

75.06/95.58 Organización de Datos
Primer Cuatrimestre 2021
Trabajo Práctico 1



Grupo:
En busca del dato perdido

Apellido/s	Nombre/s	Padrón	E-mail
Albornoz	Gonzalo Javier	103554	gjalbornoz@fi.uba.ar
Inneo Veiga	Sebastian Bento	100998	sinneo@fi.uba.ar
Massone	Mario Bernardo	102141	mbmassone@fi.uba.ar

Introducción	2
Objetivo	2
Análisis general	3
Limpieza de datos	5
Repositorio Github	5
Correlaciones	6
count_floors_pre_eq: Análisis por cantidad de pisos	6
area_percentage: Análisis por superficie normalizada ocupada por el edificio	10
height_percentage: Análisis por altura normalizada ocupada por el edificio	12
land_surface_condition: Análisis de daño para la condición de superficie n	14
position: Análisis por orientación del edificio	18
plan_configuration: Análisis por formato de construcción de la edificación (diseño sísmico)	18
age: Análisis por antigüedad	20
has_superstructure: Análisis por superestructuras	25
age vs has_superstructure	37
has_secondary_use	41
geo_level_x_id: Análisis por geolocalización	42
geo_level_1_id	42
geo_level_2_id	44
geo_level_3_id	45
Zona 17 del nivel 1(geo_level_1_id)	46
Zona 17 por antigüedad:	46
Zona 17 por cantidad de pisos:	48
Zona 17 por materiales/superestructuras:	50
Zona 6 del nivel 1(geo_level_1_id)	51
Zona 6 por antigüedad:	51
Zona 6 por cantidad de pisos:	53
Zona 6 por materiales/superestructuras:	55
Conclusión	56

Introducción

En este informe se analizará las distintas características del dataset de Richter's Predictor: Modeling Earthquake Damage de DataDriven, el cual brinda aspectos de ubicación y de construcción de edificios que sufrieron daños causados por el terremoto de Gorkha de 2015 en Nepal.

Estos datos fueron recopilados a través de encuestas realizadas por Kathmandu Living Labs y la Oficina Central de Estadísticas, que trabaja bajo la Secretaría de la Comisión Nacional de Planificación de Nepal. Esta encuesta es uno de los conjuntos de datos posteriores a un desastre más grandes jamás recopilados, que contiene información valiosa sobre los impactos del terremoto, las condiciones de los hogares y las estadísticas socioeconómicas y demográficas.

Los archivos que se utilizaron son train_labels.csv y train_values.csv.

Objetivo

¿Qué características tienen las edificaciones más dañadas?

Se plantea como objetivo hallar características relacionadas y/o aisladas que ayuden determinar el mayor daño en los edificios.

Para ellos se divide el análisis por estructura de diseño, materiales de composición de estructuras, edad y antigüedad, y geolocalización.

Análisis general

En ambos archivos hay 260601 filas, train_values tiene 39 columnas (29 son numéricas y 10 son object) y train_labels tiene 2 columnas (ambas numéricas). En ninguno de los archivos hay valores nulos. A continuación se explica qué representa cada columna:

train_labels.csv

	building_id	damage_grade
0	802906	3
1	28830	2
2	94947	3
3	590882	2
4	201944	3

building_id: identificador único de la edificación.

damage_grade: daño que recibió la edificación (1: daño menor, 2: daño medio, 3: daño mayor)

train_values.csv

	building_id	geo_level_1_id	geo_level_2_id	geo_level_3_id	count_floors_pre_eq	age	area_percentage	height_percentage	land_surface_condition
0	802906	6	487	12198	2	30	6	5	t
1	28830	8	900	2812	2	10	8	7	o
2	94947	21	363	8973	2	10	5	5	t
3	590882	22	418	10694	2	10	6	5	t
4	201944	11	131	1488	3	30	8	9	t

	foundation_type	... has_secondary_use_agriculture	has_secondary_use_hotel	has_secondary_use_rental	has_secondary_use_institution
	r ...	0	0	0	0
	r ...	0	0	0	0
	r ...	0	0	0	0
	r ...	0	0	0	0
	r ...	0	0	0	0

building_id: identificador único de la edificación.

geo_level_1_id, **geo_level_2_id** y **geo_level_3_id**: región geográfica en la cual la edificación existe desde la más general (1) hasta la más específica (3).

count_floors_pre_eq: cantidad de pisos que tenía la edificación antes del terremoto.

age: edad de la edificación en años.

area_percentage: superficie normalizada ocupada por la edificación.

height_percentage: altura normalizada ocupada por la edificación.

land_surface_condition: condición de la superficie terrestre donde el edificio fue construido.

foundation_type: tipo de cimientos usados cuando se construyó la edificación.

roof_type: tipo de techo usado cuando se construyó la edificación.

ground_floor_type: tipo de construcción usado en la planta baja cuando se construyó la edificación.

other_floor_type: tipo de construcción usado en otros pisos cuando se construyó la edificación.

position: orientación de la edificación.

plan_configuration: formato de construcción de la edificación.

has_superstructure_adobe_mud,

has_superstructure_mud_mortar_stone,

has_superstructure_stone_flag,

has_superstructure_cement_mortar_stone, **has_superstructure_mud_mortar_brick**,

has_superstructure_cement_mortar_brick,

has_superstructure_timber,

has_superstructure_bamboo,

has_superstructure_rc_non_engineered,

has_superstructure_rc_engineered y

has_superstructure_other: indican el material con el que la edificación fue construida (1 si la edificación tenía ese material, 0 en caso contrario).

legal_ownership_status: estado legal de la tierra donde la edificación fue construida.

count_families: número de familias que vivían en la edificación.

has_secondary_use: variable que indica si la edificación era usada con un uso secundario.

has_secondary_use_agriculture,

has_secondary_use_hotel,

has_secondary_use_rental,

has_secondary_use_institution,

has_secondary_use_school,

has_secondary_use_industry,

has_secondary_use_health_post,

has_secondary_use_gov_office,

has_secondary_use_use_police y

has_secondary_use_other: si las edificaciones tenían uso secundario indican cuál (1 para verdadero, 0 para falso).

Limpieza de datos

Para nuestra sorpresa todas las columnas y filas contenían datos y eran lo suficientemente variados como para no necesitar descartar ninguno así que se optó por dejar los datos tal cual estaban.

Repositorio Github

Todos los análisis realizados a continuación se encuentran en el siguiente repositorio:
<https://github.com/SBen-IV/TP-OrgaDeDatos>.

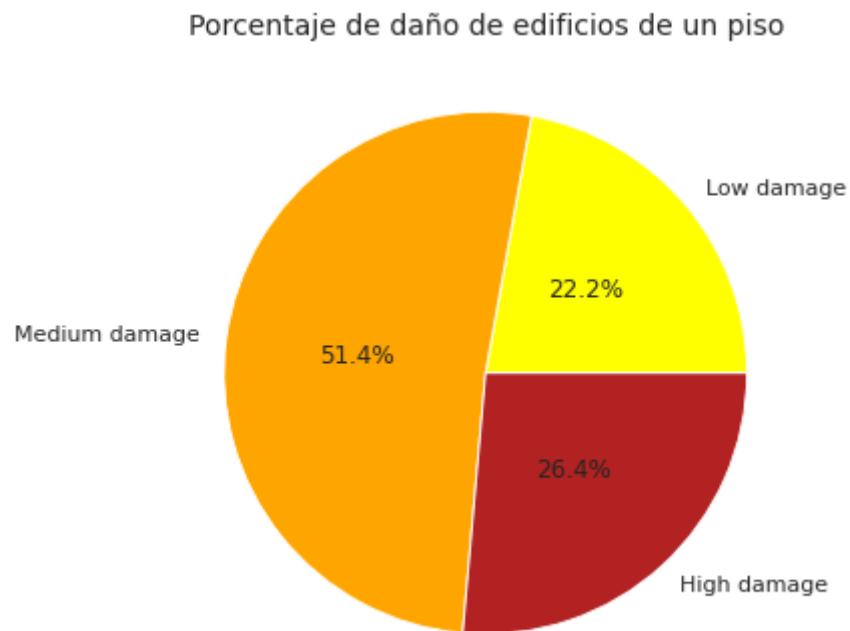
Correlaciones

Los primeros análisis hechos fueron ver las columnas por separado y luego ir juntando de a poco dependiendo de los resultados que íbamos obteniendo.

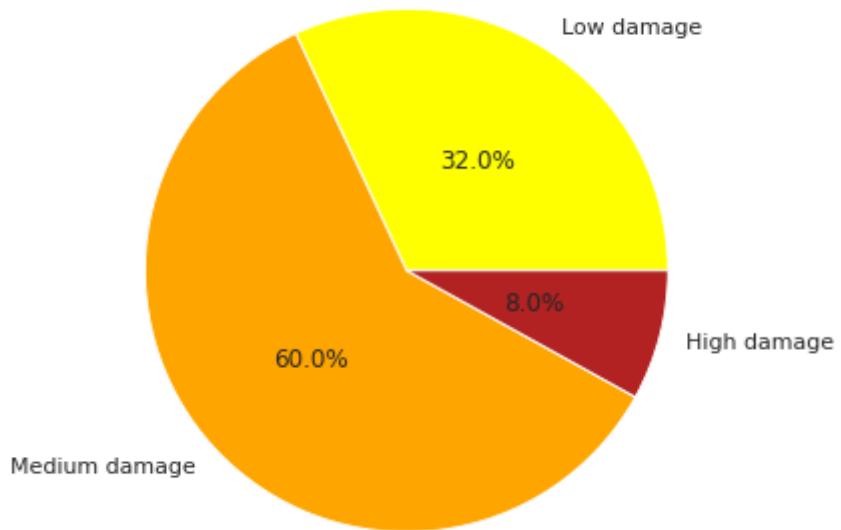
count_floors_pre_eq: Análisis por cantidad de pisos

Primeramente se observa que se cuenta con edificios de 1 hasta 9 pisos, sin embargo no hay muestras de gran cantidad para aquellos cuya cantidad de pisos es de 4 o más, por lo que se decidió agrupar a estos últimos.

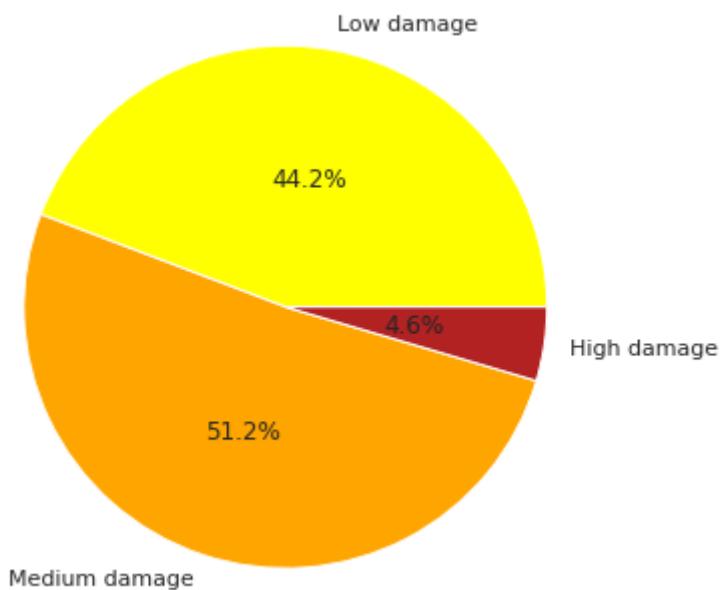
De los análisis porcentuales se obtuvo:



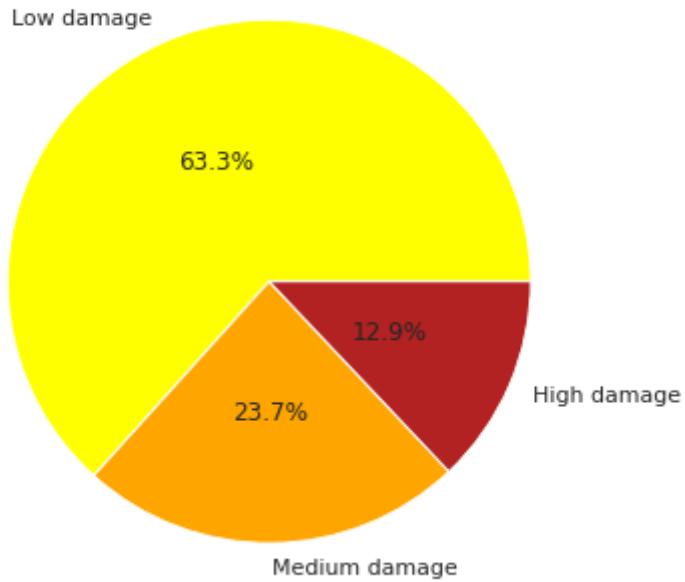
Porcentaje de daño de edificios de dos pisos



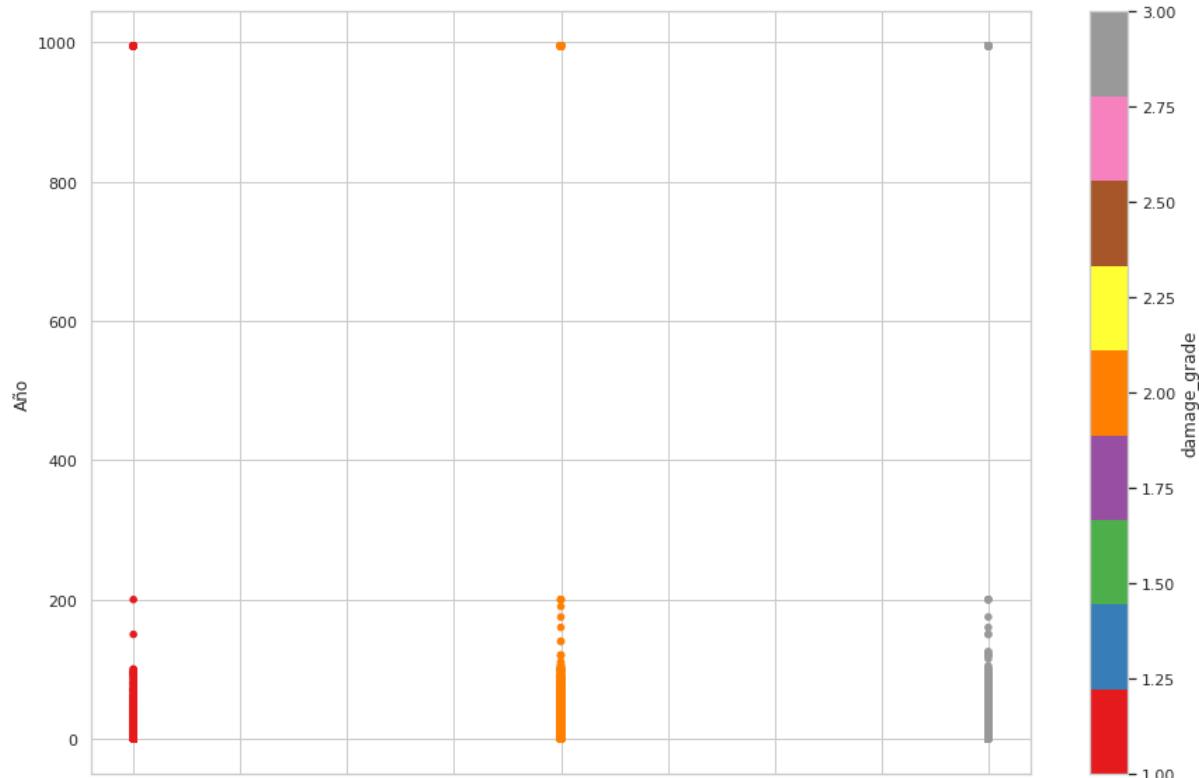
Porcentaje de daño de edificios de tres pisos



Porcentaje de daño de edificios de 4 o más pisos

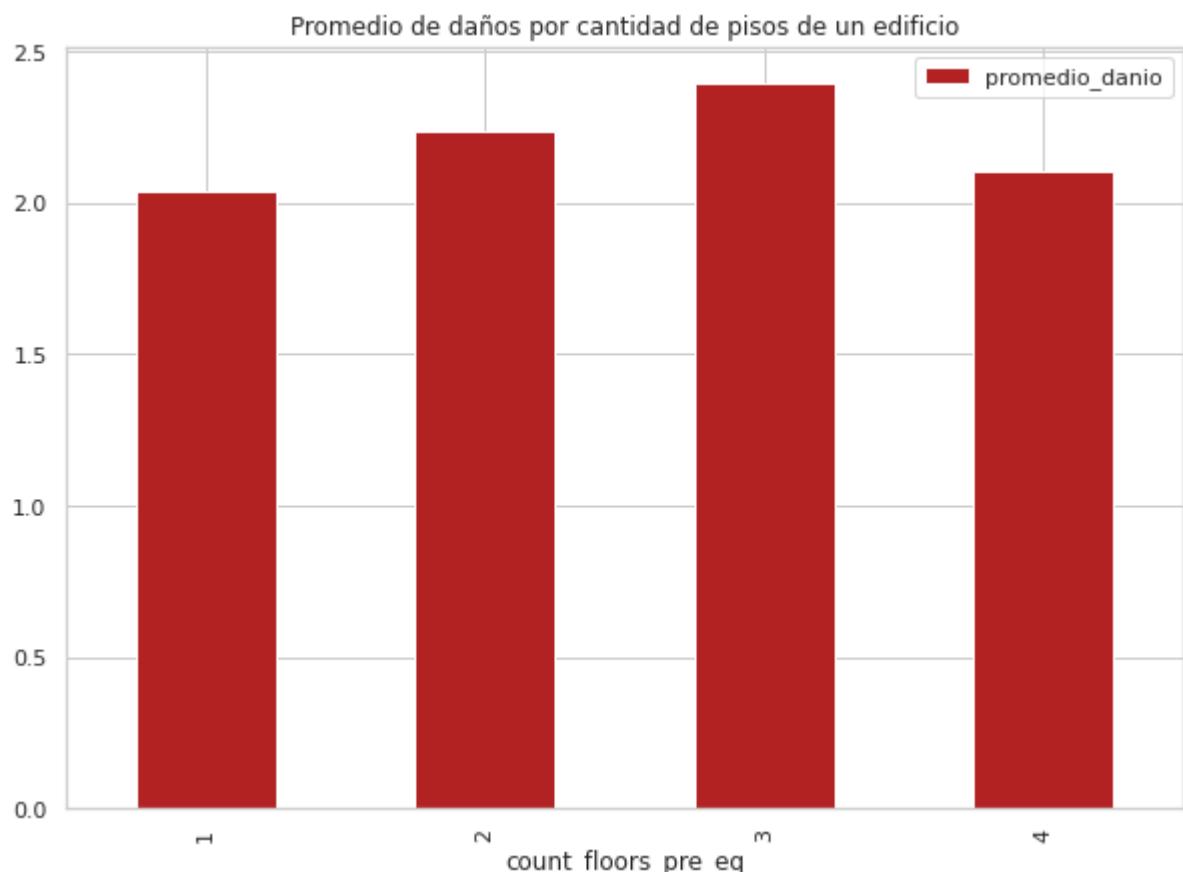


Los edificios de un solo piso son los que sufrieron mayor daño en porcentaje. Se estudia la dispersión de daño y antigüedad de estos edificios:



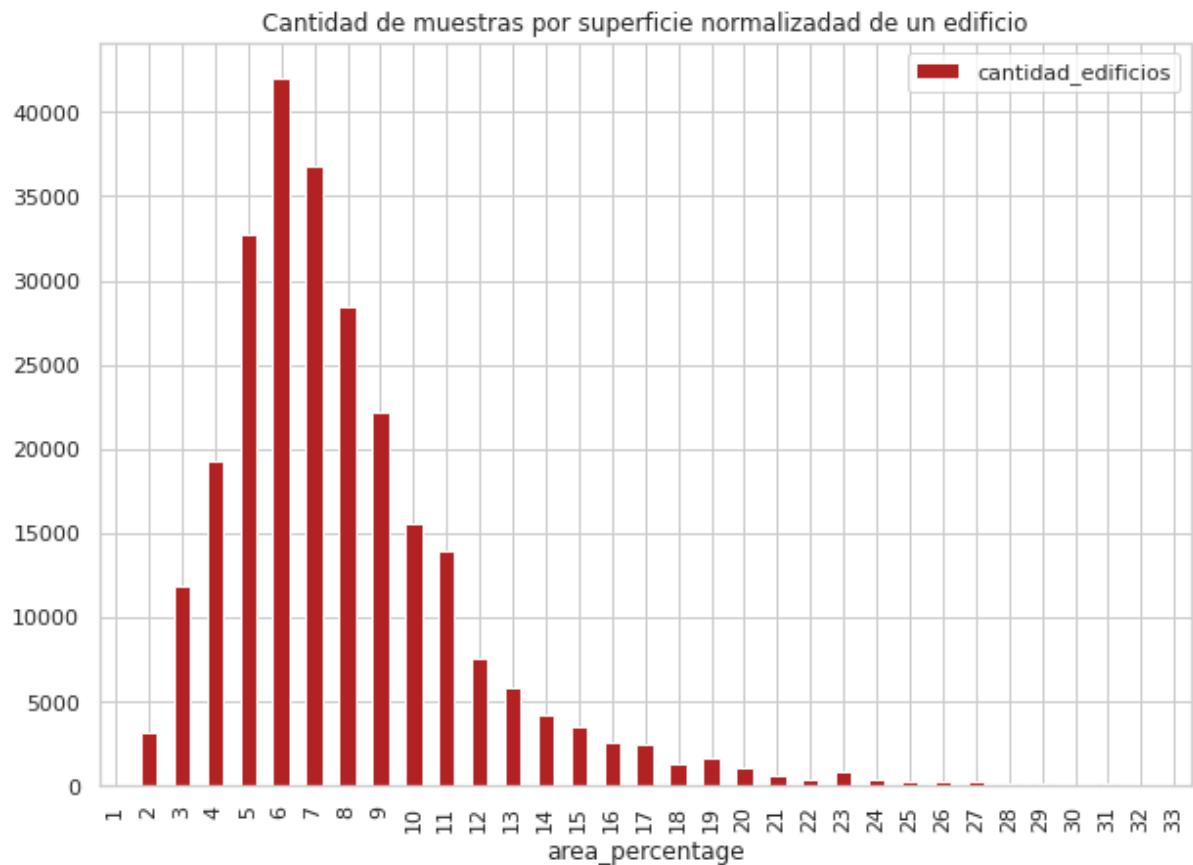
Se observa que para los edificios de un piso las edades de estos no tienen relación directa con el grado de daño, ya que se cuenta varias muestras de edificios de mismas edades con distintos tipos de daños.

También se analizó los valores promedios de daño para edificios de cada cantidad de pisos y se obtuvo:



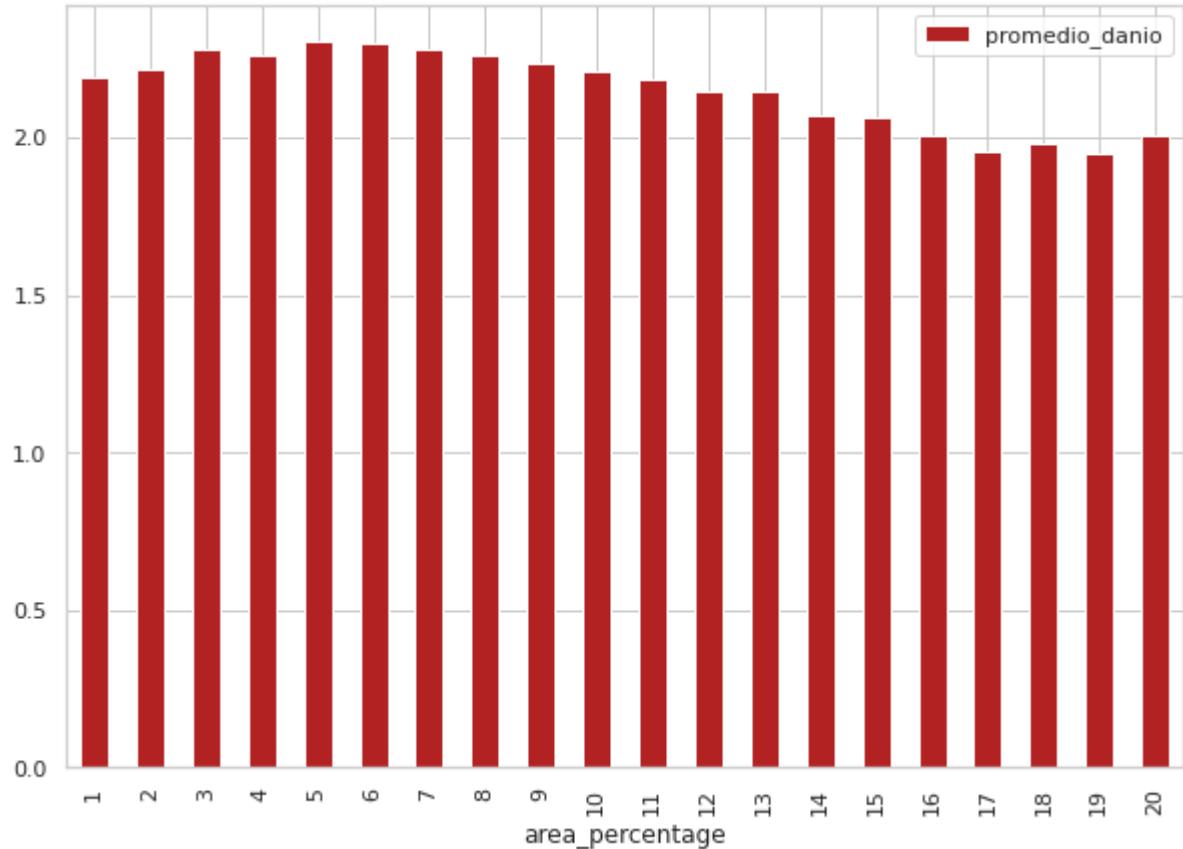
En promedio el daño de los edificios se mantiene entre medio y alto para las distintas cantidades de pisos, con una leve diferencia para los de 3 pisos.

area_percentage: Análisis por superficie normalizada ocupada por el edificio



Al analizar las muestras y sus cantidades, se observó que para edificios con superficie normalizada mayor a 20 se cuenta con escasa muestra, por lo cual se descartó para los siguientes pasos del análisis.

Promedio de daño segun superficie normalizada ocupada por el edificio

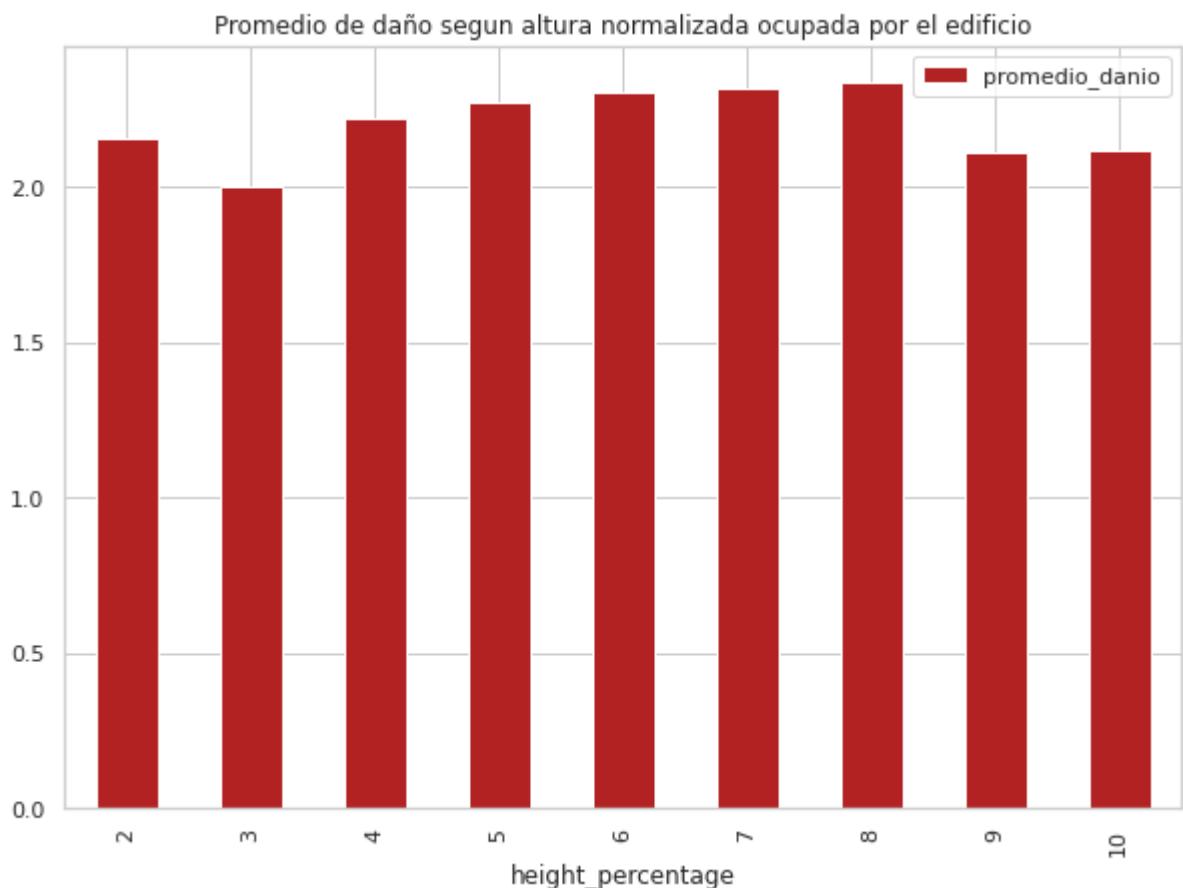


El daño promedio para las distintas cantidades de superficies normalizadas ocupadas por un edificio se encuentran en el grado medio de daño, con un leve aumento para las cantidades entre 3 a 10.

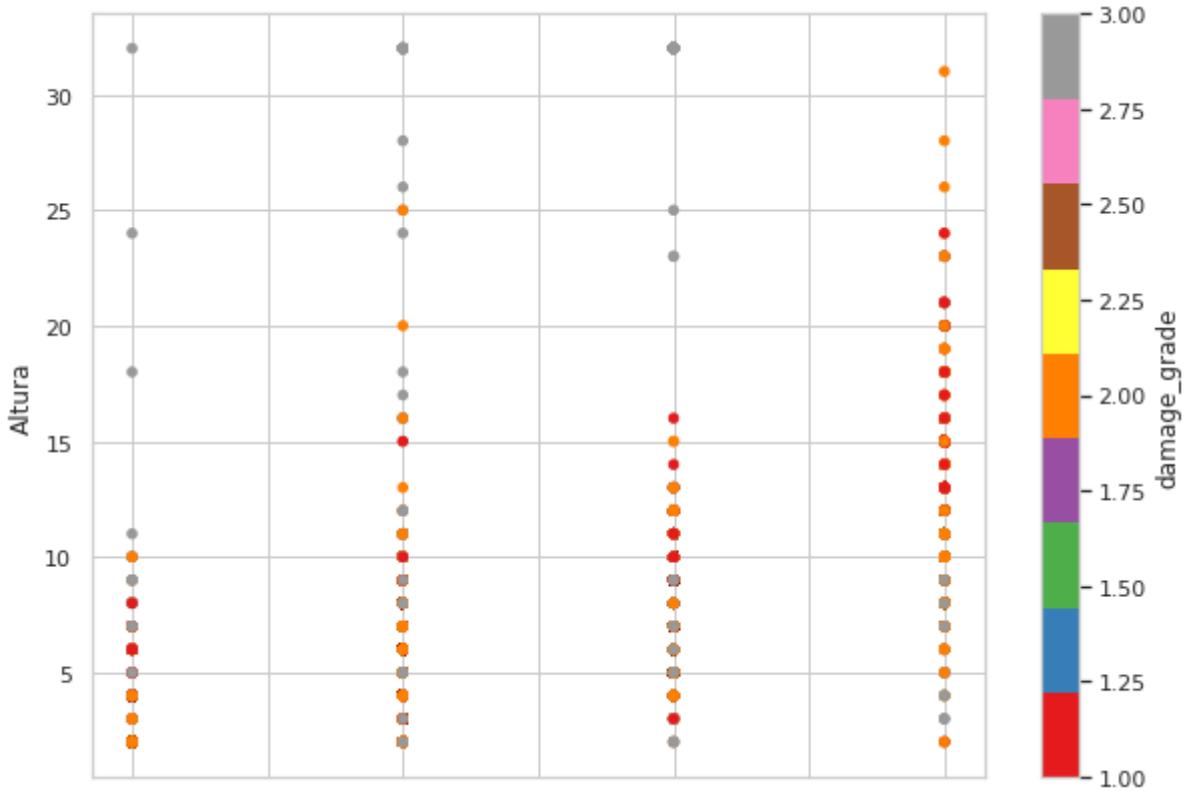
height_percentage: Análisis por altura normalizada ocupada por el edificio



Al igual que el caso de la superficie normalizada, la cantidad de muestra de edificios con altura normalizada mayor a 10 es escasa por lo que se decide descartar esos datos para la continuación del análisis.



El daño promedio para las distintas alturas normalizadas ocupadas por un edificio se encuentran entre grado medio y alto de daño, con un leve aumento para las alturas entre 4 a 8.

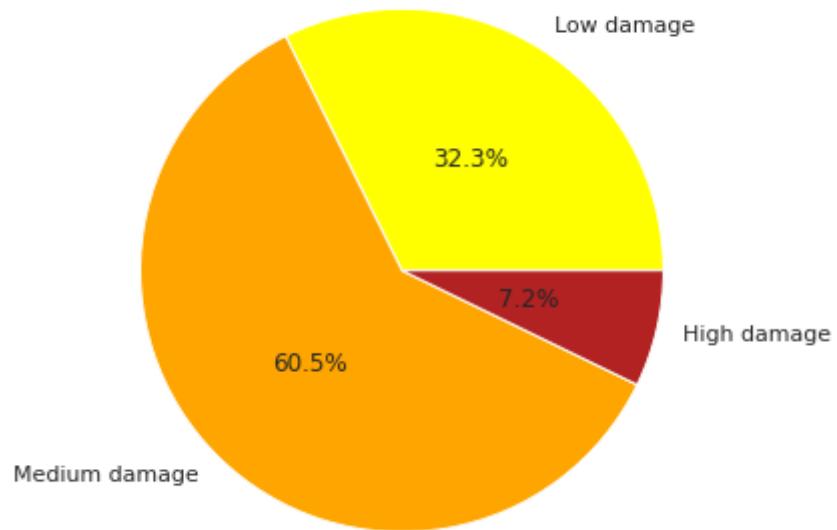


También se analizó la relación entre la cantidad de pisos, alturas normalizadas y el daño ocasionado por el fenómeno. De este último gráfico no se obtiene nada en concreto, solamente ayuda a reafirmar que la dispersión de daños no tiene una relación directa significativa con estas características de los edificios.

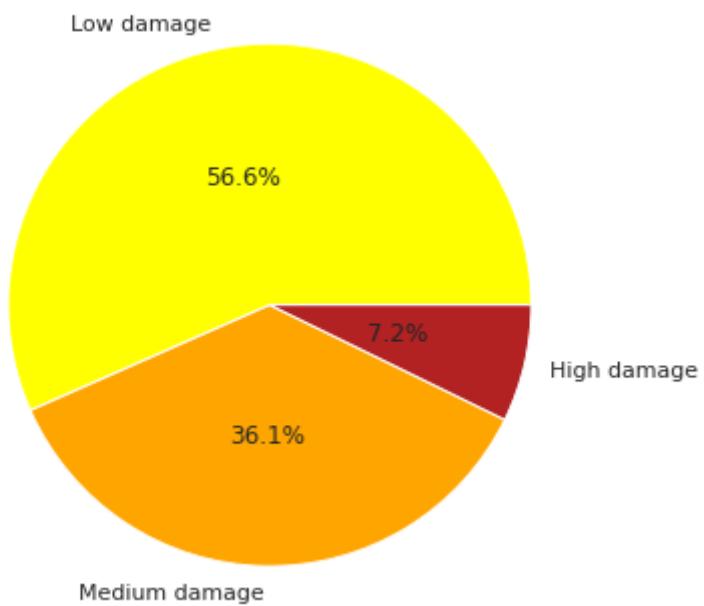
land_surface_condition: Análisis de daño para la condición de superficie n

Se analizó de forma porcentual los niveles de daños ocasionados para las diferentes condiciones de superficie en las que se construyeron los edificios obteniendo:

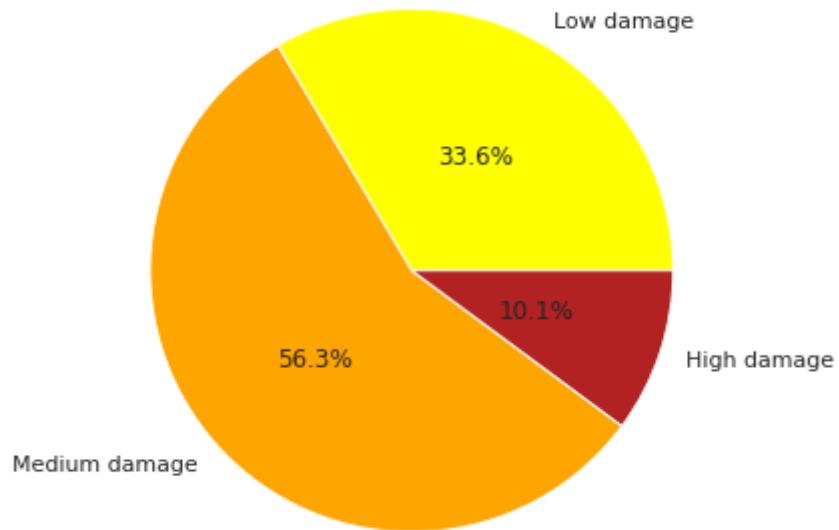
Porcentaje de daño con condicion de superficie n



Porcentaje de daño con condicion de superficie o



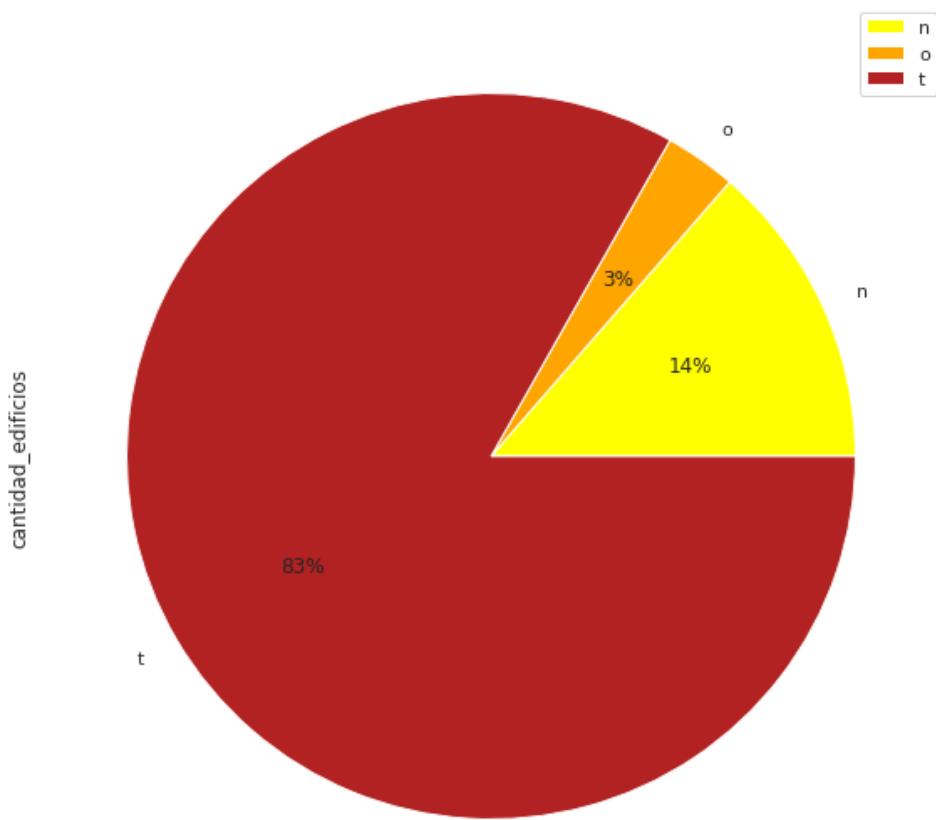
Porcentaje de daño con condición de superficie t



Por una leve diferencia, la condición de superficie de tipo "t" es la que posee mayor valor de daño grave. También se analizaron los daños de estas condiciones en relación de la antigüedad de los edificios obteniendo que no se tienen relación directa con el grado de daño, ya que se cuenta con datos de edificios de mismas edades y condiciones de superficie y con distintos tipos de daños.

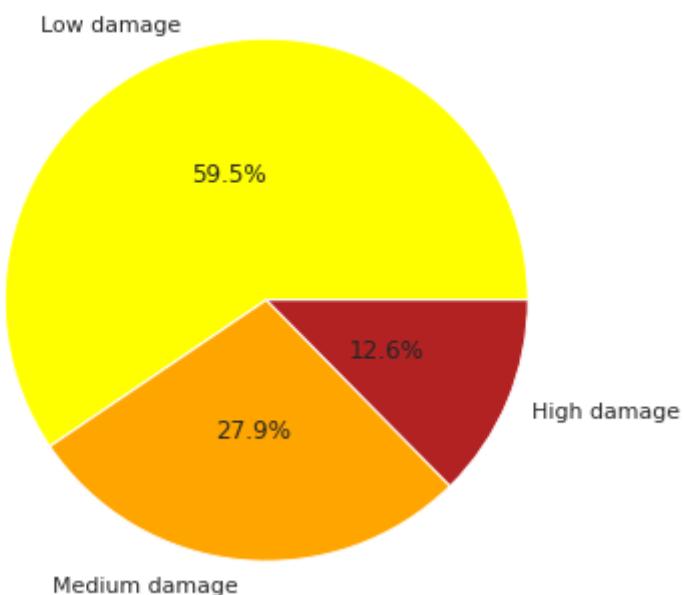
También cabe destacar que este tipo "t" de condición de superficie es el más común en las edificaciones:

Cantidad de muestras por condicion de la superficie terrestre del edificio



position: Análisis por orientación del edificio

Porcentaje de daño de edificios con orientacion j

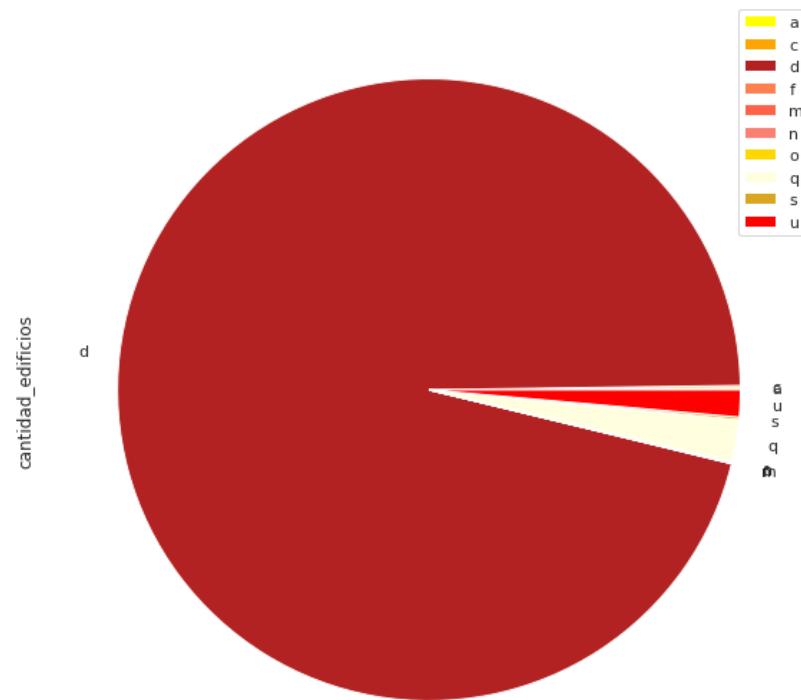


La orientación “j” de las edificaciones posee un valor porcentual mayor de daño alto pero un valor con poca diferencia al resto de las orientaciones. La orientación “o” es la que menor valor porcentual de daño alto presenta, pero se cuenta con poca cantidad de muestras de ese tipo por lo que no se puede garantizar que esa orientación haya tenido algún tipo de resguardo ante el fenómeno.

plan_configuration: Análisis por formato de construcción de la edificación (diseño sísmico)

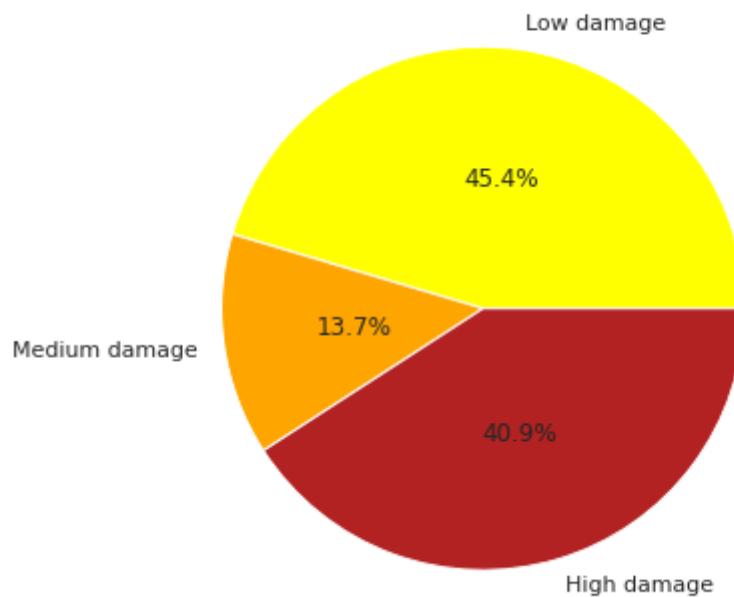
Estos datos son de sumo interés para el objetivo de este análisis, sin embargo la diferencia entre la cantidad de muestras entre los distintos tipos de formatos de construcción de diseño sísmico es importante.

Cantidad de muestras por formato de diseño sismico de construcción de la edificación



El tipo de diseño predominante en las edificaciones es el de tipo "d", y se encuentra entre los de menor valor porcentual de daño alto. El segundo tipo de diseño es el "q" el cual presenta el mayor valor porcentual (y muy significativo en comparación de los restantes).

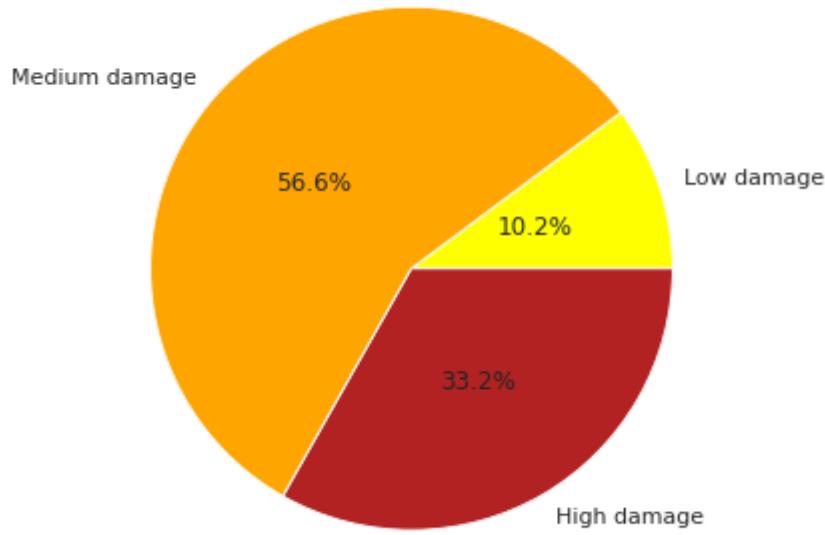
Porcentaje de daño de edificios con diseño sismico tipo q



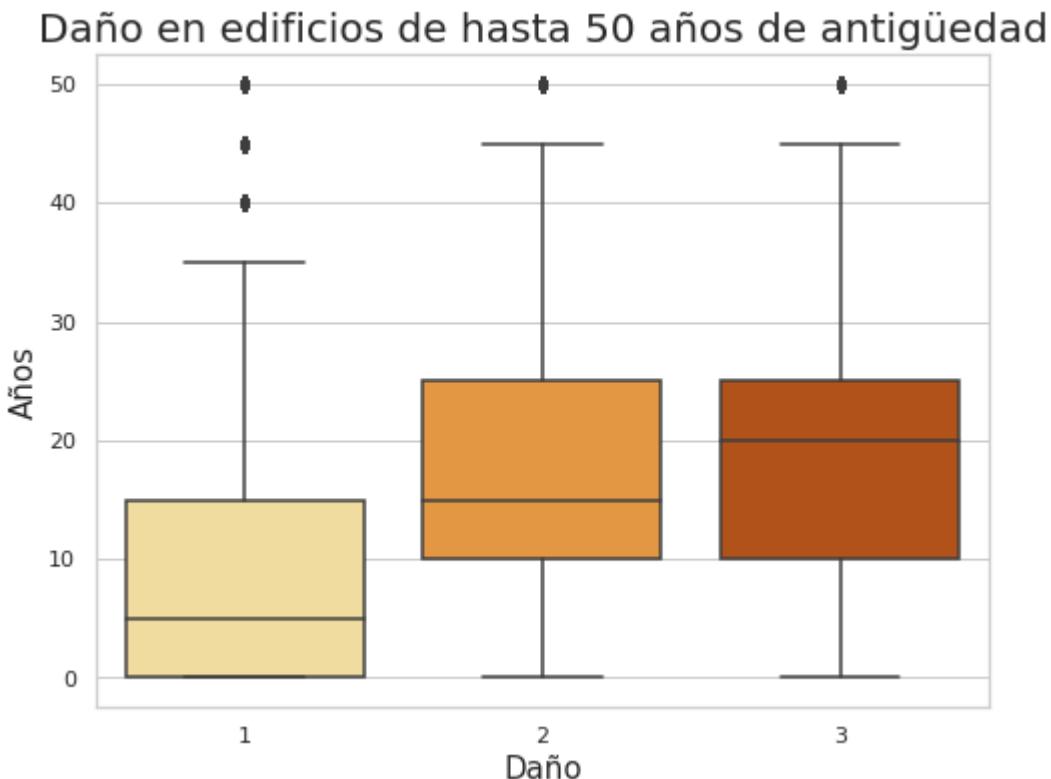
age: Análisis por antigüedad

En el caso de la columna age se decidió compararlo por rangos dado que habían muchas más edificaciones nuevas que viejas (sólo 242457 filas para las edificaciones de hasta 50 años). Para el primer rango tomamos los primeros 50 años:

Porcentaje de daño de buildings de hasta 50 años de antigüedad



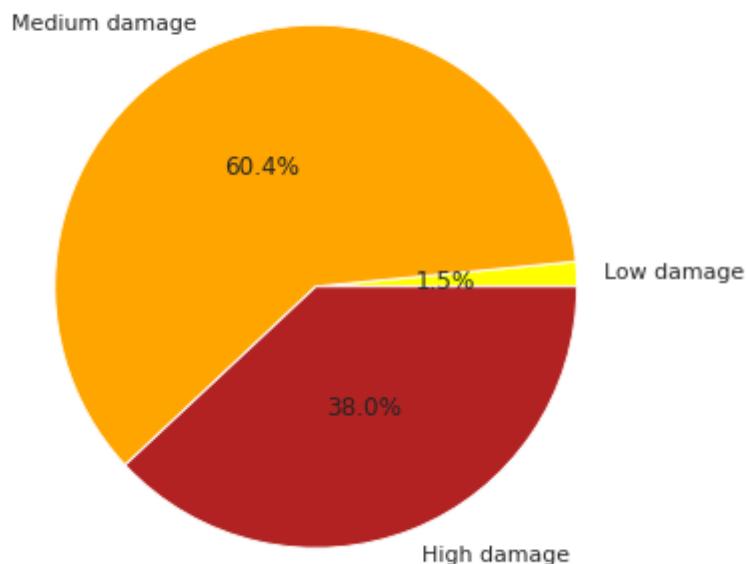
Pareciera ser que las edificaciones nuevas sufrieron de medio a mayor daño pero mirando el siguiente gráfico:



Las edificaciones de a partir de 15 años fueron las más gravemente afectadas.

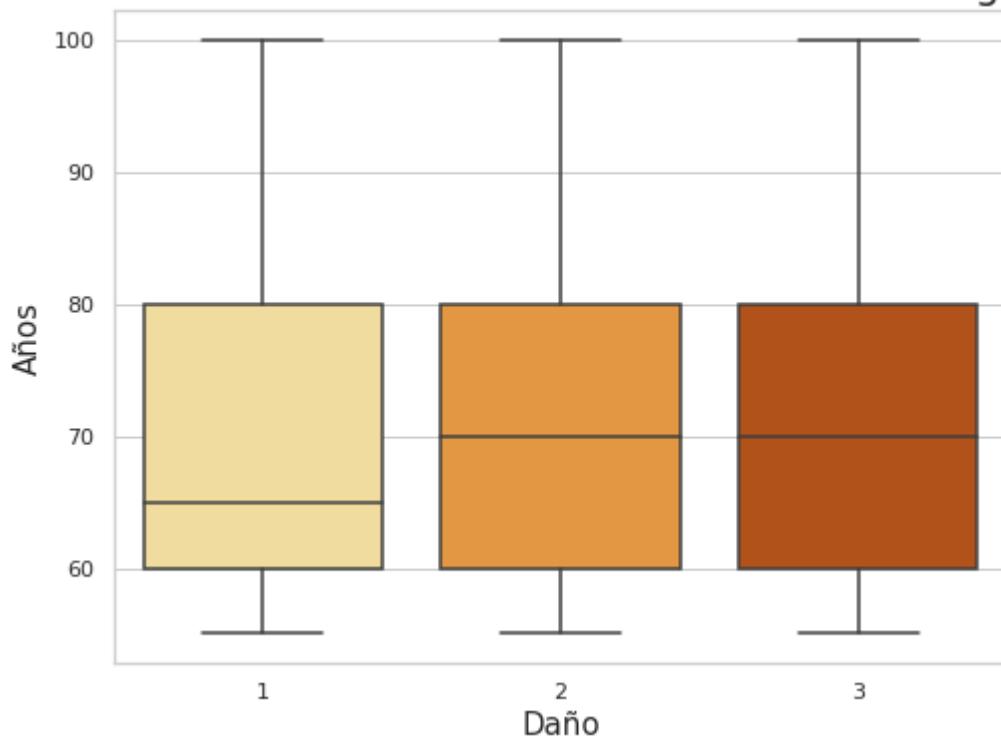
Para el siguiente rango de valores tomamos las edificaciones mayores de 50 años hasta 100 años. En este caso sólo tenemos 16020 filas para comparar el cual es un número bastante bajo comparado al rango anterior y al total de filas.

Porcentaje de daño de buildings de entre 50 a 100 años de antigüedad



Ahora las edificaciones de este rango de años parecen ser menos afortunadas.

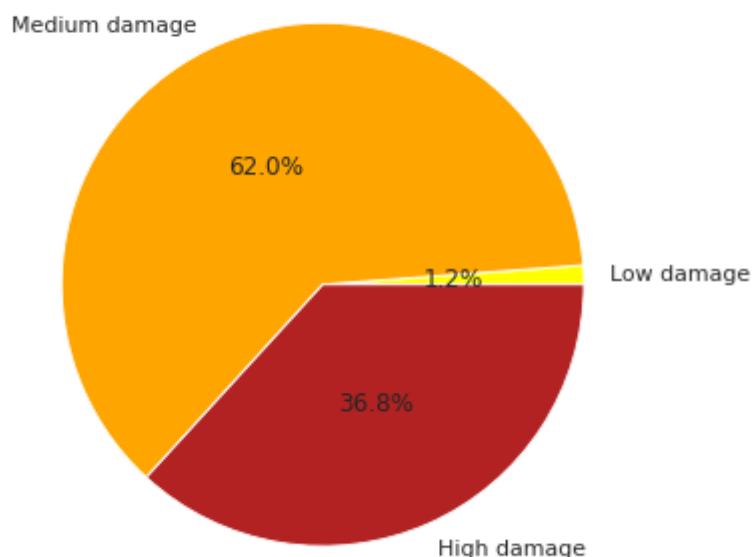
Daño en edificios de entre 50 a 100 años de antigüedad



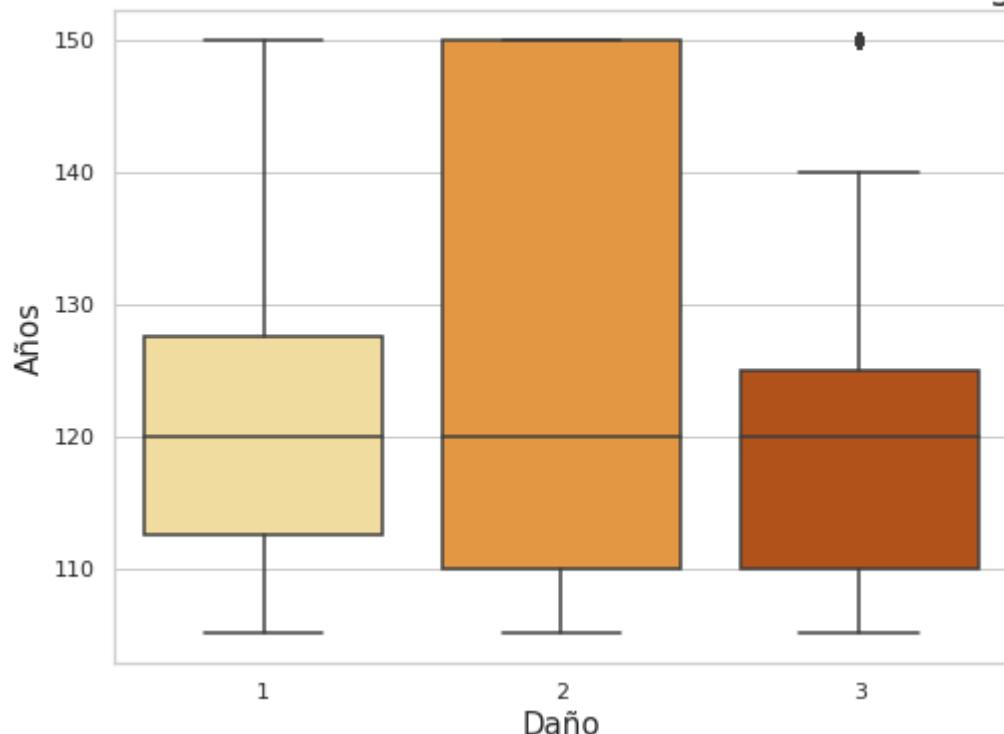
Sin embargo, la media es de 68 años entre los 3, más parejo que el rango anterior por lo que no parece ser un factor determinante en las edificaciones de más de 50 años.

Los siguientes rangos son menos datos y los gráficos no dicen mucho:

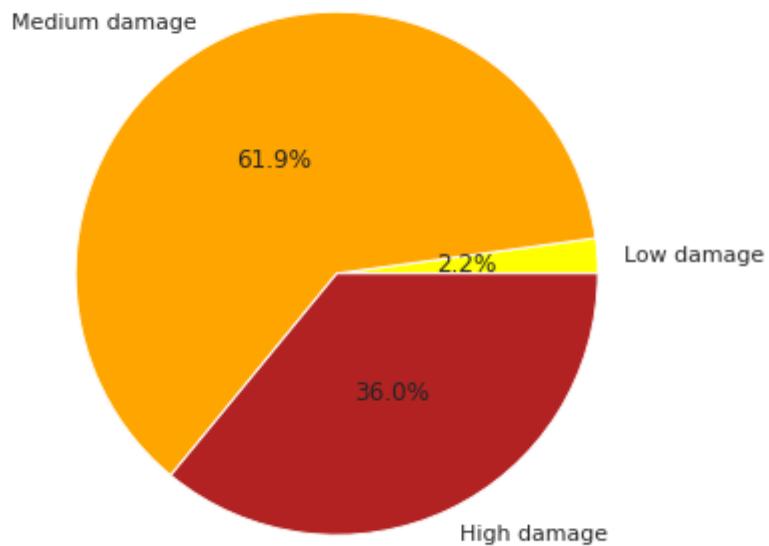
Porcentaje de daño de buildings de entre 100 a 150 años de antigüedad



Daño en edificios de entre 100 a 150 años de antigüedad

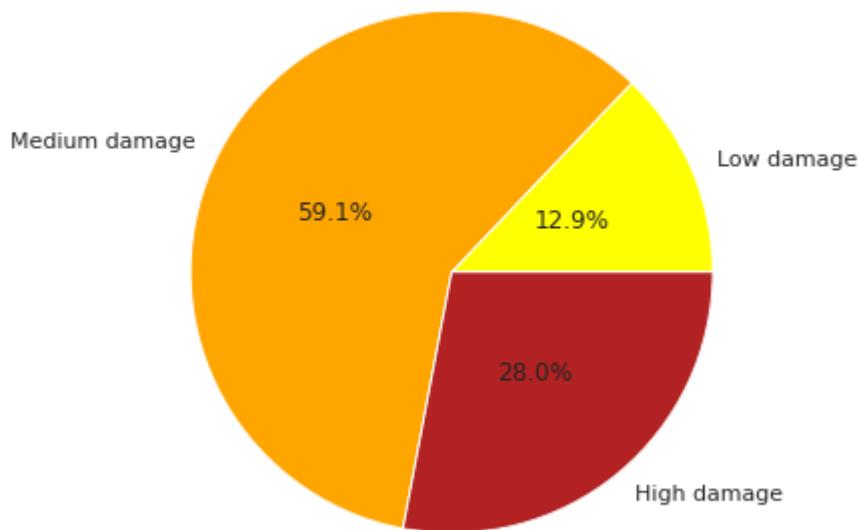


Porcentaje de daño de buildings de entre 150 a 200 años de antigüedad



Ahora hay un caso particular, hay edificaciones de 995 años, creemos que son edificaciones muy viejas cuya edad exacta se desconoce. Como son unas 1390 (más que el rango de 100 a 200 años) decidimos investigarlas para ver si habían sido igualmente afectadas:

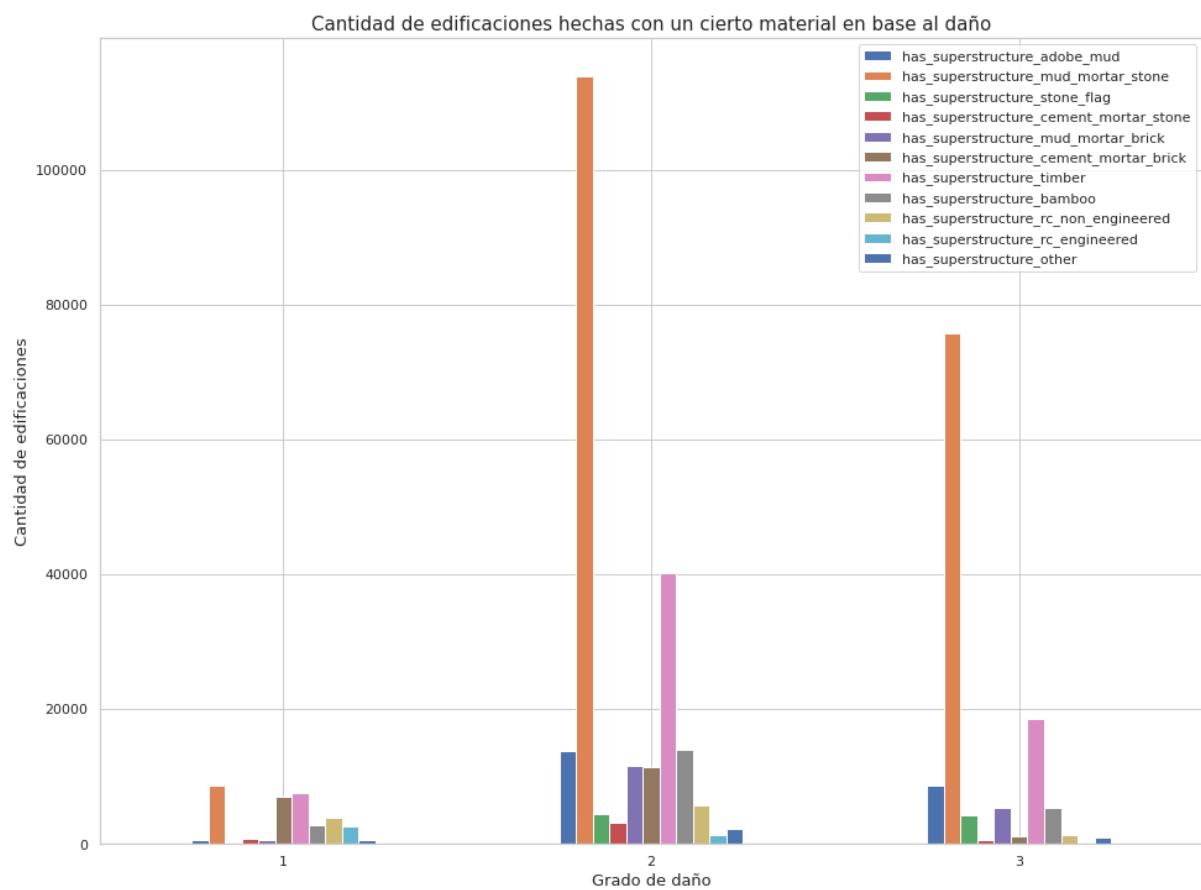
Porcentaje de daño de buildings de 995 años de antigüedad



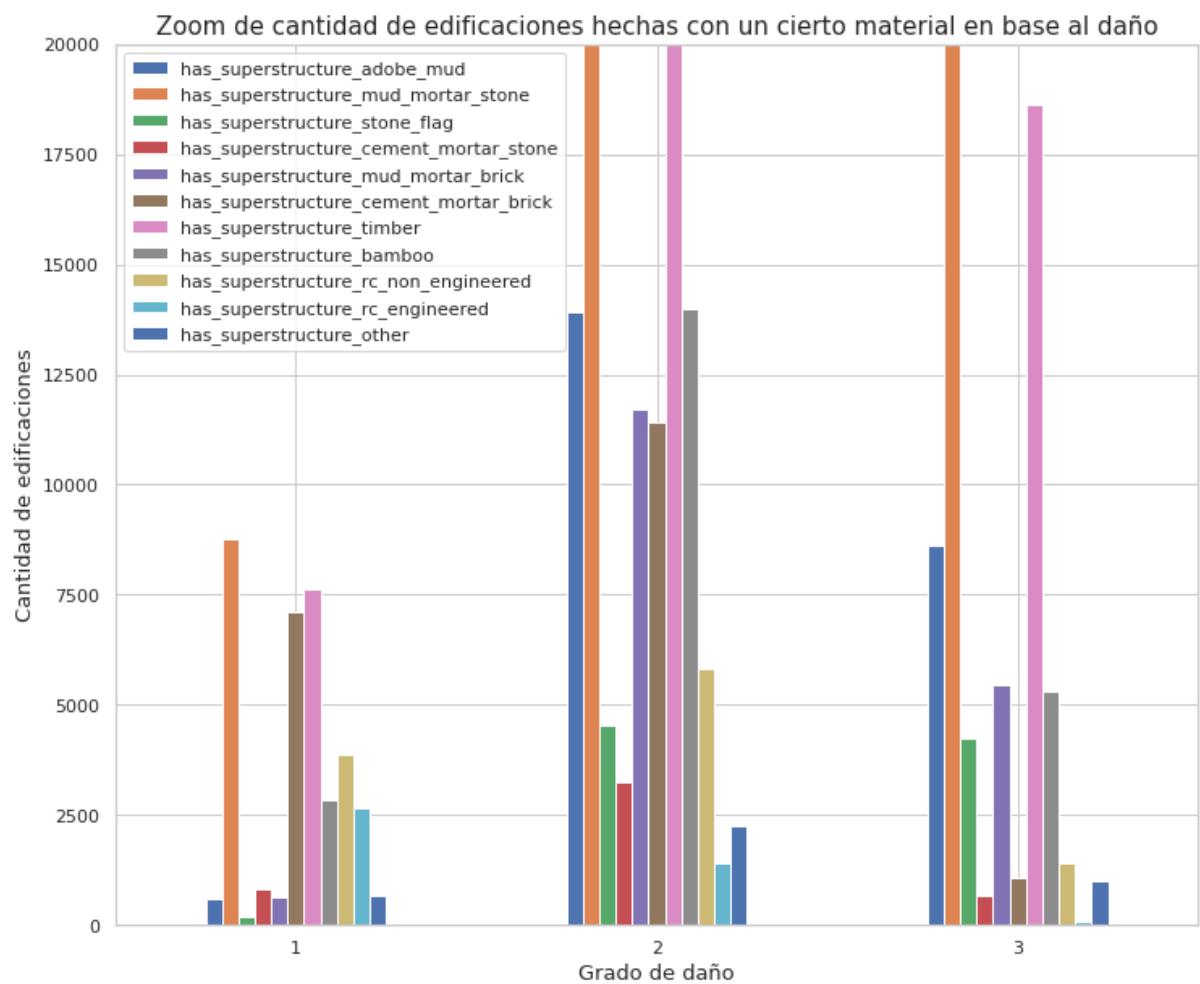
No parece ser que los años de las edificaciones sea un factor determinante en cuánto daño recibieron.

has_superstructure: Análisis por superestructuras

Para analizar el material con el que estaban hechas las edificaciones primero las agrupamos en base al daño que recibieron:

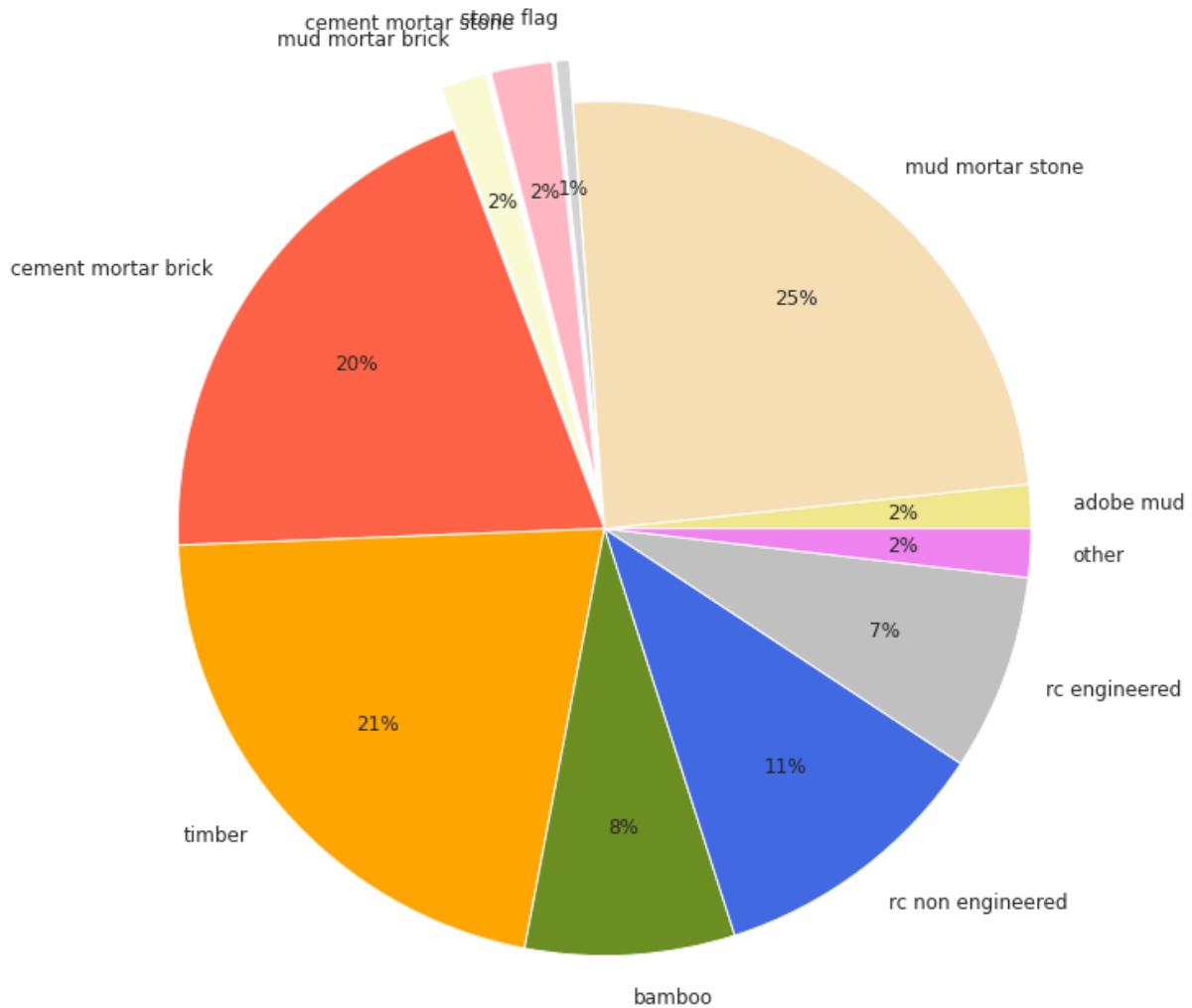


Dado que hay una clara mayoría del uso de mud mortar stone decidimos hacer un zoom para diferenciar aquellos materiales menos usados:



Tal como habíamos visto antes mud mortar stone es de los materiales más usados junto con timber. Cabe recalcar que algunas edificaciones podrían usar más de un tipo de material por lo que si bien es cierto que la mayoría de edificaciones usa mud mortar stone también podrían usar otros materiales. Dicho esto, es una buena visualización para entender de qué están hechas las edificaciones en esta zona de Nepal, por lo que proseguimos a analizar cómo se comportan en porcentaje dependiendo del daño recibido:

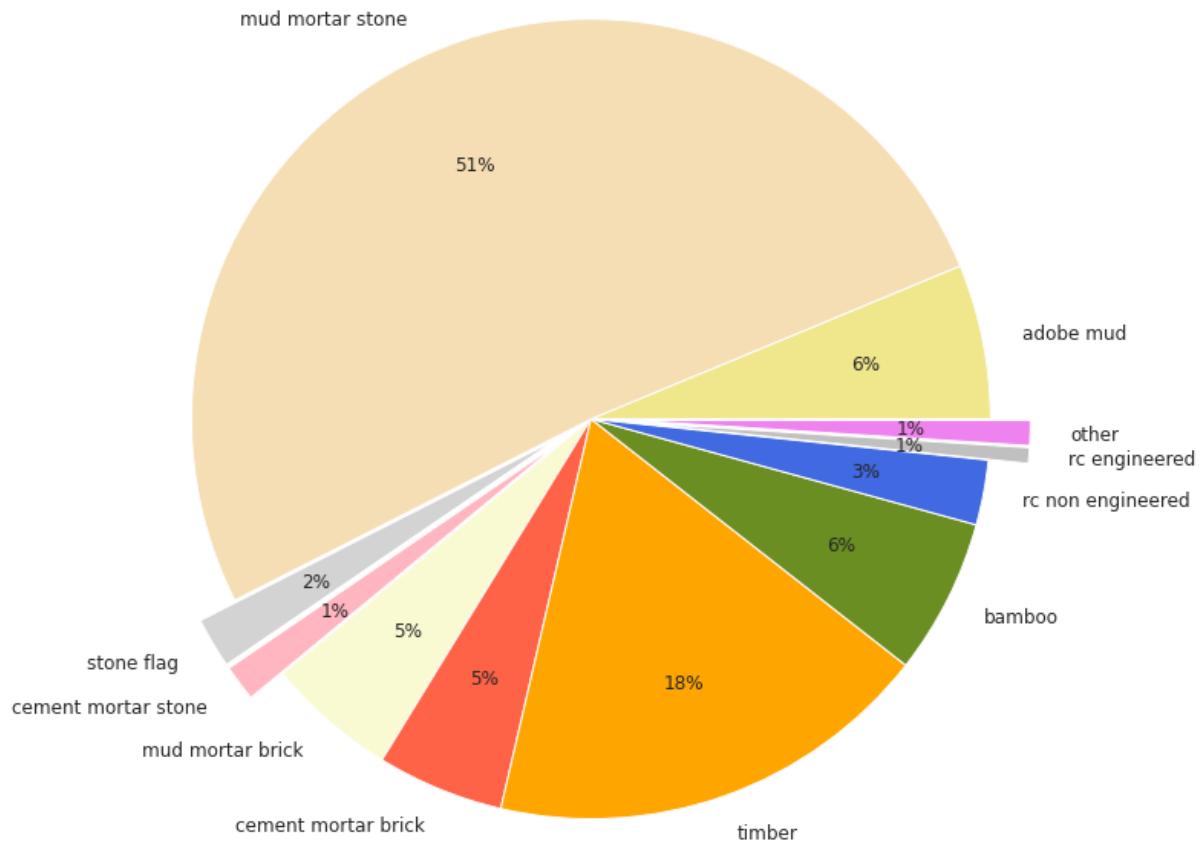
Porcentaje del tipo de edificación con daño mínimo



Este gráfico junto con el anterior nos da una relación de la proporción de los materiales utilizados en las edificaciones que sufrieron daño mínimo. En este caso hay una distribución pareja entre mud mortar stone, timber y cement mortar brick.

Para los que recibieron daño medio:

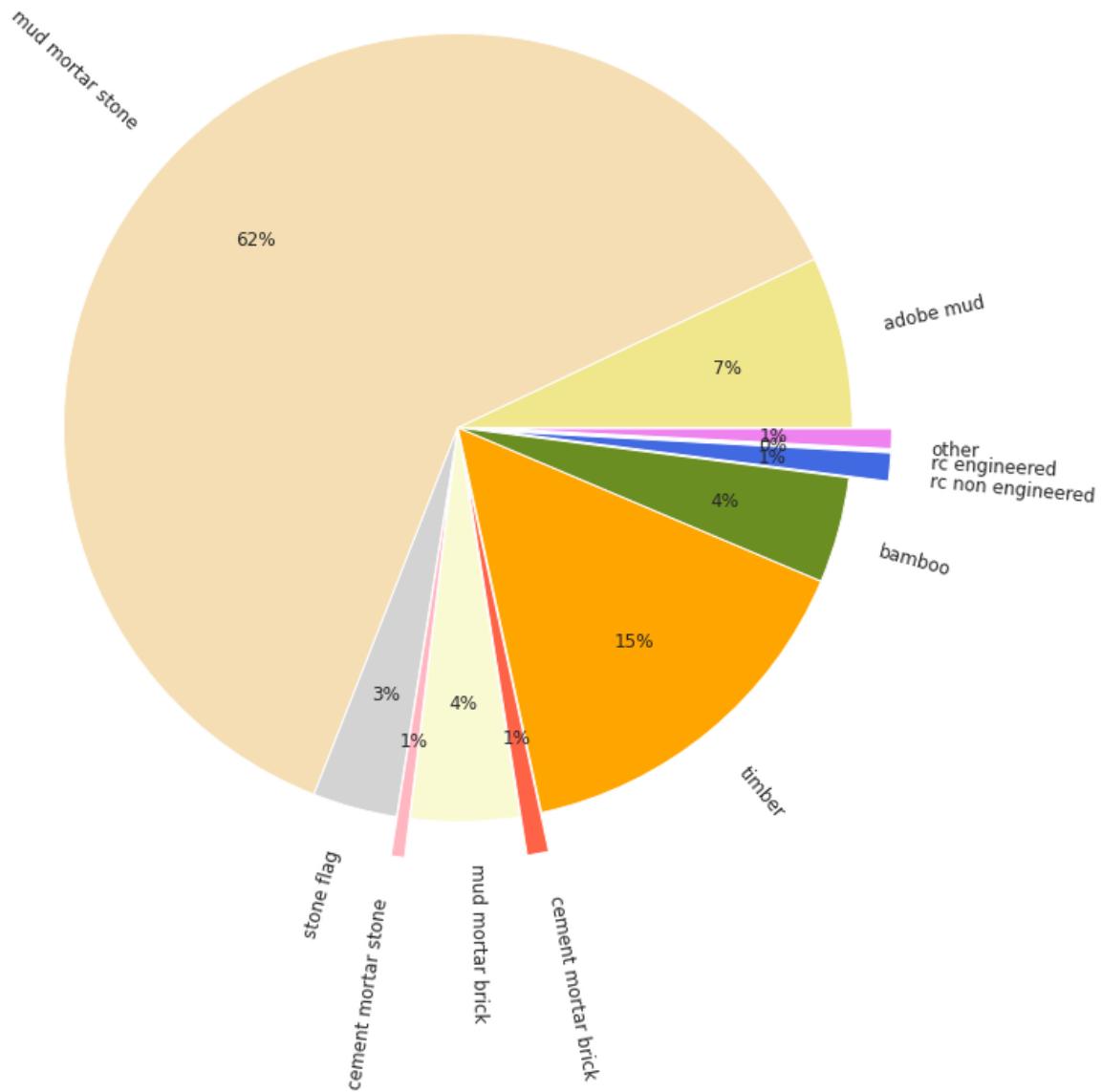
Porcentaje del tipo de edificación con daño medio



Acá ya se empieza a ver cómo mud mortar stone es de los materiales más utilizados para las edificaciones que recibieron daño medio.

En cuanto a los que recibieron daño mayor:

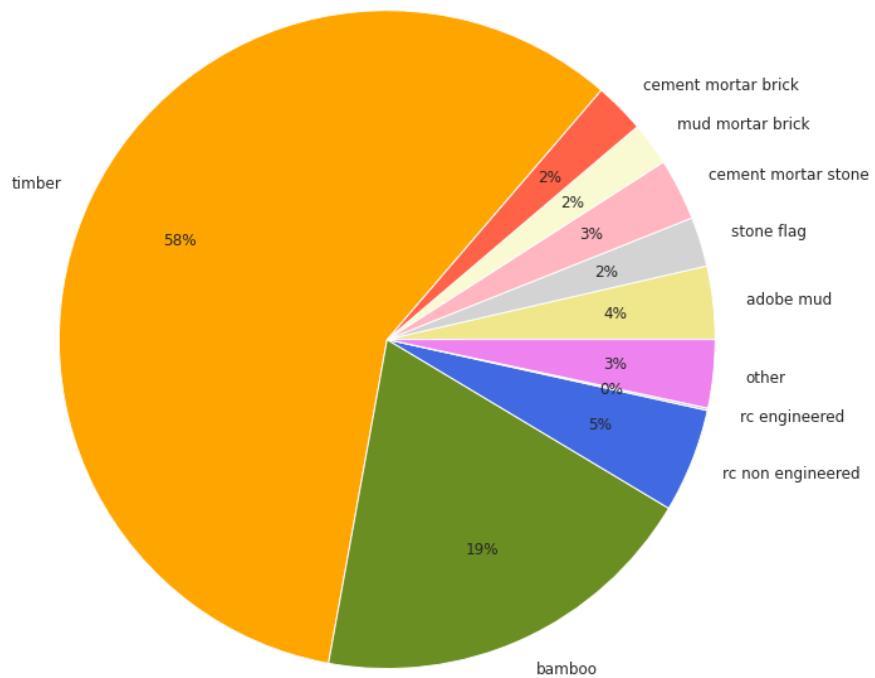
Porcentaje del tipo de edificación con daño mayor



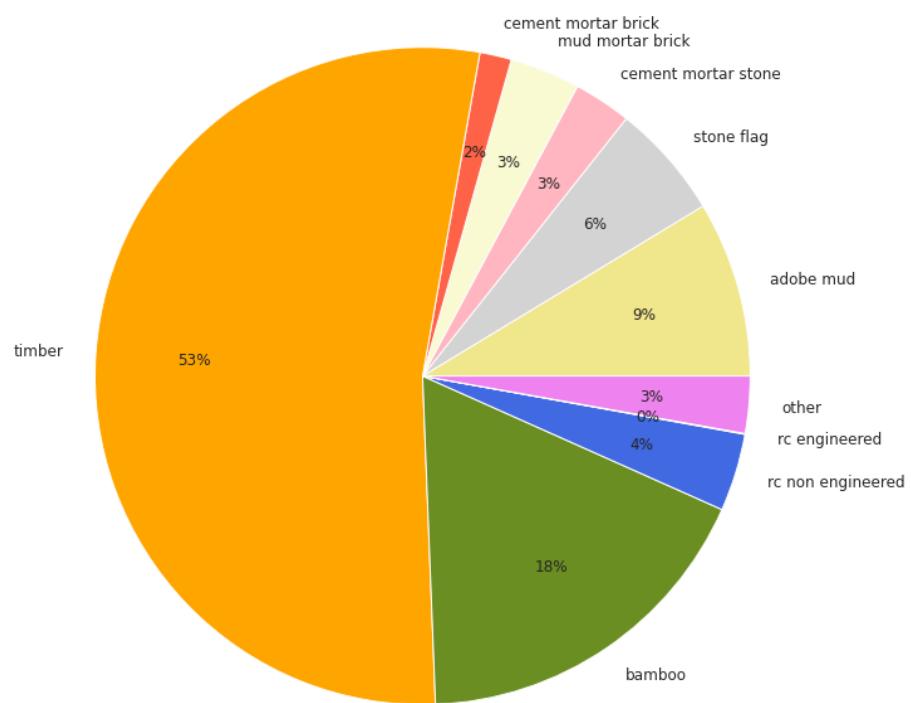
Se comporta muy igual al gráfico anterior.

Como vimos en los primeros gráficos, la mayoría de las edificaciones están hechas con mud mortar stone por lo que debe ser un material muy común en Nepal. Por ello mismo (y dado que se usan materiales en conjunto) filtramos las edificaciones que contenían este material y los que no. En el primer caso nos quedaron 198561 edificaciones (un 76,2% del total) y nos quedaron estas proporciones en base al daño que recibieron:

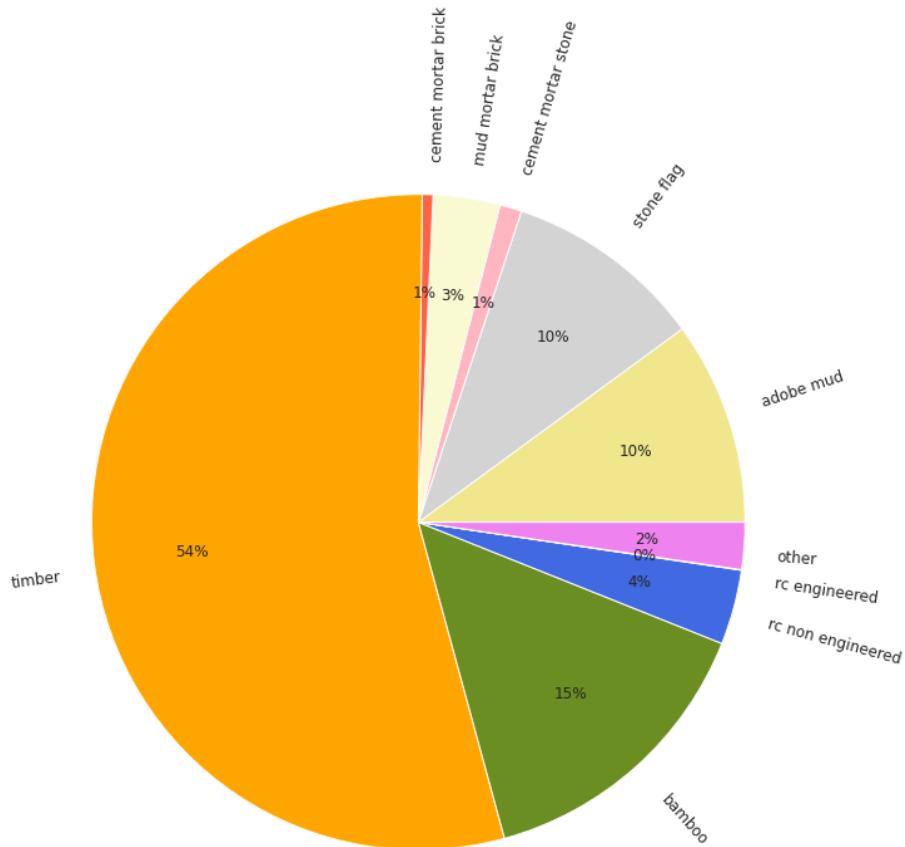
Porcentaje del tipo de edificación en base a mud mortar stone con daño menor



Porcentaje del tipo de edificación en base a mud mortar stone con daño medio

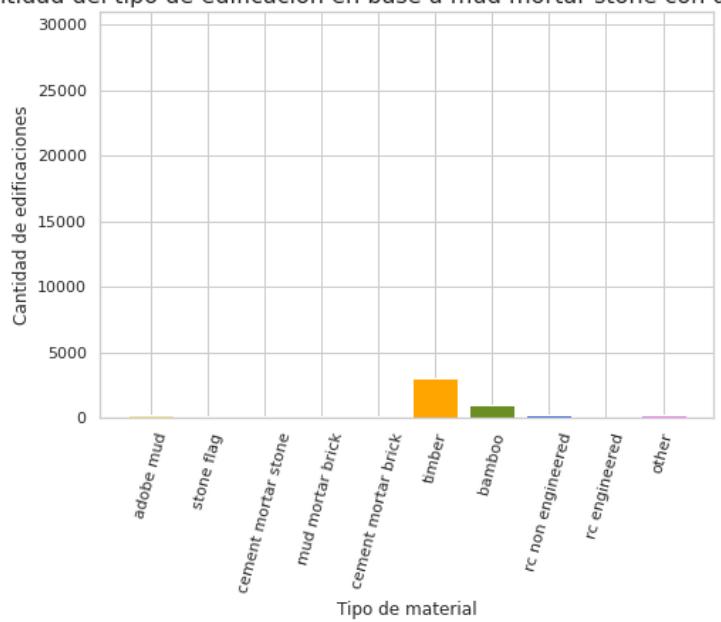


Porcentaje del tipo de edificación en base a mud mortar stone con daño mayor

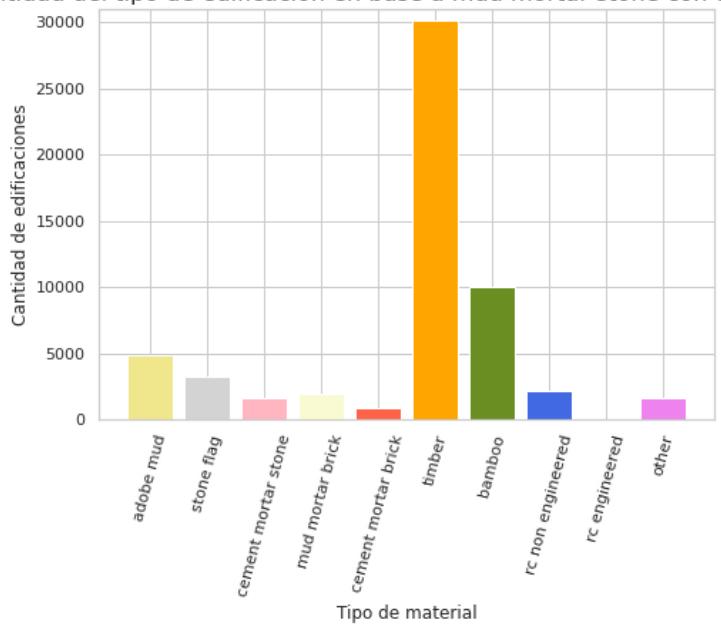


Como se puede observar los 3 gráficos muestran proporciones parecidas. Timber y bamboo son los materiales más usados con mud mortar stone. Ahora si lo vemos en cantidades:

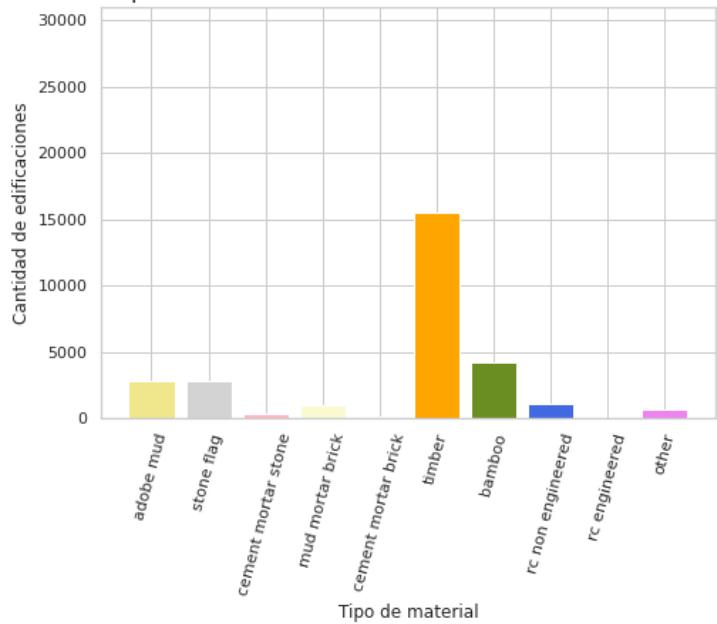
Cantidad del tipo de edificación en base a mud mortar stone con daño menor



Cantidad del tipo de edificación en base a mud mortar stone con daño medio



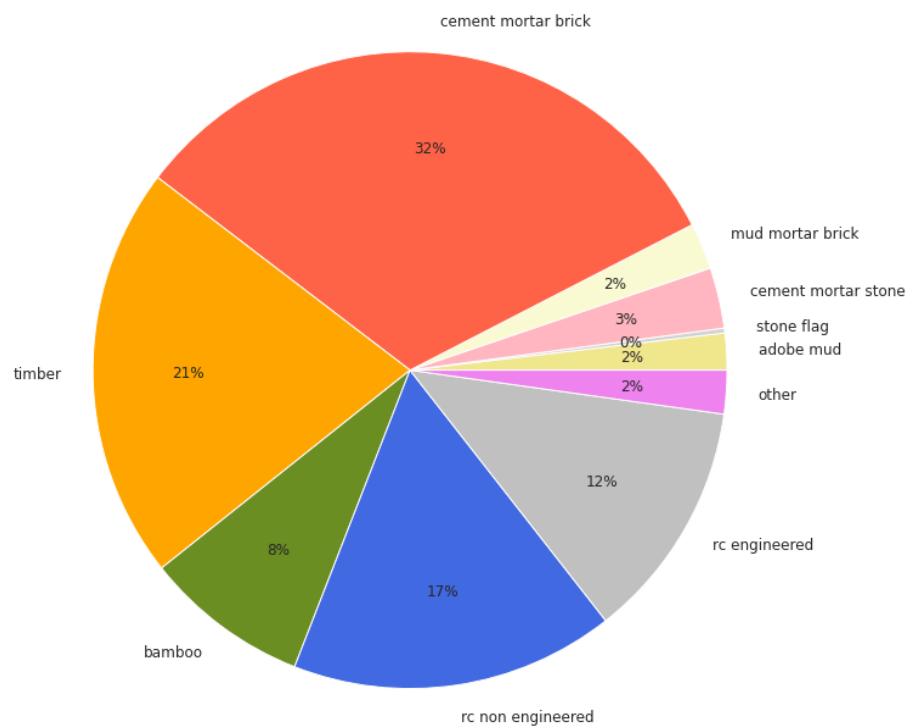
Cantidad del tipo de edificación en base a mud mortar stone con daño mayor



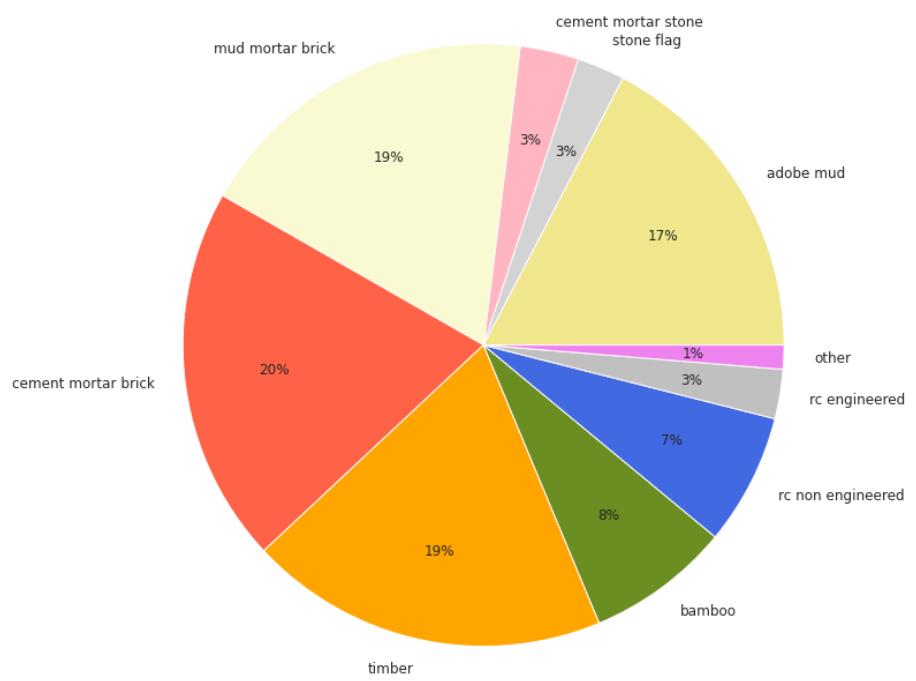
Si bien las cantidades son distintas las proporciones se mantienen como habíamos visto antes.

Ahora pasando a las edificaciones que no usan mud mortar stone separados por daño recibido, tienen una proporción más diferente comparado a los gráficos anteriores:

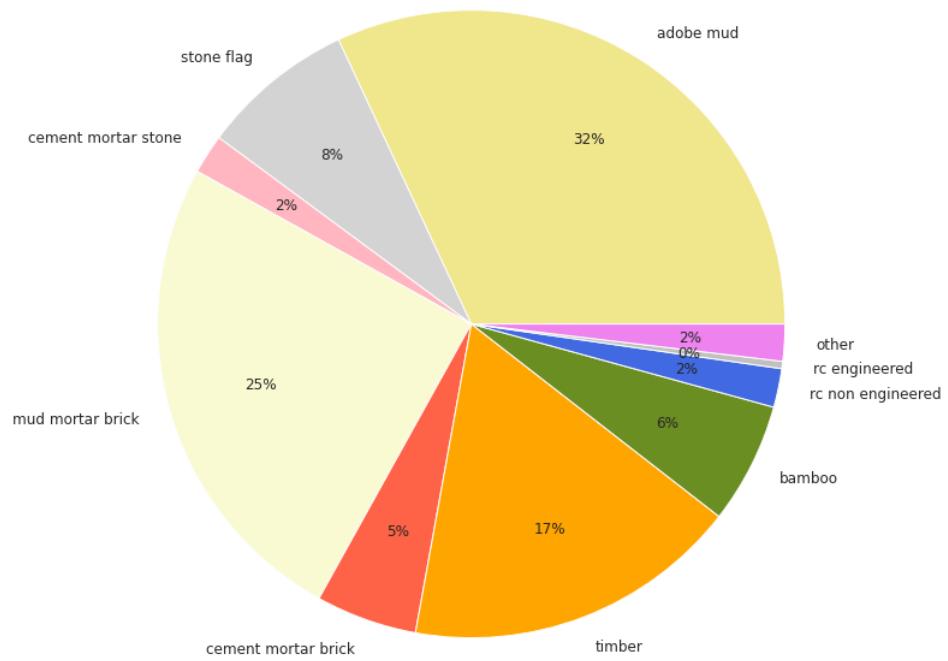
Porcentaje del tipo de edificación sin mud mortar stone con daño menor



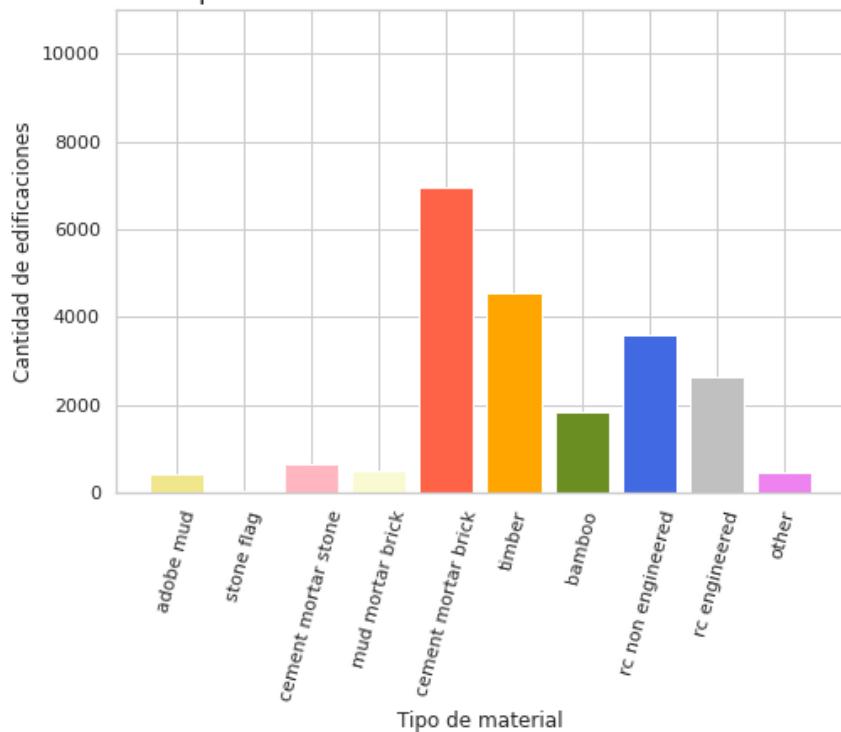
Porcentaje del tipo de edificación sin mud mortar stone con daño medio



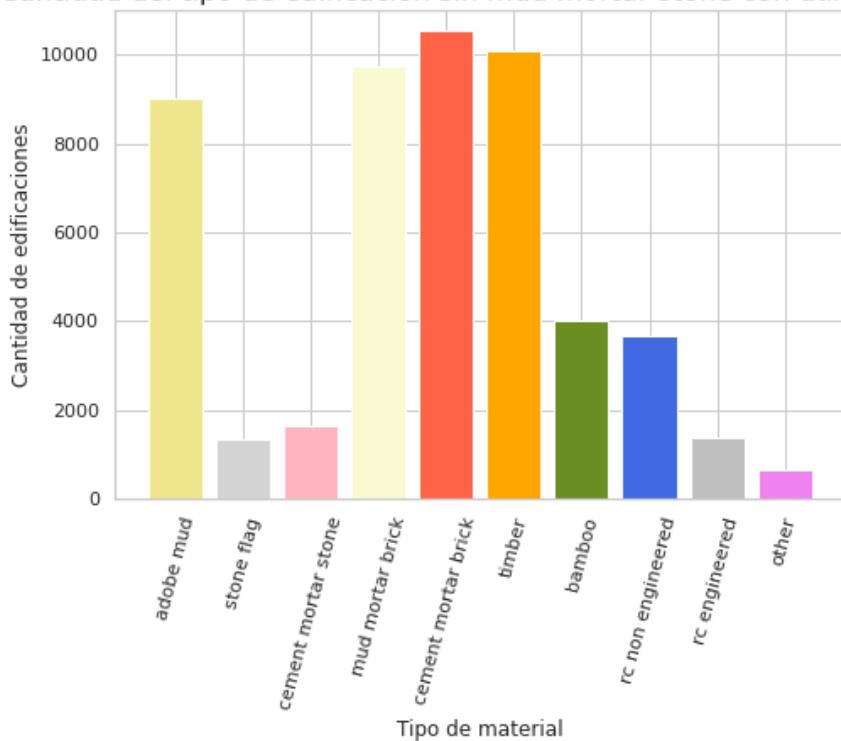
Porcentaje del tipo de edificación sin mud mortar stone con daño mayor



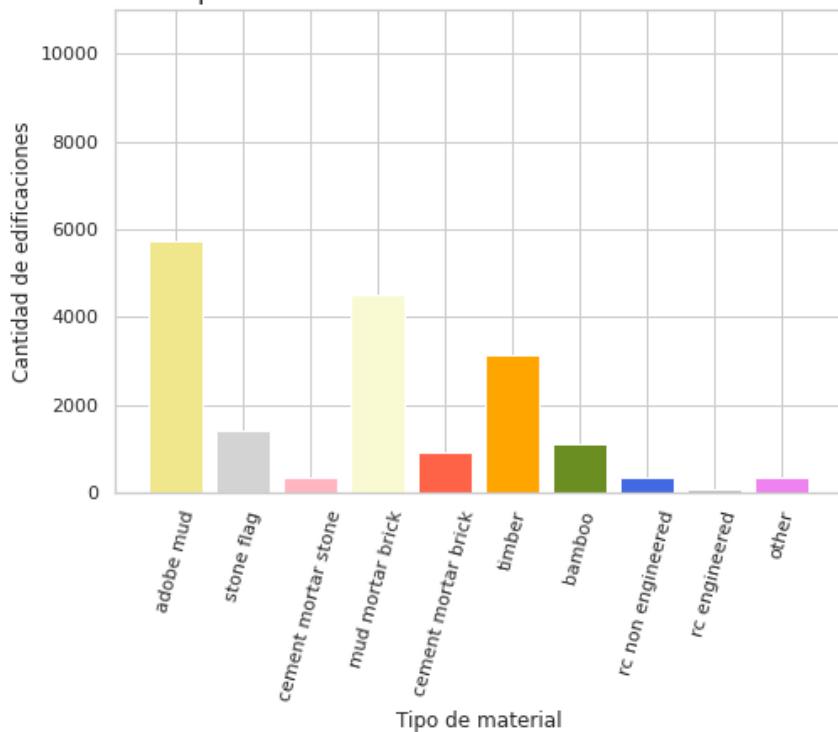
Cantidad del tipo de edificación sin mud mortar stone con daño menor



Cantidad del tipo de edificación sin mud mortar stone con daño medio



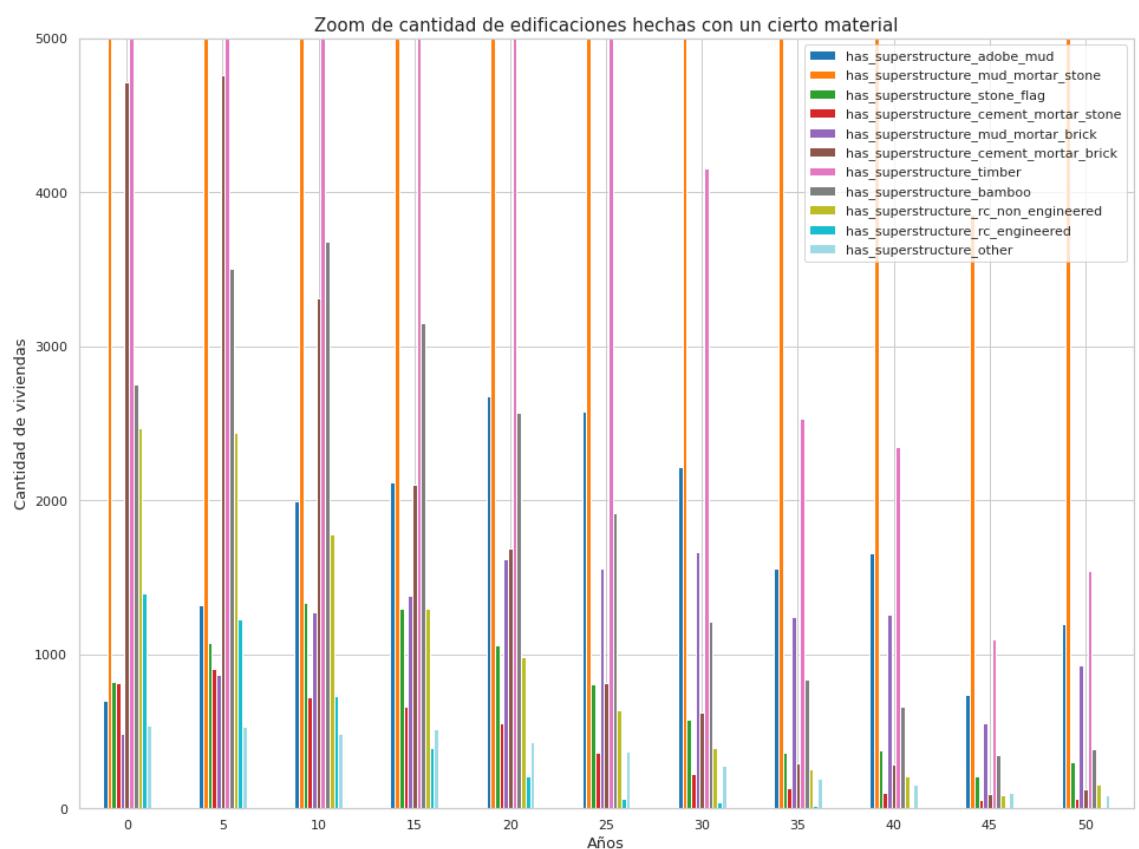
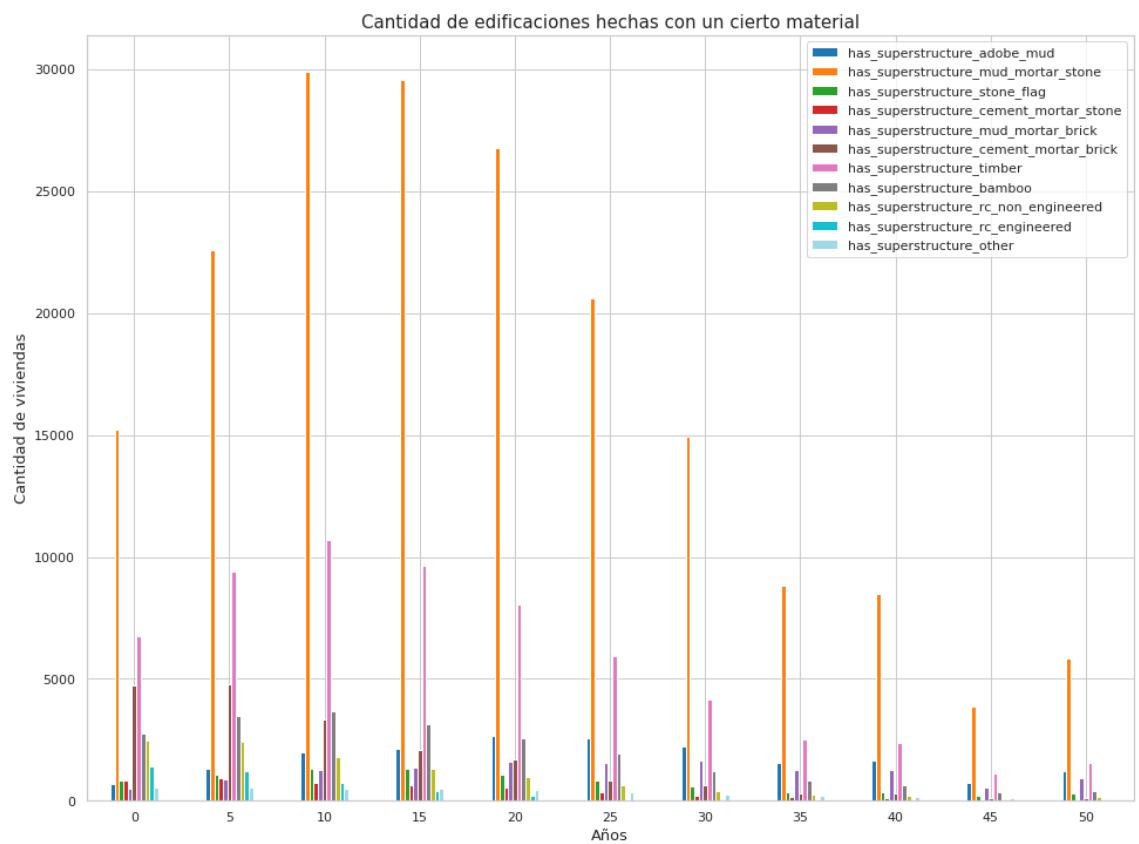
Cantidad del tipo de edificación sin mud mortar stone con daño mayor

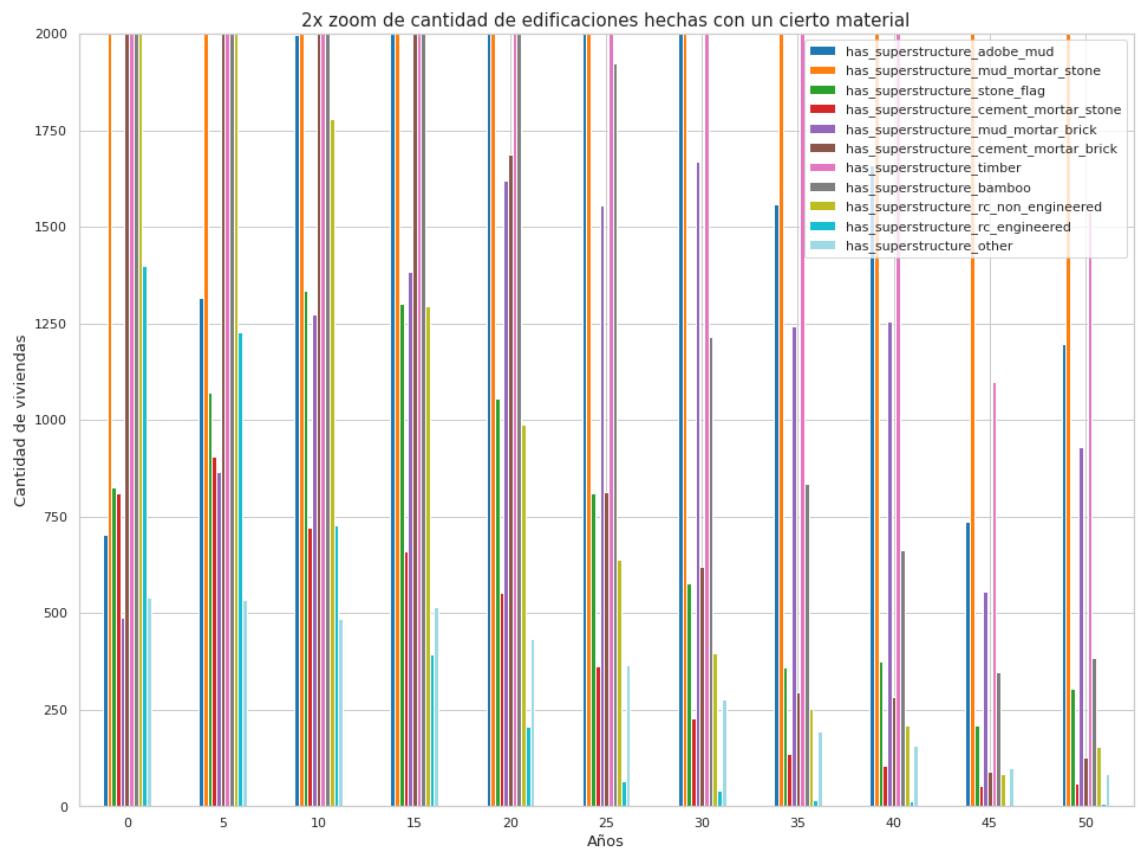


A falta de mud mortar stone, las edificaciones fueron hechas principalmente con cement mortar brick y fueron las que más resistieron (recibieron daño menor y medio). Por el contrario, las que estaban hechas con adobe mud fueron las que más daño mayor sufrieron, seguido de mud mortar brick y timber.

age vs has_superstructure

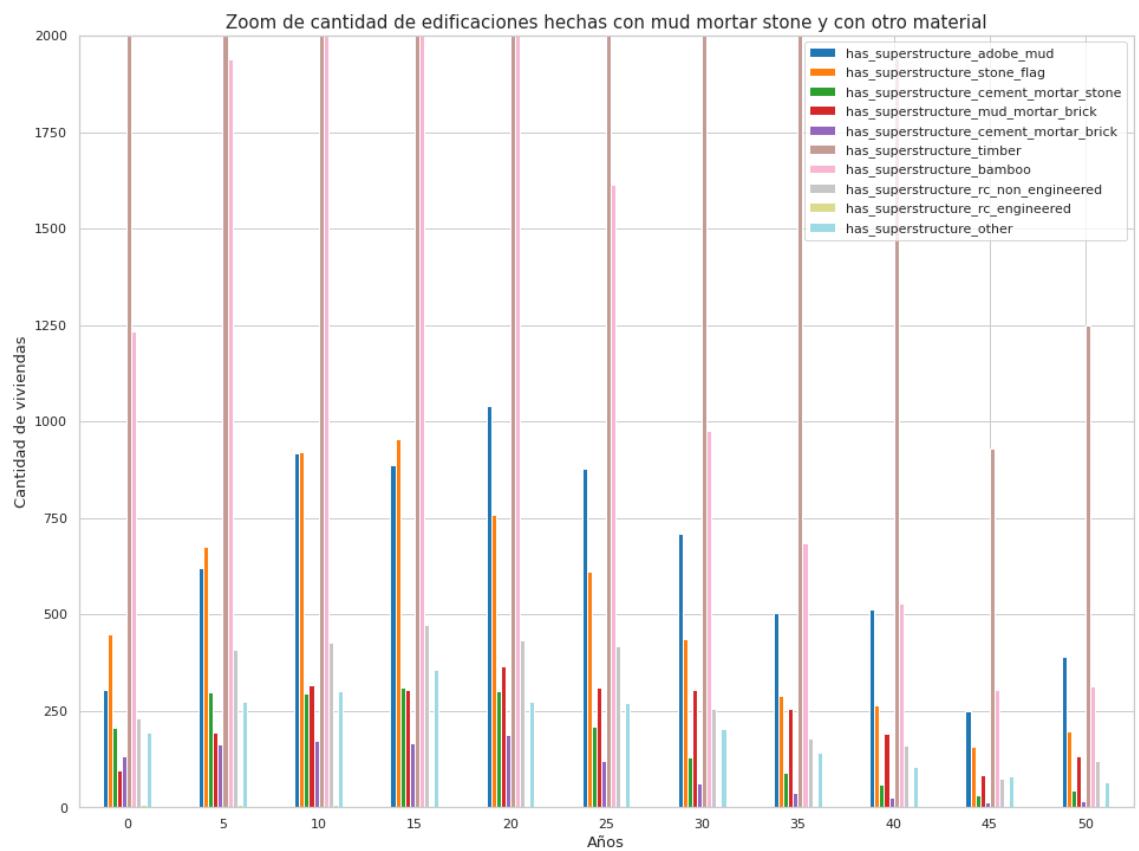
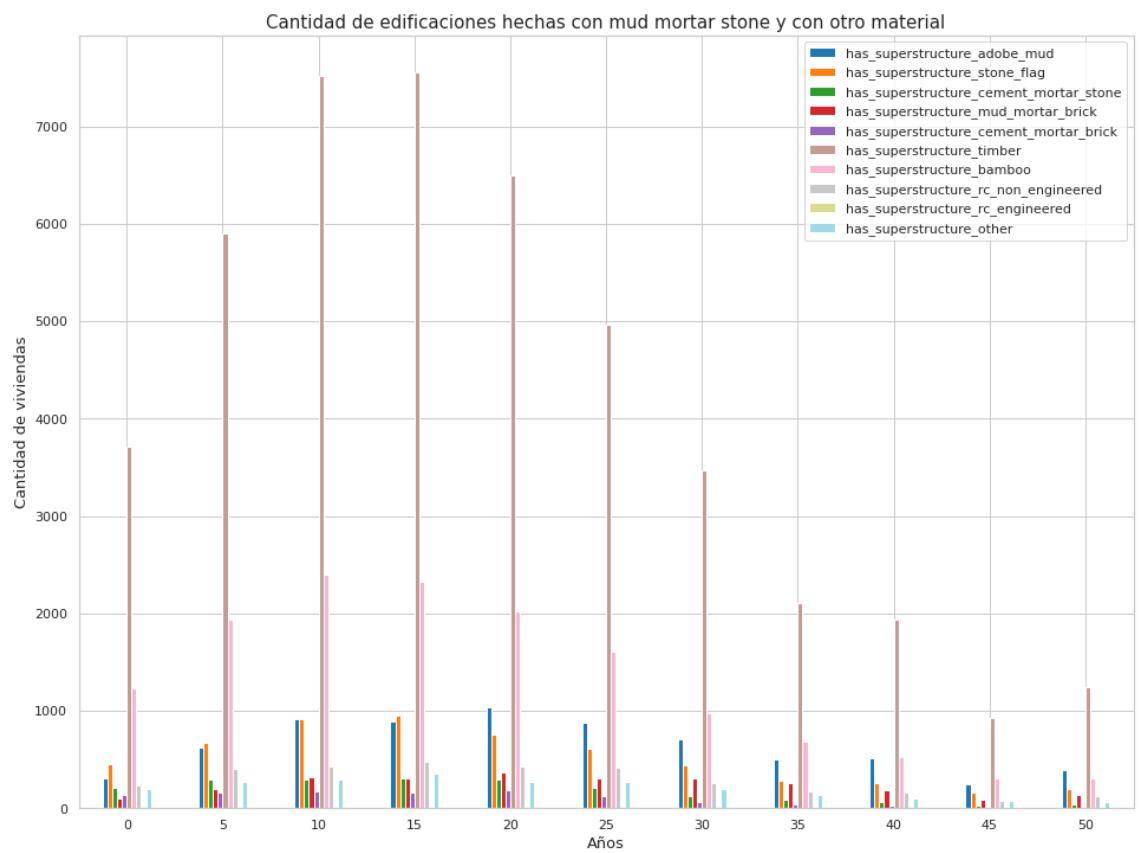
Dado los resultados obtenidos en age y has_superstructure nos preguntamos si había alguna relación entre ellos que nos llevara más cerca de encontrar las características particulares de las edificaciones dañadas. Para ello, nos enfocamos en tomar el grupo de edificaciones de 0 a 50 años (que era el más grande y variado) para ver qué características poseen respecto a los materiales con los que están hechos:





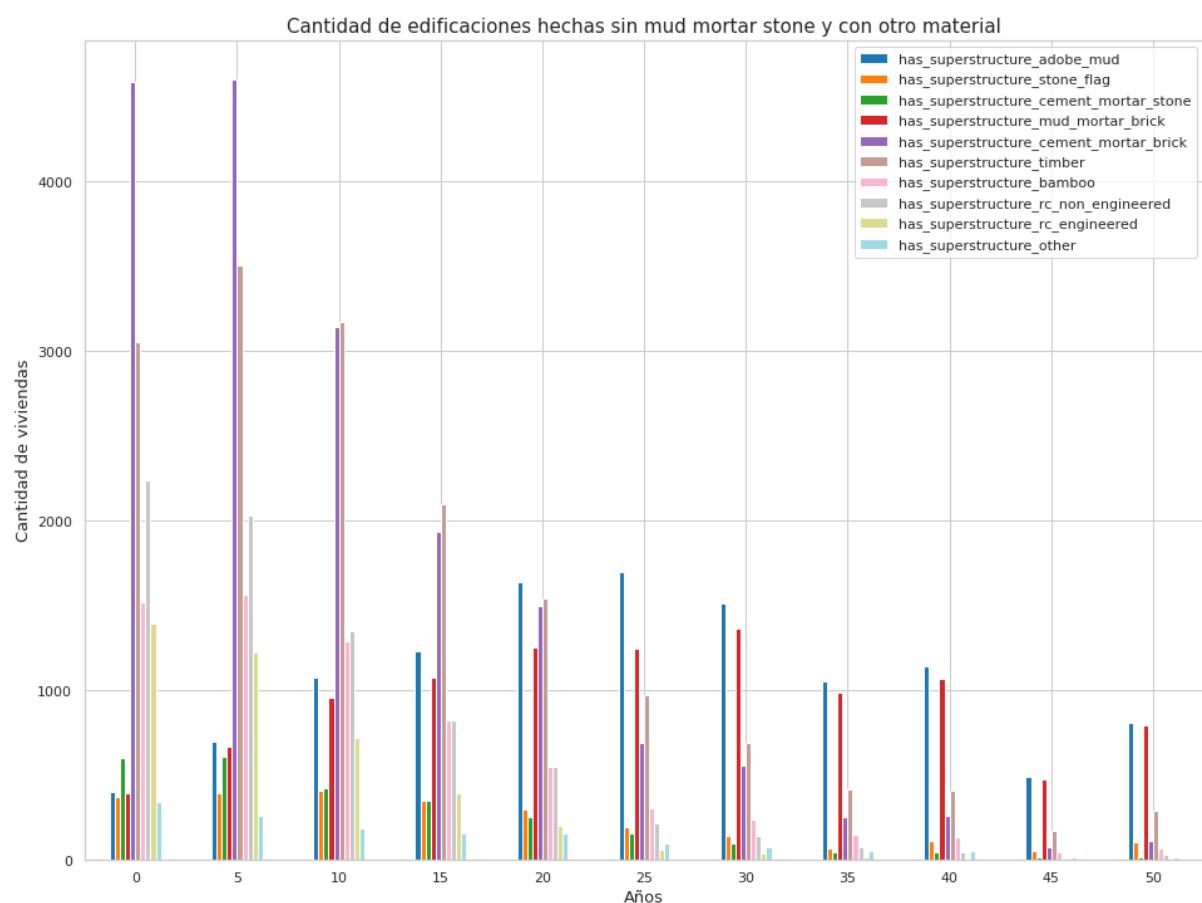
Esto solo nos vuelve a confirmar el uso de mud mortar stone en las edificaciones más nuevas y cómo se distribuyen entre los distintos períodos de antigüedad de las mismas.

Como habíamos visto en la sección de has_superstructure, mud mortar stone es de los materiales más usados así que graficamos nuevamente con qué materiales es usado en edificaciones más nuevas:



Timber seguido de bamboo como vimos antes son los más usados, seguidos en menor cantidad por adobe mud y stone flag.

En cuanto a los que no utilizan mud mortar stone:



Cement mortar brick y timber para los más nuevos y los más viejos en base a adobe mud y mud mortar brick. Esto nos lleva a pensar que para las edificaciones más nuevas se decidió cambiar los materiales con los que hacían las edificaciones.

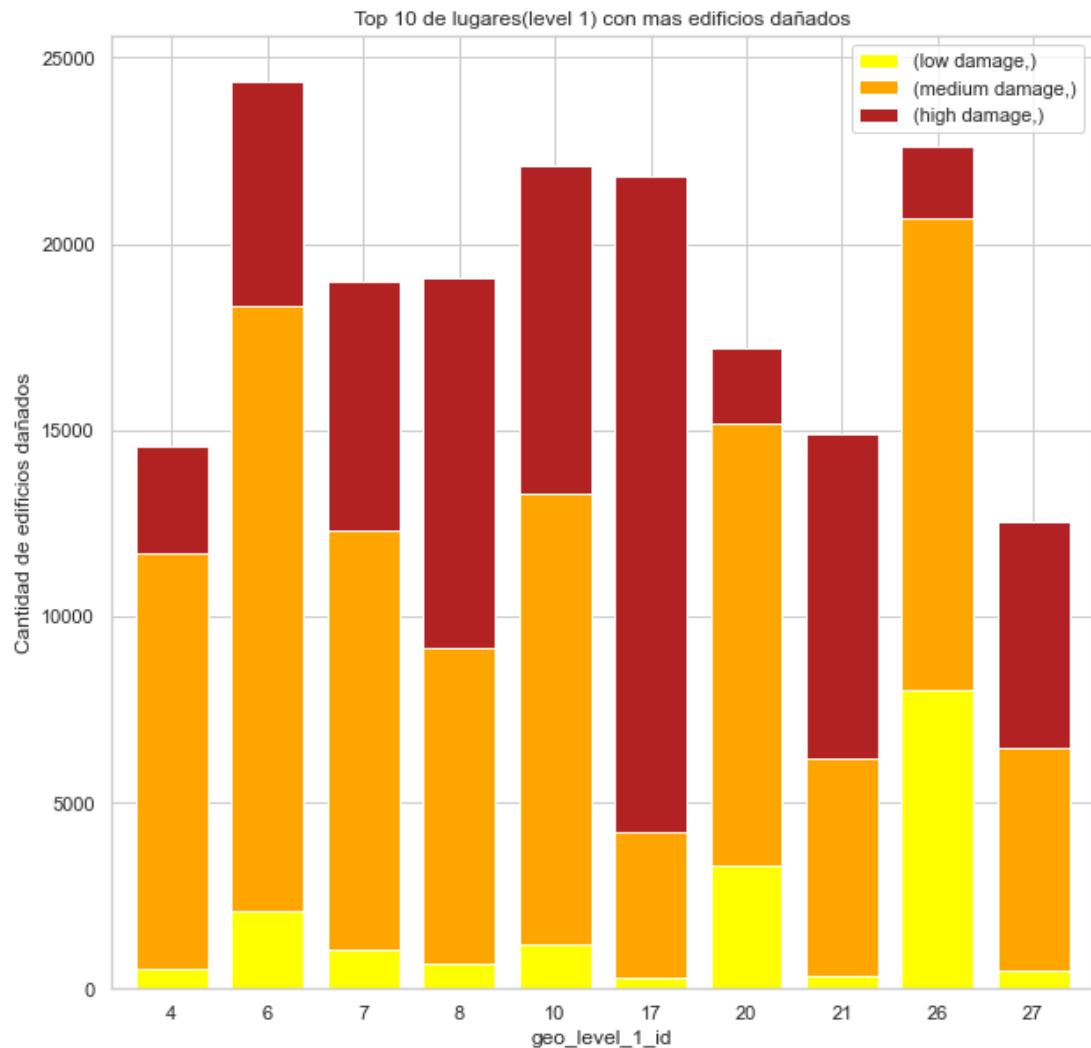
has_secondary_use

En cuanto a las columnas has_secondary_use, sólo el 11% de las edificaciones tenía algún uso secundario así que optamos por no analizarlas más a fondo ya que no representan la mayor parte de las mismas sino que serían casos más particulares.

geo_level_x_id: Análisis por geolocalización

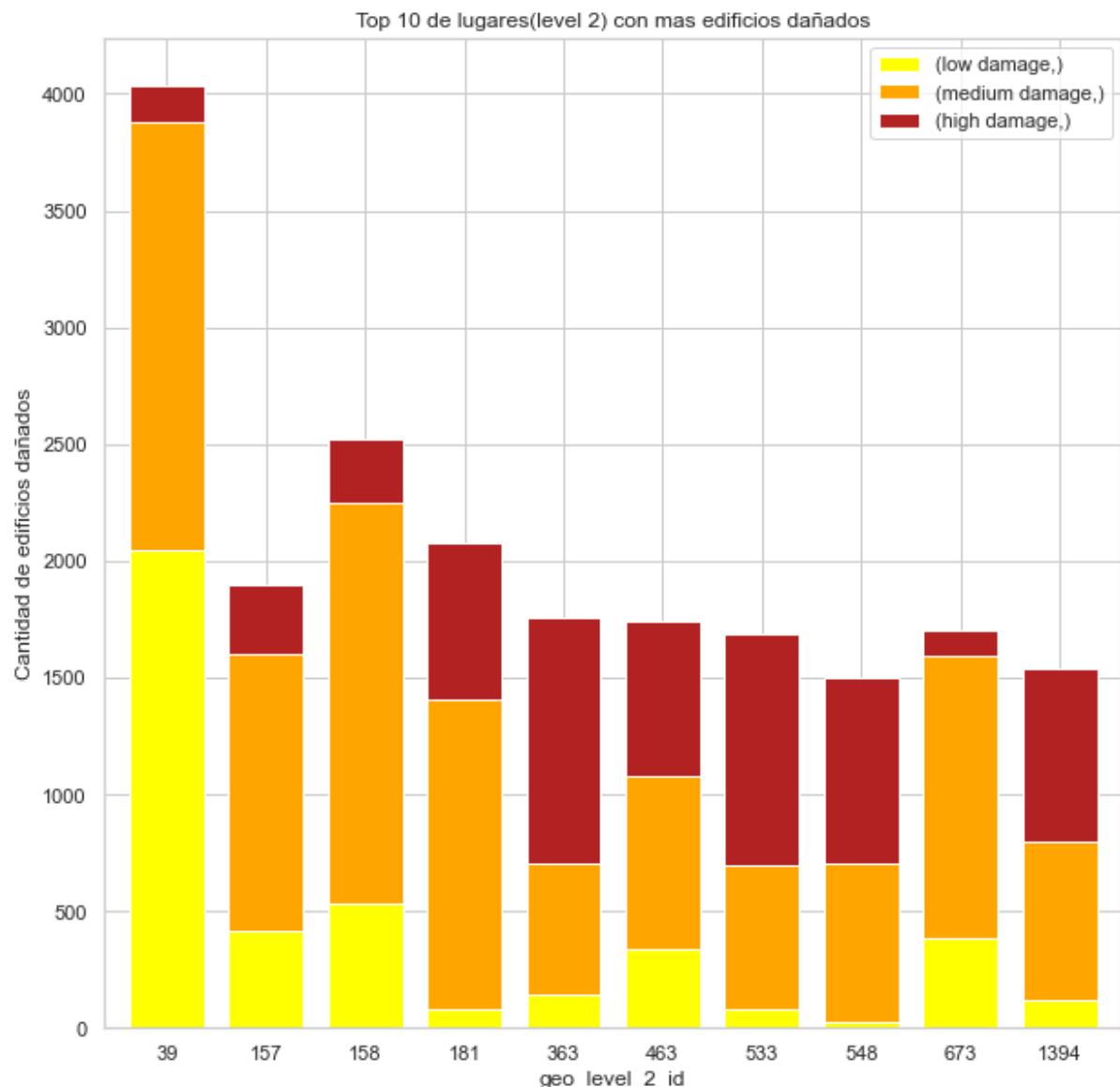
Las geo_level_id son datos que indican la zona a la que pertenece un edificio, el número de level determina el nivel de especificidad de la zona, 1 sería lo mas general, 3 la zona mas específica.

geo_level_1_id



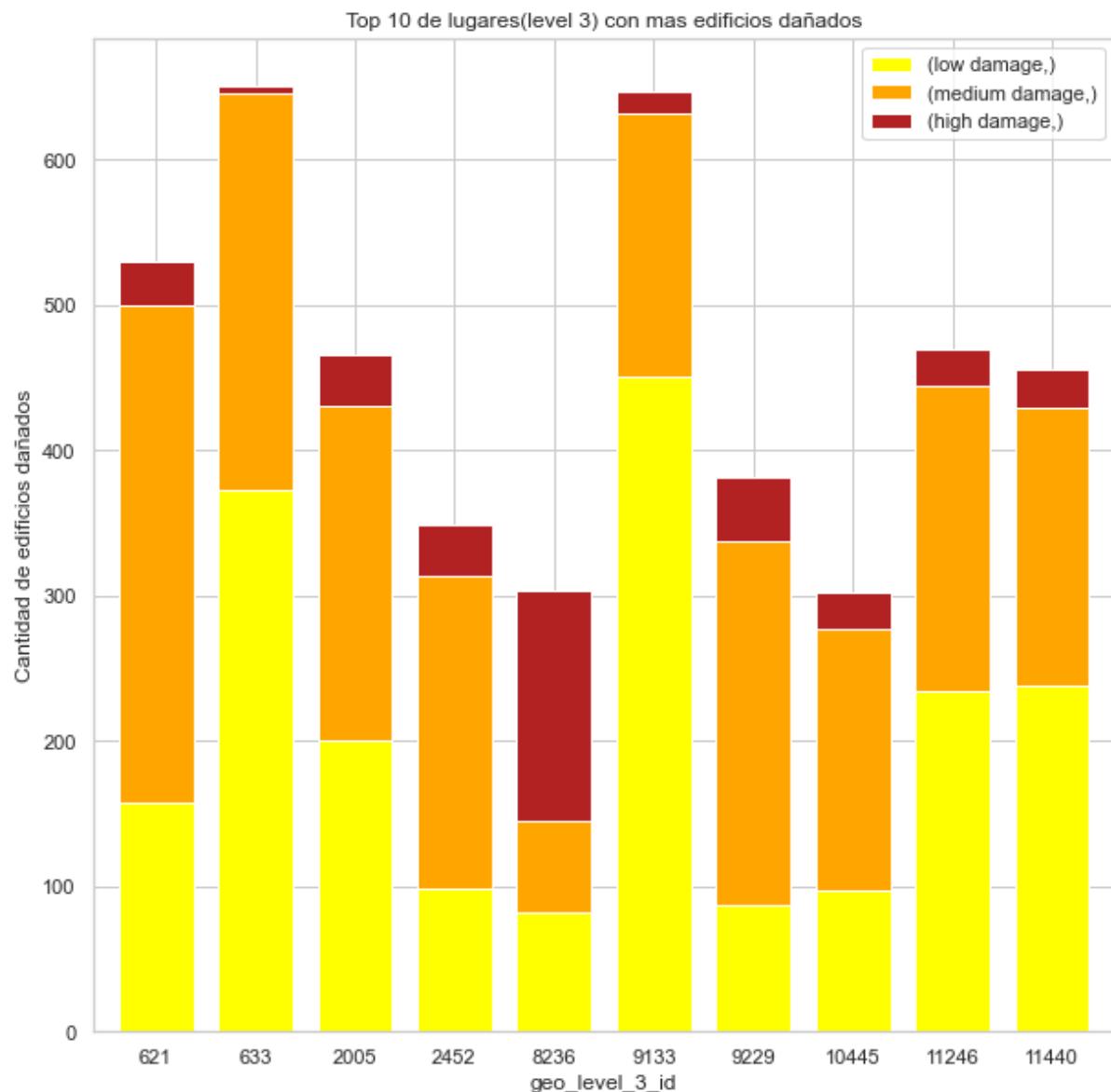
Podemos observar que entre las zonas más afectadas la gran mayoría de los daños que se encuentran son graves o medianos, esto puede deberse tanto a la localización de estos edificios como a otros factores como los materiales de los que están hechos o sus características.

geo_level_2_id



Como vemos cuando vamos a zonas más específicas dentro del grupo original(las 10 zonas de nivel 1 con más casos) los casos siguen siendo en su mayoría de mediana gravedad, con la excepción de la zona 39 donde se ve una mayor cantidad de casos leves, lo que nos hace pensar que pertenece casi con seguridad a la zona 26 de nivel 1 dado que en el resto no vemos una gran cantidad de casos leves, y las zonas 363 y 548 donde predominan los casos graves.

geo_level_3_id



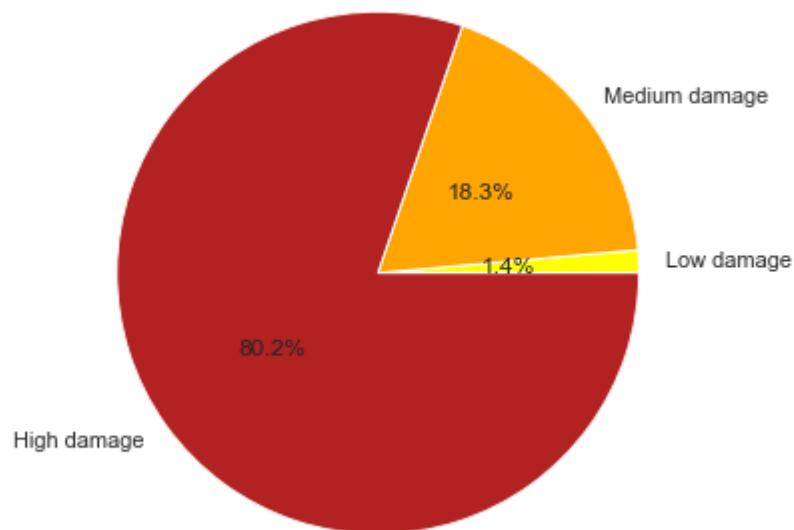
Si nos movemos a las zonas de nivel 3(de entre las zonas que se investigaron en las 2 anteriores visualizaciones) empezamos a ver que las zonas empiezan a tener un mayor cantidad de casos de daños leves siendo la predominante en casi la mitad de las zonas expuestas, esto puede ser debido a que la zona de nivel 2 más numerosa era la 39 que tenía una predominancia de casos leves.

Zona 17 del nivel 1(geo_level_1 id)

Como esta era la zona con más edificios gravemente dañados parecía interesante revisar cuál podría ser la causa y si se observaba alguna diferencia en algún apartado con respecto del resto de datos.

Zona 17 por antigüedad:

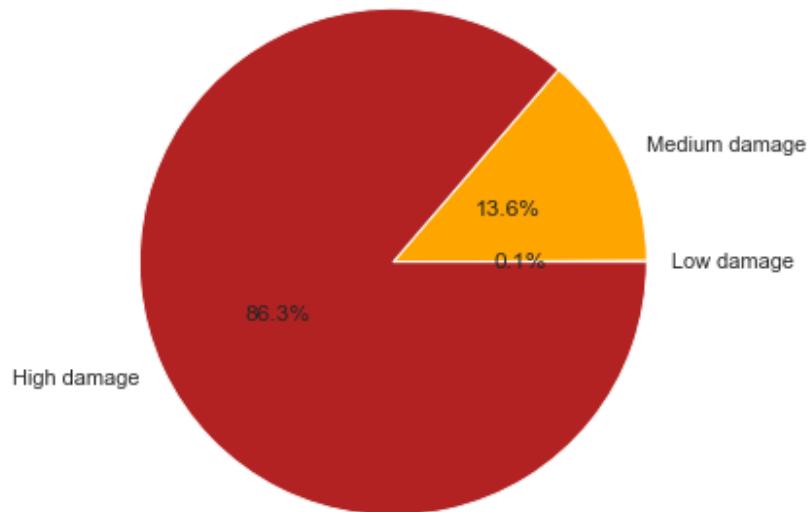
Porcentaje de daño de buildings de la zona 17(level 1) de hasta 50 años de antigüedad



Como podemos observar la gran mayoría de daños en edificios de hasta 50 años son de alta gravedad, marcando una diferencia bastante grande con el gráfico de la sección de age con estos mismos parámetros donde la mayoría de edificios habían sido dañados con daños medios, sumado a esto no hay casi edificios con daños leves.

Si intentamos observar los datos con más de 50 años encontramos esto:

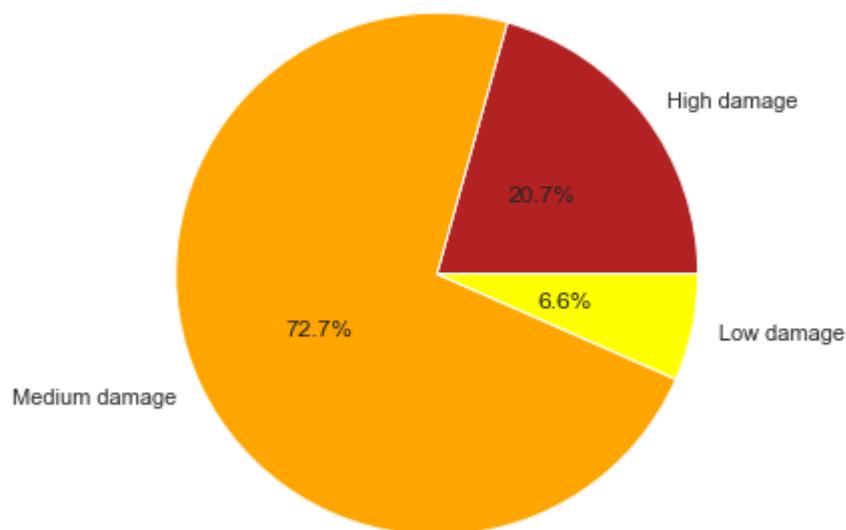
Porcentaje de daño de buildings de la zona 17(level 1) a partir de 50 años de antigüedad



Como podemos ver en estas visualizaciones los porcentajes de daño de la zona 17 son muy distintos a los de el set de datos en general, predominando los daños graves(High damage) cuando en el set de datos por lo general predominan los daños medios(Medium damage) y quedando los daños leves(Low damage) prácticamente desaparecidos. Por ahora no es posible saber si estos daños fueron por la localización de esta zona o si tuvieron algo que ver otros factores.

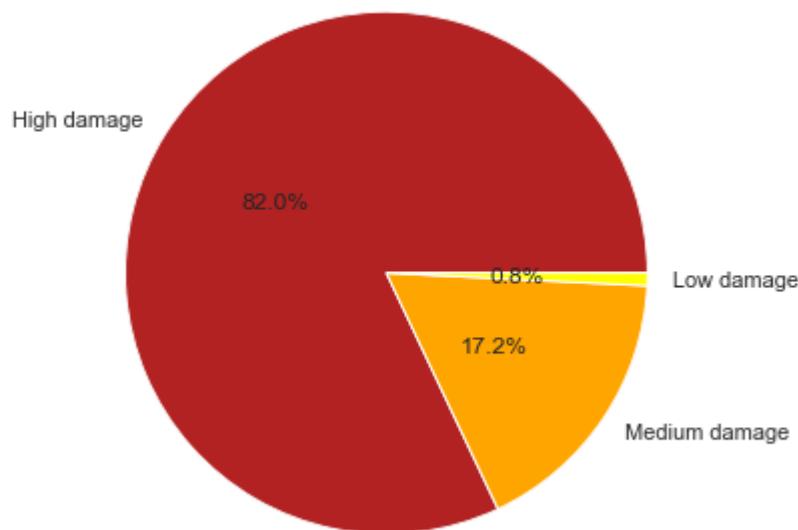
Zona 17 por cantidad de pisos:

Porcentaje de daño de edificios de un piso en la zona 17



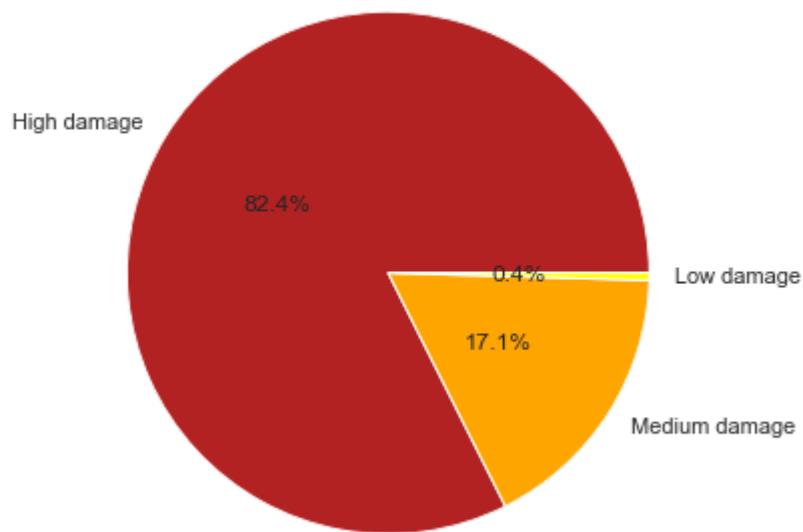
Podemos ver que entre los edificios de 1 piso de esta zona los daños fueron en su mayoría de mediana gravedad cuando en principio con lo que ya habíamos visto era de esperar que predominaran los daños graves.

Porcentaje de daño de edificios de 2 pisos en la zona 17



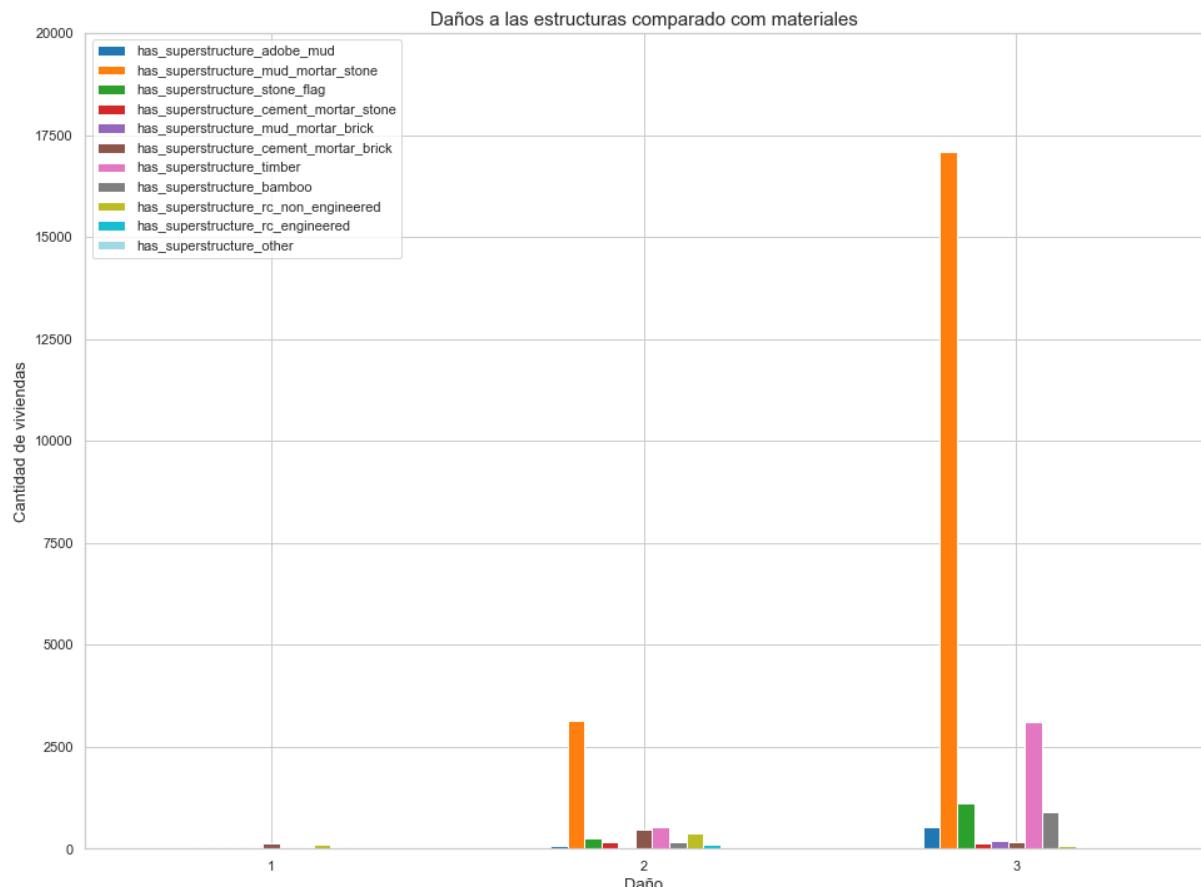
Como vemos cuando pasamos a los edificios con 2 pisos (la categoría más numerosa dentro de esta zona) los daños son mayoritariamente graves otra vez despegándose de lo que se podía ver en el análisis de todo el set de datos.

Porcentaje de daño de edificios de 3 pisos en la zona 17



Los edificios de 3 pisos presentan un gráfico muy parecido a los de 2 pisos con porcentajes casi exactamente iguales.

Zona 17 por materiales/superestructuras:



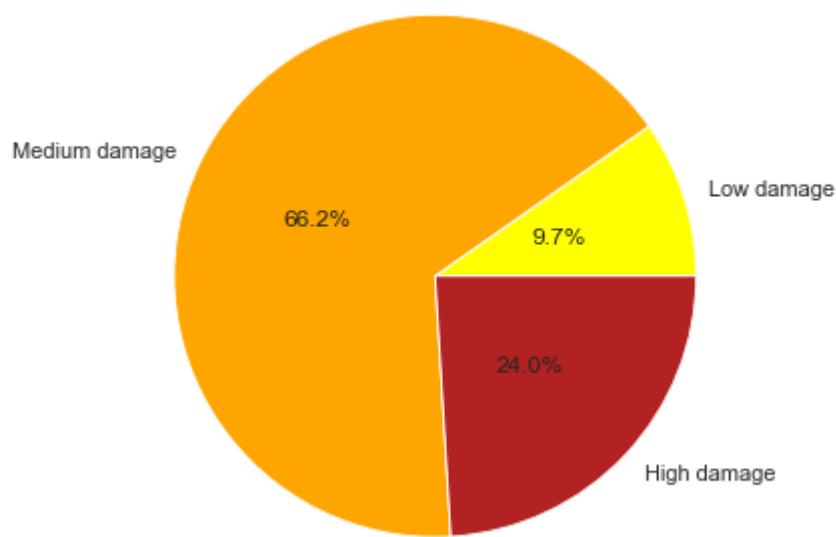
Podemos observar que en las edificaciones con daño 3 se utilizó como material mayoritariamente mud_mortar_stone, con la madera, el bambú y stone_flag siguiéndolo, en las edificaciones de daño 2 vemos que mayoritariamente se usó también mud_mortar_stone con la madera, cement_mortar_brick y con la superestructura rc_non_engineered, como vemos los edificios que fueron menos dañados casi no usaron ninguna de estas superestructuras.

Zona 6 del nivel 1(geo_level_1_id)

Como esta era la zona con más edificios parecía interesante revisar cuales eran las razones de porque tantos edificios recibieron algún daño.

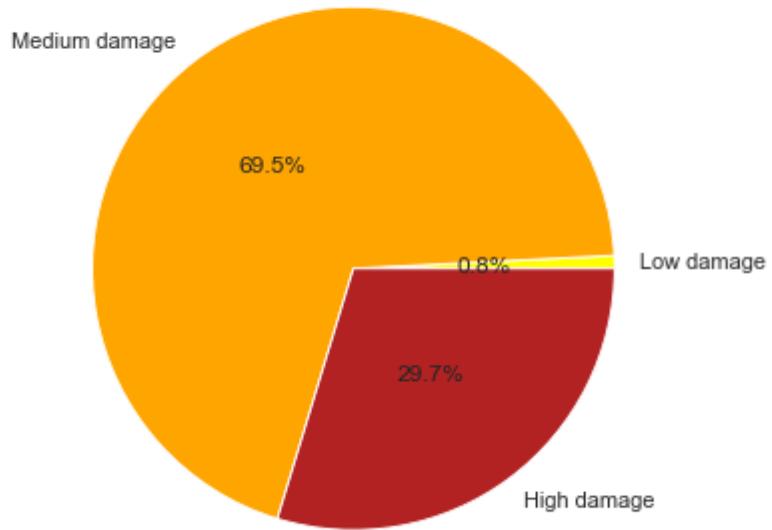
Zona 6 por antigüedad:

Porcentaje de daño de buildings de la zona 6(level 1) de hasta 50 años de antigüedad



En este gráfico a diferencia del de la zona 17 podemos ver que fue dañado de una manera más parecida a como ocurrió en el general de los datos.

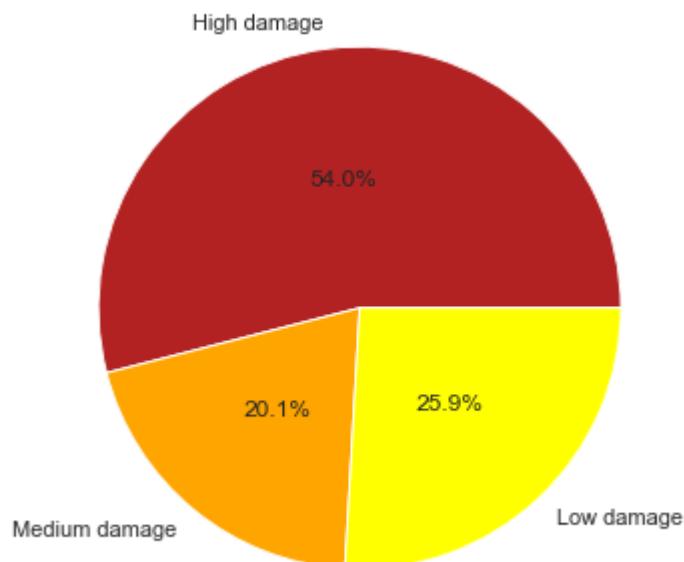
Porcentaje de daño de buildings de la zona 6(level 1) a partir de 50 años de antigüedad



Otra vez podemos observar como los daños se reparten de manera parecida al análisis general de age lo que nos puede decir que la zona 17 fue simplemente una anomalía.

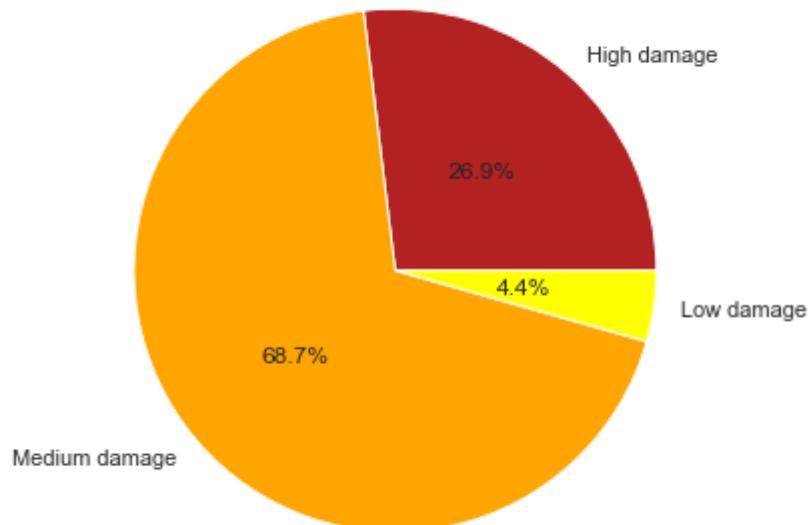
Zona 6 por cantidad de pisos:

Porcentaje de daño de edificios de un piso en la zona 6

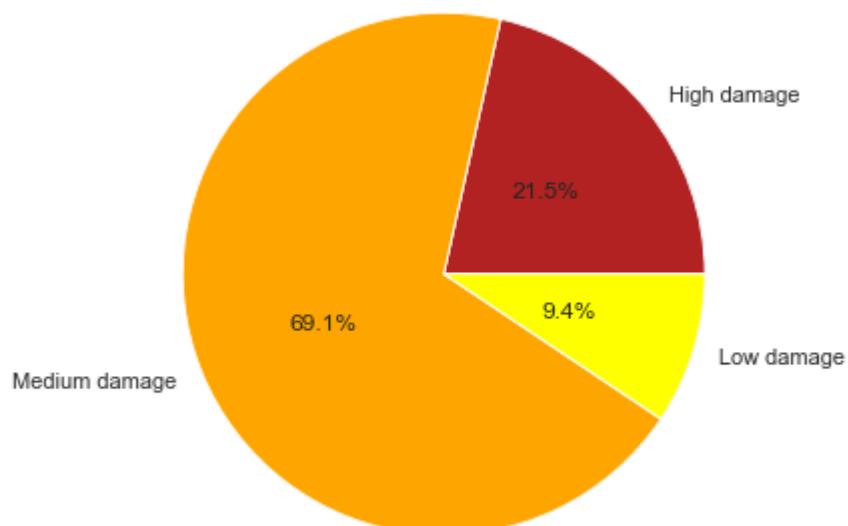


En esta zona a diferencia de la anterior(la 17) y a diferencia del general al parecer los edificios de 1 piso sufrieron una gran cantidad de daños graves aunque es cierto que la cantidad de datos de este tipo con los que contamos son bastante menos numerosos.

Porcentaje de daño de edificios de 2 pisos en la zona 6

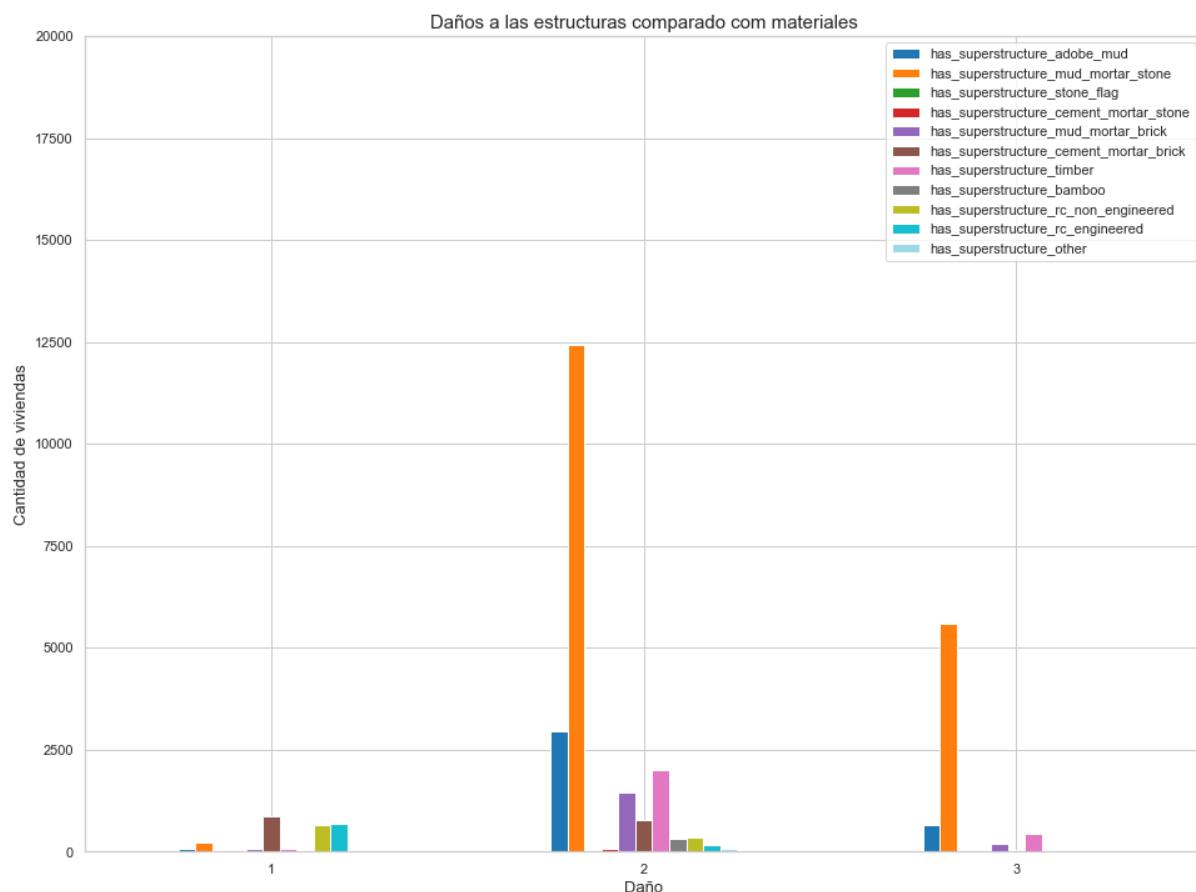


Porcentaje de daño de edificios de 3 pisos en la zona 6



Volvemos a ver una diferencia con los otros gráficos en los edificios de 2 o 3 pisos donde la mayor parte de los daños son medianamente graves.

Zona 6 por materiales/superestructuras:



Vemos de nuevo la predominancia de mud_mortar_stone como material, además en este caso se suma el adobe_mud como el segundo en los edificios más dañados, la madera vuelve a tener una gran cantidad y el mud_mortar_brick también se encuentra en gran cantidad en los edificios medianamente dañados.

Los edificios menos dañados usaron las superestructuras de rc_engineered y rc_non_engineered tambien usaron como material más predominante el cement_mortar_brick por lo que podemos intuir que estas superestructuras ayudan a disminuir la cantidad de daños que reciben los edificios.

Conclusión

Tras todo este análisis podemos concluir que:

- La antigüedad de los edificios por sí solo no es un factor determinante del daño que recibieron.
- Las zonas por geo_level_1 que más cantidad de edificios presentan también fueron las más afectadas.
- Los edificios más dañados casi no presentaban superestructura rc (rc_engineered y rc_non_engineered).
- Mud mortar stone es de los materiales más utilizados en la región. Se lo suele combinar con timber y bamboo.
- Las edificaciones que no están hechas de mud mortar stone sufrieron menos daños y estaban hechas de cement mortar brick, mud mortar brick y timber en su mayoría.
- La mayoría de las superficies edificadas poseen una condición de tipo "t" la cual fue la más afectada.
- Edificaciones cuyos valores estructurales (cantidad de pisos, superficie y altura normalizada, condición de superficie, orientación y diseño sísmico) por separado fueron los más dañados a su vez en conjunto recibieron mayor daño. Sin embargo, esto no se puede afirmar debido al escaso muestreo de las edificaciones que presentan dichos valores.