

# Trabajo Práctico 1: Análisis Exploratorio de Datos

[75.06 / 95.58] Organización de Datos  
Primer cuatrimestre de 2021

Grupo: Equipo onda maravilla lobo

Alumno	Padrón	Mail
Dituro, Celeste	104011	dituro@fi.uba.ar
Pfaab, Ivan	103862	ipfaab@fi.uba.ar
Lopez, Victoria Abril	103927	vlopez@fi.uba.ar

[https://github.com/vickylopezz/  
Organizacion-de-Datos-75.06-95.58-/tree/main/TP1](https://github.com/vickylopezz/Organizacion-de-Datos-75.06-95.58-/tree/main/TP1)

## Curso

- Argerich, Luis Argerich
- Golmar, Natalia
- Martinelli, Damian Ariel
- Ramos Mejia, Martín
- Gianmarco Cafferata
- Joaquin Torre Zaffaroni
- Julian Crespo
- Esteban Djeordjian

# Índice

<b>1</b>	<b>Introducción</b>	<b>1</b>
<b>2</b>	<b>Manipulación general de los datos</b>	<b>1</b>
2.1	Preparación de los datos . . . . .	3
2.2	Limpieza de datos . . . . .	3
<b>3</b>	<b>Pre análisis de los datos</b>	<b>3</b>
3.1	Exploración de datos . . . . .	3
3.2	Estructuras . . . . .	4
3.3	Sectorización . . . . .	6
3.4	Antigüedad . . . . .	7
<b>4</b>	<b>Análisis de daños</b>	<b>8</b>
4.1	Antigüedad . . . . .	9
4.2	Estructura . . . . .	11
4.3	Localización . . . . .	15
4.3.1	Características de la edificación . . . . .	20
4.4	Propiedades físicas . . . . .	22
4.4.1	Altura . . . . .	22
4.4.2	Área . . . . .	23
4.5	Cantidad de pisos . . . . .	24
4.6	Uso secundario . . . . .	25
4.7	Social . . . . .	28
<b>5</b>	<b>Relación entre los daños</b>	<b>28</b>
<b>6</b>	<b>Conclusiones</b>	<b>31</b>

## 1. Introducción

En el presente informe se propone analizar los datos provistos por <https://www.drivendata.org> con el objetivo de determinar características y variables importantes, descubrir insights interesantes, y analizar la estructura de los mismos sobre el terremoto Gorkha ocurrido en Nepal en el año 2015. Los datos están proporcionados por dos archivos csv:

- “train\_values.csv”: aquí encontraremos 38 features que describen a la edificación identificada por un id.
- “train\_labels.csv”: aquí encontraremos la variable ‘damage\_grade’ y el correspondiente id de la edificación.

## 2. Manipulación general de los datos

En primera instancia observamos la información proporcionada por los archivos para tener una visión general de los datos con los cuales se trabajará. El dataframe aporta:

- Datos correspondientes a edificios situados en Nepal.
- 260601 registros con 40 atributos.

Los atributos son:

- **buildin\_id**: (tipo: ID) identificador único de la edificación.
- **geo\_level\_1\_id**, **geo\_level\_2\_id**, **geo\_level\_3\_id**: (tipo: enteros): región geográfica en la cual la edificación existe, desde la más general (level 1) a la más específica (level 3). Valores posibles: level 1: 0-30, level 2: 0-1427, level 3: 0-12567.
- **count\_floors\_pre\_eq**: (tipo: entero) número de pisos en la edificación antes del terremoto.
- **age**: (tipo: entero) antigüedad de la edificación en años.
- **area\_percentage**: (tipo: entero) superficie normalizada ocupada por la edificación.
- **height\_percentage**: (tipo: entero) altura normalizada ocupada por la edificación.
- **land\_surface\_condition**: (tipo: categórico) condición de la superficie terrestre donde el edificio fue construido. Valores posibles: n, o, t.
- **foundation\_type**: (tipo: categórico) tipo de cimientos usados cuando se construyó la edificación. Valores posibles: h, i, r, u, w.
- **roof\_type**: (tipo: categórico) tipo de techo usado cuando se construyó la edificación. Valores posibles: n, q, x.

- **ground\_floor\_type:** (tipo: categórico) tipo de construcción usado en la planta baja cuando se construyó la edificación. Valores posibles: f, m, v, x, z.
- **other\_floor\_type:** (tipo: categorical) tipo de construcción usado en otros pisos cuando se construyó la edificación (exceptuando el techo). Posibles valores: j, q, s, x.
- **position:** (tipo: categórico) orientación de la edificación. Posibles valores: j, o, s, t.
- **plan\_configuration:** (tipo: categórico) formato de construcción de la edificación (para diseño sísmico). Valores posibles: a, c, d, f, m, n, o, q, s, u.
- **has\_superstructure\_adobe\_mud:** (tipo: binario) variable que indica si la edificación fue construida con adobe/barro.
- **has\_superstructure\_mud\_mortar\_stone:** (tipo: binario) variable que indica si la edificación fue construida con barro - piedra.
- **has\_superstructure\_stone\_flag:** (tipo: binario) variable que indica si la edificación fue construida con piedra.
- **has\_superstructure\_cement\_mortar\_stone:** (tipo: binario) variable que indica si la edificación fue construida con cemento - piedra.
- **has\_superstructure\_mud\_mortar\_brick:** (tipo: binario) variable que indica si la edificación fue construida con barro - ladrillos.
- **has\_superstructure\_cement\_mortar\_brick:** (tipo: binario) variable que indica si la edificación fue construida con cemento - ladrillos.
- **has\_superstructure\_timber:** (tipo: binario) variable que indica si la edificación fue construida con Timber (madera específica para la construcción).
- **has\_superstructure\_bamboo:** (tipo: binario) variable que indica si la edificación fue construida con Bambú (caña).
- **has\_superstructure\_rc\_non\_engineered:** (tipo: binario) variable que indica si la edificación fue construida con concreto reforzado no-diseñado.
- **has\_superstructure\_rc\_engineered:** (tipo: binario) variable que indica si la edificación fue construida con concreto reforzado diseñado.
- **has\_superstructure\_other:** (tipo: binario) variable que indica si la edificación fue construida con otro material.
- **legal\_ownership\_status:** (tipo: categórico) estado legal de la tierra donde la edificación fue construida. Valores posibles: a, r, v, w.
- **count\_families:** (tipo: entero) número de familias que vivían en la edificación.
- **has\_secondary\_use:** (tipo: binario) variable que indica si la edificación era usada con un uso secundario.

- **has\_secondary\_use\_agriculture**: (tipo: binario): variable que indica si la edificación era usada con propósitos de agricultura.
- **has\_secondary\_use\_hotel**: (tipo: binario) variable que indica si la edificación era usada como oficina de gobierno.
- **has\_secondary\_use\_rental**: (tipo: binario) variable que indica si la edificación se alquilaba.
- **has\_secondary\_use\_institution**: (tipo: binario) variable que indica si la edificación era usada como sede de una institución.
- **has\_secondary\_use\_school**: (tipo: binario) variable que indica si la edificación era usada como escuela.
- **has\_secondary\_use\_industry**: (tipo: binario) variable que indica si la edificación era usada con propósitos industriales.
- **has\_secondary\_use\_health\_post**: (tipo: binario) variable que indica si la edificación era usada como puesto de salud.
- **has\_secondary\_use\_gov\_office**: (tipo: binario) variable que indica si la edificación era usada como oficina de gobierno.
- **has\_secondary\_use\_police**: (tipo: binario) variable que indica si la edificación era usada como estación de policía.
- **has\_secondary\_use\_other**: (tipo: binario) variable que indica si la edificación era usada con otro uso secundario.

## 2.1. Preparación de los datos

Al tener los datos distribuidos en dos satasets los combinamos para trabajar sobre uno unico que contenga toda la información necesaria.

## 2.2. Limpieza de datos

Es en este momento del análisis donde hacemos las configuraciones necesarias sobre el set de datos para poder trabajar mejor más tarde. Comprobamos que el dataframe no posea datos nulos para descartar datos innecesarios y verificamos que no haya edificios repetidos para evitar una alteración en las estadísticas.

# 3. Pre análisis de los datos

## 3.1. Exploración de datos

- ¿Qué tipo de features contiene?

Vamos a encontrar distintas características respecto a las edificaciones. Algunas de ellas son: locación, estructuras de construcción, usos secundarios, cantidad de personas que habitaban en las mismas, antigüedad.

- ¿Cómo están representados los datos en las distintas features?

La información va a estar representada en forma numérica (int64) y categórica (object).

### 3.2. Estructuras

- Por cada tipo de estructura, ¿cuántas edificaciones hay?

Estructura	Con	Sin
adobe - barro	23101	237500
barro - piedra	198561	62040
piedra	8947	251654
cemento - piedra	4752	255849
barro - ladrillo	17761	242840
cemento - ladrillo	19615	240986
concreto reforzado no diseñado	11099	249502
concreto reforzado diseñado	4133	256468
Madera	66450	194151
Bambú (Caña)	22154	238447
Otro	3905	256696
Total	380478	-

Cuadro 1: Cantidad de edificaciones construidas con la estructura

Una observación a tener en cuenta es que una edificación puede estar construida por más de un tipo de estructura por lo tanto la suma total de todas las cantidades no coincide con el total de edificaciones. En la figura 1 se visualiza la cantidad de edificios construidos a base de cada estructura. Por otro lado, la figura 2 presenta la distribución de estructuras considerando únicamente aquellas edificaciones compuestas por una única estructura.

Podemos notar al exigir una exclusividad en la estructura de construcción que el barro y piedra continúa predominando mientras que para el resto de las estructuras las cantidades se ven modificadas considerablemente.

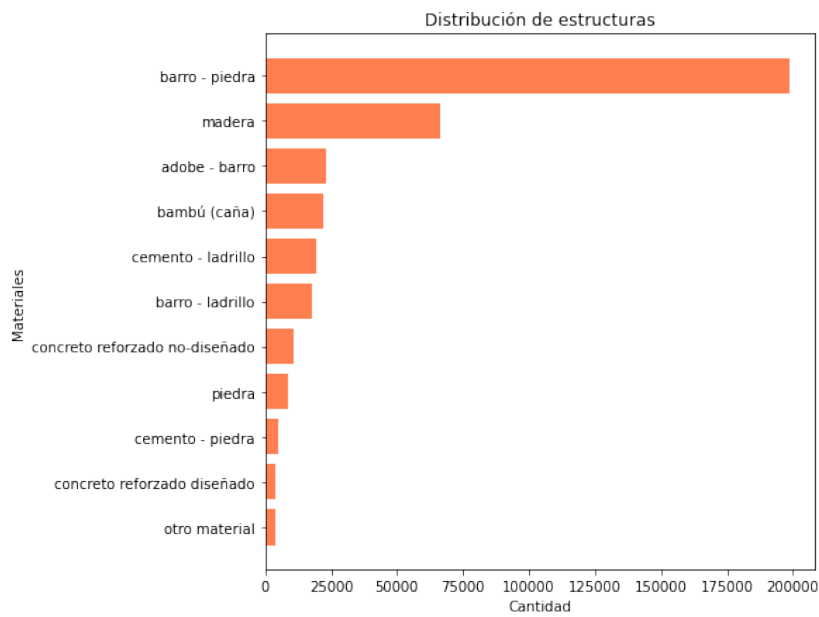


Figura 1: Distribución de estructuras totales

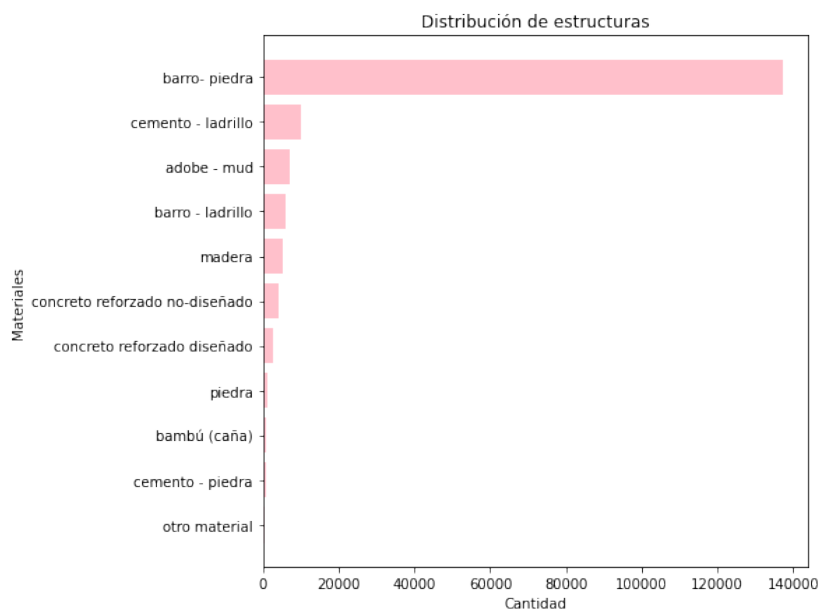


Figura 2: Distribución de materiales por edificio único

- ¿Son más las edificaciones construidas a partir de una única estructura que las compuestas por una mayor cantidad de estructuras?



Figura 3: Comparación de composiciones de estructuras

Consideramos una estructura como una composición de uno o más materiales, por ejemplo, barro-piedra, madera, etc. Podemos notar que son más las edificaciones construidas a partir una única estructura.

### 3.3. Sectorización

- ¿Existe intersección entre sectores?

Queremos probar que el geo level 3 está contenido en el geo level 2 y que el geo level 2 está incluido en el geo level 1.

Denominando al geo level 1 como G1, al geo level 2 como G2 y al geo level 3 como G3 probamos que:

- G1 tiene 31 sectores, la misma cantidad de ids que se menciona la información previa.
- G2 tiene 1414 sectores, cantidad menor que el rango de ids por lo que se deduce que no necesariamente a cada sector le corresponde un id de ese rango.
- G3 tiene 11595 sectores, cantidad menor que el rango de ids por lo que se deduce que no necesariamente a cada sector le corresponde un id de ese rango.

Dado que la cantidad de sectores que pertenecen tanto a G1 como a G2 coincide con la cantidad de sectores de G2 podemos deducir que G2 está contenido en G1 y lo mismo entre G2 y G3. Por lo cual,  $G1 \subset G2 \subset G3$ .

Teniendo en cuenta lo que se menciona anteriormente trabajaremos solo con G1 ya que agrupa por la mayor cantidad de edificios, mientras que G2 y G3 tienen una cantidad muy baja de edificios por sector la cual alteraría las estadísticas. En la figura 4 se puede observar la cantidad de edificios por sector del G1.



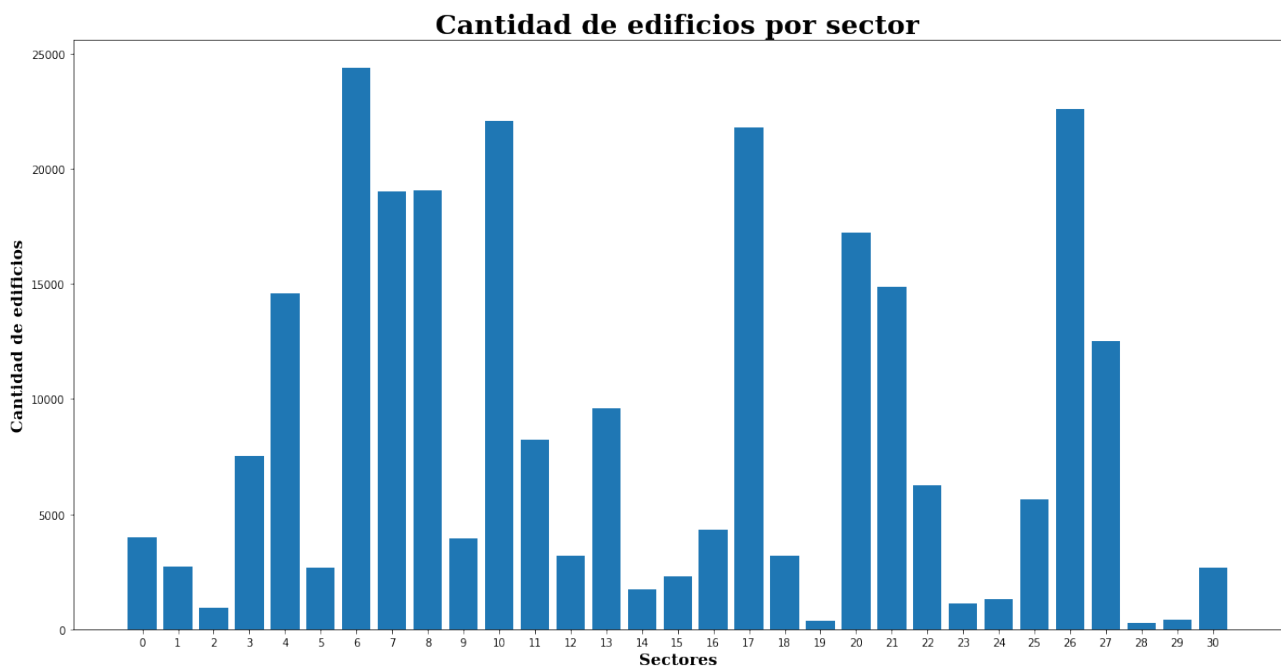


Figura 4: Cantidad de edificios por sector (geo-level 1)

### 3.4. Antigüedad

- ¿Cual es la distribución de los edificios según su antigüedad?

Podemos observar en la figura 5 que la mayoría de los edificios se encuentran entre los 0 y los 25 años de antigüedad. Notamos que existen 1390 edificaciones de 995 años de antigüedad, frente a este dato consideramos que son edificaciones historicas y que fueron correctamente consideradas dentro del dataset.

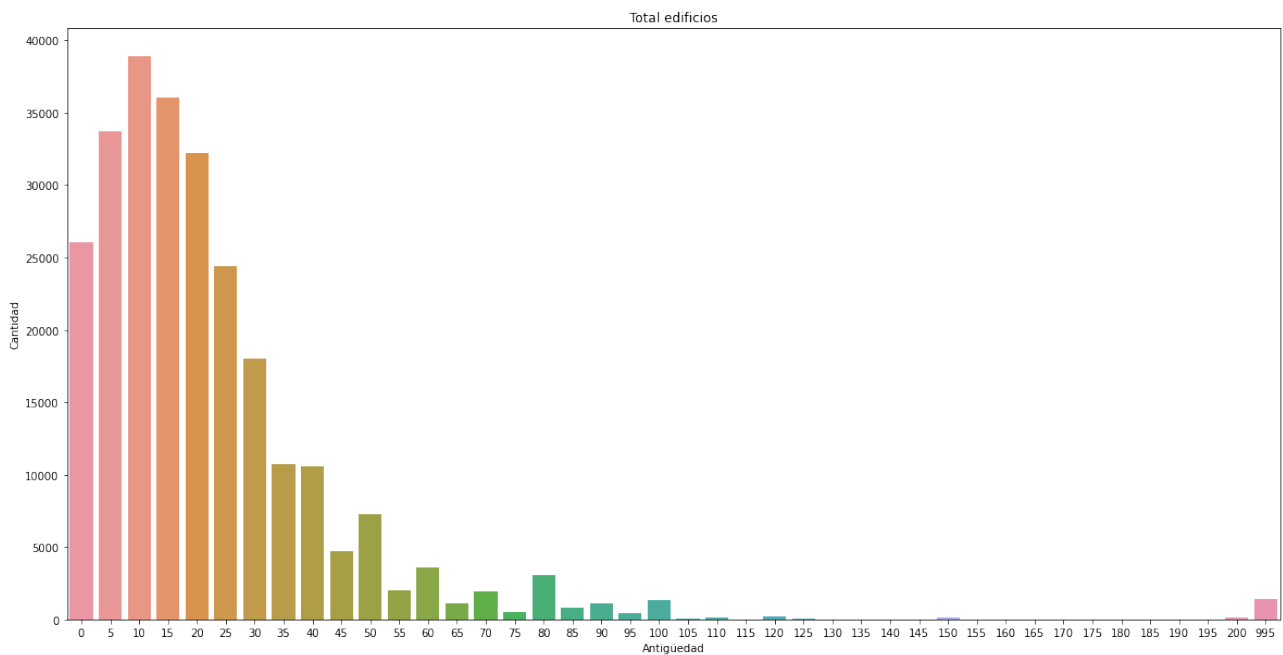


Figura 5: Cantidad de edificios por antigüedad

## 4. Análisis de daños

- ¿Cuál es el daño que predomina tras el terremoto?

Se puede observar que la mayoría de los edificios sufrieron un daño medio, seguidos por edificios con un daño serio y por último la menor cantidad sufrió un daño bajo. (Ver figuras 6 y 7)

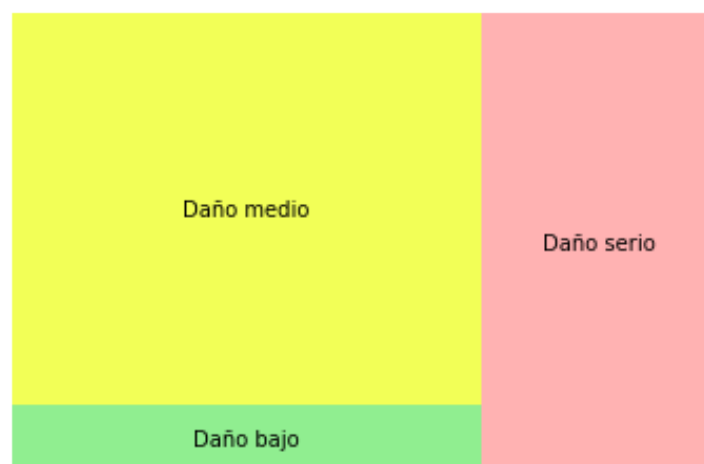


Figura 6: Cantidad de edificios por daño

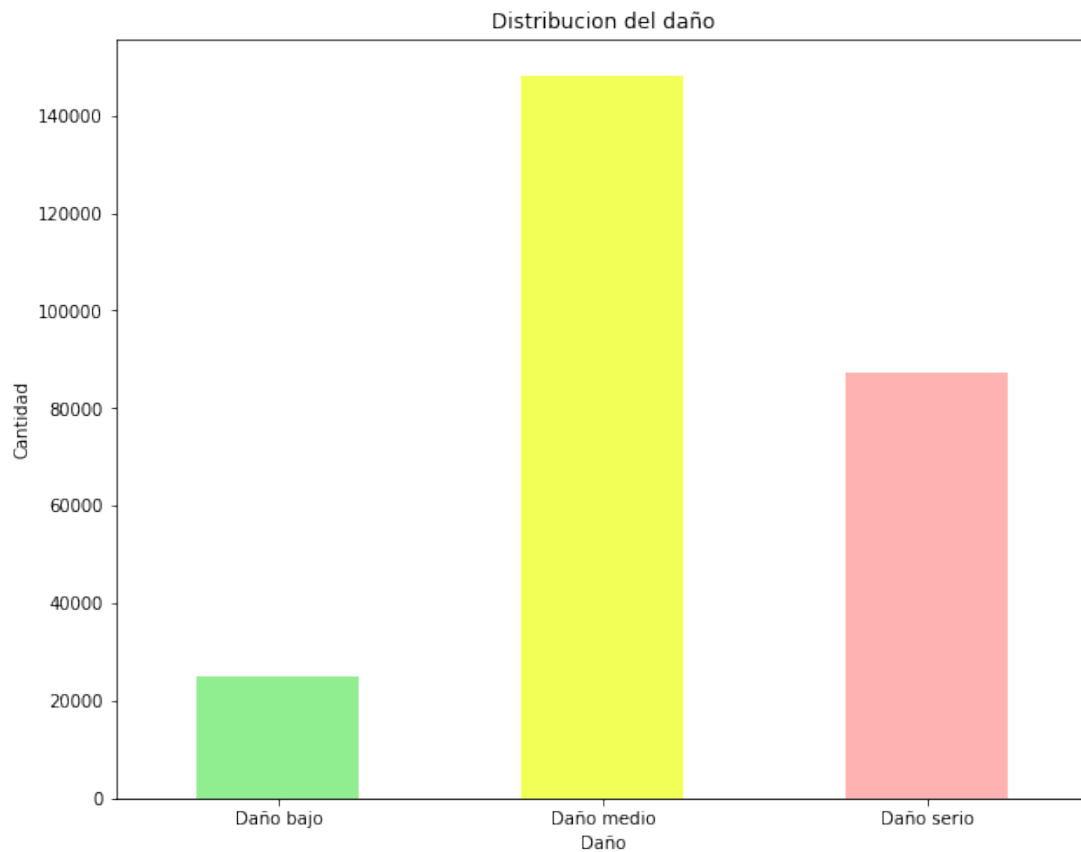


Figura 7: Cantidad de edificios por daño

#### 4.1. Antigüedad

- ¿Los edificios más antiguos son los que mayor daño sufrieron?

En las figuras 8 y 9 se analizan cuántos edificios sufrieron cada tipo de daño y su porcentaje de acuerdo a su antigüedad. Los edificios con daños más serios no son exactamente los más antiguos. El daño serio se mantuvo constante en la mayoría de los edificios, al igual que los de daño medio. A medida que la antigüedad aumenta los daños comienzan a ser mas severos y la diferencia entre el daño bajo y el daño más critico comienza a ser mas notoria.

Llegamos a la conclusión de que las edificaciones mas antiguas son las que se vieron más afectadas por el terremoto.

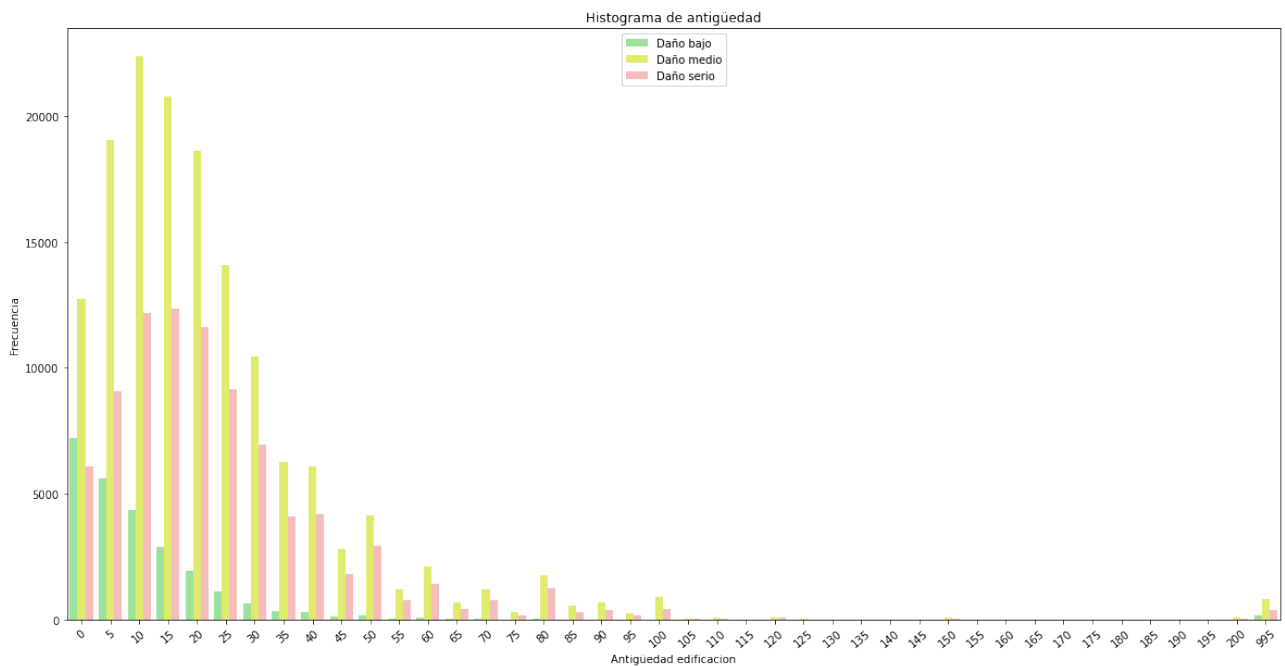


Figura 8: Cantidad de edificios por daño

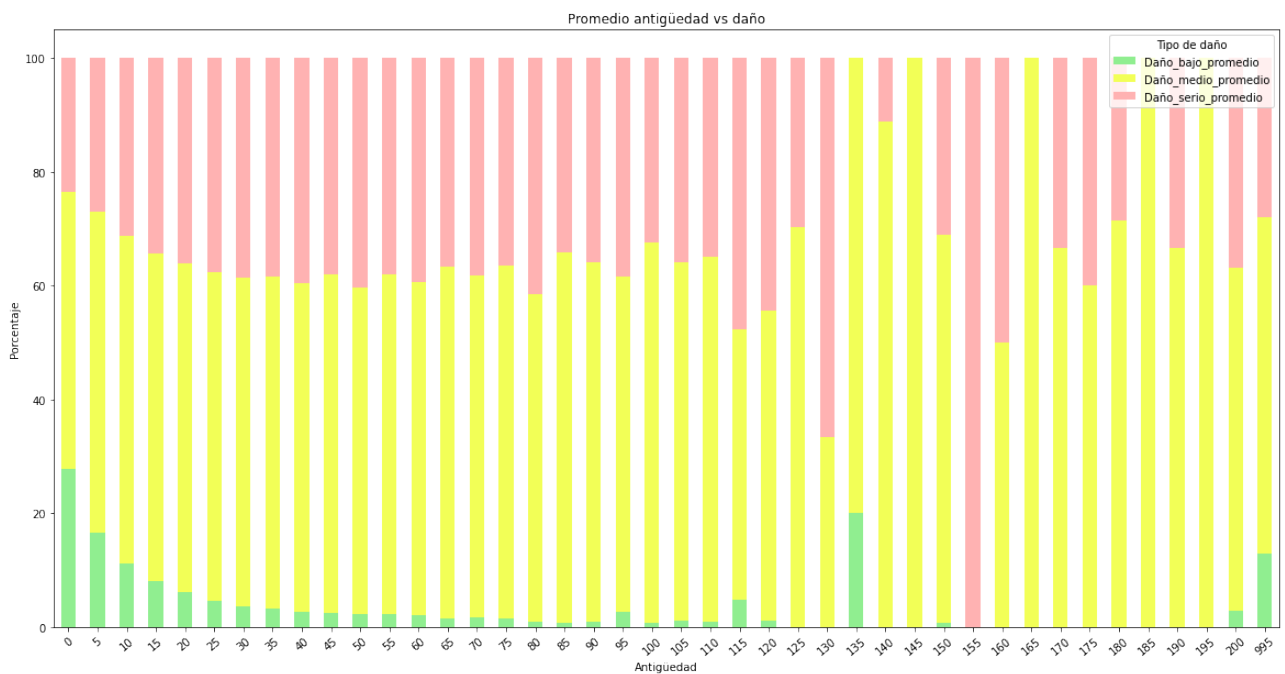


Figura 9: Cantidad de edificios por daño

## 4.2. Estructura

- ¿Qué impacto tuvo cada edificación según su estructura?

Observamos que los edificios están contruídos en mayor parte por barro y piedra y que predominan los que sufrieron daño medio. Esto se visualiza además en cada tipo de estructura. La unica esctructura que sufrió en primer lugar un daño bajo y las que menos sufrieron daño medio y serio fueron las de concreto reforzado diseñado. (Ver figura 10).

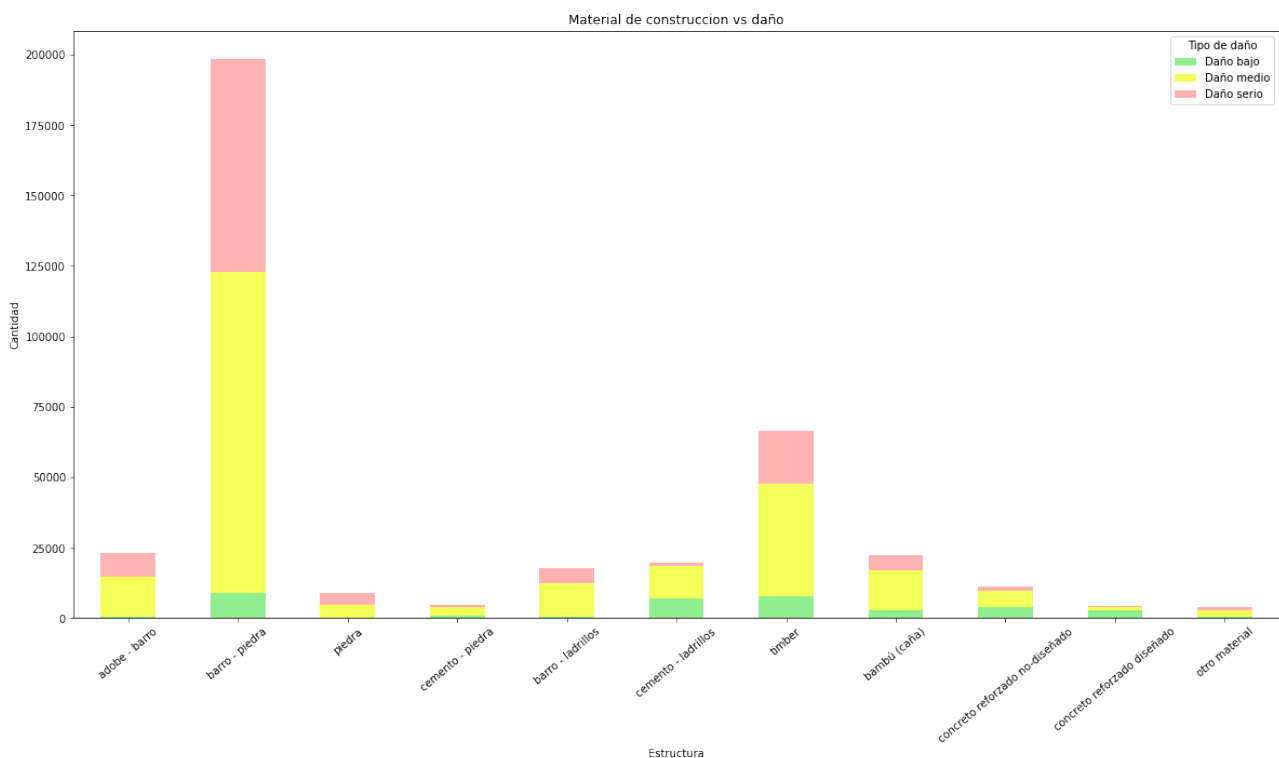


Figura 10: Cantidad de edificios por daño con una única estructura

A su vez, en las figuras 11 y 15 analizamos las edificaciones contruídas con una única estructura (solo piedra, solo adobe-barro,etc.) y observamos que el daño predominante continuaba siendo el de nivel 2. Por lo cual llegamos a la conclusión de que el grado de daño es independiente de la cantidad de estructuras que componen la edificación.

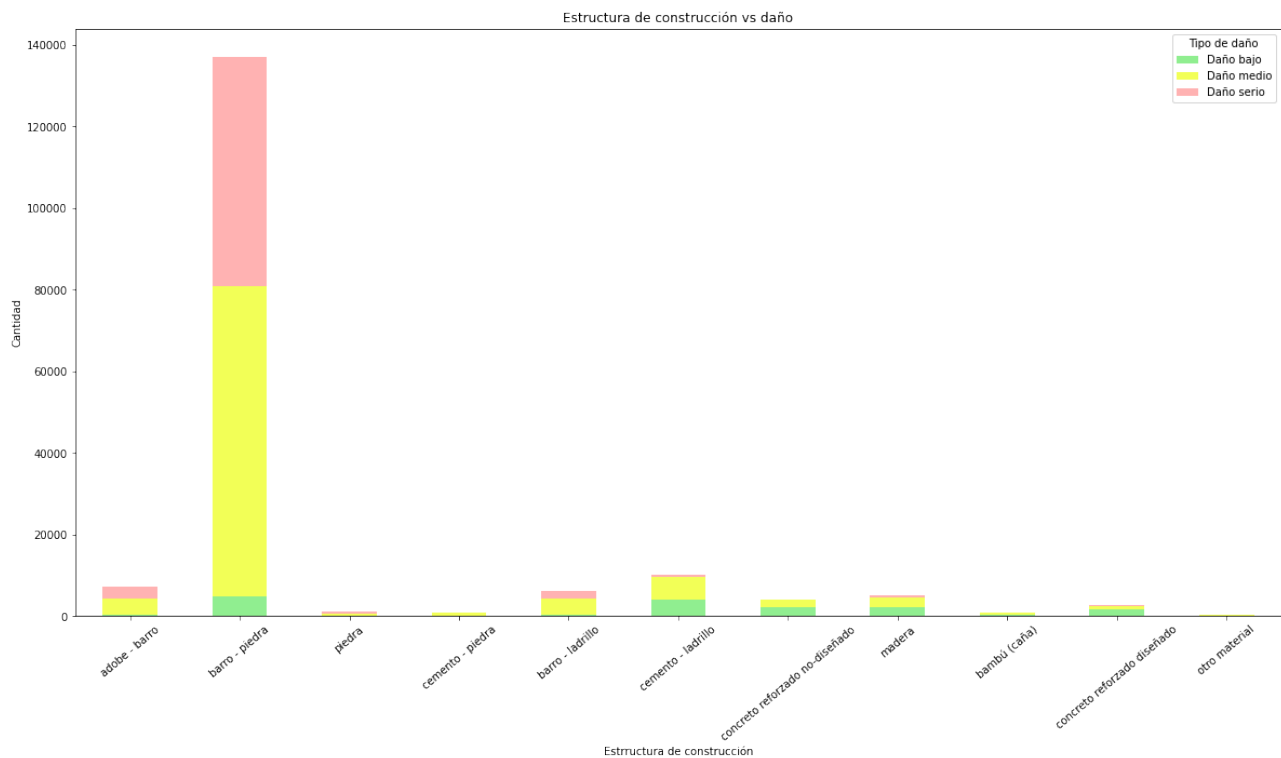


Figura 11: Cantidad de edificios por daño con una única estructura

- Teniendo en cuenta la estructura de construcción, ¿cuáles son las edificaciones que sufrieron mayor daño?

Las edificaciones construídas a base de piedra tuvieron en mayor parte daño serio. Por lo tanto podemos deducir que las edificaciones con esta estructura fueron las más vulnerables frente al terremoto. (figura 12 y cuadro 2)

Tipo de daño	Estructura: Piedra
Bajo	15
Medio	494
Serio	719

Cuadro 2: Cantidad de edificaciones con la estructura piedra según el tipo de daño

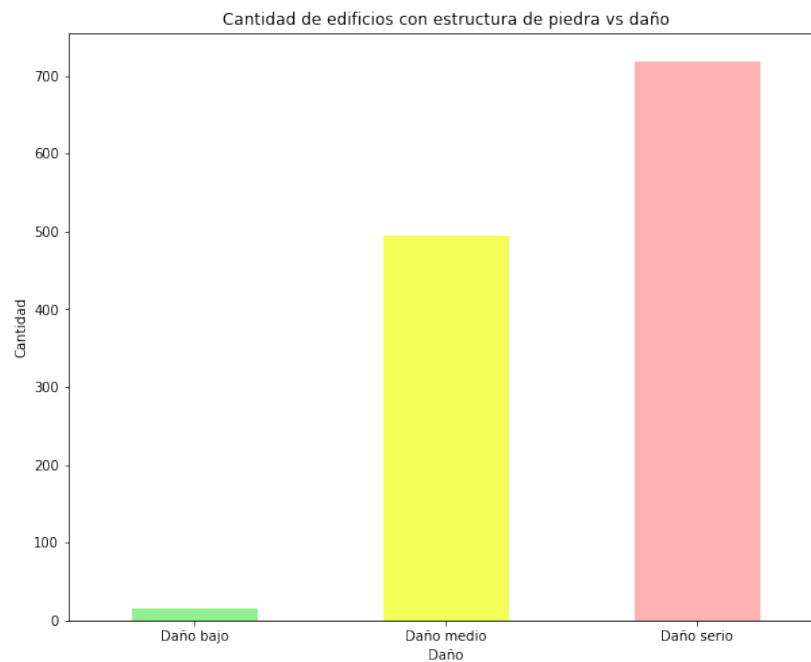


Figura 12: Cantidad de edificios con estructura piedra

■ ¿Y las que tuvieron menos impacto?

Las edificaciones con estructuras del tipo concreto reforzado diseñado y no-diseñado recibieron mayor daño bajo que cualquier otro daño. En consecuencia, observamos que estas estructuras fueron más eficientes. (figuras 13 y 14 y cuadros 3 y 4)

Tipo de daño	Estructura: Concreto reforzado diseñado
Bajo	1774
Medio	773
Serio	47

Cuadro 3: Cantidad de edificaciones con la estructura concreto diseñado según el tipo de daño

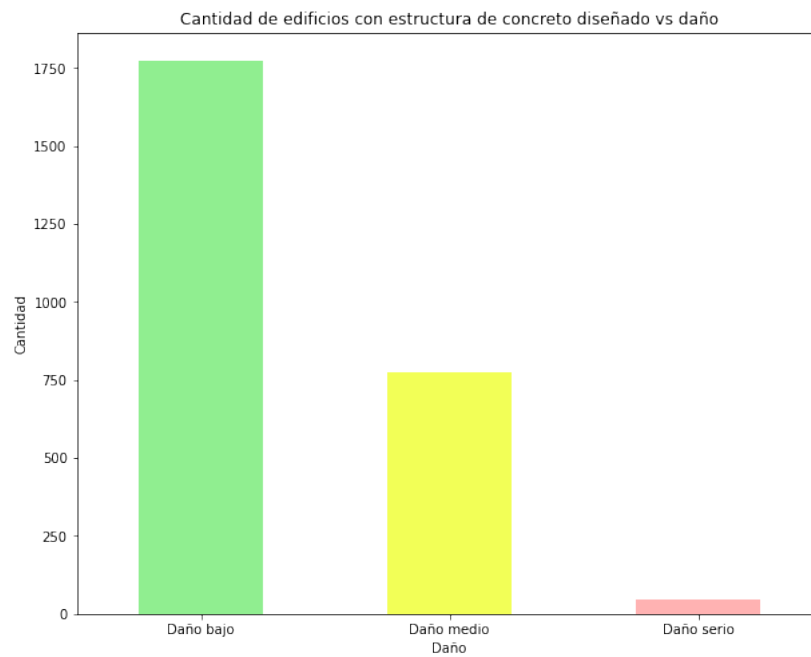


Figura 13: Cantidad de edificios con estructura concreto diseñado

Tipo de daño	Estructura: Concreto reforzado no-diseñado
Bajo	2307
Medio	1756
Serio	104

Cuadro 4: Cantidad de edificaciones con la estructura concreto no diseñado según el tipo de daño



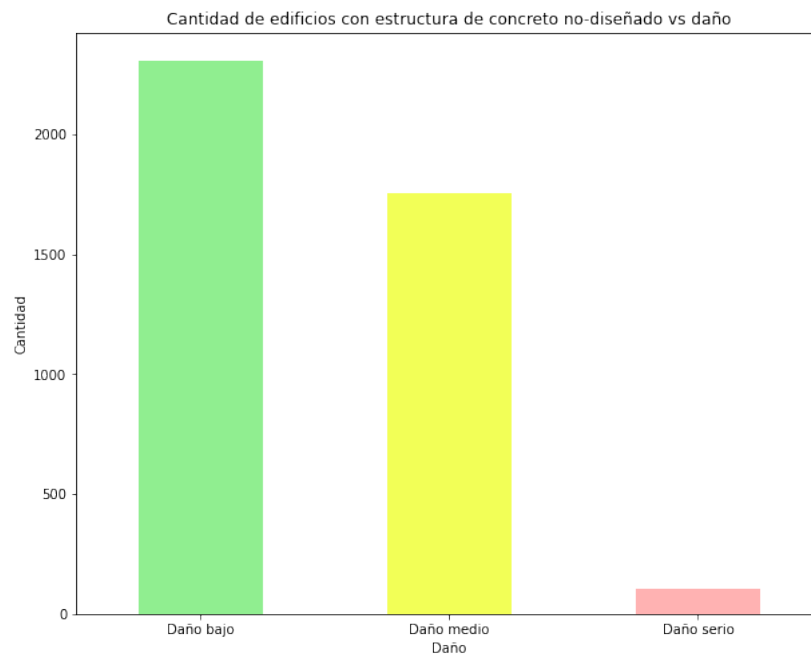


Figura 14: Cantidad de edificios con estructura concreto no-diseñado

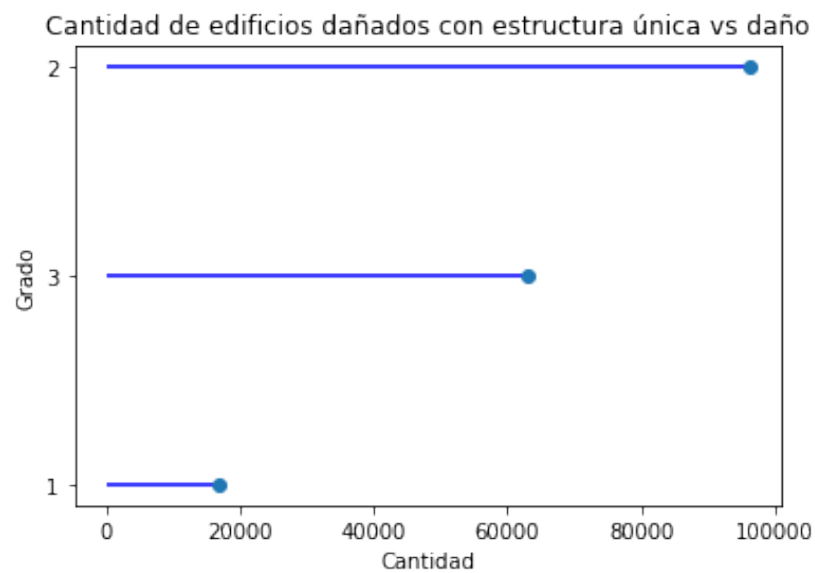


Figura 15: Cantidad de edificios dañados con estructura única

### 4.3. Localización

- ¿Cuál fue el porcentaje de destrucción para cada zona?

Como mencionamos en la sección 3.3 todos los análisis realizados sobre localizados fueron para los sectores de G1. En la figura 16 observamos que:

- El mayor porcentaje de destrucción está centrado en el daño de grado 2.
- El menor porcentaje de destrucción está centrado en el daño de grado 1.
- El porcentaje de destrucción de grado 3 tiene un comportamiento irregular y se puede observar principalmente en las zonas 17, 18 y 21.

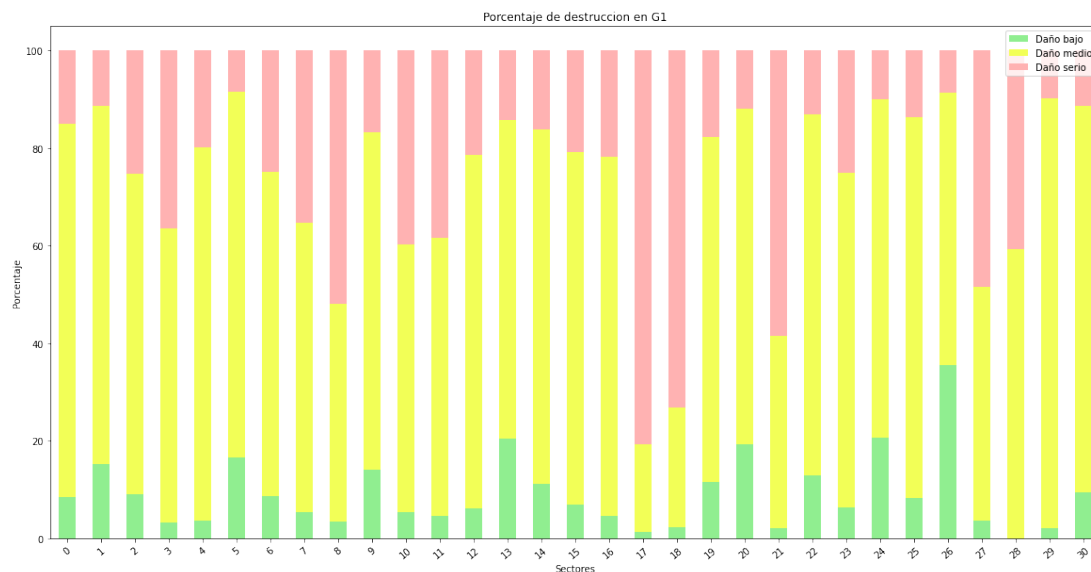


Figura 16: Cantidad de edificios dañados en G1

- Teniendo en cuenta la cantidad, ¿cuántos edificios sufrieron cada daño?

Considerando exclusivamente la cantidad de edificios que se vieron afectados por el terremoto, por la figura 17 llegamos a la conclusión de que:

- Dado que la cantidad de edificios varía según la zona, la consistencia de los porcentajes no se mantiene.
- La mayoría de los edificios sufrieron daño de grado 3 (dimensión círculo rojo), en segundo lugar daño de grado 2 (dimensión círculo amarillo) y por último daño de grado 1 (dimensión círculo verde).
- Hay sectores que se repiten. ¿Qué significa? Esos sectores tienen muchos más edificios que el resto, con lo cual su cantidad de edificios aparece como protagonista en más de una circunferencia.

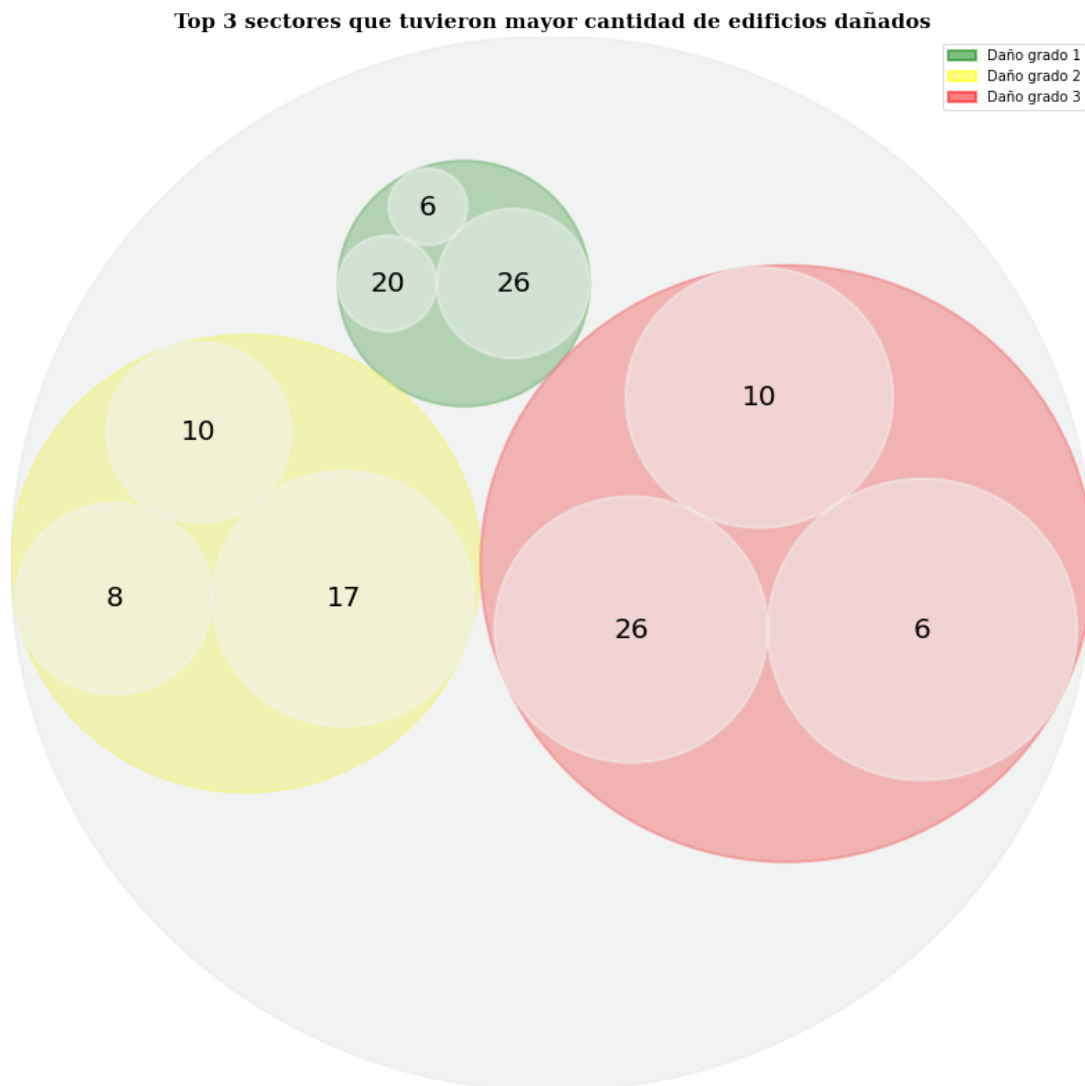


Figura 17: Top 3 sectores por daño

- ¿Existe algún contraste entre las zonas que tuvieron destrucción grado 1 (fondo verde) y las que tuvieron grado 3 (fondo rojo)? ¿Qué características se diferencian? ¿Pueden haber sido determinantes?

Se tomaron los promedios del área, la altura, la cantidad de pisos y la antigüedad de cada edificación como posibles causantes de la destrucción. Para ello, se tomaron los cinco sectores que tuvieron el mayor porcentaje de destrucción y que su cantidad de edificios totales superaba las 10000 edificaciones (figura 18).

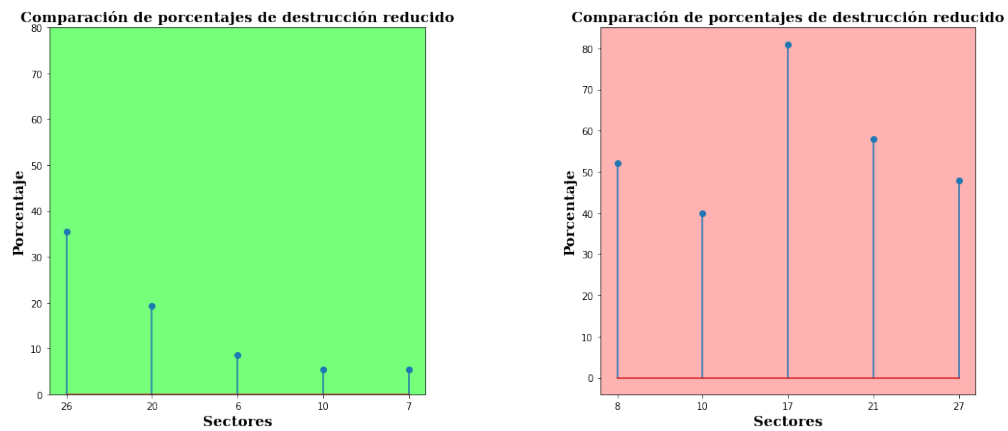


Figura 18: Top 5 porcentaje de destrucción grado 1 y grado 3 por sector

Para el caso del área no observamos una diferencia muy marcada entre ambos daños (figura 19) con lo cual en esos sectores esta característica es irrelevante.

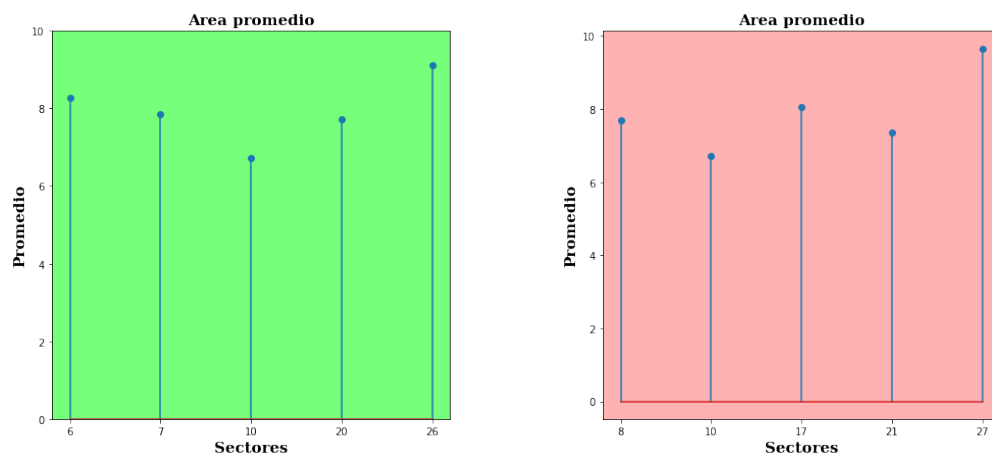


Figura 19: Top 5 área promedio de destrucción grado 1 y grado 3 por sector

Para el caso de la altura observamos en la figura 20 que existe una diferencia en la constancia de los valores. Mientras que las edificaciones que sufrieron grado 1 mantienen una altura promedio cercana al 5, las de grado 3 son mucho más cercanas al 6. Por lo tanto, una posibilidad es que a mayor altura mayor posibilidad de destrucción.

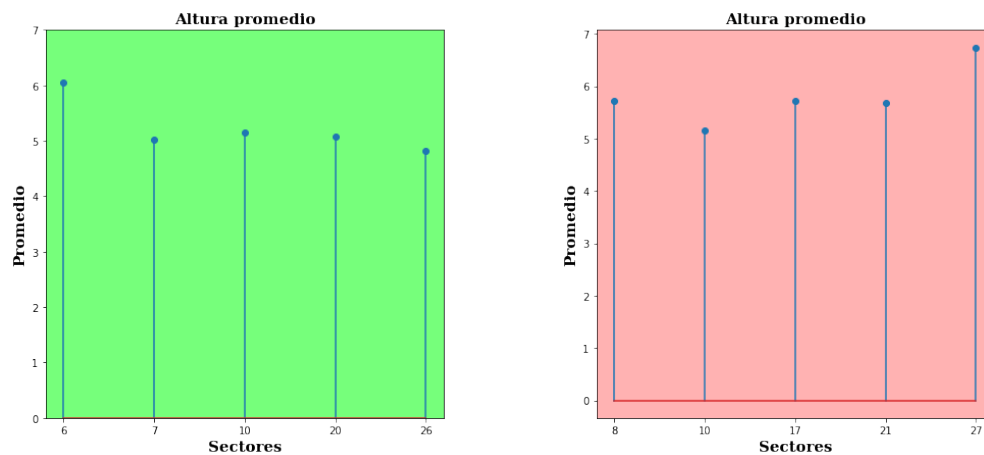


Figura 20: Top 5 altura promedio de destrucción grado 1 y grado 3 por sector

Para el caso de la cantidad de pisos promedio (figura 21) observamos que existe una diferencia más notoria. Principalmente se ve que el promedio más bajo de grado 3 alcanza un valor muy similar al máximo del promedio de grado 1. Al igual que la altura, podemos deducir que a mayor cantidad de pisos, mayor posibilidad de destrucción.

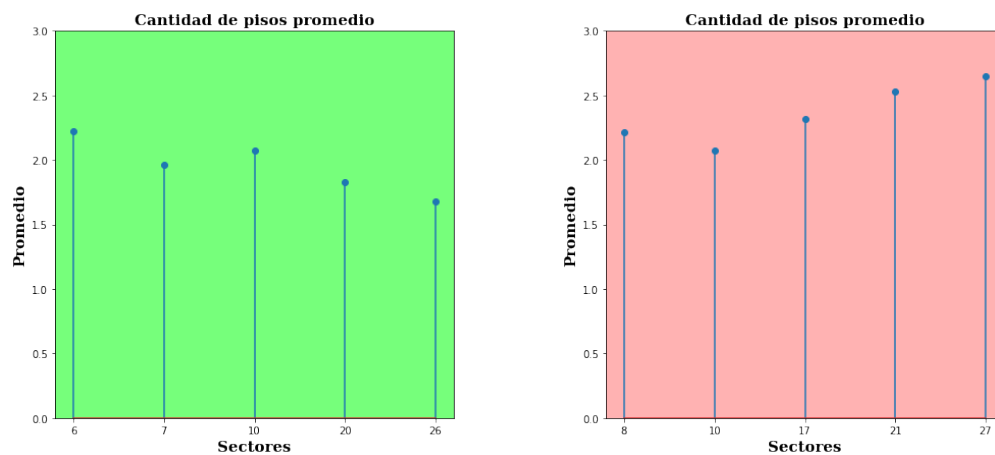


Figura 21: Top 5 cantidad de pisos promedio de destrucción grado 1 y grado 3 por sector

Para el caso de la antigüedad observamos en la figura 22 que no hay consistencia clara para comparar los datos. Sin embargo, en la zona 27 la edad promedio es elevada y, como se muestra en los gráficos 18, el promedio de destrucción serio es considerable; por lo cual no se descarta a la antigüedad como un factor determinante en la destrucción de las edificaciones.

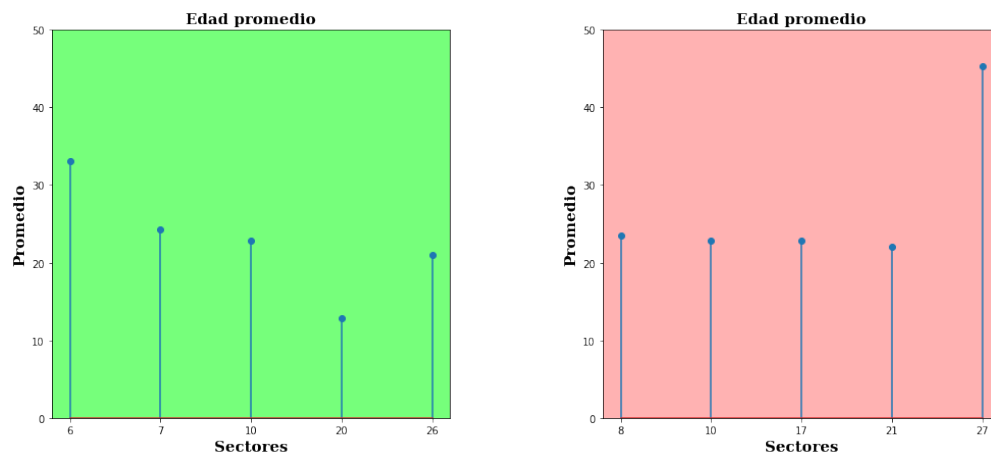


Figura 22: Top 5 antigüedad promedio de destrucción grado 1 y grado 3 por sector

#### 4.3.1. Características de la edificación

- ¿El tipo de suelo influye en el grado de destrucción?

En la figura 23 podemos notar que hay un tipo de suelo predominante en ambos daños pero como es el mismo no podemos sacar ninguna conclusión.

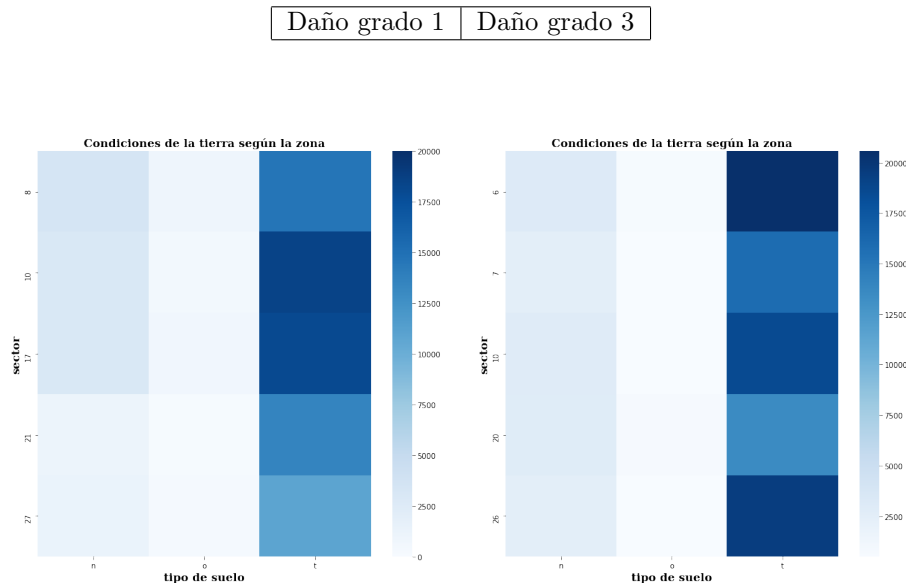


Figura 23: Top 5 tipo de suelo de destrucción grado 1 y grado 3 por sector

- ¿El tipo de cemento influye en el grado de destrucción?

Podemos observar en la figura 24 que a pesar de que se mantiene la consistencia en cuanto al tipo de cemento predominante, existen otros dos en el grado

3 que son considerablemente mayor al grado 1. Con lo cual podemos estimar que el tipo de cemento influyó en alguna medida en la destrucción.

Daño grado 1	Daño grado 3
--------------	--------------

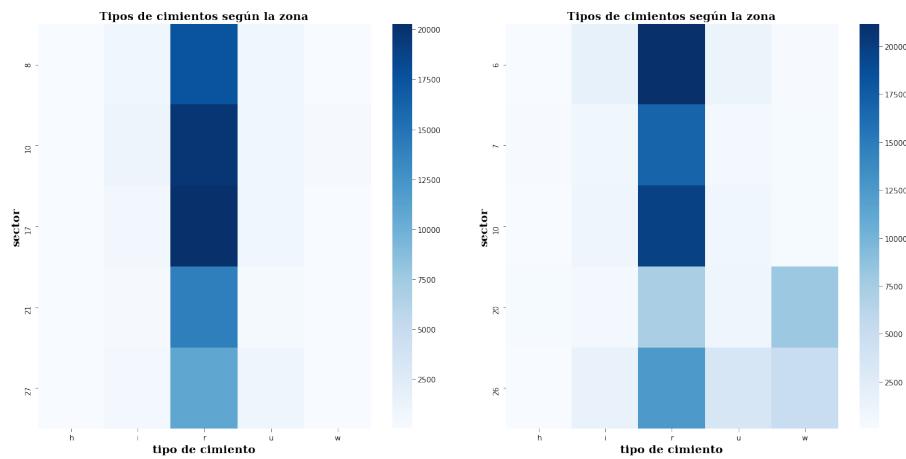


Figura 24: Top 5 tipo de cimiento de destrucción grado 1 y grado 3 por sector

- ¿El tipo de construcción influye en el grado de destrucción?

Podemos observar en la figura 25 que se mantiene la consistencia en cuanto al tipo de construcción predominante, con lo cual podemos estimar que el tipo de construcción no fue determinante en la destrucción.

Daño grado 1	Daño grado 3
--------------	--------------

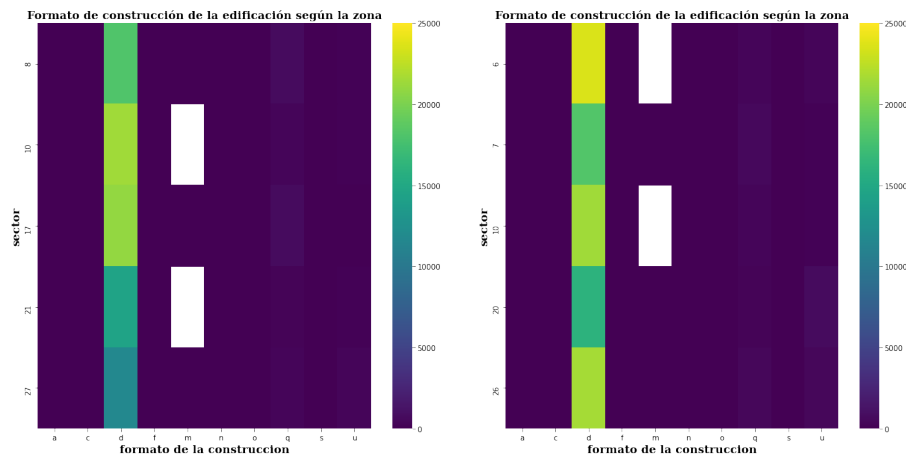


Figura 25: Top 5 formato de construcción de destrucción grado 1 y grado 3 por sector

## 4.4. Propiedades físicas

### 4.4.1. Altura

- ¿Los edificios más altos sufrieron peores daños?

Se puede observar en la figura 26 que a medida que aumenta la altura de la edificación, la diferencia entre el daño serio y el bajo se va incrementando, predominando el daño serio sobre el bajo. La mayoría de los edificios sufrieron un impacto medio.



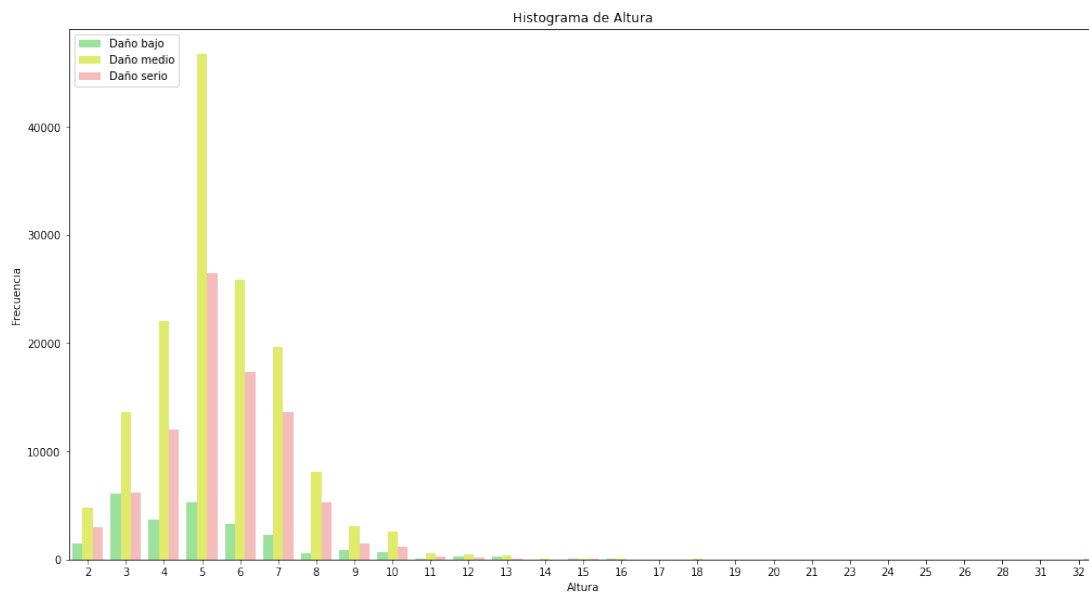


Figura 26: Cantidad de edificios por altura

#### 4.4.2. Área

- ¿Los edificios de mayor superficie recibieron peor impacto?

Observamos en la figura 27 un efecto campana en los valores de las áreas más pequeños. En los extremos de la misma la diferencia entre los tipos de daños no es tan notoria, en contraposición con lo que ocurre en el centro. Por lo cual deducimos que el área no necesariamente es un factor clave para determinar el tipo de daño.

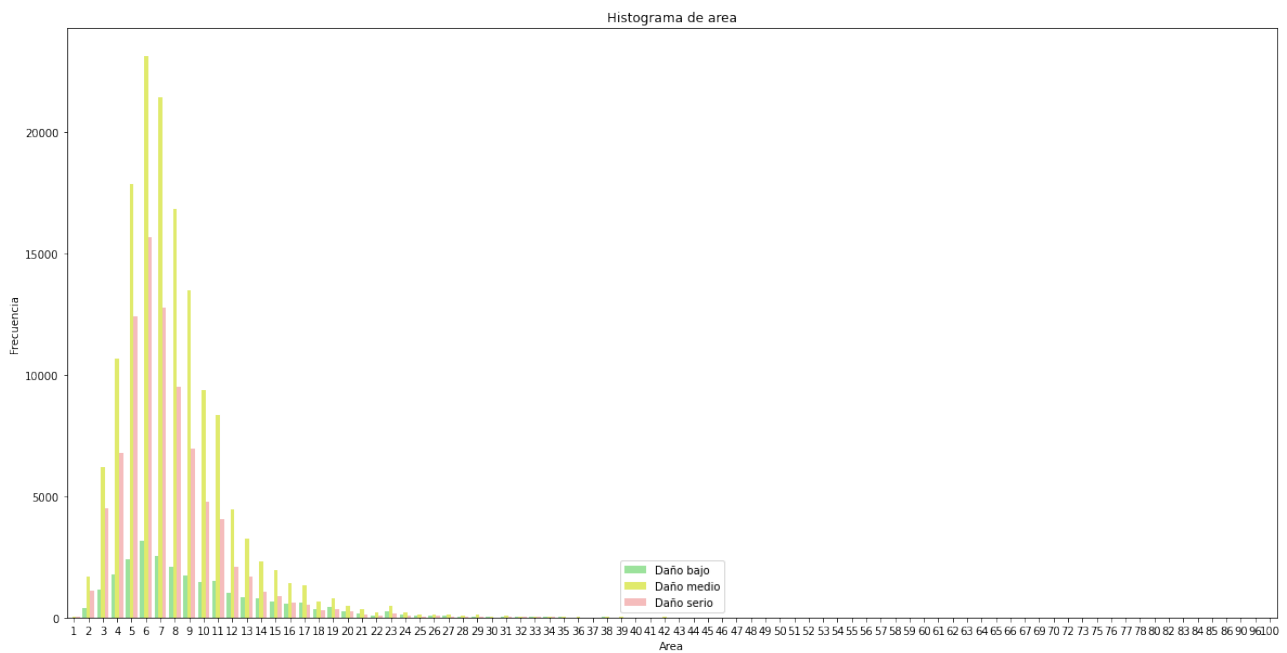


Figura 27: Cantidad de edificios por área

#### 4.5. Cantidad de pisos

- ¿Los edificios con mayor cantidad de pisos recibieron peor impacto?

Observamos en la figura 28 que a medida que la edificacion tiene más pisos el impacto es más fuerte. A mayor cantidad de pisos mayor daño serio.

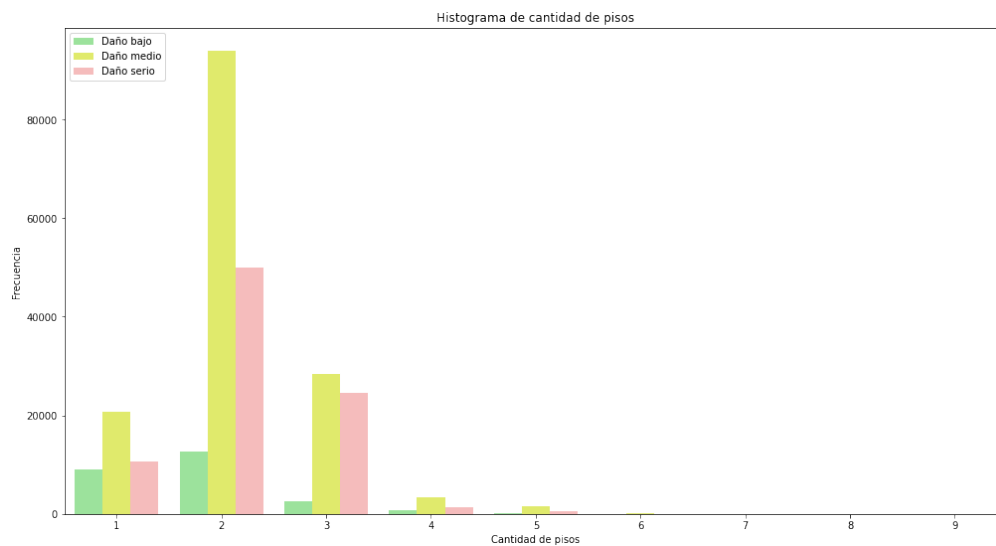


Figura 28: Cantidad de edificios por área

#### 4.6. Uso secundario

- ¿Cómo se distribuyó el daño según el uso secundario de la edificación?

Visualizamos en la figura 29 que en los tres grados de daño el porcentaje de edificios con uso secundario es muy bajo en relación a los que no tienen, por lo cual no lo consideramos un factor influyente.

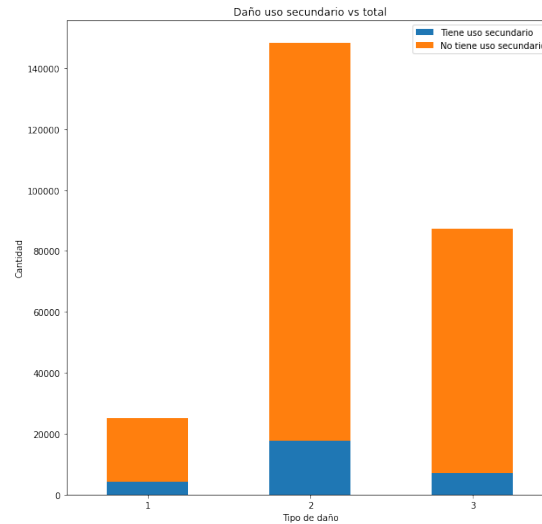


Figura 29: Cantidad de edificios por daño dependiendo su uso

En las figuras 30, 31 y 32 analizamos la relación entre el uso secundario de los edificios y el tipo de daño que sufrieron, teniendo en cuenta la cantidad y el porcentaje total de los mismos. Notamos que el daño medio predomina en todos los usos, sin embargo, en algunos casos la cantidad de daños serios es mayor a la de daños bajos y viceversa.

Notamos que las edificaciones con uso secundario para alquiler e institución fueron las que sufrieron daños bajos, en cambio las destinadas a agricultura, estación de policía e industria tuvieron un daño más severo.

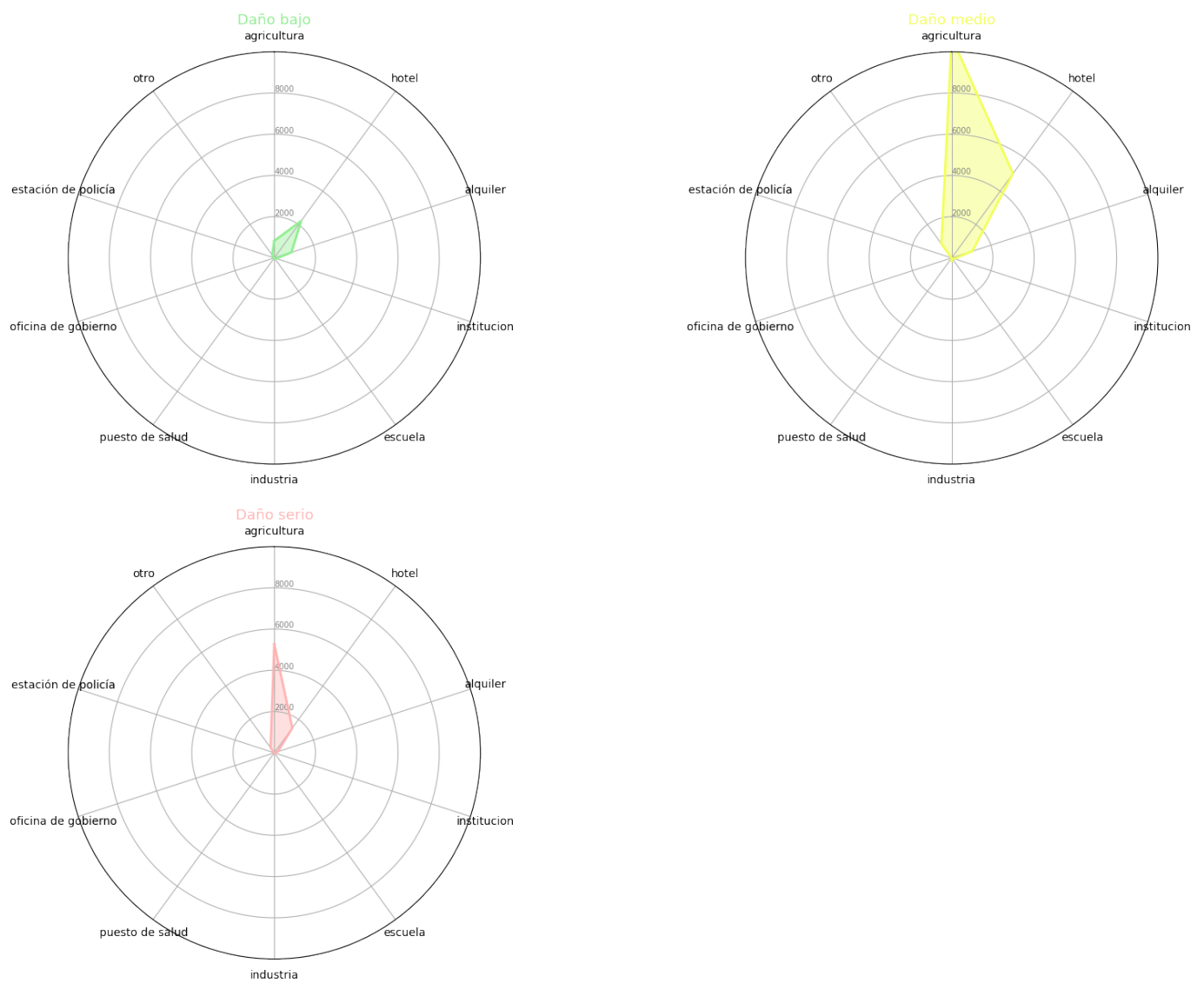


Figura 30: Cantidad de edificios por daño dependiendo su uso secundario

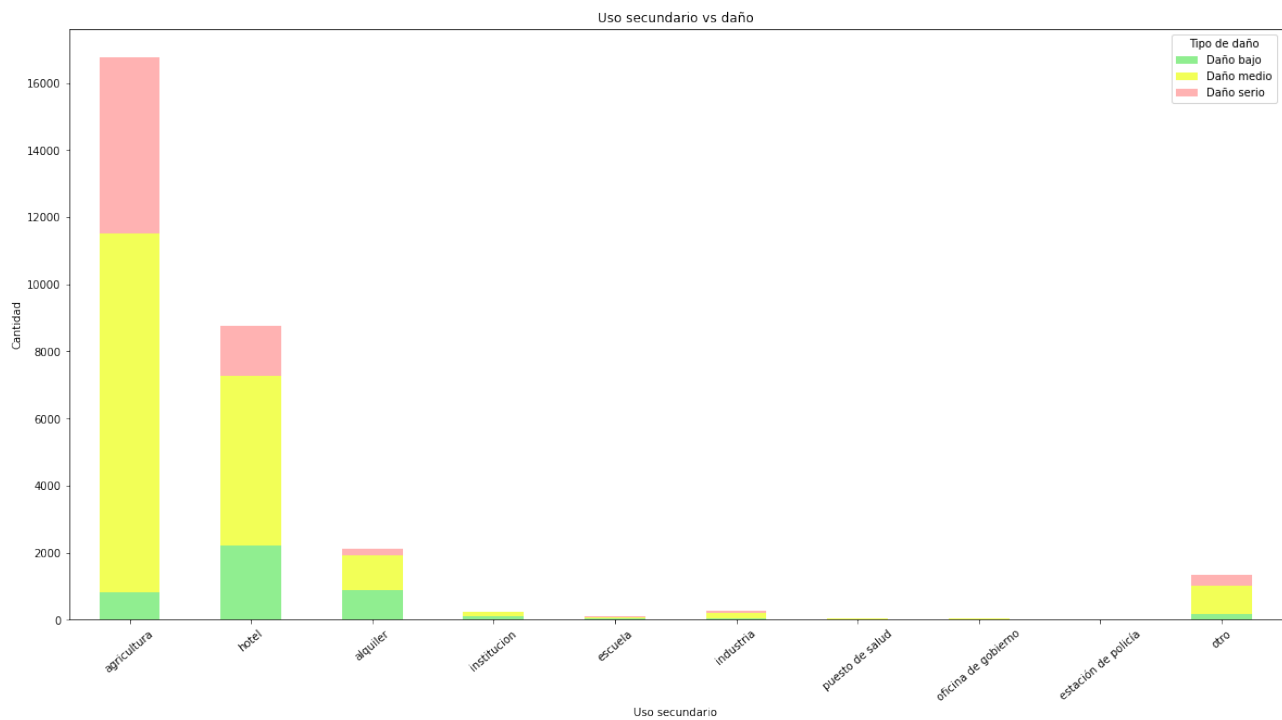


Figura 31: Cantidad de edificios por daño dependiendo su uso secundario

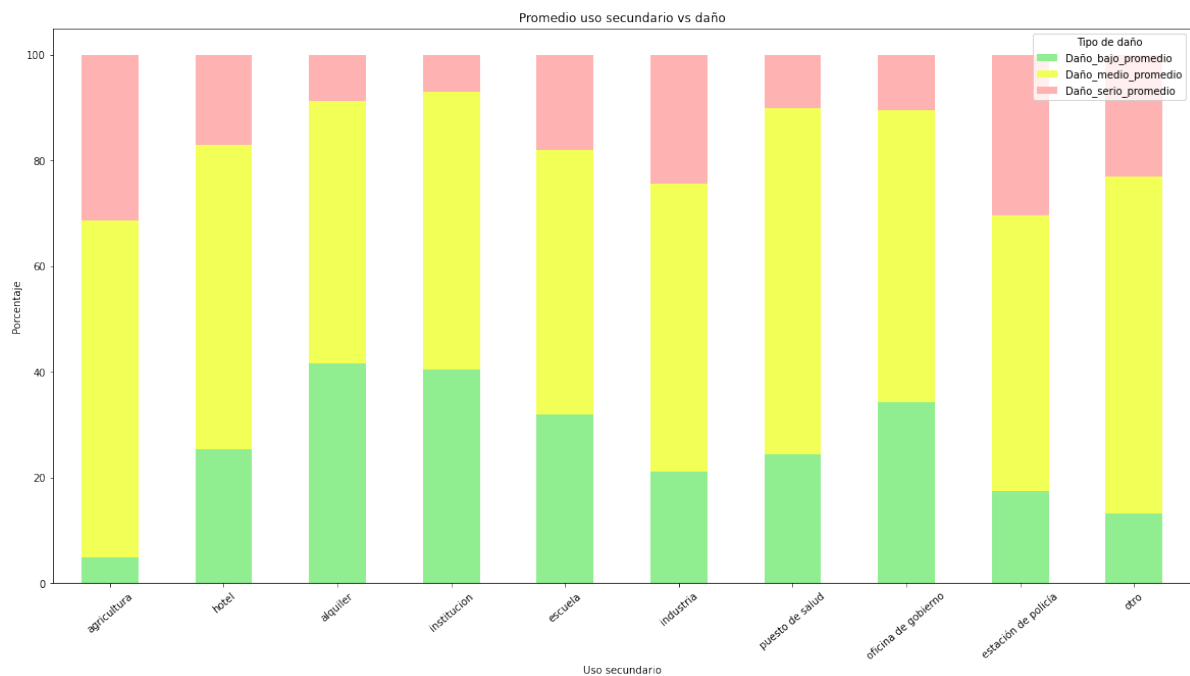


Figura 32: Cantidad de edificios por daño dependiendo su uso secundario

## 4.7. Social

- ¿Cuántas familias perdieron su hogar?

Podemos observar que fueron muchas las familias afectadas tras el fenómeno ocurrido.

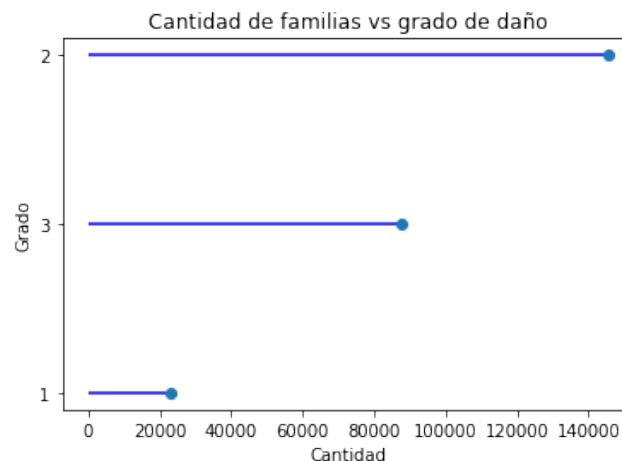


Figura 33: Cantidad de edificios por daño dependiendo su uso secundario

## 5. Relación entre los daños

Una vez realizado un análisis previo de cada característica por separado y su relación con el daño, buscamos deducir si su combinación es determinante en el resultado de la destrucción de la edificación.

Tomamos como factores claves la antigüedad, la altura, la cantidad de pisos y el tipo de estructura a partir de las conclusiones de la sección previa.

Llegamos a que:

- A partir de la altura 4 la diferencia entre cantidad de edificios dañados de grado 1 y grado 3 comienza a tener una brecha amplia (figura 26) y el daño de grado 1 es decreciente.
- A partir de los 2 pisos la diferencia entre cantidad de edificios dañados de grado 1 y grado 3 comienza a tener una brecha amplia (figura 28) y el daño de grado 1 es decreciente.
- A medida que la antigüedad del edificio aumenta, la cantidad de edificios con daño serio comienza a tener una brecha ancha con respecto al grado 1, el cual va bajando a medida que suben los años (figura 8). El momento donde la diferencia se ve mas asentada es a partir de los 35 años de antigüedad.
- Como ilustra la figura 12 los edificios construidos a base de piedra fueron los que mas daños sufrieron, por lo tanto podemos deducir que esta estructura fue la mas debil frente a este tipo de fenómeno.

Comenzamos filtrando las edificaciones que tenían una altura mayor a 4 y una cantidad de pisos mayor a 2. La cantidad total de edificios que cumplen con estas condiciones es de 218067, este valor se ve reflejado en el gráfico 34.

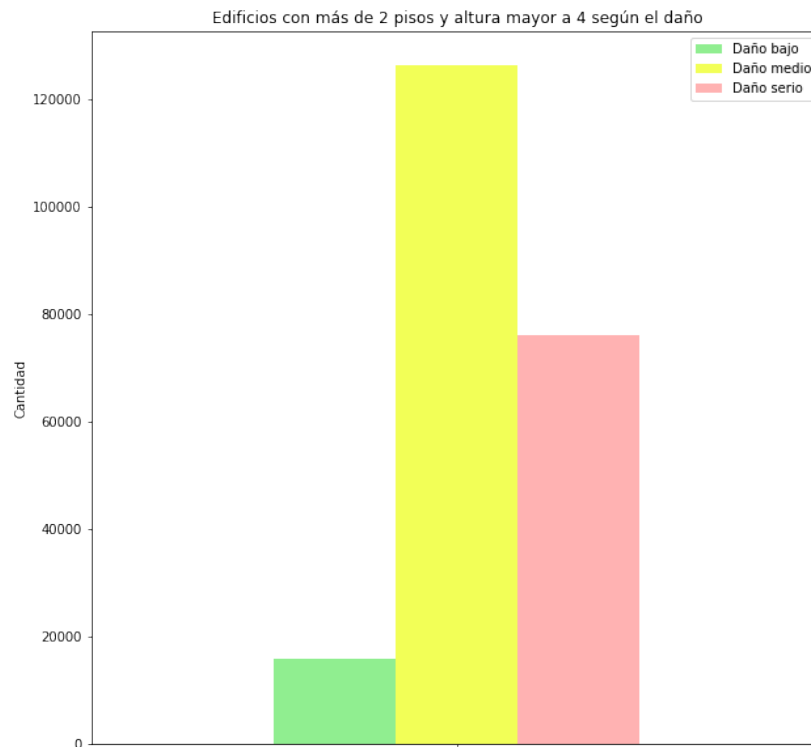


Figura 34: Cantidad de edificios con más de 2 pisos y altura mayor a 4 según el daño

Manteniendo las dos características anteriores y agregando la condición de antigüedad mayor a 35 obtenemos un total de 47159 edificaciones, cuyo desempeño se puede ver en la figura 35

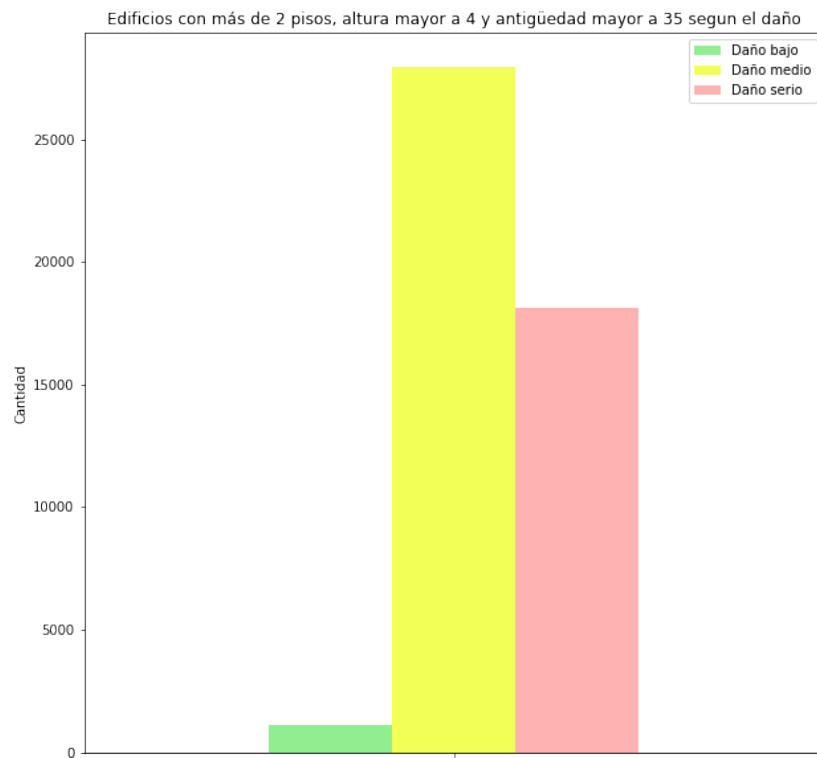


Figura 35: Cantidad de edificios con más de 2 pisos, altura mayor a 4 y antigüedad mayor a 35 según el daño

Por último, manteniendo los 3 aspectos anteriores, incorporamos la restricción de que la estructura sea exclusivamente de piedra. Obtuvimos un total de 1673 edificaciones (figura 36)



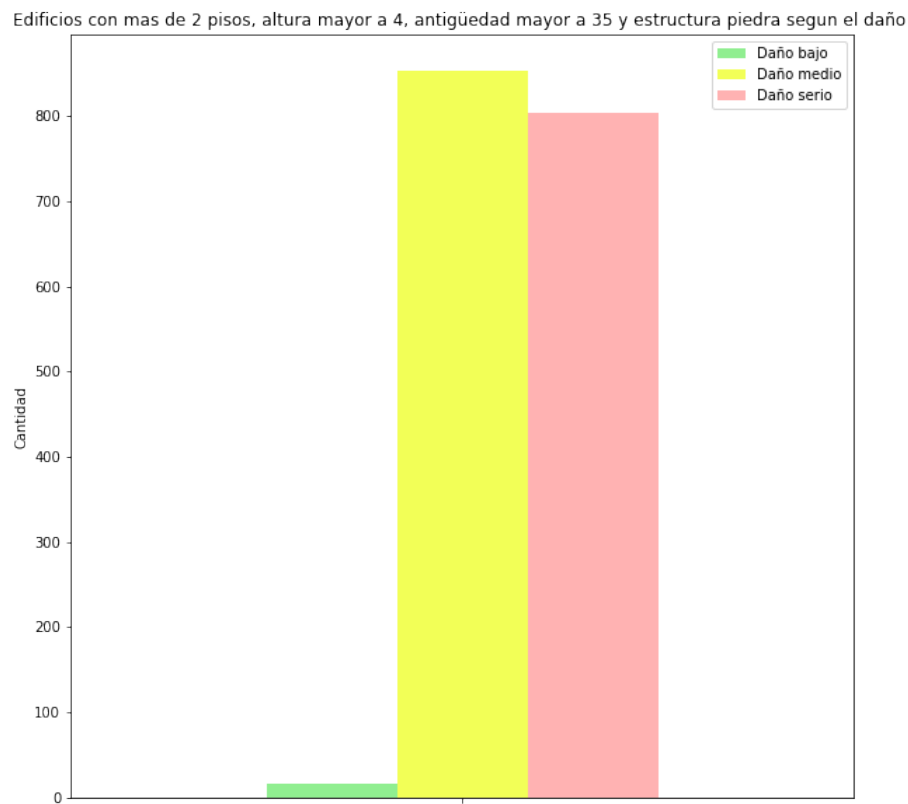


Figura 36: Cantidad de edificios con más de 2 pisos, altura mayor a 4, antigüedad mayor a 35 y estructura piedra según el daño

En resumen, vemos como a medida que las características que consideramos como determinantes se incorporan al filtrado de las edificaciones, la cantidad de edificios con daño bajo se reduce considerablemente hasta volverse despreciable en comparación a las de daño serio. Por otro lado, la brecha entre las edificaciones con daño medio y daño serio decrece.

## 6. Conclusiones

En síntesis, el trabajo práctico propuso realizar un análisis exploratorio de un set de datos relacionados con el terremoto ocurrido en Nepal en el año 2015. Para comenzar a elaborar diferentes hipótesis acerca de por qué las edificaciones sufrieron mayor o menor grado de destrucción, seccionamos el set de datos de manera tal que las características puedan haber influido de manera directa en el resultado.

Trabajamos sobre cada aspecto de las edificaciones de manera particular y específica con el objetivo de ver si contribuían con el tipo de daño causado. Encontramos datos que presentaban patrones similares para los tres tipos de daños, los cuales nos llevaron a concluir que no fueron determinantes. Por otro lado, otros set de datos sí tenían una diferencia marcada y nos ayudo a generar

una posible causa de la destrucción.

Principalmente utilizamos representaciones gráficas para visualizar los datos de forma mas simple y práctica.

Para nuestro análisis tomamos aspectos tanto propios de la edificación como externos y llegamos a la conclusión de que factores como la altura, la cantidad de pisos, la antigüedad y el tipo de estructura fueron relevantes a la hora de hacerle frente al fenomeno natural. Además, las edificaciones que combinaban todas ellas fueron las mas afectadas por el terremoto.