

# Herramientas computacionales para la investigación

## Tarea: Data visualization

Tomás Pacheco y Abigail Riquelme

Fecha de entrega: 31/07/2022

### Tarea 1

En el paper de [Schwabish \(2014\)](#) se presentan tres principios básicos que se deben cumplir al realizar gráficos. El primero de ellos dice que se debe mostrar los datos de modo que el lector comprenda la historia que se está contando. Por esto, hay que mostrarlos de la forma más clara posible. Además, no es necesario mostrar todos los datos, sino aquellos que son más relevantes.

El segundo principio dice que hay que ser ordenados. El autor recomienda evitar el uso de elementos que sean innecesarios o distractivos para el lector. El tercer principio plantea que se debe integrar el texto y el gráfico, es decir, hacer que el gráfico esté autocontenido.

Ahora bien, con el objetivo de cumplir con los tres principios anteriormente nombrados, en esta tarea, corregiremos tres gráficos realizados en clase.

En la Figura (1) se presentan los siete histogramas realizados en clase. Podemos decir que no se está cumpliendo el primer principio del *paper* dado que son varios gráficos para incluir en un texto, esto podría resumirse en una menor cantidad de figuras y así mostrar la misma información, con esto estaríamos reduciendo la carga de información visual del lector. Teniendo en cuenta esto decidimos unificar los siete gráficos en uno solo (Figura 2). Además de reducir la información visual para el lector, el tener todos los histogramas en un mismo gráfico permite comparar de mejor manera el monto de los préstamos para diferentes grados de crédito. En este caso elegimos colores con cierta transparencia para que se puedan diferenciar los siete histogramas.

También podemos decir que el gráfico original, Figura (1), no cumple con el tercer principio dado que no tiene título y los nombres de los ejes no son claros, esto también fue corregido en la Figura (2).

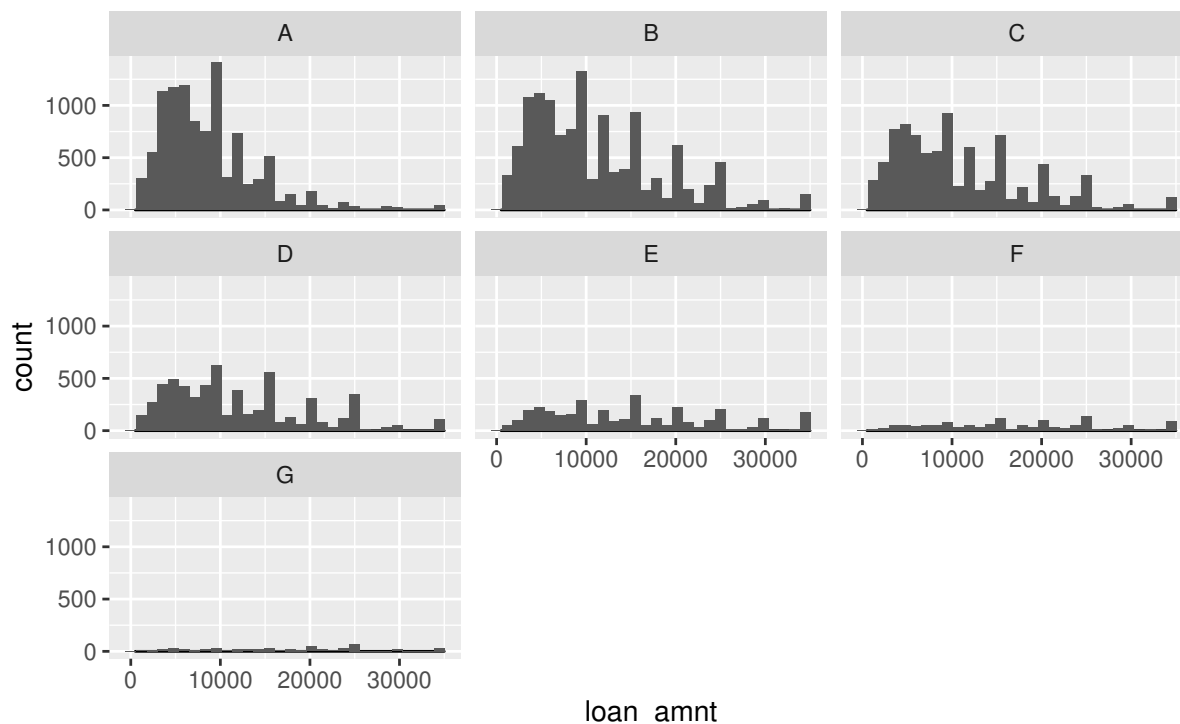
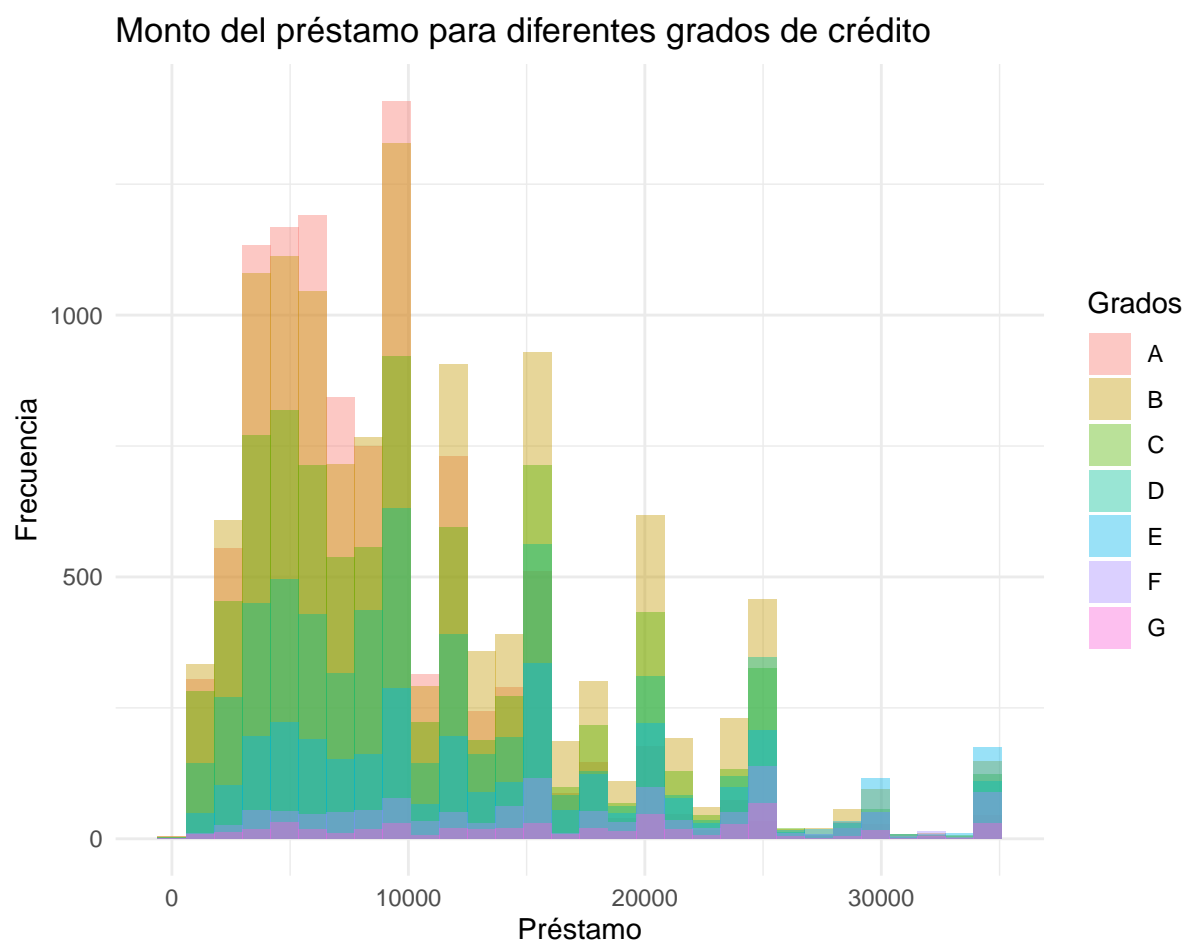


Figura 1: Primer gráfico (original)

En la Figura (3) se observa un gráfico de dispersión entre el consumo de electricidad per cápita y el PBI per cápita. Se diferencia mediante colores entre cinco países distintos. Además se presentan las líneas de tendencia para estos países. Podemos decir que este gráfico cumple con el segundo principio dado que no se identifican elementos distractivos. Ahora bien, al igual que el gráfico anterior, no cumple con el primer ni tercer principio. Con respecto al primer principio podemos decir que la Figura (3) no lo cumple dado que, si suponemos que son de interés los países China e India, no se pueden observar de manera clara estos datos. Con respecto al tercer principio, podemos decir que no lo cumple debido a que no tiene título y los títulos de los ejes no son claros.

Teniendo en vista lo especificado en el párrafo anterior construimos la Figura (4). Para poder cumplir con el primer principio decidimos hacer un zoom para los datos con menor PBI per cápita, de esta forma se pueden observar de manera mas clara los datos de China e India. Luego, para poder cumplir con el tercer principio agregamos un título que sea informativo y corregimos los nombres de los ejes.

En la Figura (5) presentamos el tercer gráfico de clase elegido para este trabajo. Este es



Fuente: elaboración propia

Figura 2: Primer gráfico (modificado)

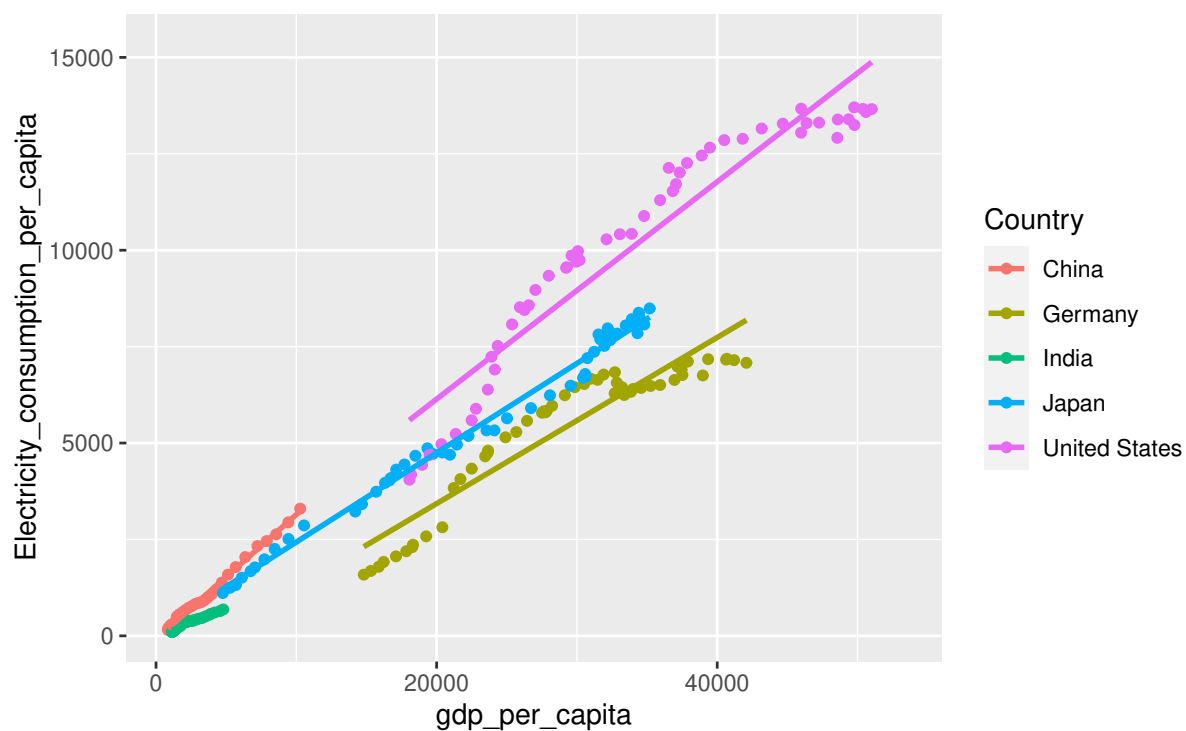


Figura 3: Segundo gráfico (original)

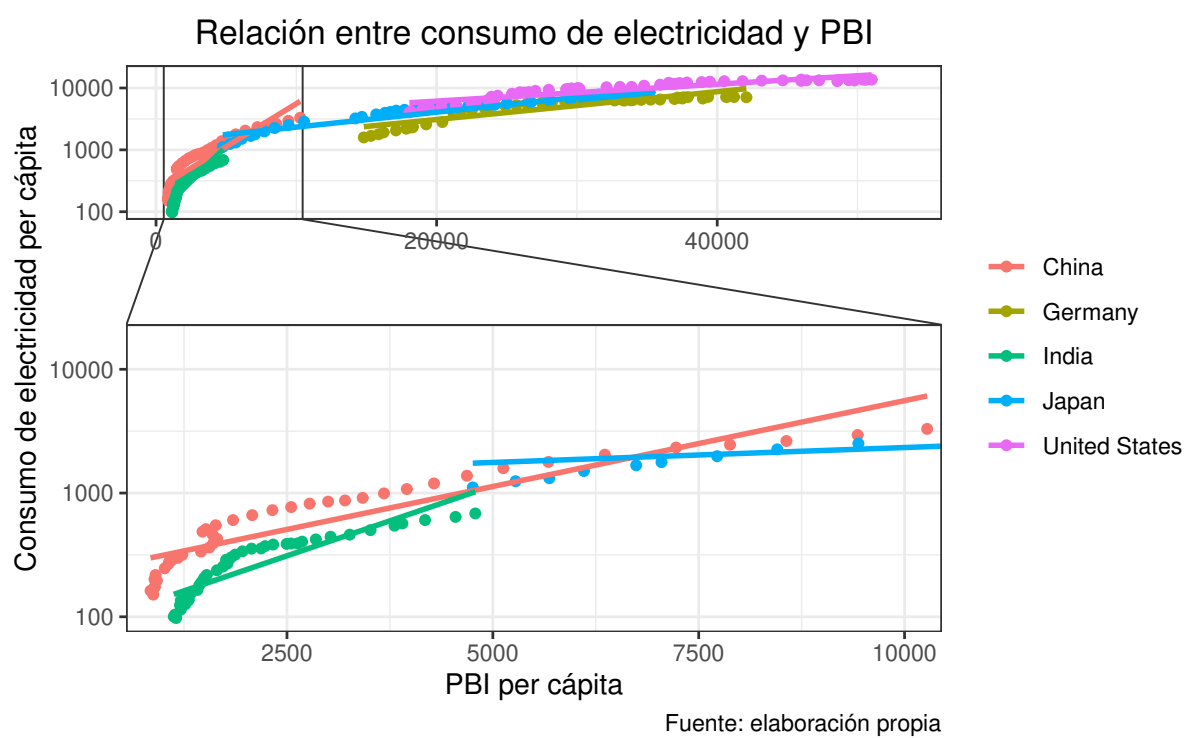


Figura 4: Segundo gráfico (modificado)

un gráfico de dispersión entre el precio de cierre de Facebook (USD) y el mes del año. Además, se agrega el promedio en cada mes. En primer lugar, es importante notar que en ningún momento se aclara que tenemos datos de los años 2017 y 2018, es por esto que podemos decir que dicha figura no cumple con el primer principio. En cuanto al segundo y tercer principio es posible ver que se cumplen dado que no hay elementos distractivos y el gráfico está, en general, autocontenido.

Ahora, para mejorar la claridad de la Figura (5) decidimos diferenciar mediante colores aquellos datos que corresponden al año 2017 y aquellos que corresponden al año 2018. Además, incluimos el dato del promedio del precio de cierre de Facebook para cada año.

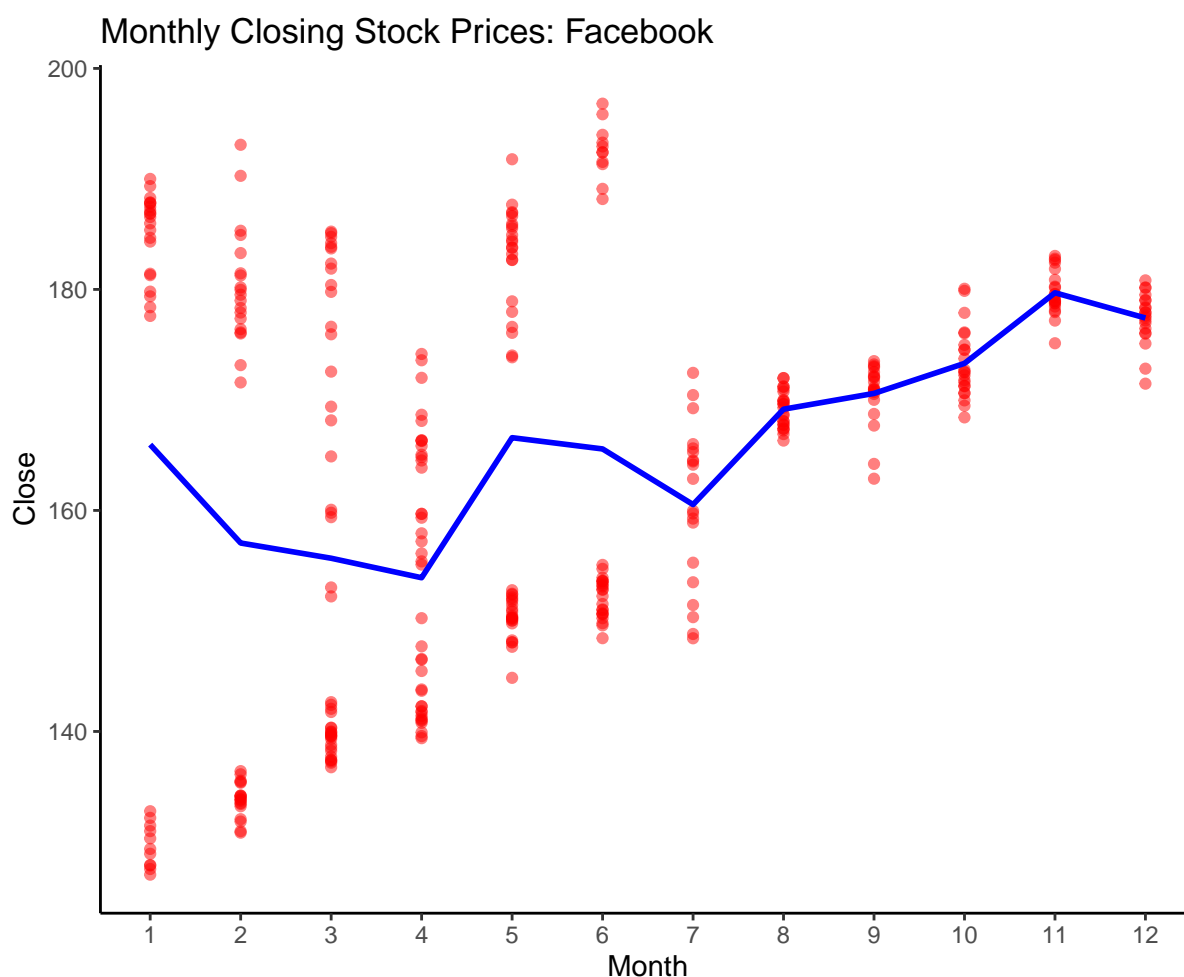


Figura 5: Tercer gráfico (original)

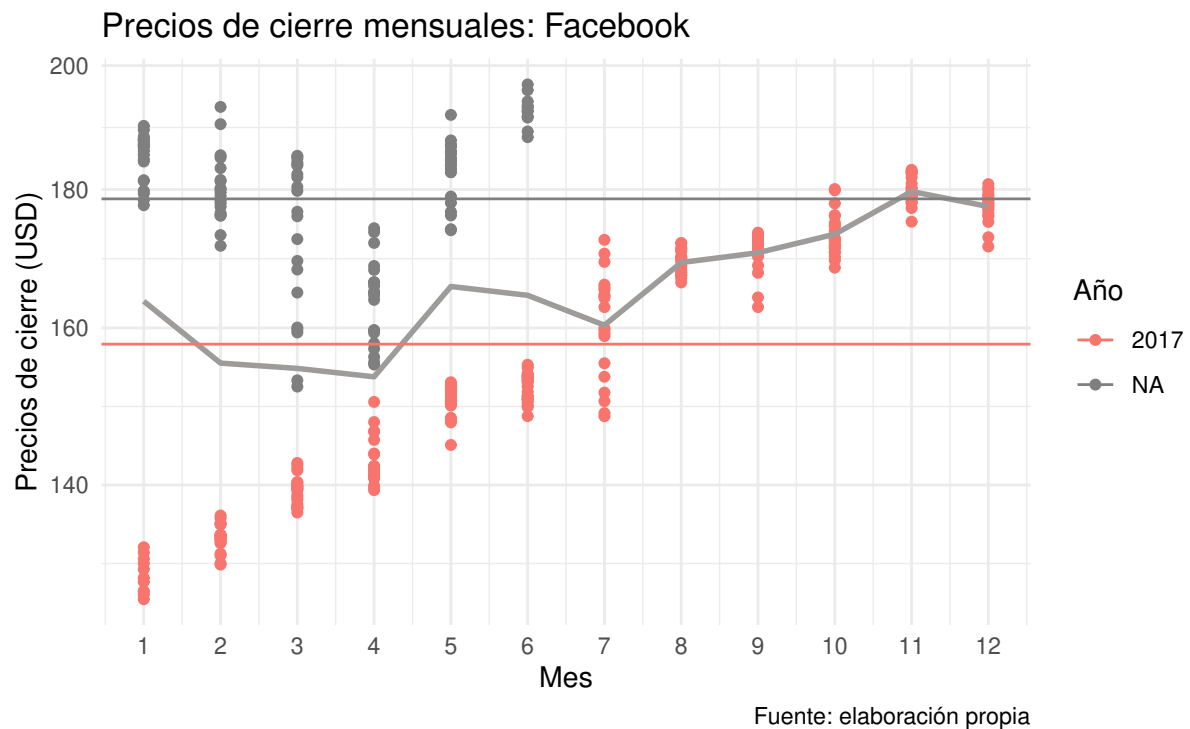


Figura 6: Tercer gráfico (original)

## Tarea 2

La segunda tarea de esta semana tiene como objetivo que confeccionemos un mismo mapa de tres formas distintas. Lo que vamos a mostrar es la cantidad de robos cada 1000 habitantes para la ciudad de Londres, desagregado por barrio. Como se dijo, la única diferencia que tienen los mapas es que fueron hechos de distintas maneras en cuanto a paquetes. La Figura (7) muestra el mapa hecho con ggplot2 en R, la Figura (8) muestra el mapa hecho con el paquete tmap en R y la Figura (9) muestra el mapa hecho con spmap en Stata.

Robos cada 1000 habitantes en cada barrio de Londres

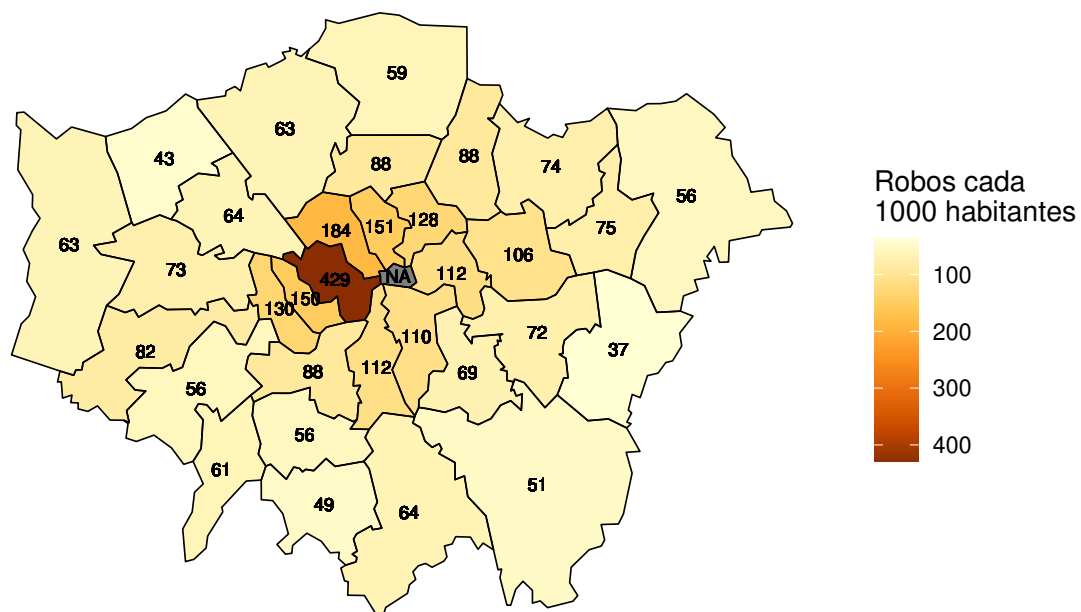


Figura 7: Mapa realizado con ggplot (R)

## Robos cada 1000 habitantes en cada barrio de Londres

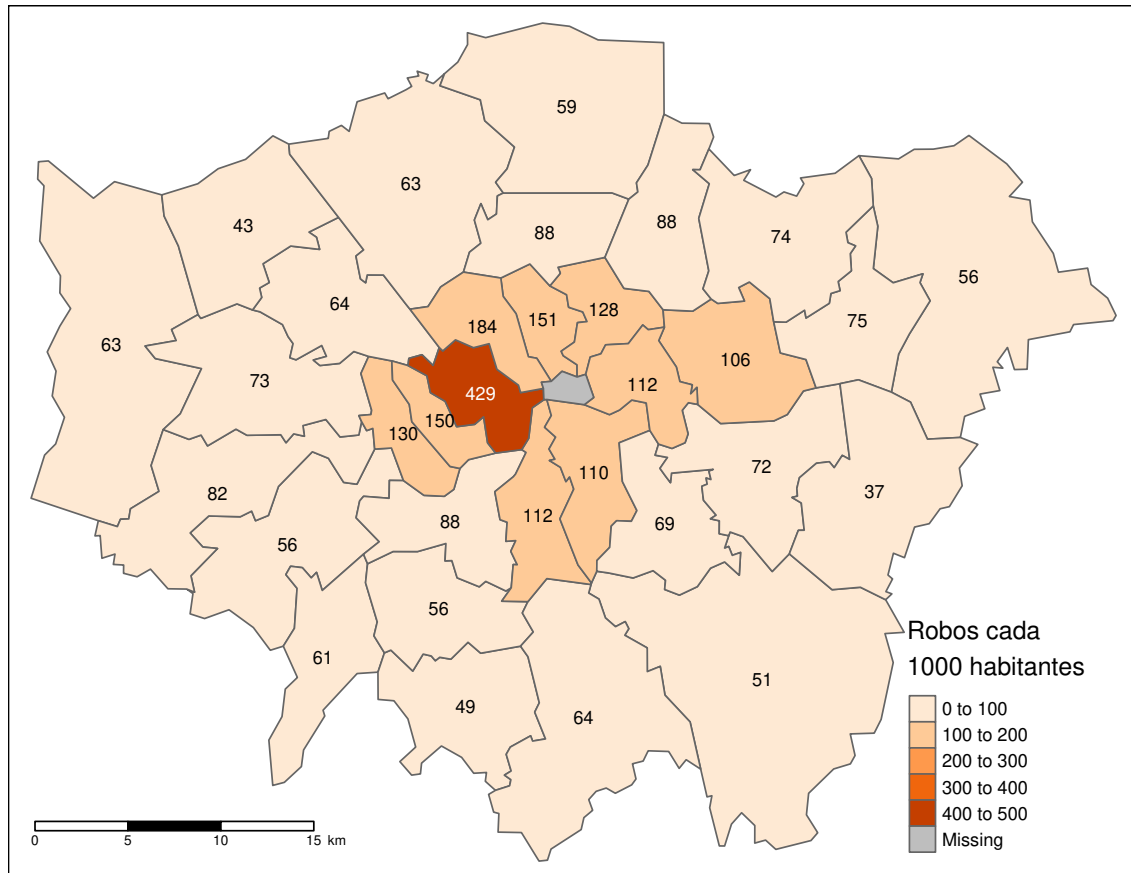


Figura 8: Mapa realizado con tmap (R)

Como se puede observar en estos dos mapas, los robos cada 1.000 habitantes son mayores en el centro de la ciudad. Se muestra que Westminster es el barrio con mayor criminalidad, con aproximadamente 429 crímenes cada 1.000 habitantes. En las áreas periféricas de la ciudad, se observa una menor cantidad de crímenes. En el centro se puede ver un barrio, City of London, con el que no contamos con información. La escala del gráfico es un gradiente que toma colores claros para los valores menores y colores oscuros para los valores mayores.



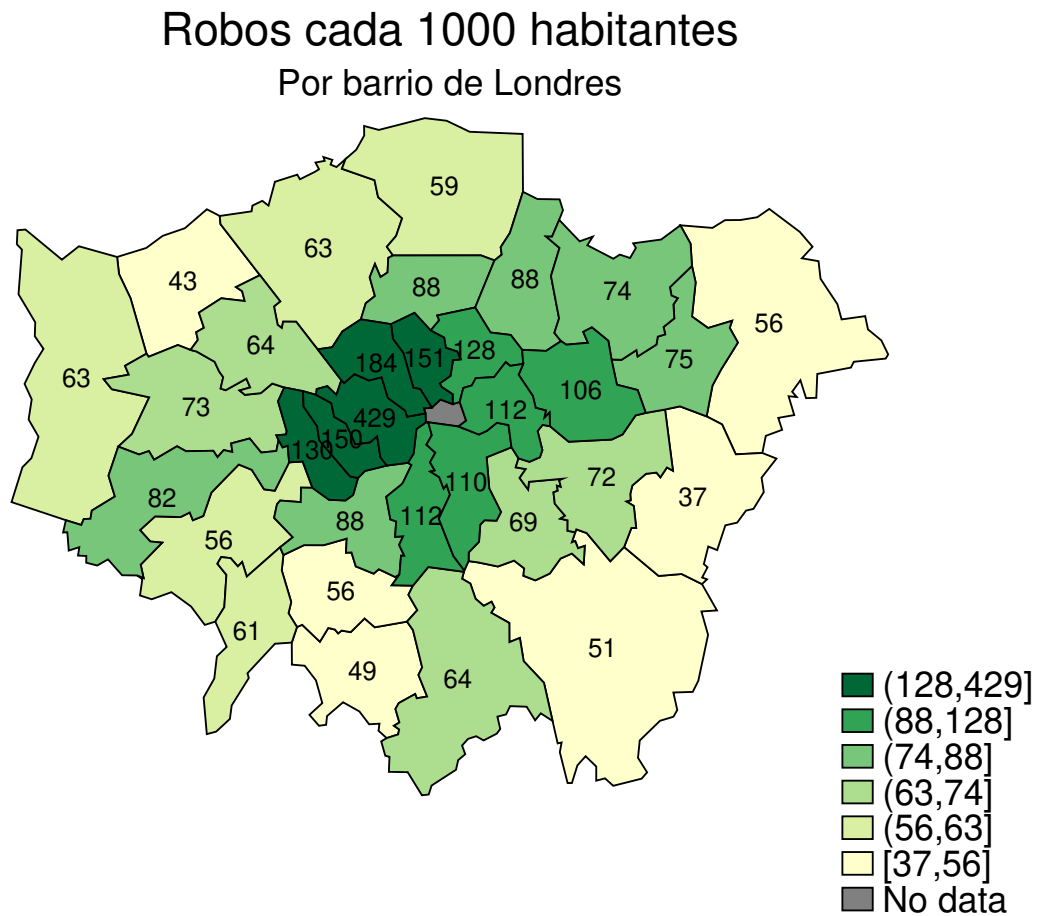


Figura 9: Mapa realizado con spmap (Stata)

Las conclusiones para este mapa en términos de la distribución de los crímenes es exactamente igual a los anteriores. La única diferencia es que la escala está dividida en seis cuantiles. Gracias a esta división, se puede apreciar un tanto mejor el hecho de que la mayor cantidad de crímenes se concentran en torno al centro de la ciudad.

## Referencias

Schwabish, J. A. (2014). An economist's guide to visualizing data. *Journal of Economic Perspectives*, 28(1), 209–34.