

Trabajo Práctico I

Organización de Datos [75.06]

Primer cuatrimestre de 2021

Betz, Ghersi, Fontela, Francetich, Rosenblatt

104348, 104330, 103924, 104048, 104105

Índice

1. Introducción	3
1.1. Terremoto	3
1.2. Motivaciones	3
1.3. Preprocesamiento de Datos	4
1.3.1. Información	4
1.3.2. Consumo de Memoria	4
2. Desarrollo	6
2.1. Introducción al Análisis de Daño	6
2.2. Daño Social	10
2.2.1. Promedio de Edad	13
2.3. Características Principales de las Estructuras	14
2.3.1. Antigüedad	15
2.3.2. Altura	17
2.3.3. Pisos/Niveles	20
2.3.4. Área	22
2.4. Geografía	24
2.5. Materiales	26
2.6. Algunos gráficos mas avanzados	29
2.6.1. Concentración de Alturas	29
2.6.2. Daño en Función de la Edad y Altura	30
2.6.3. Altura y Edad	32
2.6.4. Destrucción por Material en Función de la Edad	33
2.6.5. Materiales más Viejos	36
2.6.6. Daño Promedio	36
3. Conclusiones	39
3.1. El terremoto y sus consecuencias	39
3.2. Edificaciones	40
3.3. Ubicación Geográfica	40
3.4. Materiales	41

4. Repositorio

42

1. Introducción

En el presente se desarrollará un análisis exploratorio acerca de las diferentes variables que intervienen en el impacto de los terremotos sobre diferentes tipos de estructuras. Los datos fueron obtenidos como consecuencia del terremoto de [Gorkha en Nepal del año 2015](#).

1.1. Terremoto

El 25 de abril de 2015, en Nepal, ocurrió un sismo de magnitud 7,8 a 8,1 en la escala sismológica de magnitud. Se registró a las 6:11 GMT (11:57 hora local) del sábado 25 de abril de dicho año y que afectó a ese país del Asia del Sur. Tuvo su hipocentro a una profundidad de 15 km y su epicentro se localizó a 81 km al noroeste de la capital Katmandú, concretamente en el distrito de Lamjung.

Ha sido el terremoto más grave que ha sufrido Nepal desde el acontecido en 1934. Con casi 9000 víctimas mortales, aproximadamente 22.000 heridos y altos costos producto del daño material, entre ellos la plaza Basantapur Durbar de Katmandú y otros edificios emblemáticos, como la torre Dharahara y el templo Manakamana esta resulta ser una gran catástrofe.

1.2. Motivaciones

El análisis que se realiza en este trabajo se basa en los diferentes aspectos de la ubicación de las edificaciones, además cuenta con diferentes análisis respecto de el material con el que está construido el mismo, la antigüedad de la estructura al momento de sufrir el accidente, la utilidad que se le daba a la estructura, la existencia de familias en la estructura.

A su vez, se hace hincapié en la ubicación geográfica del inmueble, así-

mismo como las magnitudes físicas de este; el área que el mismo abarca, la altura que alcanza a medir y la superficie del mismo, todos normalizados.

Puede resultar de gran utilidad usar las conclusiones del mismo para evitar futuras catástrofes causadas por una tragedia como esta.

1.3. Preprocesamiento de Datos

1.3.1. Información

El DataSet cuenta con 260601 edificaciones particulares de las cuales se tienen registro único. Para estas se hallan datos específicos que se utilizan para el análisis de los desastres ocurridos.

Los dos DataFrames no cuentan con datos *null*.

1.3.2. Consumo de Memoria

En el proyecto podemos visualizar que el DataFrame original ocupa 77.5MB. Debemos, antes de comenzar con el análisis, reducir el tamaño del DataFrame a utilizar para no cachear tanta información innecesaria.

```
dtypes: int64(31), object(8)
memory usage: 77.5+ MB
```

Primero transformamos los datos, de tipo *'Object'* a *'Category'* lo cual hace que decrezca el tamaño total del DataFrame a 63.6MB.

```
dtypes: category(8), int64(31)
memory usage: 63.6 MB
```

Luego debemos procesar los datos numéricos, ya que valores que son flags binarios (es decir, 1s o 0s) pueden ocupar un espacio mucho menor al que

ofrece un `np.int64`.

```
dtypes: category(8), int16(3), int32(1), int8(27)  
memory usage: 11.2 MB
```

Y el tamaño total del DataFrame pasa a ser de 11.2MB.

Todo esto lo volvemos a realizar para el DataFrame que únicamente mapea valores de `'building_id'` a `'damage_grade'`. El mismo pasa de pesar 4MB a 1.2MB.

2. Desarrollo

2.1. Introducción al Análisis de Daño

Cuando se produce un terremoto u otro desastre natural, las instituciones gubernamentales de socorro, las aseguradoras y otros organismos de respuesta convergen para hacer un balance de las muertes y las lesiones, y para evaluar el alcance y el costo de los daños a la infraestructura pública y los bienes personales.

Este tipo de desastres naturales vienen acompañados de incendios, colapso de construcciones que resultaron dañadas durante el evento principal por la ocurrencia de réplicas, saqueos, interrupción de actividades económicas, entre otras.

Debido a esto, tras este acontecimiento se tiene registro de un grado de daño para cada una los inmuebles. El grado de daño es el alcance o extensión de los daños producidos por un siniestro.

En este caso, nos referimos a los daños producidos tras el terremoto. El siniestro puede ser considerado en 3 grados: 1, 2 y 3. Véase el 1 siendo el mas leve y 3 como el mas agravado. Sobre todas las estructuras que sufrieron daños, el grado 2 concentra una densidad mayor en comparación a los grados 1 y 3.

El resto de las columnas del DataFrame son valores 1/0 en función de si se cumplen diferentes flags o valores alfa-numéricos que pueden representar cosas como la ubicación, propiedades físicas del inmueble, etc.

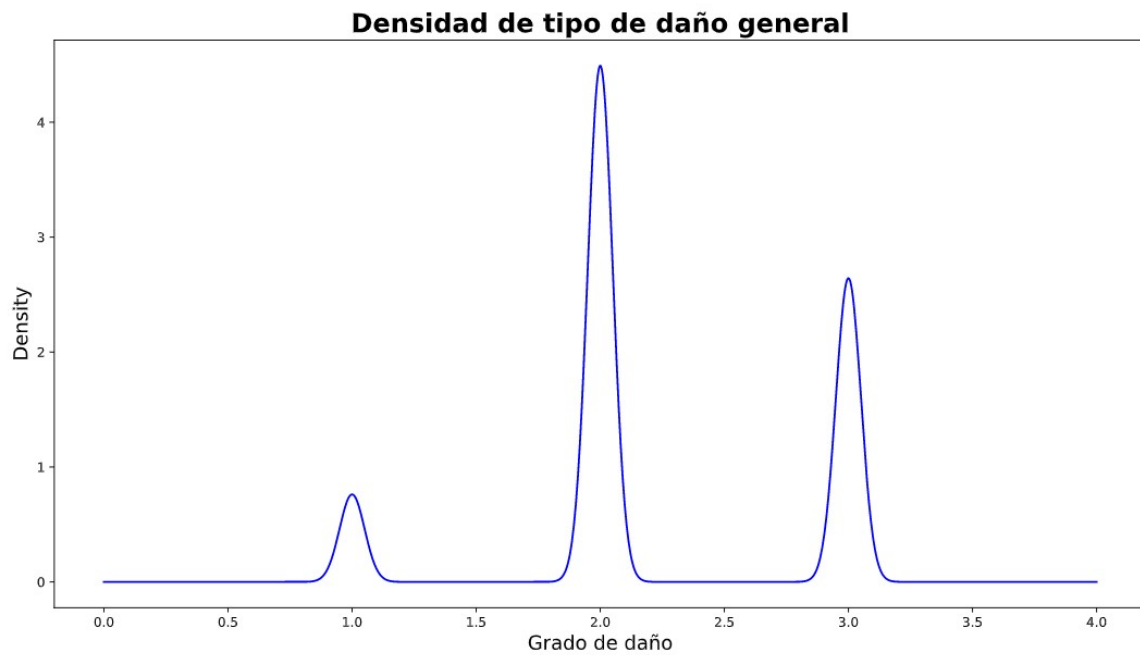


Gráfico de densidad de daño

Este ponderamiento del daño de grado 2 se puede apreciar detalladamente por sobre la cantidad total de estructuras, siendo este grado el mas representativo de los 3.

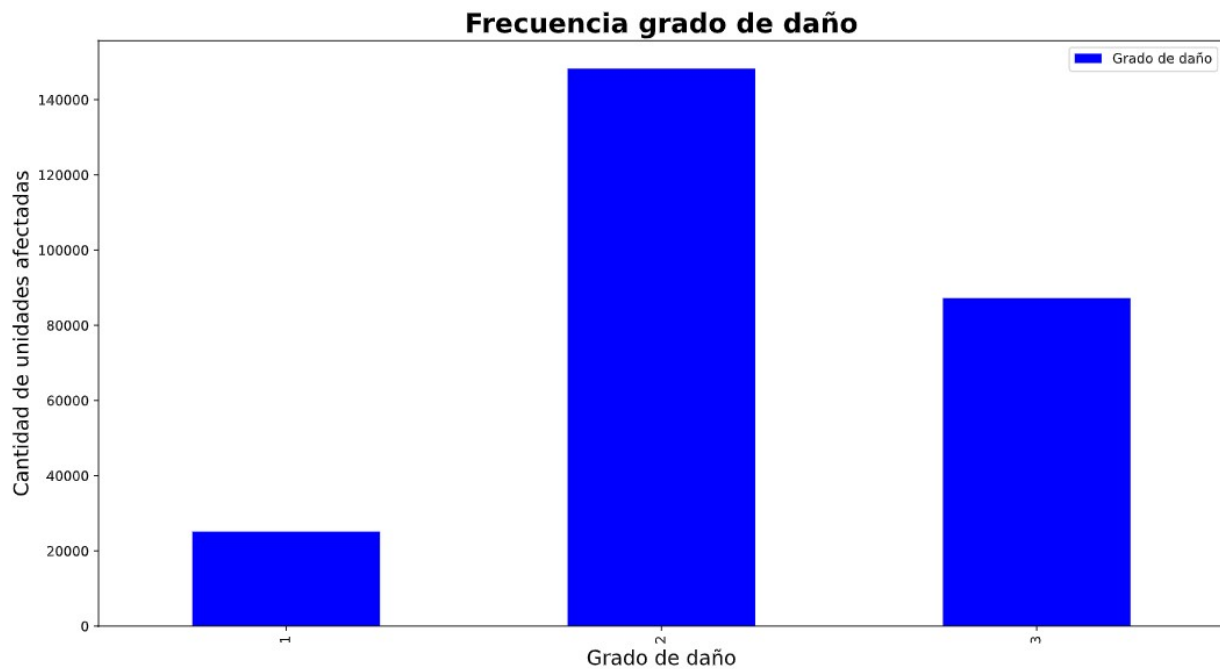
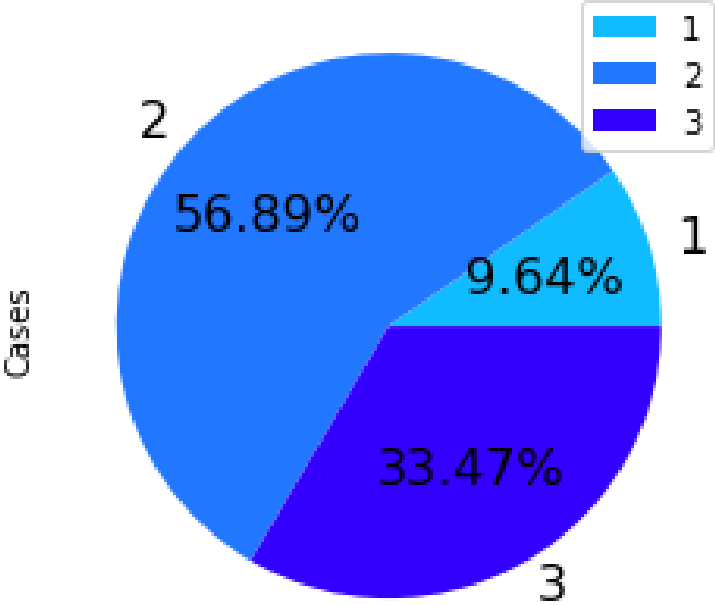


Gráfico de frecuencia de daño por estructura

Además se puede deducir en base a estos gráficos que el terremoto del 25 de abril de 2015 causo en promedio un grado de daño 2 sobre la gran mayoría de edificaciones que se vieron afectadas por este, dejando lógicas secuelas socio-económicas por los daños ocasionados.

Por ultimo les presentaremos este gráfico que representa el porcentaje del total de edificios dañados por grado de daño, este junto al gráfico anterior nos servirá para visualizar la cantidad de apariciones de los datos.

Porcentaje General de Grado de daño

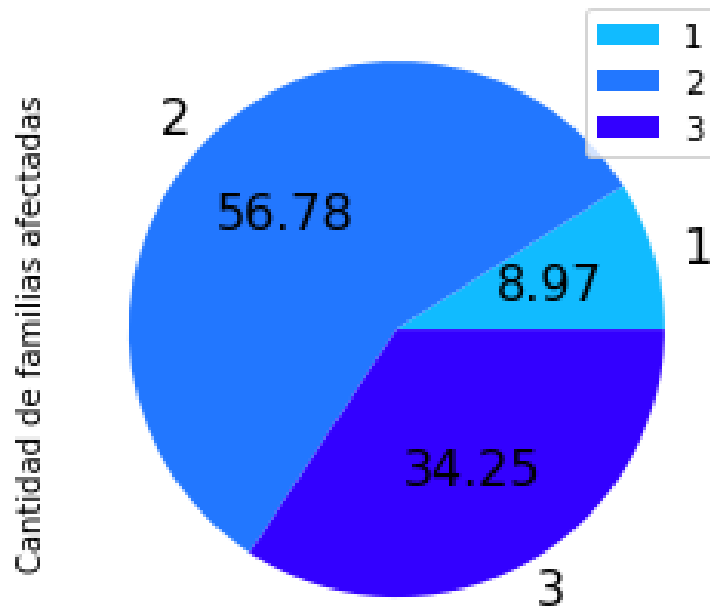


Porcentaje General por Grado De Daño

2.2. Daño Social

En esta sección nos encargaremos de analizar el daño que le ha generado el terremoto a la sociedad según los distintos sectores de esta.

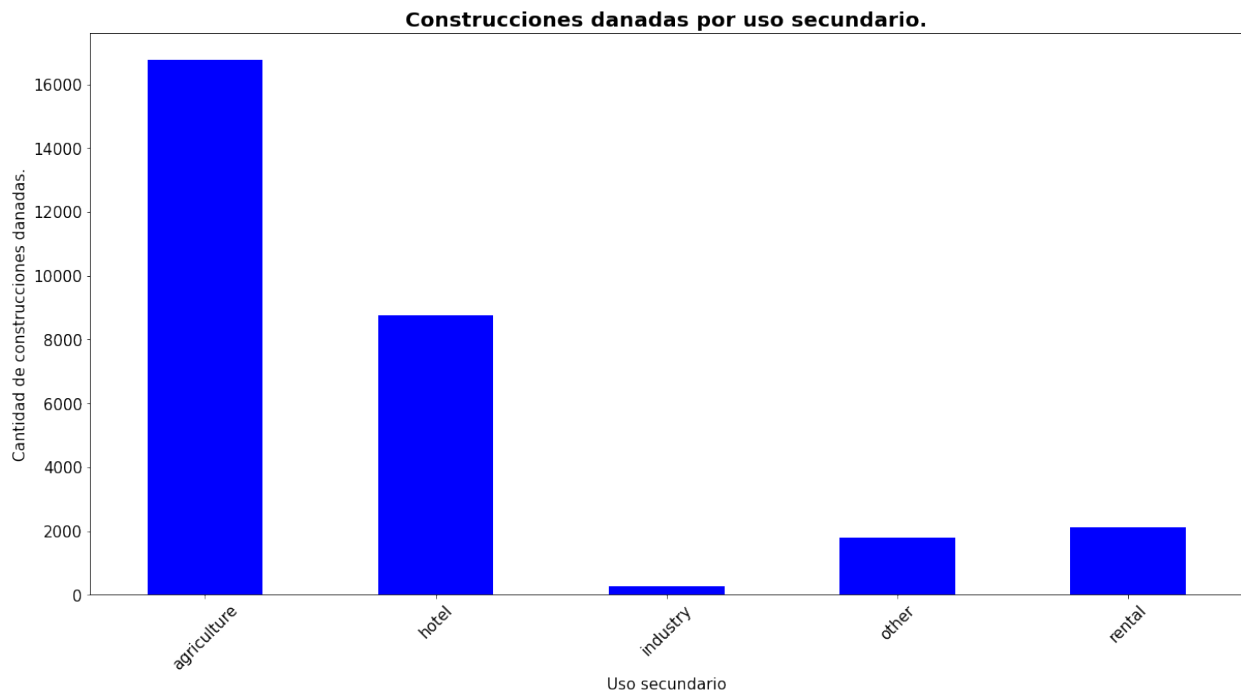
Familias afectadas por grado de daño.



Acá se puede observar que predomina el grado de daño 2 por sobre las familias afectadas.

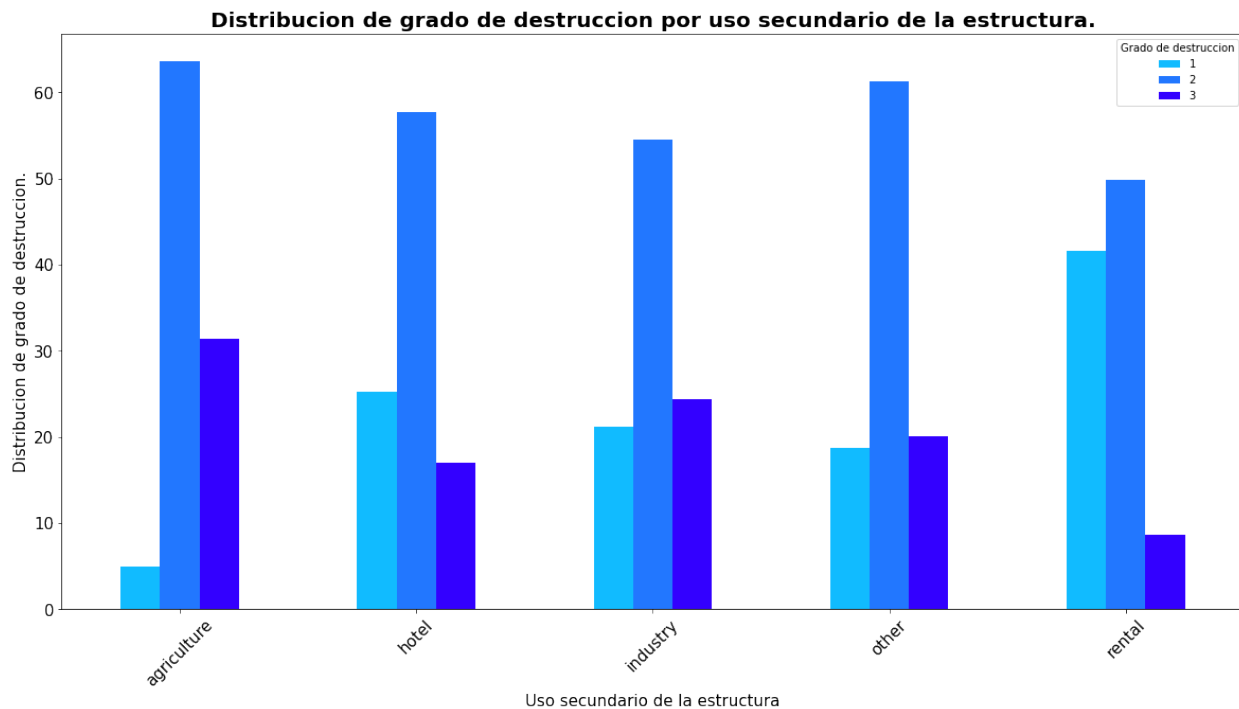
Este dato es de mayor relevancia a nivel social para obtener un panorama y tener un plan de reconstrucción en base a lo afectados que se pudieron ver los habitantes.

Y como segunda parte de esta sección, veremos que sectores fueron los mas afectados por el fenómeno. Analizaremos esto mediante el uso secundario de las estructuras dañadas:



Este primer gráfico presenta el uso secundario de las estructuras (eje x) y la cantidad de estructuras dañadas por cada categoría (eje y).

Podemos observar que las dos áreas mas afectadas por gran diferencia son: en primer lugar, agricultura; y en segundo lugar, hotelería. También han sido afectados menormente los sectores de alquileres e industrial. Finalmente, el conjunto de otros sectores que por separado no parecen significativos en la gráfica, han sido agrupados en la categoría 'other' ('otros') y alcanzan la misma cantidad de construcciones dañadas que alquileres.



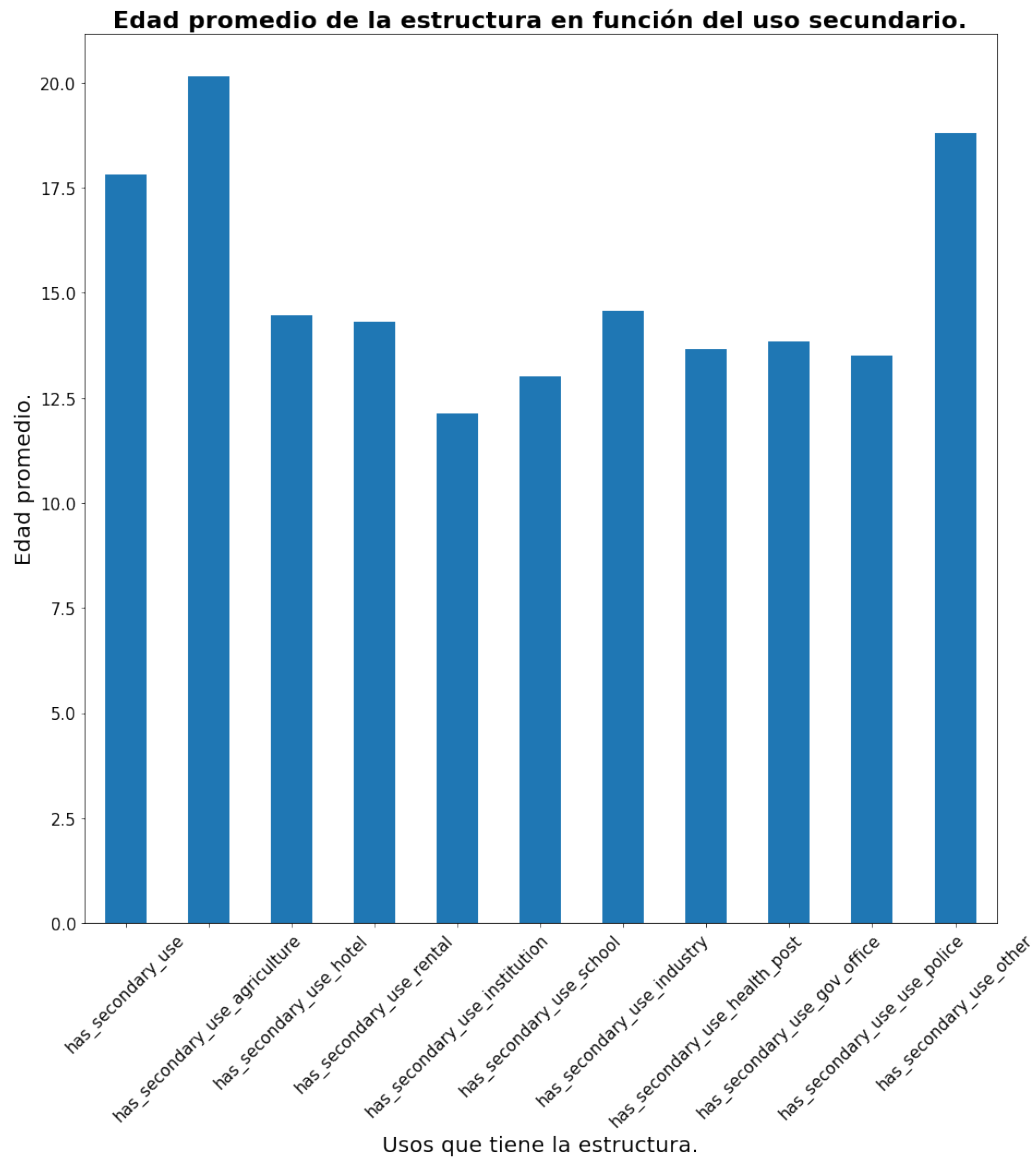
Por ultimo, en este segundo gráfico relacionado al uso secundario, podemos observar la relación entre esta clasificación (eje x) y, para cada grupo, la proporción de estructuras que recibió cada nivel de daño.

Como conclusiones podemos decir que:

- Hotelera, industria y otros tuvieron un destino bastante parecido, con un poco mas del 50 % de las construcciones dañadas con grado 2, y los grados 1 y 3 parejos en alrededor del 25 %.
- Alquileres es el que se ve mas favorecido por esta gráfica, ya que mas del 40 % de las destrucciones fueron de grado 1 y menos del 10 % de grado 3.
- El mas afectado (no solo en cantidad como habíamos visto, sino ahora también en daño) fue agricultura, sector en el cual casi el 30 % de las estructuras dañadas recibió daño de grado 3.

2.2.1. Promedio de Edad

Dentro de los promedios de edad, los inmuebles de utilización institucional (probablemente de privados), las escuelas y estaciones de policía tienden a ser las más nuevas, mientras que las más viejas son que se utilizan para la agricultura.



Esto puede ser una consecuencia directa del hecho de que Nepal es un país del tercer mundo con una falta de infraestructura moderna.

2.3. Características Principales de las Estructuras

Puede interesarnos ahora buscar relaciones entre ciertas características de las estructuras de nuestro conjunto de datos. Para este apartado, nos vamos a centrar en el análisis de la edad/longevidad de las construcciones, las alturas de ellas y el área de cada una (estos dos últimos valores están reducidos a un valor que lo único que nos permite es comparar una estructura con otra, ya que no están en ninguna unidad conocida - o por lo menos no se indica de tal manera en la documentación -).

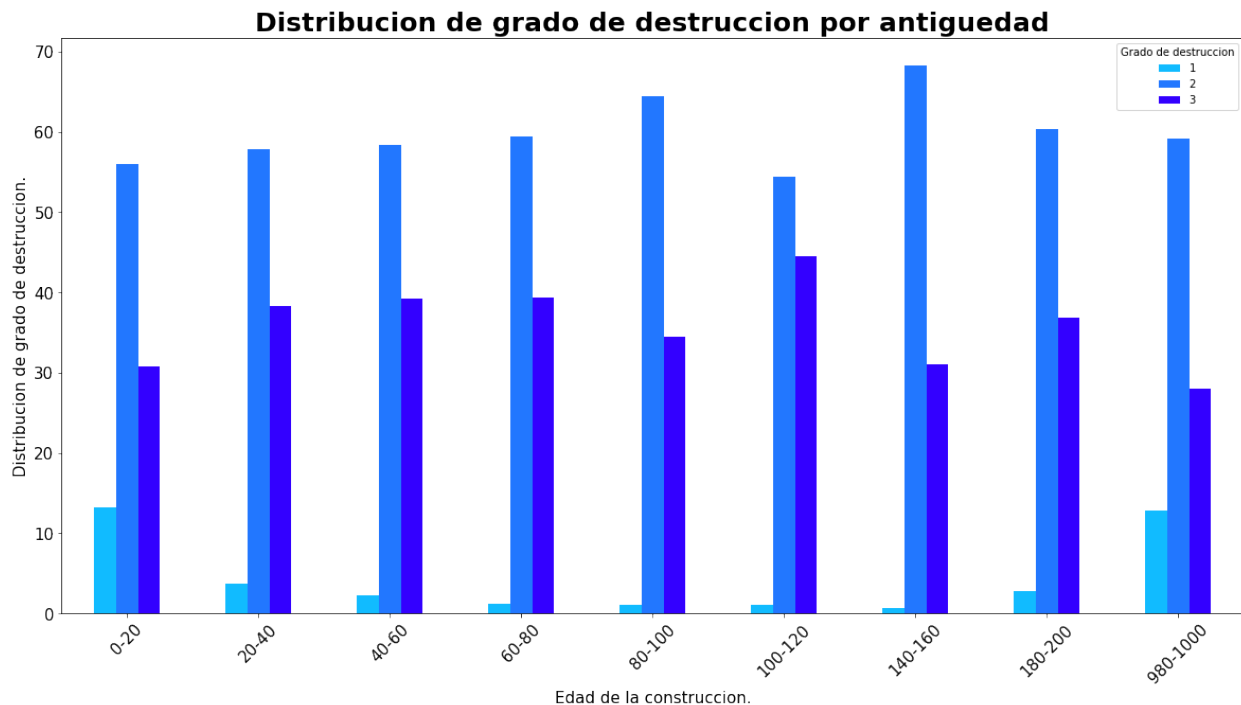
2.3.1. Antigüedad

En los dos gráficos a continuación podremos observar el análisis que realizamos en la relación antigüedad-daño.



En este primer gráfico podemos observar la relación entre la edad de las construcciones, y su media de nivel de destrucción. Aquí podemos sacar dos conclusiones:

- Cuando una estructura esta en el rango de 0 a 25 años de antigüedad, se nota un punto a favor de aquellas construcciones mas recientes (la media de destrucción oscila desde valores inferiores a 2.0 para construcciones nuevas hasta aproximadamente 2.35 para construcciones de 25 años de antigüedad).
- Para estructuras de antigüedad mayor a 25 años, la media de destrucción oscila aproximadamente entre 2.3 y 2.4, sin definirse un patrón.



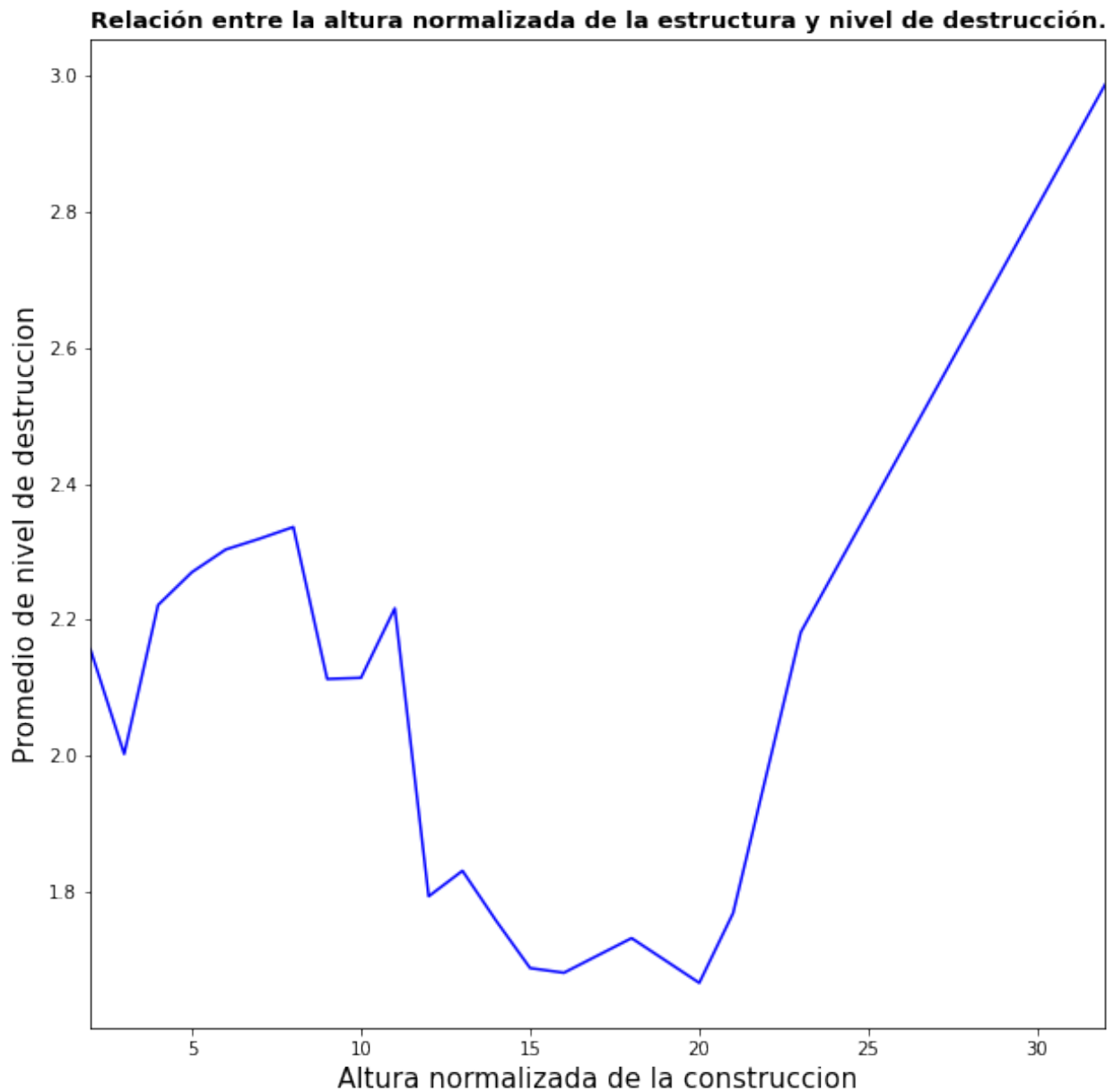
Por otro lado, en este segundo gráfico podemos analizar, para cada rango de 20 años (en los que encontramos un numero de datos lo suficientemente grande como para ser analizado), el porcentaje de construcciones que recibió cada nivel de daño.

Creemos que aquí no hay mayores patrones a destacar: aproximadamente en todos los rangos el daño de grado 1 oscila entre 0 y 15 % siendo máximo para aquellas construcciones o muy recientes o muy antiguas; además, en

todos los rangos el daño de grado 2 oscila entre 55-65 % y el grado 3 entre 30-40 %.

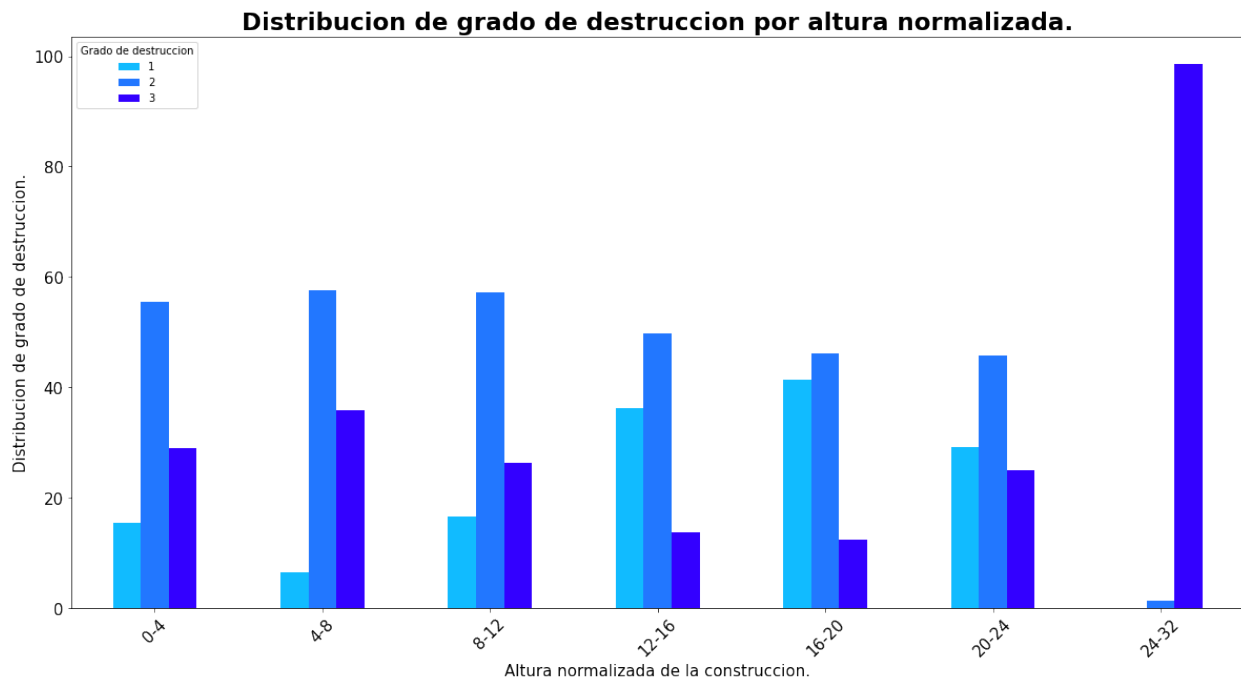
2.3.2. Altura

Ahora analizaremos, con el mismo formato que en los dos gráficos anteriores, la relación entre la altura de las edificaciones y el daño que recibieron como consecuencia del terremoto.



Podemos observar en el primer gráfico - que relaciona la altura de la construcción (en el eje x) con la media de grado de destrucción para esa altura (eje y) - que los valores intermedios de altura (entre 12 y 20) tuvieron los mejores resultados lidiando con el terremoto, con medias por debajo de 1.7 en los mejores casos. Para las edificaciones mas pequeñas este valor oscilo entre 1.8 y 2.3. Finalmente, las mas afectadas fueron aquellas mas altas: podemos observar que, en el rango de altura de entre 20 y 30, a medida que aumenta

esta, aumenta también la media de daño, alcanzando un máximo de 3.0.



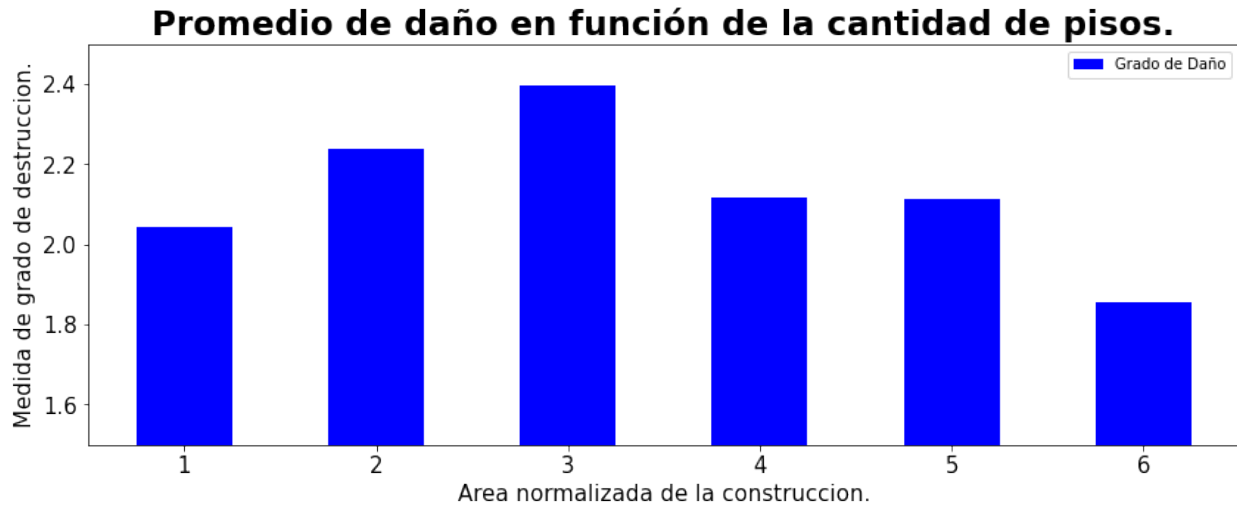
Por otro lado, en este ultimo gráfico podemos observar lo mismo al gráfico similar en la sección de antigüedad: para cada rango de altura, obtenemos el porcentaje de edificaciones que recibió cada uno de los tres niveles de daño.

En los valores menores a 20 podemos observar cierta tendencia de aumento de porcentaje de destrucción de grado 1 (compensado por una disminución en la destrucción de grado 3) a medida que crece la altura de las edificaciones. El porcentaje para el grado 2 presenta oscilaciones menores.

Para los valores mayores a 20 podemos ver, por un lado, en el rango de altura 20-24 una ruptura de la tendencia antes mencionada, y finalmente para el rango 24-32, que casi el 100 % de las destrucciones dañadas presenta daño de grado 3.

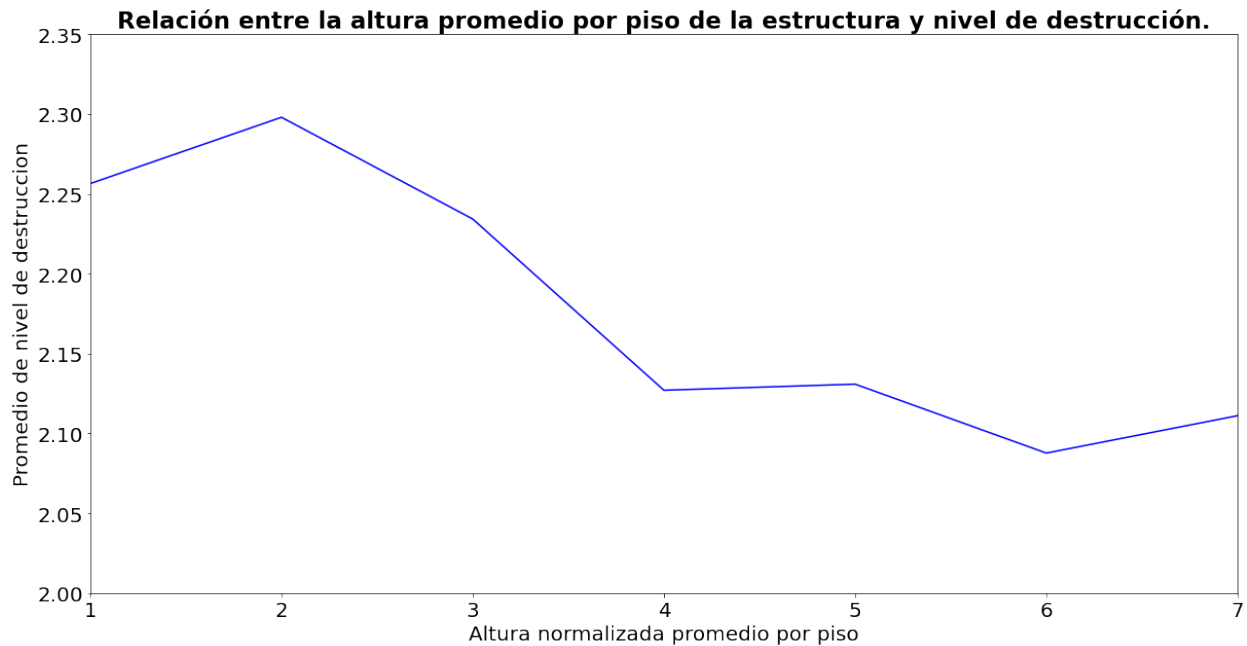
2.3.3. Pisos/Niveles

En esta tercera sección de características observaremos dos relaciones entre los pisos de las diferentes edificaciones y la destrucción recibida.



En el gráfico de arriba se puede observar la relación entre la cantidad de pisos de las edificaciones (eje x) y la media de daño recibido para esa cantidad de pisos (eje y).

Notamos un pico de destrucción en las edificaciones de tres pisos, y una disminución tanto hacia izquierda como hacia derecha. Por lo tanto, como conclusión obtuvimos que aquellas edificaciones mas dañadas fueron aquellas cuya cantidad de pisos mas se acercaba a 3, disminuyendo a medida que la cantidad de pisos se alejaba de este valor.



En este gráfico podemos analizar la relación entre la altura promedio por piso (eje x) y la media de grado de destrucción para esos valores (eje y).

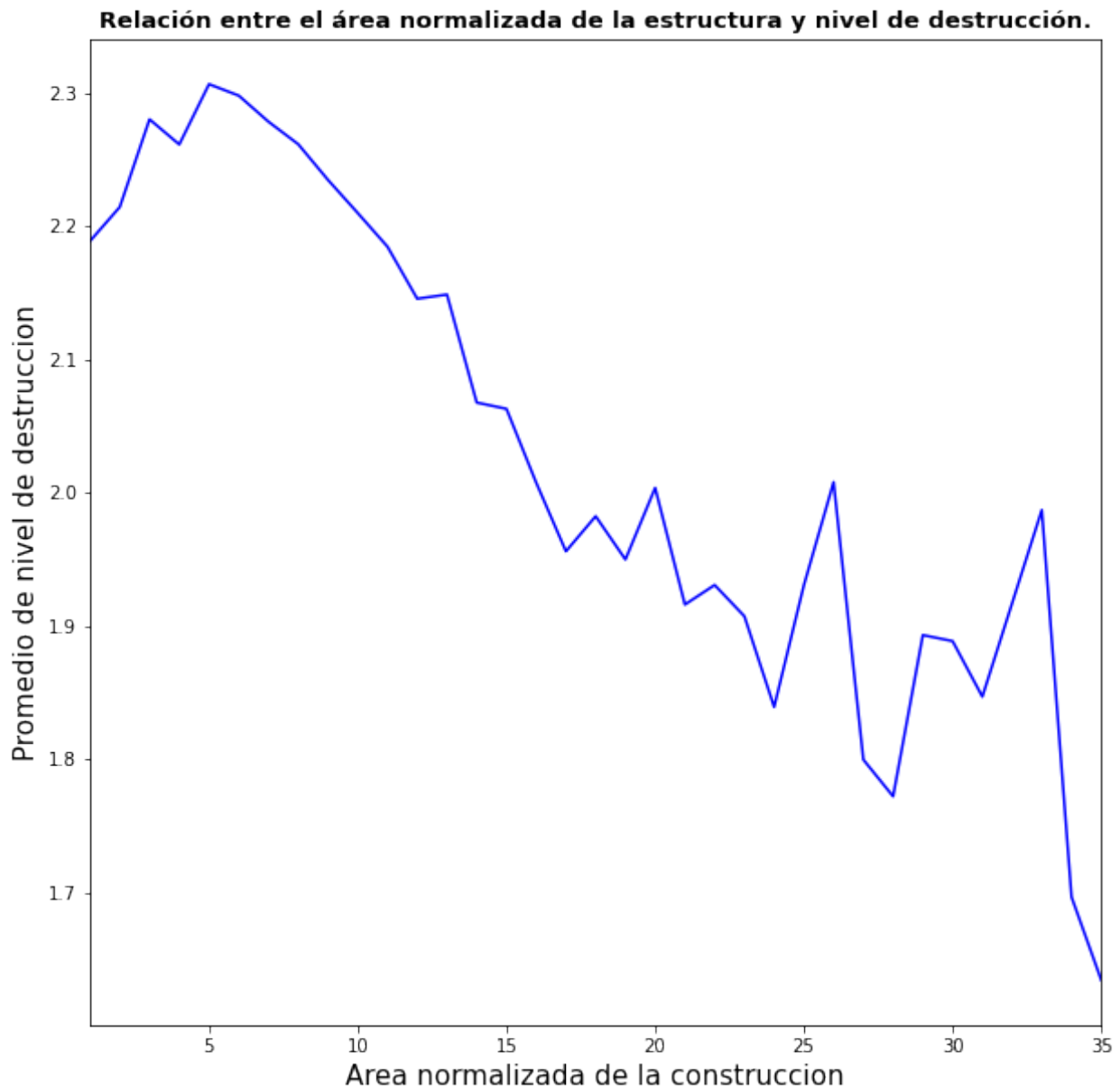
No se tuvo en cuenta el rango total al que se tenía acceso mediante el conjunto de datos dado ya que para alturas por piso mayores a 8, se presentaban pocos datos con oscilaciones muy grandes.

Podemos ver que hay una tendencia descendente, con la media de grado de destrucción oscilando entre 2.05 y 2.30, de tal manera que para aquellas edificaciones con una altura por piso mas baja se nota un aumento de la destrucción con respecto a aquellas que presentan mas altura por cada piso.

No sabemos a ciencia cierta si esta correlación implica causa (ya que no logramos encontrar la razón) pero consideramos valido mostrarla como base para futuros análisis.

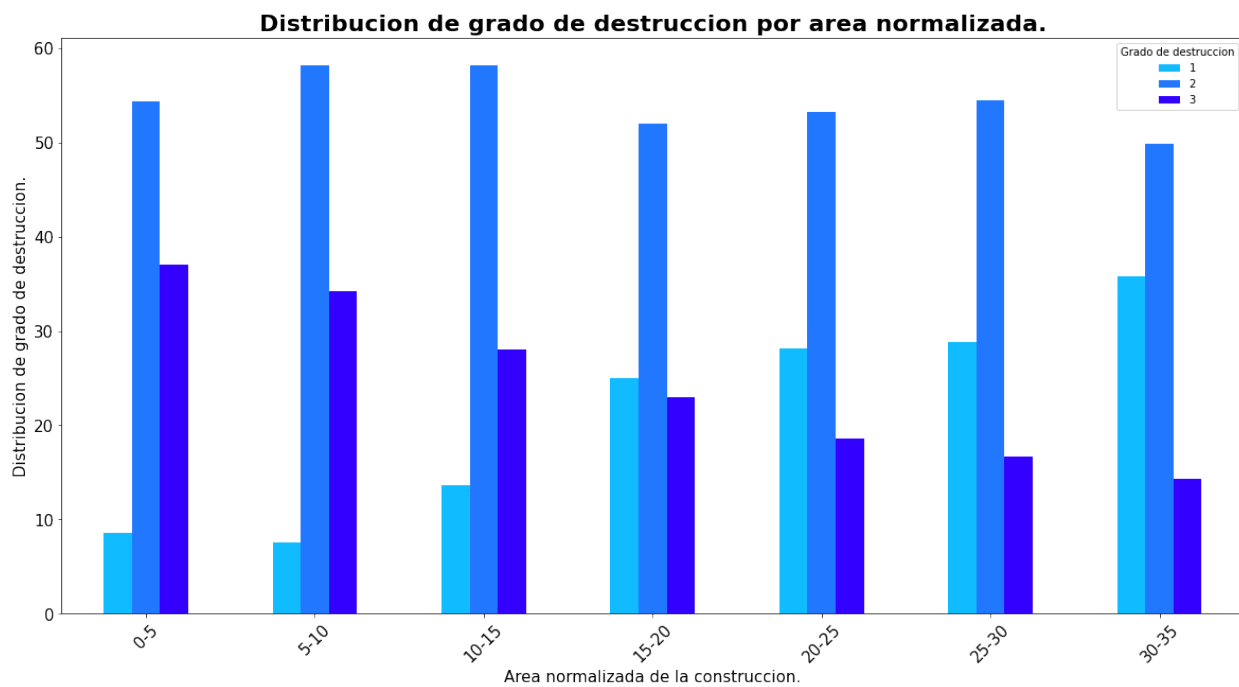
2.3.4. Área

En este ultimo apartado del análisis de las características principales de las estructuras analizaremos el área. Como se menciona mas arriba al principio de esta unidad, la documentación no especifica la unidad en la que esta representado el valor del área ni al área de que se refiere.



Para el primer gráfico (con los valores de área normalizada en el eje x, y

la media de grado de destrucción en el eje y) podemos llegar a una conclusión evidente con el solo hecho de verlo: las construcciones que presentan un mayor valor en "área" han resultado considerablemente menos dañadas. Así lo demuestra el pico de media de daño (2.3) para un valor de área cercano a 5, como también el mínimo del gráfico (una media por debajo de 1.7 de grado de daño) alcanzado en un valor de área de 35 (el máximo del conjunto de datos).



En este otro gráfico podemos observar la relación entre distintos conjuntos de construcciones clasificados por rangos de área de tamaño 5 (eje x) y el porcentaje de construcciones de ese rango que recibió cada grado de daño (eje y).

En todo el gráfico, podemos ver que el porcentaje para el grado 2 permanece casi constante, con oscilaciones menores que resultan en un descenso (a medida que aumenta el área) muy controlado.

Por otro lado, podemos ver como el área aumenta junto con el porcentaje de construcciones con daño de grado 1, que se ven compensadas por su contraparte (la disminución del porcentaje de daño de grado 3).

Por lo tanto, como conclusión podemos decir que la disminución de la media de grado de destrucción a medida que aumenta el área observada en el primer gráfico esta principalmente dada por las variaciones en las destrucciones de grado 1 y 3 observadas en el segundo gráfico.

2.4. Geografía

Aunque la interacción entre Placas Tectónicas es la principal causa de los sismos no es la única. Cualquier proceso que pueda lograr grandes concentraciones de energía en las rocas puede generar sismos cuyo tamaño dependerá, entre otros factores, de qué tan grande sea la zona de concentración del esfuerzo.

Hay regiones donde la actividad sísmica es casi nula o desconocida lo que pone de manifiesto que el peligro representado por los temblores es muy grande en ciertas regiones y casi nulo o insignificante en otras. Estudiando la distribución de este terremoto podemos determinar los niveles geográficos y los grados de daño que sufrieron.

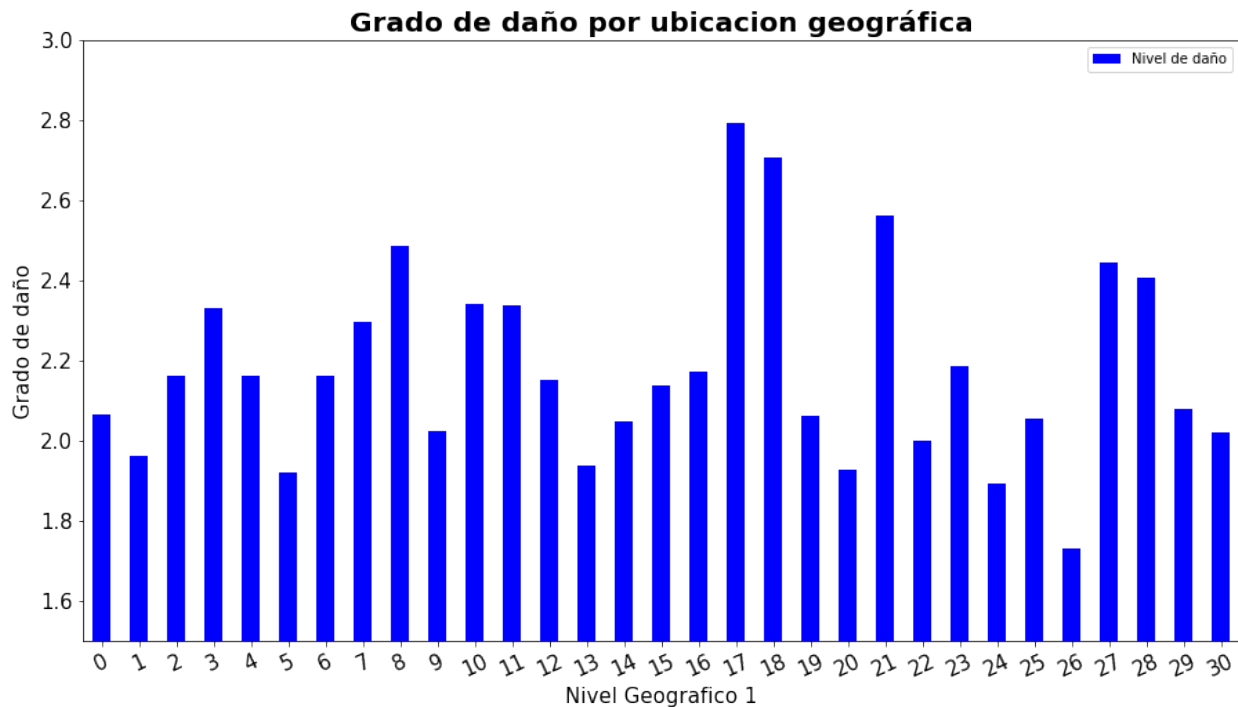


Gráfico de grado de daño por ubicación

El nivel geográfico se divide en 3: 1, 2 y 3, siendo 1 el mas abarcativo y el 3 el mas especifico. Determinamos que para llegar a un veredicto utilizar el nivel geográfico 1 iba a facilitar la deducción y obtención de datos. El nivel geográfico 1 se divide en 30 regiones, determinando el promedio de grado de daño por cada región de este nivel geográfico.

Podemos observar que las regiones mas dañadas son, por un margen considerable, la numero 17 y la numero 18. Luego, les siguen de relativamente cerca la región 7, la 21, la 27 y la 28.

Por el otro lado, podemos ver que han sido poco afectadas regiones como la 1, 5, 13, 20, 22, 24 y 26.

Para poder obtener conclusiones mas precisas acerca de las causas de estas diferencias entre regiones, deberíamos tener acceso a datos relacionados a la geografía de estas, los cuales no están al alcance del set de datos que se

utilizo para este informe.

El daño también se secciona por posición, se puede determinar que sigue habiendo un dominio del grado 2 por sobre todas las posiciones.

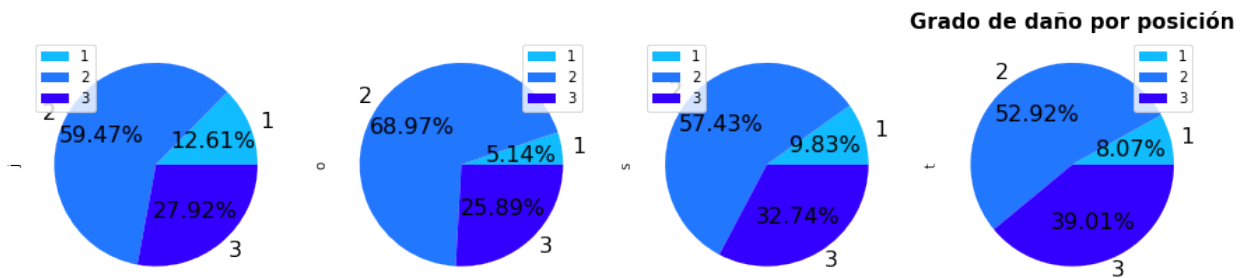


Gráfico de densidad de daño

Se tiene también que la posición que menos daño de grado 3 recibió fue la denominada como 'o', esto se puede traducir en daño mas leve.

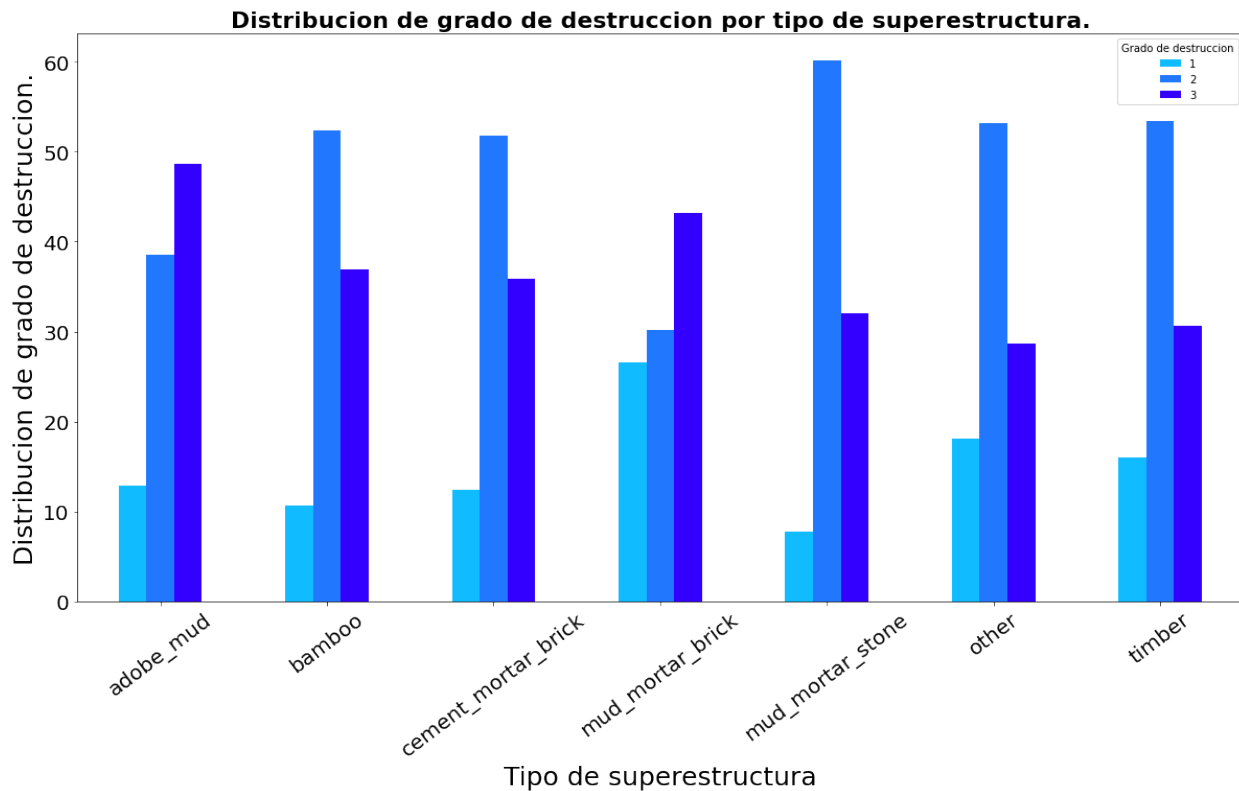
2.5. Materiales

Para esta sección, hemos decidido enfocarnos en los materiales con los cuales se realizó la construcción, para ver si podemos obtener conclusiones acerca de cuales son mas o menos resistentes a este tipo de fenómenos naturales.

Presentamos dos gráficos relacionando las categorías de superestructura mas afectadas con el nivel de destrucción:



En esta primera gráfica podemos observar que la gran mayoría de las construcciones dañadas estaban hechas de mortero de barro. Por supuesto, no podemos por esto decir que es el peor material, ya que puede ocurrir que sea también por mucho el material mas utilizado (y sobre esto no poseemos datos al respecto). En segundo lugar se encuentra la madera, con una diferencia considerable tanto con el primero (el recién mencionado barro) pero también con el resto de los materiales. Tenemos también valores de destrucción significativos para el adobe, bambú y ladrillo, entre otros.



Esta segunda gráfica si puede darnos datos mas útiles al momento de evaluar que material es mas resistente. Relaciona los tipos de superestructura con el porcentaje de estructuras que recibieron cada tipo de daño para cada superestructura. Podemos obtener las siguientes conclusiones:

- El material menos resistente parece ser el adobe: mas del 50 % de las destrucciones de este material son de grado 3. El ladrillo tampoco parece ser un gran material, aunque compensa con una buena proporción de destrucciones de grado 1.
- A diferencia de lo que podríamos haber pensado mirando solo el primer gráfico, el mortero de barro tiene buena resistencia y presenta uno de los valores mas bajos de destrucción de grado 3 junto con la madera y otros.

Como conclusión podemos incitar al lector a no quedarse con el primer gráfico que se le entregue en mano, y si es lo único que se tiene, analizarlo detalla-

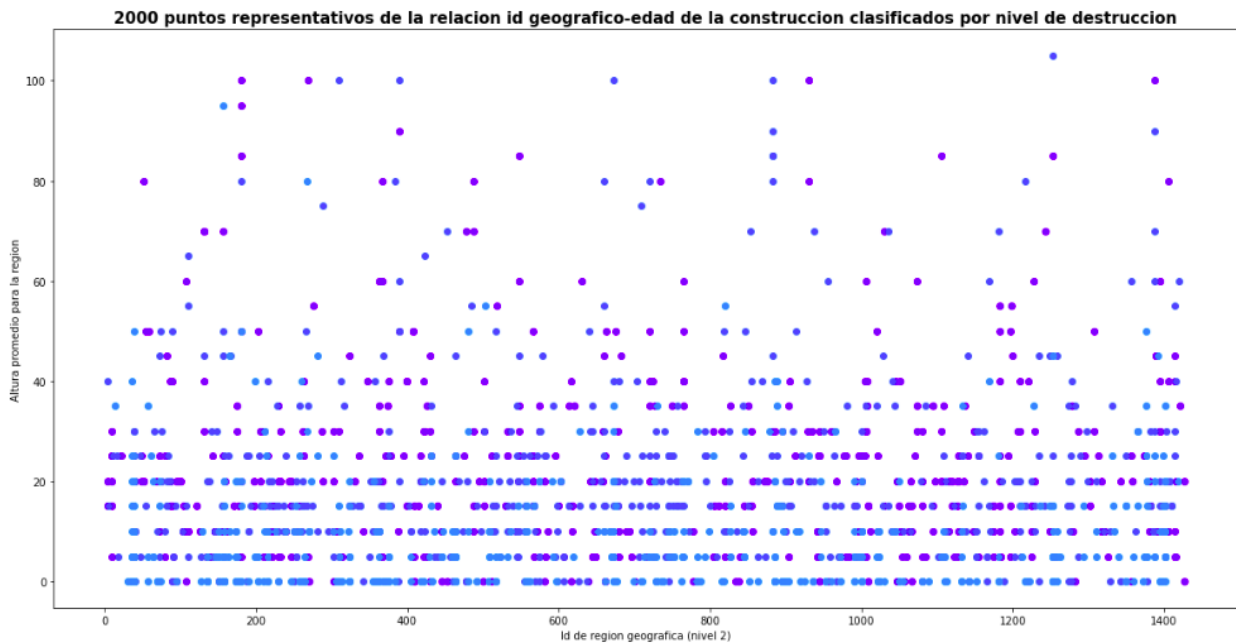
damente, y no sacar conclusiones rápidamente: observar exactamente cual es el dato que nos están dando y comprenderlo.

2.6. Algunos gráficos mas avanzados

En esta ultima sección del desarrollo de este informe, le dedicaremos un espacio a algunos gráficos mas avanzados.

2.6.1. Concentración de Alturas

El primero de ellos es un "scatter plot"(gráfico de dispersión):



En este podemos observar 2000 muestras del conjunto de datos obtenidas al azar. El gráfico nos indica, por cada una, el id de la región (eje x), la altura de la edificación (eje y) y el grado de daño para la edificación en cuestión representado por el color (celeste = 1, azul = 2, violeta = 3).

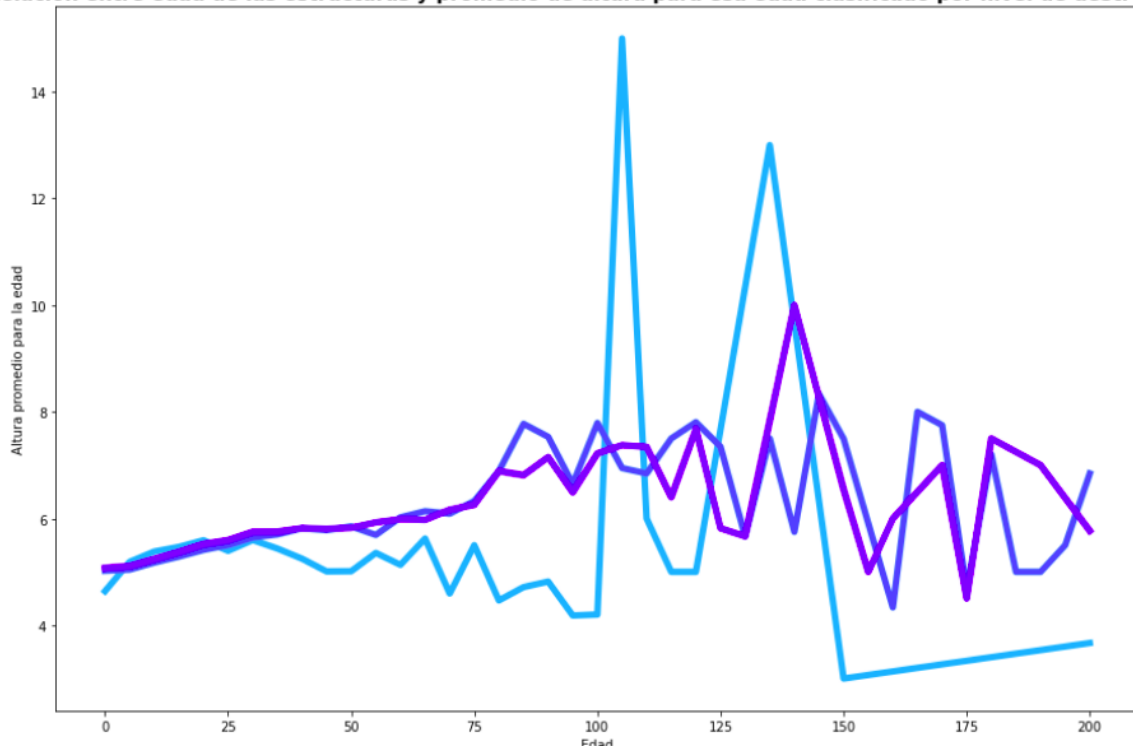
Este es uno de los gráficos que consideramos fallidos, ya que no presenta ninguna correlación entre los distintos valores. Esto es porque, según enten-

demos, no hay ninguna relación entre los valores de id de la región cercanos entre si. Por lo tanto, aunque pueda existir algún patrón entre algunas regiones, no lo podemos observar en este gráfico ya que las tenemos ordenadas por id, el cual es un valor irrelevante.

La única conclusión que se puede obtener de la dispersión de puntos es que el conjunto de datos posee muchos mas registros para edificaciones de poca antigüedad, y por lo tanto ha habido mas construcciones de pocos años de vida que han sido dañadas. Sin embargo, este análisis ya se ha realizado mas exhaustivamente arriba, y esta conclusión puede estar sesgada por el hecho de que, por ejemplo (y esto es solo una hipótesis) en los últimos años se construyo mas.

2.6.2. Daño en Función de la Edad y Altura

Luego tenemos un segundo gráfico:

Relacion entre edad de las estructuras y promedio de altura para esa edad clasificado por nivel de destruccion.

Este nos muestra la relación entre la edad de las estructuras (eje x) y la altura promedio para la edad en cuestión (eje y) en tres funciones, cada una representando un grado de daño y diferenciadas por color (celeste = 1, azul = 2, violeta = 3).

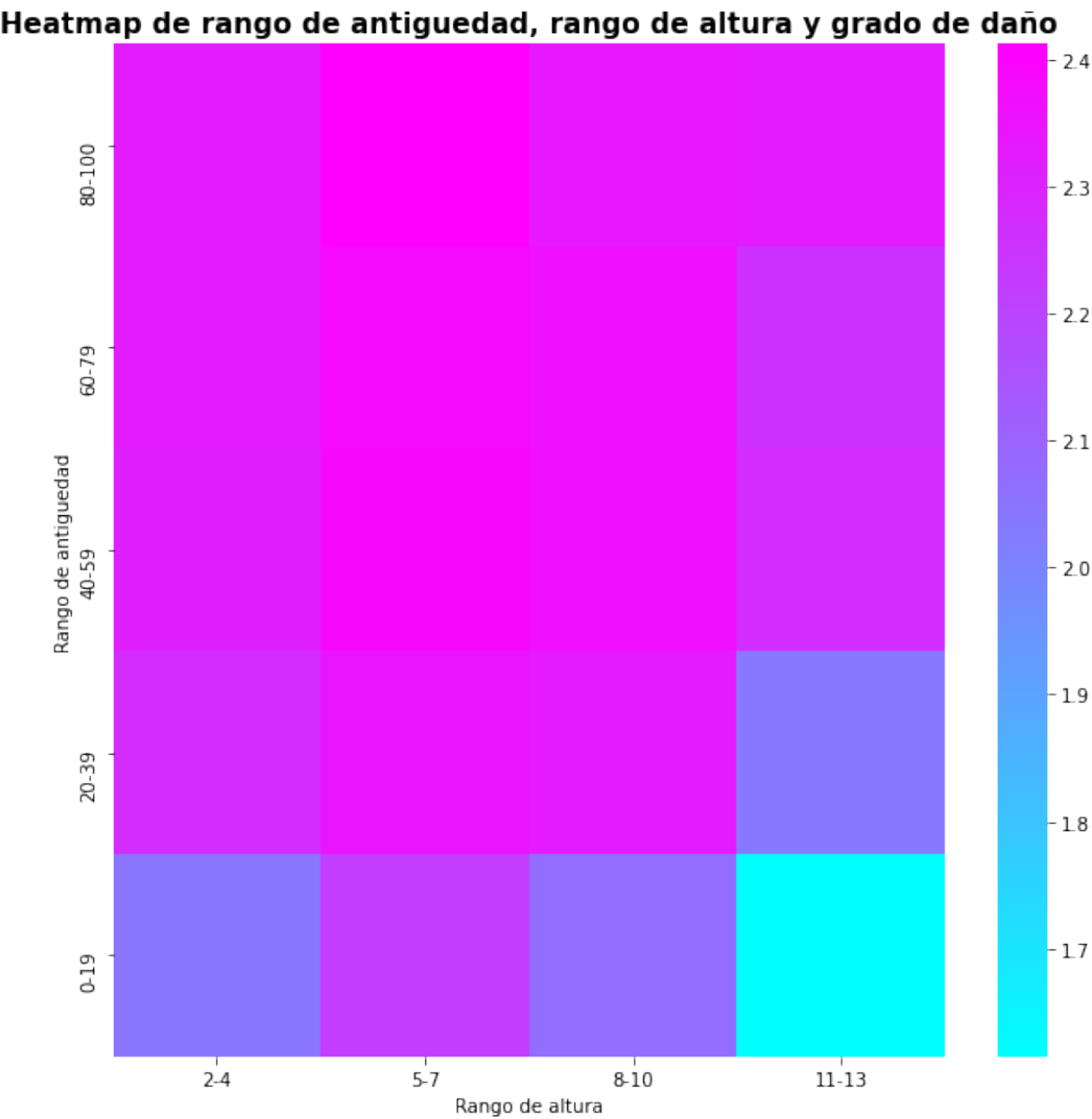
Podemos obtener varias conclusiones:

- Por un lado, la correlación entre las gráficas de daño de grado 2 y daño de grado 3, parece ser mucho mayor que la correlación de cada una de estas con la gráfica de grado 1.
- Por otro lado, la gráfica de grado 1 presenta mas oscilaciones ya que, si hablamos de altura promedio, podemos observar que el rango de esta contiene al de las otras dos.
- Mientras que las gráficas de grado 2 y 3 tienen una mínima tendencia al alza (es decir, cuanto mas antigüedad tiene una estructura, mas altura tiene), para la gráfica de grado 1 (excepto por los picos producidos

aproximadamente en 105 y 135 años de antigüedad) se da lo contrario: tiene una tendencia a la baja.

2.6.3. Altura y Edad

Finalmente podemos observar un "heatmap"(mapa de calor):



Este gráfico relaciona cinco conjuntos de construcciones clasificadas por rangos de antigüedad (de 20 años de tamaño cada uno), con cuatro conjuntos de construcciones clasificadas por rangos de altura (de 3 unidades cada uno) para formar un total de veinte conjuntos realizando las intersecciones de cada uno de los conjuntos clasificados por antigüedad, con cada uno de los conjuntos clasificados por altura.

Para cada conjunto de estos veinte, se calculo el grado de destrucción promedio y se representaron los valores en el mapa de calor visto arriba.

Las conclusiones son las siguientes:

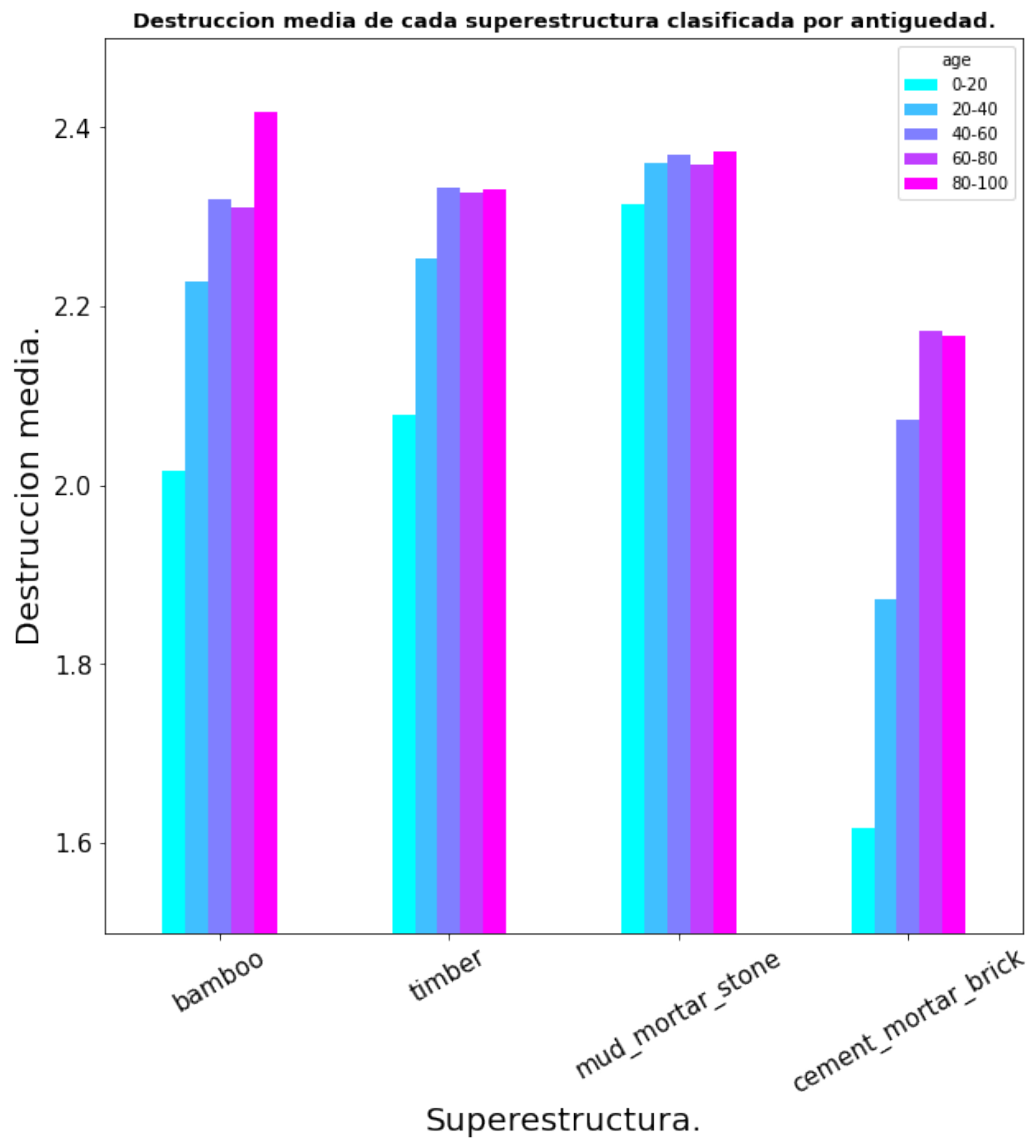
- Podemos ver, por un lado, que los edificios que a su vez son altos y modernos (pocos años de antigüedad) presentan los valores mas bajos de destrucción.
- También podemos observar que en esto ultimo influye principalmente la antigüedad (como habíamos visto anteriormente en el análisis), ya que toda la columna del rango de 0-19 años presenta un grado de destrucción bajo.
- El rango de altura 5-7 presenta los mayores daños, y se ve aun mas afectado cuando la construcción es de gran antigüedad.
- La evolución de la media de grado de destrucción según antigüedad y altura parece presentar una buena correlación, ya que en el mapa de calor no se ven puntos rosas rodeados de puntos celestes o viceversa.

2.6.4. Destrucción por Material en Función de la Edad

En el próximo gráfico podremos visualizar el daño medio (eje y) de cada grupo de estructuras clasificadas por superestructura (eje x) y por antigüedad (colores, donde el rango que representan esta indicado arriba a la izquierda en el gráfico).

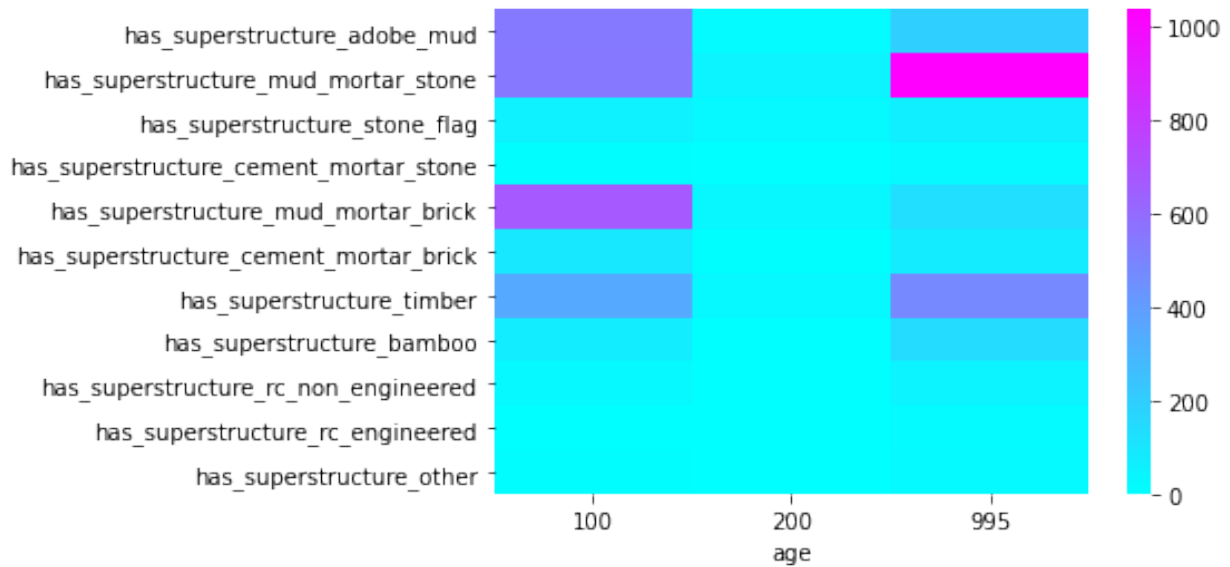
Podemos obtener varias conclusiones al respecto:

- El ladrillo con cemento (cuarta columna) es el material mas resistente de los cuatro analizados. Sin embargo, podemos notar que se degrada mas rápido que otros materiales: para estructuras con este material de menos de 20 años de antigüedad tenemos un mínimo de media de daño de 1.6, mientras que para estructuras del mismo material de entre 80-100 años, el daño medio alcanza el valor de aproximadamente 2.2.
- El bambú (primera columna) parece mostrar un comportamiento similar al ladrillo con cemento, aunque con dos diferencias: por un lado, el piso de daño medio es mas alto, ya que para el rango de 0-20 años de antigüedad el valor es de 2.0; por el otro, mientras que en el ladrillo con cemento, el aumento de la media del daño parece ir disminuyendo, en el caso del bambú se asemeja mas a ser constante.
- Para el caso de la madera (segunda columna) podemos decir que la degradación se produce principalmente en los primeros años (0-40) ya que luego de este periodo podemos observar que el grado de daño permanece casi constante, con oscilaciones despreciables.
- Finalmente, el mortero de piedra y barro (tercera columna) es el que menos degradación parece presentar a lo largo del tiempo, ya que la variación del daño medio para estructuras en un rango de 100 años no llega a 0.1 (y hemos visto casos con mucha mayor variación). Sin embargo, debemos tener en cuenta que arranca de un piso mucho mas alto, ya que para estructuras modernas (0-20 años) ya presenta un daño medio de 2.3.



2.6.5. Materiales más Viejos

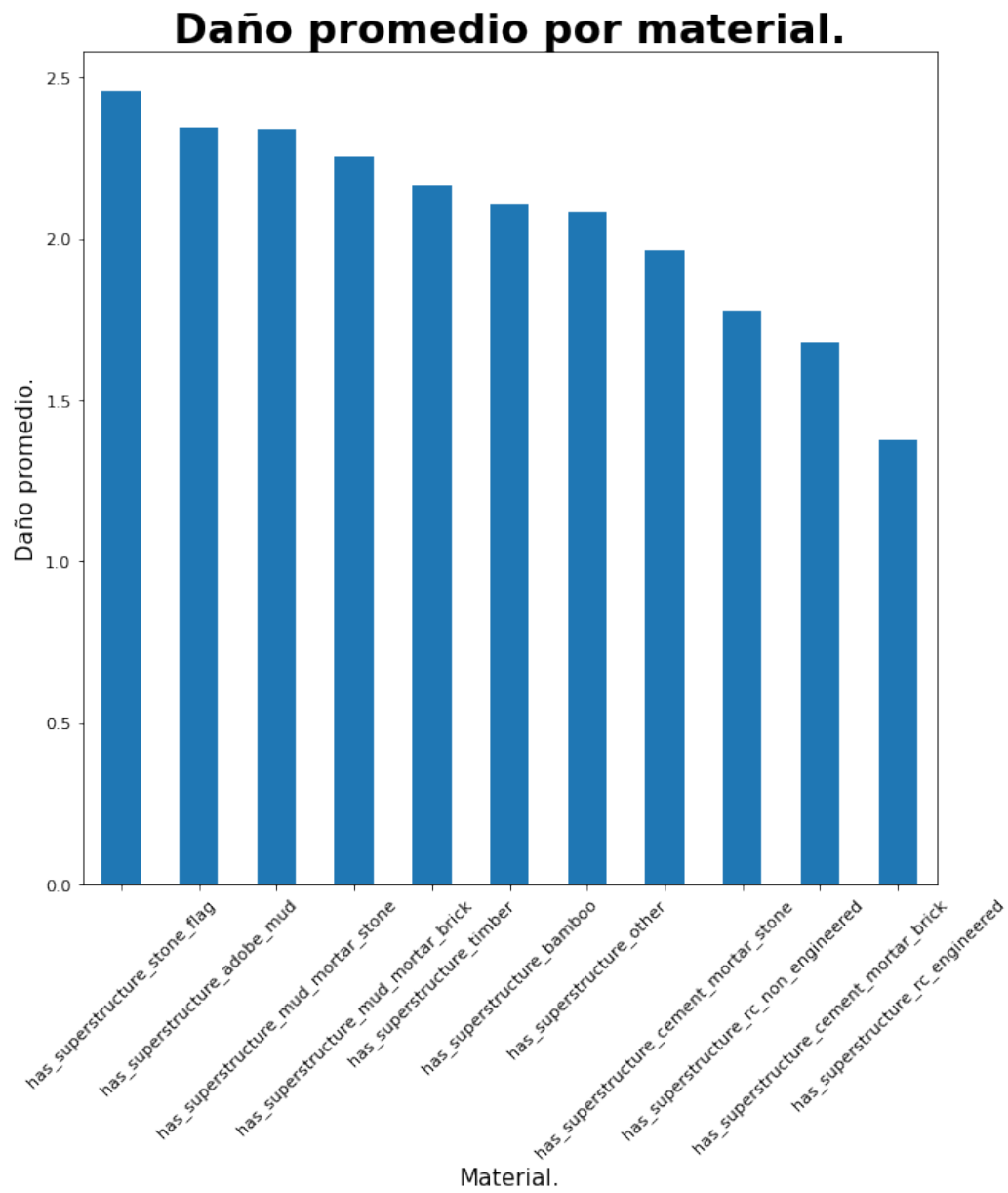
En el próximo mapa de calor:



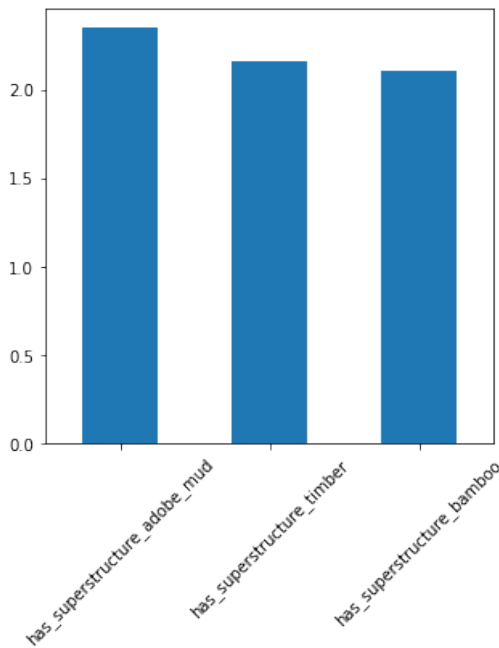
Entendemos que, cuanto más viejo es una estructura, más tiende a estar construido con: adobe de lodo, piedras de lodo y mortero (particularmente cuando la construcción es milenaria), ladrillo de lodo y mortero y timber.

2.6.6. Daño Promedio

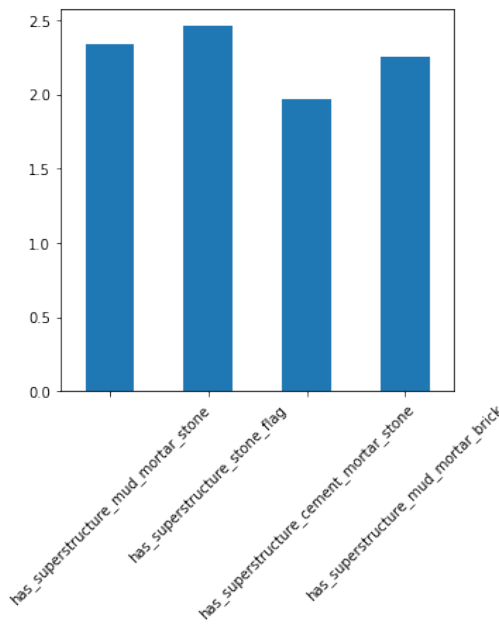
Podemos visualizar como, en función del daño promedio, el peor material de construcción para tener en una estructura es la piedra y los mejores son los que utilizan concreto y cemento.



Además podemos ver como se comparan los datos en función de si los materiales son orgánicos:



E inorgánicos:



3. Conclusiones

A lo largo de este informe se explicaron diversos temas acerca del terremoto en Nepal del 25 de abril de 2015.

Desde principio a fin se han llegado a diversas conclusiones, tanto de la temática de análisis como en la esencia de lo que es hacer un análisis exploratorio en si.

Luego de parsear la data y verificar que esta esté preparada. Se paso a verificar que los tipos de datos sean válidos y no existan datos nulos. Además se debe minimizar el tamaño que ocupa el DataFrame en caché, así los algoritmos tardan menos en ejecutarse y podemos expandir el rango de máquinas que poseen los requerimientos mínimos para trabajar con el set de datos.

3.1. El terremoto y sus consecuencias

Después de diversos análisis se pueden llegar a varias conclusiones en lo que al terremoto respecta. En primer lugar, las zonas afectadas en su mayoría sufrieron un grado de daño de nivel 2, dejando solo una minoría de establecimientos con daños de grado 1.

En cuanto a lo social, aquí también se vio reflejado esta ponderación, con un 56,78 % de las familias que sufrieron el grado de daño 2 en sus establecimientos.

La agricultura fue la que mas sufrió en cuanto a lo que daño causado por el terremoto respecta.

3.2. Edificaciones

Las edificaciones cumplen un rol principal en este informe, ya que de ellas se pueden tomar diversos datos que apoyan a un análisis en profundidad del terremoto ocurrido.

La edad de la estructura en su mayoría determina cual es el daño que recibe, las que tienen mas de 25 años tienen en porcentaje una mayor cantidad de casos con grado de daño 2, mientras que las que no superan esta cifra su media de destrucción oscila valores inferiores.

Por otro lado, la altura que la edificación tenga influye en el grado de daño que esta recibe. A partir de los 20 metros de altura los daños se disparan de manera exponencial, lo cual tiende a disminuir en el rango de estructuras de 5 a 20 metros de altura. Esto se ve ligado a la cantidad de pisos que estos tengan, siendo 3 pisos el caso con mayor promedio de daño. También, la altura de estos pisos tiende a ponderar el daño por encima de 2, 10 cuando estos miden entre 1, 5 y 3 metros de altura.

Se puede determinar que el área de la estructura también influye en el deterioro que esta sufre. Es decir, a mayor área menor sera el grado de destrucción.

3.3. Ubicación Geográfica

El análisis de estos datos dejan como resultado que la zona geográfica 17 y 18 tienen el promedio mas alto de daño registrado.

El daño por posición deja en evidencia a la zona "O" como la que mas daño de grado 2 sufrió.

3.4. Materiales

Se analizaron los diversos casos ocurridos y los materiales con los que están construidas las estructuras. Las construcciones que están compuestas de roca son las mas afectadas por este desastre natural, mientras que las que están hechas de lodo de adobe y de ladrillo sufrieron en su mayoría daños de grado 3.

Los ladrillos, según su especie, hacen variar la posibilidad de recibir daños de grado 2 y 3, siendo los de cemento el que disminuye la probabilidad de recibir daños de grado 2. Para el resto de los tipos, los daños son en su mayoría de grado 3.

El concreto se ve ausente en poco menos del 80 % de las edificaciones que reportan los daños de grado 1 y se ve presente en la gran mayoría de las que sufrieron grado de daño 3, pudiendo concluir que es un material que no debería usarse en las construcciones y debiendo reemplazarse por otros que ayuden a minimizar los daños causados por un posible desastre natural similar en el futuro.

Además, la edad de estos elementos cumple un papel principal en la resistencia de este, ya que mientras mas se avejenta la estructura su material se vuelve mas frágil, lo que lo hace propenso a subir la probabilidad de una destrucción de grado 2 o 3.

4. Repositorio

El código fuente del trabajo se encuentra en el [siguiente repositorio](#).