

Entry Name: **Edwin Rueda MC1**  
**VAST Challenge 2019**  
**Mini-Challenge 1**

**Team Members:**

Edwin Rueda, Universidade Federal do Pará  
Belém-Pará, ejrueda95g@gmail.com

**Student Team: YES**

**Tools Used:**

- *Python 3.6.5*
- *BokehJS 1.1.0*
- *Matplotlib*
- *Pandas*

**Approximately how many hours were spent working on this submission in total?**

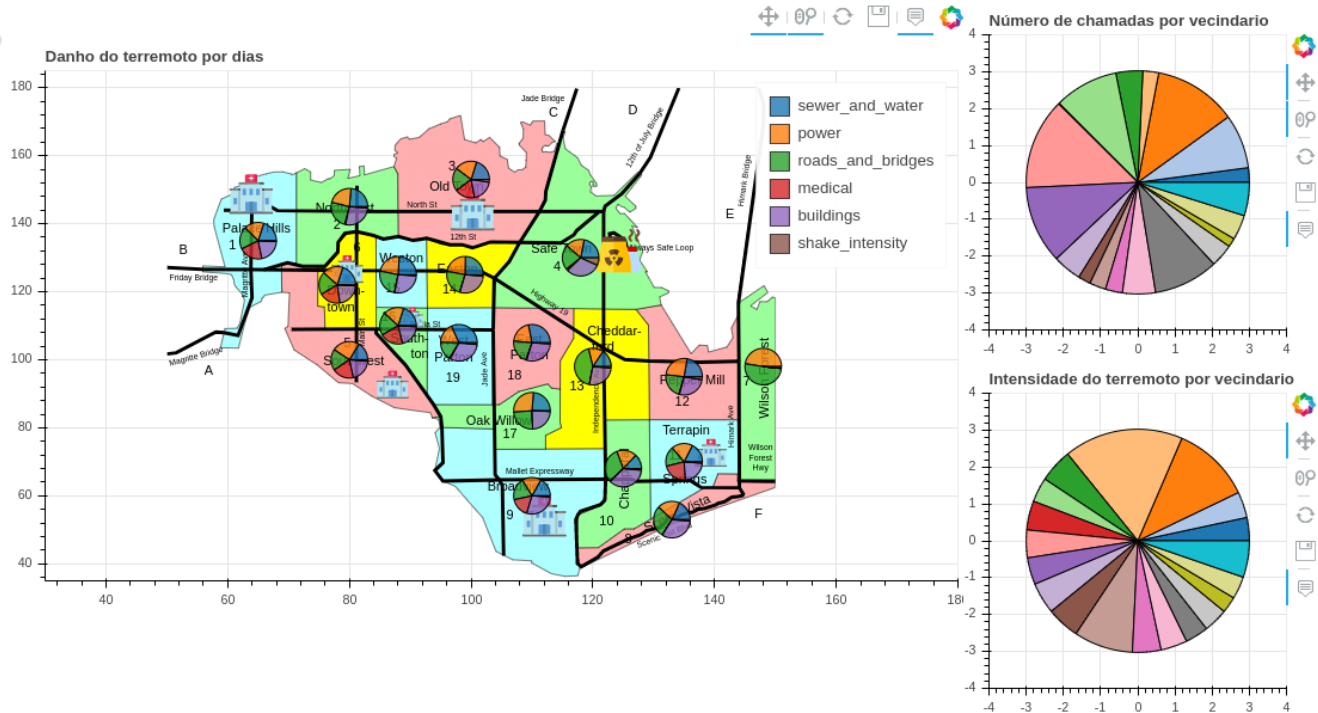
*Provide an estimate of the total number of hours worked on this submission by your entire team. 2H, 2h, 3h, 4h, 3h, 4h*

**Questions**

**1-** Emergency responders will base their initial response on the earthquake shake map. Use visual analytics to determine how their response should change based on damage reports from citizens on the ground. How would you prioritize neighborhoods for response? Which parts of the city are hardest hit? Limit your response to 1000 words and 10 images.

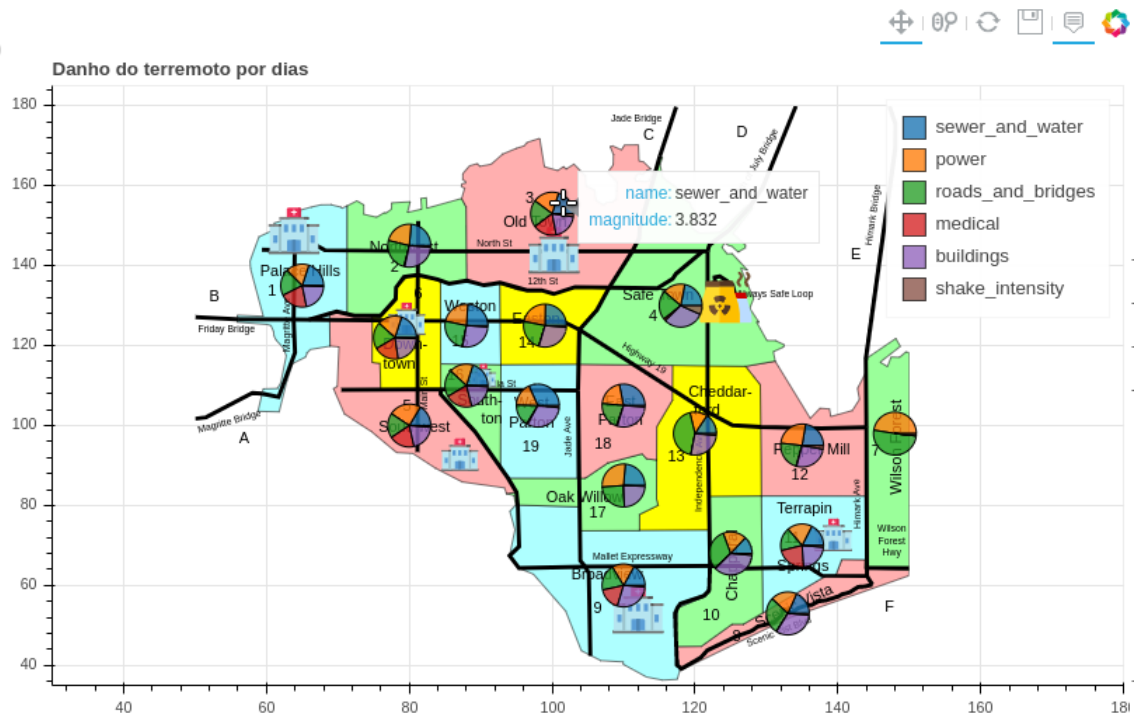
Las siguientes ilustraciones se basan en la interacción con el gráfico 1, este presenta en cada vecindario un diagrama de pizza el cual indica la magnitud del daño en cada tipo reportado, pudiendo iterarse por cada día:

registro de acordo com o dia: 6



En primer lugar se tiene en cuenta una visión global del mapa, la figura 1 muestra la magnitud del daño por cada tipo en cada vecindario, dando a conocer que cosas fueron más afectadas en cada vecindario, sin olvidar, que dicha información es la media diaria del daño presentado.

registro de acordo com o dia: 6



También se puede observar mediante la Figura 2 y 3 el número de llamadas que realizó cada vecindario y la magnitud del terremoto reportada por cada vecindario, teniendo así una relación de llamadas con la magnitud del daño reportado.

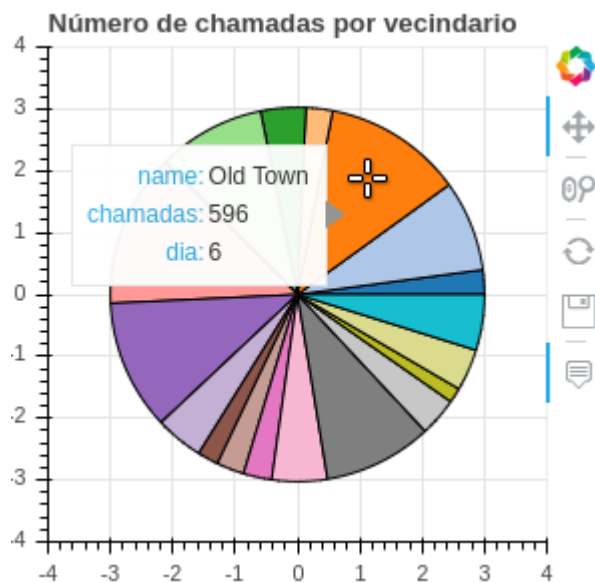


Figura 2. Número de llamadas por vecindario.

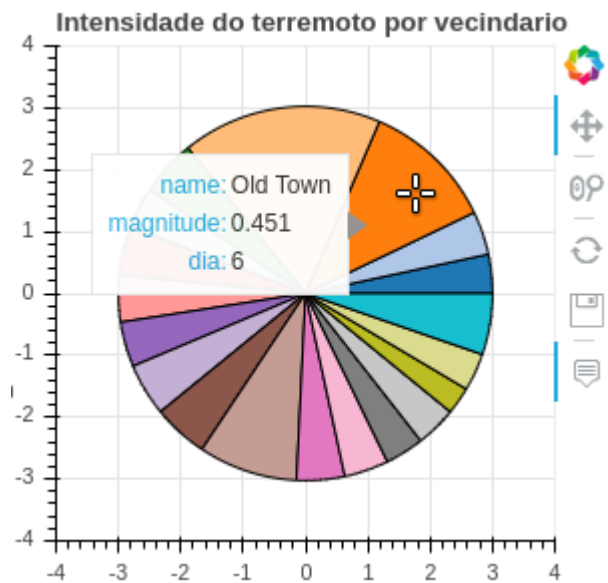


Figura 3. Magnitud del terremoto reportada por vecindario.

Haciendo un zoom a cada día, como lo muestra la Figura 4, se puede ver que cada diagrama de pizza representa el daño que fue causado en cada tipo, viéndose así la relación de daño causado entre cada uno de los tipos en el vecindario.

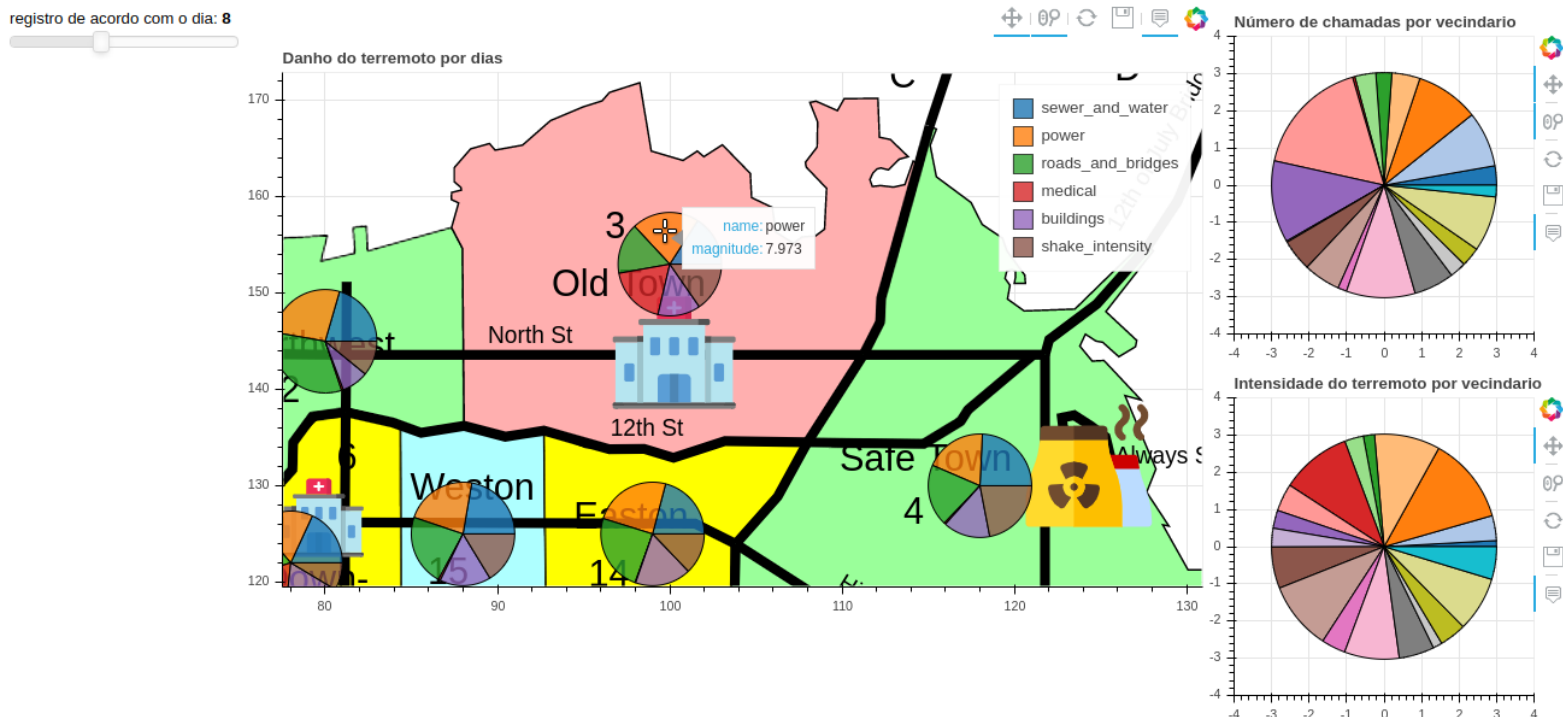


Figura 4. Magnitud de daño por tipo en cada vecindario.

Se podría asociar en un principio el número de llamadas con la urgencia de ser atendidos, la Figura 2 muestra esa proporción, mas sin embargo, se puede ver en la figura 5 que aunque en Wilson Forest la magnitud del terremoto reportada fue alta, no hubieron demasiadas llamadas ya que es una reserva natural, entonces se puede decir que no hay muchas personas, por ende la prioridad es menor.

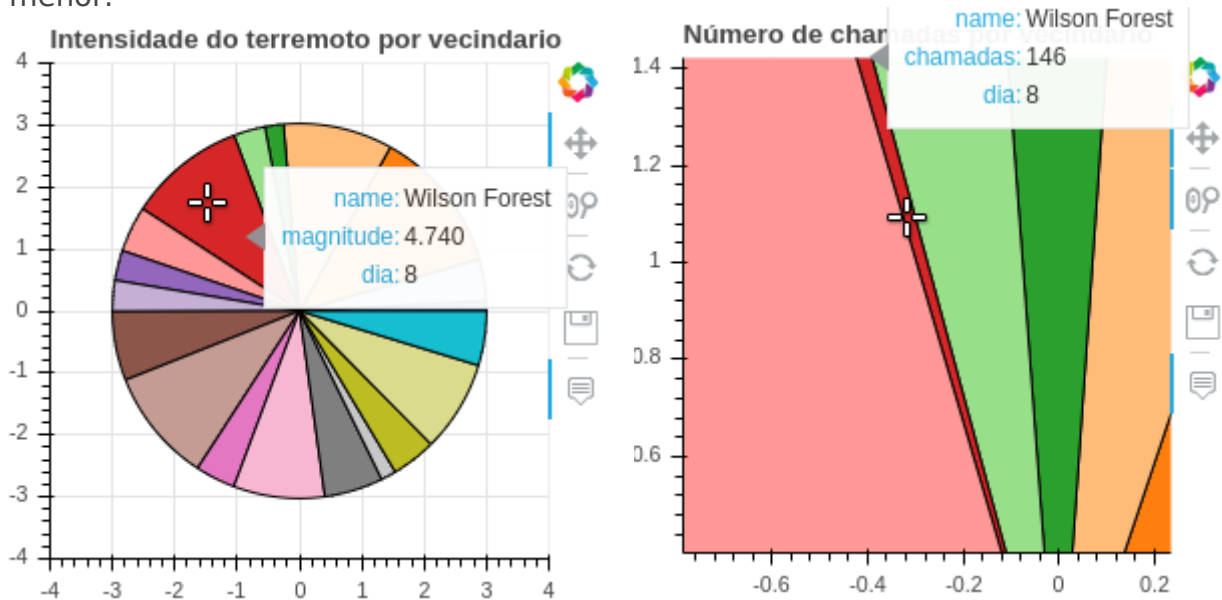


Figura 5. Número de llamadas y magnitud del terremoto para Wilson Forest en el día 8.

También se puede observar, por la figura 6 que algunas zonas, como en Safe Town, aunque el número de llamadas no fue tan elevado como en relación a otros vecindarios, se puede decir que tiene una alta importancia, ya que la magnitud del terremoto ahí fue alta, y en ese lugar se encuentra la planta nuclear.

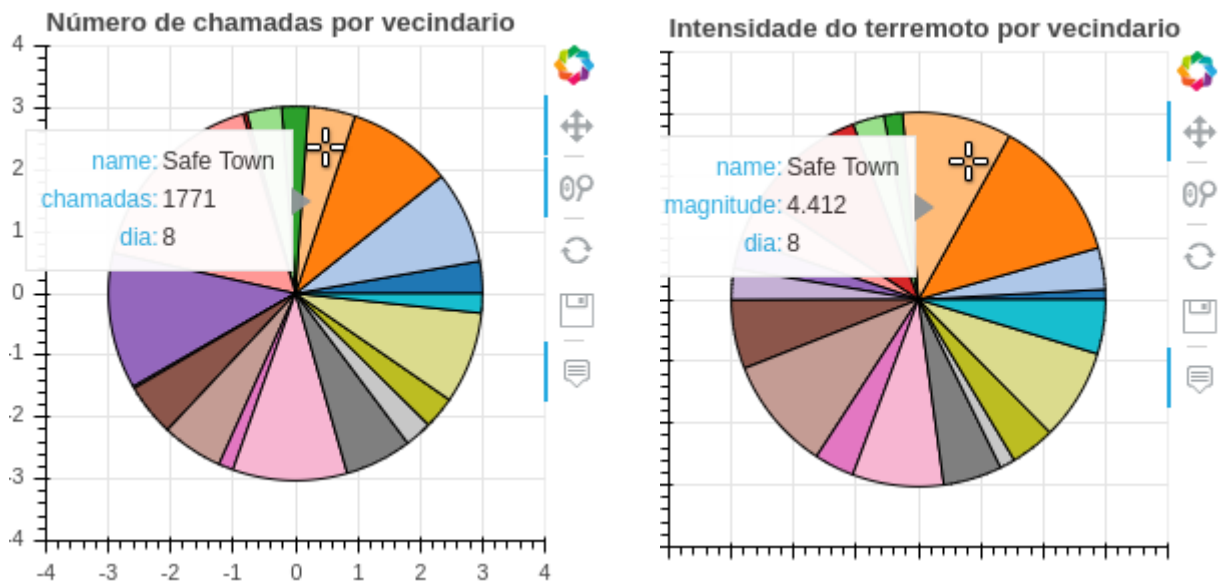


Figura 6. Relación de llamadas con la magnitud del terremoto en Safe Town.

Por último, la figura 7 muestra el daño promedio de cada medición por cada vecindario, dando a conocer, como lo muestra la figura 8, que “old Town” fue el vecindario mas afectado.

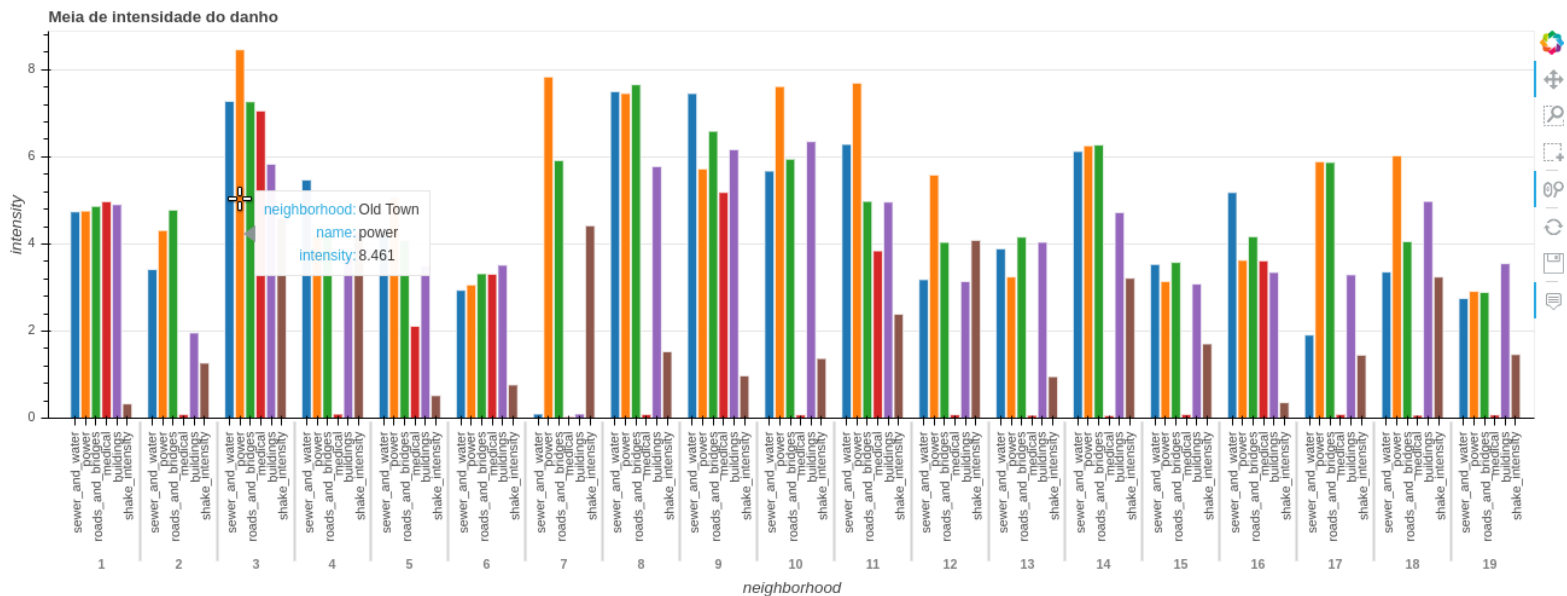


Figura 7. Promedio de daño por cada medición en cada vecindario.

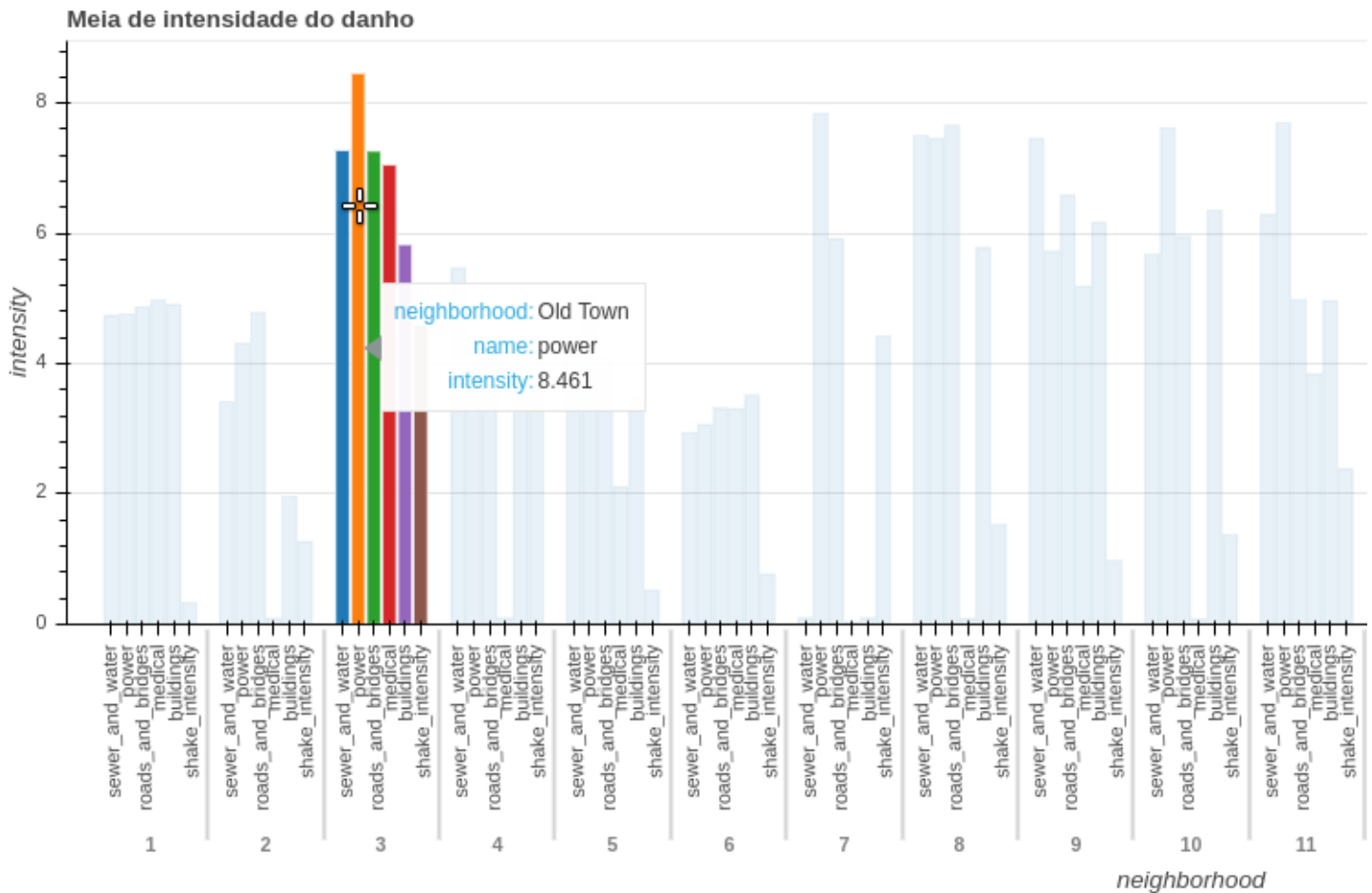


Figura 8. Old Town el cual es el que presenta mayores daños.

Tomando como base las figuras mencionadas anteriormente e interactuando con ellas, se puede decir que los vecindarios mayormente afectados según la media de daños reportados son *Old Town*, *Safe Town* y *Broadview*. Por otro lado, según estas medidas se crea el siguiente orden de atención a los llamados, teniendo como prioridad aquellos vecindarios en los cuales las infraestructuras médicas tuvieron un daño relevante:

- *Safe Town*, *Old Town*, *Palace Hills*, *Broadview*, *Terrapin Springs*, *Southon*, *Downtown*, *Southwest*.

Luego de dar prioridad a los vecindarios con mayor daño médico, se daría prioridad ha aquellos vecindarios con daños elevados en su alumbrado.

- *Northwest*, *Scenic Vista*, *Chapparral*, *Easton*, *Pepper Mill*, *CheddardFord*, *East Parton*, *Oak Willow*, *West Parton*, *Weston*, *Wilson Forest*

**2** - Use visual analytics to show uncertainty in the data. Compare the reliability of neighborhood reports. Which neighborhoods are providing reliable reports? Provide a rationale for your response. Limit your response to 1000 words and 10 images.

Las siguientes ilustraciones se basan en la interacción con el gráfico 2, el cual presenta una ilustración en forma de *boxplot* en la cual se podrán ver los reportes que están alejados de los valores medios. Por otro lado, el gráfico de barras de la izquierda representa la cantidad de *outliers* por tipo de daño.

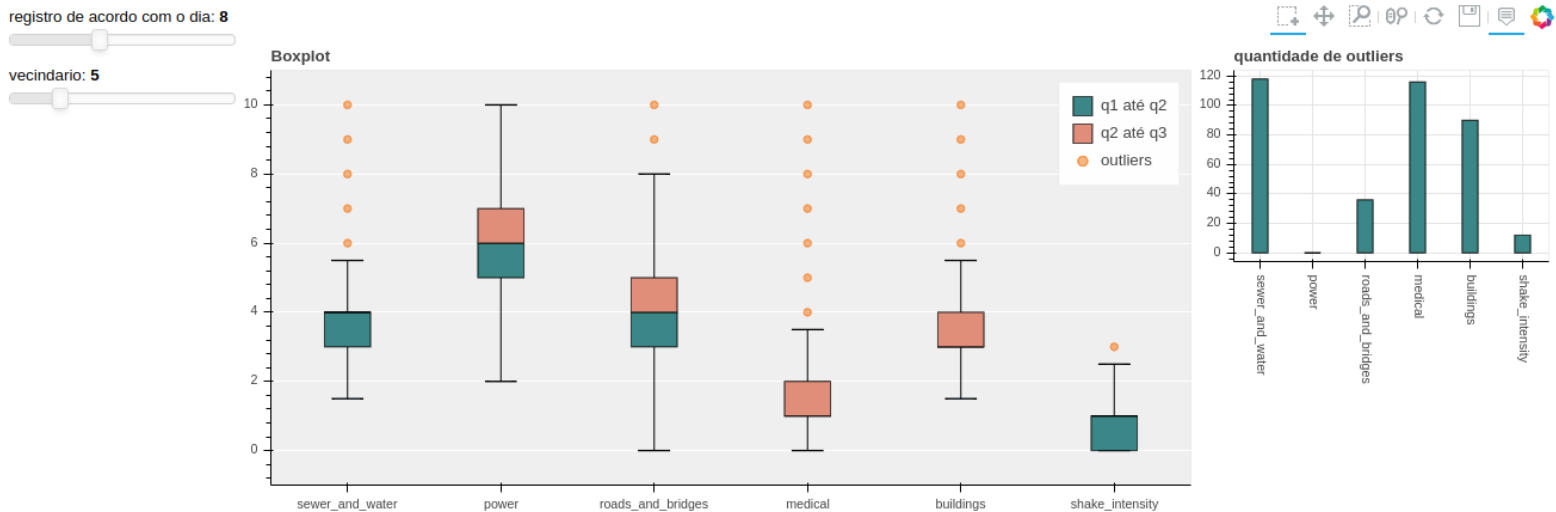


Gráfico 2. Panel de interacción diaria por vecindario con los datos del terremoto.

En principio los datos presentan valores "nan" los cuales se consideraron como si el usuario no hubiera querido reportar ningún daño de ese tipo, por lo tanto se tomaron como cero.

Algunos vecindarios reportan daños en infraestructura médica sin ellos tener hospitales, La figura 9 es un ejemplo de ello, por ende los usuarios no están dando datos verídicos.

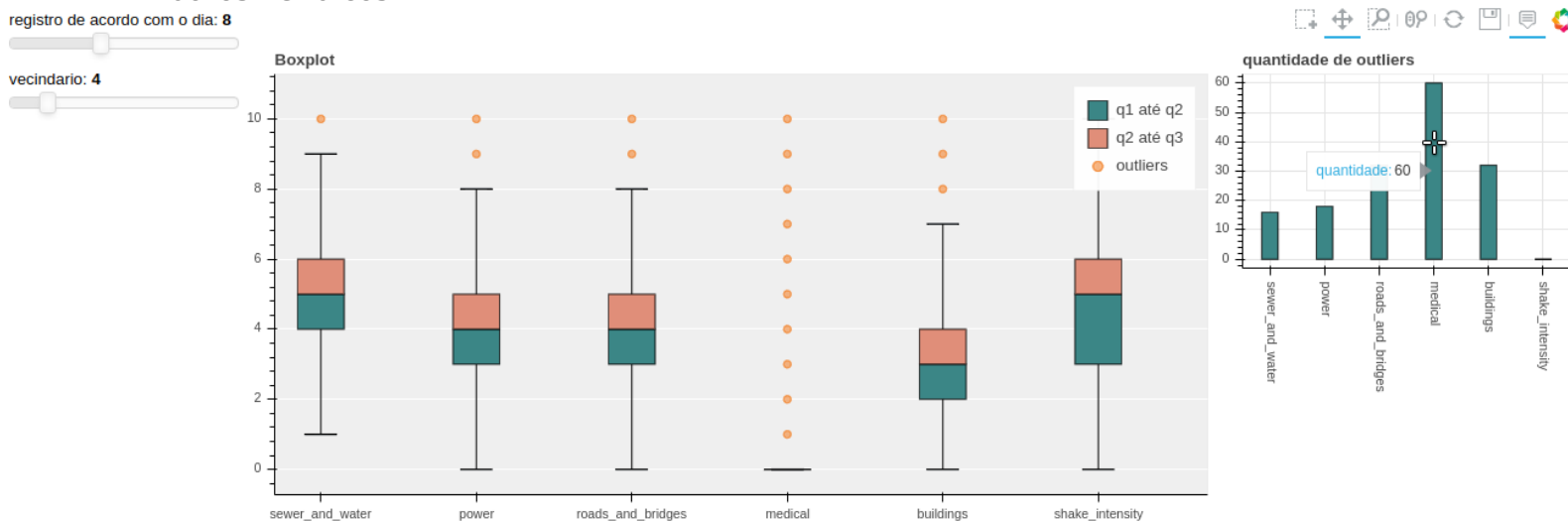


Figura 9. registro del vecindario cuatro el día ocho.

Por otro lado, hay vecindarios que presentan valores muy parecidos a lo largo del tiempo, presentan muy pocos datos con una dispersión alta, la figura 10 muestra uno de los vecindarios, el cual no presenta tantos datos inconsistentes.

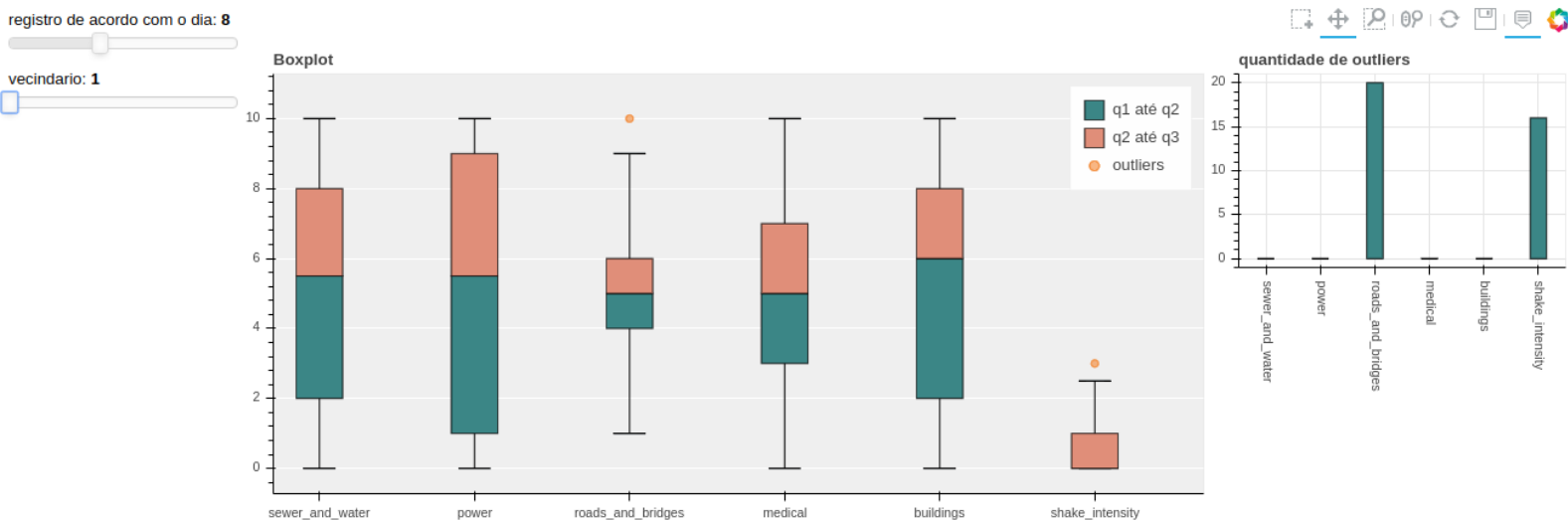


Figura 10. registro del vecindario uno el día ocho, se evidencian muy pocos registros fuera de lo normal.

Aunque también hay vecindarios en los que se reportaron valores muy altos en el daño, los cuales se salen del patrón normal de reporte, por ende se consideran que son datos exagerados. Como lo muestra la figura 11 muchos de los datos reportados son iguales, por ello el cuartil uno o el dos no se observan, ya que no hay una variación significativa de cuartil a cuartil.

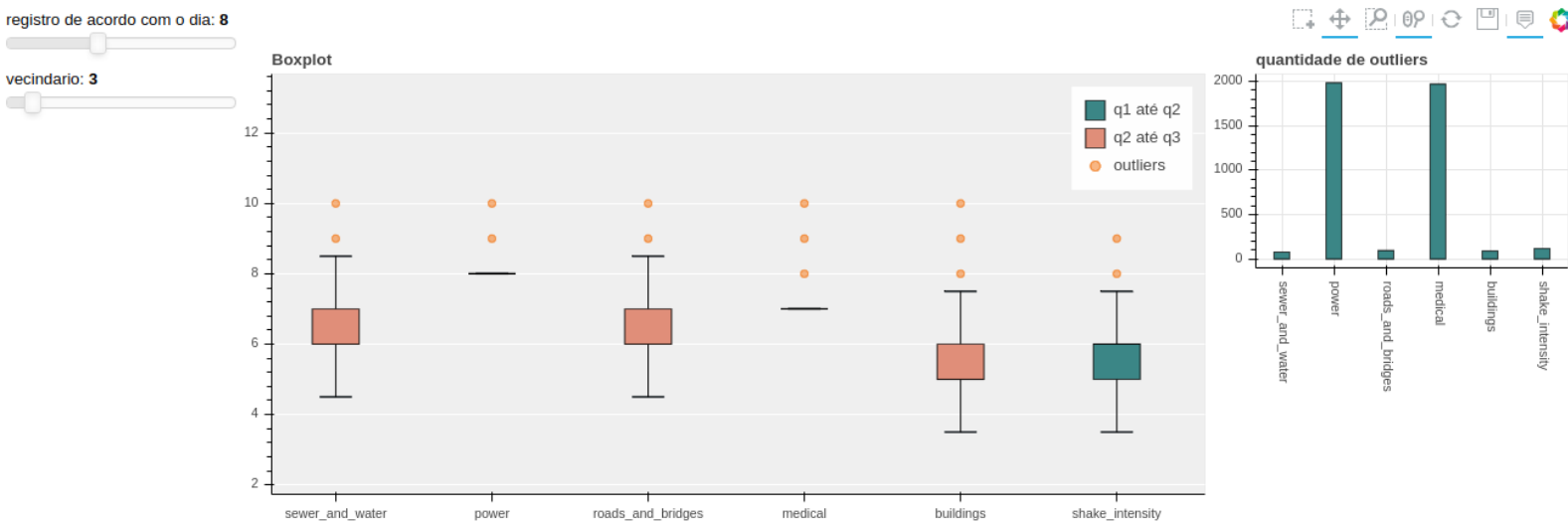


Figura 11. Outliers para el vecindario tres el día ocho.

También se puede observar que debido a que del día 11 se tienen muy pocos registros y estos no son muy dispersos, no se presentan *outliers* en ningún vecindario. La figura 12 ilustra lo mencionado.



registro de acordo com o dia: 11

vecindario: 2

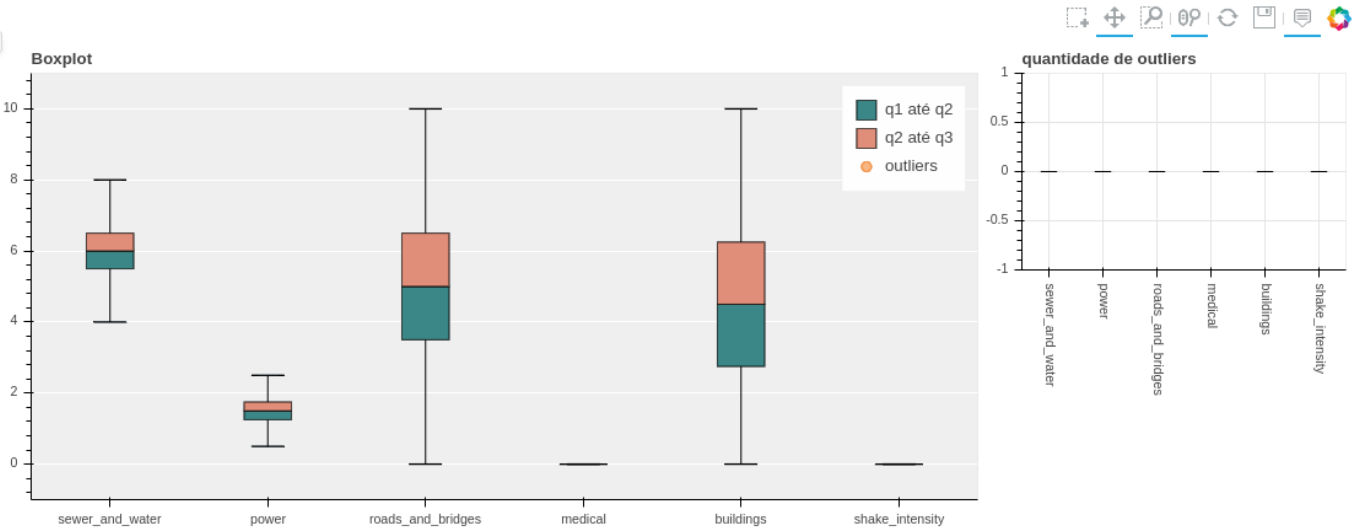


Figura 12. Registro de daño del vecindario dos el día once, se nota que no hay presencia de outliers.

Después de interactuar con el gráfico 2, se puede encontrar un patrón, el cual muestra que los datos que tienen un nivel de dispersión mas grande se encuentran en los días 7 y 8, en estos días también fue más alta la magnitud reportada para el terremoto, por ende, se puede decir que muchas personas se asustaron y reportaron datos un poco elevados, solo los vecindarios 1 y 10 mantuvieron unos reportes menos dispersos, como lo muestran las figuras 13 y 14.

registro de acordo com o dia: 8

vecindario: 1

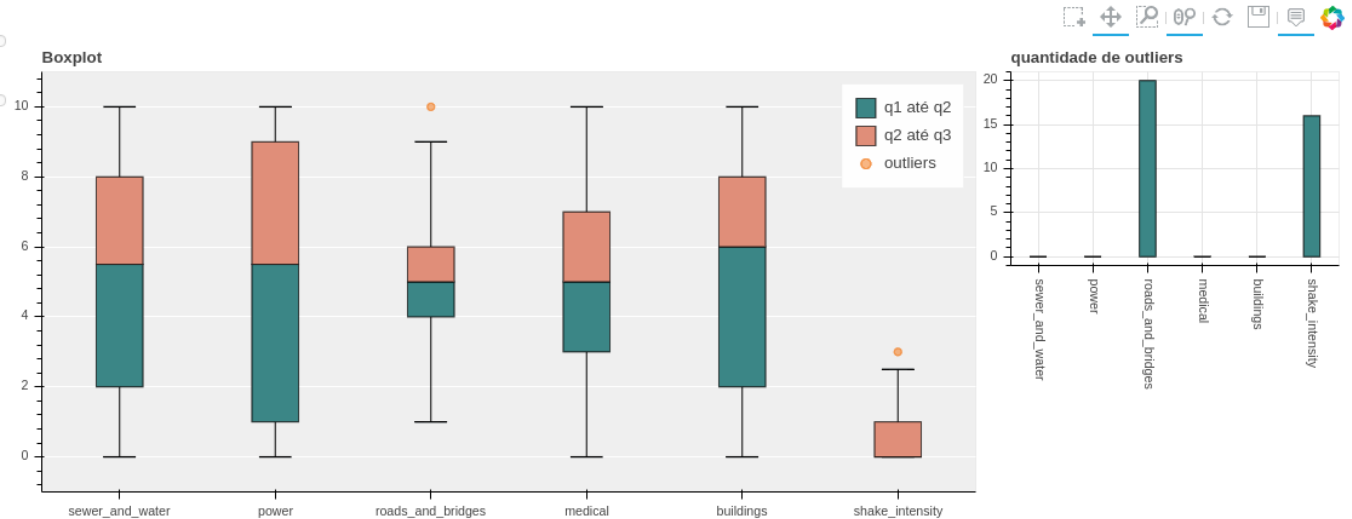


Figura 13. Registro de daño del vecindario uno el día ocho, solo presenta 36 outliers en total.

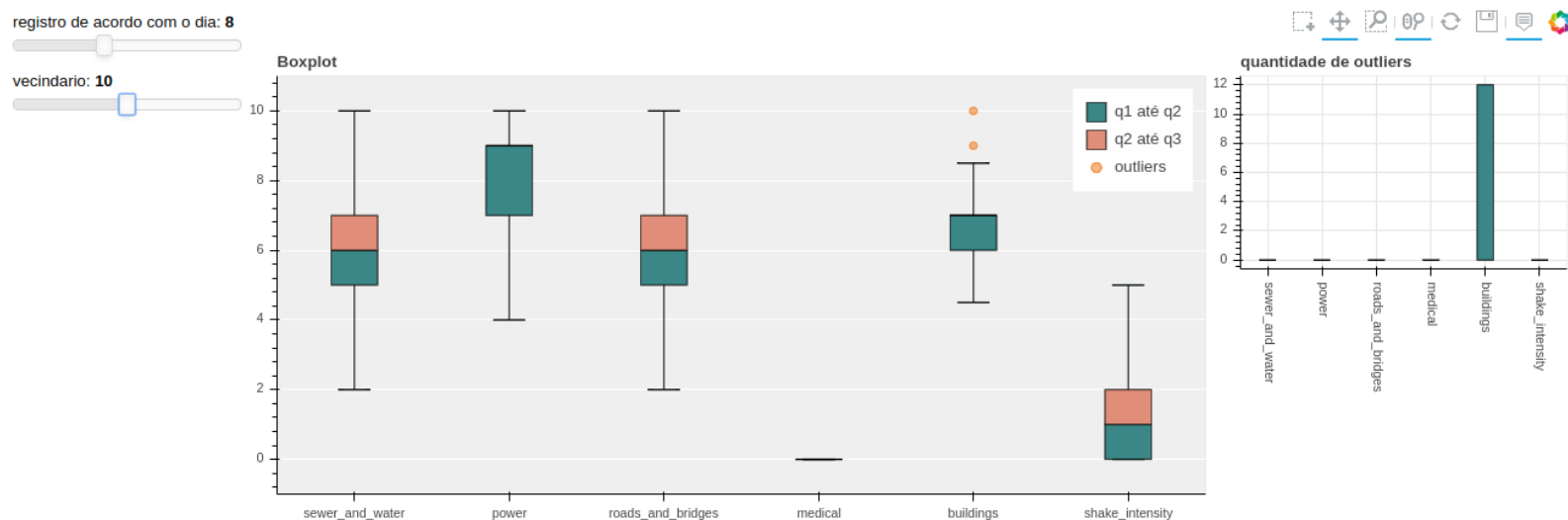


Figura 14. Registro de daño del vecindario diez el día ocho, solo presenta 12 outliers en total.

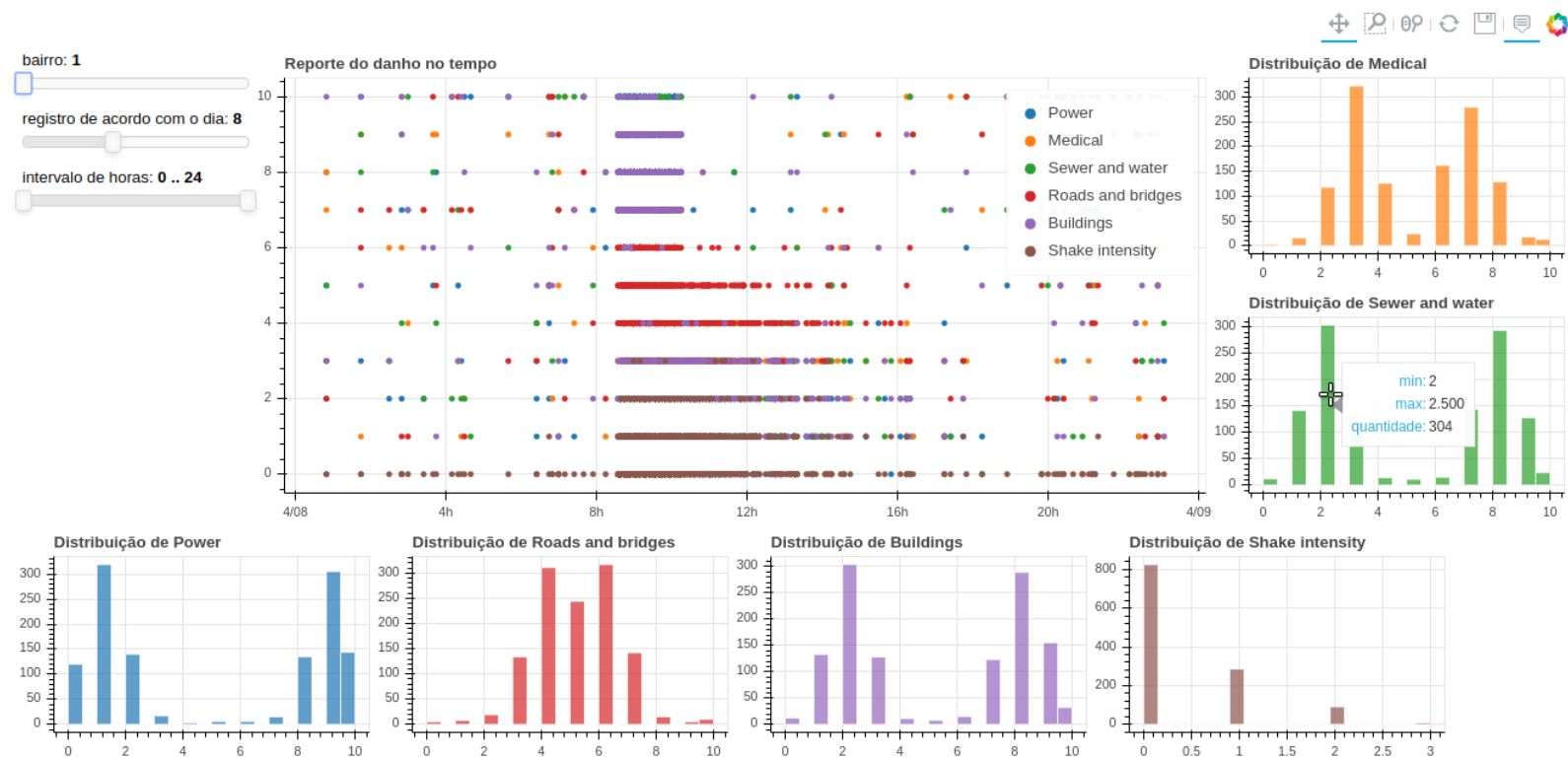
Luego de analizar todos los posibles casos, se puede decir que los siguientes vecindarios son los que reportan dados más confiables, entendiendo como confiable que los valores reportados tienen una menor dispersión.

- Palace Hills, Safe Town, Southwest, Wilson Forest, Chapparal, Cheddarford

**3** – How do conditions change over time? How does uncertainty in change over time? Describe the key changes you see. Limit your response to 500 words and 8 images.

¿Cómo cambian las condiciones con el tiempo? ¿Cómo cambia la incertidumbre en el tiempo? Describe los cambios clave que ves. Limita tu respuesta a 500 palabras y 8 imágenes.

Las siguientes ilustraciones se basan en la interacción con el gráfico 3, el cual presenta una ilustración central la cual contiene todos los reportes por día de cada barrio y para cada tipo de daño, pudiendo así, observar como cambian los reportes a lo largo del tiempo. También al rededor se presentan las distribuciones que va tomando cada variable a lo largo de tiempo. En la parte izquierda superior el usuario puede interactuar seleccionando el barrio, el día y el intervalo de horas que quiere analizar, a su vez, al hacer click en el "label" del gráfico central, ese atributo se ocultará para mejorar la visualización.



Una de las primeras cosas que se puede observar, es que los días 8 y 9 son en los cuales ocurre el pico del terremoto, en estos días es en los cuales mas se reporta la intensidad del terremoto, siendo para el día 8 el lapso entre las 8h y las 16h donde la gente percibió el terremoto. Mientras que el día 9 los reportes de intensidad se concentraron entre las 15h y las 19h. La figura 15 y 16 muestran esto respectivamente. Además, dejan ver que ciertos tipos de daños tienen una misma distribución.

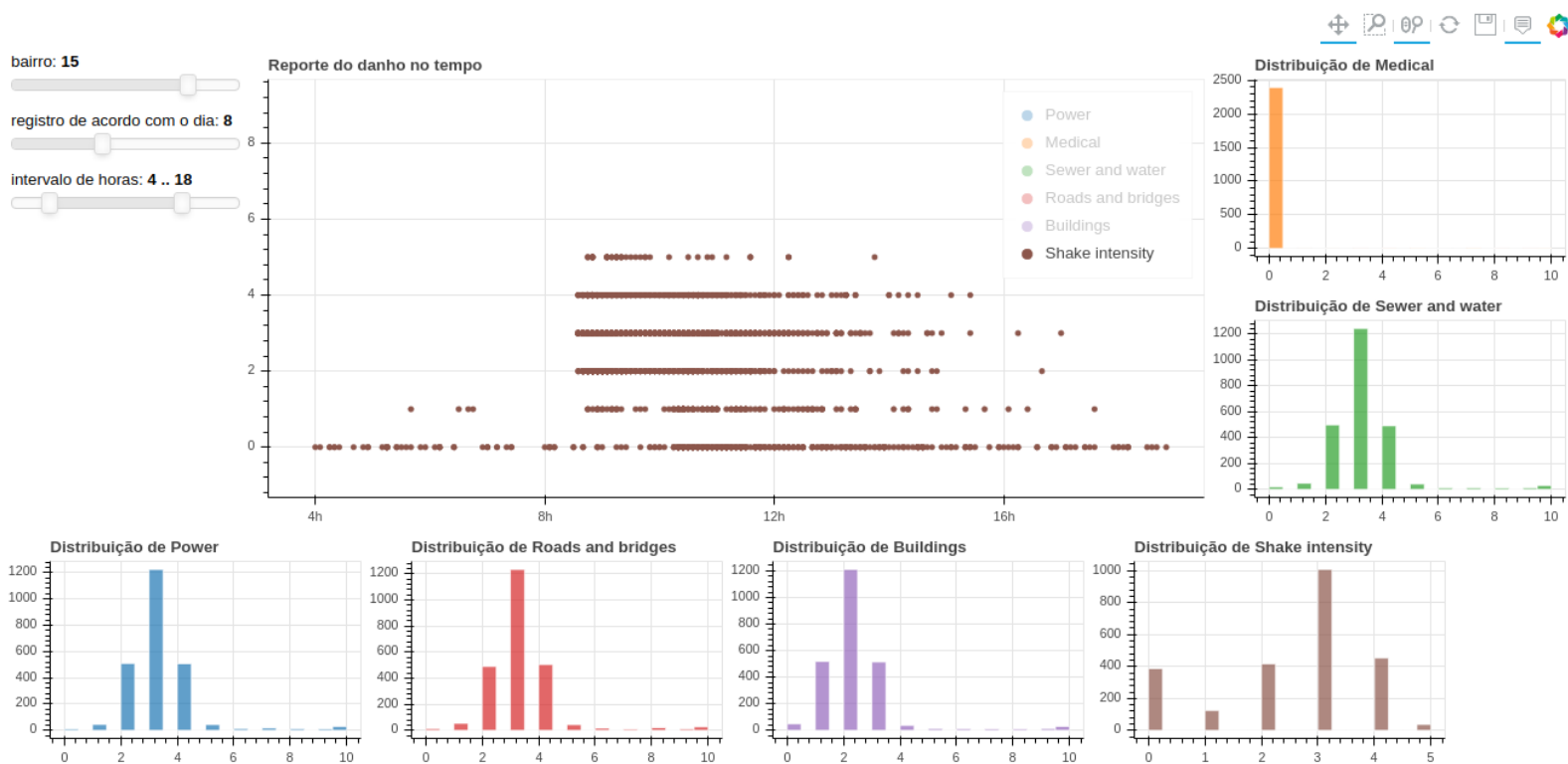


Figura 15. Intensidad del terremoto reportada por el barrio 15 el día 8, todos los barrios presentan ese intervalo horario de reporte, aunque solo algunos, como este, tienen relacionados sus tipos de daños, se podría decir que iguales.

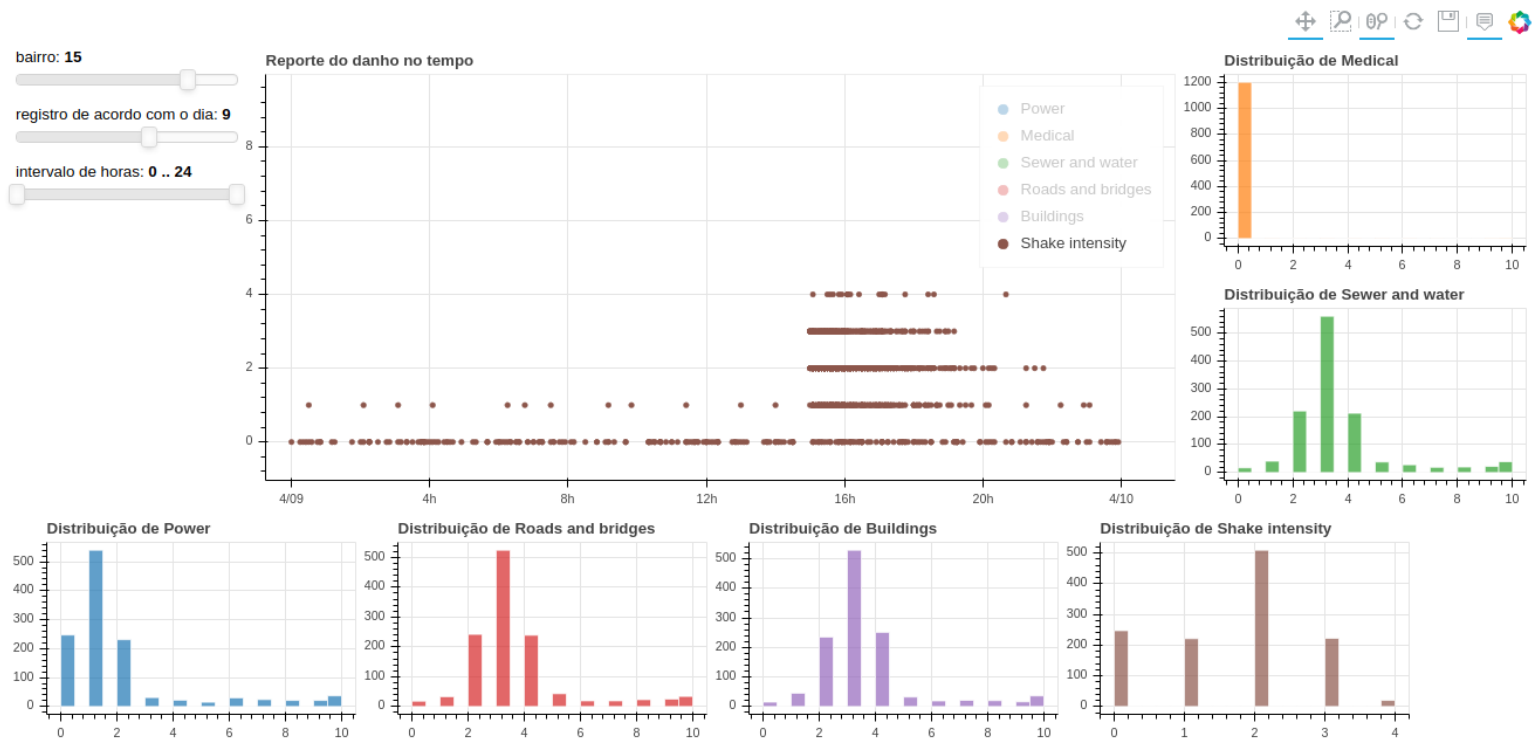


Figura 16. Intensidad del terremoto reportada por el barrio 15 el día 9, todos los barrios presentan ese intervalo horario de reporte, aunque solo algunos, como este, tienen relacionados sus tipos de daños, se podría decir que iguales.

A medida que se interacciona con el gráfico 3, se puede ver que se presentan ciertas anomalías en los datos, ya que presentan una incertidumbre grande, la figura 17 muestra como para el daño reportado en “buildings” hay dos picos de concentraciones, lo cual en un principio uno creería que los datos deberían seguir una distribución normal, aunque para ese día en ese barrio, la mayoría sigue esa anomalía.

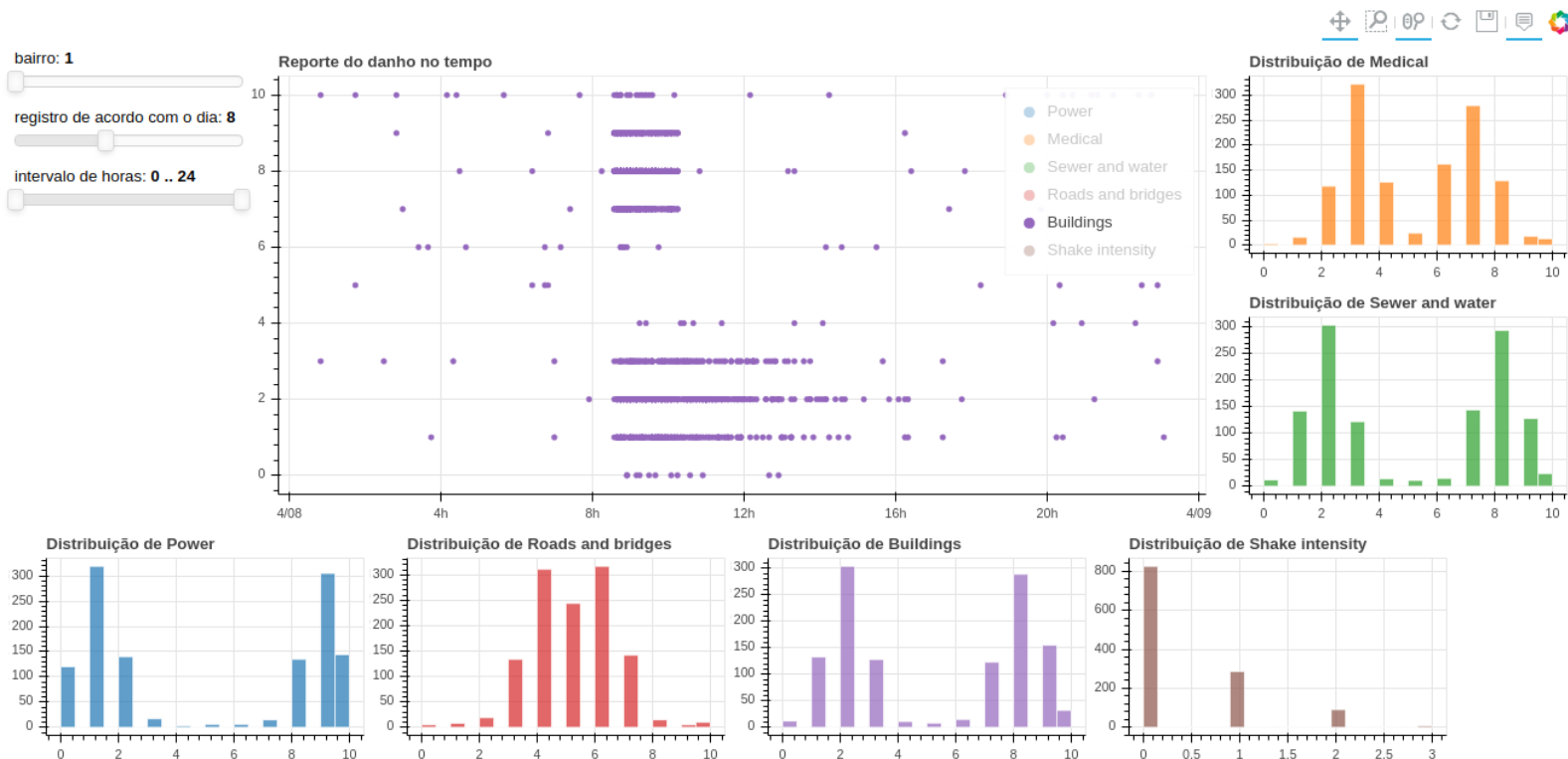


Figura 17. Reporte de daño en *buildings* por el barrio 1 el día 8. Se puede evidenciar que casi todos los daños reportados siguen una distribución a dos picos, algunas con los picos mas apartados que otras.

**4** -- The data for this challenge can be analyzed either as a static collection or as a dynamic stream of data, as it would occur in a real emergency. Describe how you analyzed the data - as a static collection or a stream. How do you think this choice affected your analysis? Limit your response to 200 words and 3 images.

***Provide your answer and corresponding images here.***