

TIPOLOGÍA DE DATOS

PRÁCTICA 1

CONTEXTO

El tema de la seguridad vial y consecuencias ha ido tomando importancia a lo largo de los años, y es que como lo indica la Organización Panamericana de la Salud en conjunto con la OMS¹, el 49% de las personas que mueren en las vías del mundo son peatones, ciclistas y motociclistas, lo que los convierte en las principales víctimas fatales de estos hechos, los cuales les cuestan a los países alrededor del 3% de su PIB. La OPS también advierte que las muertes y las lesiones producto de accidentes de tránsito son, ahora mismo, un problema de salud pública en la región de las Américas, lo cual se traduce en una posible carga para el sistema de salud de cada país a medida que las cifras aumenten.

En Estados Unidos de América, la cantidad de accidentes automovilísticos se mantuvo estable durante aproximadamente dos décadas (1990-2010). Sin embargo, entre el 2015 y 2019 esta cifra ha ido aumentando, aunque los índices de fatalidad han disminuido.

La llegada del coronavirus, provocó importantes reducciones en el tráfico vehicular en el último año, por lo cual se podría pensar que la frecuencia de los accidentes de tránsito, así como de las fatalidades disminuiría. Por esto, el propósito de la conformación de esta base de datos es promover una investigación que responda a la pregunta: ¿han disminuido los accidentes de tránsito en New York City a causa de la pandemia? De todos los accidentes reportados en la plataforma NYC Open Data, se recolectaron datos de aquellos que ocurrieron entre octubre de 2019 y enero 13 de 2020, y entre octubre de 2020 y enero 13 de 2021. La razón, es que según las cifras arrojadas diariamente por The New York Times², el primer caso de coronavirus llegó a la ciudad en marzo de 2020 por lo que, antes de este suceso había una afluencia relativamente normal de vehículos y la actividad económica no se había visto tan afectada, además, entre octubre y enero de 2021 se presentó un aumento progresivo de los casos de coronavirus y el pico de la pandemia llegó justo el 13 de ese mismo mes. Así, un dataset que contenga datos de este período de tiempo podría ser de gran ayuda para responder a la pregunta planteada, y adicionalmente visualizar patrones en los accidentes presentados y sus consecuencias para así posibilitar la toma de acciones para reducir estos sucesos en un futuro.

Adicionalmente, se recolectó información acerca de los 5 distritos que conforman Nueva York con el fin de averiguar el tamaño de cada uno, la cantidad de habitantes con la que cuentan, y su relación con los accidentes de tránsito presentados. Esto, ya que se podrían manejar las frecuencias en términos de proporciones y todo sería mucho más comparable.

Para la extracción de los datos de accidentes se eligió la página web NYC Open Data³, ya que es un sitio muy completo que ofrece información de distintas temáticas en la ciudad de Nueva York. Además, se visitó la página de United States Census Bureau⁴ para descargar la información de la densidad poblacional según los resultados del censo de 2019. Y de la página oficial del Departamento de Salud del Estado de Nueva York⁵ se extrajeron los datos del área del suelo de cada distrito

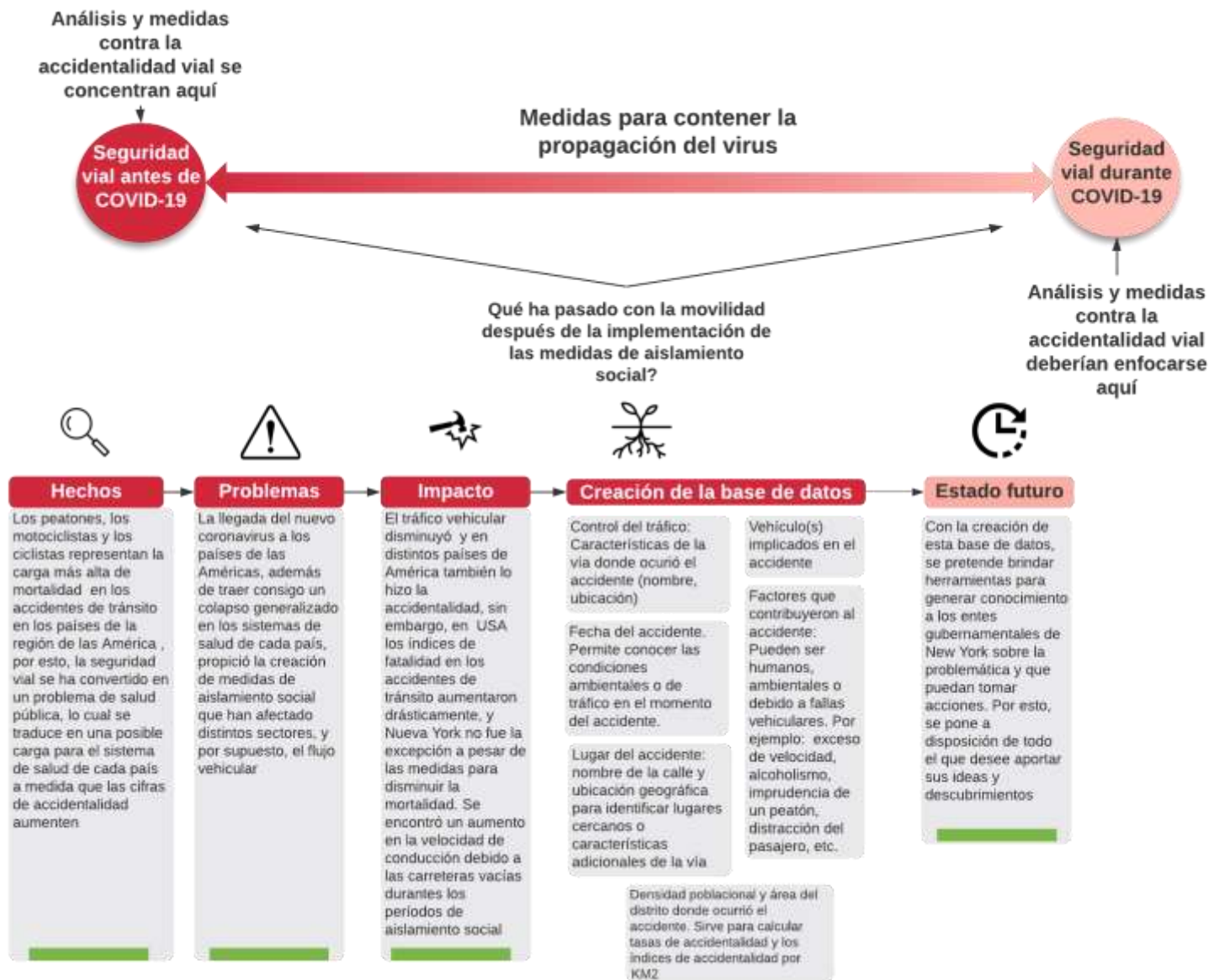
TÍTULO

Accidentes de tránsito en Nueva York. Efectos de una pandemia global.

DESCRIPCIÓN

El dataset: Accidentes de tránsito en Nueva York. Efectos de una pandemia global, contiene detalles sobre las colisiones de vehículos de motor presentadas en la ciudad de Nueva York durante el período de tiempo comprendido entre octubre y 13 de enero de los años 2020 y 2021. Cada fila, representa el día en que se presentó el accidente. Los períodos de tiempo elegidos, facilitan la comparación de los eventos ocurridos antes y durante la pandemia del coronavirus, posibilitando la realización de análisis detallados acerca los efectos de la pandemia sobre la movilidad y la accidentalidad vial en la ciudad. Los datos se conformaron a nivel de distrito, por lo que además se encuentra información sobre el área en KM cuadrados de cada uno y su densidad poblacional.

REPRESENTACIÓN GRÁFICA



CONTENIDO

Los campos mostrados en el dataset, corresponden al período de tiempo comprendido entre octubre y 13 de enero de los años 2020 y 2021 son los siguientes:

- **ID_ACCIDENTE:** Un identificador único para esta colisión.
- **DISTRITO:** El distrito en el que se produjo el accidente
- **LUGAR_ACCIDENTE:** Nombre de la calle donde se produjo el accidente
- **LATITUD:** Latitud del lugar del accidente. Ubicación geográfica
- **LONGITUD:** Longitud del lugar del accidente. Ubicación geográfica
- **DIA:** Día del accidente
- **MES:** Mes del accidente
- **AÑO:** Año del accidente
- **DIA_SEMANA:** Día de la semana en que ocurrió el accidente
- **HORA:** Hora del accidente
- **TIPO_VEHÍCULO_(1, 2, 3, 4, 5):** Tipos de vehículos implicados en el accidente
- **FACTOR_CONTRIBUYENTE_VEHÍCULO_(1, 2, 3, 4, 5):** Motivos del accidente por implicados.
- **TOTAL_HERIDOS:** Número de heridos en el accidente
- **TOTAL_MUERTOS:** Número de muertos en el accidente
- **CATEGORIA_HERIDOS:** Categoría de los heridos según si fueron: 1 = ciclista; 2 = motorista; 3 = peatón
- **CATEGORIA_MUERTOS:** Categoría de los muertos según si fueron: 1 = ciclista; 2 = motorista; 3 = peatón
- **DENSIDAD_POB:** Número de habitantes estimado para 2019
- **AREA_DISTRITO_KM2:** Área del distrito en KM cuadrados

Los datos se recogieron de la plataforma de datos abiertos de Nueva York (NYC Open Data)³. Dicha plataforma permite descargarlos en el formato que se desee, sin embargo, la descarga manual en un formato conocido y fácil de manejar como el CSV, toma mucho tiempo ya que puede resultar demasiado pesada la información. Es así como se decidió descargarlos en formato JSON usando Web Scraping con BeautifulSoup en Python, lo cual fue mucho más sencillo y rápido. De la página de United States Census Bureau⁴ se descargó mediante la misma técnica, la información de la densidad poblacional según los resultados del censo de 2019. Y la de página oficial del Departamento de Salud del Estado de Nueva York⁵ se extrajeron los datos del área del suelo de cada distrito.

AGRADECIMIENTOS

A todas las organizaciones que han aportado a la consolidación de las bases de datos publicadas en el sitio NYC Open Data, en especial al Departamento de Policía de Nueva York quienes en 1998 se aventuraron en la implementación de los principios CompStat con el fin de garantizar la seguridad en las carreteras de la ciudad, recolectando información precisa de cada accidente de tránsito ocurrido en un formato específico, lo que facilita el posterior análisis del comportamiento vial de los ciudadanos. También, al Departamento de Salud de Nueva York y al United States Census Bureau por ofrecer las estadísticas vitales de esa ciudad en forma actualizada cada cierto periodo de tiempo.

ANÁLISIS PREVIOS Y MOTIVACIÓN

Según la agencia de noticias Reuters⁶, aunque las medidas impuestas por los gobiernos a nivel mundial para contener la propagación del virus del COVID-19, han traído como consecuencia una importante reducción en el tráfico vehicular y también en los accidentes en las carreteras, se ha generado un incremento en las cifras de accidentes mortales, ya que las vías están más libres y la velocidad de conducción tiende a aumentar. En Nueva York, los análisis mostraron que en abril de 2020 la proporción de accidentes de tránsito mortales con respecto a todas las demás colisiones incrementaron un 167% comparado con el mismo período del año anterior, y que, en todo Estados Unidos, aunque el tráfico disminuyó, las tasas de mortalidad por milla recorrida incrementaron un 37%. Lo anterior sugiere que un aumento de velocidad puede ser la causa de estos sucesos.

Lo anterior, contrasta con las reducciones en las cifras de accidentes mortales de tránsito que se venían presentando desde el año 2014, cuando se empezó a implementar un plan para reducir las víctimas mortales en las carreteras de la ciudad. El plan llamado: Vision Zero, tiene como objetivo bajar a cero las cifras de fallecimiento por accidentes de tránsito⁷. Según el sitio de la Administración Nacional de Tráfico en Carretera (NHTSA), en su investigación sobre las colisiones en Estados Unidos⁸, la cantidad de accidentes de tránsito fatales en NYC disminuyó un 3.4% entre 2018 y 2019.

Es así como por medio del presente dataset se reconoce la importancia de analizar el tema de la seguridad vial en la ciudad de Nueva York, especialmente en época de pandemia y en el marco de las medidas de aislamiento social obligatorio que se han presentado a lo largo de este último año. Resulta interesante identificar patrones en los accidentes de tránsito ocurridos antes y durante de la entrada en vigencia de estas medidas, así como sus causas, sus desenlaces, su frecuencia y la tasa de mortalidad en estos tipos de incidentes.

LICENCIA

El presente dataset se publica bajo la licencia: CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International). La razón, es que gran parte de la información que contiene, viene del Departamento de Policía de Nueva York y ha sido publicada en la plataforma de datos abiertos de la ciudad bajo la modalidad de uso y visualización pública, por consiguiente, se deben conceder estos mismos permisos a la hora de modificarla o adaptarla y dejarla a la libre disposición de los usuarios bajo los términos de atribución y mantener abierta la información.

CONTRIBUCIONES

Investigación previa	Maria Victoria Diaz López
Redacción de las respuestas	
Desarrollo código	

REFERENCIAS

1. Organización Panamericana de la Salud. (s.f.). Seguridad Vial [Ficha informativa]. Recuperado de: <https://www.paho.org/es/temas/seguridad-vial#:~:text=La%20seguridad%20vial%20se%20refiere,muertes%20causadas%20por%20el%20tr%C3%A1nsito.&text=El%2090%25%20de%20las%20muertes,d el%203%25%20de%20su%20PIB>
2. The New York Times. (2021). Tracking the Coronavirus in New York [Artículo]. Recuperado de: <https://www.nytimes.com/interactive/2021/us/new-york-covid-cases.html>
3. NYC Open Data. (2021). Recuperado de: <https://data.cityofnewyork.us/>
4. United States Census Bureau. (s.f.). QuickFacts [Database]. Recuperado de: <https://www.census.gov/quickfacts/fact/table/newyorkcitynewyork,bronxcountybronxboroughnewyork,kingscountybrooklynboroughnewyork,newyorkcountymanhattanboroughnewyork,queenscountyqueensboroughnewyork,richmondcountystatenislandboroughnewyork/POP060210>
5. New York Department of Health. (2011). Vital Statistics of New York State 2010 [Database]. Recuperado de: https://www.health.ny.gov/statistics/vital_statistics/2010/table02.htm
6. Masri, Lena. (26 de junio de 2020). Car crashes deadlier as drivers speed during lockdowns. Reuters. Recuperado de: <https://www.reuters.com/article/us-health-coronavirus-traffic-casualties-idUSKBN23X1OR>
7. NYC Vision Zero. (s.f.). Recuperado de: <https://www1.nyc.gov/content/visionzero/pages/>

María Victoria Díaz López
Maestría en Ciencia de Datos

8. National Highway Traffic Safety Administration. (2020). Overview of Motor Vehicle Crashes in 2019 [Artículo en línea]. Recuperado de:
<https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813060>