**MARCO Y JUSTIFICACIÓN DE LA RELEVANCIA Y NOVEDAD DEL TRABAJO REALIZADO**

<https://www.malaga.es/noticias/com1_md-3/com1_ct-0/com1_fb-0/com1_cb-0/com1_md3_cd-50341/el-sector-turistico-de-la-costa-del-sol-evidencia-el-peso-del-turismo-sobre-otros-sectores-y-lanza-un-potente-mensaje-contra-la-turismofobia>

<https://cadenaser.com/andalucia/2024/05/08/el-empleo-del-sector-turistico-crece-casi-un-8-en-malaga-con-125000-ocupados-en-la-hosteleria-alojamientos-y-agencias-ser-malaga/>

https://www.malaga.es/noticias/com1\_prt-1/com1\_md3\_cd-50000/la-costa-del-sol-continua-despuntando-en-el-primer-trimestre-de-2024-con-cerca-de-millon-y-medio-de-viajeros#:~:text=As%C3%AD%2C%20de%20este%20%C3%BAltimo%20informe,en%20la%20cifra%20de%205.423.

El sector turístico es una de los pilares más fuertes de la provincia de Málaga, y de su ciudad.

Durante el primer cuatrimestre de 2024, la costa del sol registró un incremento del 8,1% en la estancia de viajeros en establecimientos reglados, y un 13,5% en alojamientos turísticos durante el primer cuatrimestre de 2024. Es sobre ese 13,5% donde va a centrarse este análisis.

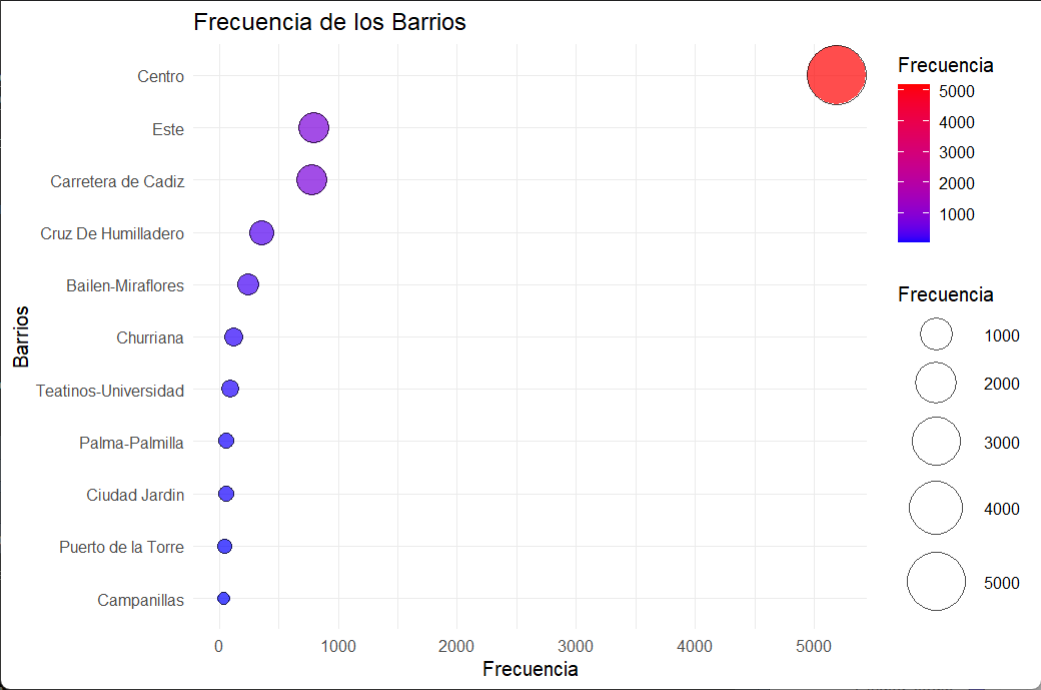
La inmensa mayoría de estos alojamientos turísticos, cuyo número en la ciudad de Málaga no para de ir en crecimiento, se encuentran anunciados bajo la plataforma AirBNB.

El objetivo de este trabajo es, a partir de los datos ofrecidos por Airbnb de las viviendas que se anuncian en su plataforma, hacer un estudio del mercado desde el punto de vista de la oferta y la demanda.

**PRESENTACIÓN DE LOS DATOS**

Los datos encontrados se agrupan en barrios, esto puede llevar a la pregunta, para comenzar el análisis ¿Se agrupan de forma diferente los pisos ofertados por su zona?, o expresado de una forma más directa, ¿Hay barrios más interesantes para los turistas y para los inversores?

Para comenzar un acercamiento a la posible resolución de esta pregunta, se ha calculado la frecuencia de anuncios por barrio, es decir, el número de anuncios por barrio, lo cual permite saber cuáles son los barrios en que más anuncios hay.

Estos son: 

Destaca con mucha diferencia el barrio “Centro” de Málaga, con una amplia diferencia con respecto a al segundo y al tercero en posición, los cuales son “Este” y “Carretera de Cádiz”.

De forma específica, el número de anuncios por cada barrio es:

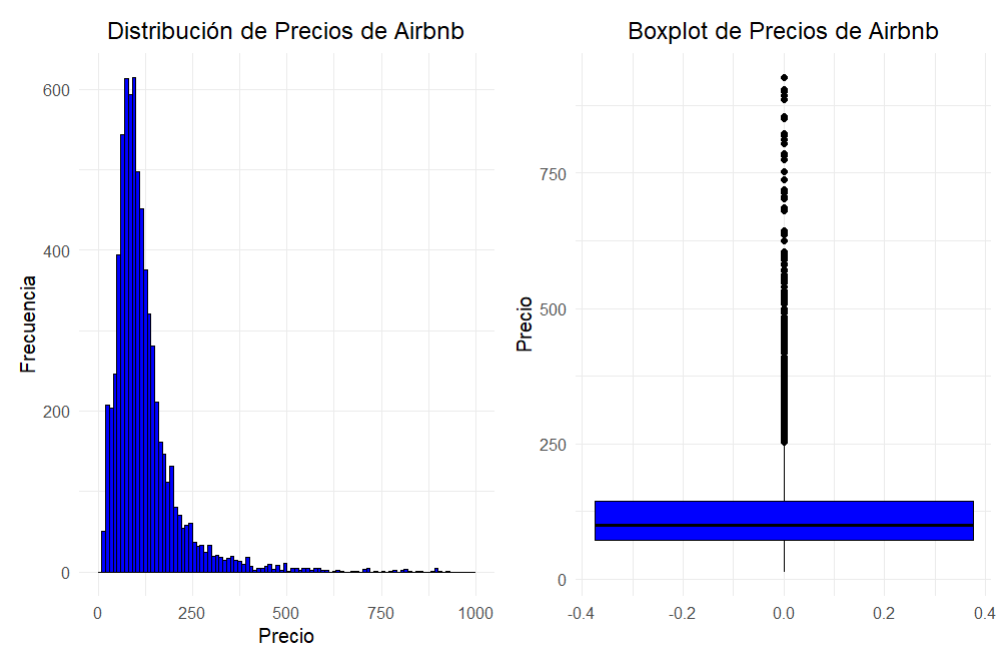


A primera vista, por la amplia diferencia entre el barrio “Centro” con los demás, ya da una primera indicación de que este puede ser el barrio más atractivo para los turistas, y, por ende, para los inversores en vivienda turística.

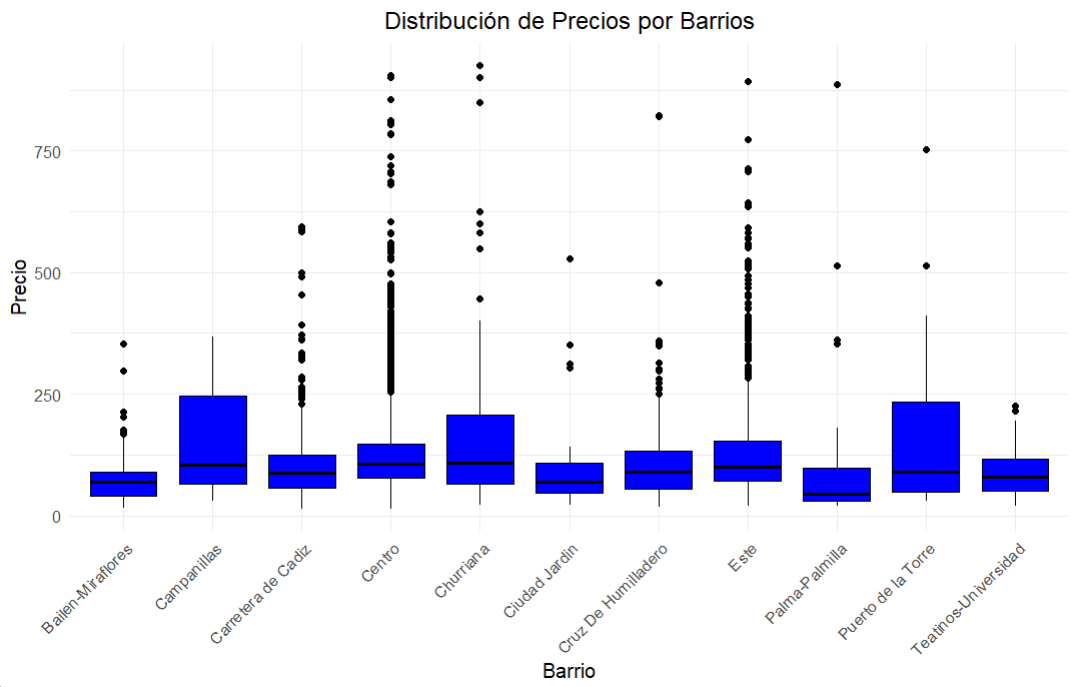
**Análisis de la demanda**

Habiendo observado que existe una amplia diferencia entre el número de anuncios del barrio “Centro” con el resto de los demás, es natural hacerse la siguiente pregunta ¿Se da esta diferencia solo en el número de anuncios? Dado que la naturaleza de este trabajo es de índole económica, el precio será el principal factor a estudiar.

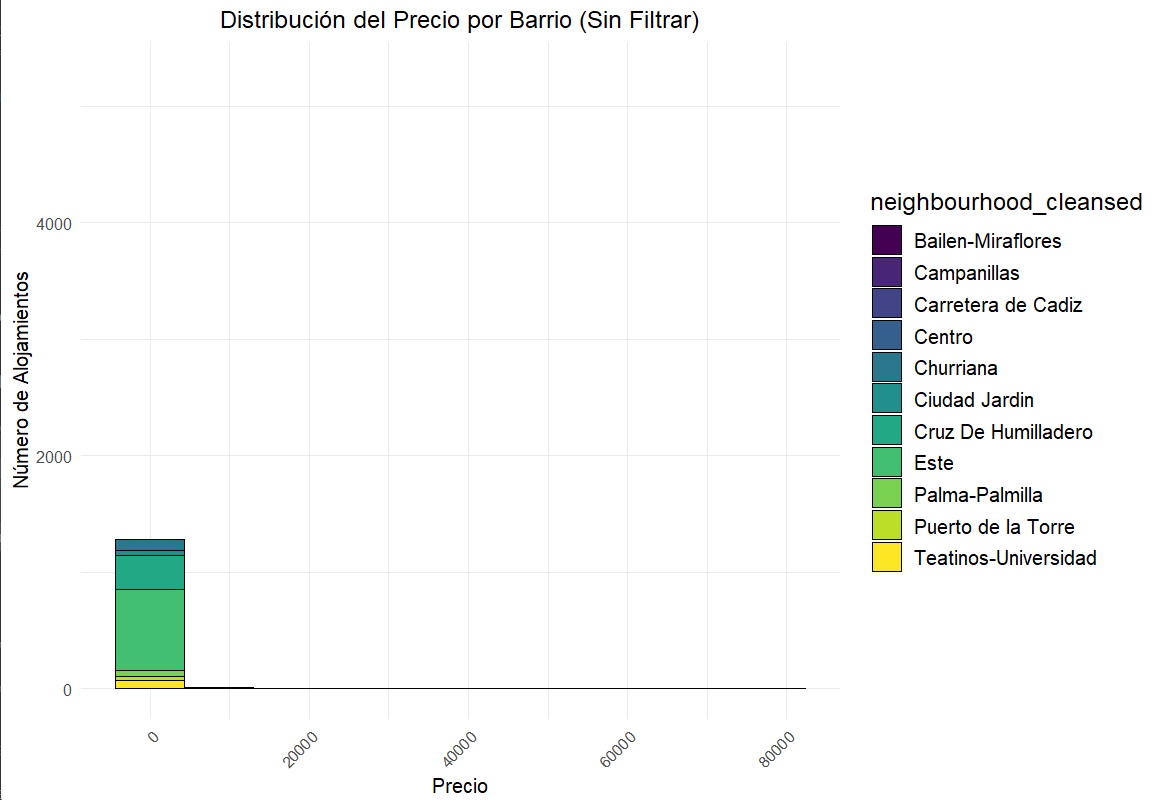
Se comenzará estudiando la distribución del precio ¿Cuáles son los rangos de precios más frecuentes?



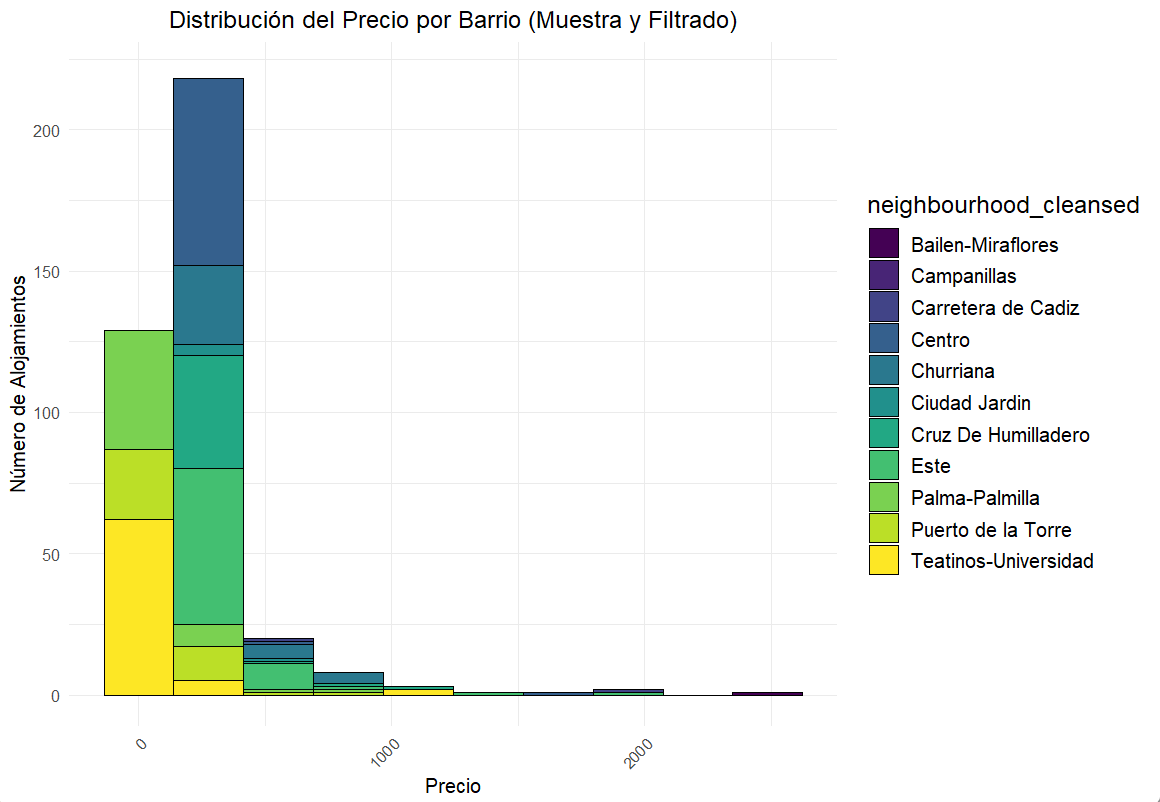
Particularizando a nivel de cada barrio:



Se procede ahora mostrar el precio de las propiedades en cada barrio en función, a su vez del número de anuncios (llamados a este respecto “Alojamientos”)[[1]](#footnote-1)

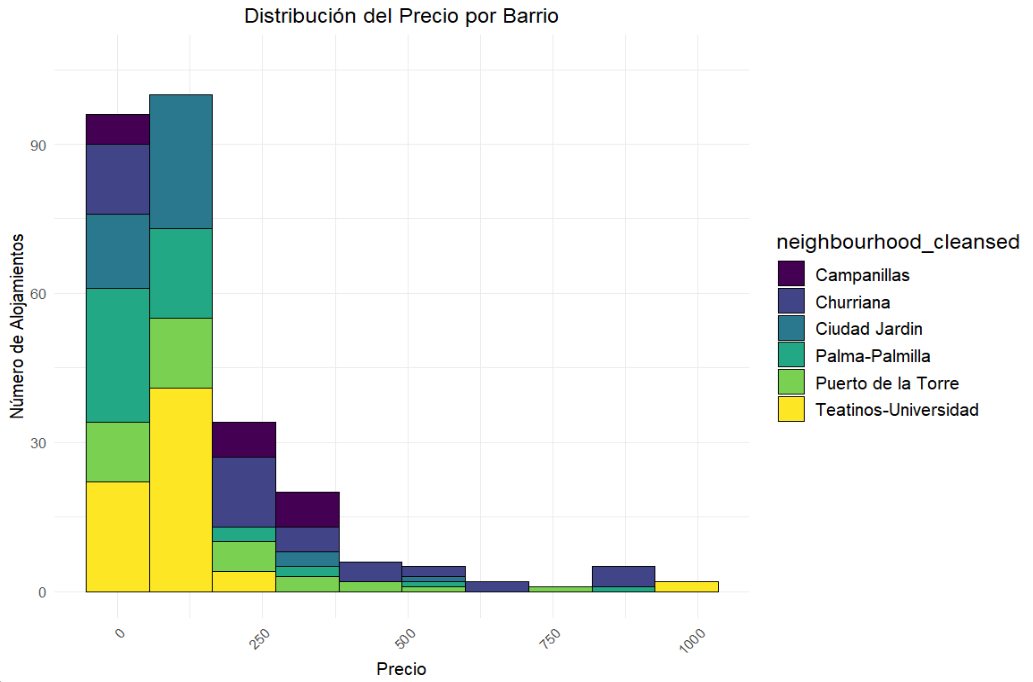


Poniendo un límite de 200 alojamientos:



Destaca aquí el caso de Teatinos-Universidad, que es a su vez uno de los barrios con precios más bajos y altos.

Si se quitan los barrios con más de 200 propiedades se aprecia lo siguiente:



Mientras que incluyendo los barrios con más de 200 propiedades el precio podía llegar a ser, en ocasiones, superior a 2000 euros, retirándolos se da que apenas supera los 1000, así como un suavizado de la distribución, por lo que se puede deducir que no solo que en los barrios con más de 200 alojamientos ofertados es donde hay más oferta, sino que los precios son más altos.

Aun así, en todos los casos las distribuciones son muy marcadas a la izquierda. No olvidar que se está hablando del precio de los alojamientos por noche.

**Modelo Logit: Considerando el precio alto**

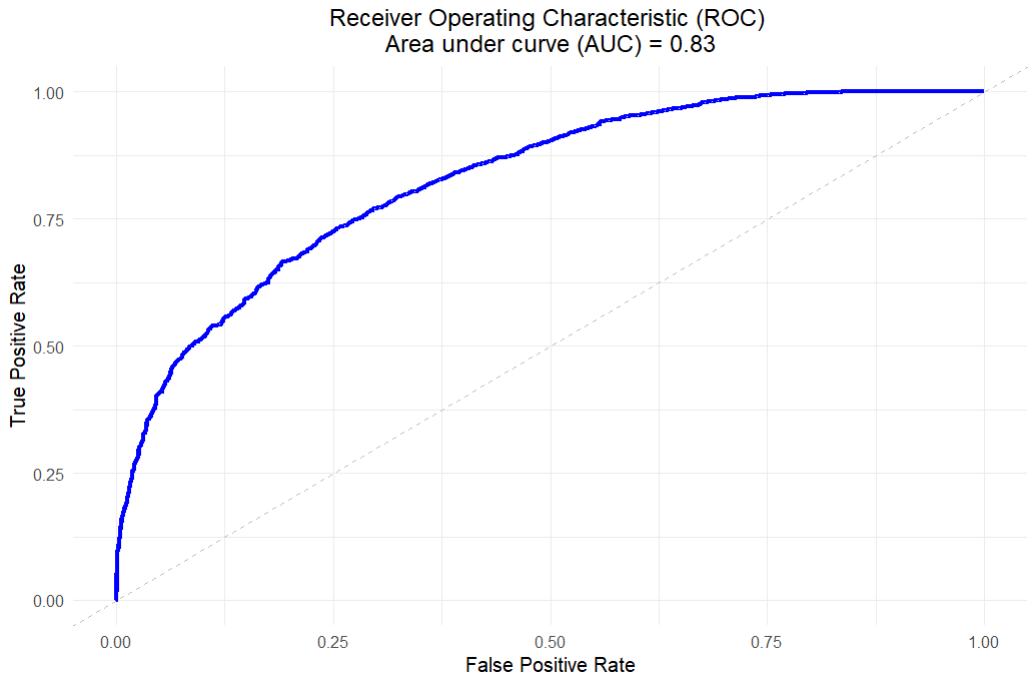
Se procede a crear un modelo Logit, considerando como variable binaria “Precio alto”.

La variable la defino de forma que vale 1 si el precio del alojamiento es mayor a la mediana del precio y 0 si es menor al mismo.[[2]](#footnote-2)

Las variable significativas que devuelve son:

* Accomodates
* Host\_response\_time
* Bathrooms
* Beds
* Host\_listings\_count
* Host\_has\_profile\_pic
* Minimum\_nights
* Maximum\_nights

Para saber si este modelo es un modelo logit correcto, tras calcular sus predicciones grafico su curva de ROC:

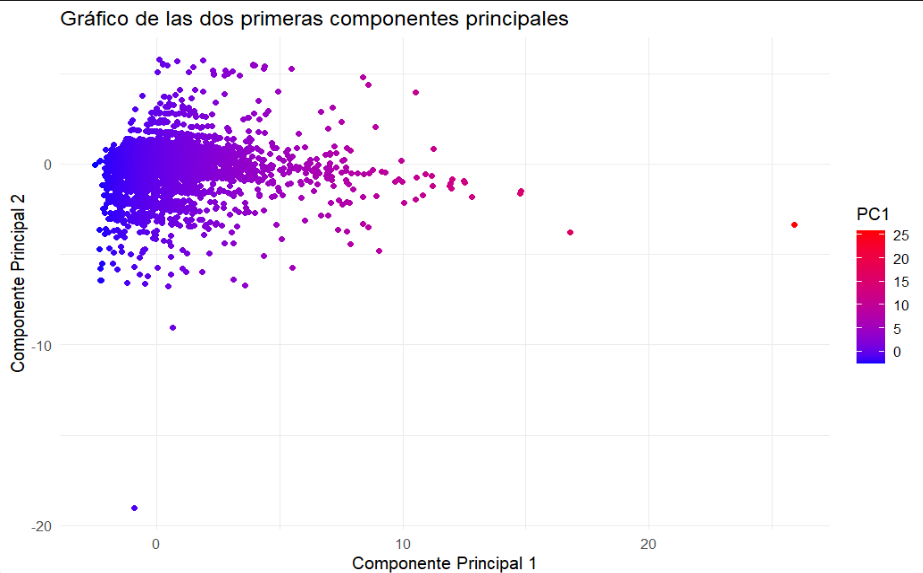


Un AUC de **0.83** indica que el modelo tiene una buena capacidad para distinguir entre las clases, aunque no es perfecto. Significa que, en promedio, el modelo ordenará una instancia positiva por delante de una negativa en el 83% de los casos.

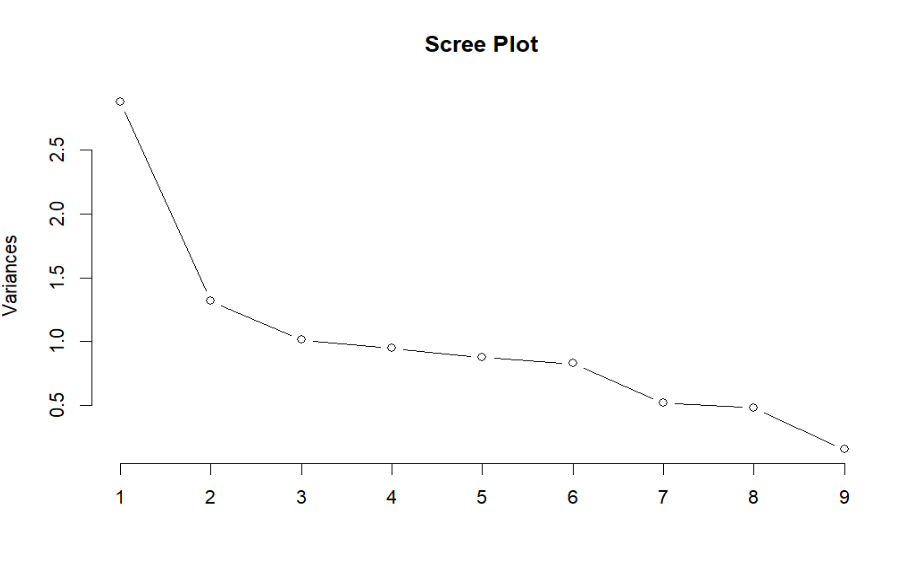
**Segmentación del mercado**

Las variable significativas devueltas por el modelo Logit también me pueden dar una indicación de cuales son los clúster en los que se acumulan los precios más altos.

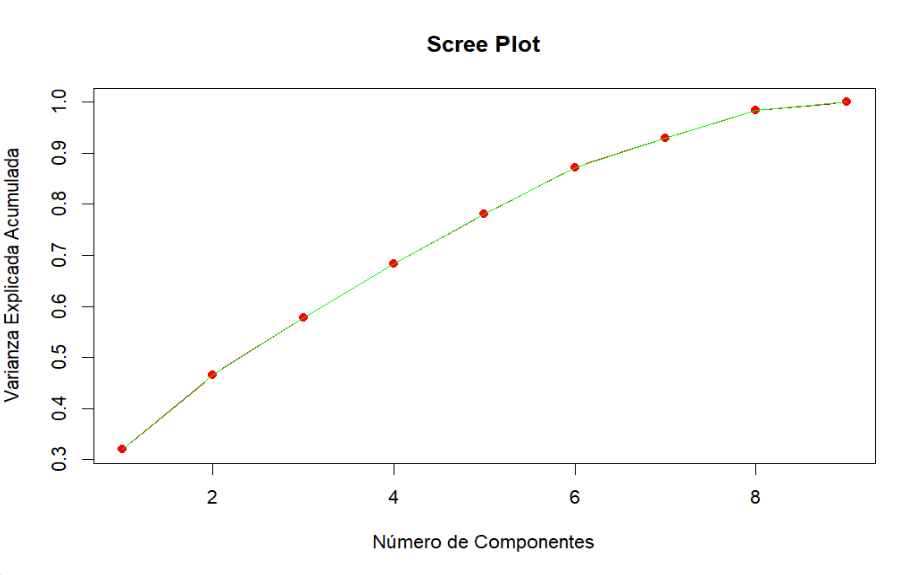
Seleccionándolas para sacar los clúster y sus componentes principales se tiene:



Varianza explicada

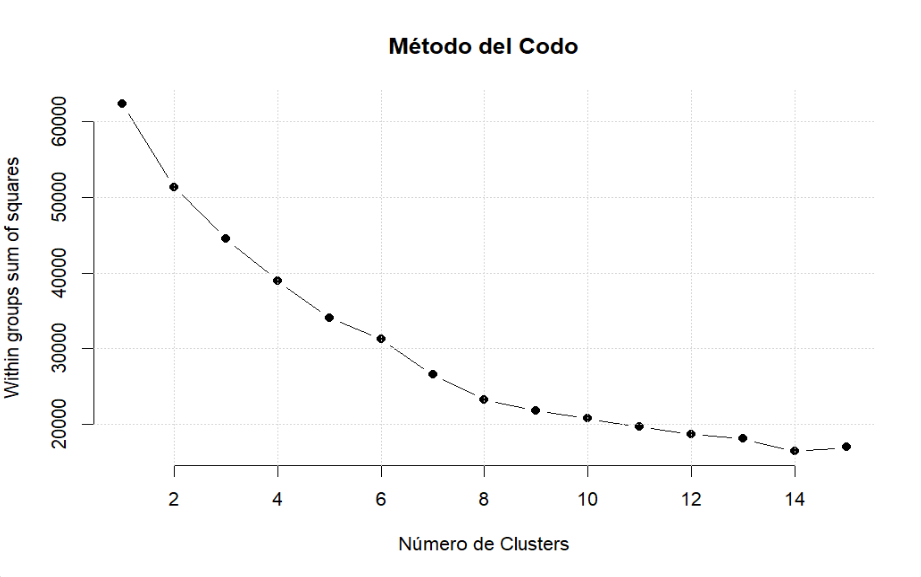


Varianza explicad acumulada



**Clústers**

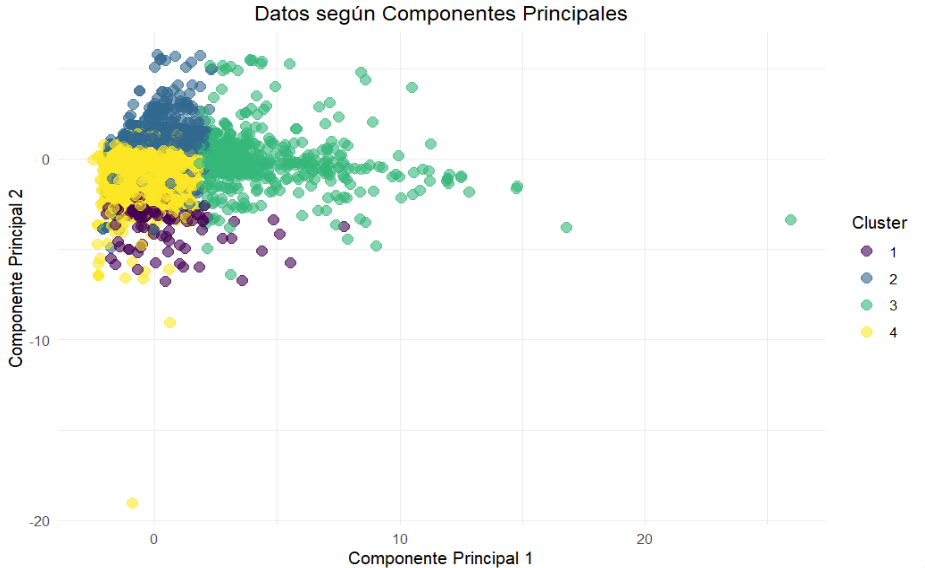
Con el método del codo, procedo a calcular el número óptimo de clústers en que agrupar los datos.



Según el método del codo, el número óptimo de clústers es 1. Sin embargo, con fines de continuación del trabajo se harán 4.

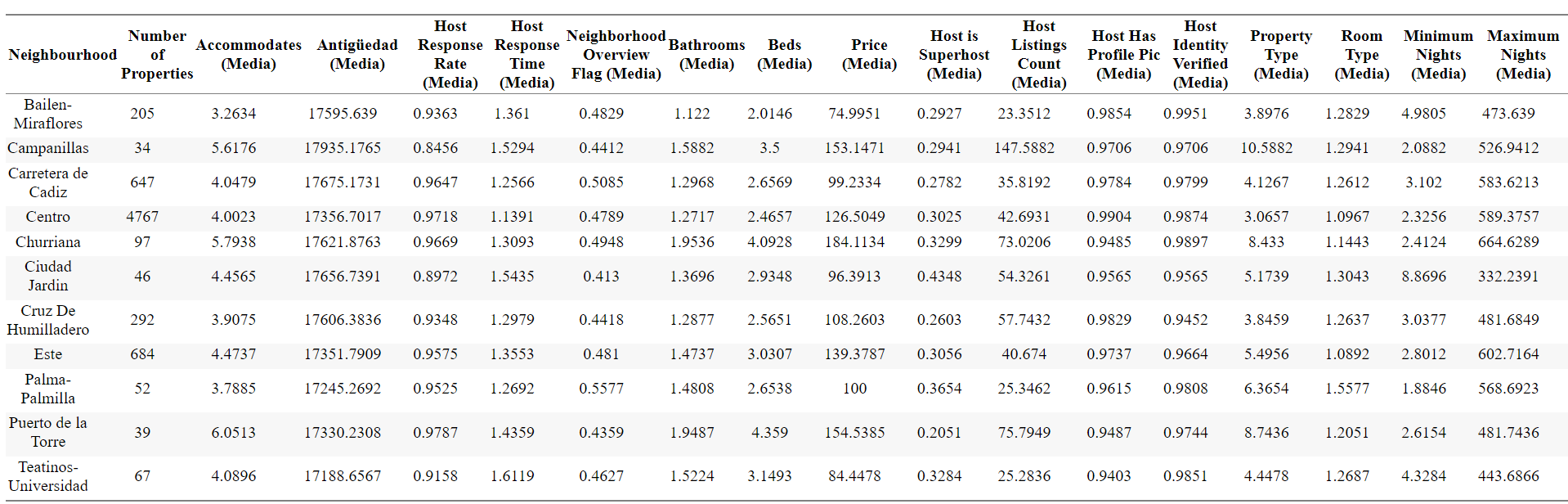
La influencia del barrio “Centro”, siendo el que dispone de un número muy superior de propiedades con respecto a los demás barrios en pferta puede ser una de las explicaciones a que se agrupen los datos en un solo clúster de forma óptima. El peso de este arrastra todos los demás.

Dividiendo en 4 clúster se tiene:

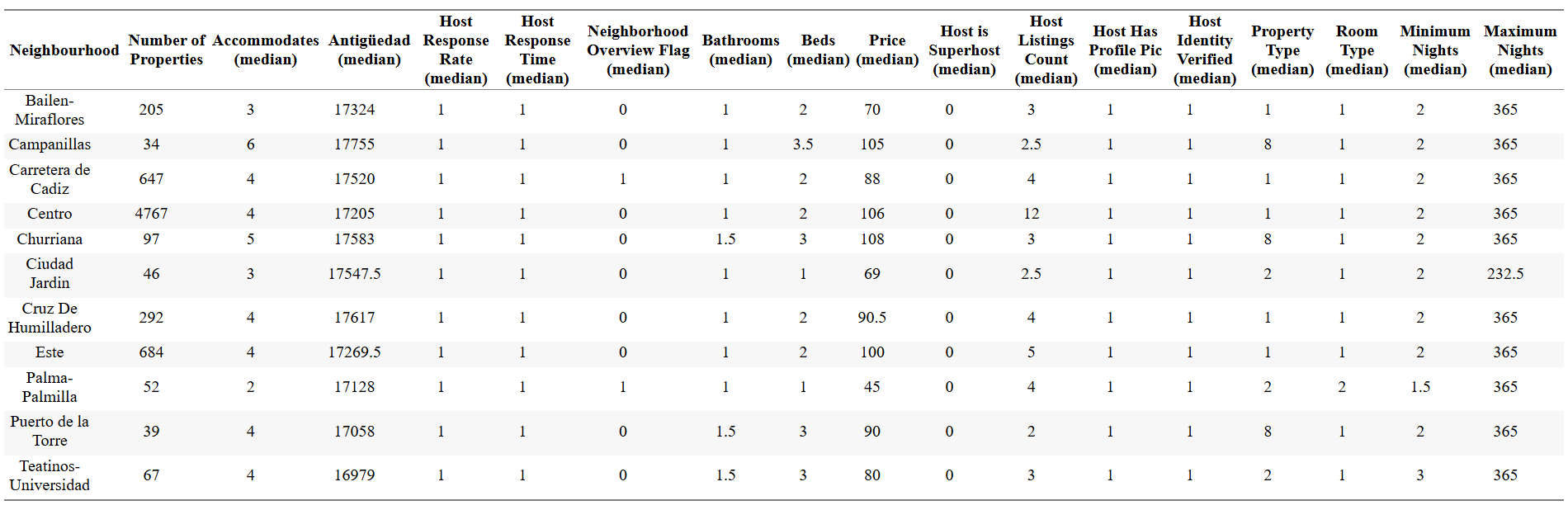


ANEXO I – DESCRIPTIVOS POR BARRIOS

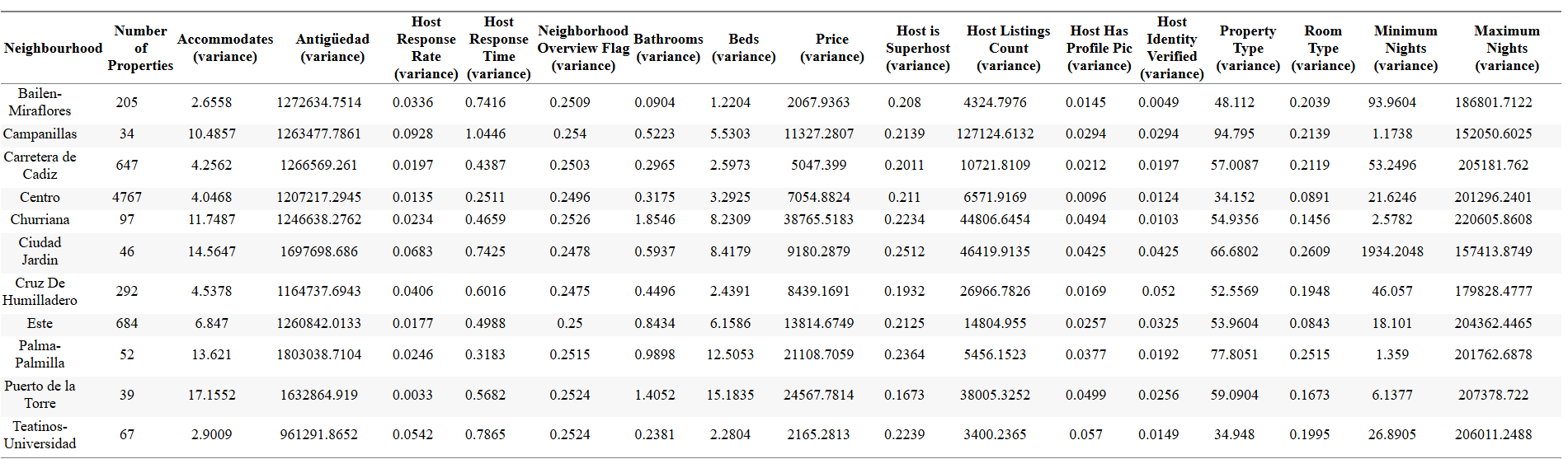
Media por barrios de cada variable



Mediana por barrios de cada variable



Varianza por barrios de cada variable



1. Se ha establecido un límite de 200 alojamientos, por lo que de los barrios que se exceden se toma ha tomado una muestra aleatoria de 200 y es con lo que se trabaja. [↑](#footnote-ref-1)
2. El summary del modelo se encuentra en el archivo logit\_model\_summary.txt que se puede encontrar en el github. [↑](#footnote-ref-2)