente, con las variables que resulten tener más relación y teniendo en cuenta su evolución a lo largo de los años, se aplicará un algoritmo de aprendizaje no supervisado

(*k-means*) con el objetivo de agrupar los barrios en distintas clases que puedan identificar el grado de gentrificación de los barrios, para poder así detectar qué barrios están sufriendo esta transformación. Como se ha visto en los distintos análisis, al no tener los barrios etiquetados según si están gentrificados o no, la elección de escoger un algoritmo no supervisado es clara, además de las discrepancias que habría si se etiquetasen los barrios de forma manual, ya que la gentrificación es un proceso en constante cambio y no hay un consenso sobre el tema debido a la abstracción del concepto y sus particularidades geográficas.

2.4. Decisiones tecnológicas

En este proyecto se utilizarán las herramientas y tecnologías con las que suelo trabajar en mi día a día y que se han trabajado en el máster. El lenguaje de programación principal será Python, con el que se harán los análisis de datos, limpieza y modelado. Al ser un proyecto serio y con cierta entidad, se utilizará PyCharm como IDE principal, aunque para ciertas tareas también es posible que se utilice Visual Studio Code, respaldando el código en todo momento por el control de versiones Git.

Para la herramienta SIG se ha escogido Carto, con la que gracias a la licencia de estudiante se podrán aprovechar todas las funcionalidades que ofrece.

Capítulo 3

Diseño e implementación del trabajo

En este apartado se detalla la fase del desarrollo e implementación del proyecto, desde la obtención de los datos hasta la obtención de una visualización de las distintas variables por barrios y un modelo que clasifica los barrios según su nivel de gentrificación.

No obstante, se dividirá el diseño e implementación según las distintas etapas o módulos necesarios que converjan en el producto final de este proyecto. Concretamente, primero se explicará en detalle la etapa de extracción, limpieza y análisis estadísticos simples de los datos. En segundo lugar se explicará la generación de la visualización en Carto que permitirá explorar las variables que afectan a la gentrificación y por último, la generación del modelo de aprendizaje no supervisado para clasificar los barrios según su nivel de gentrificación.

3.1. Extracción, limpieza y análisis

3.1.1. Lugares y ubicaciones de interés de los barrios (por año)

Tal como se ha estudiado en el estado del arte, el número de localizaciones de los distintos tipos de negocio o actividades es muy importante para identificar la gentrificación de los barrios. Por este motivo, en esta primera parte se va a extraer la información y ubicación de estos lugares que pueden estar relacionados con la gentrificación. A continuación, se muestra un esquema de los procesos realizados desde la extracción de los datos hasta su análisis, pasando por la etapa de limpieza y transformación de los datos.

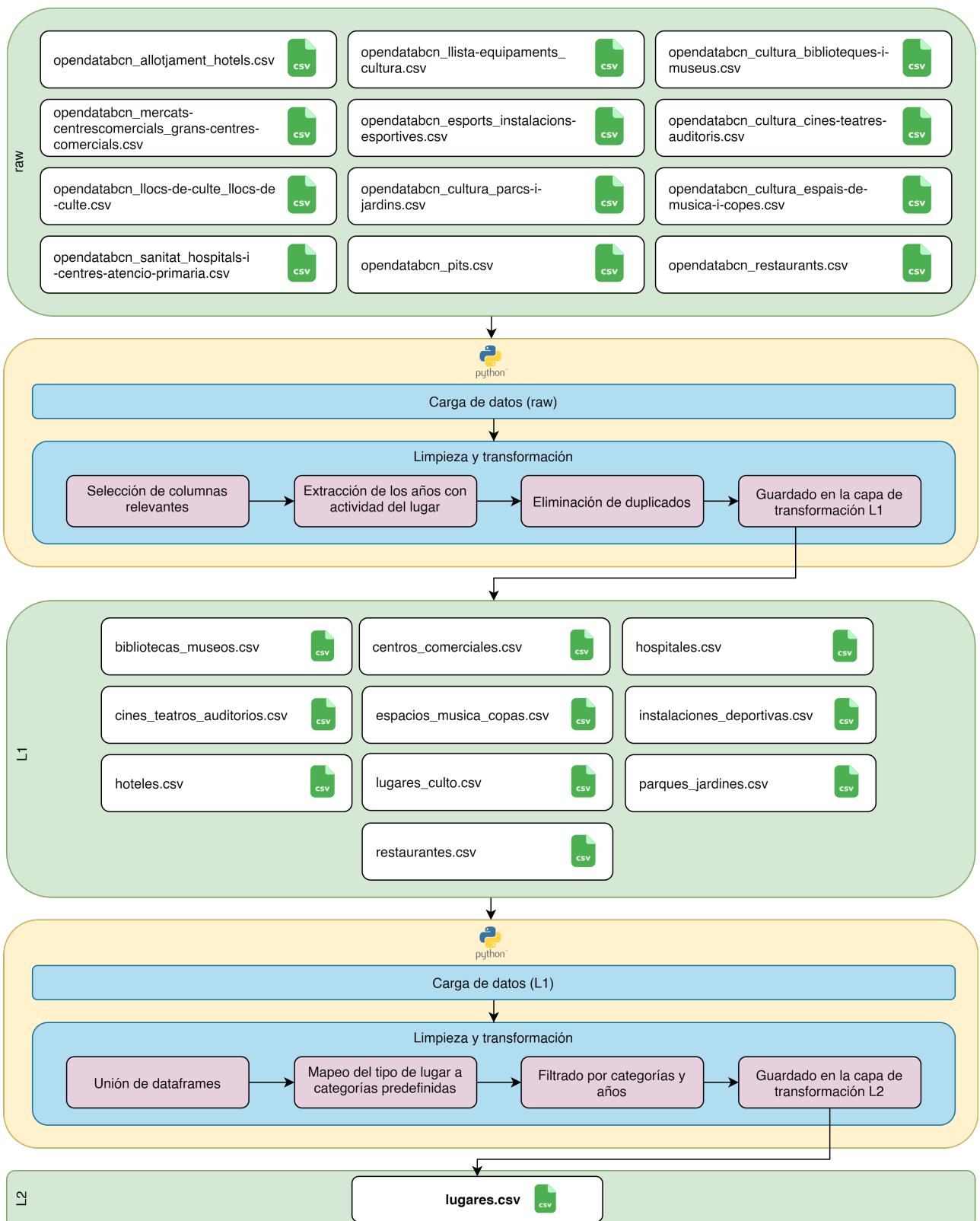


Figura 3.1: Diagrama del proceso de extracción, limpieza y transformación de los lugares de interés.

3.1.1.1. Fuentes de datos

Todos los conjuntos de datos que se han utilizado para obtener los lugares relevantes de la ciudad de Barcelona se han extraído de la web *Open Data BCN* [12]. Concretamente, los ficheros descargados son los siguientes:

- **opendatabcn_allotjament_hotels.csv**: Contiene información sobre los hoteles y alojamientos disponibles.
- **opendatabcn_cultura_biblioteques-i-museus.csv**: Datos sobre las bibliotecas y museos de la ciudad.
- **opendatabcn_cultura_cines-teatres-auditoris.csv**: Detalles sobre lugares de entretenimiento cultural, como cines, teatros y auditorios.
- **opendatabcn_cultura_espais-de-musica-i-copes.csv**: Información sobre los espacios de música, de copas y discotecas.
- **opendatabcn_cultura_parcs-i-jardins.csv**: Datos sobre las ubicaciones de zonas verdes: parques y jardines.
- **opendatabcn_esports_instalacions-esportives.csv**: Detalles sobre las distintas instalaciones deportivas de la ciudad.
- **opendatabcn_llista-equipaments_cultura.csv**: Datos sobre más equipamiento cultural y de ocio, como centros cívicos, casales de personas mayores a pesar de volver a encontrar los mismos lugares que en las fuentes de datos anteriores relacionadas con el ocio y la cultura.
- **opendatabcn_llocs-de-culte_llocs-de-culte.csv**: Ubicación de os lugares de culto de diferentes religiones.
- **opendatabcn_mercats-centrescomercials_grans-centres-comercials.csv**: Detalle sobre mercados y grandes centros comerciales
- **opendatabcn_pits.csv**: Información sobre más puntos de interés de la ciudad.
- **opendatabcn_restaurants.csv**: Detalles sobre los restaurantes disponibles.
- **opendatabcn_sanitat_hospitals-i-centres-atencio-primaria.csv**: Ubicación de los centros hospitalarios y de atención primaria.

3.1.1.2. Limpieza y transformación de los datos

La primera parte del código consiste en cargar todas las fuentes mencionadas en el apartado anterior. Debido a las peculiaridades de cada *dataset* se ha diseñado un fichero de configuración de carga (*config_paths.yaml*) totalmente personalizable que permite ajustarse a estas necesidades. Primero de todo, se puede escoger si los ficheros se cargarán desde el ordenador local (los ficheros deberán haber sido descargados previamente) o desde la URL del *dataset*. Entrando más en las configuraciones disponibles, para cada fuente se permite especificar el tipo de los datos, las columnas a seleccionar, un mapeo de estas columnas a los nombres deseados, el esquema del fichero *csv* y otras configuraciones de carga (cabecera, separador o codificación del CSV). Este fichero se puede consultar en el repositorio de Github de este proyecto: [config_paths.yaml](#)

Además, en caso de tratar con conjuntos de datos que están en diferentes ficheros según el año o cualquier otra partición, especificando todos los ficheros que formarán el *dataset* se concatenarán automáticamente.

Todos estos *datasets* tienen la misma estructura, por lo que el procesado será común para todos. Hay que remarcar que todos estos datos no están por defecto a nivel anual, si no que hay que realizar una pequeña operación para obtener una fila por año. En estos conjuntos hay dos columnas relacionadas con fechas. Una es la fecha de creación del registro y otra la última fecha de modificación. Por tanto, se ha supuesto que la fecha de creación indica la fecha de apertura del lugar, mientras que la de modificación indica la última actualización de los datos. Es decir, si por ejemplo existe un restaurante con valores `created=2017-01-12` y `updated=2021-08-17`, entonces se replicará la fila por cada año, el cual se especifica en una nueva columna. Sin embargo, de todas estas filas se eliminan las que tienen un año inferior a una fecha mínima, la cual es en este caso 2015. A continuación, se muestra un pequeño ejemplo de esta transformación.

name	created	modified	addresses.neighborhood_id
La Porca	2018-07-16	2021-03-03	11
secondary_filters.name	geo.epgs_4326.x	geo.epgs_4326.y	
Restaurants	41.3730	2.1708	

Cuadro 3.1: Ejemplo de un restaurante del dataset original con una selección de columnas.

nombre	id_barrio	nom_barrio	tipo_local	geo_epgs_4326_x	geo_epgs_4326_y	anyo
La Porca	11	el Poble-sec	Restaurants	41.3730	2.1708	2018
La Porca	11	el Poble-sec	Restaurants	41.3730	2.1708	2019
La Porca	11	el Poble-sec	Restaurants	41.3730	2.1708	2020
La Porca	11	el Poble-sec	Restaurants	41.3730	2.1708	2021

Cuadro 3.2: Ejemplo de la transformación del mismo restaurante en el que se obtiene una fila por cada año del intervalo definido por los años de las columnas originales `created` y `modified`.

De esta forma, se podrá obtener más tarde el número de lugares de cada tipo en cada año. El siguiente paso es eliminar duplicados, ya que un mismo lugar puede estar categorizado en el mismo dataset con dos valores distintos, como el caso en el que un restaurante está categorizado como por ejemplo "coctelería" además de "restaurante". Posteriormente, se guarda cada uno de estos *dataframes* que pertenecen a distintos tipos de lugares en un CSV. En el último paso, se leen todos estos ficheros transformados y se concatenan entre ellos. Se vuelve a hacer una eliminación de duplicados por si el mismo lugar está en diferentes ficheros de origen. De hecho en muchas ocasiones el *dataset* de puntos de interés contenía lugares del resto de fuentes. Una vez realizada esta acción, se mapean las categorías predeterminadas de los lugares a unas que se han determinado. Por ejemplo, todos los lugares con `categoria_lugar` ∈ [Museus, Biblioteques, Sales d'estudi, Teatres, Cinemes, Biblioteques municipals, Auditoris] se mapean al valor cultura.

Finalmente se eliminan las filas cuyos barrios o años de vigencia son desconocidos y se filtran los años según el intervalo de años de estudio, el cual va de los años 2015 a 2018, ya que fuera de este periodo los datos disponibles del resto de fuentes que se explicarán luego eran muy limitados. Después de estas transformaciones, se guarda la versión definitiva del *dataset* de lugares en el *path* `data/L2/lugares.csv`, el cual cuenta con 16268 filas y 11 columnas, las cuales son las siguientes:

- **nombre:** Nombre del lugar (*str*).
- **id_barrio:** Identificador del barrio (*int*).
- **nom_barrio:** Nombre del barrio (*str*).
- **tipo_lugar:** Tipo del lugar (*str*).
- **geo_epgs_25831_x:** Latitud en el sistema de coordenadas de proyección ETRS89 (zona 31N). Se usará para la representación posterior en Carto. (*float*)

- **geo_epgs_25831_y:** Longitud en el sistema de coordenadas de proyección ETRS89 (zona 31N). Se usará para la representación posterior en Carto. (*float*).
- **geo_epgs_4326_x:** Latitud en el sistema geodésico mundial WGS84. Se usará para la representación posterior en Carto. (*float*)
- **geo_epgs_4326_y:** Longitud en el sistema geodésico mundial WGS84. Se usará para la representación posterior en Carto. (*float*).
- **anyo:** Año en el que el lugar está vigente (*int*).
- **categoria_lugar:** Categoría del lugar generada a partir del mapeo de la columna tipo_lugar (*str*).
- **anyo_date:** Fecha del primer día del año especificado en la columna anyo. Necesario para filtrar por año en la línea temporal más tarde en Carto. (*date*)

nombre	id_barrio	nom_barrio	tipo_lugar	geo_epgs_25831_x	geo_epgs_25831_y	geo_epgs_4326_x	geo_epgs_4326_y	anyo	categoria_lugar	anyo_date
Bar Café del Born Nou	4	Sant Pere, Santa ...	Bars i pubs musicales	431744.772458145	4581835.2235519	41.38514950183597	2.183651690111131	2017	ocio	2017-01-01
Bar Café del Born Nou	4	Sant Pere, Santa ...	Bars i pubs musicales	431744.772458145	4581835.2235519	41.38514950183597	2.183651690111131	2018	ocio	2018-01-01
Bar musical Summum ...	8	l'Antiga Esquerra ...	Bars i pubs musicales	429687.7983651348	4582262.613959856	41.38882170822292	2.1590027056469863	2015	ocio	2015-01-01
Bar musical Summum ...	8	l'Antiga Esquerra ...	Bars i pubs musicales	429687.7983651348	4582262.613959856	41.38882170822292	2.1590027056469863	2016	ocio	2016-01-01
Bar musical Summum ...	8	l'Antiga Esquerra ...	Bars i pubs musicales	429687.7983651348	4582262.613959856	41.38882170822292	2.1590027056469863	2017	ocio	2017-01-01
Bar musical Summum ...	8	l'Antiga Esquerra ...	Bars i pubs musicales	429687.7983651348	4582262.613959856	41.38882170822292	2.1590027056469863	2018	ocio	2018-01-01
Bar musical Número 9	4	Sant Pere, Santa ...	Bars i pubs musicales	431682.1890884668	4581727.213593327	41.38417137430599	2.182915430465904	2015	ocio	2015-01-01
Bar musical Número 9	4	Sant Pere, Santa ...	Bars i pubs musicales	431682.1890884668	4581727.213593327	41.38417137430599	2.182915430465904	2016	ocio	2016-01-01
Bar musical Número 9	4	Sant Pere, Santa ...	Bars i pubs musicales	431682.1890884668	4581727.213593327	41.38417137430599	2.182915430465904	2017	ocio	2017-01-01
Bar musical Número 9	4	Sant Pere, Santa ...	Bars i pubs musicales	431682.1890884668	4581727.213593327	41.38417137430599	2.182915430465904	2018	ocio	2018-01-01
Sala Tarantos	2	el Barri Gòtic	Discoteques	431039.5634355078	4581240.772210454	41.3797353120014	2.1752857129194934	2015	ocio	2015-01-01
Sala Tarantos	2	el Barri Gòtic	Discoteques	431039.5634355078	4581240.772210454	41.3797353120014	2.1752857129194934	2016	ocio	2016-01-01
Sala Tarantos	2	el Barri Gòtic	Discoteques	431039.5634355078	4581240.772210454	41.3797353120014	2.1752857129194934	2017	ocio	2017-01-01
Sala Tarantos	2	el Barri Gòtic	Discoteques	431039.5634355078	4581240.772210454	41.3797353120014	2.1752857129194934	2018	ocio	2018-01-01
Restaurant Tablao de ...	11	el Poble-sec	Tablaos flamencs	428641.3109980409	4580088.501961428	41.36914809976753	2.1467430809493484	2015	ocio	2015-01-01
Restaurant Tablao de ...	11	el Poble-sec	Tablaos flamencs	428641.3109980409	4580088.501961428	41.36914809976753	2.1467430809493484	2016	ocio	2016-01-01
Restaurant Tablao de ...	11	el Poble-sec	Tablaos flamencs	428641.3109980409	4580088.501961428	41.36914809976753	2.1467430809493484	2017	ocio	2017-01-01
Restaurant Tablao de ...	11	el Poble-sec	Tablaos flamencs	428641.3109980409	4580088.501961428	41.36914809976753	2.1467430809493484	2018	ocio	2018-01-01
Discoteca Boulevard C...	1	el Raval	Discoteques	431037.2844809817	4581127.035320146	41.37871072162061	2.1752714032628853	2015	ocio	2015-01-01
Discoteca Boulevard C...	1	el Raval	Discoteques	431037.2844809817	4581127.035320146	41.37871072162061	2.1752714032628853	2016	ocio	2016-01-01
Discoteca Boulevard C...	1	el Raval	Discoteques	431037.2844809817	4581127.035320146	41.37871072162061	2.1752714032628853	2017	ocio	2017-01-01
Discoteca Boulevard C...	1	el Raval	Discoteques	431037.2844809817	4581127.035320146	41.37871072162061	2.1752714032628853	2018	ocio	2018-01-01

Figura 3.2: Muestra del *dataset* de lugares.

3.1.2. Lugares y ubicaciones de interés de los barrios (año 2019)

En el apartado anterior se ha generado un conjunto de datos con distintas ubicaciones de interés que se pueden encontrar por cada año de 2015 a 2018. No obstante, hay muchos más lugares como múltiples tipos de negocio que no están en ningún *dataset* historificado anualmente. No obstante, existe un *dataset* que tiene un listado completo de todos los negocios de cada barrio para el año 2019, incluyendo establecimientos relevantes como tiendas de ropa,

de alimentación, centros educativos, etc., que no se incluían en el anterior y que por lo tanto, contendrá muchos más datos. Como después se verá, en el *dataset* anterior hay muy poca variabilidad entre las ubicaciones a lo largo de los años, por lo que se considera que este *dataset* de negocios del 2019 puede ser más relevante para detectar los barrios gentrificados debido a su mayor riqueza de ubicaciones.

El procedimiento de transformación será parecido pero bastante más simple, ya que solo será necesario leer un solo fichero CSV que contiene todas las ubicaciones. Igual que antes, se leen las columnas relevantes con el esquema adecuado y se guarda en la carpeta *data/L1*. Posteriormente, se le aplicarán transformaciones simples para mapear y resumir todas las categorías de los lugares y poder representarlos en el mapa, así como generar gráficos que permitan ver la distribución de los tipos de lugares en cada barrio.

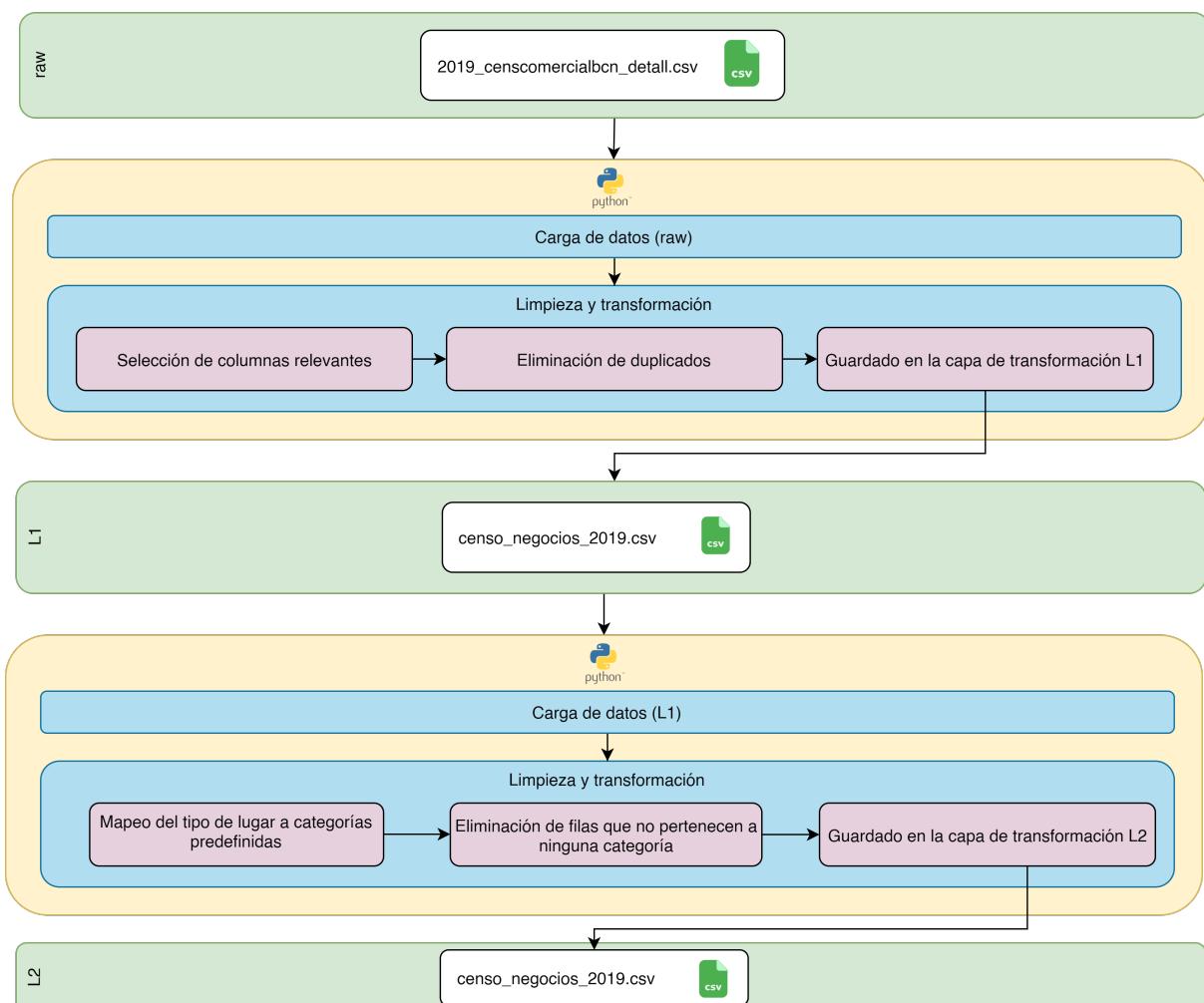


Figura 3.3: Diagrama del proceso de extracción, limpieza y transformación de los lugares de interés del año 2019.

3.1.2.1. Fuentes de datos

Como se ha comentado previamente, para construir este conjunto de datos solo es necesario un único fichero CSV, el cual de nuevo se ha extraído de la web *Open Data BCN* [12].

- **2019_CensComercialBCN_Detall.csv:** Contiene el censo de locales de 2019 de la ciudad de Barcelona con el objetivo de identificar todos los locales con actividad económica.

3.1.2.2. Limpieza y transformación de los datos

De igual forma, la transformación de datos a realizar es también más sencilla al tratarse de una única fuente de datos. Una vez leído el fichero crudo, en la primera transformación se leen solo las columnas de interés y se eliminan duplicados. En este caso, no habían duplicados explícitos, ya que todas las filas tenían identificadores diferentes. No obstante, sí que había locales con mismo nombre y mismas coordenadas, por lo que en estos casos es donde se ha aplicado la eliminación de duplicados. No se ha detectado un motivo claro para explicar qué causa estos duplicados, ya que absolutamente todos los campos excepto el identificador son exactos en todas las filas replicadas y tampoco se menciona este detalle en la descripción del fichero original. Después de este tratamiento inicial, se guarda el fichero resultante en *data/L1*.

En la siguiente transformación, se analizan todos los tipos de lugares que hay y se construye como antes un mapeo para relacionar dichos lugares con unas categorías predefinidas. En este caso, al tener más lugares disponibles, las nuevas categorías mapeadas son más ricas y ahora hay nuevos tipos de lugares que antes no habían. Por ejemplo, se ha creado una nueva categoría que une los centros comerciales con las tiendas de ropa, se han añadido también las supermercados y tiendas de alimentación, negocios tradicionales (mercerías, tintorerías, ferreterías, etc.), centros de estética y peluquerías, y finalmente centros educativos.

Finalmente, una vez realizado el mapeo de valores, se eliminan aquellos lugares que no han podido ser mapeados porque no estaban especificados. El resultado de esta transformación se guarda en *data/L2*, que ya contiene todo lo necesario para ser representado en el mapa. Este *dataset* está formado por 41687 filas y 8 columnas, que se especifican a continuación:

- **grupo_actividad:** Tipo del grupo de la actividad realizada en el lugar (*str*).
- **nom_actividad:** Tipo específico de la actividad realizada en el lugar (*str*).

- **nombre:** Nombre del lugar (*str*).
- **lat:** Latitud en el sistema geodésico mundial WGS84 (*float*).
- **lon:** Longitud en el sistema geodésico mundial WGS84 (*float*).
- **id_barrio:** Identificador del barrio (*int*).
- **nom_barrio:** Nombre del barrio (*str*).
- **categoria_lugar:** Categoría del lugar generada a partir del mapeo de la columna nom_actividad (*str*).

grupo_actividad	nom_actividad	nombre	lat	lon	id_barrio	nom_barrio	categoria_lugar
Altres	Perruqueries	SD & GAR	41.43002297	2.18513399	60	Sant Andreu	estetica_peluquerias
Restaurants, bars i hotels (Inclòs ho...)	Bars / CIBERCAFÉ	JAMMING LOUNGE	41.40351545	2.1579625	31	la Vila de Gràcia	restaurantes
Altres	Basars	BASAR SPER GANGA	41.38150396	2.16388285	1	el Raval	negocios_tradicionales
Oci i cultura	Joguines i esports	PESCA WORLD	41.38067178	2.16627806	1	el Raval	deporte
Restaurants, bars i hotels (Inclòs ho...)	Restaurants	CUINA D'AMICS	41.394433759	2.19866171	67	la Vila Olímpica del Poblenou	restaurantes
Restaurants, bars i hotels (Inclòs ho...)	Restaurants	BAR HERRANZ	41.44269019	2.17558265	51	Verdun	restaurantes
Restaurants, bars i hotels (Inclòs ho...)	Restaurants	COMO EN CASA	41.39031823	2.14062274	19	les Corts	restaurantes
Ensenyament	Ensenyament	LALAL'ART	41.40692224	2.1373491	25	Sant Gervasi - la Bonanova	educacion
Sanitat i assistència	Sanitat i assistència	ORTOPEDIA	41.41281637	2.19729111	72	Sant Martí de Provençals	sanidad
Quotidià alimentari	Pa, pastisseria i làctics	FORN DE PA SABADELL	41.40413491	2.16526553	32	el Camp d'en Grassot i Grà...	minoristas_alimentacion_supermercado
Equipaments culturals i recreatius	Equipaments culturals i recreatius	EL PUNT DE REUNIO	41.37263717	2.15957657	11	el Poble Sec	ocio_cultura
Equipaments culturals i recreatius	Equipaments culturals i recreatius	CASAL CONCORDIA	41.3727056	2.15963289	11	el Poble Sec	ocio_cultura
Restaurants, bars i hotels (Inclòs ho...)	Bars / CIBERCAFÉ	BREAKFAST & LUNCH	41.38364303	2.17785883	2	el Barri Gòtic	restaurantes
Sanitat i assistència	Sanitat i assistència	DRA BERTA LLORENS	41.40352432	2.12844416	24	les Tres Torres	sanidad
Restaurants, bars i hotels (Inclòs ho...)	Bars / CIBERCAFÉ	THE SWEETOPHELIA CAFE	41.39049632	2.17641681	7	la Dreta de l'Eixample	restaurantes
Ensenyament	Ensenyament	CENTRE D'ESTUDIS MUSICAL...	41.39680993	2.17650306	7	la Dreta de l'Eixample	educacion
Altres	Souvenirs	AMOR BARCELONA	41.38362211	2.17759028	2	el Barri Gòtic	negocios_tradicionales
Oci i cultura	Llibres, diaris i revistes	DIT I FET LLIBRERIA	41.43544045	2.19020873	60	Sant Andreu	negocios_tradicionales
Equipament personal	Joieria, rellotgeria i bijuteria	SIMO	41.43000387	2.18738973	60	Sant Andreu	negocios_tradicionales
Altres	Perruqueries	CREARTE	41.397937	2.167193	7	la Dreta de l'Eixample	estetica_peluquerias
Altres	Grans magatzems i hipermercats	EL CORTE INGLES	41.3925356	2.14635499	8	l'Antiga Esquerra de l'Eixa...	ropa_cc
Sanitat i assistència	Sanitat i assistència	CENTRE DIAGNOSTIC DR. MA...	41.41094258	2.13630304	25	Sant Gervasi - la Bonanova	sanidad
Ensenyament	Ensenyament	ESCOLA CALDERON DE LA BA...	41.43368575	2.17158315	46	el Turó de la Peira	educacion
Quotidià alimentari	Pa, pastisseria i làctics	MASSLNARA	41.38097113	2.15753899	10	Sant Antoni	minoristas_alimentacion_supermercado
Quotidià alimentari	Autoservei / Supermercat	SPAR	41.40140856	2.16790981	32	el Camp d'en Grassot i Grà...	minoristas_alimentacion_supermercado

Figura 3.4: Muestra del *dataset* completo de lugares de 2019.

3.1.3. Indicadores de gentrificación de los barrios

Hasta ahora, el foco de la primera parte estaba en obtener los lugares de las diferentes tipologías que pueden estar relacionadas con la gentrificación. Sin embargo, tal como se comentaba en capítulos anteriores, hay más factores que pueden influir en la gentrificación de los barrios aparte de simplemente los tipos de negocios o ubicaciones que se localizan en cada uno de ellos. A continuación, se muestra un diagrama resumen de transformaciones de las fuentes de datos hasta generar el *dataset* con todos los indicadores de los barrios que potencialmente están relacionados con la gentrificación.

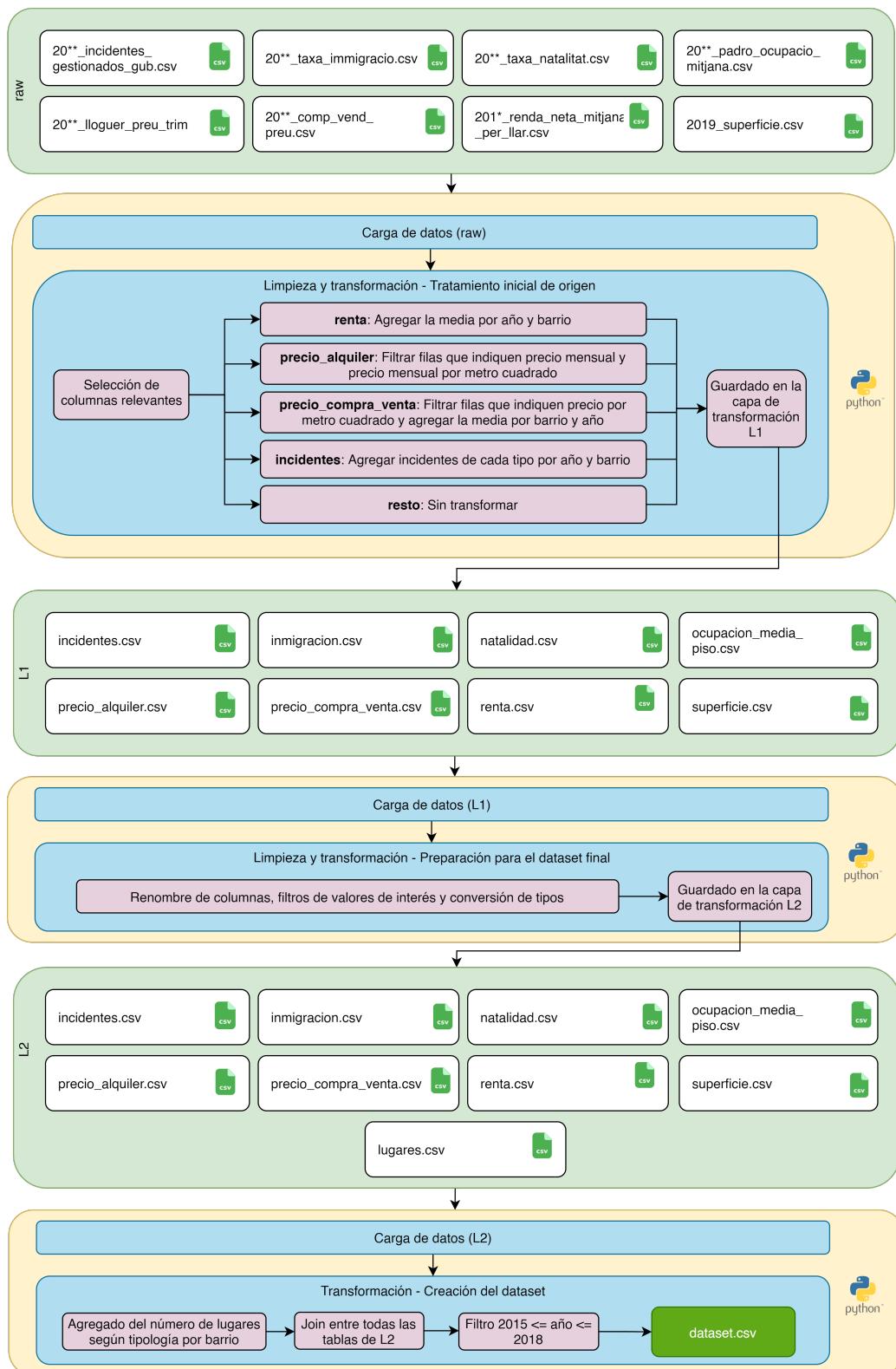


Figura 3.5: Diagrama del proceso de extracción, limpieza y transformación de los indicadores de cada barrio potencialmente relacionados con la gentrificación.

3.1.3.1. Fuentes de datos

Todos los datos de indicadores han sido extraídos también de la web *Open Data BCN* [12], ya que es el único portal que dispone gran cantidad de datos de la ciudad que estén a nivel de barrio. Concretamente, todas las fuentes de datos que se han utilizado para generar el *dataset* de indicadores para cada barrio son las siguientes.

- **20**_incidents_gestionats_gub.csv:** Incidentes por barrio de cada tipo gestionados por la Guardia Urbana de Barcelona. Un fichero CSV para cada año desde 2014 a 2020.
- **20**_taxa_immigracio.csv:** Tasa de inmigración por barrio. Un fichero CSV para cada año desde 2014 a 2020.
- **20**_taxa_natalitat.csv:** Tasa de natalidad por barrio. Un fichero CSV para cada año desde 2014 a 2020.
- **20**_padro_ocupacio_mitjana.csv:** Número medio de personas por vivienda de cada barrio. Un fichero CSV para cada año desde 2014 a 2020.
- **20**_lloguer_preu_trim.csv:** Precios del alquiler por barrio y trimestre. Un fichero CSV para cada año desde 2014 a 2021.
- **20**_comp_vend_preu.csv:** Precios de compra venta medios de una vivienda por barrio. Un fichero CSV para cada año desde 2014 a 2020.
- **201*_renda_neta_mitjana_per_llar.csv:** Renta neta media por sección censal de cada barrio. Un fichero CSV para cada año desde 2015 a 2018. Desgraciadamente, no hay datos de 2014 ni 2020 como en el resto de fuentes, por lo que al ser la renta una variable muy importante deberemos adaptar el estudio para contemplar únicamente los años de 2015 a 2018, periodo que puede ser suficiente para analizar la gentrificación de los barrios.
- **2019_superficie.csv:** Superficie de cada barrio en hectáreas.

No obstante, existen otras variables muy interesantes para el análisis que se han visto en el estudio del arte, como el nivel de estudios de la población o la edad de media de los habitantes, las cuales no se han podido encontrar por barrio.

3.1.3.2. Limpieza y transformación de los datos

Para la generación del dataset de indicadores de barrios, se sigue el mismo procedimiento que se hizo anteriormente para el dataset de lugares. Por tanto, todos los ficheros fuentes se cargan de la misma manera. La única diferencia es que la mayoría de las fuentes de datos están particionadas por año o trimestre, traduciéndose en un fichero CSV por cada una de estas particiones. Lo que se hace para estos casos es indicar para cada fichero nuevo que queremos generar (se guardará en la carpeta *data/L1*) el listado de ficheros que se han de leer y concatenar para tener todos los datos en un único fichero. Por ejemplo, para generar el fichero CSV con las tasas de natalidad, en el fichero de configuración de lectura (*config_paths.yaml*) se definen estos orígenes de la siguiente manera:

```
natalidad:  
  local:  
    - data/raw/natalidad/2020_taxa_natalitat.csv  
    - data/raw/natalidad/2019_taxa_natalitat.csv  
    - data/raw/natalidad/2018_taxa_natalitat.csv  
    - data/raw/natalidad/2017_taxa_natalitat.csv  
    - data/raw/natalidad/2016_taxa_natalitat.csv  
    - data/raw/natalidad/2015_taxa_natalitat.csv  
    - data/raw/natalidad/2014_taxa_natalitat.csv
```

Sin embargo, no es trivial hacer la concatenación de los datos, ya que en algunos casos las columnas tienen diferentes nombres o el orden de las columnas tampoco es igual. Para estos casos, ya se definen acciones alternativas de selección de columnas en caso de que alguno de los ficheros no encaje con la estructura esperada que está definida en la configuración de lectura de ficheros.

Una vez concatenadas todas las particiones, se realizan pequeñas transformaciones en algunos datasets para adecuarlos a la granularidad y valores esperados en la siguiente etapa de transformaciones. Por ejemplo, en el *dataframe* de renta se agrega por la media de cada año y barrio, ya que las fuentes respectivas están a nivel de sección censal. Para el caso del precio del alquiler, hay una columna que indica para cada año y barrio, la métrica correspondiente de precio de alquiler, de las cuales solo interesan las que indican el precio del alquiler por metro cuadrado y el precio mensual absoluto. Una vez aplicado el filtro, se tendrían dos filas por cada año y barrio, por lo que se transforma para tener dos columnas (una para el precio por metro cuadrado y otra para el absoluto) por cada año y barrio. Luego, para el *dataframe* de precio de compra-venta, se hace el mismo tratamiento que para el precio de alquiler, con

el único detalle añadido que primero se realiza también una agregación por barrio y año, ya que hay alguna partición que tiene los datos a nivel trimestre. Finalmente, para los incidentes, al tenerlos originariamente por barrio, mes y tipología, una vez concatenados se agregan por barrio, año y tipo de incidente. El resto de ficheros se guardarán también en la carpeta *data/L1* tal cual una vez concatenados con la configuración de lectura definida en el fichero mencionado anteriormente.

Una vez tratados mínimamente los ficheros de origen, en el siguiente nivel de procesado, se preparan los *datasets* para generar el *dataset* final de indicadores. Es decir, se renombran algunas columnas y se eliminan algunas que no son relevantes. El único caso donde se realiza una transformación algo más compleja es en los incidentes. Para los incidentes, realmente no interesa el número de cada tipo, si no los incidentes totales que pueden estar relacionados con violencia, robos, vandalismo, etc. Es decir, aquellos que estarían relacionados con la inseguridad de los barrios. Concretamente, se toman como incidentes aquellos que tienen algunas de las siguientes tipologías: convivencia vecinal, atentados, peleas, vandalismo, actividades molestas en espacios públicos, agresiones, alarmas de incendio o robo, ocupaciones ilícitas, delitos contra el orden público, conflictos en locales, actos contra la propiedad privada, violencia doméstica, estupefacientes y contra la libertad sexual. Para ello, se filtran aquellos tipos de incidentes pertenecientes a los tipos que se acaban de listar y se agregan por año y barrio. El resultado de estas transformaciones se guarda en ficheros CSV en la carpeta *data/L2*.

Por último, para generar el dataset final se debe hacer *join* de todas las tablas que se han generado, las cuales ya se han preparado para que tengan los datos a nivel año y barrio. Además, se recupera el *dataset* de lugares generado anteriormente y se genera para cada año y barrio una columna con el número de lugares de cada tipo. Finalmente, se filtra el *dataset* final para que solo contenga datos de los años 2015 a 2018, que son en los que se centrará el análisis. Este dataset final de indicadores se guarda en el path *data/dataset/dataset.csv*, el cual tiene 293 filas y 23 columnas, las cuales son las siguientes:

- **id_barrio:** Identificador del barrio (*int*).
- **nom_barrio:** Nombre del barrio (*str*).
- **superficie_ha:** Superficie del barrio en hectáreas (*float*).
- **anyo:** Año al que hace referencia los indicadores del barrio (*int*).
- **num_incidentes:** Número de incidentes. (*int*).

- **inmigracion_mil_hab:** Tasa de inmigración por mil habitantes (*float*).
- **tasa_natalidad_mil_habitantes:** Tasa de natalidad por mil habitantes (*float*).
- **poblacion:** Población de cada barrio (*int*).
- **num_domicilios:** Número de domicilios (*int*).
- **num_personas_por_domicilio:** Número medio de personas residentes en el mismo domicilio (*float*).
- **precio_alquiler_mes_m2:** Precio medio del alquiler por metro cuadrado en euros (*float*).
- **precio_compra_venta_m2:** Precio medio de compra-venta por metro cuadrado en euros (*float*).
- **renta:** Renta media de los habitantes del barrio en euros (*float*).
- **num_ubic_centros_comerciales:** Número de centros comerciales (*int*).
- **num_ubic_cultura:** Número de lugares relacionados con la cultura (cines, teatros, etc.) (*int*).
- **num_ubic_deporte:** Número de centros deportivos (gimnasios, polideportivos, etc.) (*int*).
- **num_ubic_gastronomia:** Número de locales relacionados con la gastronomía (restaurantes) (*int*).
- **num_ubic_hoteles:** Número de hoteles (*int*).
- **num_ubic_lugares_culto:** Número de lugares de culto (*int*).
- **num_ubic_ocio:** Número de lugares relacionados con el ocio (discotecas, karaokes, etc.) (*int*).
- **num_ubic_parques_jardines:** Número de parques y jardines (*int*).
- **num_ubic_sanidad:** Número de centros sanitarios (hospitales, CAPs) (*int*).
- **anyo_date:** Fecha del primer día del año especificado en la columna anyo (*date*).

3.1. Extracción, limpieza y análisis

40

id_barrio	nom_barrio	superficie_ha	anyo	num_incidentes	inmigracion_mil_hab	tasa_natalidad_mil_habitantes	poblacion	num_domicilios	num_personas_por_domicilio	precio_alquiler_mes_m2	precio_compra_venta_m2
1	el Raval	110.0	2015	8111	113.2	9.1	47617	16776	2.84	11.0	2401.9
1	el Raval	110.0	2016	8698	105.5	9.6	47274	16851	2.81	12.27	2805.2
1	el Raval	110.0	2017	9346	113.4	10.5	47986	17090	2.81	13.9	3469.9
1	el Raval	110.0	2018	10788	121.39	9.29	47605	17212	2.77	14.02	3969.8
2	el Barri Gòtic	81.6	2015	6925	142.2	6.4	15269	6461	2.36	11.29	3193.4
2	el Barri Gòtic	81.6	2016	8130	140.7	6.2	15729	6450	2.44	13.01	4149.4
2	el Barri Gòtic	81.6	2017	7654	167.2	6.7	16240	6537	2.48	14.19	4565.8
2	el Barri Gòtic	81.6	2018	7997	206.32	7.53	17035	6565	2.59	13.77	5162.7
3	la Barceloneta	109.5	2015	2770	103.3	6.5	15036	6820	2.2	16.33	3474.4
3	la Barceloneta	109.5	2016	2699	103.0	6.1	15068	6850	2.2	18.21	4151.9
3	la Barceloneta	109.5	2017	2603	117.2	6.7	15101	6911	2.19	19.84	4501.1
3	la Barceloneta	109.5	2018	2759	114.05	5.56	14893	6917	2.15	16.92	4905.4
4	Sant Pere, Santa Caterina i la Ribera	111.0	2015	3727	113.6	7.3	22305	9869	2.26	12.74	3479.4
4	Sant Pere, Santa Caterina i la Ribera	111.0	2016	4611	106.5	7.7	22380	9954	2.25	14.06	3646.5
4	Sant Pere, Santa Caterina i la Ribera	111.0	2017	4897	121.6	7.9	22923	10154	2.26	15.7	4593.7
4	Sant Pere, Santa Caterina i la Ribera	111.0	2018	4824	127.06	6.55	22605	10238	2.21	15.02	5169.4
5	el Fort Pienc	92.9	2015	1539	63.5	7.5	31645	12880	2.46	10.89	3602.0
5	el Fort Pienc	92.9	2016	1810	62.9	8.0	31693	12805	2.48	11.74	3762.4
5	el Fort Pienc	92.9	2017	1959	69.8	7.2	32048	12838	2.5	13.08	4296.6
5	el Fort Pienc	92.9	2018	1855	75.47	8.36	32080	12831	2.5	13.14	4650.6
6	la Sagrada Família	104.2	2015	2166	63.9	7.6	51347	21850	2.35	11.2	2983.9
6	la Sagrada Família	104.2	2016	2555	61.4	7.7	51349	21844	2.35	12.19	2785.8
6	la Sagrada Família	104.2	2017	2661	69.6	7.5	51651	21920	2.36	13.33	3494.5
6	la Sagrada Família	104.2	2018	2753	74.24	7.42	51654	21981	2.35	13.72	3833.1

renta	num_ubic_centros_comerciales	num_ubic_cultura	num_ubic_deporte	num_ubic_gastronomia	num_ubic_hoteles	num_ubic_lugares_culto	num_ubic_ocio	num_ubic_parques_jardines	num_ubic_sanidad	anyo_date
23379.57	0	45	3	101	45	21	26	7	3	2015-01-01
23866.76	0	46	3	114	45	21	27	7	3	2016-01-01
24383.95	0	47	3	119	46	22	28	7	3	2017-01-01
25485.81	0	47	3	145	47	22	29	7	3	2018-01-01
26896.56	3	25	3	115	51	13	21	1	1	2015-01-01
27044.11	3	25	3	119	52	13	22	1	1	2016-01-01
27260.0	3	25	3	122	54	13	22	1	1	2017-01-01
28267.56	3	25	3	158	57	13	22	2	1	2018-01-01
23761.18	0	8	9	66	5	4	6	2	3	2015-01-01
24206.36	0	8	8	70	5	4	6	2	3	2016-01-01
24598.45	0	8	8	77	5	4	6	2	3	2017-01-01
25867.0	0	8	8	90	5	4	6	2	3	2018-01-01
26178.46	0	33	1	98	19	11	20	1	1	2015-01-01
27142.0	0	34	1	102	19	11	20	1	1	2016-01-01
27412.62	0	35	1	112	20	11	20	1	1	2017-01-01
29068.0	0	36	1	141	23	11	21	1	1	2018-01-01
37344.5	0	10	5	18	10	3	2	2	2	2015-01-01
37836.45	0	10	5	20	10	3	2	2	2	2016-01-01
38720.1	0	10	5	22	10	3	3	2	2	2017-01-01
40134.55	0	10	5	28	10	3	3	2	2	2018-01-01
36138.29	0	8	2	38	4	5	3	1	2	2015-01-01
36530.32	0	8	2	38	4	5	3	2	2	2016-01-01
37014.79	0	8	2	42	4	5	3	2	2	2017-01-01
38589.24	0	8	2	56	4	5	3	2	2	2018-01-01

Figura 3.6: Muestra del *dataset* de indicadores.

Sin embargo, de cara a representar todos estos datos en un mapa, se necesita disponer de la geometría de los barrios de Barcelona. Esta se puede obtener también del mismo portal de datos abiertos y se puede [descargar](#) en distintos formatos, como *geojson*, o *shapefile*, entre otros.

A través de la librería de Python *geopandas*, se puede leer la geometría de los barrios y hacer *join* con el *dataset* de indicadores previamente generado. Una vez se ha realizado esta operación, se obtiene un nuevo *dataset* exactamente igual pero con los datos geométricos de los barrios, que será el que se subirá a Carto para realizar las visualizaciones.

3.2. Generación de las visualizaciones en Carto

Una vez ya se tienen todos los datos limpiados y preparados, el siguiente paso es generar una visualización para analizar las distintas variables en cada barrio. Como se comentó al inicio, la herramienta seleccionada para realizar este análisis es Carto.

El primer paso es subir los ficheros necesarios en el apartado de *Datasets*. Se suben por lo tanto el fichero *geojson* con la geometría de los barrios, el *dataset* de indicadores (que contiene también la geometría) y por último los *datasets* de lugares, los cuales contienen las ubicaciones de lugares o negocios que pueden ser relevantes para identificar los barrios gentrificados.

Your Datasets		Last Modified	Rows	Size	Builder Usage	Privacy
...	censo_negocios_2019	Updated 46 minutes ago	41.69k	7.17 MB	1 map	Public
...	lugares	Updated 4 days ago	15.92k	2.75 MB	1 map	Public
...	barris	Updated 4 days ago	73	656 kB	2 maps	Public
...	dataset_inidcadores_barrios	Updated 5 days ago	292	2.43 MB	1 map	Public

Figura 3.7: Datasets en Carto en los que se basarán las visualizaciones. Fuente: elaboración propia.

Para no sobrecargar en un solo mapa toda la información, se ha decidido realizar tres visualizaciones. Las dos primeras permiten explorar las ubicaciones de cada barrio. La primera de las dos, contendrá los lugares obtenidos en la primera parte, los cuales están a nivel anual de 2015 a 2018 y pertenecen solo a unas cuantas categorías, mientras que la segunda será exclusivamente del año 2019 y su contenido se aproximará a la totalidad real de todos los lugares y negocios que hay en Barcelona. Por último, la tercera visualización se centra en visualizar los indicadores de cada barrio que no están relacionados con las ubicaciones que se encuentran en ellos.

3.2.1. Visualización de ubicaciones relevantes

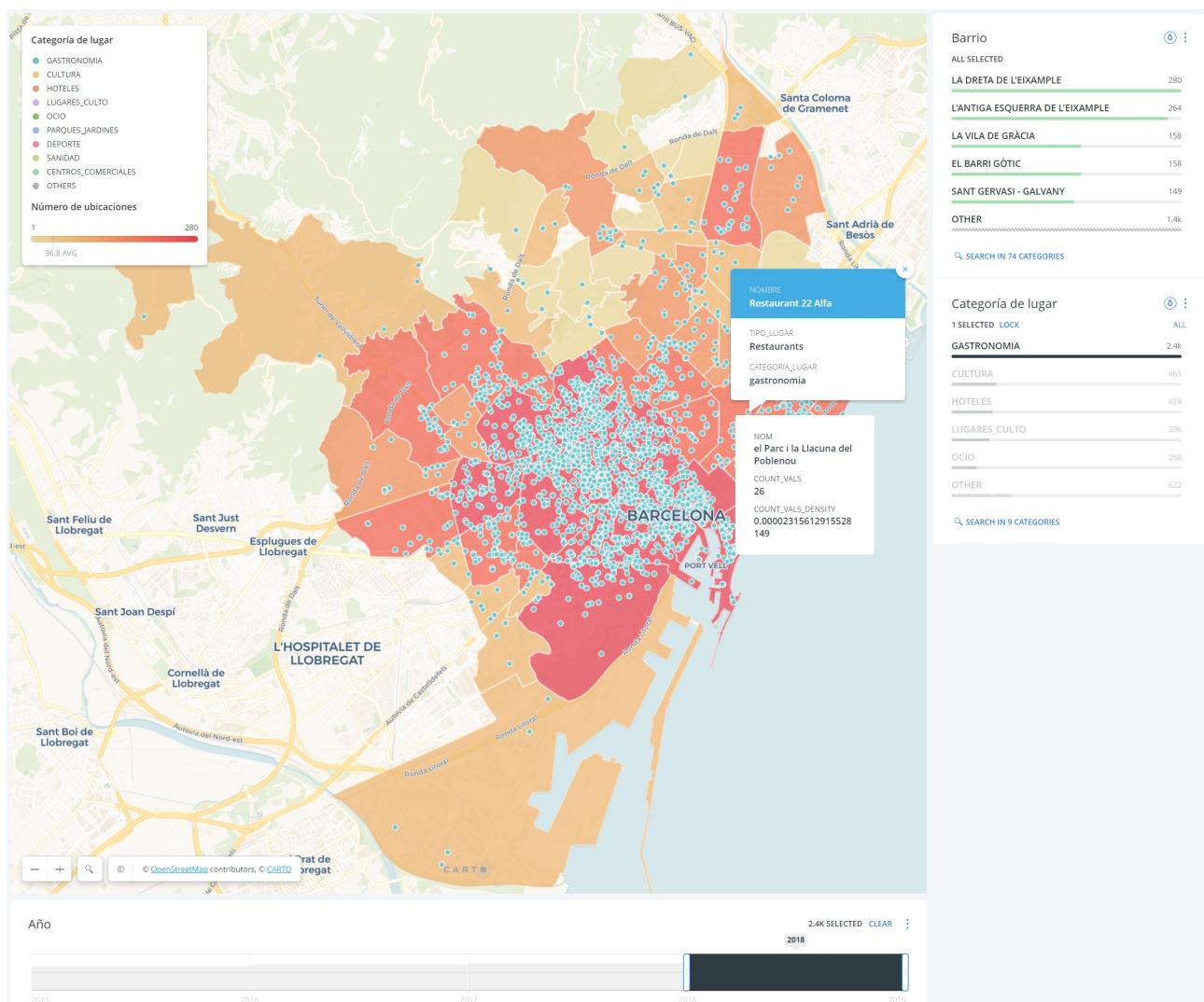


Figura 3.8: Visualización de los lugares relevantes anualmente por barrio con filtro por categoría gastronomía y año 2018. Fuente: elaboración propia. Accesible en [Carto](#).

En esta primera visualización, se muestran todas las ubicaciones o lugares que pueden estar relacionados con la gentrificación de Barcelona de los cuales existen datos históricos. Para ver el aspecto general que tiene el mapa, se puede consultar la Figura 3.8, o se puede acceder al mapa interactivo completo haciendo clic en el enlace disponible en el pie de foto de dicha imagen. En el mapa, se puede observar un punto por cada ubicación, cuyo color dependerá de la categoría del lugar, el cual se puede ver en la leyenda superior izquierda. También se puede observar que las delimitaciones de los barrios son claramente diferenciables, y que cada barrio tiene un color distinto de acuerdo a una escala de color, el cual depende de la densidad de puntos que haya en cada barrio. Esta operación se realiza mediante un

análisis nativo de Carto, que dada una capa de geometría (límites de los barrios) y otra con coordenadas (lugares) calcula la intersección y hace un recuento de los puntos que caen en cada barrio. Si se observa el panel derecho, aquí es donde se permite interactuar con los barrios y las categorías de los lugares para ver el mapa de coropletas según las variables seleccionadas, además de poder visualizar el número de lugares que cumplen la categoría o que pertenecen a cada barrio. La operativa habitual que se hará para el análisis será hacer clic en una de las categorías, de forma que se filtrarán las ubicaciones y solo se mostrarán los puntos pertenecientes a la categoría seleccionada. Consecuentemente, esta acción provocará un nuevo cálculo de las intersecciones con la capa de los barrios y por tanto se volverán a pintar los barrios según las ubicaciones de la variable escogida.

Por otro lado, en la parte de abajo, se muestra una línea temporal que permite filtrar por año. Esto es importante, ya que si no se selecciona ningún año, el conteo total de ubicaciones por barrio será el acumulado. Es decir, si un lugar ha estado vigente entre 2015 y 2018, se contaría cuatro veces en el total. Como se verá luego, las diferencias de ubicaciones entre años son poco significativas, por lo que con la visualización es complicado ver las evoluciones de los lugares en distintos años, por lo que lo más sencillo es seleccionar un solo año para realizar el análisis visual.

Por último, si se pasa el cursor por encima del mapa, se muestra el nombre del barrio sobre el que se está, el total de puntos que hay visibles en él y la densidad de puntos. También, si se quiere ver más información de un punto concreto, haciendo clic sobre él se muestra el nombre del lugar, su tipo y categoría.

Debido a estas pocas variaciones entre años y a la dificultad de verlas en el mapa, dará más valor la segunda visualización que tiene todos los lugares relevantes de Barcelona para el año 2019. El procedimiento de generarla es exactamente el mismo a la anterior, con la única diferencia de que ya no hay un filtro temporal a nivel anual en la parte inferior, puesto que todos los datos pertenecen al año 2019. Esta visualización se puede ver en la figura 3.9 y se puede acceder también a su mapa interactivo en Carto haciendo clic en el enlace del pie de foto.

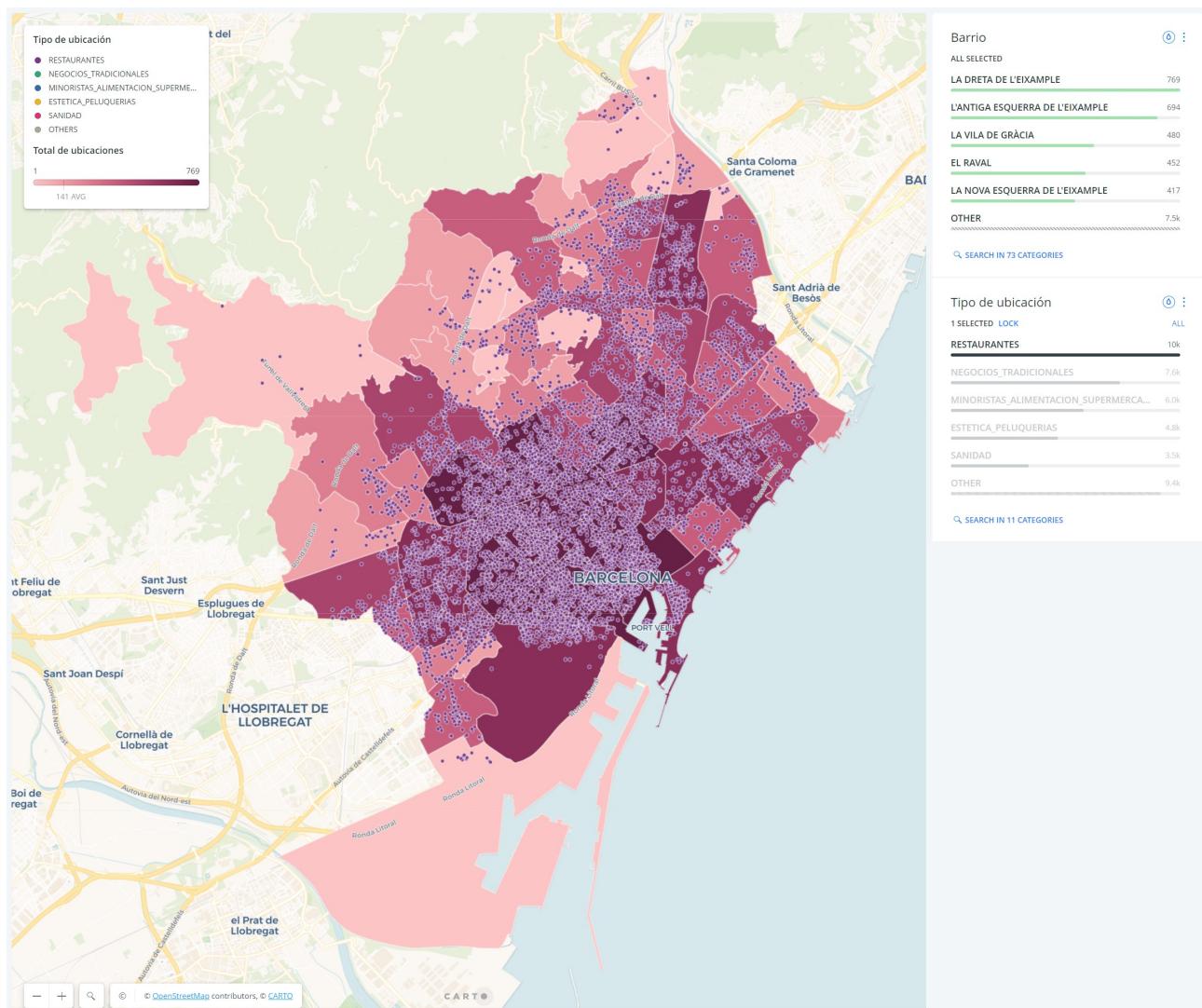


Figura 3.9: Visualización completa de los lugares relevantes por barrio en 2019 con el filtro de restaurantes. Fuente: elaboración propia. Accesible en [Carto](#).

En el apéndice que hay al final de este documento, se muestran en capturas todos los mapas de coropletas para todos los tipos de lugares, de forma que se puede ver de forma inmediata en qué zonas geográficas se concentran la mayoría de negocios de cada tipo. Además, se incluyen diagramas de radar para todos los barrios que muestran la distribución de los distintos tipos de ubicaciones en cada barrio, pudiendo complementar la información que ofrecen los mapas de coropletas. Estos diagramas se pueden consultar en el apéndice A.1 y los mapas de coropletas en A.2.

3.2.2. Visualización de indicadores relevantes

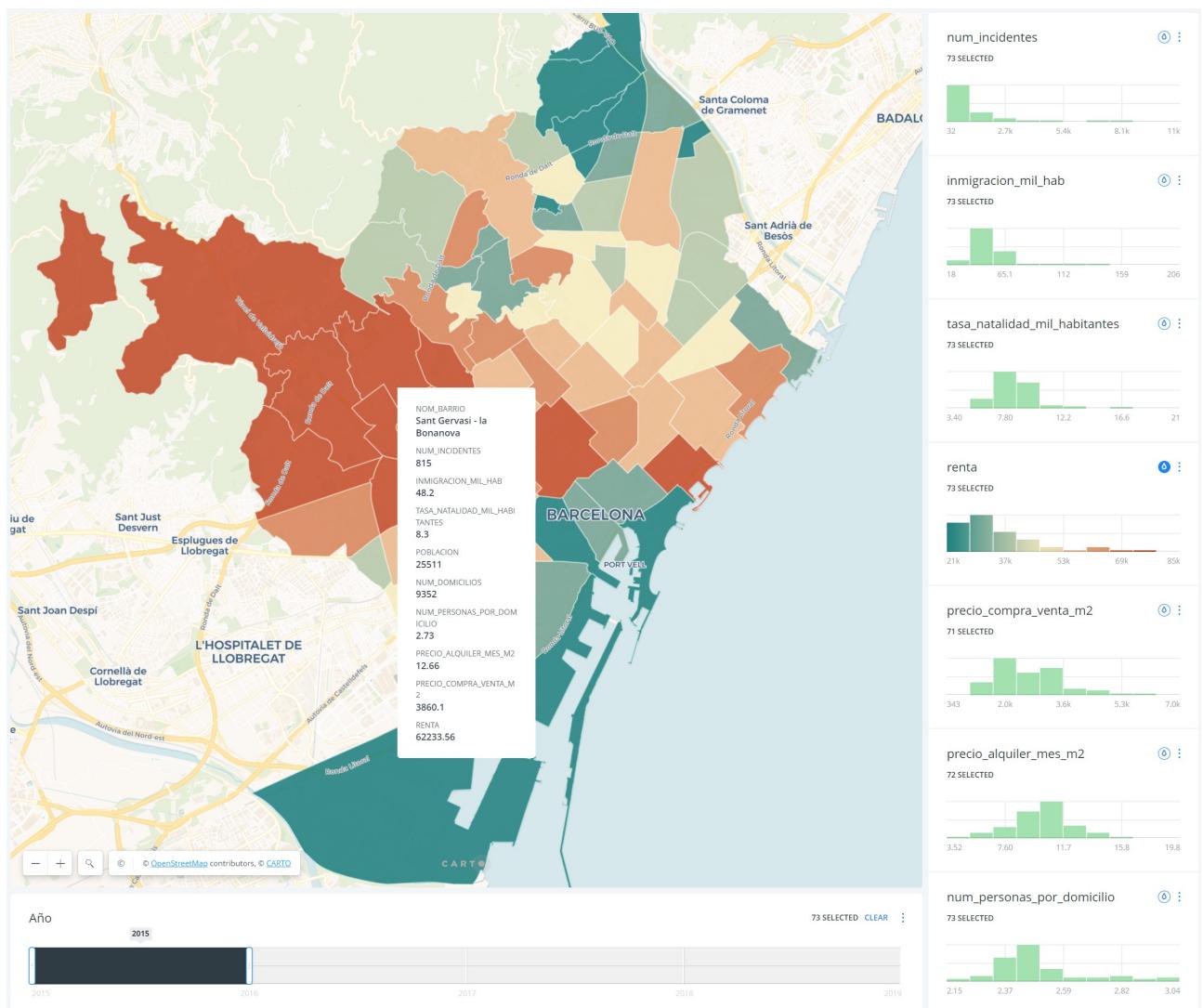


Figura 3.10: Visualización de los indicadores relevantes por barrio, filtrando por renta del año 2015. Fuente: elaboración propia. Accesible en [Carto](#).

Esta segunda visualización está centrada en mostrar también en forma de mapa de coropletas los valores de las distintas variables que pueden estar relacionadas con la gentrificación, con el objetivo de identificar las más relacionadas con este fenómeno. El aspecto de este mapa es muy parecido al anterior, solo que este no hace foco en los lugares sino en los demás indicadores, por lo que no muestra ningún punto en el mapa. Este puede ser consultado en la Figura 3.10, o clicar como antes en el enlace a Carto del pie de la imagen para abrir la visualización interactiva en la web. En el panel derecho, se puede visualizar la distribución de cada una de estas variables, así como seleccionar cualquiera de ellas para que se represente

en el mapa de coropletas según el valor en cada barrio. Para hacer dicha acción, hay que pulsar en ícono de *Apply Auto Style* (símbolo de gota en el extremo superior derecho de la caja de la variable) e inmediatamente se repinta el mapa para mostrar los valores por barrio de la variable escogida.

Como antes, se puede filtrar por año en el panel inferior. En este caso, esta funcionalidad tiene mucho más sentido, pues se puede ver con más facilidad la evolución de cada variable a lo largo de los cuatro años del estudio, pudiendo ver mejor si la evolución de alguna variable está relacionada con el proceso de gentrificación.

De igual forma que antes, si se pasa el ratón por encima de un barrio se muestra toda la información del barrio, por lo que a parte del nombre se muestra también el valor de todos los indicadores.

En el apéndice [A.3](#) se pueden ver todos los mapas de coropletas para todas las variables y todos los años, además de una serie de gráficos de líneas que muestran la evolución de cada variable por años, por lo que con la información de ambas visualizaciones se podrá analizar la evolución de todas las variables y detectar posibles barrios gentrificados o en proceso de gentrificación.

3.3. Análisis y resultados del número de ubicaciones relevantes y evolución temporal de indicadores

3.3.1. Análisis de ubicaciones por barrio

En esta primera parte de análisis, se van a estudiar los distintos tipos de ubicaciones que hay en cada barrio (usando los datos completos de 2019) con el fin de intuir cuáles son los barrios que parecen estar gentrificados. Para ello, el análisis se basará en todas las visualizaciones generadas con estos datos, las cuales son los diagramas de radar y mapas de coropletas que se pueden encontrar en los apéndices [A.1](#) y [A.2](#).

Para recordar qué se busca en este análisis, los barrios gentrificados se caracterizan por tener un gran número de restaurantes, lugares de ocio y cultura, y tiendas de ropa, centros comerciales y alojamientos. Por tanto, los barrios que destaque por tener un número elevado de este tipo de ubicaciones son los que tienen más posibilidades de estar gentrificados.

Si se empieza estudiando el número de restaurantes, se puede comprobar que estos están focalizados en la zona más céntrica de Barcelona, lo que se traduce en los barrios de La

Dreta de l'Eixample, L'Antiga Esquerra de l'Eixample, La Vila de Gràcia, El Raval, La Nova Esquerra de l'Eixample, El Barri Gòtic, Sant Gervasi - Galvany y Sant Antoni.

Si se comprueba ahora el número de lugares de ocio y cultura (cines, teatros, discotecas, etc.), se puede ver que están localizados aproximadamente en la misma zona que los restaurantes, es decir, de nuevo la zona centro de Barcelona. En este caso, hay menos diferencias entre barrios que con los restaurantes, pero de nuevo se localizan mayoritariamente en los barrios de La Dreta de l'Eixample, La Vila de Gràcia, L'Antiga Esquerra de l'Eixample, El Raval, El Barri Gòtic, Sant Pere y El Poble Sec.

Seguidamente, viendo el número de alojamientos, volvemos al mismo caso de concentración en la zona centro de la ciudad, destacando con diferencia el barrio de La Dreta de l'Eixample en cuanto al número de alojamientos. Por detrás, pero con una cantidad importante también, se sitúan los barrios de El Raval, El Barri Gòtic y l'Antiga Esquerra de l'Eixample.

Por último, si se analiza el número de tiendas de ropa y centros comerciales, se ve de nuevo que los barrios con más ubicaciones de este tipo se encuentra en el centro, aunque en este caso se pueden observar barrios periféricos que también disponen de un gran número de tiendas y centros comerciales, como Sant Andreu, Bon Pastor, Horta, Sant Gervasi - La Bonanova o Sarriá, aunque los que más tienen son de nuevo La Dreta de l'Eixample, Sant Gervasi - Galvany, El Barri Gòtic, La Vila de Gràcia y Sant Pere.

No obstante, solo con el número de locales de estas tres categorías, no se puede afirmar que ningún barrio de los mencionados esté gentrificado. Para tener un poco más de contexto, se debe comparar la cantidad de estos lugares respecto a la del resto y a la media de Barcelona. Con este objetivo están las visualizaciones del apéndice A.1, donde se puede observar para cada barrio un diagrama de radar que muestra las proporciones de ubicaciones de cada tipo junto a la media de Barcelona, por lo que se puede obtener una visión más amplia y acertada sobre los barrios que parecen estar gentrificados o en proceso de estarlo.

Por tanto, lo que se busca ahora en los diagramas, es encontrar barrios que estén por encima de la media en número de restaurantes, lugares de ocio y cultura, tiendas de ropa y centros comerciales, y alojamientos.

El primer barrio que destaca por cumplir estas condiciones es el **Barri Gòtic**, donde estas variables están por encima de la media de Barcelona, especialmente en tiendas de ropa y centros comerciales. Sin embargo, la cantidad de negocios tradicionales es también superior a la media, variable que indicaría una no gentrificación, pero que sería muy interesante tener datos históricos para saber si la cantidad de este tipo de negocios se ha reducido con el paso de los años e indicar en este caso la presencia de un proceso de gentrificación. Otro barrio ya

mencionado que también podría estar gentrificado es **la Dreta de l'Eixample**. En este caso no se observan diferencias extremadamente grandes respecto a la media pero sí se ve que las variables indicadoras están algo por encima de la media de la ciudad, sobre todo también en tiendas de ropa. Muy parecido a este diagrama es el del barrio de **Sant Pere**, por lo que se sugiere que tiene un comportamiento similar. Por tanto, basándose solo en el número de locales se podría decir que estos tres barrios están gentrificados.



Figura 3.11: Diagramas de radar con los barrios que parecen estar más gentrificados según el número de ubicaciones de distintos tipos del año 2019. Fuente: Elaboración propia.

Analizando otros barrios de los mencionados anteriormente, se puede observar que están en general sobre la media y que destacan ligeramente en solo algún tipo de ubicación, como sucede con los barrios de **la Vila de Gràcia** y **Sant Gervasi - Galvany**, donde solo destacan por sus tiendas de ropa y centros comerciales o **l'Antiga Esquerra de l'Eixample** por su número de restaurantes. Es algo más curiosa la situación del barrio del **Poble Sec**, ya que destaca en restaurantes y especialmente en lugares de ocio y cultura, pero está bastante por debajo de la media en cuanto a tiendas de ropa. Otros de los barrios que parece mostrar un pequeño signo de gentrificación es **el Raval**, que pese a estar muy cerca de la media en la mayoría de ubicaciones, los alojamientos y los restaurantes son ligeramente superiores a la media de la ciudad. En resumen, los barrios mencionados en este párrafo, podrían estar en una fase inicial de gentrificación, siempre recordando que se está basando esta conclusión en el número de negocios de diferentes tipos en solo el año 2019 sin poder analizar la evolución años atrás.

3.3. Análisis y resultados del número de ubicaciones relevantes y evolución temporal de indicadores

49

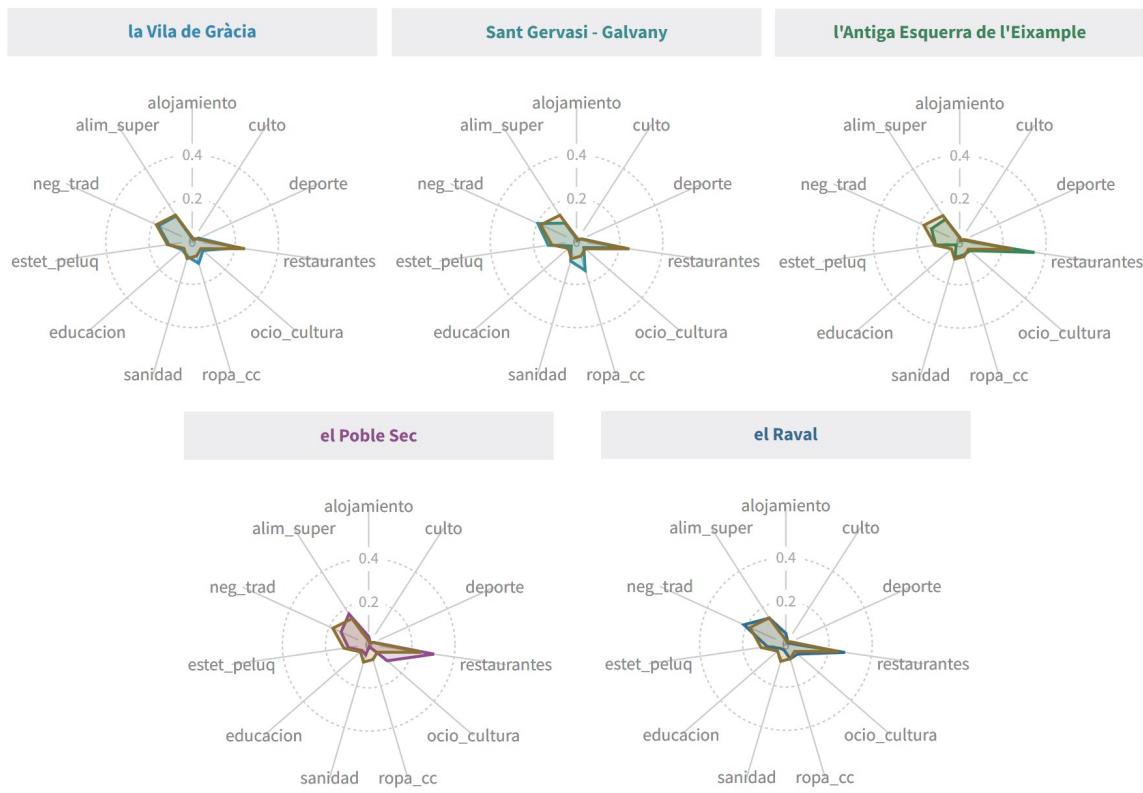


Figura 3.12: Diagramas de radar con los barrios que parecen estar en una fase inicial de gentrificación según el número de ubicaciones de distintos tipos del año 2019. Fuente: Elaboración propia.

También, se pueden utilizar los diagramas para ver qué barrios están en una situación estable y por tanto están en bajo de riesgo de sufrir la gentrificación. Este tipo de barrios tiene un diagrama lo más parecido a la media de Barcelona. Por ejemplo, estarían en esta situación los barrios de la Nova Dreta de l'Eixample, Sants, o el Clot, entre algunos otros.

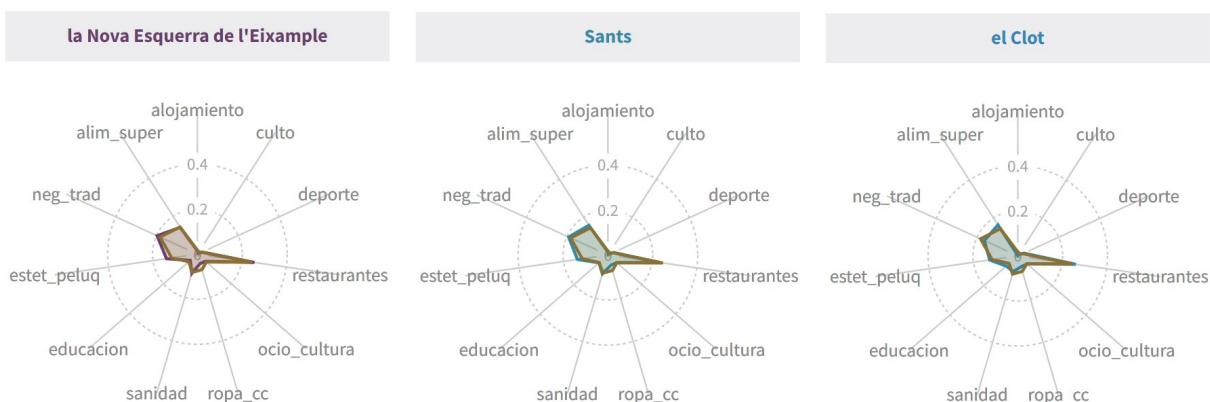


Figura 3.13: Diagramas de radar con los barrios que parecen estar en situación estable según el número de ubicaciones de distintos tipos del año 2019. Fuente: Elaboración propia.

También se puede encontrar el caso opuesto, o barrios que no están en ningún riesgo de ser gentrificados. Esto en cuanto al número de ubicaciones se traduciría como barrios con número de locales que indican gentrificación por debajo de la media y número de locales que no indican gentrificación (tiendas de alimentación, centros educativos y de sanidad, etc.) por encima de la media. Por ejemplo, los barrios de Vallbona, el Coll o Vallcarca i els Penitents pertenecen a este grupo de barrios.

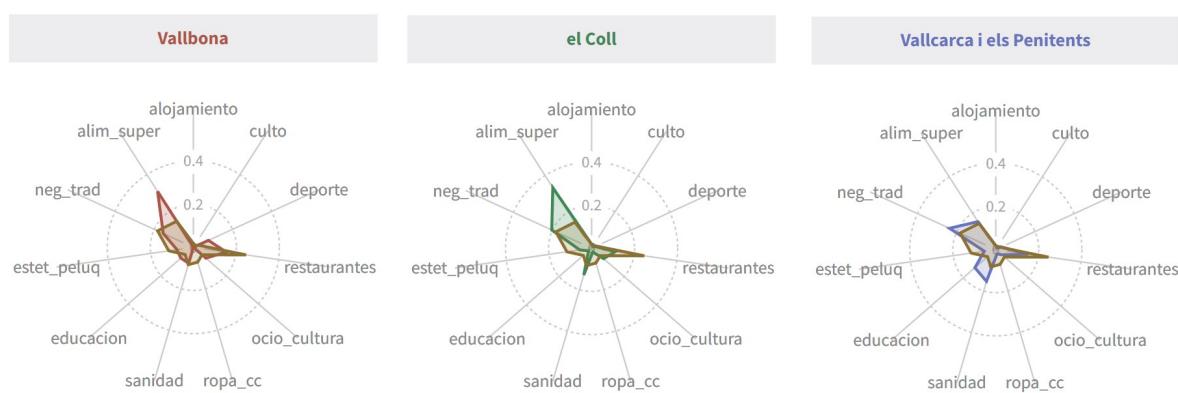


Figura 3.14: Diagramas de radar con los barrios que no tienen ningún riesgo de ser gentrificados según el número de ubicaciones de distintos tipos del año 2019. Fuente: Elaboración propia.

3.3.2. Análisis de la evolución de los indicadores por barrio

Una vez ya se han analizado los tipos de ubicaciones y se han obtenido algunos barrios sospechosos de estar gentrificados, ahora se va a proceder a analizar la evolución del resto de indicadores que pueden estar relacionados con la gentrificación. Estos análisis se harán a partir del estudio de los mapas de coropletas y gráficos de líneas disponibles para cada variable en el apéndice A.3.

3.3.2.1. Renta

La renta es un indicador relevante para determinar la posible gentrificación de un barrio. Normalmente, en un barrio gentrificado, la renta del barrio no suele ser ni muy baja ni muy alta, si no que se debe situar aproximadamente entre el 75 % y el 125 % respecto de la media de Barcelona en el caso general. Además, en los barrios gentrificados la renta no suele apenas evolucionar, puesto que una característica de este proceso es justamente el incremento del nivel de vida del barrio. Por tanto, la renta por si sola no puede dar pistas sobre la gentrificación, pero sí si se combina con el resto de variables, ya que por ejemplo

si el alquiler aumentase considerablemente en un barrio determinado, se podría comprobar cómo ha evolucionado la renta. Si esta aumentase proporcionalmente a la subida de los alquileres, seguramente no se trataría de un caso de gentrificación. En caso positivo, podría ser que sí lo sea, aunque tampoco se puede confirmar con certeza ya que esta situación podría haberse dado por otro motivo.

Si se visualizan los gráficos (figuras A.5 y A.6), se puede observar que la mayoría de barrios se mantienen estables a lo largo de los cuatro años, a excepción de Pedralbes, Vallvidrera, Sarrià, les Tres Torres, Sant Gervasi - la Bonanova y Sant Gervasi - Galvany que se ve más claramente un aumento de la renta y son los menos sospechosos de estar gentrificados.

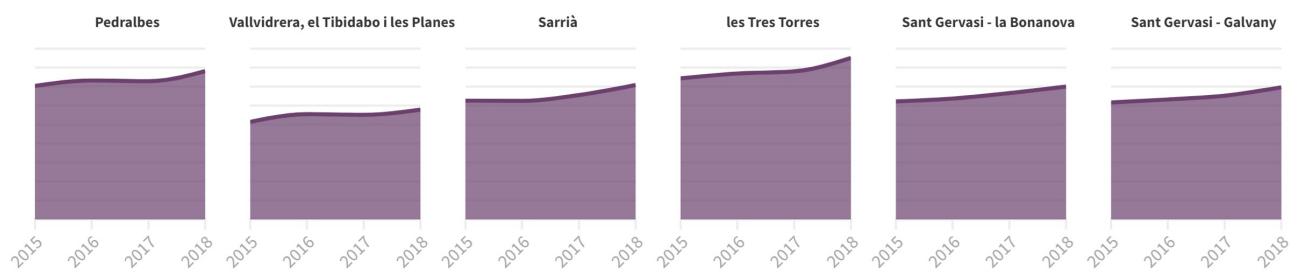


Figura 3.15: Gráficos de líneas de los barrios con más evolución de la renta. Fuente: Elaboración propia.

3.3.2.2. Precio de compra-venta por metro cuadrado

El precio de compra-venta de las viviendas, es una de las variables más importantes para comprobar si el nivel de vida de un barrio es alto o no. En este caso, lo que se busca es encontrar barrios en los que durante los años analizados el precio haya aumentado. Como se puede ver en los gráficos (figuras A.7 y A.8), la mayoría de los barrios tienen una evolución al alza, por lo que lo más interesante es ver cuáles de ellos han variado más en los 4 años que se comparan. La siguiente tabla y figura muestran justamente los diez barrios donde el crecimiento ha sido mayor. Como se puede ver, aparecen barrios que ya fueron mencionados previamente, como el Poble Sec, el Barri Gòtic o el Raval, pero aparecen otros como les Roquetes o Verdun. Sin embargo, cabe destacar que estos últimos dos eran en 2015 de los más baratos de la ciudad, por lo que es normal que con los años el precio haya subido. Destacan especialmente los barrios de Navas y el Putxet i el Farró, con incrementos muy notables y con una evolución parecida a la de los barrios de el Barri Gòtic o Poble Sec, los cuales ya resultaron ser sospechosos de estar gentrificados anteriormente.

Barrio	Evolución
les Roquetes	+130.26 %
Sants - Badal	+81.82 %
el Poble Sec	+75.46 %
el Putxet i el Farró	+69.26 %
Verdun	+68.3 %
Navas	+66.62 %
el Raval	+65.28 %
el Barri Gòtic	+61.67 %
el Coll	+56.35 %
Can Baró	+55.44 %

Cuadro 3.3: Barrios con más crecimiento del precio de compra-venta de 2015 a 2018.

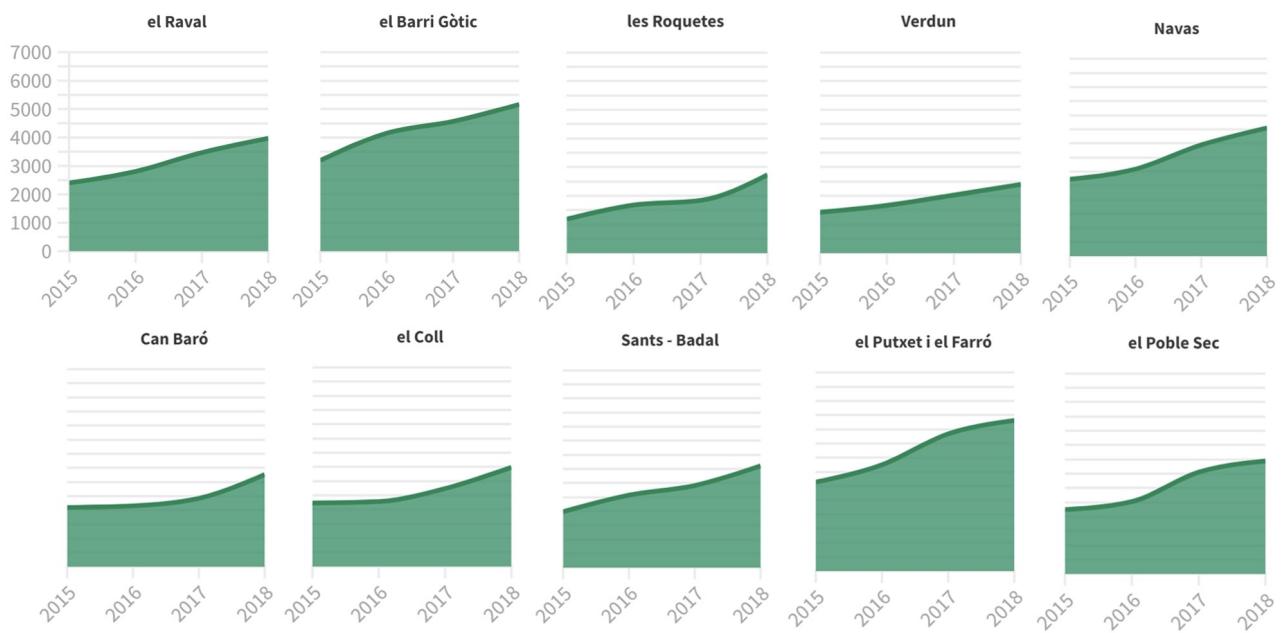


Figura 3.16: Gráficos de líneas de los barrios con más crecimiento del precio de compra-venta de viviendas de 2015 a 2018. Fuente: Elaboración propia.

3.3.2.3. Precio del alquiler al mes por metro cuadrado

El precio de los alquileres a priori es otro indicador importante para detectar la gentrificación de los barrios al estar relacionado con el precio de compra-venta. No obstante, a diferencia de esta variable, el precio del alquiler tiene un comportamiento bastante más parecido entre barrios, siendo de forma mayoritariamente creciente a lo largo de los años analizados pero

claramente de forma más lineal y suave (figuras A.9 y A.10). En este caso, el grueso de los aumentos de alquileres se sitúa aproximadamente entre el 20 y 35 % de subida, destacando solamente por encima el barrio de Baró de Viver con un crecimiento del 55 %.

Barrio	Evolución
Baró de Viver	+55.42 %
Provençals del Poblenou	+36.6 %
la Trinitat Nova	+36.02 %
el Turó de la Peira	+34.64 %
Ciutat Meridiana	+34.36 %
el Besòs i el Maresme	+32.53 %
Canyelles	+28.67 %
les Roquetes	+28.6 %
la Marina de Port	+28.55 %
Diagonal Mar i el Front Marítim del Poblenou	+28.26 %

Cuadro 3.4: Barrios con más crecimiento del precio del alquiler de viviendas de 2015 a 2018.

No obstante, si se ve la evolución completa del barrio de **Baró de Viver**, se puede ver cómo en el año 2016 se produjo una bajada importante, y fue solo de 2017 a 2018 cuando se produjo la gran subida que provoca que esté el primero de la tabla, tal como se ve en la siguiente gráfica. Sin embargo, la subida del alquiler en este barrio no es coherente con la subida del precio de compra-venta, la cual es solo de un 16 % contra el 55 % del alquiler. De todas formas, será necesario analizar este barrio desde la perspectiva de otras variables para ver si es posible que esté gentrificado o no.

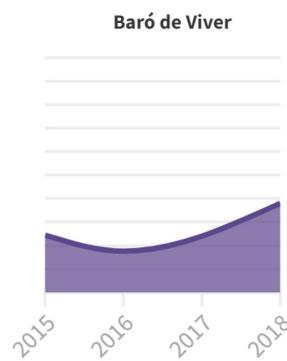


Figura 3.17: Gráfico de líneas de la evolución del precio del alquiler del barrio de Baró de Viver de 2015 a 2018. Fuente: Elaboración propia.

Sin embargo, parece que esta variable no ayuda demasiado para distinguir los barrios como

sí lo hacía el precio de compra-venta, puesto el comportamiento es muy parecido en casi todos los barrios.

3.3.2.4. Número de personas por domicilio

Según el estado del arte, los barrios gentrificados se caracterizan por ser un lugar de residencia para solteros o parejas que empiezan a asentarse y buscan un primer hijo. Por tanto, el número medio de personas por domicilio de un barrio puede ser una variable interesante para controlar este factor. No obstante, si se observa en el apéndice la página con todos los barrios (figura A.11), no se ve prácticamente evolución ni positiva ni negativa a excepción del Barri Gòtic, que aumenta casi un 10 %. Que este barrio aumente el número de personas por domicilio puede ser debido al hecho de que se encuentra en una fase avanzada de gentrificación, por lo que es probable que sea debido a que muchas de estas parejas ya han tenido su primer hijo. Esta suposición luego se reafirmará al analizar las tasas de natalidad de los barrios. Dejando este caso un poco aparte, lo que se busca aquí por tanto son barrios con un número de personas por domicilio bajo. En el siguiente mapa de coropletas se muestran precisamente en color verde oscuro los barrios con menor número de personas por domicilio para 2018. Los mapas para el resto de años pueden encontrarse en la figura A.12.

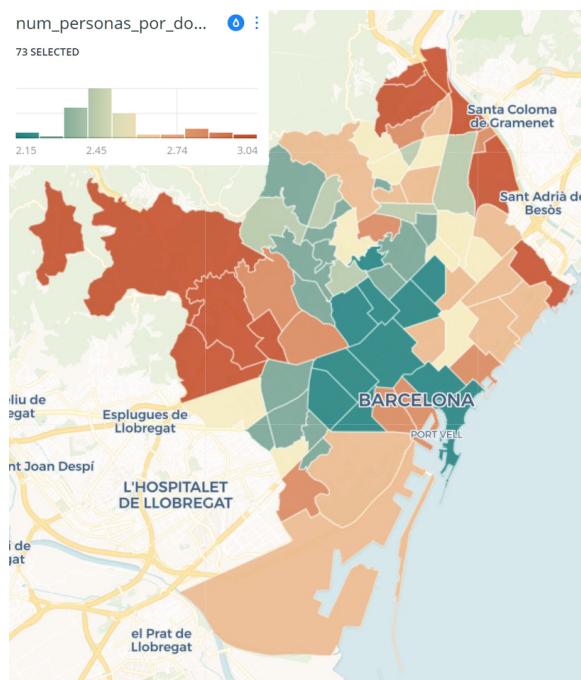


Figura 3.18: Mapa de coropletas con el número de personas por domicilio en 2018. Fuente: Elaboración propia.

Estos barrios con un valor menor, son la **Barceloneta** (2.15 habitantes/domicilio), **Sant Pere** (2.21), **La Vila de Gràcia** (2.21) y de los que en algún momento han aparecido como posibles barrios gentrificados están **La Nova Esquerra de l'Eixample**, **l'Antiga Esquerra de l'Eixample**, **la Dreta de l'Eixample** y **Sant Antoni**.

3.3.2.5. Número de incidentes

Cuando un barrio está en proceso de gentrificación, se suelen realizar políticas de mejora del barrio, siendo la reducción de la criminalidad una de ellas. Se ha querido medir si este hecho se ha producido en Barcelona mediante la evolución del número de incidentes relacionados con el crimen y violencia. En este caso, se han filtrado los barrios que tienen de media más de 1000 incidentes, puesto que es posible que barrios con número de incidentes casi despreciables puedan desvirtuar los resultados con porcentajes de crecimiento extremadamente altos.

En este caso, se han seleccionado los barrios que tienen una variación de más del 20 %, ya que la tendencia de los incidentes es que suban a lo largo de los años.

Barrio	Evolución
Sant Antoni	+46.9 %
Hostafrancs	+37.64 %
el Raval	+33.0 %
Sant Pere, Santa Caterina i la Ribera	+29.43 %
Diagonal Mar i el Front Marítim del Poblenou	-27.33 %
la Sagrada Família	+27.1 %
el Parc i la Llacuna del Poblenou	+26.73 %
el Fort Pienc	+20.53 %

Cuadro 3.5: Barrios con más evolución del número de incidentes de 2015 a 2018.

Como se puede observar en esta tabla, solo el barrio de Diagonal Mar i el Front Marítim del Poblenou reduce de una forma considerable el número de incidentes, aunque no es demasiado significativo al tener pocos incidentes de partida (unos 1000). No obstante, de los barrios que han ido saliendo como posibles candidatos a ser gentrificados, solo salen en esta lista el Raval (número de incidentes muy elevado y a la alza) y Sant Pere, Santa Caterina i la Ribera. Sin embargo, es interesante de ver que el resto de barrios que parecen estar también en fase de gentrificación según los análisis anteriores (Barri Gòtic, la Dreta de l'Eixample, l'Antiga Esquerra de l'Eixample, Poble Sec, Vila de Gràcia), no aparecen entre los que más

ha evolucionado el número de incidentes, si no que el número se ha mantenido bastante estable o ha subido solo ligeramente. Este hecho puede significar un intento de reducir la criminalidad en estos barrios y por tanto refuerza la teoría de que puedan estar gentrificados. A continuación, se muestran los gráficos de evolución para estos barrios candidatos a estar gentrificados que han ido saliendo durante los diferentes análisis. Las evoluciones de todos los barrios pueden encontrarse en las figuras [A.13](#) y [A.14](#).

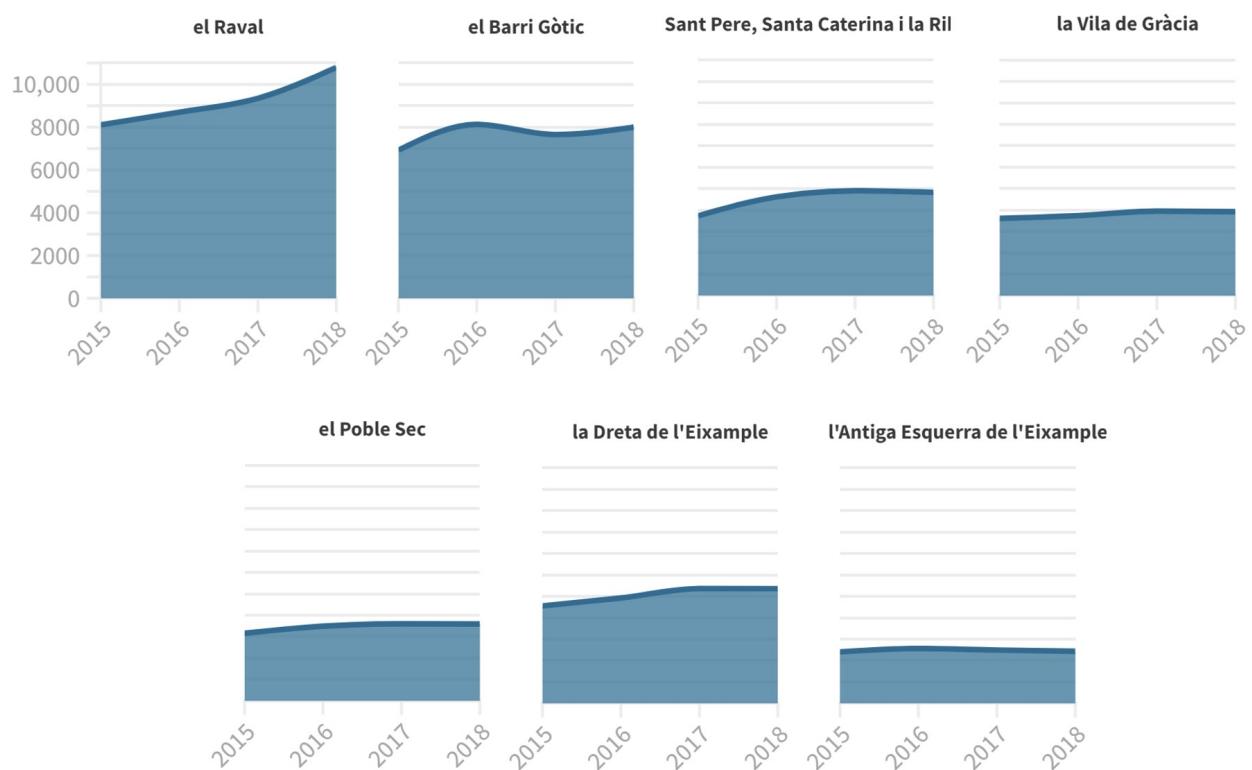


Figura 3.19: Gráficos de líneas de la evolución del número de incidentes de los barrios sospechosos de estar gentrificados según análisis previos. Fuente: Elaboración propia.

3.3.2.6. Tasa de inmigración

Los barrios gentrificados se suelen caracterizar por atraer a gente de distintas culturas y nacionalidades, por lo que se busca en este punto son barrios que hayan sufrido un aumento relevante de inmigración durante los años del análisis. En la siguiente tabla, se muestran los diez barrios que más variación han sufrido en cuanto a la tasa de inmigración. De nuevo, en las figuras [A.15](#) y [A.16](#) se puede ver las visualizaciones con las evoluciones de todos los barrios.

Barrio	Evolución
Porta	+72.63 %
Sant Genís dels Agudells	+67.31 %
Horta	+50.78 %
el Barri Gòtic	+45.09 %
la Trinitat Nova	+43.22 %
el Besòs i el Maresme	+38.43 %
Montbau	+35.40 %
Can Peguera	+34.72 %
el Bon Pastor	+32.89 %
el Congrés i els Indians	+32.83 %

Cuadro 3.6: Barrios con más evolución de tasa de inmigración de 2015 a 2018.

Observando la lista, el primer barrio que llama la atención es Porta, ya que no había sido mencionado hasta ahora y sorprende por su incremento en inmigración. Sin embargo, no acaba de ser sospechoso de estar gentrificado, puesto que se produjo un decremento del precio de compra-venta de 2017 a 2018, por lo que este incremento de inmigración podría darse por otros factores sociales distintos. En las siguientes posiciones, se encuentran los barrios de Sant Genís dels Agudells y Horta, los cuales tampoco habían sido mencionados. A pesar de ello, ninguna otra variable ni el análisis de ubicaciones han mencionado estos barrios como candidatos a estar gentrificados, por lo que el motivo de este aumento podría estar relacionado con otro factor externo. No obstante, **el Barri Gòtic** pese a tener un incremento algo menor, en valor absoluto es un crecimiento muy grande y reafirma de nuevo la gentrificación de este barrio. El resto de barrios de la tabla, de bastante menos importancia y de la periferia de la ciudad, ya se acercan a cifras más razonables de aumentos, que se pueden entender por la localización de estos barrios y al no tener un coste de vida tan elevado.

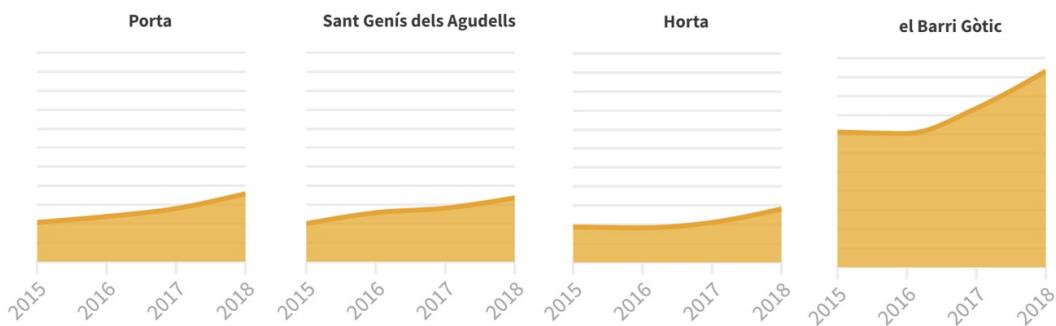


Figura 3.20: Gráficos de líneas con los barrios con mayor evolución en la tasa de inmigración. Fuente: Elaboración propia.

3.3.2.7. Tasa de natalidad

La tasa de natalidad puede dar pistas sobre la gentrificación de un barrio según su tendencia, aunque no de forma clara ya que las variaciones pueden darse por muchos motivos. Un decremento de la tasa de natalidad puede verse como una migración hacia el barrio de solteros o parejas gentrificadoras. Pero un aumento puede significar también que estas parejas comienzan a tener sus primeros hijos, lo que indicaría a la vez un estado avanzado de gentrificación.

Sin embargo, observando todos los gráficos de la Figura A.18 no se puede llegar a ninguna conclusión clara, ya que hay muy pocas variaciones entre los años 2015 y 2018 y sin seguir ningún patrón entre barrios sospechosos de estar gentrificados y los que no. Por ejemplo, a continuación se muestran los gráficos para los barrios sospechosos.



Figura 3.21: Gráficos de líneas de la evolución de la tasa de natalidad de los barrios sospechosos de estar gentrificados según análisis previos. Fuente: Elaboración propia.

De estos barrios mostrados, los únicos que aumentan son el Barri Gòtic y la Vila de Gràcia, aunque este último solo muy ligeramente y entre 2017 y 2018. El caso del Barri Gòtic puede explicarse por su fase avanzada de gentrificación, tal como se ha comentado anteriormente.

3.3.3. Conclusiones del análisis y barrios gentrificados

Una vez se ha efectuado el análisis de todos los barrios tanto de las ubicaciones y negocios que se pueden encontrar como de la evolución de indicadores, es necesario resumir los puntos claves que se han ido encontrando, además de concretar cuáles han sido los barrios detectados como gentrificados y en qué nivel están según estos análisis realizados.

Empezando por el análisis de ubicaciones, se ha visto que las variables más relevantes para descubrir barrios en gentrificación son la abundancia en restaurantes, en tiendas de ropa y centros comerciales, y en alojamientos, así como la escasez de tiendas de alimentación y supermercados. Algo menos relevante pero también con cierta importancia es la abundancia de lugares de ocio y cultura.

En cuanto al estudio del resto de indicadores y su evolución entre 2015 y 2018, se ha visto que no es suficiente analizar una sola variable para detectar la gentrificación, si no que es necesario combinar más de una. En este caso, se ha observado que la combinación de va-

riables que indican barrios sospechosos de gentrificación son incrementos en los precios de compra-venta de vivienda, número bajo de personas por domicilio, estabilidad o incremento suave de incidentes y tasa de inmigración creciente, así como estabilidad de la renta.

Por tanto, para decidir qué barrios están gentrificados es necesario combinar los aprendizajes de los análisis de las ubicaciones con los de los indicadores. De esta forma, a continuación se muestra una lista de los barrios que según estos criterios pueden estar gentrificados, así como los motivos que han llevado a tomar esta decisión.

- **el Barri Gòtic:** Claramente un barrio gentrificado en una fase avanzada. En cuanto a ubicaciones destaca por su cantidad de tiendas de ropa y centros comerciales, restaurantes y alojamientos, así como por su escasez en tiendas de alimentación o supermercados. Su precio de compra-venta ha incrementado hasta convertirse en un barrio donde este valor es más caro y es con diferencia el barrio en el que ha habido más incremento de la inmigración, siendo en 2018 claramente en el barrio con más inmigrantes. Además, es el que más incrementó el número medio de personas por domicilio.
- **Sant Pere, Santa Caterina i la Ribera:** Este barrio vuelve a destacar como el anterior por las ubicaciones que tiene. Dispone de un gran número de restaurantes, tiendas de ropa y lugares de ocio y cultura pero también pocos supermercados y tiendas de alimentación. En este caso por alojamiento no destaca, pues se sitúa alrededor de la media de Barcelona. Además, en cuanto a los indicadores analizados, su precio de compra-venta evolucionó al alza de forma clara y sufrió un incremento relevante de la tasa de inmigración. Además, es el segundo barrio de la capital catalana con el menor número medio de habitantes por domicilio. Por el resto de variables, no destaca tanto, pero están todas en línea con lo esperado en un barrio gentrificado. Por tanto, se puede hablar de un barrio que está gentrificado.
- **la Dreta de l'Eixample:** De forma muy parecida al Barri Gòtic en cuanto a lugares, destaca igual por su número de tiendas de ropa, restaurantes y alojamientos y por sus pocas tiendas de alimentación y supermercados. No obstante, por indicadores no se ve tan claramente la gentrificación. Solo destaca por su bajo número de personas por domicilio, siendo de los más bajos de Barcelona. El resto de indicadores presenta unos valores que encajan con los supuestos de gentrificación, aunque de forma poco notable. Sin embargo, el precio de compra-venta sufrió una ligera bajada de 2017 a 2018, cosa que sería el único factor que puede contradecir el concepto de gentrificación. Sin embargo, se puede decir también está gentrificado, aunque no de forma tan clara como el Barri Gòtic o Sant Pere.

- **el Poble Sec:** Es otro de los barrios que destaca por el número de ubicaciones de distintos tipos. Tiene más restaurantes y lugares de ocio y cultura que la media, pese a tener menos tiendas de ropa que la media, aunque en este caso el número de supermercados está algo por encima de la media, mientras que en los casos anteriores estaba notablemente por debajo. En cuanto al resto de indicadores, es uno de los barrios con más evolución al alza del precio de compra-venta. Por el resto de variables no destaca demasiado, pero todos sus valores o tendencias van acorde a las de un barrio gentrificado. Igual que con la Dreta de l'Eixample, no se puede decir que esté tan gentrificado como el Barri Gòtic.
- **el Raval:** Por ubicaciones no destaca demasiado, ya que todos sus valores son muy cercanos a la media. No obstante, destaca un poco en alojamientos y restaurantes, estando por encima de la media en estos dos tipos de lugares. En cuanto al resto de variables, siguen la mayoría la tendencia general de un barrio gentrificado, pese a que el número de incidentes ha incrementado de forma notable. Sin embargo, pese a que el número medio de personas por habitante es elevado, ha ido decrementando con más velocidad que el resto de barrios. Por tanto, se puede considerar que este barrio está en una fase inicial de gentrificación.
- **l'Antiga Esquerra de l'Eixample:** En ubicaciones tiene valores muy cercanos a la media, pero destaca notablemente en el número de restaurantes, siendo bastante superior a la media. Es además uno de los barrios con menos habitantes por domicilio. En el resto de indicadores no destaca, pero de nuevo todos siguen las tendencias propias de un barrio gentrificado.

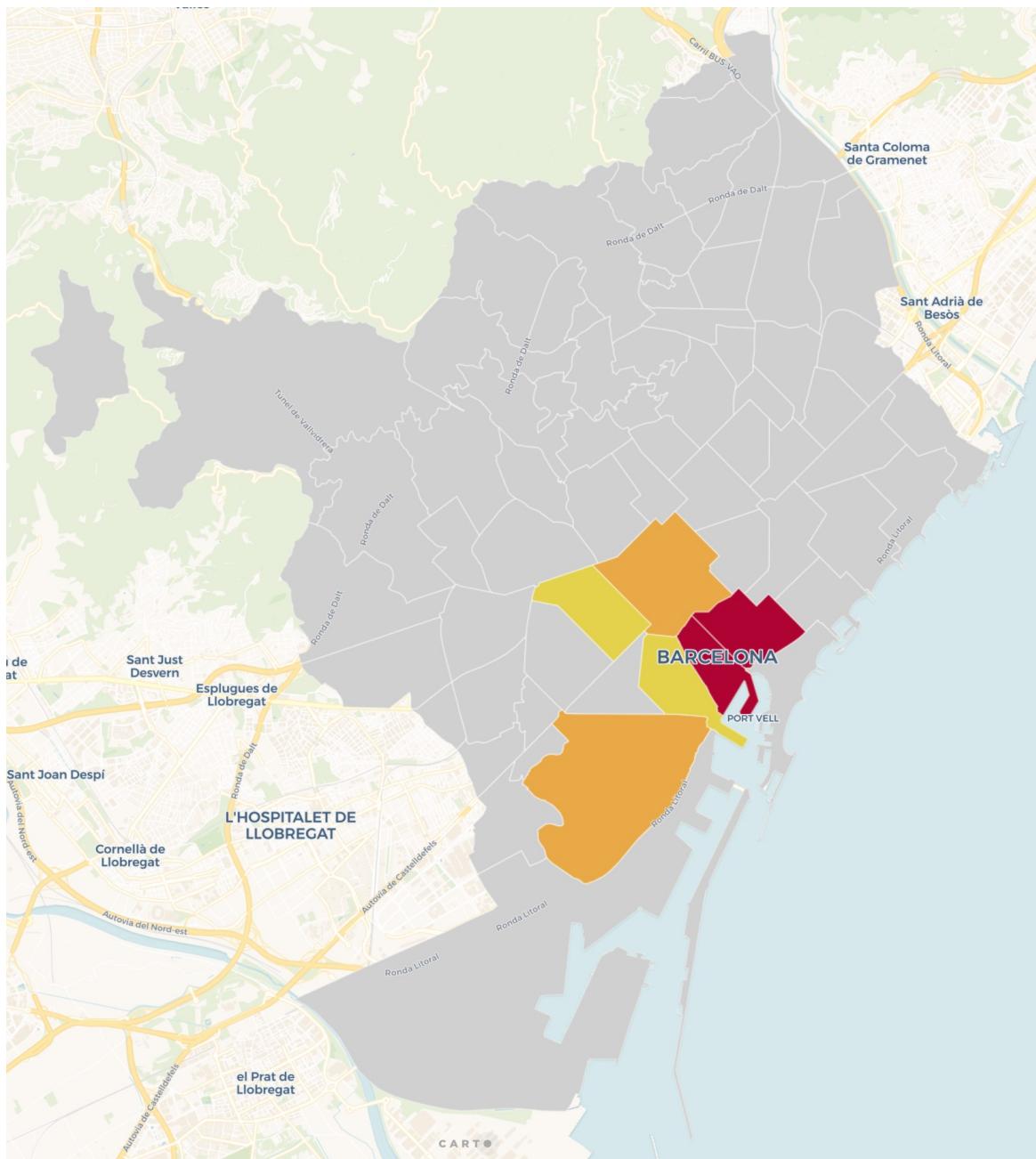


Figura 3.22: Mapa con los barrios gentrificados según el grado de gentrificación a partir de los análisis de ubicaciones y el resto de indicadores. De color amarillo se muestran los barrios en una fase inicial de gentrificación (el Raval y l'Antiga Esquerra de l'Eixample), en naranja los que están en una fase intermedia de gentrificación (la Dreta de l'Eixample y el Poble Sec) y en rojo los que están en una fase avanzada (Sant Pere y el Barri Gòtic - muy gentrificado) Fuente: Elaboración propia.

3.4. *k-means* para clasificar el grado de gentrificación

En esta última parte del proyecto, se intenta construir un modelo de aprendizaje automático no supervisado que permita clasificar los distintos barrios en diferentes grupos según el nivel de gentrificación que tenga cada uno de estos barrios. El algoritmo escogido por su simplicidad es *k-means*, aunque según el estado del arte el algoritmo de agrupación jerárquica da también unos resultados muy parecidos.

El primer paso será realizar una transformación de los datos para prepararlos para el algoritmo. No obstante, la estructura actual del *dataset* es una fila por barrio por cada año entre 2015 y 2018. Como lo que realmente interesa son las variaciones de las variables, se han transformado los datos para que el valor de cada variable sea la diferencia de los valores entre 2015 y 2018. Además, como variables solo se van a coger los indicadores que no están relacionados con los lugares, puesto que el *dataset* completo de ubicaciones no está historificado y por tanto no daría una visión temporal de cómo ha evolucionado cada barrio. Por tanto, el *clustering* se realizará utilizando las siguientes variables: número de incidentes, tasa de inmigración, tasa de natalidad, número de personas por domicilio, precio del alquiler, precio de compra-venta y renta.

Como las magnitudes de las variables son diferentes, antes de nada es necesario normalizar las variables para que estén todas en el mismo rango de magnitud. Una vez escaladas las variables, se buscan las correlaciones entre variables, con el fin de eliminar alguna que esté muy correlacionada con otra. Al ver que la variable de tasa de inmigración tiene una correlación moderada con el número de personas por domicilio (0.53), además de dar resultados algo peores en el *clustering*, se decide eliminar esta primera variable de las que se usarán en el modelo final. También, se ha intentado hacer una reducción de dimensiones con PCA, pero los componentes principales no conseguían explicar demasiada varianza. El componente principal solo lo hacía en un 20 % y se necesitaban entre 5 y 6 dimensiones para explicar una varianza mayor del 80 %, por lo que al final se ha decidido no reducir las dimensiones.

El siguiente paso es escoger el número óptimo de *clusters*, teniendo en cuenta que se quiere obtener grupos con distintos niveles de gentrificación. Para elegir este valor, se utilizan los valores de la silueta y método del codo.

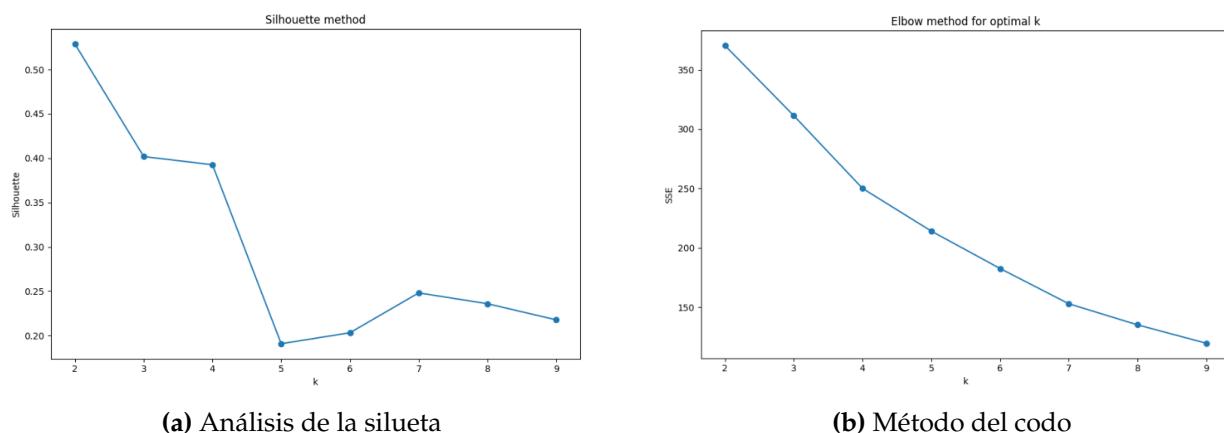


Figura 3.23: Selección de la k óptima con el análisis de la silueta y método del codo. Fuente: Elaboración propia.

Viendo los resultados de ambos métodos en los gráficos anteriores, se ve que con el método de la silueta el valor más grande se consigue claramente con $k = 2$. Sin embargo, al no querer buscar solo dos grupos, se busca el siguiente valor, el cual pertenece a $k = 3$, teniendo en cuenta que el valor para $k = 4$ es solo ligeramente inferior. Con el método del codo, pese a que no es fácil ver con claridad cual es el valor óptimo, parece que está en $k = 4$. Por tanto, después de observar ambos métodos, se ejecuta el algoritmo *k-means* con $k = 3$ y $k = 4$. Una vez etiquetado cada barrio, se genera un fichero *geojson* con el resultado del *clustering* para ambas *k*'s y la geometría de los barrios y se suben a Carto para crear los mapas y visualizar geográficamente la etiqueta con la que se ha categorizado cada barrio. A continuación, se muestran estos resultados, de los cuales se puede consultar la versión publicada en Carto siguiendo los enlaces que hay en el pie de cada imagen.

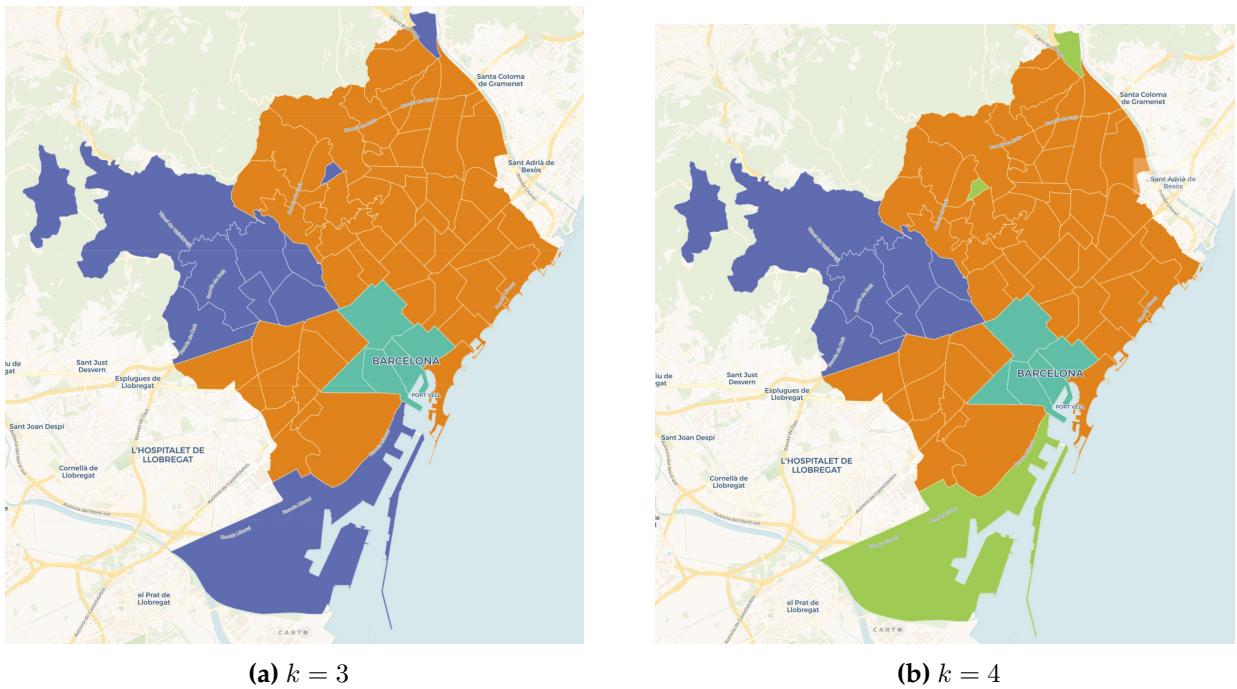


Figura 3.24: Resultado del *clustering* utilizando *k-means* con $k = 3$ (Carto) y $k = 4$ (Carto).

3.5. Análisis de los resultados del *k-means*

Lo primero que se puede observar al mirar el resultado con $k = 3$ es que este mapa resulta algo familiar, puesto que el cluster de color aguamarina tiene casi todos los barrios de los que habían salido gentrificados en los resultados del análisis anterior (Figura 3.22), por lo que este cluster hace referencia a los barrios gentrificados. La única diferencia que hay, es que en este grupo se incluye el barrio de Sant Antoni, el cual no se acabó categorizando como gentrificado, y se excluyen los barrios de el Poble Sec y l'Antiga Esquerra de l'Eixample. Estas diferencias pueden explicarse con el hecho de que el modelo no tiene en cuenta las ubicaciones que hay en cada uno de estos barrios. Por ejemplo, se pudo ver que el barrio de Sant Antoni tenía una distribución de lugares prácticamente igual a la media, mientras que el Poble Sec destacaba bastante por tener un gran número de restaurantes y lugares de ocio y cultura. No obstante, el barrio de l'Antiga Esquerra de l'Eixample, salió como gentrificado, aunque con bastantes más dudas puesto que sus análisis no lo hacían destacar especialmente pese a que sus variables iban acorde con la gentrificación, por lo que es explicable que el *clustering* no lo haya incluido en el grupo de barrios gentrificados.

No obstante, para diferenciar los otros dos clusters obtenidos, es necesario hacer un análisis más específico, por lo que se ha realizado una visualización con las evoluciones medias de

los barrios que han sido clasificados en cada cluster.

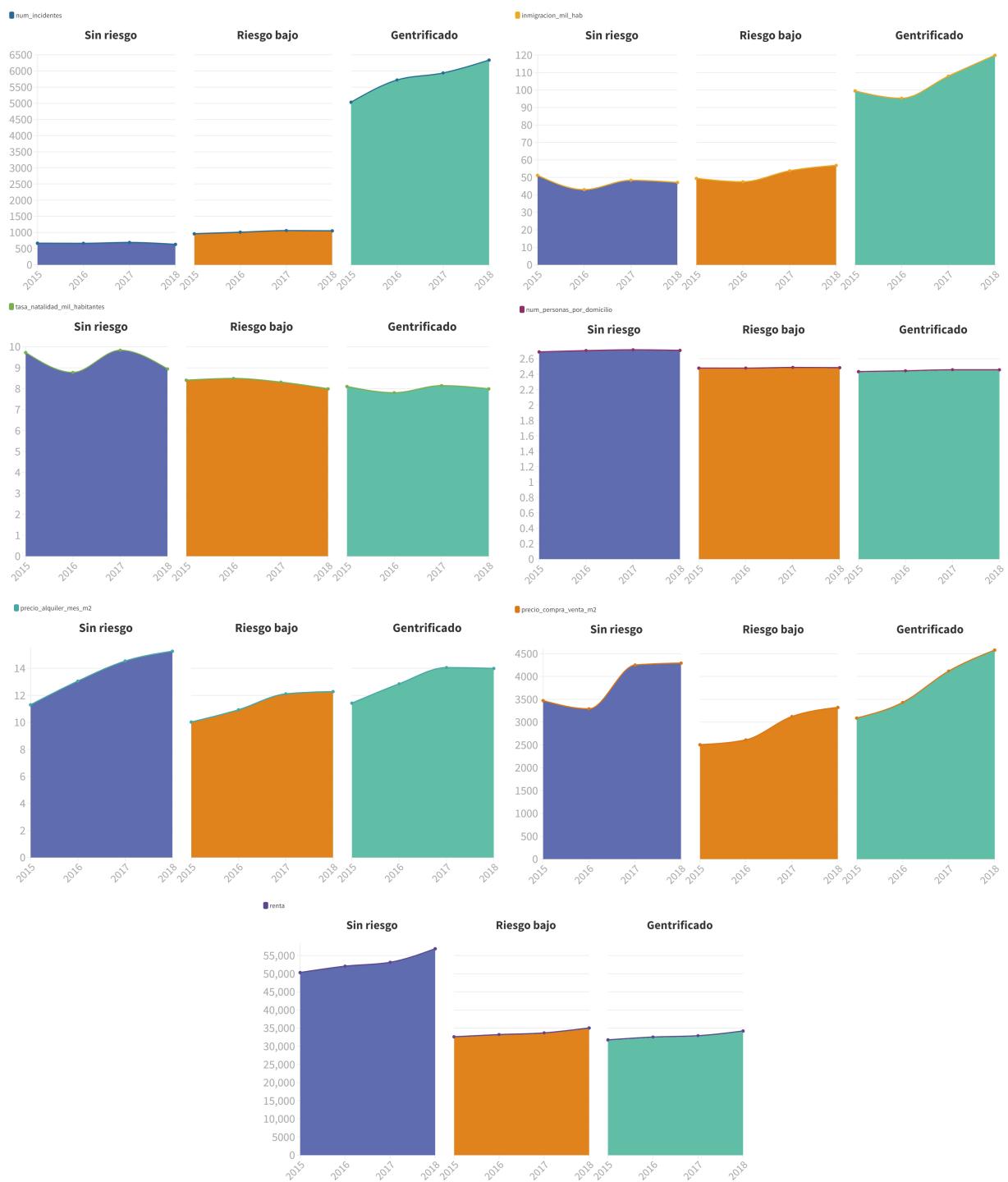


Figura 3.25: Evolución temporal media de las variables por cluster para $k = 3$. Versión interactiva en [Flourish](#). Fuente: Elaboración propia.

Tal como se ve en los gráficos, para cada variable hay una barra por cluster encontrado. La barra de la izquierda hace referencia al grupo morado, la central al naranja y la derecha al aguamarina, la cual se ha visto fácilmente que era la correspondiente al grupo de barrios gentrificados. Como se ve, el color de las barras corresponde al color del cluster representado en el mapa para facilitar la relación entre los gráficos. Para poner nombre a las dos primeras barras o clusters según su nivel de gentrificación, hay que inspeccionar detalladamente los gráficos, pese a que el título de las barras que han sido puestos a posteriori ya indican cuáles han sido las conclusiones.

Si se analiza la primera variable con el número de incidentes, se puede ver claramente las diferencias entre los barrios gentrificados y el resto. En los barrios gentrificados se ve una clara progresión de los incidentes, además de tener un valor absoluto medio mucho mayor que en las dos primeras barras. La primera de ellas, tiene el número de incidentes más bajos y un valor muy estable a lo largo de los años, mientras que la segunda tiene una evolución ligeramente creciente y con un número de incidentes algo mayor.

En cuanto a la inmigración, la barra derecha (gentrificado) tiene un valor considerablemente mayor que el resto de grupos, así como la evolución positiva, a excepción del año 2016 que se produce un decremento general de la inmigración en todos los clusters. Comparando las dos primeras barras, se puede ver que son parecidas, pero que la central tiene una evolución más creciente que para la primera barra.

Si se observa la tasa de natalidad, la barra morada tiene un comportamiento bastante irregular con bajadas y subidas notables, además de ser el grupo con valor mayor. En cambio, las columnas naranja y aguamarina tienen un comportamiento parecido en cuanto a evolución y valores, solo con la diferencia de que la barra central tiene una evolución decreciente más consistente.

Viendo el número de personas por domicilio, no hay ninguna evolución clara en ninguna de las barras, pero sí se puede ver de forma más fácil que la primera barra morada es la que tiene en todos los años el número más alto, mientras que en la segunda es bastante más bajo y en la tercera aún más bajo, pero con mucha menos diferencia respecto la barra naranja, por lo que encaja perfectamente que la barra derecha (aguamarina) es el grupo de barrios gentrificados, tal como se analizó previamente.

En la evolución del precio del alquiler, se observa en la barra morada una evolución bastante lineal, siendo además la de valor mayor. La segunda y tercera barra, tienen valores muy parecidos entre ellas, incluso por el suavizado de la evolución de los años 2016 a 2017, aunque el valor total de esta última barra que pertenece a los barrios gentrificados es algo mayor

que la del cluster central.

En el precio de compra-venta, de nuevo se puede observar que la primera barra tiene un comportamiento diferente a las dos siguientes, pues tiene una evolución irregular a lo largo de los años. Además, las barras central y derecha tienen una evolución creciente bastante parecida, aunque la correspondiente a la gentrificada tiene más pendiente y valores considerablemente mayores que los de la columna naranja, además de tener el valor medio más alto en 2018.

Finalmente, si se observa la renta, claramente la primera barra tiene un valor mucho mayor que el resto, además de ser la que más evolución tiene. Por otro lado, las columnas central y derecha tienen valores prácticamente calcados, por lo que facilita darse cuenta de la clara gentrificación del grupo de la derecha, puesto que con rentas similares el precio de compra-venta es considerablemente mayor en el grupo de la derecha.

En general, después de haber analizado todas las variables con sus barras, se puede ver que la barra central tiene tendencias y valores más parecidos a la correspondiente a los barrios gentrificados (derecha), por lo que todos estos barrios que se han clasificado en el cluster naranja, son aquellos barrios que tienen un **riesgo bajo** de ser gentrificados. Por tanto, los de la barra izquierda (cluster morado) son aquellos que no siguen las características propias de la gentrificación y por lo tanto se puede decir que no tienen **ningún riesgo** de ser gentrificados. Finalmente, la barra derecha de color aguamarina pertenece tal como se ha comentado a los barrios que **ya están gentrificados**.

Si se analiza ahora el caso para $k = 4$, se puede observar que el resultado de los barrios es el mismo que para $k = 3$ con la diferencia de que ahora tres barrios que estaban en el cluster de color morado (sin riesgo de gentrificación) se han movido al nuevo cluster de color verde, por lo que a priori ya se puede intuir que el nuevo cluster será una nueva casuística dentro de los barrios sin riesgo de gentrificación. Si se observan de nuevo las diferencias entre clusters en la figura 3.26, se puede ver como en el nuevo cluster (segunda barra) las variables tampoco siguen comportamientos propios de barrios gentrificados y son además diferentes respecto el cluster de la primera columna (morado). Lo que más destaca sin duda de este nuevo cluster es el precio de compra-venta. No solo es el más bajo de todos los clusters si no que además de 2017 a 2018 tiene una tendencia negativa. Además, a pesar del valor nulo en el precio del alquiler de 2018, se ve una clara bajada del precio de 2016 a 2017. También, cabe destacar que la renta es también la más baja respecto el resto de clusters y que la tasa de natalidad es muy irregular, siendo en este caso la mayor de todas.

3.5. Análisis de los resultados del k -means

69

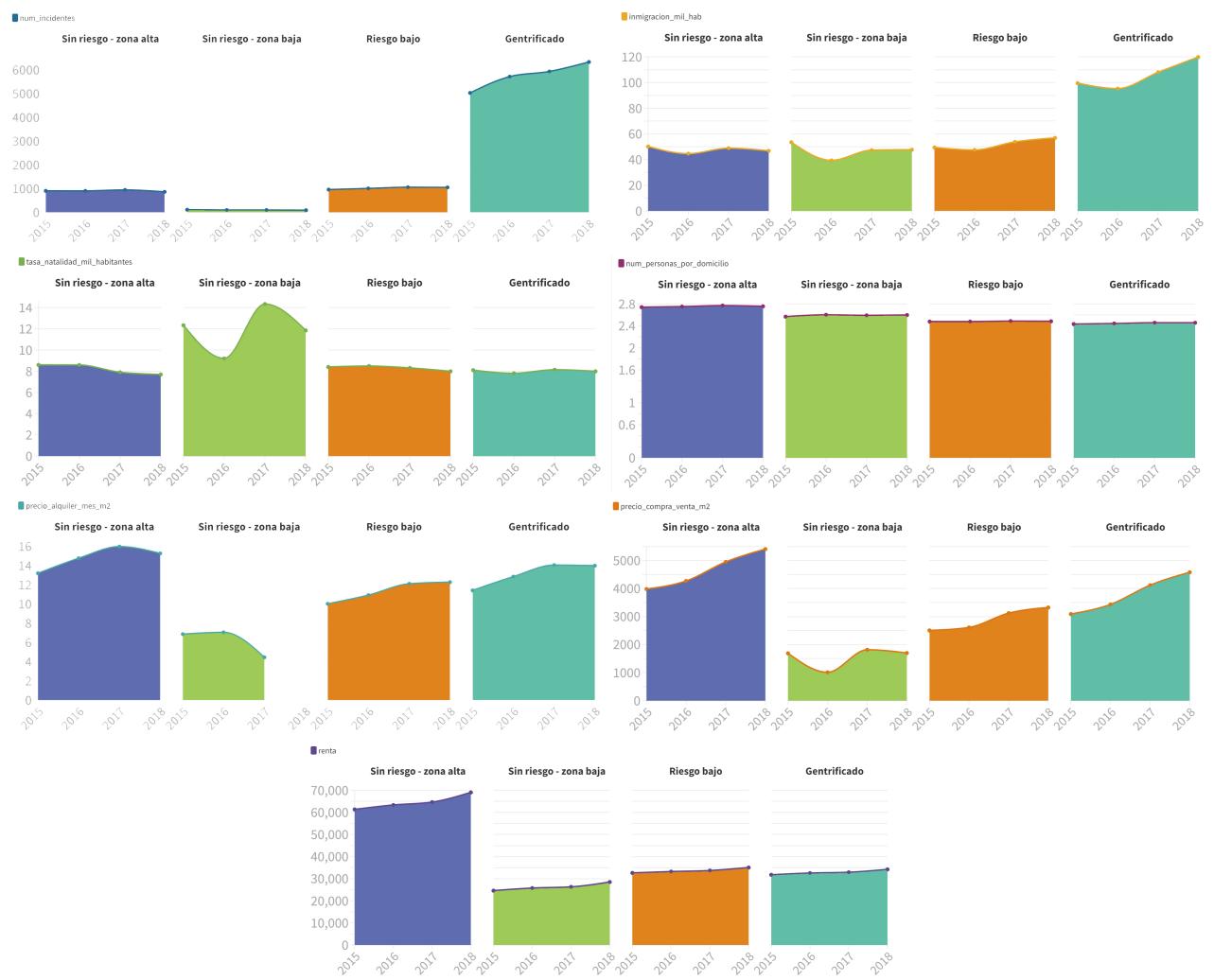


Figura 3.26: Evolución temporal media de las variables por cluster para $k = 4$. Versión interactiva en [Flourish](#). Fuente: Elaboración propia.

En resumen, se ve que el cluster morado del caso $k = 3$ se ha dividido en dos grupos de barrios. El primero, que sigue siendo de color morado y hace referencia a la primera barra de los gráficos, pertenecería al conjunto de barrios que están sin riesgo de ser gentrificados pero que pertenecen a la clase más alta de la ciudad, puesto que las rentas y precios son con diferencia los más altos de Barcelona. Por otra parte, el segundo cluster en el que se ha dividido, que ahora es verde y hace referencia a la segunda barra de los gráficos, pertenece también a los barrios que no están en riesgo de gentrificación, solo que en este caso son barrios de zona más baja en los que las rentas son bajas y los precios de alquiler y compra también. Los dos últimos clusters, que equivaldrían a las barras naranja y aguamarina no se han visto afectados, puesto que la clasificación ha sido exactamente igual que en el caso de $k = 3$.

Capítulo 4

Conclusiones

En este proyecto se ha realizado un análisis de la gentrificación de los barrios de Barcelona basándose en los factores que según el estado del arte podían estar más relacionados con la gentrificación con el objetivo de descubrir cuál es el grado de gentrificación de cada barrio.

Con esta meta, se ha comenzado haciendo un análisis sobre las ubicaciones y negocios presentes en cada barrio que están relacionadas con la gentrificación (restaurantes, tiendas de ropa, etc.) y luego se ha continuado analizando la evolución de otros factores también relacionados con la gentrificación entre los años 2015 y 2018 (renta, inmigración, etc.). De estos estudios se ha podido concluir que los barrios con un nivel avanzado de gentrificación son el Barrí Gòtic y Sant Pere, Santa Caterina i la Ribera. Otros barrios que también han salido como gentrificados, aunque en una fase más intermedia, son la Dreta de l'Eixample y el Poble Sec. Finalmente, los barrios del Raval y l'Antiga Esquerra de l'Eixample se han categorizado como barrios en fase inicial de gentrificación.

En una segunda parte, se ha aplicado un algoritmo de aprendizaje no supervisado (*k-means*) para clasificar los barrios según el nivel de gentrificación y se han obtenido tres categorías las cuales se traducen en barrios gentrificados, barrios estables con riesgo bajo y barrios sin riesgo. Entre los barrios gentrificados el resultado ha sido parecido al del análisis, estando en este grupo el Barri Gòtic, Sant Pere, Santa Caterina i la Ribera, el Raval, Sant Antoni y la Dreta de l'Eixample.

A nivel personal, este proyecto me ha hecho darme cuenta de que la gentrificación, de la cual conocía muy poco, es un hecho que ocurre en Barcelona y como se ha podido ver no en lugares puntuales, sino que parece irse extendiendo moderadamente por los barrios más céntricos de Barcelona y que por tanto, creo que debería controlarse para evitar las consecuencias más negativas, como por ejemplo el hecho de que muchas familias de toda la vida

de un barrio se vean obligadas a desplazarse a otro por la subida del nivel de vida.

En cuanto a nivel más técnico, tengo mucha más visión de las implicaciones que existen en un proyecto de ciencia de datos, especialmente de la importancia de la preparación y limpieza de los datos, la cual es la base de todo proyecto y he necesitado dedicarle bastante más tiempo del esperado para garantizar la calidad y usabilidad de los datos. Acerca de los objetivos del proyecto, considero que se han cumplido, ya que se ha podido identificar factores que caracterizan barrios gentrificados y se ha podido identificar los barrios con diferentes niveles de gentrificación mediante los análisis y el algoritmo *k-means*.

Sobre el seguimiento y desarrollo del proyecto, la planificación y metodología se ha seguido bastante bien en general, aunque en la parte de extracción y preparación de datos he tenido que dedicar más tiempo del esperado, ya que he necesitado hacer varias iteraciones sobre este proceso para añadir fuentes de datos o realizar algunos cambios en las transformaciones. Otra parte que se ha demorado un poco más de lo previsto es la de los análisis, ya que no esperaba tener que analizar tantos datos como realmente he hecho, teniendo que inspeccionar más de una decena de variables en distintas visualizaciones y en los 73 barrios de Barcelona. No obstante, la metodología ha sido adecuada y no se ha necesitado modificar la planificación para garantizar el éxito del trabajo.

Finalmente, al ser un proyecto con un tiempo tan acotado, no se ha podido entrar en demasiado detalle en algunos aspectos en los que sería interesante entrar si se dispusiera de más tiempo.

Por ejemplo, el primer paso a realizar sería considerar o buscar más variables que puedan estar relacionadas con la gentrificación de las que se han podido encontrar para este proyecto, como las ubicaciones de los barrios en años anteriores, las edades medias de los habitantes de cada barrio, los niveles de educación, las obras públicas de mejoras de los barrios, entre otras que se han mencionado en el análisis del estado del arte.

Otra línea de trabajo podría ser la elaboración de un modelo de aprendizaje supervisado que intentase predecir el grado de gentrificación de un barrio a lo largo de los años, con el objetivo de estimar de forma precisa cómo va a evolucionar cada barrio y poder tomar medidas para evitar ciertos fenómenos a tiempo.

Además, sería interesante ver si los modelos obtenidos pueden generalizarse a otras ciudades similares, ya sea a nivel nacional o continental, ya que en el estado del arte se vio que los aprendizajes de Madrid no servían para predecir Nueva York.

En definitiva, el análisis y predicción de la gentrificación es un tema complejo debido a

su abstracción por definición. Es un tema que todavía no está demasiado estudiado y que tiene un gran potencial, por lo que seguramente existan muchas más líneas de trabajo muy interesantes aparte de las que se han mencionado.

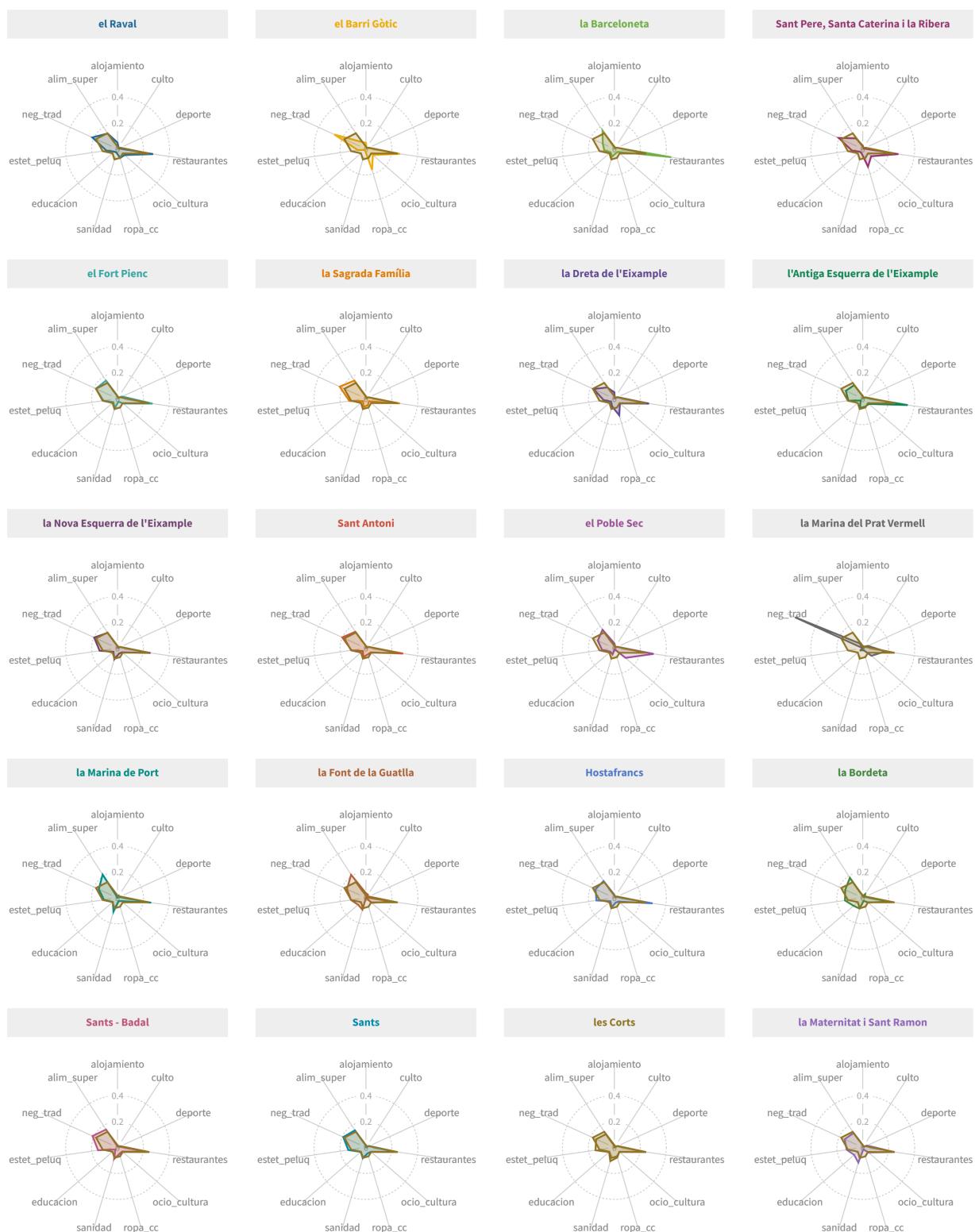
Apéndice A

Gráficos y mapas

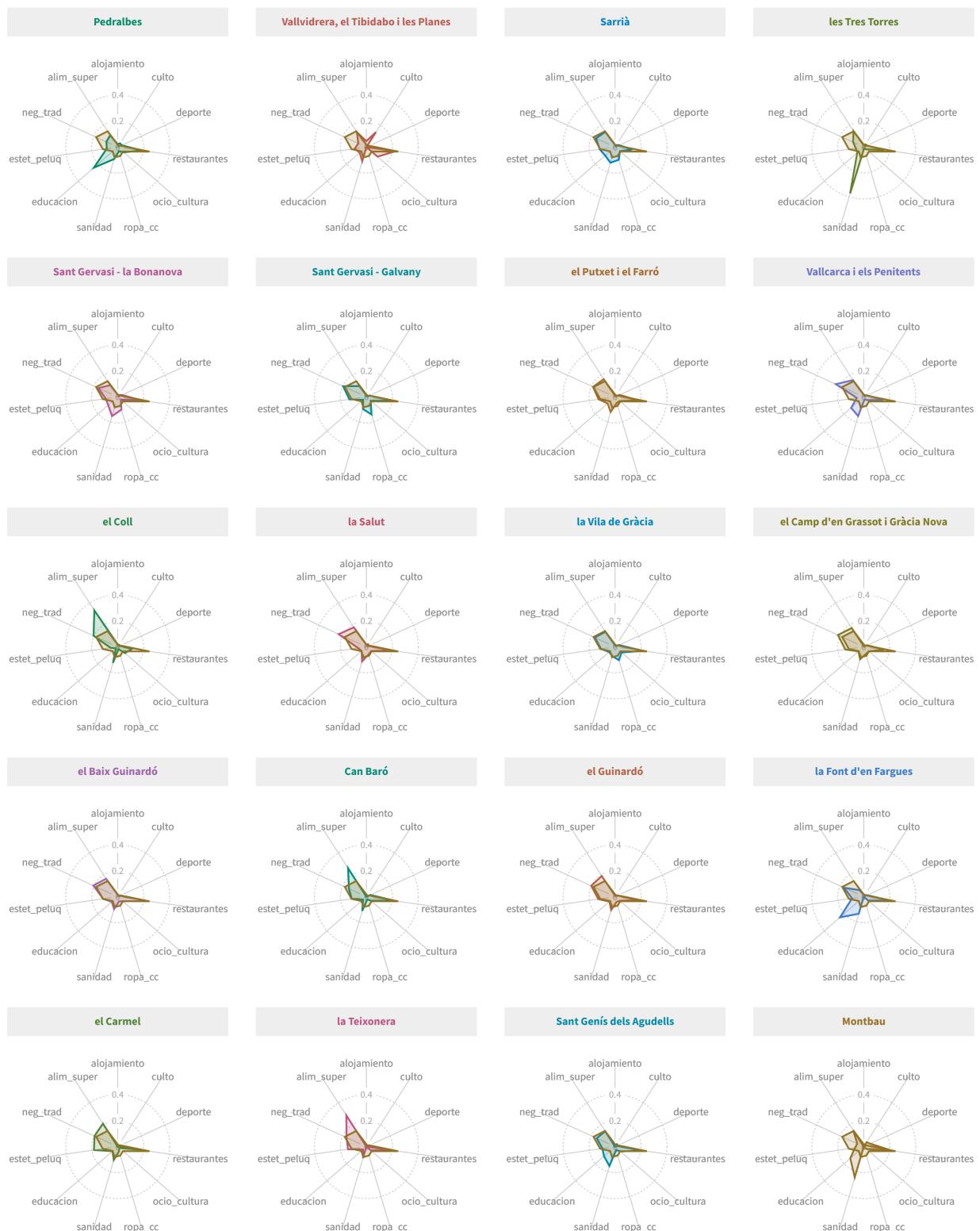
Este apéndice contiene en primer lugar los gráficos de radar que muestran la distribución de lugares o negocios de en cada barrio. Para cada uno de ellos, se ven dos líneas o áreas: una que indica los valores para el barrio en cuestión y otra que indica la media de la ciudad entera para poder comparar los valores particulares y ver si la proporción de negocios de cada tipo en un barrio está por encima o por debajo de la media. La visualización interactiva es accesible en [Flourish](#). Seguidamente, se muestra la misma información pero en mapas de coropletas, teniendo un mapa por cada variable en el que el color de cada barrio dependerá del número de lugares de cada tipo presentes en cada barrio, los cuales están representados por puntos en el mapa, el cual está accesible en [Carto](#).

En segundo lugar, se encontrarán los gráficos de la evolución de 2015 a 2018 de los distintos indicadores en cada barrio representados en mapas de coropletas y en gráficos de línea. Recordar que las gráficos interactivos pueden ser consultados en [Carto](#) y en [Flourish](#).

A.1. Gráficos de radar de ubicaciones



A.1. Gráficos de radar de ubicaciones



A.1. Gráficos de radar de ubicaciones



A.1. Gráficos de radar de ubicaciones

77

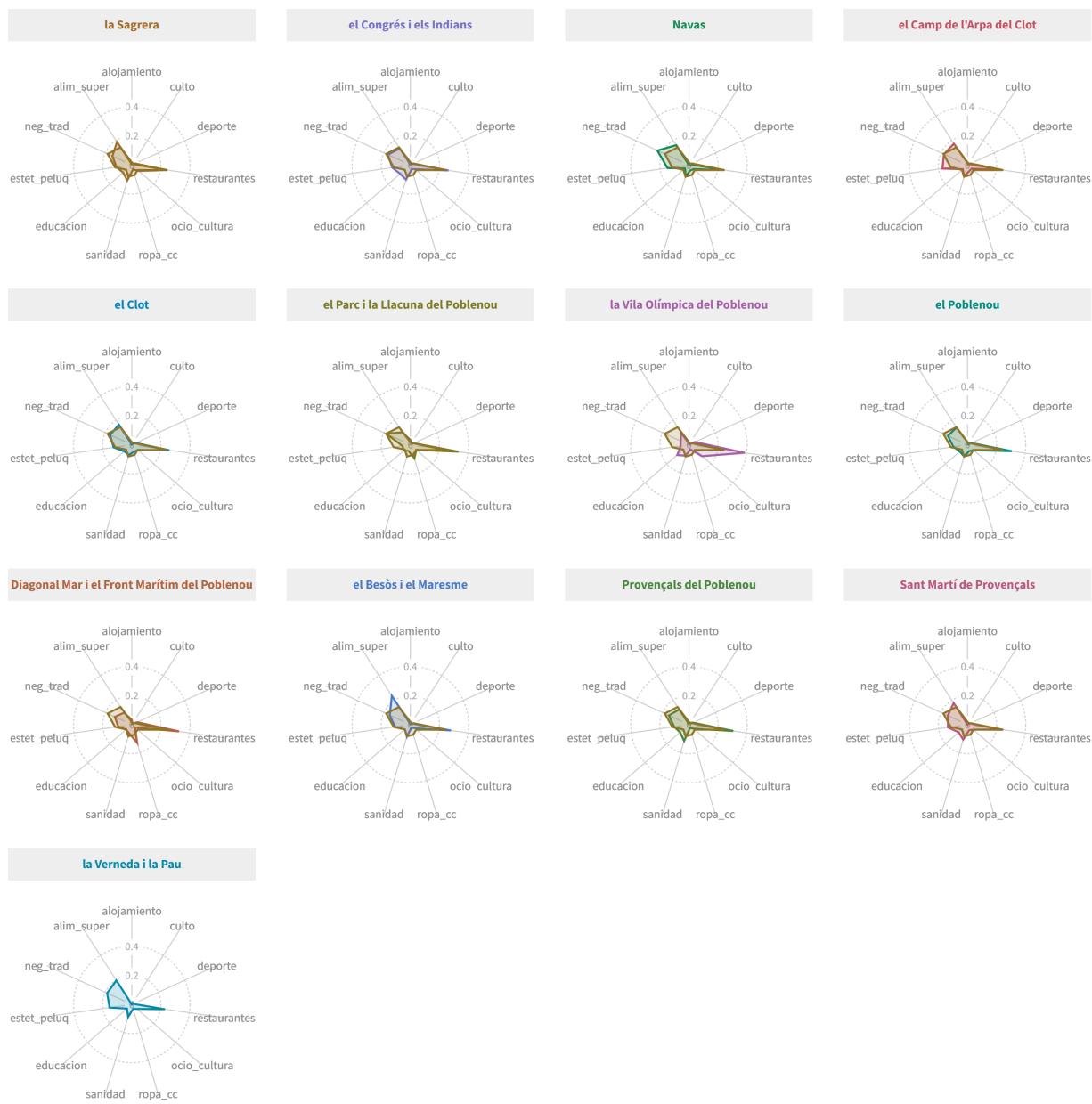
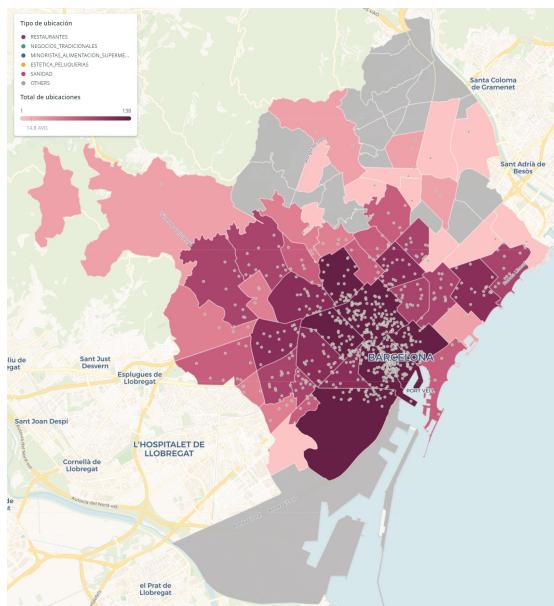
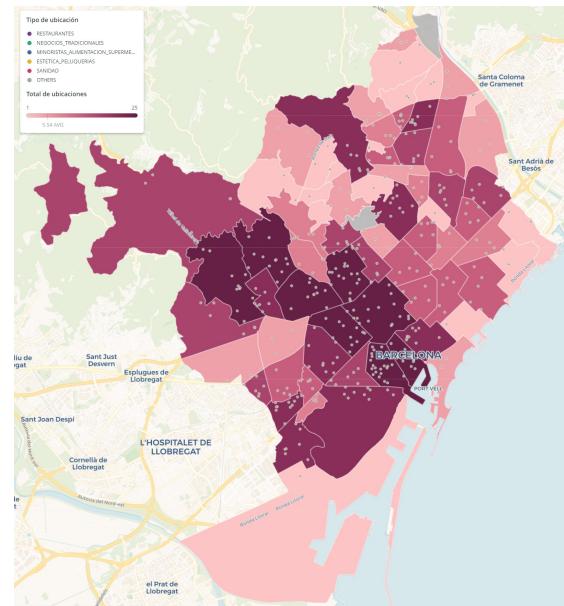


Figura A.1: Diagramas de radar que muestran la distribución de las distintas categorías de negocios o ubicaciones por barrio. La línea marrón de cada gráfico muestra la media de Barcelona, mientras que la de color distinto muestra la propia del barrio. Fuente: Elaboración propia.

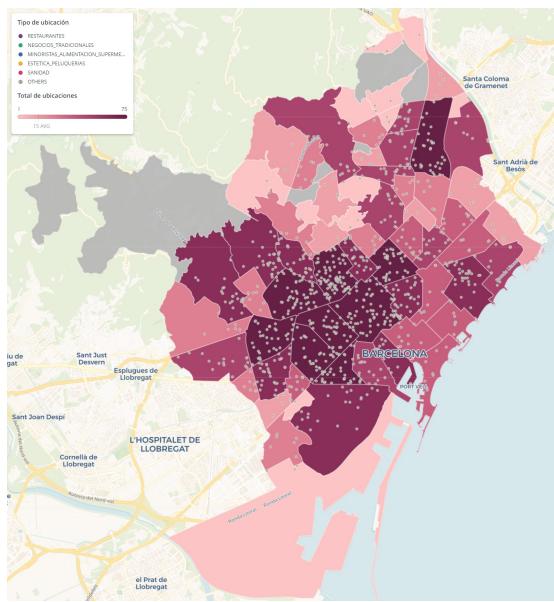
A.2. Mapas de coropletas de número de ubicaciones de cada tipo



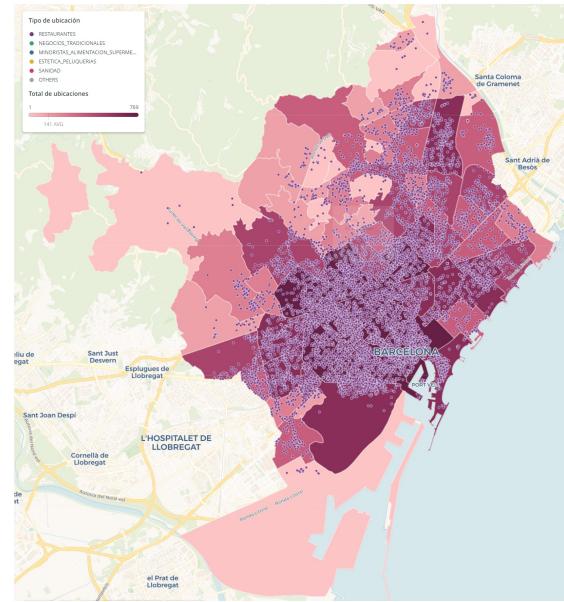
(a) Alojamientos



(b) Lugares de culto

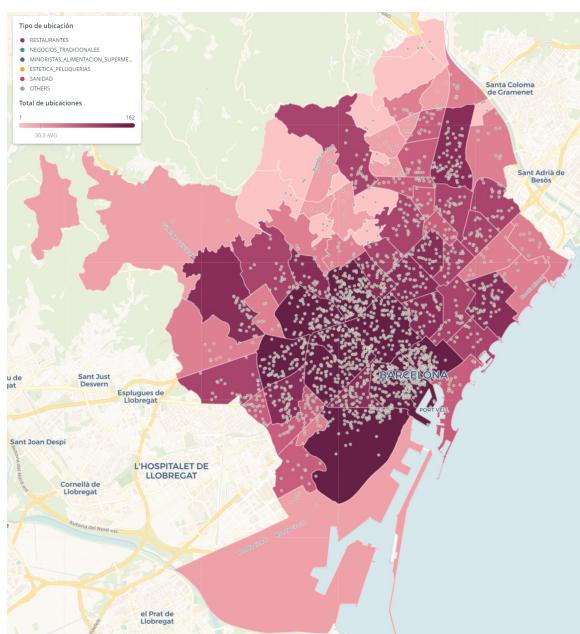


(c) Instalaciones deportivas / gimnasios

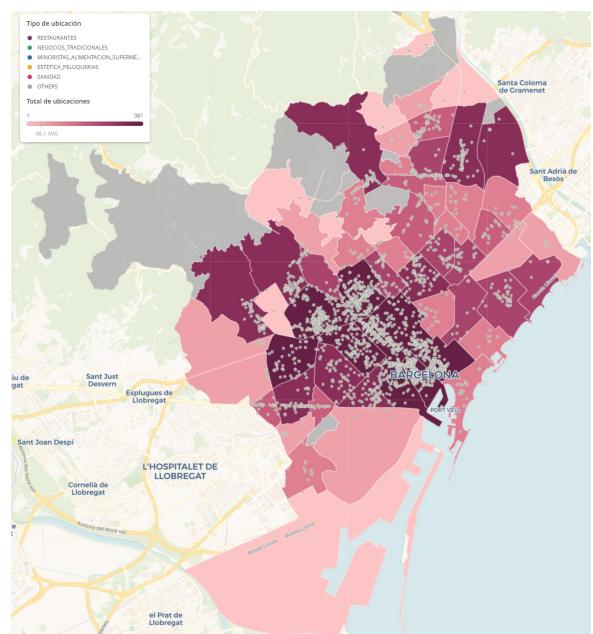


(d) Restaurantes

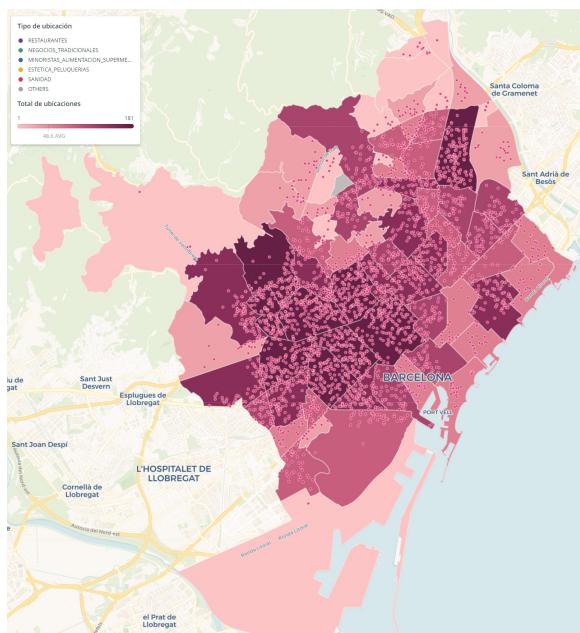
Figura A.2: Mapa de coropletas para el número de alojamientos, lugares de culto, instalaciones deportivas y restaurantes. Fuente: Elaboración propia.



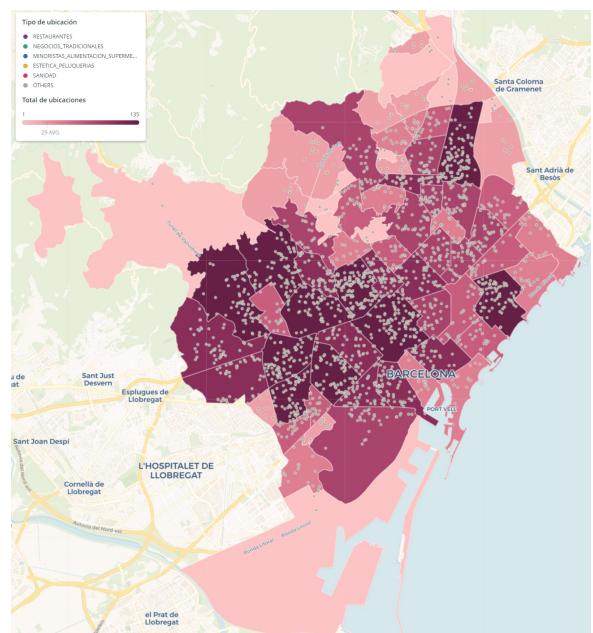
(a) Ocio y cultura



(b) Tiendas de ropa y centros comerciales



(c) Sanidad



(d) Educación

Figura A.3: Mapa de coropletas para el número de ubicaciones de ocio y cultura, tiendas de ropa y centros comerciales, centros sanitarios y farmacias y centros educativos. Fuente: Elaboración propia.

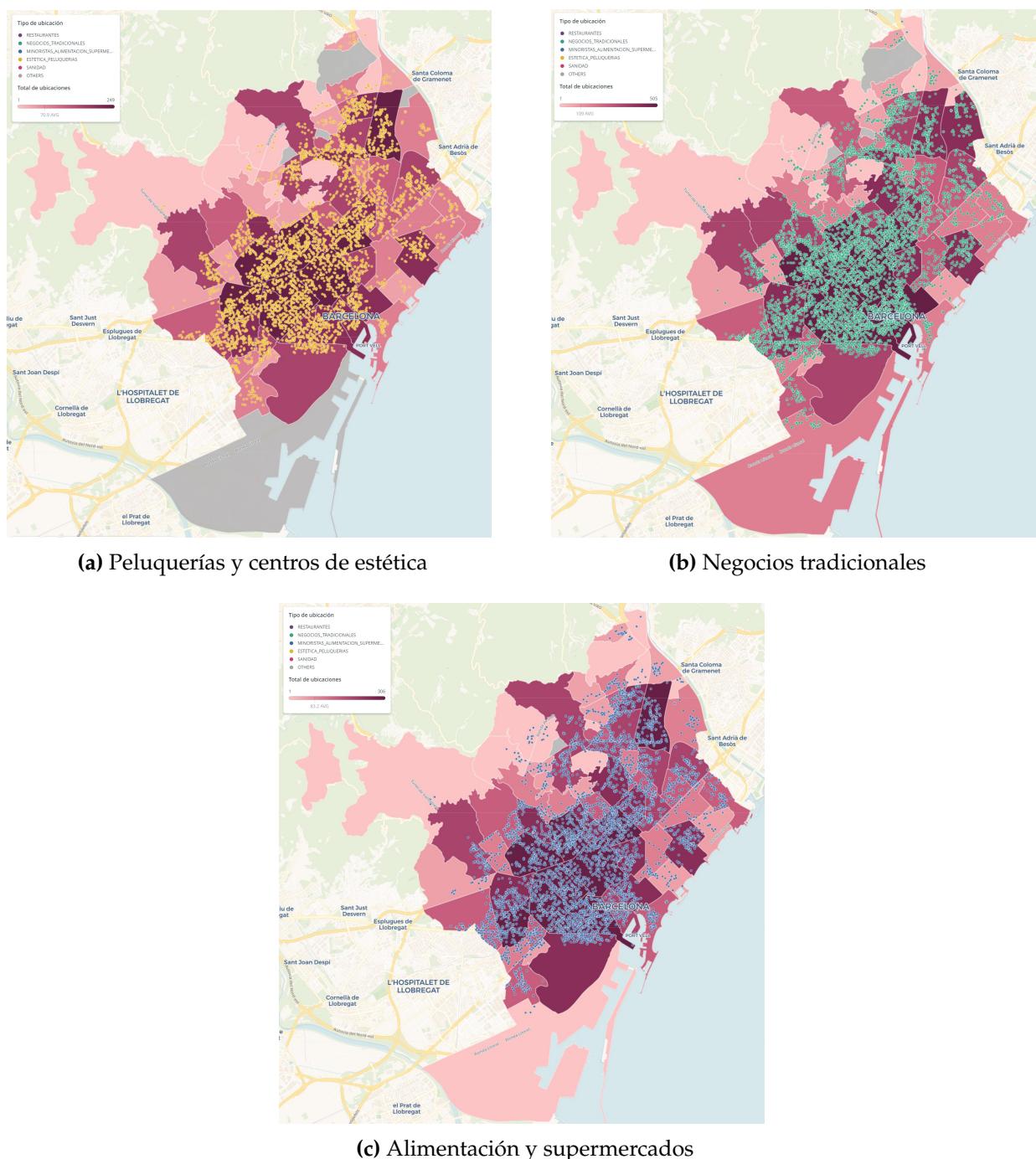


Figura A.4: Mapa de coropletas para el número de ubicaciones de peluquerías y centros de estética, negocios tradicionales y alimentación y supermercados. Fuente: Elaboración propia.

A.3. Evolución de indicadores de 2015 a 2018

A.3.1. Renta

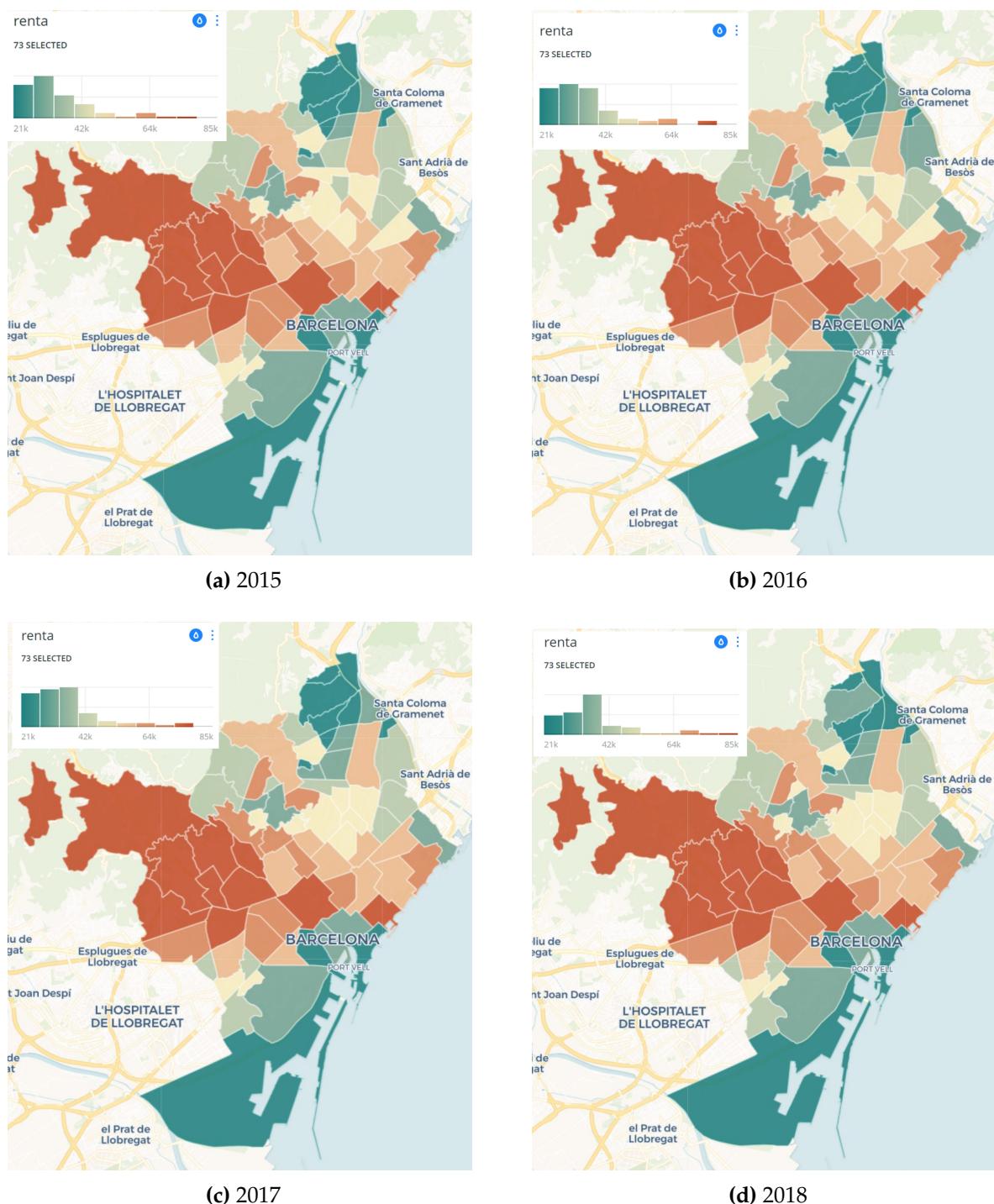


Figura A.5: Evolución de la renta de los años 2015 a 2018. Fuente: Elaboración propia.

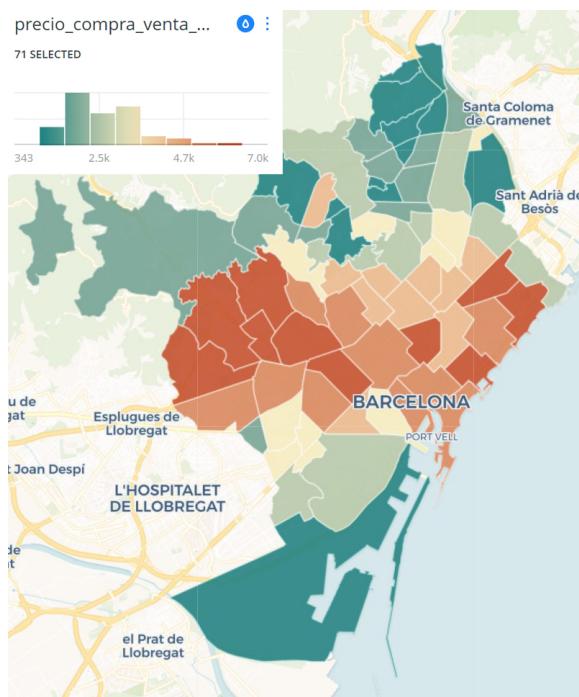
A.3. Evolución de indicadores de 2015 a 2018

82

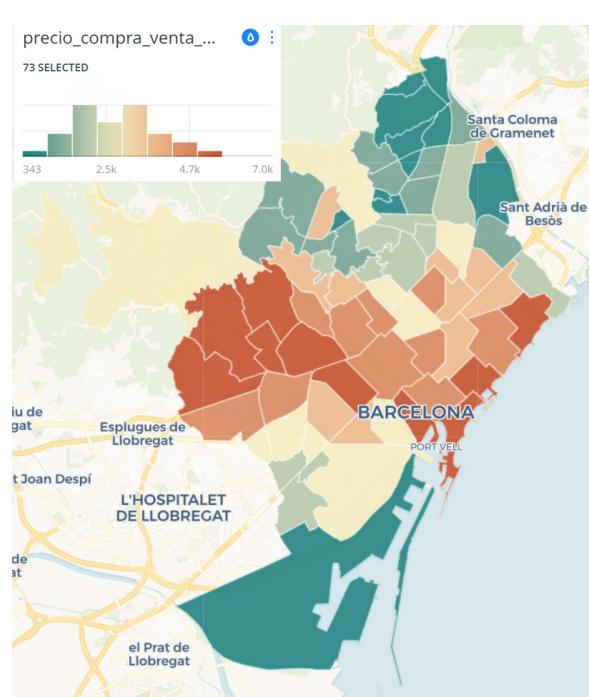


Figura A.6: Evolución de la renta de los años 2015 a 2018. Fuente: Elaboración propia.

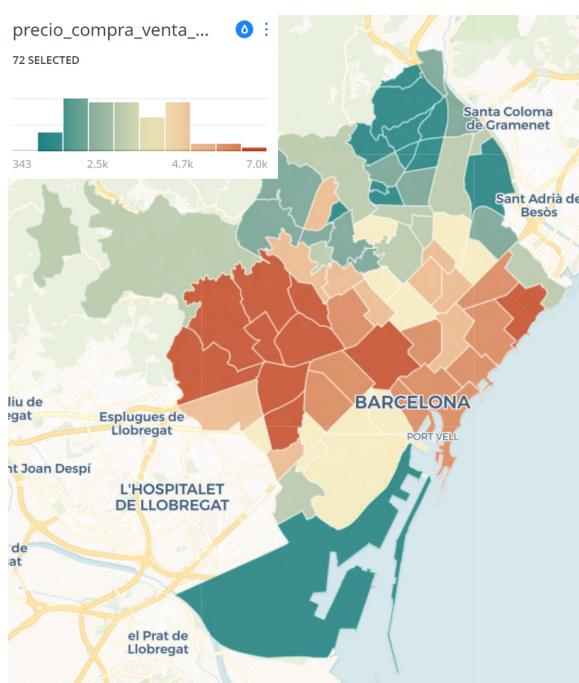
A.3.2. Precio de compra-venta por metro cuadrado



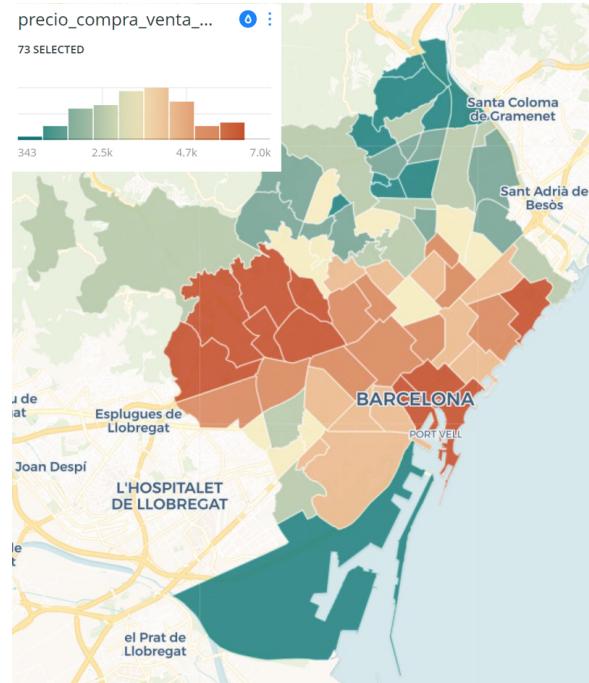
(a) 2015



(b) 2016



(c) 2017



(d) 2018

Figura A.7: Evolución de los precios de compra-venta por metro cuadrado de los años 2015 a 2018.
Fuente: Elaboración propia.

A.3. Evolución de indicadores de 2015 a 2018

84



Figura A.8: Evolución de los precios de compra-venta por metro cuadrado de los años 2015 a 2018.
Fuente: Elaboración propia.

A.3.3. Precio del alquiler al mes por metro cuadrado

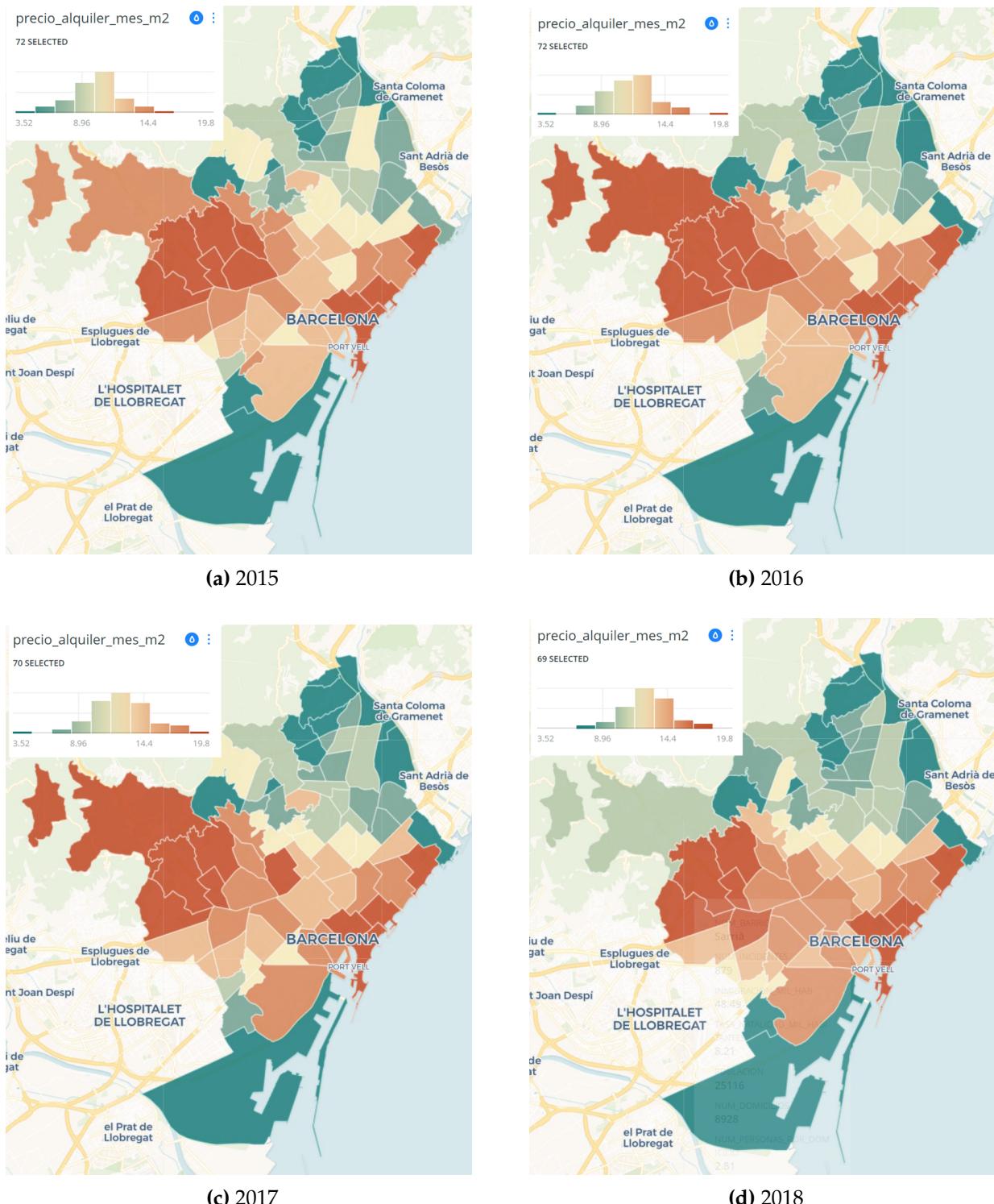


Figura A.9: Evolución de los precios mensuales de alquiler por metro cuadrado de los años 2015 a 2018. Fuente: Elaboración propia.

A.3. Evolución de indicadores de 2015 a 2018

86



Figura A.10: Evolución de los precios mensuales de alquiler por metro cuadrado de los años 2015 a 2018. Fuente: Elaboración propia.

A.3.4. Número de personas por domicilio

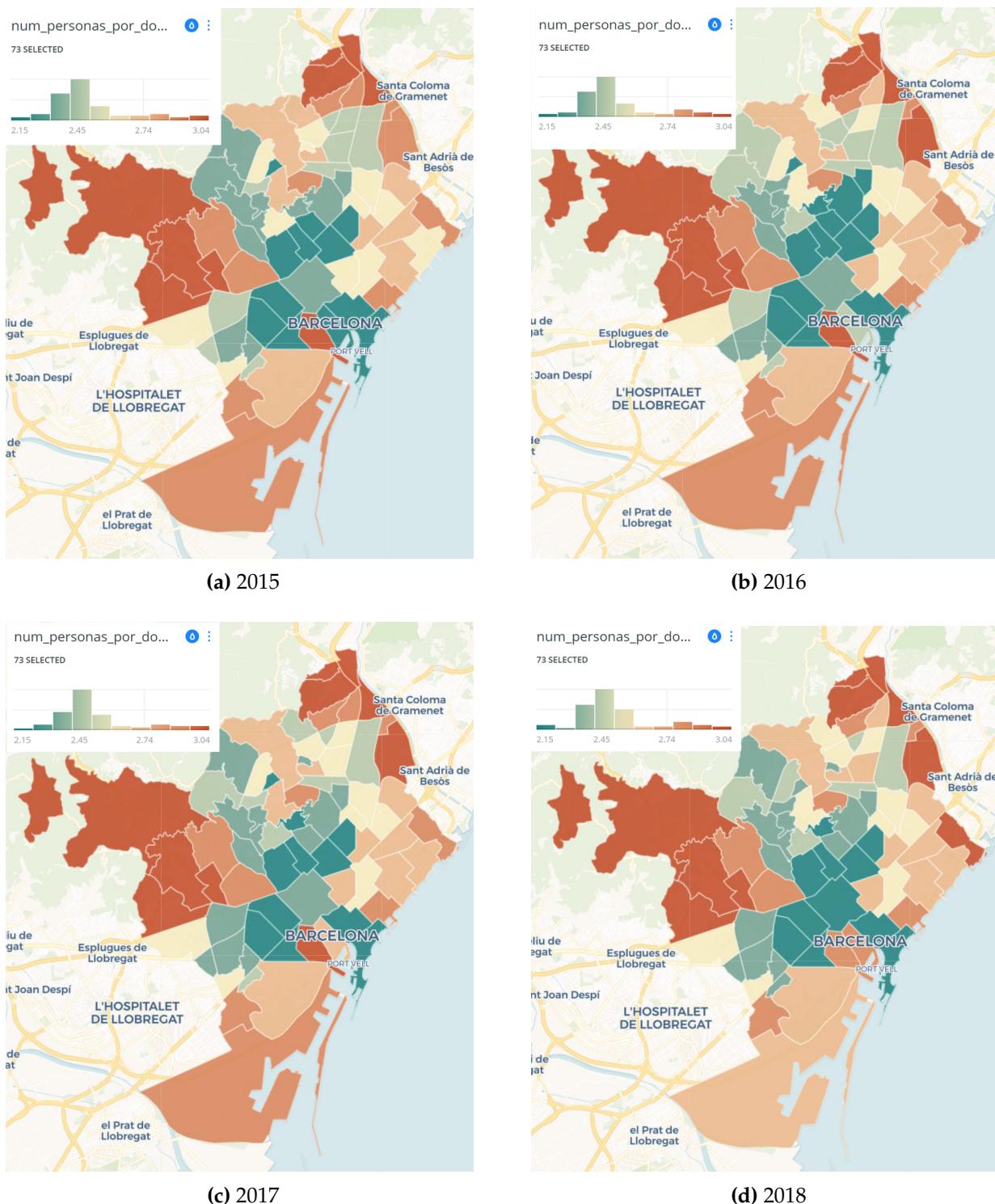


Figura A.11: Evolución del número de personas por domicilio de los años 2015 a 2018. Fuente: Elaboración propia.

A.3. Evolución de indicadores de 2015 a 2018

88



Figura A.12: Evolución del número de personas por domicilio de los años 2015 a 2018. Fuente: Elaboración propia.

A.3.5. Número de incidentes

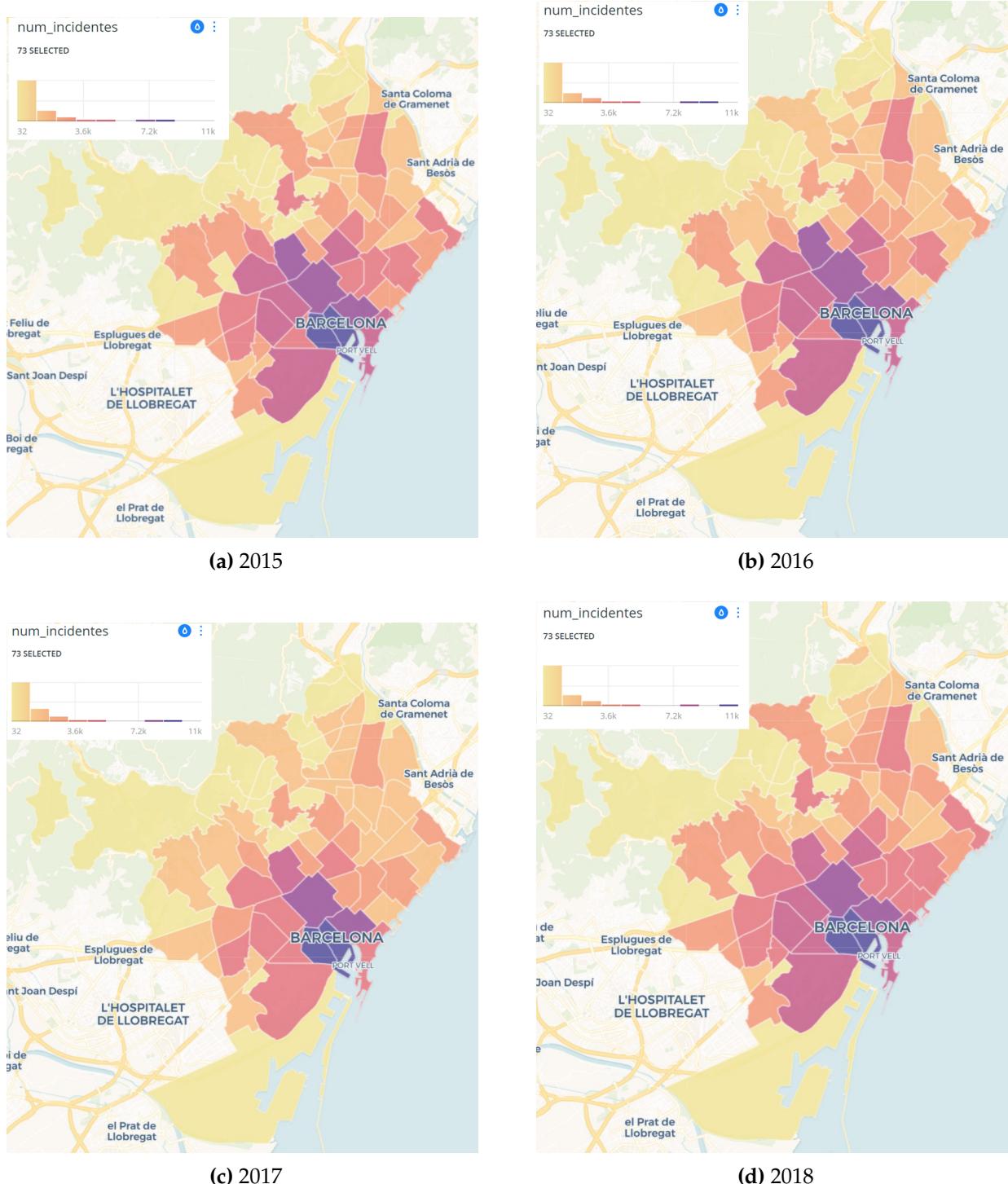


Figura A.13: Evolución del número de incidentes de los años 2015 a 2018. Fuente: Elaboración propia.

A.3. Evolución de indicadores de 2015 a 2018

90

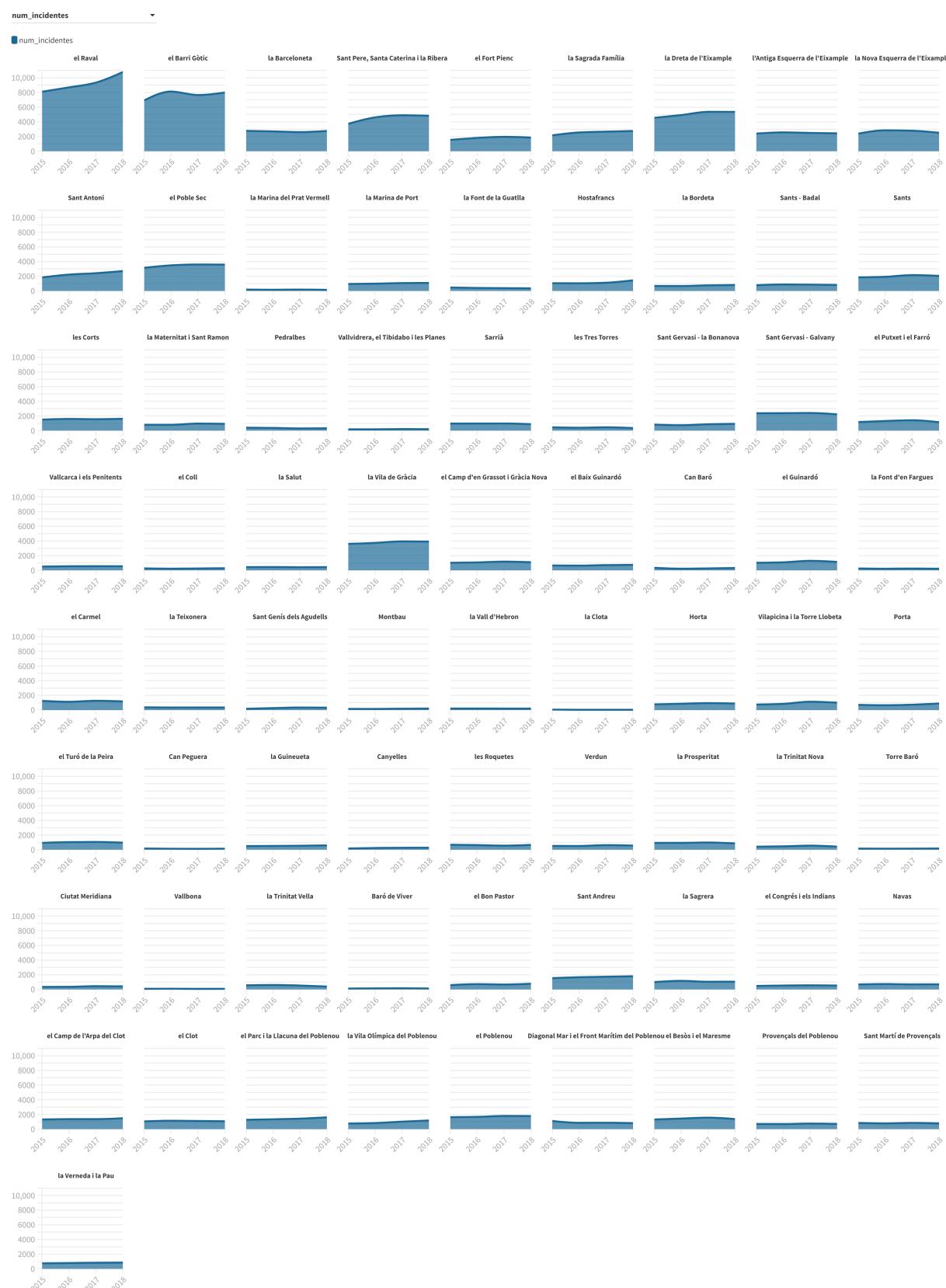


Figura A.14: Evolución del número de incidentes de los años 2015 a 2018. Fuente: Elaboración propia.

A.3.6. Tasa de inmigración

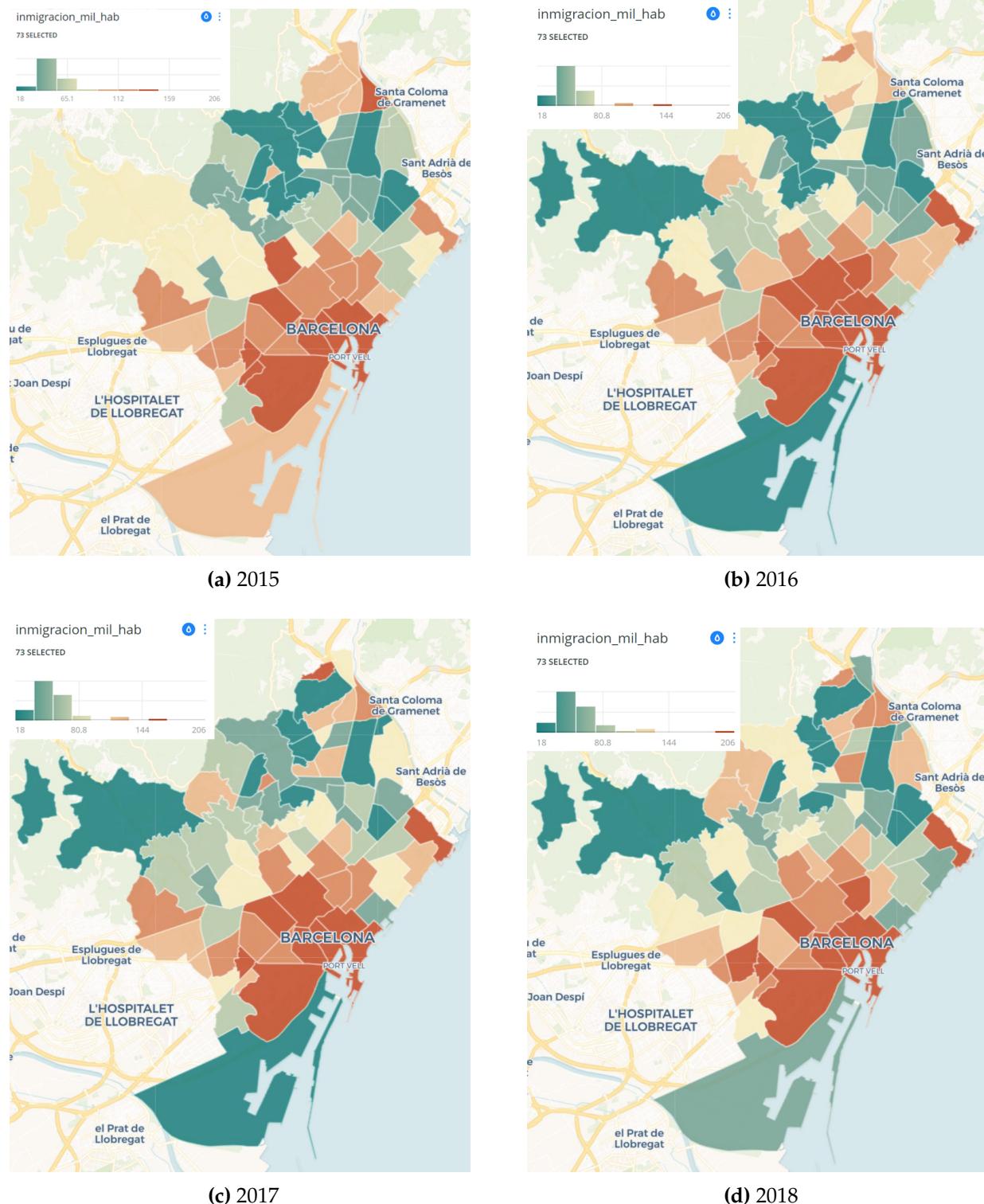


Figura A.15: Evolución de la tasa de inmigración de los años 2015 a 2018. Fuente: Elaboración propia.

A.3. Evolución de indicadores de 2015 a 2018

92



Figura A.16: Evolución de la tasa de inmigración de los años 2015 a 2018. Fuente: Elaboración propia.

A.3.7. Tasa de natalidad

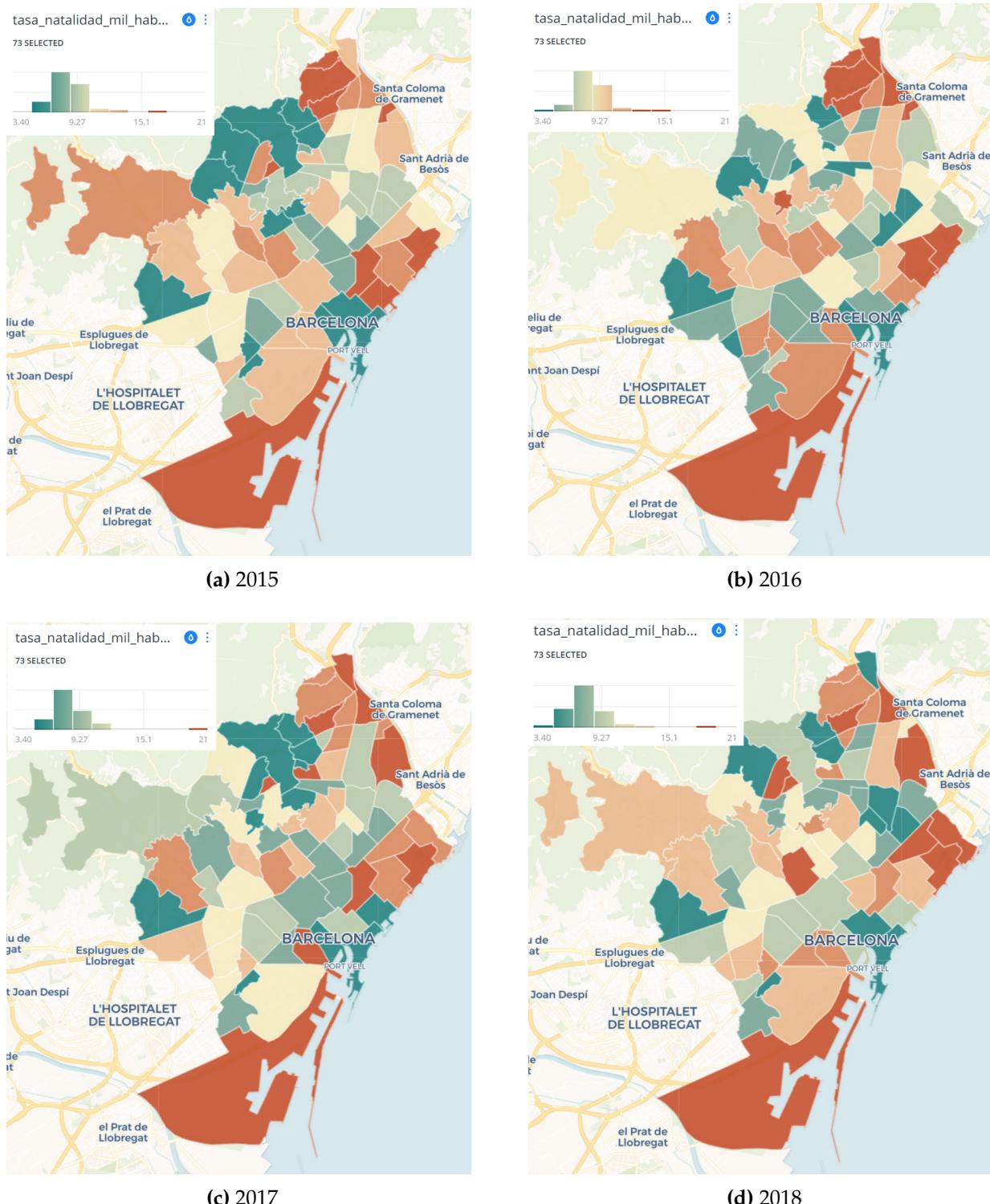


Figura A.17: Evolución de la tasa de natalidad de los años 2015 a 2018. Fuente: Elaboración propia.

A.3. Evolución de indicadores de 2015 a 2018

94



Figura A.18: Evolución de la tasa de natalidad de los años 2015 a 2018. Fuente: Elaboración propia.

A.4. Mapa de barrios de Barcelona

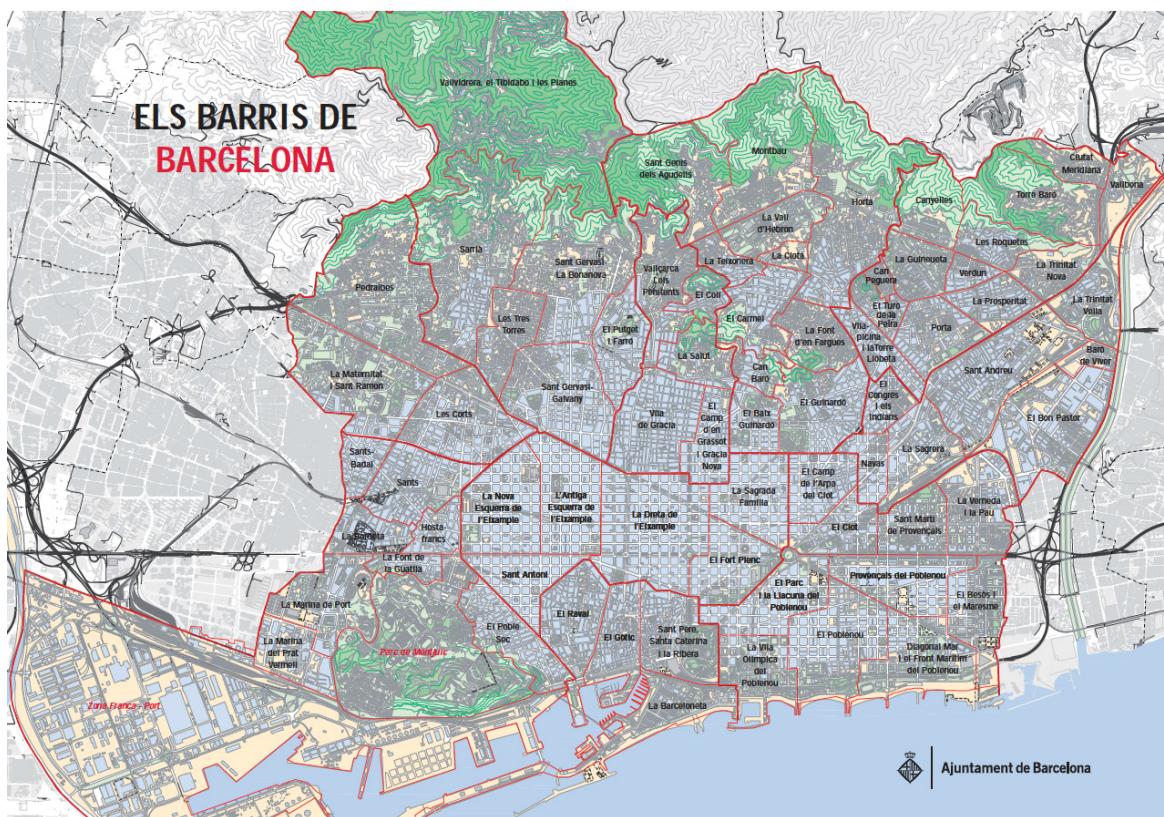


Figura A.19: Mapa de los barrios de Barcelona. Fuente: [21]

Apéndice B

Código

El código Python con el que se ha desarrollado el proyecto está subido en un repositorio de [Github](#). En el fichero README . md está explicado el proceso necesario para ejecutar el código en cualquier máquina.

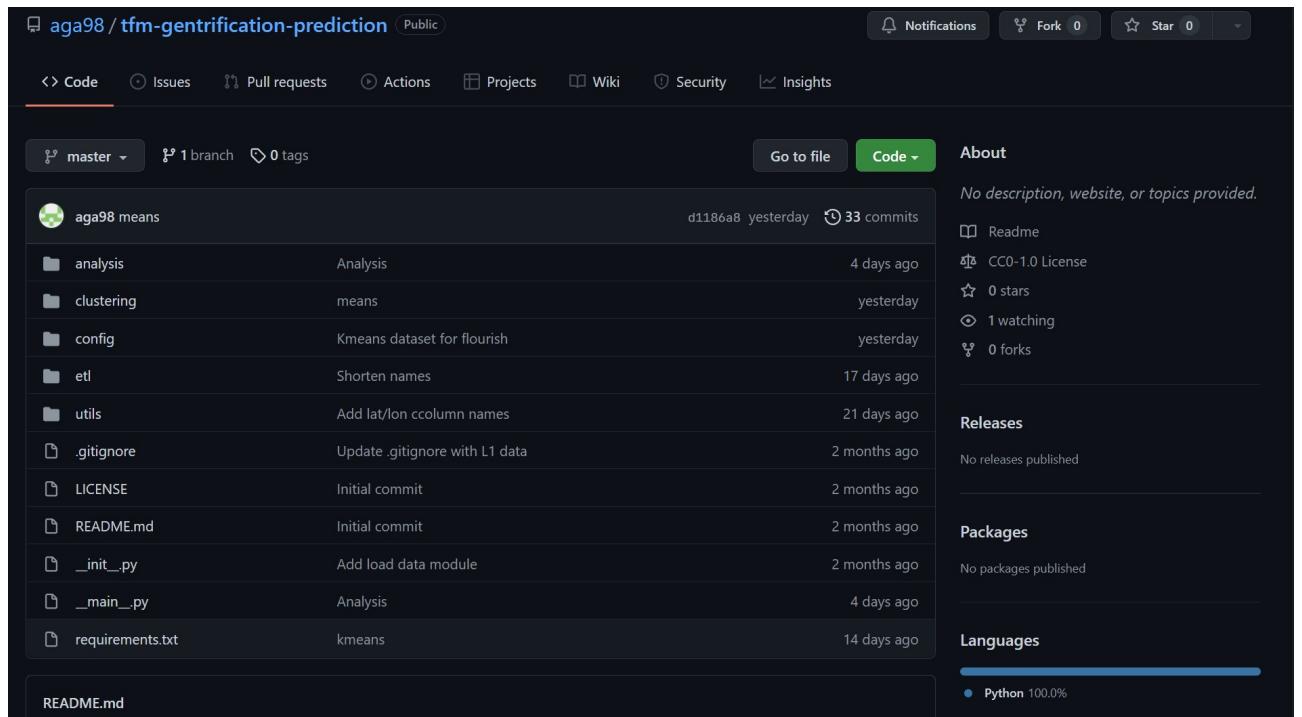


Figura B.1: Captura del repositorio en Github. Fuente: Elaboración propia

Glosario

gentrificación Proceso de transformación de un espacio urbano deteriorado o en declive, a partir de la preconstrucción o rehabilitación edificatoria con mayores alturas que las existentes, lo que provoca un aumento de los alquileres o del coste habitacional en estos espacios. [2–8, 11, 16–22, 24–26, 34, 36, 42, 45–52, 54, 55, 57–61, 63, 65, 67–71](#)

GIS Sistema de Información Geográfico (SIG). [11](#)

k-means Método de agrupamiento, que tiene como objetivo la partición de un conjunto de n observaciones en k grupos en el que cada observación pertenece al grupo cuyo valor medio es más cercano . [8, 22, 23, 25, 63, 64, 70, 71](#)

open data Filosofía y práctica que persigue que determinados tipos de datos estén disponibles de forma libre para todo el mundo, sin restricciones de derechos de autor, de patentes o de otros mecanismos de control . [7](#)

Random Forest Combinación de árboles predictores tal que cada árbol depende de los valores de un vector aleatorio probado independientemente y con la misma distribución para cada uno de estos. Es una modificación sustancial de bagging que construye una larga colección de árboles no correlacionados y luego los promedia . [24](#)

turistificación Conjunto de consecuencias que experimenta una población, o parte de ella, al recibir una afluencia masiva de turistas . [3, 4](#)

Bibliografía

- [1] Juan Fernández. *Gentrificación: así cambia nuestros barrios- En Naranja*, ING. Oct. de 2019.
URL: <https://www.ennaranja.com/hipotecas/gentrificacion-barrios-ciudades/> (visitado 17-08-2021).
- [2] Alyssa Wiltse-Ahmad. *Study: Gentrification and cultural displacement most intense in America's largest cities, and absent from many others* » NCRC. Ago. de 2019.
URL: <https://ncrc.org/study-gentrification-and-cultural-displacement-most-intense-in-americas-largest-cities-and-absent-from-many-others/> (visitado 18-08-2021).
- [3] *Gentrification - Wikipedia*. URL:
<https://en.wikipedia.org/wiki/Gentrification> (visitado 18-08-2021).
- [4] *Gentrificación: descubre los barrios más transformados de España*. Nov. de 2019.
URL: <https://blog.haya.es/gentrificacion-descubre-los-barrios-mas-transformados-de-espana/> (visitado 17-08-2021).
- [5] *Understand the relationship between gentrification and small music venues - Live DMA*.
URL: <https://www.live-dma.eu/understand-the-relationship-between-gentrification-and-small-music-venues/> (visitado 19-08-2021).
- [6] Jose Mansilla. *Consideraciones sobre la relación entre turistificación y gentrificación — Artículo de opinión en Hosteltur*. Abr. de 2018. URL:
https://www.hosteltur.com/comunidad/005997_consideraciones-sobre-la-relacion-entre-turstificacion-y-gentrificacion.html (visitado 17-08-2021).
- [7] LUCA Talk: *Open Data para entender la transformación de los barrios. - Think Big Empresas*. 2019. URL: <https://empresas.blogthinkbig.com/luca-talk-open-data-gentrificacion/> (visitado 17-08-2021).

- [8] *Agile-Waterfall Hybrid: Is It Right for Your Team?* — Lucidchart Blog.
URL: <https://www.lucidchart.com/blog/is-agile-waterfall-hybrid-right-for-your-team> (visitado 14-09-2021).
- [9] *Top 4 software development methodologies* — Synopsys. 2017.
URL: <https://www.synopsys.com/blogs/software-security/top-4-software-development-methodologies/> (visitado 14-09-2021).
- [10] *Agile becomes mainstream.* 2018.
URL: <https://www.computerweekly.com/opinion/Agile-becomes-mainstream> (visitado 14-09-2021).
- [11] *The Cascading Costs of Waterfall. The Waterfall model has traditionally... — by Jones + Waddell* — Medium. Abr. de 2019.
URL: <https://medium.com/@joneswaddell/the-cascading-costs-of-waterfall-5c3b1b8beaec> (visitado 14-09-2021).
- [12] *Open Data BCN — Servicio de datos abiertos del Ajuntament de Barcelona.* URL: <https://opendata-ajuntament.barcelona.cat/es/> (visitado 22-11-2021).
- [13] Joe Cortright. *Everything that causes gentrification, from A to Z* — City Observatory.
URL: http://cityobservatory.org/everything-that-causes-gentrification-from-a-to-z/?utm_campaign=Bass%20Newsletter&utm_source=hs_email&utm_medium=email&utm_content=74087731 (visitado 30-09-2021).
- [14] Morgan Simon. *What Do Hipsters And Banks Have In Common? Gentrification.* 2019.
URL: <https://www.forbes.com/sites/morgansimon/2019/01/31/what-do-hipsters-and-banks-have-in-common-gentrification/?sh=4410a64f559f> (visitado 01-10-2021).
- [15] Jesse Keenan, Thomas Hill y Anurag Gumber.
«Climate gentrification: From theory to empiricism in Miami-Dade County, Florida». En: *Environmental Research Letters* 13 (mayo de 2018), pág. 054001.
DOI: [10.1088/1748-9326/aabb32](https://doi.org/10.1088/1748-9326/aabb32).
- [16] Ingrid Gould Ellen y Keren Horn.
Has Falling Crime Invited Gentrification? – NYU Furman Center. 2019.
URL: <https://furmancenter.org/research/publication/has-falling-crime-invited-gentrification> (visitado 01-10-2021).

- [17] *The club next door: as MLS expands, concerns over gentrification loom — Orlando City — The Guardian.*
URL: <https://www.theguardian.com/football/2017/may/18/mls-stadium-gentrification-orlando-city> (visitado 01-10-2021).
- [18] Jackelyn Hwang y Jeffrey Lin. *What Have We Learned About the Causes of Recent Gentrification? by Jackelyn Hwang, Jeffrey Lin :: SSRN*. Jul. de 2016. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2812045 (visitado 30-09-2021).
- [19] David Christopher Knorr. «Using Machine Learning to Identify and Predict Gentrification in Nashville, Tennessee.» En: (jul. de 2019).
URL: <https://ir.vanderbilt.edu/handle/1803/13285>.
- [20] *K-means: Elbow Method and Silhouette — by Jonathan Ramirez — Medium.*
URL: <https://medium.com/@jonathanrmzg/k-means-elbow-method-and-silhouette-e565d7ab87aa> (visitado 12-10-2021).
- [21] *Mapa y plano de 10 distritos y barrios de Barcelona.*
URL: <https://mapabarcelona360.es/mapa-barrios-barcelona> (visitado 16-12-2021).