

Movilidad Medellín: Accidentalidad en 2016

Daniel Montoya CC: 1.128.414.907

21 de marzo de 2018

1. Resumen

El presente trabajo tuvo como objetivo acercarse a los datos de movilidad disponibles en Medellín, para aportar en conocimiento que sirva para tomar acciones que mitiguen los problemas de accidentalidad de la ciudad. Se realizaron análisis descriptivos sobre accidentalidad, y análisis bivariados cruzando variables de interés. Se encuentran resultados respecto a qué días, horas, barrios y comunas se relacionan con cantidades de accidentes y la gravedad de dichos accidentes. Se incluye un análisis con geolocalización con la gravedad y cantidad de accidentes en Medellín. En cada sección se dan recomendaciones con base en la información obtenida.

2. Introducción

Un reto que tienen las ciudades es la movilidad de la gente que las habita[1]. Gran parte de esta movilidad se da en vehículos motorizados. Un problema que enfrenta constantemente la movilidad en vehículos es el de accidentalidad, pues en los accidentes se presentan diversas consecuencias negativas como daños materiales, heridas a humanos o animales, y muertes, además en general los accidentes generan obstaculización de las vías, con lo cual se incrementan las dificultades para movilizarse. Por esto, tomar acciones que busquen reducir la accidentalidad es necesario, y estas decisiones deben tomarse con base en conocimiento del fenómeno [2].

En este documento se analizan los datos de accidentalidad en Medellín en 2016. se comienza por un análisis exploratorio, y luego se hacen análisis bivariados. Las preguntas que se buscan responder son: Hay días, horas, y lugares que se relacionan con mayores tasas de accidentalidad y mortalidad?. En cada sección se anotan observaciones que pueden ser de utilidad para las autoridades de la ciudad.

3. Análisis de resultados

3.1 Análisis exploratorio

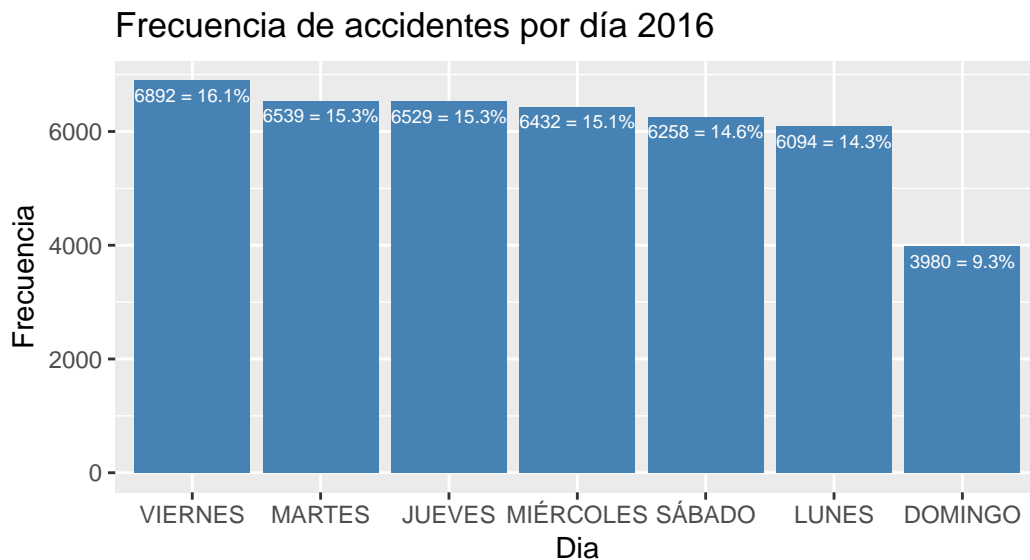
Primero comenzaremos por conocer la base de datos y cómo se distribuyen cada uno de los datos en ella. La base de datos original tiene 42841 entradas, es decir, esta es la

cantidad de accidentes registrados en el año. Esto implica un promedio de 117.3 accidentes diarios en la ciudad. Para cada accidente se tienen las variables: Localización geográfica (coordenadas latitud y longitud), Radicado, Fecha, Hora, Día, Clase, Dirección, CBML, Tipo de geocodificación, Gravedad, Barrio, Comuna, y Diseño.

De estas variables no son de particular interés el Radicado, CBML, tipo de geocodificación, y dirección, pues Radicado, tipo de geocodificación y CBML son códigos administrativos de identificación del accidente. La dirección se hace poco importante al contar con las coordenadas geográficas de cada accidente.

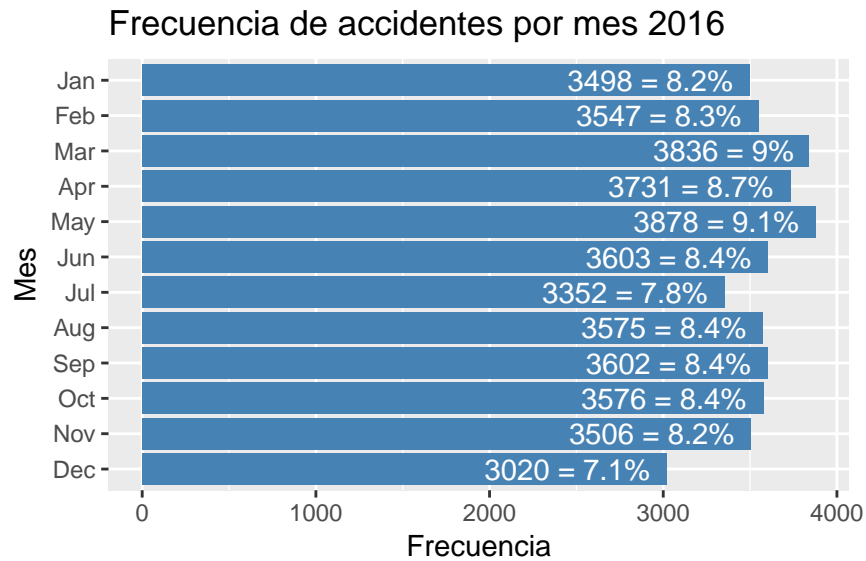
Comenzaremos por un análisis exploratorio de la distribución de las demás variables.

3.1.1 Variable: Día



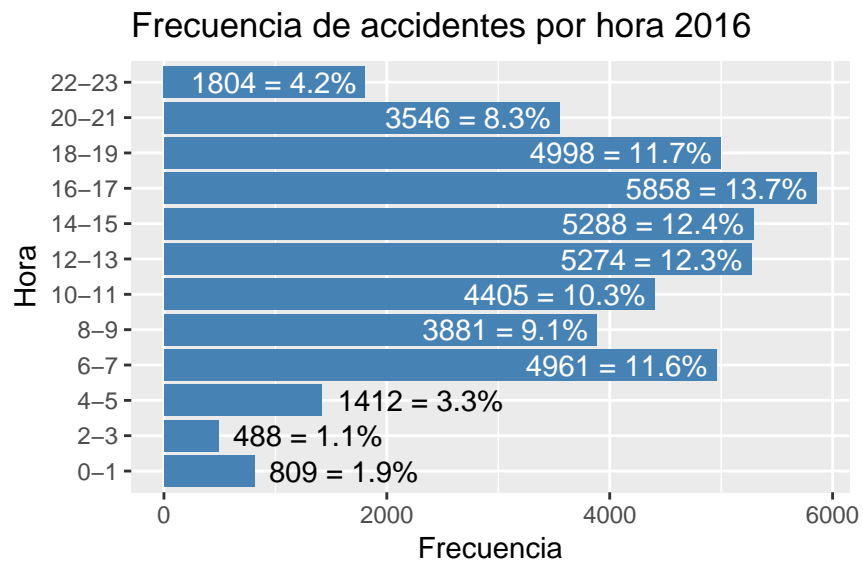
Vemos que el viernes presenta mayor accidentalidad, aunque la frecuencia parece estar distribuida de manera uniforme, excepto por el domingo, día en el cual se presenta menor accidentalidad. Esta diferencia es significativa, como lo respalda la prueba Chi cuadrado ($\chi^2 = 923.02$, $df=6$, $p<0.001$). Este se puede considerar como un resultado esperado, pues tradicionalmente el domingo es un día de descanso, donde menos personas se desplazan a lugares de estudio o trabajo.

3.1.1 Variable: Mes



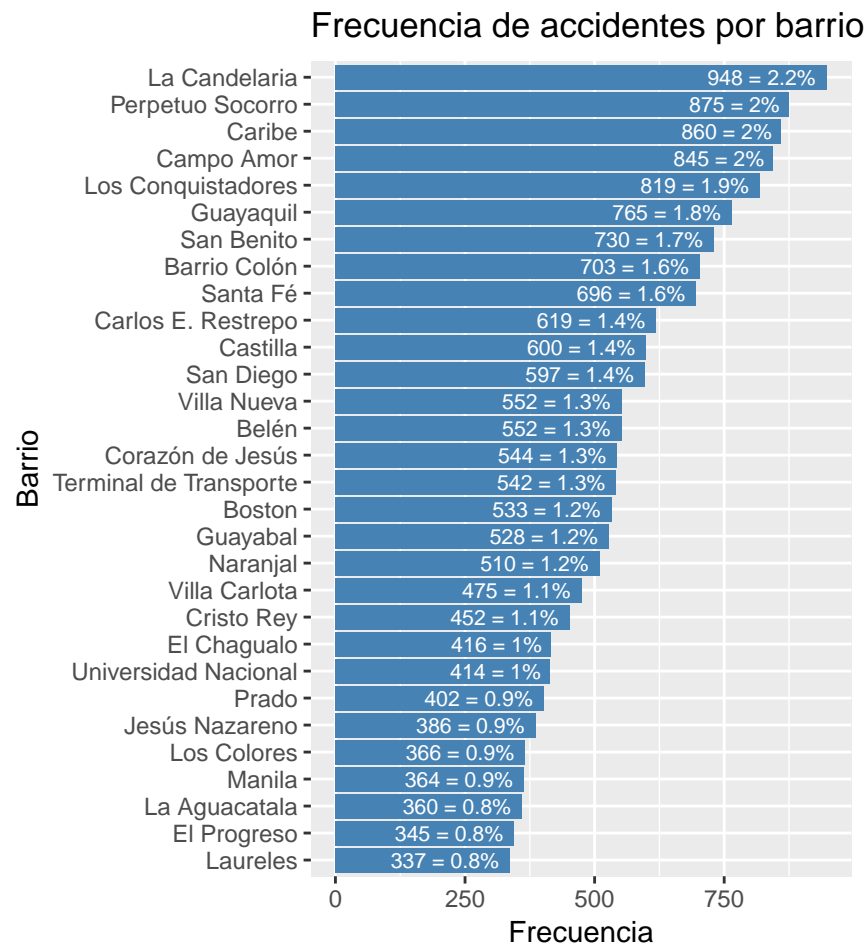
Vemos que la frecuencia por mes presenta también una distribución aproximadamente uniforme, con picos altos en mayo y marzo, y picos bajos en julio y diciembre. Esto puede estar relacionado también con el hecho de que diciembre y julio son tradicionalmente meses de vacaciones, con menos personas yendo a estudiar. Lo que parece menos esperable son los picos altos, sin embargo puede que dichas fluctuaciones se deban simplemente al azar. Esta hipótesis se descarta debido a que al aplicar la prueba Chi cuadrado, se encuentra que esta alcanza significancia ($\chi^2 = 155.16$, $df=11$, $p<0.001$). Por tanto las diferencias entre meses no se deben simplemente al azar. Aun excluyendo los meses de Diciembre y Julio, la prueba Chi cuadrado alcanza significancia ($\chi^2 = 44.293$, $df=9$, $p<0.001$). Por tanto entre los otros meses las fluctuaciones no parecen ser debidas solo al azar.

3.1.3 Variable: Hora



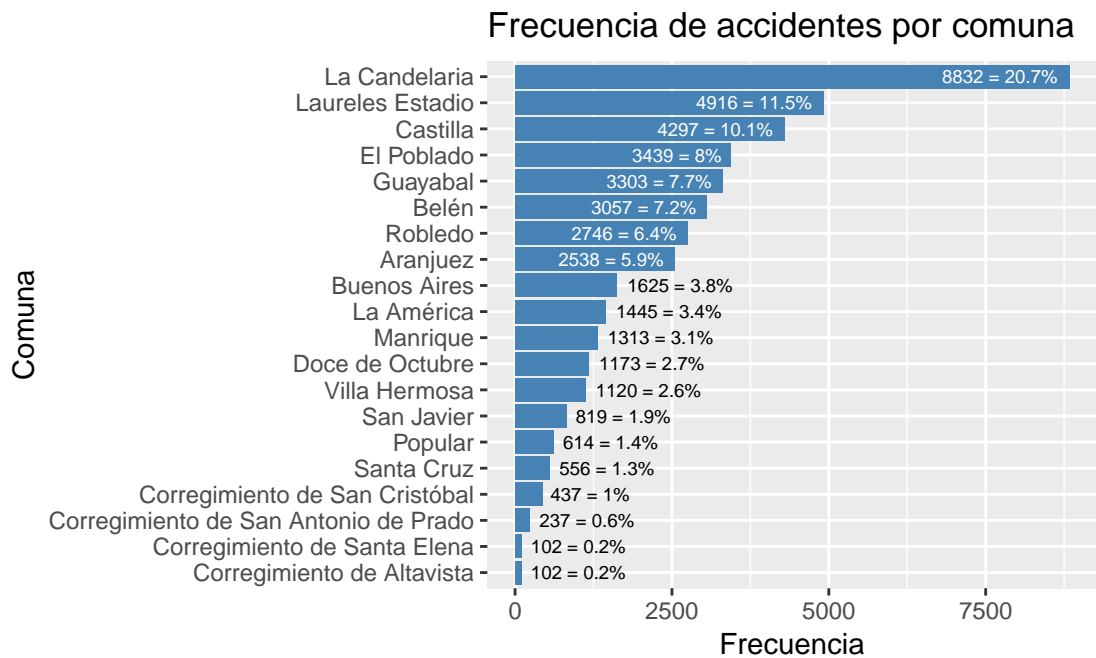
Se puede apreciar que las horas de mayor accidentalidad son las de la tarde, con un máximo entre 4 y 6 p.m.. Se observa un segundo pico entre 6 y 7 a.m., se puede decir que el resultado es esperado, pues estos máximos corresponden a las horas pico.

3.1.4 Variable: Barrio



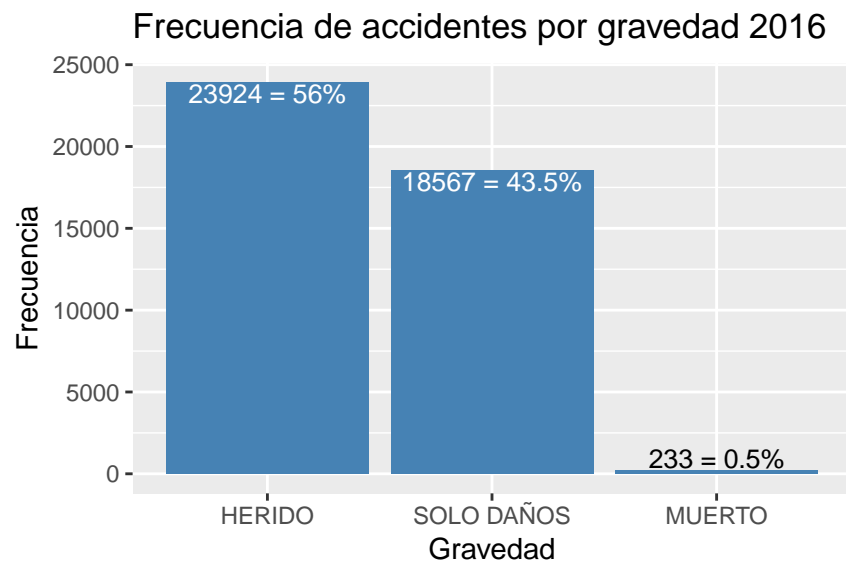
Como Medellín tiene 312 barrios, no es conveniente mostrarlos todos en un gráfico, por tanto se seleccionaron los 30 barrios con mayor accidentalidad. En primer lugar se encuentra el barrio La Candelaria, que corresponde al centro de Medellín. Luego se encuentra el barrio Perpetuo Socorro que corresponde a la zona entre las calles 30 y 33, y la avenida oriental y la avenida del río. El barrio Caribe es el que comprende entre la Universidad Nacional y la terminal de transporte del norte.

3.1.5 Variable: Comuna



Vemos que la comuna con mayor accidentalidad es La Candelaria, como era de esperarse, pues corresponde a las zonas alrededor del centro de Medellín. A continuación se encuentran Laureles, Castilla y El Poblado. Las comunas con menor accidentalidad son los corregimientos de la ciudad.

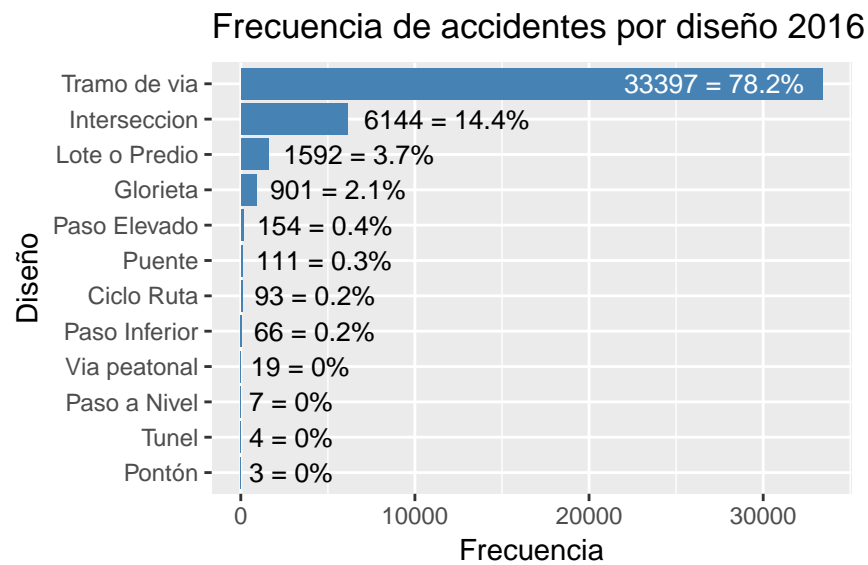
3.1.6 Variable: Gravedad



Este gráfico nos indica que la gran mayoría de accidentes tienen como consecuencia daños

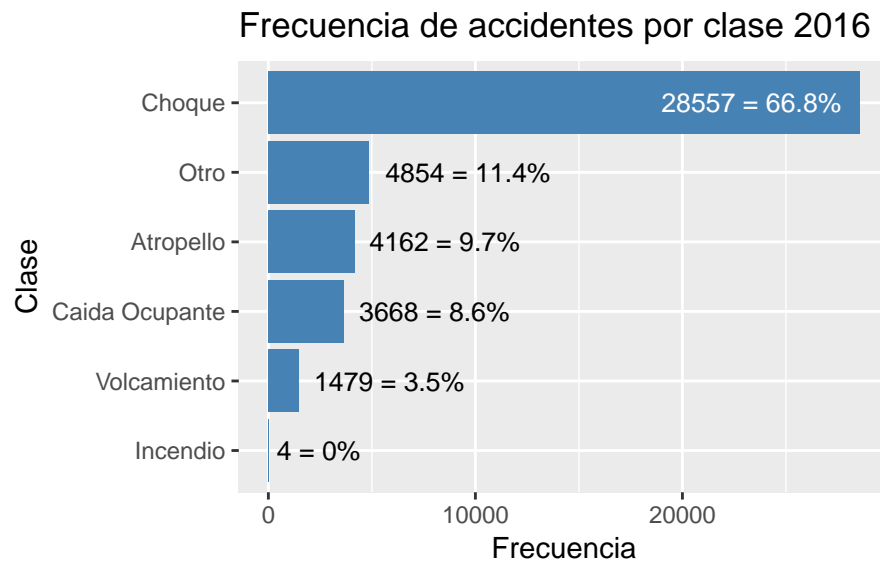
materiales, y heridas, mientras que la muerte es una consecuencia que se da en alrededor del 0.5% de los casos. Sin embargo, por ser la consecuencia de mayor gravedad, esta variable es importante en el análisis. Igualmente importante es destacar que la mayoría de accidentes presentan heridos, por lo cual se debe indagar por la causa de las heridas en los accidentes, para saber si es posible mejorar la seguridad de vehículos o infraestructura vial. Adicionalmente, sería interesante conocer la cantidad de muertos, heridos, y un estimativo del valor de los costos de daños en los accidentes para hacer análisis con estos datos también.

3.1.7 Variable: Diseño



Esta variable se refiere al sitio de la vía donde ocurrió el incidente. Vemos que ésta no presenta mucha variabilidad, pues el valor más común abarca cerca del 80% del total de los datos.

3.1.8 Variable: Clase



Vemos que la variable “clase” se refiere al tipo de accidente, en esta la mayoría son choque, seguido de la clase “Otra clase”. Esta variable podría ser más útil si se descompusiera la categoría “otra clase” en varias de sus subclases pues esta es la segunda en frecuencia.

3.1.9 Variable: Geolocalización

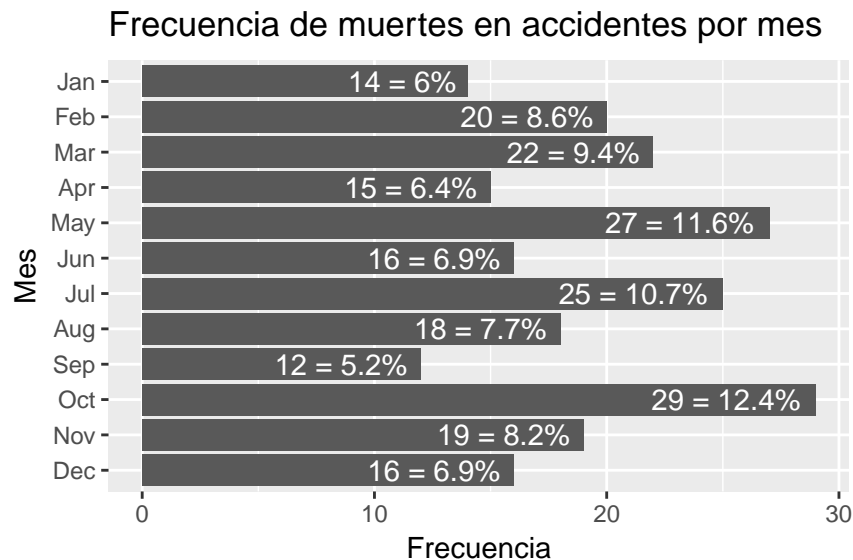
Para esta variable, en este link: [<http://rpubs.com/danielm322/372929>] se podrá encontrar un mapa interactivo con clustering geolocalizado para conocer las localizaciones de accidentes con muertes y accidentes de heridos y solo daños.

En estos mapas, se evidencia que gran cantidad de accidentes con muertos se dan en las vías regional sur y norte, cerca a las zonas centrales. Así mismo, el centro es un lugar de alta densidad de accidentes con muertes.

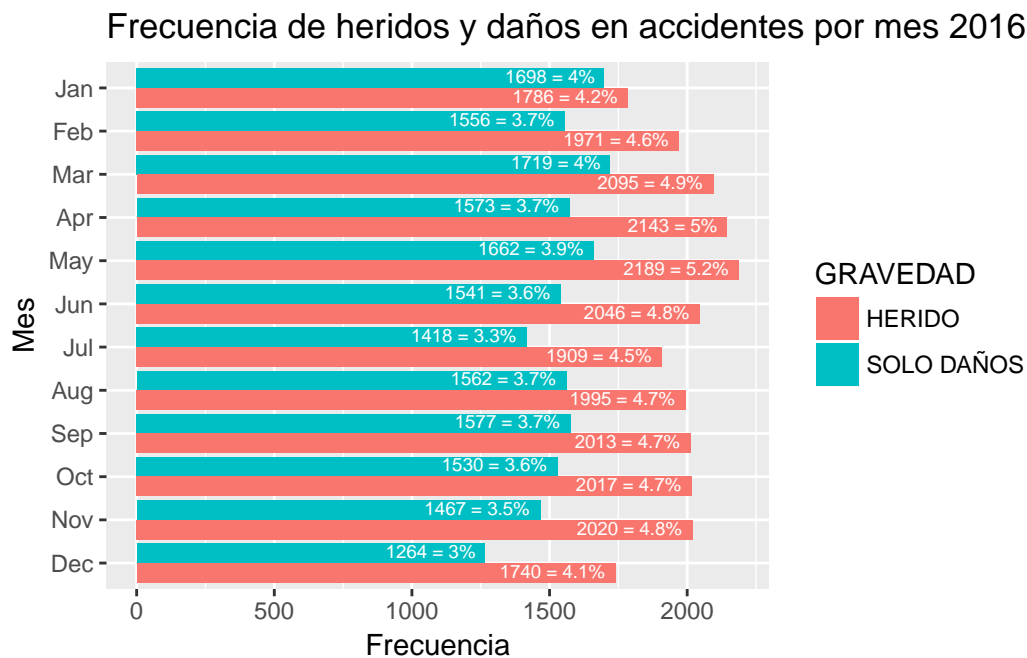
3.2 Análisis bivariado

En esta sección se cruzan variables para conocer con mayor detalle el comportamiento de la variable gravedad del accidente según meses, días, horas, y comunas.

3.2.1 Gravedad por mes



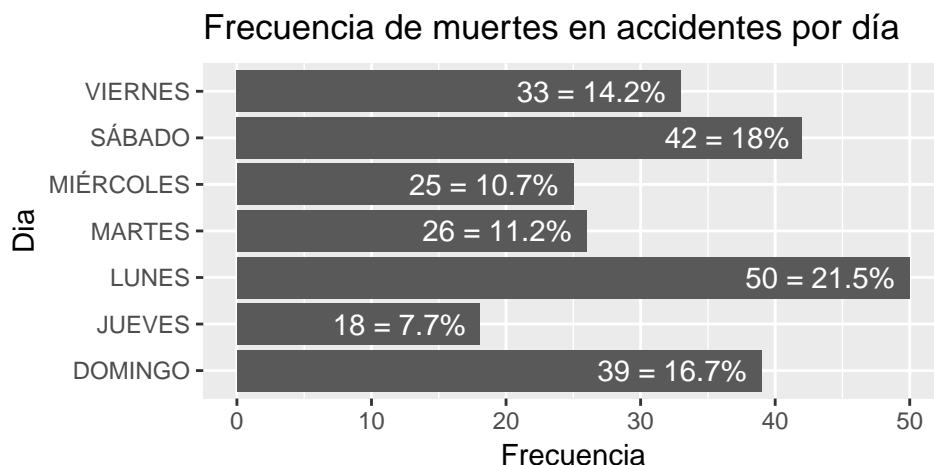
A pesar de que la distribución de muertes por mes a simple vista no parece uniforme, las fluctuaciones parecen ser debidas al azar, pues al comparar esta gráfica con las de distribuciones uniformes generadas aleatoriamente, se encuentra que son muy similares. Adicionalmente, al aplicar la prueba Chi cuadrado se observa que la prueba no alcanza significancia ($\chi^2 = 16.322$, $df=11$, $p=0.1296$) Por tanto no es posible concluir que haya meses que presenten diferencias significativas con otros meses en cuanto a esta variable.



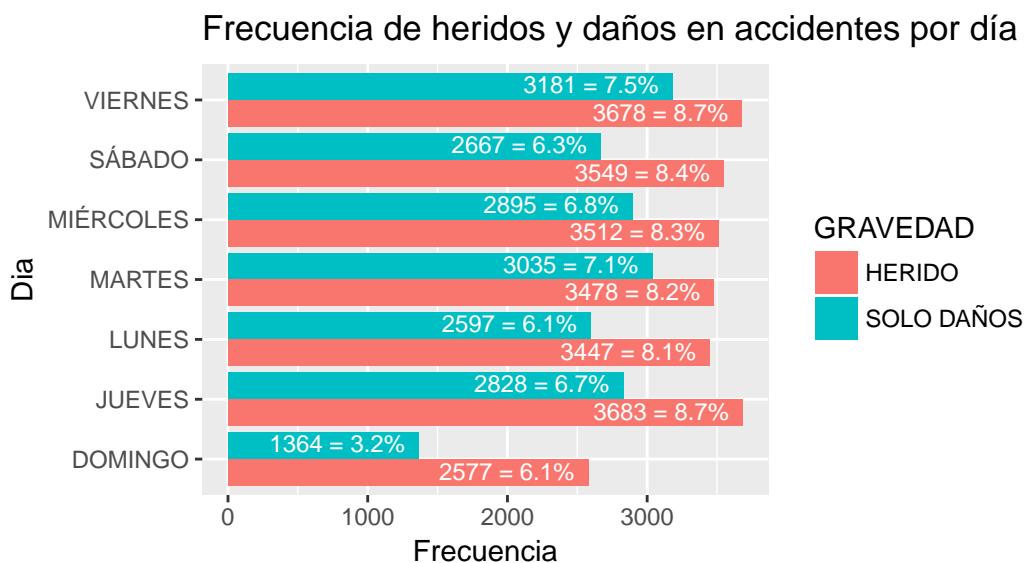
La variable de accidentes con herido y solo daños materiales tambien parece seguir una distribucion uniforme, excepto por el mes de diciembre. En este mes se presentan menores

accidentes de estos tipos. Las diferencias son significativas, como es indicado por el resultado de la prueba Chi cuadrado de Pearson ($\chi^2 = 52.337$, $df=11$, $p<2.3e-7$). Es posible que se relacione que con el hecho de que es un mes de vacaciones escolares. Cruzando esta informacion con la de otras bases de datos, como la de aforos, puede saberse si en este mes tambien hay menos flujo vehicular, con lo cual se apoyaría la validez de esta hipótesis.

3.2.2 Gravedad por día



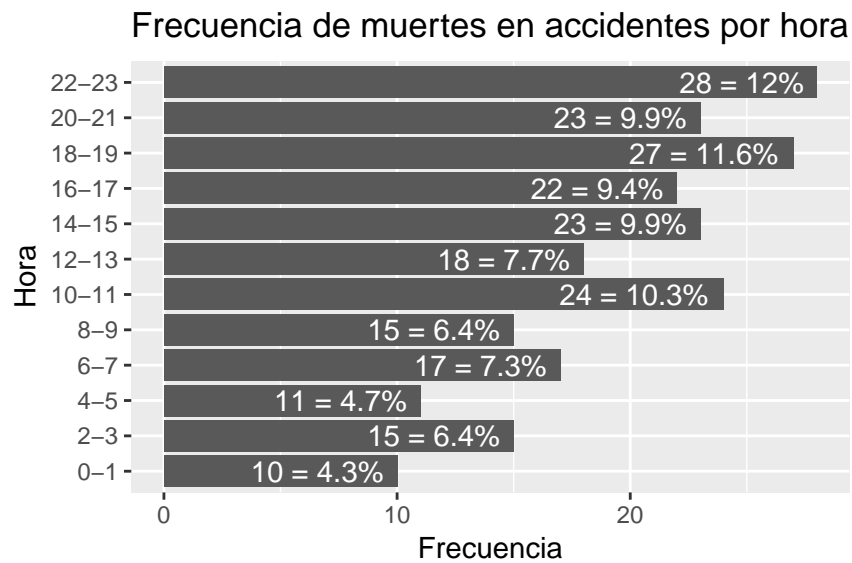
Las frecuencias de muertes por día no parecen seguir una distribucion uniforme, con un máximo el lunes. Las diferencias son significativas, segun indicado por la prueba Chi cuadrado ($\chi^2 = 22.335$, $df=6$, $p=0.001$). Sin embargo el nivel de significancia no es muy alto, por lo que no parece sensato descartar que las diferencias entre días se deban al azar.



En Diciembre se presentan menores frecuencias de accidentes con muertos y daños. Se encuentran diferencias significativas entre meses según prueba Chi cuadrado ($\chi^2 = 183.41$, $df=6$, $p<2.2e-16$). Por tanto no se pueden considerar como debidas al azar las diferencias

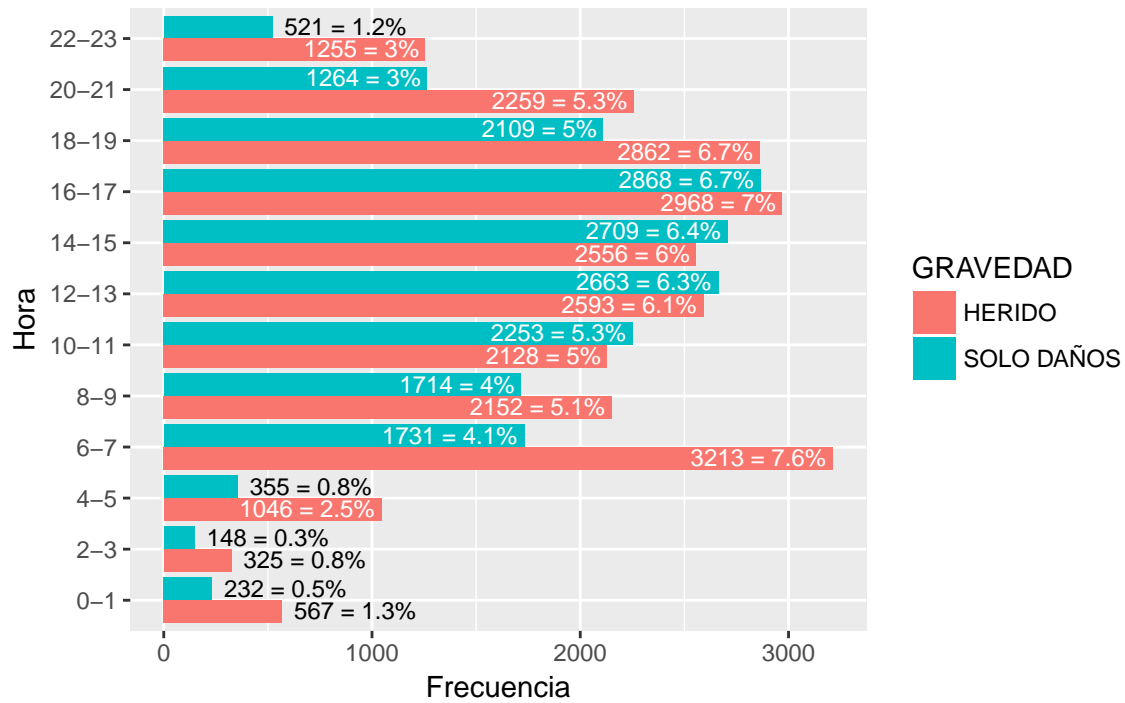
entre meses. De nuevo, es importante destacar que las frecuencias de accidentes con heridos son mayores que las de accidentes con daños materiales, lo cual invita a indagar por las causas de heridas en dichos accidentes, para mejorar la seguridad de los vehículos, o infraestructura vial.

3.2.3 Gravedad por hora



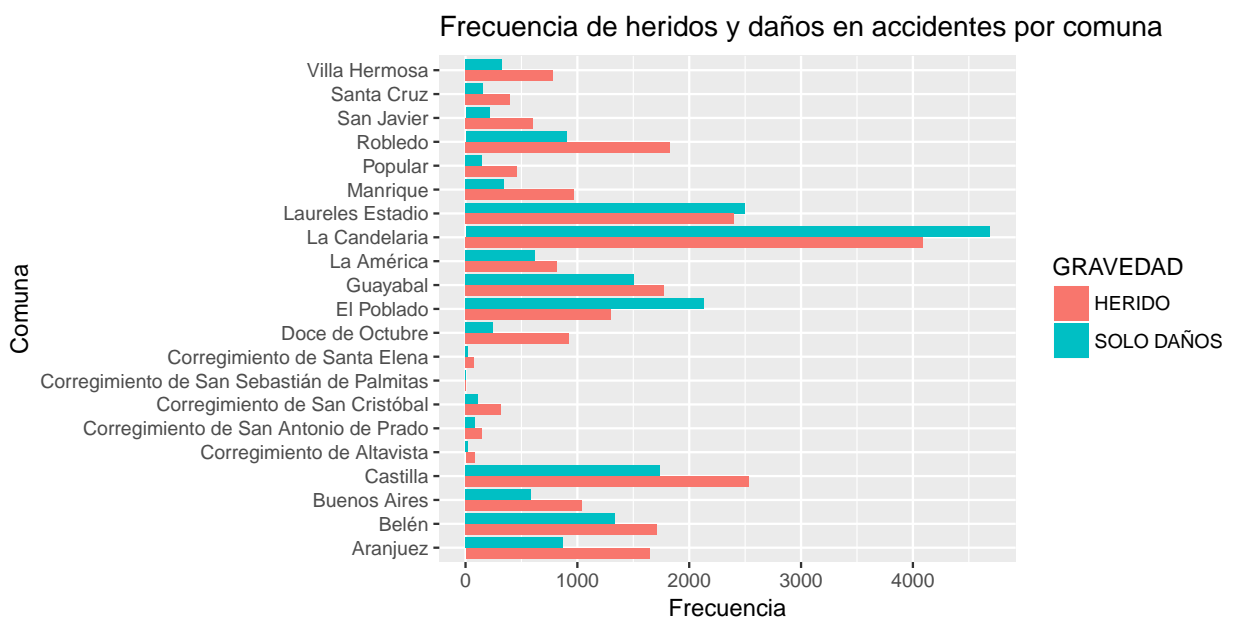
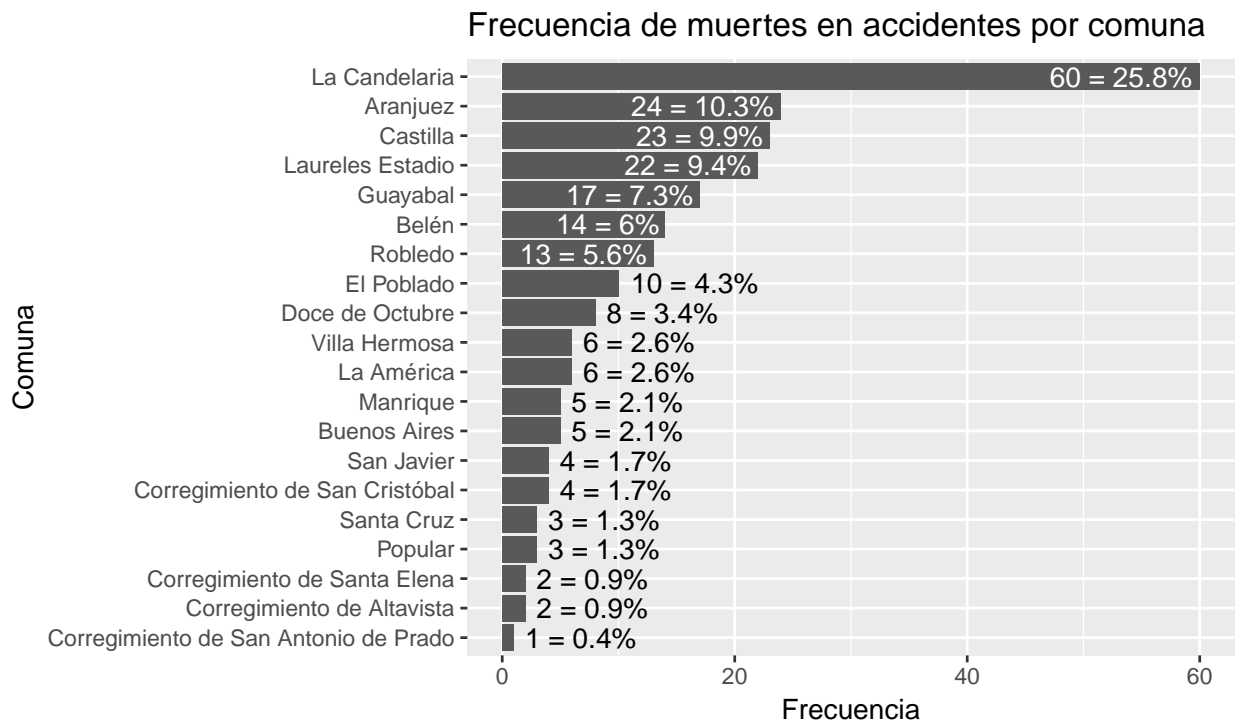
Claramente la distribución de muertes en accidentes según hora no es uniforme, y el pico se encuentra entre las 11 y 12 de la noche, seguido de entre las 6 y 7 p.m. Por esto deben tomarse medidas de control o concientización para generar precaución entre conductores. Entre 6 y 7 p.m. es hora pico, por tanto hay más flujo vehicular, y más probabilidades de generar accidentes, sin embargo, entre 11 y 12 de la noche, el flujo vehicular es menor, pero muchas personas pueden haber ingerido bebidas alcohólicas, o simplemente no respetan semáforos o cruces. Esto indica que es necesario prevenir estos accidentes mediante la generación de conciencia de respetar normas de tránsito, especialmente a estas horas.

Frecuencia de heridos y daños en accidentes por hora



La distribución de accidentes con heridos y daños claramente no es uniforme, con un pico de heridos entre 6 y 7 a.m., y segundo pico entre 4 y 5 p.m. El pico en la mañana es tan pronunciado que requiere atención, ya sea para verificar si este dato es correcto, como para tomar medidas de prevención.

3.2.4 Gravedad por comuna



Se observa que la mayoría de accidentes con muertes, heridos y daños materiales suceden en la comuna de la Candelaria, por lo que esta requiere especial atención, además porque la movilidad en esta zona es crucial para el resto de la ciudad. Se destacan la cantidad de heridos en la comuna de Castilla, y de heridos y daños en la comuna de Laureles-Estadio.

4. Conclusiones

La mayoría de accidentes en 2016 tuvieron como consecuencia heridos, lo cual invita a indagar por las causas de estas heridas para conocer qué acciones son pertinentes tomar.

Los días viernes se presentan mayores tasas de accidentalidad, y menos los domingos.

Las horas pico generan mayor cantidad de heridos y daños, pero el pico de muertes es entre 11 y 12 p.m.

Las zonas de mayor accidentalidad son las zonas del centro, y las zonas de la autopista norte y sur.

5. Bibliografía

[1] MOLLINEDO, Carmen Lizárraga. Movilidad urbana sostenible: un reto para las ciudades del siglo XXI. Economía Sociedad y Territorio, 2006.

[2] ASCHER, François. Ciudades con velocidad y movilidad múltiples: un desafío para los arquitectos, urbanistas y políticos. ARQ (Santiago), 2005, no 60, p. 11-19.