

# Herramientas computacionales para la investigación

## Tarea: Scraping & QGIS

Tomás Pacheco y Abigail Riquelme

Fecha de entrega: 17/07/2022

### Ejercicio 1

En el problem set de esta semana utilizamos distintos tipos de datos para el estado de Maryland. En primer lugar, obtuvimos datos sobre crímenes para este estado desegregado a nivel de condado. En segundo lugar, obtuvimos datos del clima para cada uno de los condados durante un año a través de la técnica de scraping. El desafío de este trabajo fue interactuar datos obtenidos de diferentes maneras para colapsarlos en una base y trabajar con ellos.

Primeramente, la consigna solicitaba que grafiquemos las precipitaciones y los cuatro tipos de crímenes que tiene nuestra base de datos. Contamos con los crímenes categorizados de la siguiente manera: robos, que incluyen aquellos actos en los que se sustrae la propiedad de otra persona utilizando la violencia; hurtos, que incluye la sustracción de la propiedad sin el uso de la violencia; entraderas, es decir, la entrada al domicilio de otra persona; y asaltos que son directamente un ataque físico.

En los siguientes gráficos, lo que haremos es presentar la correlación que hay entre la cantidad de robos de cada tipo y la cantidad de milímetros que llovió en cada condado por mes. Nuestra base de datos está a nivel condado-mes y solo tenemos información para el año 2015. En otras palabras, tenemos doce observaciones para cada uno de los condados.

En la Figura (1) podemos observar un gráfico de dispersión entre la cantidad de hurtos cada 100.000 habitantes y los milímetros llovidos en promedio por mes. Al visualizar la

distribución de los puntos podemos notar que, si bien hay mucha varianza, es posible identificar hay una relación positiva entre estas dos variables, es decir, a mayor nivel de precipitaciones, mayor cantidad de hurtos cada 100.000 habitantes.

### Relación entre precipitaciones y hurtos por condado-mes en Maryland

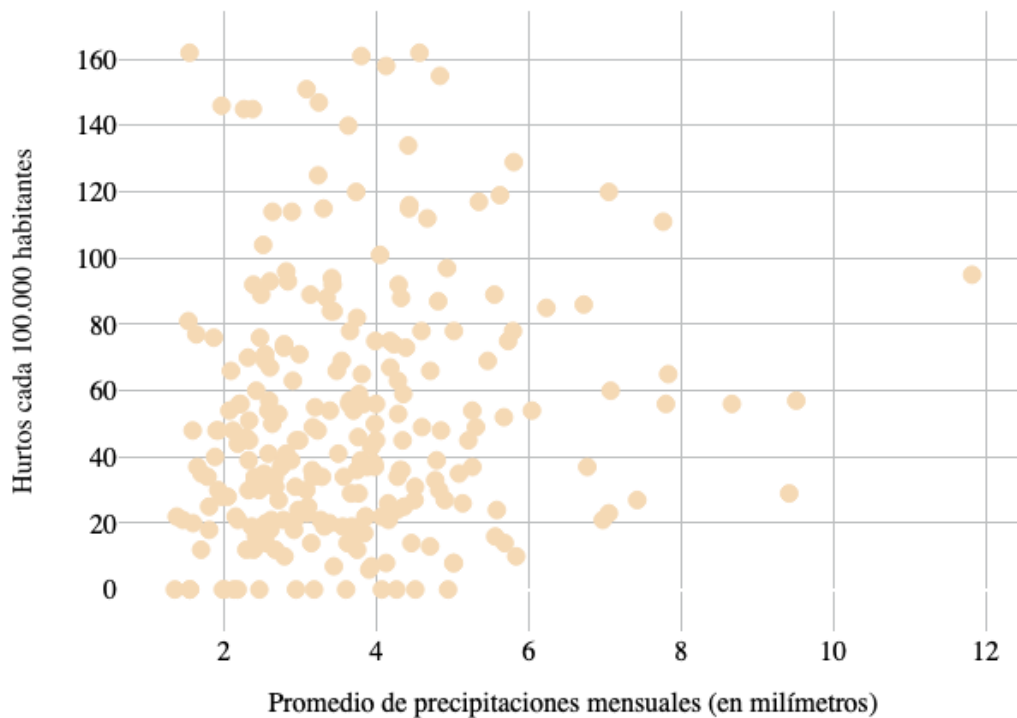


Figura 1: Relación entre precipitaciones y hurtos

En el segundo gráfico, Figura (2), mostramos la relación entre los milímetros de precipitaciones y la cantidad de asaltos cada 100.000 habitantes. Esto lo hicimos a través de un histograma 2D. Este tipo de gráfico lo que hace es mostrar la frecuencia de observaciones para un nivel de precipitaciones, dado un nivel de asaltos. Podemos ver que hay mayor concentración de observaciones en los menores niveles. En otras palabras, las observaciones se concentran, en su mayoría, entre los 0 y 6 milímetros de precipitaciones y entre los 0 y 40 asaltos cada 100.000 habitantes. Esto nuevamente nos está dando una idea de una relación positiva entre ambas variables, dado que si a menor precipitación, menor la cantidad de asaltos, entonces, es una relación positiva. Claramente con este gráfico no podemos observar la magnitud de esta relación, pero si podemos interpretar, muy visualmente, la dirección.

## Relación entre precipitaciones y asaltos por condado-mes en Maryland

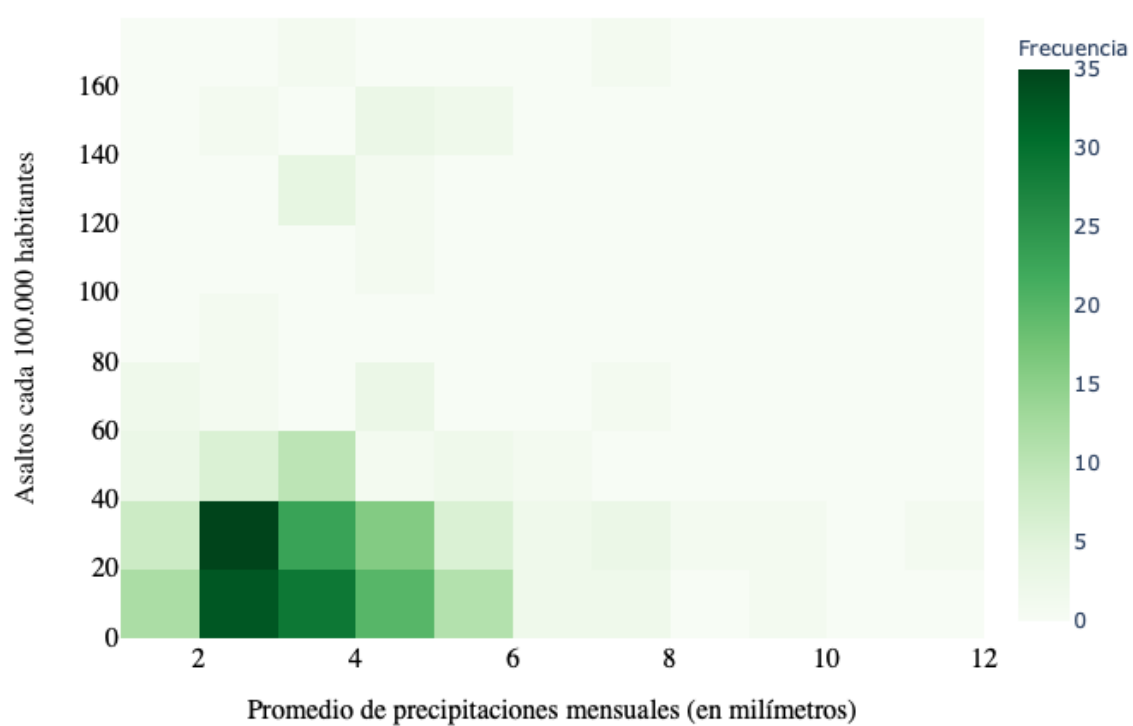


Figura 2: Relación entre precipitaciones y asaltos

La Figura (3) muestra la relación entre las precipitaciones y entraderas. En este caso, dada la dispersión de los datos no podemos sacar muchas conclusiones. No parecería haber una relación clara entre las dos variables.

Relación entre precipitaciones y entraderas por condado-mes en Maryland

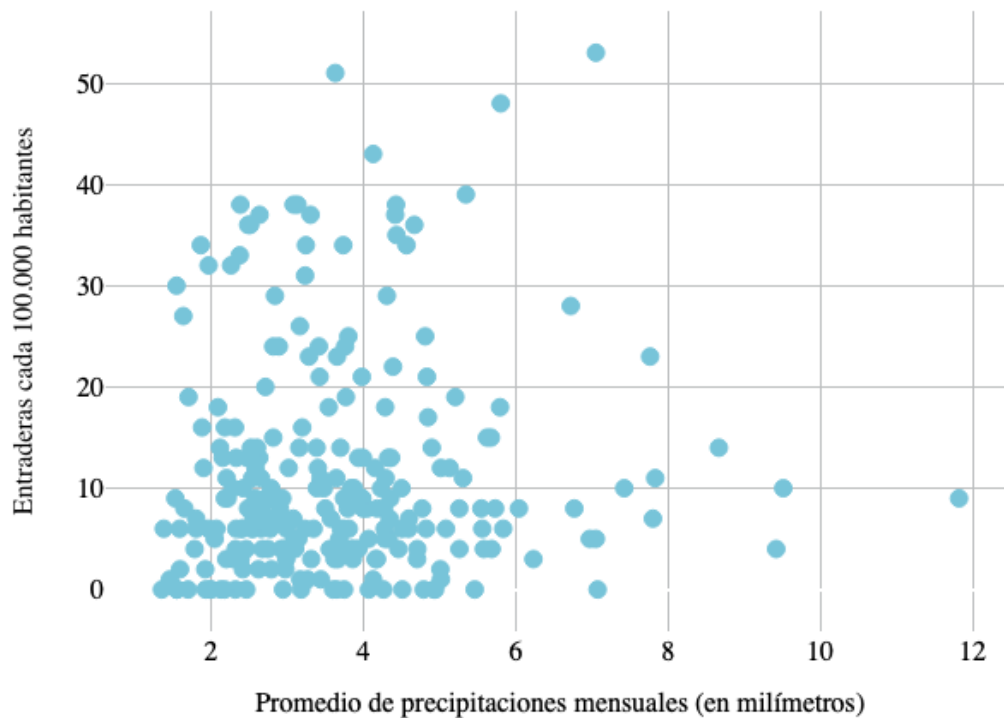


Figura 3: Relación entre precipitaciones y entraderas

Finalmente, en la Figura (4), presentamos un histograma 2D de las precipitaciones y la cantidad de robos cada 100.000 habitantes por condado. Al igual que en la Figura (2), la mayor cantidad de observaciones se encuentra en el segundo cuadrante del gráfico. Esto nos estaría sugiriendo la presencia de una relación positiva. Sin embargo, como los valores se concentran entre 0 y 6 para precipitaciones y 0 y 10 para robos, es decir, hay una mayor dispersión, no se puede asegurar una relación directa positiva.

Si bien reconocemos las limitaciones de estos gráficos (dado que son solo gráficos) y también sabemos que estamos trabajando con variabilidad de panel pero graficando un cross-section, podemos decir que se observa la relación contraria a la mostrada en la literatura existente. Artículos como los de [Cohn \(1990\)](#), [Field \(1992\)](#) y [Jacob, Lefgren, y Moretti \(2007\)](#) explican que generalmente más lluvia está relacionada con menos

## Relación entre precipitaciones y robos por condado-mes en Maryland

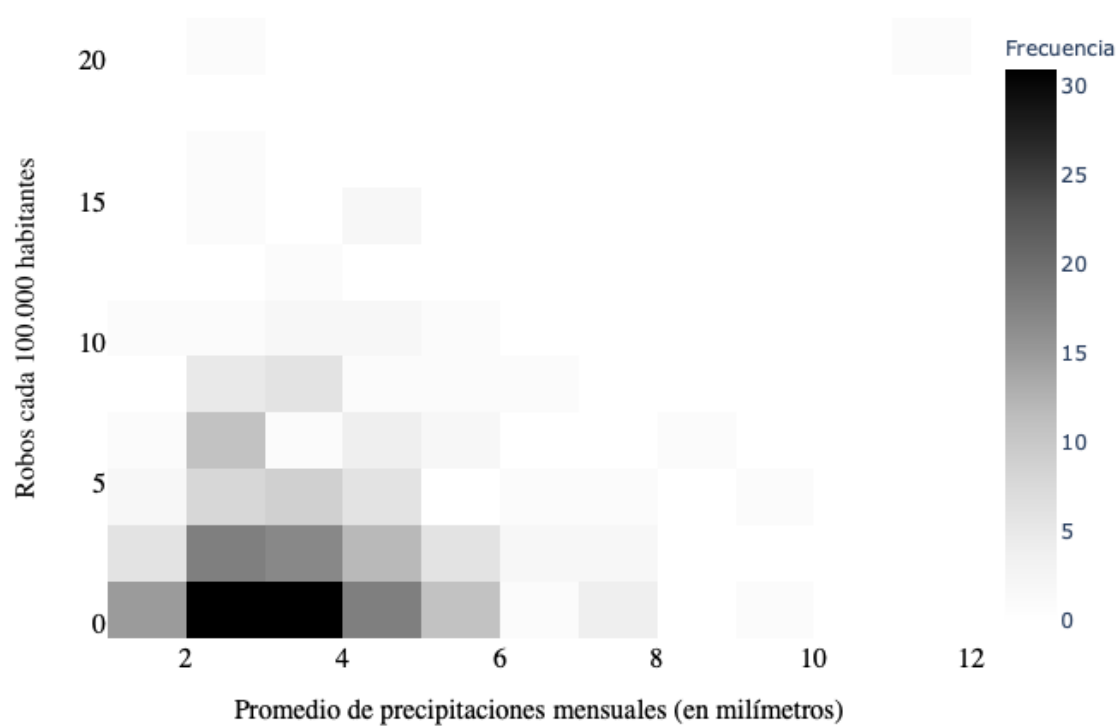


Figura 4: Relación entre precipitaciones y robos

crimen.

## Ejercicio 2

Nuestra base de datos cuenta con una variable que nos dice la cantidad de gente que pertenece a raza negra para cada uno de los condados de Maryland. Este ejercicio lo que pedía es que hagamos un mapa en el que incluyamos esta variable de forma tal que se llegue a una conclusión relevante.

Es por esto que nosotros elegimos graficar el cuartil de la cantidad de crímenes cada 100.000 habitantes y el cuartil de cantidad de habitantes negros cada 100.000 habitantes. Para hacer esto, comenzamos generando cada una de las variables mencionadas. Luego, para cada una identificamos los cuartiles de la distribución y distinguimos a cuál de ellos pertenece cada uno de los condados. Algo que consideramos importante destacar es que la base de datos con la que estamos trabajando es un panel. Entonces, para poder graficar algo estático lo que hicimos fue agregar a nivel estado sumando la cantidad de crímenes. De este modo, tenemos una sola observación para cada uno de los condados. Con esta información, finalmente, clasificamos a cada condado de acuerdo a los cuartiles de ambas variables. De esta forma, nos quedaron un total de 16 categorías. Con este gráfico esperamos ver si hay alguna relación geográfica de una presencia de mayor (o menor) población negra y mayor cantidad de crímenes. El resultado se encuentra en la Figura (5).

Lo que primero se puede ver en el gráfico es que hay una distribución relativamente homogénea en relación a los cuartiles de cada una de las variables. En otras palabras, no se observa un patrón geográfico en donde se vean mayor proporción de crímenes acompañados de una mayor proporción de población negra. Solo seis de veintitres condados de la muestra<sup>1</sup> (Baltimore County, Charles, Howard, Montgomery, Somerset y Wicomico) tienen una cantidad de crímenes y población negra mayor a la media. El valor que más se repite (Calvert, Caroline, St. Mary's y Washington) es aquel en donde los estados pertenecen al cuartil 2 de la cantidad de personas negras y al cuartil 3 de la cantidad de asaltos. En resumen, el gráfico nos muestra que no hay una relación evidente entre mayor criminalidad y mayor cantidad de población negra. Si esto fuera así, observaríamos una mayor presencia de estados que pertenecen a los cuartiles más altos.

---

<sup>1</sup>Maryland tiene 24 condados pero no contamos con información sobre los crímenes para Garret.

Relación entre cantidad de crímenes cada 100.000 y población negra cada 100.000 habitantes por condado en Maryland

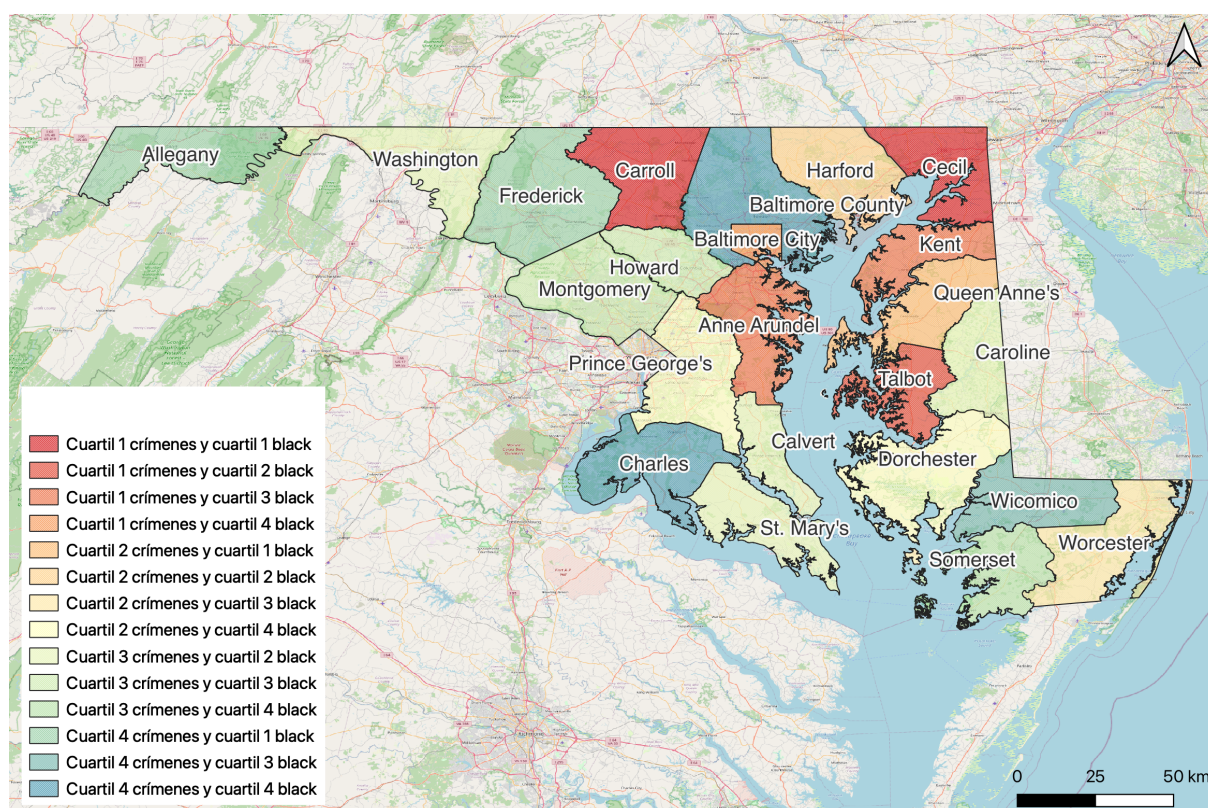


Figura 5: Cuartiles para cada uno de los condados

## Ejercicio 3

Este último ejercicio solicitaba que hagamos dos gráficos dinámicos, que se encuentra en el repositorio de Github. El primero de ellos es el que muestra el avance de las precipitaciones mes a mes durante 2015 para cada uno de los condados. Para esta figura decidimos graficar en seis intervalos fijos de a dos milímetros, para poder identificar la variabilidad a través del tiempo. En cambio, el segundo mapa dinámico muestra la evolución de los robos cada 100.000 habitantes. Dado que esta es una variable que no posee mucha variabilidad, optamos por hacer los cortes de acuerdo al criterio de Jenks que explicamos en el [problem set anterior](#).

## Referencias

- Cohn, E. G. (1990). Weather and crime. *The British Journal of Criminology*, 30(1), 51–64.
- Field, S. (1992). The effect of temperature on crime. *The British Journal of Criminology*, 32(3), 340–351.
- Jacob, B., Lefgren, L., y Moretti, E. (2007). The dynamics of criminal behavior evidence from weather shocks. *Journal of Human resources*, 42(3), 489–527.