

Caracterización física, demográfica y socioeconómica de los cantones de Costa Rica a partir de las tasas de usuarios lesionados o fallecidos en carretera en el año 2018

Physical, demographic, and socioeconomic characterization of the cantons of Costa Rica based on the rates of users injured or killed on the road in 2018


Caracterização física, demográfica e socioeconômica dos cantões da Costa Rica com base nas taxas de usuários da estrada feridos ou mortos em 2018

Henry Hernández-Vega¹
Universidad de Costa Rica, Costa Rica

Diana Sanabria-Barboza²
Universidad de Costa Rica, Costa Rica

Resumen

Este trabajo presenta la aplicación de técnicas relacionadas con Análisis de Componentes Principales (ACP) que buscó caracterizar en limitadas dimensiones a los 81 cantones de Costa Rica con base en indicadores de exposición a heridas leves, graves o fallecimiento (en sitio) en accidentes de tránsito de distintos usuarios de la vía, con el fin de analizar posibles variables que podrían estar relacionadas a su incidencia. Con el análisis de agrupamientos jerárquicos se hallaron tres grupos de cantones con condiciones diferenciadas en dos componentes construidos y en términos de las

1 Máster, Escuela de Ingeniería Civil de la Universidad de Costa Rica. Correo electrónico: henry.hernandez@ucr.ac.cr  <https://orcid.org/0000-0002-4765-7320>

2 Licenciada, Escuela de Ingeniería Civil de la Universidad de Costa Rica. Correo electrónico: diana.sanabriabarboza@ucr.ac.cr  <https://orcid.org/0000-0001-6262-8483>

tasas de incidencia en análisis. Al asociar variables físicas, se halló que los cantones con menores índices de desarrollo y mayores carencias presentan mayor riesgo de victimización en ocupantes de automóviles de cuatro o más ruedas.

Palabras clave: APC, agrupación, accidentes, tránsito

Abstract

This work presents the application of techniques related to Principal Component Analysis (PCA) to characterize the 81 cantons of Costa Rica in limited dimensions, based on indicators of road users' exposure to minor, serious injuries, or death (on-site) in traffic accidents, to analyze possible variables that could be related to their incidence. Through the analysis of hierarchical groupings, three groups of cantons were found with differentiated conditions in two constructed components, and in terms of the incidence rates in analysis. With the association of physical and sociodemographic variables, it was found that the cantons with the lowest development indices, the greatest deficiencies, and the highest proportion of young population have a higher risk of road victimization in occupants of cars with four or more wheels.

Keywords: PCA, aggrupation, accidents, transit

Resumo

Este trabalho apresenta a aplicação de técnicas relacionadas à Análise de Componentes Principais (ACP) que buscaram caracterizar em dimensões limitadas os 81 cantões da Costa Rica com base em indicadores de exposição a ferimentos leves ou graves ou morte (no local) em acidentes de trânsito de diferentes usuários da via, a fim de analisar possíveis variáveis que possam estar relacionadas à sua incidência. Com a análise de agrupamento hierárquico, foram encontrados três grupos de cantões com condições diferenciadas em dois componentes construídos e em termos das taxas de incidência em análise. Ao associar variáveis físicas, verificou-se que os cantões com menores índices de desenvolvimento e maiores deficiências apresentam maior risco de vitimização em ocupantes de carros com quatro ou mais rodas.

Palavras-chave: Acidentes; Grupo; APC; Trânsito.

Introducción

La investigación en torno al análisis de personas afectadas en accidentes de tránsito responde a una problemática vigente en torno a la movilidad: según la [Organización Mundial de la Salud \(2018\)](#), el número anual de muertes por accidentes de tránsito es superior a 1,35 millones y cuestan a la mayoría de los países el 3% de su Producto Interno Bruto (PIB), además, son ahora la principal causa de muerte de las personas de entre 5 y 29 años. Se destaca la afección desproporcionada a los peatones, ciclistas y motociclistas que viven en países en desarrollo. Adicionalmente, [Jaramillo, Lizárraga y Grindlay \(2012, p. 342\)](#) indican que es importante considerar las necesidades sociales de la población al momento de tomar decisiones en la planificación del transporte.

Por otro lado, la variación en los patrones de movilidad debido al origen de la emergencia sanitaria a causa de la pandemia del Covid-19 en marzo del 2020 tuvo efectos diversos en la incidencia de lesiones y defunciones por accidentes de tránsito. En el caso de Estados Unidos, si bien se redujo en 13,2% la conducción en el 2020 debido a la pandemia, se estima que alrededor de 38680 personas murieron en accidentes de tránsito automovilísticos, lo cual representa un aumento del 7,2% con respecto al año 2019 y la mayor cantidad proyectada de muertes desde 2007 ([National Highway Traffic Safety Administration, 2021](#)).

Particularmente en Costa Rica sí se observó una disminución de este tipo de sucesos. Según el Consejo de Seguridad Vial ([COSEVI, 2021](#)), la tasa de mortalidad en sitio en accidentes de tránsito durante el 2020 fue de 6,08 personas por cada 100000 habitantes, mientras que la tasa de lesionados fue de 287,1 en el mismo año, cuando en el año 2019 se estimaron tasas de 8,92 y 378,7 respectivamente. La estacionalidad generada por la pandemia no inhibe la situación nacional en cuanto a accidentes de tránsito, la cual obligó a varias instituciones gubernamentales a crear una estrategia en el año 2019 para plantear una medida contra los accidentes de tránsito ([Ministerio de Salud, 30 de octubre de 2019](#)).

En cuanto a la localización geográfica de estos eventos y su asociación a las características de los entornos donde la frecuencia es mayor, los hallazgos son necesarios para orientar la aplicación de medidas estratégicas que busquen contrarrestar este fenómeno epidémico.

El problema radica en que en el país no existe evidencia de que existan relaciones entre variables geográficas que tengan relación con el desempeño de la seguridad vial. El presente estudio tiene como objetivos agrupar a los cantones de Costa Rica con base en tasas de exposición de riesgo comparativo de sufrir heridas o fallecer (en sitio) en un accidente de tránsito durante el año 2018, y caracterizar a los grupos de cantones resultantes con base en variables externas que son determinantes en la dinámica de la movilidad cantonal. Para lograr lo anterior, se propone agrupar a los cantones nacionales según tasas de exposición construidas para eventos del año 2018, las cuales indican el riesgo comparativo de las personas de sufrir heridas o fallecer (en sitio) en un accidente de tránsito, además de su caracterización según otras variables características consideradas como significativas en la dinámica de la movilidad cantonal.

Se aclara que el propósito del presente estudio es el de añadir insumos adicionales con base en técnicas multivariantes de componentes principales y análisis de conglomerados, que permitan un mejor entendimiento y caracterización a nivel cantonal de un fenómeno complejo como lo es la accidentabilidad, en el cual intervienen una serie de diferentes factores.

Marco teórico-conceptual

Variedad de autores han estudiado la relación entre accidentabilidad y características socioeconómicas a nivel de las personas usuarias de las vías, de los hogares de ellas o de una región. En relación a los individuos, por ejemplo, en una investigación realizada en los Estados Unidos, se comparó a los peatones con otros pacientes con lesiones no intencionales. En dicho estudio, se encontró que los peatones atendidos tenían una probabilidad significativamente mayor de ser afroamericanos, solteros, desempleados, bebedores compulsivos, dependientes del alcohol, dependientes de las drogas, de presentar concentración de alcohol en la sangre mayor a 200 mg/dl, de tener bajos ingresos, bajo rendimiento educativo, de ser más joven y no tener una licencia de conducir o tenerla revocada o caducada (Ryb, Dischinger, Kufera, y Soderstrom, 2007). Especialmente, con respecto a la característica de edad, Agüero Valverde y Jovanis (2006, p. 620) determinaron que los condados en Pensilvania que tenían un porcentaje más alto de población por debajo del nivel de pobreza y en los grupos de edad de 0 a 14 años, 15 a 24 años y mayores de 64 años y aquellos con un mayor kilometraje y densidad de carreteras presentaban un mayor riesgo de accidentes.

Por otro lado, al estudiar el vínculo entre los accidentes de tránsito y la caracterización de los hogares de las personas involucradas, Pirdavani, Daniels, Van Vlierden, Brijs y Kochan (2017) encontraron que el ingreso promedio del hogar y la tenencia de vehículo son predictores significativos de cierto tipo de accidente: a menor nivel de ingreso del hogar hay una mayor cantidad de accidentes viales, específicamente. También encontraron que esta magnitud varía entre géneros, aspecto que puede ser de especial interés en el futuro del país. De manera similar, Brown (2016) halló relaciones entre el riesgo de conductores de estar involucrados en colisiones con muertos o heridos y las características sociodemográficas de las áreas residenciales donde habitan los conductores en Carolina del Sur, Estados

Unidos. Anteriormente, [Fleury, Peytavin, Alam, y Brenac \(2010\)](#) encontraron un mayor riesgo de sufrir un accidente de tránsito si se vive en una zona con viviendas en mal estado o con un gran desbalance entre la cantidad de empleos y la cantidad de personas que habitan en la zona.

Las condiciones del entorno son determinantes para la ocurrencia de accidentes. Por ejemplo, recientemente, [Hu, Wu, Huang, Peng y Liu \(2020\)](#) confirmaron que la frecuencia de atropellos se encuentra relacionada con las características del ambiente de la vía, con el uso de suelo urbano y con la distribución de la red de carreteras. [Cubbin y Smith \(2002\)](#) mencionaron que las comunidades de mayores estatus sociales tienden a vivir en zonas con mejor espacio público y vías seguras, además de que usualmente tienen acceso restringido al tráfico y son menos tolerantes a comportamientos que se apartan de ciertas normas sociales; factores que podrían ser protectores ante las lesiones. En un estudio reciente se analizó la influencia de la pandemia del COVID-19 en los accidentes de tránsito en Los Ángeles y Nueva York, en la cual se encontró que, dado que ha afectado de manera desproporcionada a ciertos grupos de edad, razas y géneros, los “puntos calientes” de accidentes de tránsito se han desplazado tanto en el tiempo como en el espacio ([Lin, Shi y Li, 2020](#)).

En cuanto al estudio del entorno, en una investigación reciente de [Yang, Chung y Kim \(2021\)](#) se utilizó aprendizaje de máquinas, del cual se concluyó que la presencia de instituciones educativas, de salud y de servicios, junto con la presencia de semáforos, tienen una influencia positiva significativa en la ocurrencia de accidentes. La relación entre el paisaje urbano, su transición y la accidentabilidad fue analizada por [Poulos, Chong, Olivier y Jalaludin \(2012\)](#), quienes determinaron que localidades con mayor riesgo peatonal tienden a ser áreas metropolitanas urbanizadas con mayor densidad poblacional y el riesgo ciclista tiende a aumentar en regiones urbanas o en el margen de áreas urbanizadas metropolitanas de New South Wales en Australia.

Las características de los sitios inmediatos a los eventos de accidentes se han estudiado como potenciales causas de los accidentes. En un estudio desarrollado en Inglaterra y Gales se determinó que variables relacionadas con el número y las características de la población, la exposición al tráfico, la longitud de la carretera, la curvatura y la densidad de las intersecciones, el uso del suelo, la elevación y el tipo de terreno y el

clima como posibles variables explicativas de la cantidad de víctimas relacionadas con eventos de tránsito (Jones *et al.*, 2008). De manera similar, el estudio de Hashimoto (2005), realizado en el condado de Hillsborough en Florida, Estados Unidos, determinó que el uso de suelo y la densidad de la red vial estatal afecta el número de atropellos; por otro lado, determinó que el ingreso del hogar y el porcentaje de adultos mayores reduce la cantidad de atropellos.

No necesariamente el sitio donde ocurre una colisión corresponde con el sitio donde reside esa persona, según Hewson (2005). Además, indicó que una persona de menor nivel socioeconómico tiene menos posibilidades de viajar en automóvil, lo que establece una relación entre el nivel socioeconómico y probabilidad de sufrir un atropello. Este autor señala el problema con respecto a la subestimación de reportes de heridos. Por ejemplo, en la Dirección General de Policía de Tránsito de Costa Rica, no existe un protocolo, al menos que sea de conocimiento de los autores, para clasificar a los heridos en carretera. Se ha conocido de manera puntual de casos de heridos leves (principalmente usuarios vulnerables) que no son reportados en las estadísticas de accidentes de tránsito en el país.

En el ámbito costarricense, Vargas Aguilar y Agüero Valverde (2020) elaboraron modelos estadísticos de estimación de choques mediante la aplicación de la distribución binomial negativa y el uso de suelo, además de variables socioeconómicas y de la red vial. Los autores también determinaron los distritos con mayor exceso de frecuencia, en términos de exceso de choques sobresalió Alajuela, Uruca, Hospital, San Nicolás, Río Segundo y Heredia. En un estudio de Agüero Valverde (2013) basado en modelos univariados y multivariados derivados de un enfoque jerárquico de Bayes incluyó diferentes variables por cantón para determinar los cantones con exceso de colisiones. Las variables consideradas fueron la cantidad de colisiones (mortales, con heridos y con daños materiales), la cantidad de vehículos-kilómetro en rutas nacionales, longitud de la red vial, el porcentaje de población bajo la línea de pobreza, el porcentaje de personas por rango de edad (entre 0 y 14 años, entre 15 y 24 años, y mayores de 65 años), porcentaje de mujeres y el porcentaje de población urbana. En la capital, San José, se presentó el mayor exceso de choques, seguida de Desamparados, cantón muy poblado cercano a la capital y Alajuela, la segunda ciudad en importancia de Costa Rica (Agüero Valverde, 2013).

En el sector transporte en Costa Rica, [Magaña, Hernández y Jiménez \(2014\)](#), aplicaron métodos de conglomerados para caracterizar las variaciones temporales del tráfico en diferentes zonas del país.

Por otro lado, a partir de un análisis de conglomerados de colisiones con ciclistas involucrados en la ciudad de Bogotá, [Torres-Barragan, Cotrill y Beecroft \(2020\)](#) indican que existe concentración de colisiones que resultan de muertes de ciclistas en zonas de bajos ingresos y una asignación de infraestructura desigual. Además, [Martínez, Contreras y Moreno \(2020\)](#), generaron un índice de privación utilizando variables socioeconómicas en un estudio realizado en Chile, con el fin de analizar las desigualdades socioeconómicas. En dicho estudio concluyeron que, a mayor privación socioeconómica a nivel regional, se tienen mayores tasas de accidentes en adultos mayores. Los autores indican que es necesaria más investigación relacionada con las desigualdades socioeconómicas en las tasas de accidentes en los países en desarrollo, especialmente, enfocada en adultos mayores. En una investigación realizada por [Shen, Lu, Long & Chen \(2019\)](#), se combinó el uso de análisis de componentes principales, de conglomerados entre otras técnicas, para generar identificación de puntos negros en vías rurales en China.

Materiales y métodos

El presente estudio se enfoca más en las víctimas de los accidentes de tránsito que los accidentes en sí mismos. Esto porque en un accidente de tránsito pueden verse afectados más de un tipo de persona usuaria de la vía, por ejemplo, personas usuarias de vehículos motorizados (tanto de transporte público como particular), ciclistas o peatones. Debido a lo anterior se generan tasas a nivel cantonal que muestran la exposición al riesgo de los diferentes tipos de personas usuarias del sistema de transporte.

Así, como variables medulares para la evaluación de los cantones en la temática, se construyeron tasas cantonales relacionadas con la exposición de las personas a heridas o decesos, derivadas principalmente de datos demográficos del X Censo Nacional de Población y VI Censo Nacional de Vivienda de Costa Rica, publicados en el año 2011 por el Instituto Nacional de Estadística y Censos ([INEC, 2011](#)), en conjunto con datos de accidentes del Observatorio Costarricense de Seguridad Vial del Consejo de Seguridad Vial ([COSEVI, 2020](#)).

Para el trabajo de construir tasas poblacionales que relacionen la exposición de usuarios de la vía a lesiones o muerte en un accidente de tránsito con el total de habitantes del cantón en el periodo de estudio, que relacionen así la ocurrencia de accidentes y unidades geográficas, es necesario destacar la limitante de que la ocurrencia geográfica de un evento de accidente de tránsito no coincide necesariamente con el lugar de origen o residencia de las personas involucradas en el accidente. Si bien se esperaría que esta limitación sea menor para los accidentes con ciclistas y peatones involucrados, para los casos de accidentes con vehículos el sesgo debe ser mayor dado que este medio de transporte se asocia a traslados mayores entre el origen y el sitio del accidente.

Para cada uno de los 81 cantones existentes, en el año 2018, se estimaron tasas asociadas al riesgo de lesionarse leve o gravemente o fallecer (en sitio) en un accidente de tránsito bajo el rol de conductores, motociclistas y peatones, según las ecuaciones (1), (2) y (3) respectivamente.

$$T_C = \frac{A_{C-2018}}{V_{veh} * P_{2018}} * 1000 \quad (1)$$

Donde:

- T_C : tasa conductores accidentados con lesiones o fallecidos en el año 2018 por cada 1000 habitantes del cantón.
- A_{C-2018} : conductores accidentados en el año 2018.
- V_{veh} : porcentaje de viviendas del cantón con vehículo en el año 2011.
- P_{2018} : población del cantón en el año 2018.

$$T_M = \frac{A_{M-2018}}{V_{mot} * P_{2018}} * 1000 \quad (2)$$

Donde:

- T_M : tasa de motociclistas accidentados con lesiones o fallecidos en el año 2018 por cada 1000 habitantes del cantón.
- A_{M-2018} : motociclistas accidentados en el año 2018.
- V_{mot} : porcentaje de viviendas del cantón con motocicleta en el año 2011.
- P_{2018} : población del cantón en el año 2018.

$$T_p = \frac{A_{P-2018}}{P_{2018}} * 1000 \quad (3)$$

Donde:

- T_p : tasa de peatones accidentados con lesiones o fallecidos en el año 2018 por cada 1000 habitantes del cantón.
- A_{P-2018} : peatones accidentados en el año 2018.
- P_{2018} : población del cantón en el año 2018.

Para las tasas calculadas se supuso que los porcentajes V_{veh} y V_{mot} son equivalentes a la proporción poblacional del cantón que conduce cada tipo de vehículo. La población P_{2018} corresponde a las proyecciones de población cantonales para dicho año, elaboradas por el Centro Centroamericano de Población (CCP, 2011) en colaboración con el INEC (2011). Se aclara, además, que no se realizó una ponderación por gravedad del accidente.

Por otro lado, se recopiló y se construyó una serie de variables demográficas, sociales y físicas, específicamente sobre la composición etaria poblacional de cada cantón, el desarrollo humano, y condiciones geográficas y de infraestructura vial. Se buscó teorizar para el caso de Costa Rica si es posible que estas variables caractericen distintivamente a los grupos de cantones encontrados que tienen similares patrones de ocurrencia de víctimas accidentadas en la vía. Seguidamente se detallan las variables consideradas para este efecto:

- **Densidad de vías (km/km²):** Se trata de una variable cantonal construida al dividir la cantidad de kilómetros de red vial nacional y cantonal entre el área del cantón. Las longitudes de vías de la red vial nacional fueron obtenidas a partir de datos geográficos del Instituto Geográfico Nacional (2020) y se estimó la longitud de vías cantonales a partir de información vectorial extraída de la plataforma Open Street Map (2020). Para el procesamiento se utilizó el software QGIS.
- **Densidad poblacional (habitantes/Ha):** Se trata de una variable cantonal construida, al dividir la variable P_{2018} entre el área del cantón.
- **Proporciones de habitantes en rangos etarios (%):** Con base en la información del Censo 2011, por cada cantón se estimaron proporciones de habitantes para los siguientes rangos etarios: menores de 14 años, 15 a 24 años, 25 a 64 años y mayores de 65 años.

- **Proporciones de habitantes con necesidades básicas insatisfechas (NBI):** A través de la consulta de la información por cantón del Censo 2011 se obtuvo la proporción de la población con 0, 1, 2, 3 o 4 NBI. Según el INEC (2015), se trata de la única medición multidimensional de pobreza que se ha realizado oficialmente en Costa Rica, la cual tuvo como objetivo identificar a los hogares pobres según un conjunto de 4 necesidades socialmente consideradas básicas: acceso a albergue digno, acceso a vida saludable, acceso al conocimiento y acceso a otros bienes y servicios.
- **Índice de Desarrollo Humano (IDH):** Se recopiló el IDH de cada cantón para el año 2018, según el Programa de las Naciones Unidas para el Desarrollo (2020). El IDH mide el promedio de los logros de un área geográfica específica en tres dimensiones básicas del desarrollo humano: una vida larga y saludable (medida por la esperanza de vida al nacer), el conocimiento, (medido por los años esperados de escolaridad y los años promedio de escolaridad) y un nivel de vida digno (medido por el Índice de Bienestar Material a partir del consumo de electricidad residencial por cliente).

En términos generales, para alcanzar el objetivo planteado se utilizó la técnica Análisis de Componentes Principales (ACP), para combinar las tres tasas en estudio mencionadas en este apartado (T_C , T_{MP} , T_P) y reducirlas a un número menor de combinaciones lineales o dimensiones. A la vez, se plantea también la aplicación del método no paramétrico de análisis de agrupamientos para detectar posibles relaciones con las variables sociales, demográficas y físicas presentadas con las dimensiones sugeridas.

Para la relación de las tasas construidas como variables cuantitativas originales y las dimensiones por construir, se utilizó la matriz de correlaciones, dado que se trata de un subgrupo de variables que se puede considerar estandarizado: se derivaron con base en datos de accidentes del año 2018 y cuentan con la misma unidad. La matriz de correlaciones es de común aplicación para un caso como este.

Con base en el Cuadro 1 y en cuanto a la dispersión estadística de los valores, las tasas presentan varianzas en diferentes órdenes de magnitud. Esta condición es favorable para la aplicación de la técnica ACP, puesto que si se tratara de información muy redundante puede hacer que el método sobre enfatice la contribución de las variables a las dimensiones por

desarrollar. A esto también favorece la inspección de la correlación lineal entre las tasas, dado que se encontraron bajas relaciones directamente proporcionales entre los pares T_C y T_M (0,01) y T_C y T_P (0,17), la correlación más alta se observó entre las tasas T_M y T_P (0,61).

Cuadro 1. Descripción estadística de las tasas cantonales de víctimas de accidentes de tránsito, Costa Rica, 2018.

Valor estadístico	T_C	T_M	T_P
Mínimo	1,53	7,19	0,00
Q1	4,22	30,12	0,19
Mediana	7,19	41,88	0,26
Q3	12,78	61,64	0,40
Máximo	28,10	156,01	0,91
Media	9,39	50,24	0,30
Varianza	46,08	934,74	0,03
Desviación estándar	6,79	30,57	0,17

Nota: T_C , T_M y T_P corresponde a tasa de conductores, motociclistas y peatones, respectivamente, accidentados con lesiones o fallecidos en el año 2018 por cada 1000 habitantes del cantón.

Fuente: Elaboración propia.

La técnica ACP se fundamenta en la propiedad de que la varianza de cada componente determinado es la estimación del valor característico asociado, y la sumatoria de todos los valores característicos es equivalente a la suma de las varianzas de las variables originales. Esta propiedad permite calcular la porción de la varianza original que cada uno de los componentes generados explica. Se propuso la combinación de las tasas cantonales relacionadas con el riesgo de lesión o muerte en accidentes de tránsito, a través de su nexo con la matriz de correlaciones se encontraron los componentes principales cuyas características se presentan en el Cuadro 2.

Cuadro 2. Datos asociados a la importancia de los componentes definidos mediante el análisis ACP a través de la matriz de correlaciones

Valor estadístico	Componente 1	Componente 2	Componente 3
Desviación estándar	1,28	1,00	0,60
Variancia	1,64	1,00	0,37
Proporción de la varianza explicada	0,55	0,33	0,12
Proporción de la varianza acumulada	0,55	0,88	1,00

Fuente: Elaboración propia.

A partir del Cuadro 2 es posible aplicar análisis elementales para seleccionar una cantidad de componentes apropiada: es adecuado utilizar únicamente los componentes para los cuales se estimó un valor característico (variancia) mayor a la unidad, con base en la información del Cuadro 2 se elegirían sólo los primeros dos componentes porque tienen esta condición.

Considerando que se busca un alto porcentaje de varianza explicada, con base en las propiedades intrínsecas del método, aunque no se cuenta con un criterio técnico definido sobre cuánto es el límite inferior conveniente, se aprecia en el Cuadro 2 que la proporción de varianza acumulada en los primeros dos componentes supera el 80%, lo cual refuerza la decisión de limitarse a ellos.

En el Cuadro 3 se presenta la matriz de contribuciones, con la cual es posible cuantificar cuál es el aporte de cada una de las tasas a la construcción de los componentes generados, en este caso a los dos elegidos. Además, el plano factorial, representado en el Gráfico 1, presenta el vínculo entre variables y dimensiones. Con base en la interpretación de ambos insumos se proponen las siguientes asociaciones.

En términos generales las tasas victimización de motociclistas como de peatones contribuyen en alto grado e igual proporción a la conformación de la Dimensión 1, ambas se ubican en el eje positivo del Gráfico 1 por lo que la Dimensión 1 puede considerarse como una función creciente de ambas tasas.

Cuadro 3. Indicadores de relación de las variables y los componentes definidos mediante el análisis de componentes principales a través de la matriz de correlaciones

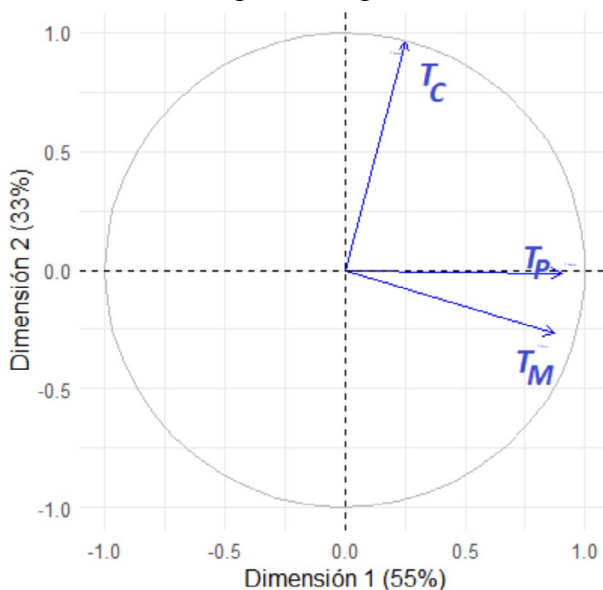
Variable	Contribución		Representación (Cos ²)	
	Dimensión 1	Dimensión 2	Dimensión 1	Dimensión 2
T_C	3,93	92,83	6,44	92,37
T_M	46,27	7,16	75,84	7,12
T_P	49,80	0,02	81,63	0,02

Nota: T_C , T_M y T_P corresponde a la tasa de conductores, motociclistas y peatones, respectivamente, accidentados con lesiones o fallecidos en el año 2018 por cada 1000 habitantes del cantón.

Fuente: Elaboración propia.

En el Gráfico 1 se aprecia que el vector de la tasa de peatones se encuentra casi completamente en la dirección del eje, lo cual es consecuente con la nula contribución y representación de esta variable en la Dimensión 2, componente en el cual sobresale significativamente el aporte de la tasa de conductores (92,83%) y un aporte de la tasa de motociclistas, en considerable menor medida.

Gráfico 1. Plano factorial de relación de las variables originales con los dos componentes generados



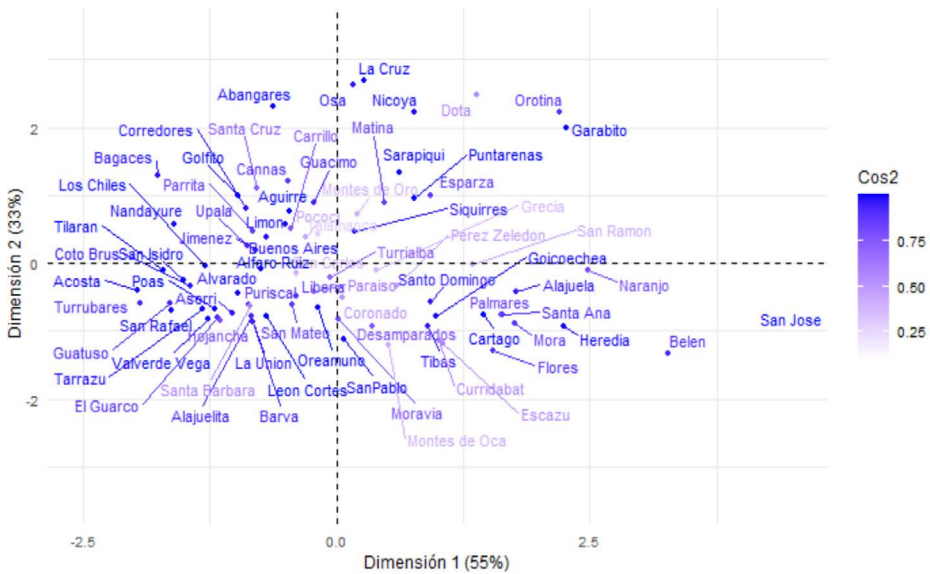
Fuente: Elaboración propia.

Nota: T_C , T_M y T_P corresponde a la tasa de conductores, motociclistas y peatones, respectivamente, accidentados con lesiones o fallecidos en el año 2018 por cada 1000 habitantes del cantón.

Con las asociaciones realizadas se interpreta con confianza que la Dimensión 1 está regida por las tasas relativas a los usuarios con roles expuestos a mayor riesgo de sufrir lesiones o fallecer en accidentes de tránsito, por las condiciones físicas en las que se desplazan por las vías, por lo cual se define como la Dimensión Usuarios vulnerables. Consecuentemente, la Dimensión 2 corresponde a Vehículo privado (de al menos cuatro ruedas), la cual está relacionada con los usuarios expuestos a heridas o muerte en accidentes de tránsito que conducen los vehículos.

El Cuadro 3 también presenta los valores del indicador cos^2 , estos son útiles para determinar si las variables originales sometidas a la técnica ACP presentan un ajuste adecuado, o en su defecto no se ajustan a la información que explican las dimensiones sugeridas. Se aprecia que los valores de representación muestran la misma tendencia que los valores de contribución, con lo cual las tasas cantonales de conductores se encuentran muy bien representadas en la Dimensión 2, y en segundo plano las tasas de motociclistas y peatones en la Dimensión 1.

Gráfico 2. Plano de cantones en los componentes principales construidos e indicador cos^2



Fuente: Elaboración propia.

El Gráfico 2 permite el análisis integral de dos elementos: el par de componentes principales en cada eje lo que conforma un plano y la dispersión de los individuos en el mismo. Con base en este gráfico y al inspeccionar ciertos datos descriptivos cantonales, se teoriza que los cantones ubicados en el cuadrante inferior derecho presentan características sociodemográficas más privilegiadas y mayor urbanización o densidad, en contraposición a los del cuadrante superior izquierdo.

Resultados

Seguidamente, se presentan los resultados de la aplicación del método de análisis de agrupamientos jerárquicos con la finalidad de detectar patrones, es decir, grupos de cantones que sean heterogéneos entre sí, pero homogéneos en sí mismos con respecto a las dimensiones propuestas. Posteriormente se relacionan las variables de caracterización cantonal seleccionadas (físicas, demográficas y socioeconómicas) con los grupos de cantones y las dimensiones previamente determinadas.

Es destacable mencionar que se aplicarán las dimensiones resultantes de la técnica ACP como variables sustitutivas de las originales. Esas dimensiones explican la variabilidad que el conjunto original presentaba, pero con una condición estandarizada, ya que se derivan de una carga factorial.

Con base en las cargas factoriales que tiene los cantones en los nuevos componentes se hizo una exploración inicial aplicando el método del codo (*elbow method* en inglés), en donde se representa la proporción de varianza explicada en función del número de grupos, para plantear una cantidad de agrupamientos conveniente. Se observó que sobresale la desviación en la curva cuando se acumulan 3 agrupamientos.

Así, se analizó primero la generación de tres clústeres jerárquicos. Para generarlos se empleó el método de agrupación de Ward, que utiliza un criterio de varianza mínima, aplicado mediante un algoritmo que busca minimizar la varianza dentro de cada conglomerado. Se trata de un método iterativo, en cada iteración se encuentra un par de conjuntos que al mezclarse incrementan mínimamente del total de la varianza. El método de Ward se caracteriza por su insensibilidad a valores extremos o atípicos. Para efectos comparativos, también se generaron agrupaciones con el método similar denominado K-Medias que se basa en el algoritmo de [Hartigan y Wong \(1979\)](#) para la generación de conglomerados a través de un proceso de 100 iteraciones, que busca maximizar la diferencia entre los grupos y minimizando la variabilidad dentro de cada uno. Los resultados se resumen en el Cuadro 4.

Cuadro 4. Promedios de las cargas factoriales en las dimensiones en tres grupos jerárquico de cantones

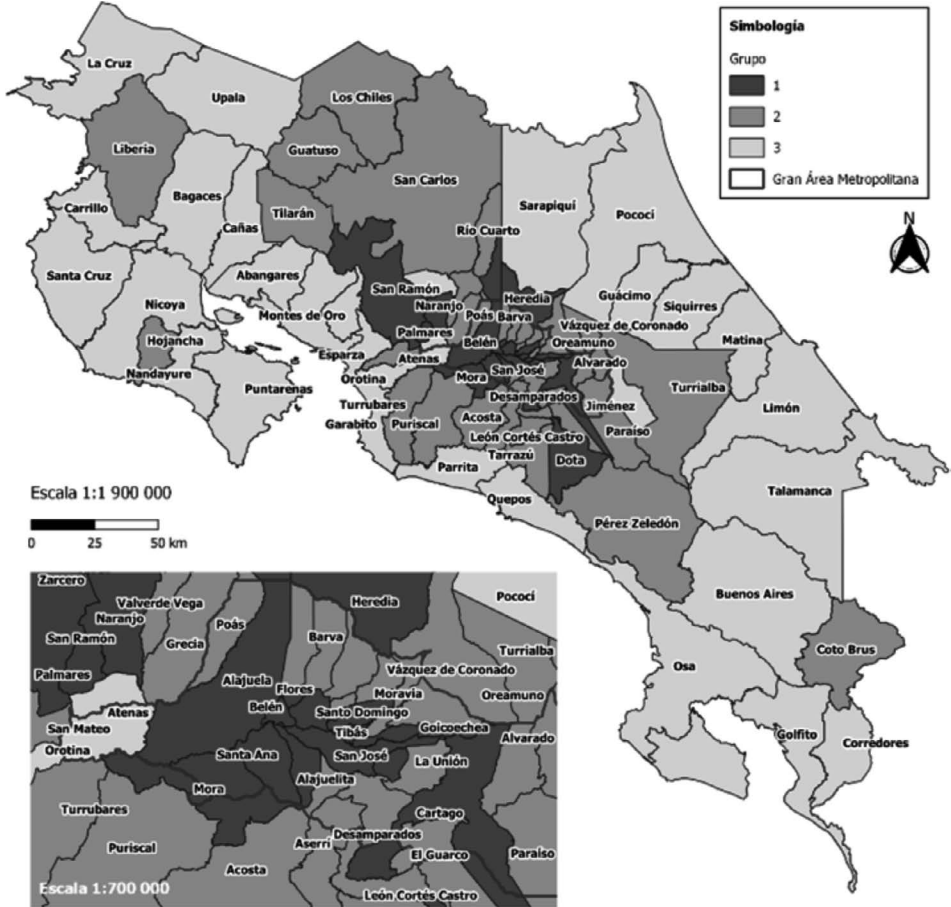
Grupo	Agrupación de Ward		Algoritmo de Hartigan-Wong	
	Dimensión 1	Dimensión 2	Dimensión 1	Dimensión 2
1	1,71	-0,63	1,60	-0,80
2	-0,80	-0,57	-0,81	-0,18
3	-0,14	1,01	0,62	1,68

Fuente: Elaboración propia.

Los resultados de ambos métodos son completamente comparables en magnitud y orden interno de los grupos en cada dimensión. La variabilidad de los promedios de las tres agrupaciones en cada dimensión ofrece una cómoda interpretabilidad de su significado, lo cual refuerza la elección de ese número de grupos.

Al buscar comprender el perfil de cada grupo, especialmente en relación las características físicas y sociales de cada cantón, se esperan diferencias relevantes, aun considerando las asociaciones subyacentes entre subgrupos de cantones individuales que pertenecen a un mismo grupo. Primeramente, se inspeccionó la composición de cada grupo en los cantones específicos que conforman cada uno de los tres clústeres jerárquicos generados. El resultado se presenta en la Figura 1 y la identificación de los cantones se enlista en el Cuadro 5.

Figura 1. Grupos de cantones según análisis de agrupamiento jerárquico en torno a tasas de personas heridas o fallecidas en accidentes de tránsito



Fuente: Elaboración propia con información del IGN (2020).

Como segundo objetivo, se buscó caracterizar a los grupos de cantones identificados con base en las dimensiones construidas. Con el vínculo de la información del Cuadro 4 y del Cuadro 5, se puede puntualizar que el grupo 1 tiene una media mayor asociada al componente Usuarios vulnerables (1,71), lo que indica mayores tasas cantonales de peatones y motociclistas lesionados o fallecidos por cada 1000 habitantes, mientras que los grupos 3 y 2 presentarían, en comparación con el componente Usuarios vulnerables, tasas cantonales decrecientes respectivamente (-0,80 y -0,14).

En cuanto al componente Vehículo privado, con base en los resultados del método K-Medias, el grupo 3 presenta una media considerablemente mayor (1,68) a la de los grupos 1 y 2 (-0,80 y -0,18).

Cuadro 5. Cantones agrupados en cada conglomerado generado a partir de la accidentabilidad de usuarios vulnerables (Dimensión 1) y vehículo privado (Dimensión 2).

Grupo	Observaciones	Cantones
1	18	San José, Escazú, Mora, Goicoechea, Santa Ana, Tibás, Montes de Oca, Dota, Curridabat, Alajuela, San Ramón, Naranjo, Palmares, Cartago, Heredia, Santo Domingo, Belén, Flores
2	33	Desamparados, Puriscal, Tarrazú, Aserri, Alajuelita, Coronado, Acosta, Moravia, Turrubares, Pérez Zeledón, León Cortés, Grecia, San Mateo, Poás, San Carlos, Valverde Vega, Los Chiles, Guatuso, Paraíso, La Unión, Turrialba, Alvarado, Oreámuño, El Guarco, Barva, Santa Bárbara, San Rafael, San Isidro, San Pablo, Liberia, Tilarán, Hojancha, Coto Brus
3	30	Atenas, Orotina, Alfaro Ruíz, Upala, Jiménez, Sarapiquí, Nicoya, Santa Cruz, Bagaces, Carillo, Cañas, Abangares, Nandayure, La Cruz, Puntarenas, Esparza, Buenos Aires, Montes de Oro, Osa, Aguirre, Golfito, Parrita, Corredores, Garabito, Limón, Pococí, Siquirres, Talamanca, Matina, Guácimo

Fuente: Elaboración propia.

Finalmente, se buscó caracterizar a las tres agrupaciones de cantones considerando variables descriptivas externas a la construcción de las dimensiones para conocer si existen también patrones identificables. Así, se realizó la comparación entre las características de los tres grupos generados para las dos dimensiones construidas y las variables que fueron elegidas para distinguir a los cantones.

El Cuadro 6 incluye los promedios de las variables externas, pero agrupadas en cada uno de los grupos con el fin de caracterizarlos a través de una medida de tendencia central sobre las variables de interés.

El Gráfico 3 presenta diagramas de caja que representan las series de datos cantonales de las variables, también reunidas en los grupos sugeridos. Dicho gráfico tiene el objetivo de mostrar la distribución de los datos de las variables externas en cada grupo, a simple vista. En el Gráfico 3 se observan distribuciones simétricas, donde la media de cada variable y su mediana son

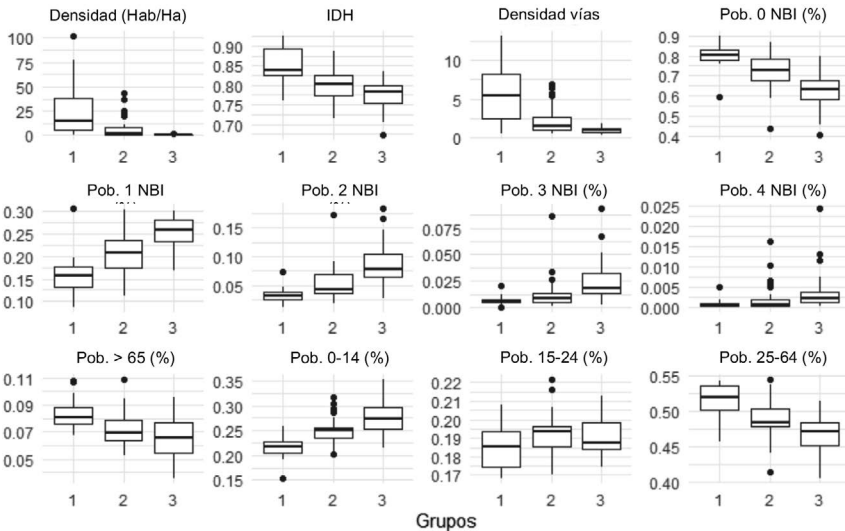
similares. Además, con base en la inspección de los rangos intercuartiles, comparativamente se observan pocas dispersiones resaltables en las variables en cada grupo (la densidad poblacional y la densidad de vías en los cantones del grupo), de los cuales predominan las bajas dispersiones.

Cuadro 6. Promedios de las variables de caracterización física, demográfica y socioeconómica según agrupamiento jerárquico de cantones en tres grupos, Costa Rica, 2018

Grupo	Km de vías por Km ²	Hab. por Ha.	Proporción poblacional por rango etario				Proporción poblacional por número de NBI					IDH
			0-14	15-24	25-64	>65	0	1	2	3	4	
1	5,57	24,99	0,22	0,19	0,51	0,08	0,80	0,16	0,03	0,01	0,00	0,85
2	2,26	6,47	0,25	0,19	0,49	0,07	0,73	0,21	0,05	0,01	0,00	0,80
3	0,94	0,63	0,28	0,19	0,47	0,07	0,63	0,25	0,09	0,03	0,00	0,77

Fuente: Elaboración propia.

Gráfico 3. Diagramas de cajas de las variables cantonales de caracterización física, demográfica y socioeconómica según agrupamiento jerárquico de 3 grupos



Fuente: Elaboración propia.

En el Cuadro 6 y en el Gráfico 3 sobresale que los cantones del grupo 1, caracterizados previamente con mayores valores medios en la Dimensión Usuarios vulnerables, también tienen condición de mayor densidad media poblacional (24,99 habitantes por Ha) y de vías (5,57 Km de vías por Km² de área). Además, pertenecen en su mayoría al Gran Área Metropolitana (ver la Figura 1), en su mayoría. Además, se trata del grupo con una mayor proporción media poblacional sin NBI y con el IDH medio más favorable (80%), a la vez que los cantones de este grupo cuentan con mayor población media adulta (51%) y adulta mayor (8%).

El grupo 2 corresponde a cantones que se encuentran geográficamente en un anillo alrededor de los cantones del grupo 1 y comparativamente cuentan con densidades intermedias. Por su parte, el grupo 3 presenta las densidades más bajas y agrupa a los cantones costeros.

El grupo 3 es sobresaliente en las medias calculadas para la Dimensión Vehículo privado, con mayor incidencia de heridas o muertes durante accidentes de tránsito a las personas en el rol de conductores. Los cantones del grupo 3 muestran condiciones sociodemográficas opuestas a los del grupo 1: menor IDH medio (0,77) y mayor proporción media de población con NBI (25% con 1 NBI), además de mayor proporción de población joven (28% son personas con edad entre 0 y 14 años).

Dado que la agrupación de los cantones del grupo 3 se basa en las tasas de accidentes de conductores con base en la población residente de cada cantón, se plantea la interrogante de si las características asociadas a partir de los resultados del Cuadro 6 y del Gráfico 3 estarán realmente relacionadas con los accidentes ocurridos en estas unidades geográficas o si corresponden a personas externas al cantón que fueron afectadas por accidentes ocurridos en estos cantones.

Discusión de resultados

Los resultados de este estudio son satisfactorios para el cumplimiento del objetivo planteado: el análisis exploratorio arrojó una agrupación de los cantones de Costa Rica en categorías con variabilidad homogénea dentro de cada grupo, pero heterogénea entre grupos, con base en los indicadores de cada unidad geográfica de exposición de los usuarios de las vías a presentar lesiones o fallecer.

La construcción de relaciones entre usuarios accidentados que presentaron lesiones o muerte y la proporción poblacional expuesta a accidentes de tránsito, para cada una de las unidades geográficas y en el periodo de estudio, permitió su combinación lineal para describir el conjunto original de indicadores en dos nuevas variables no correlacionadas: una asociada a los usuarios más vulnerables y otra a los usuarios conductores de vehículos privados.

El agrupamiento jerárquico de los cantones en tres conglomerados con base en las nuevas variables sustitutivas funcionó para unir grupos de los que se esperan similares patrones de ocurrencia de víctimas accidentadas en la vía. El grupo 1 de cantones que sobresale por presentar mayores tasas de peatones y motociclistas lesionados o fallecidos. A la vez, otra agrupación de cantones se destaca por las tasas comparativamente altas de personas conductoras lesionadas o fallecidas. Comparativamente, por ejemplo, entre los distritos con exceso de frecuencia de choques totales resaltados por Vargas Aguilar y Agüero Valverde (2020), destacan Alajuela, Río Segundo y San Joaquín, que pertenecen a los cantones Alajuela y Flores respectivamente. En el presente estudio ambos cantones pertenecen al grupo 1.

Fue posible delimitar un perfil común de cada grupo de unidades territoriales al considerar las variables demográficas, físicas y sociales propuestas. Se encontró que el grupo de cantones sobresalientes en la Dimensión Usuarios vulnerables son también comparativamente densos en población e infraestructura, con mayor desarrollo y con población mayor, mientras que el grupo de cantones que destacan en la Dimensión Vehículo privado tienen características opuestas.

Especialmente, se observan coincidencias entre estos hallazgos y las conclusiones de Hu, Wu, Huang, Peng y Liu (2020), quienes indican que las densidades de la población y de la red vial y la abundancia de servicios y actividades sociales están estrechamente relacionadas con la ocurrencia de choques de peatones. Dichos autores destacan las áreas con actividades y servicios sociales intensivos, como áreas comerciales y residenciales, con frecuencia de choques de peatones debido a las altas concentraciones de población y tráfico, así como áreas públicas en general. La concentración excesiva de servicios y actividades sociales ejercerá presión sobre la planificación histórica regional y de la red de carreteras.

Se ofrecen las siguientes recomendaciones que principalmente podrían ayudar a profundizar la aplicación de las técnicas utilizadas en esta temática particular, mediante la adición de otras variables relacionadas que pueden ser determinantes en la conformación de conglomerados y su caracterización.

El presente estudio podría complementarse en caso de tener información acerca de los viajes entre los diferentes cantones del país, o bien, de variables relacionadas con infraestructura medidas y estandarizadas a nivel cantonal (por ejemplo, que esos cantones tengan carreteras de mayor velocidad o carreteras de alto tránsito y velocidad que tengan más interacción con comunidades). Además, estudios adicionales podrían integrar los resultados obtenidos al agregar una mayor cantidad de variables socioeconómicas por cantón.

Es recomendable complementar el presente estudio con análisis que indiquen el cantón de residencia de las personas afectadas en colisiones de tránsito, con el fin de superar la limitación del presente estudio que plantea que la ubicación del evento de tránsito coincide con el cantón de residencia de las víctimas de estos eventos.

Para evaluar de manera más integral y efectiva las temáticas de densidad territorial y de infraestructura como determinantes en la incidencia de personas accidentadas, puede ser conveniente incluir variables asociadas al uso de suelo con el fin de contar con mejor una caracterización de las unidades geográficas, lejos del estudio de indicadores generales como los elegidos.

Es necesaria la generación de políticas públicas con el fin de mejorar la seguridad vial en los cantones de mayor riesgo. Por ejemplo, estudios como el de Jones, Lyons, John y Palmer (2005) muestran la efectividad de intervenciones de tráfico calmado que se asocia con reducciones absolutas en las lesiones de peatones infantiles tasas y reducciones en las desigualdades relativas en las tasas de lesiones de peatones infantiles.

Referencias

- Agüero Valverde, J. (2013). Multivariate spatial models of excess crash frequency at area level: Case of Costa Rica. *Accident Analysis & Prevention*. 59, 365-373.
- Agüero Valverde, J., y Jovanis, P. P. (2006). Spatial analysis of fatal and injury crashes in Pennsylvania. *Accident Analysis & Prevention*. 38 (3), 618-625.

- Brown, K. T. (2016). A safety analysis of spatial phenomena about the residences of drivers involved in crashes. en: https://tigerprints.clemson.edu/cgi/viewcontent.cgi?article=2747&context=all_dissertations
- Centro Centroamericano de Población. (2011). Estadísticas demográficas. 2011 – 2025. Proyecciones nacionales. CCP-INEC. <https://www.inec.cr/poblacion/estimaciones-y-proyecciones-de-poblacion>
- Consejo de Seguridad Vial. (2020). Datos para tabla interactiva de personas involucradas en accidentes con víctimas. <https://datosabiertos.csv.go.cr/datasets/193470/consolidado-de-personas-involucradas-en-accidentes-con-victimas/>
- Consejo de Seguridad Vial (CONAVI). (2021). Anuario estadístico de accidentes de tránsito con víctimas en Costa Rica 2020. <https://www.csv.go.cr/estad%C3%ADsticas>
- Cubbin, C., y Smith, G. (2002). Socioeconomic Inequalities in Injury: Critical Issues in Design and Analysis. *Annual Review of Public Health*. 23 (1), 349-375.
- Fleury, D., Peytavin, J. F., Alam, T. y Brenac, T. (2010). Excess accident risk among residents of deprived areas. *Accident Analysis & Prevention*. 42 (6), 1653-1660.
- Hartigan, J. A., y Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, 28(1), 100-108.
- Hashimoto, T. (2005). Spatial analysis of pedestrian accidents.
- Hewson, P. J. (2005). Epidemiology of child pedestrian casualty rates: can we assume spatial independence? *Accident Analysis & Prevention*. 37 (4), 651-659.
- Hu, L., Wu, X., Huang, J., Peng, Y. y Liu, W. (2020). Investigation of clusters and injuries in pedestrian crashes using GIS in Changsha, China. *Safety science*. 127, 104-105.
- Jaramillo, C., Lizárraga, C. y Grindlay, A. L. (2012). Spatial disparity in transport social needs and public transport provision in Santiago de Cali (Colombia). *Journal of Transport Geography*. 24, 340-357
- Instituto Geográfico Nacional. (2020). Nodos del SNIT: servicios WMS del Ministerio de Obras Públicas y Transportes. https://www.snitr.go.cr/ico_servicios_ogc_info?k=bm9kbzo6MTE=&nombre=MOPT

- Instituto Nacional de Estadística y Censos. (2015). Índice de Pobreza Multidimensional (IPM) - Metodología. INEC. https://www.inec.cr/sites/default/files/documentos/pobreza_y_presupuesto_de_hogares/pobreza/metodologias/mepobrezaenaho2015-01.pdf
- Instituto Nacional de Estadísticas y Censos. (2011). X Censo Nacional de Población y VI Censo Nacional de Vivienda de Costa Rica. <https://www.inec.cr/sistema-de-consultas>
- Jones, A. P., Haynes, R., Kennedy, V., Harvey, I. M., Jewell, T. y Lea, D. (2008). Geographical variations in mortality and morbidity from road traffic accidents in England and Wales. *Health & place*. 14 (3), 519-535.
- Jones, S. J., Lyons, R. A., John, A. y Palmer, S. R. (2005). Traffic calming policy can reduce inequalities in child pedestrian injuries: database study. *Injury Prevention*. 11 (3), 152-156.
- Lin, L., Shi, F. y Li, W. (2020). Assessing Road Traffic Safety During COVID-19: Inequality, Irregularity, and Severity. arXiv preprint arXiv:2011.02289.
- Magaña, J., Hernández H. y Jiménez, D. (2014). *Aplicación del Análisis de Conglomerados para la caracterización de factores temporales de tránsito para Costa Rica*. Congreso de Ingeniería Civil 2014. Colegio de Ingenieros Civiles.
- Martínez, P., Contreras, D. y Moreno, M. (2020). Safe mobility, socioeconomic inequalities, and aging: A 12-year multilevel interrupted time-series analysis of road traffic death rates in a Latin American country. *PLoS one*. 15 (1) 125-132.
- Ministerio de Salud. (30 de octubre de 2019). Se intensifica ofensiva contra accidentes de tránsito. Noticias. <https://www.ministeriodesalud.go.cr/>
- National Highway Traffic Safety Administration. (2021). Early estimates of motor vehicle traffic fatalities and fatality rate by sub-categories in 2020. (Crash•Stats Brief Statistical Summary. Report No. DOT HS 813 118).
- Open Street Map. (2020). Transport Map. <https://www.openstreetmap.org/export#map=17/9.93658/-84.06899&layers=T>.
- Organización Mundial de la Salud. (2018). Global Status Report on Road Safety. <https://www.who.int/publications/i/item/9789241565684>

- Pirdavani, A., Daniels, S., Van Vlierden, K., Brijs, K. y Kochan, B. (2017). Socioeconomic and sociodemographic inequalities and their association with road traffic injuries. *Journal of Transport & Health*. 4, 152-161.
- Poulos, R. G., Chong, S. S., Olivier, J. y Jalaludin, B. (2012). Geospatial analyses to prioritize public health interventions: a case study of pedestrian and pedal cycle injuries in New South Wales, Australia. *International journal of public health*. 57 (3), 467-475.
- Programa de las Naciones Unidas para el Desarrollo. (2020). Índice de Desarrollo Humano cantonal <https://www.cr.undp.org/content/costarica/es/home/atlas-de-desarrollo-humano-cantonal.html#:~:text=Como%20parte%20de%20la%20tendencia,más%20de%2032%2C5%25>.
- Ryb, G. E., Dischinger, P. C., Kufera, J. A. y Soderstrom, C. A. (2007). Social, behavioral, and driving characteristics of injured pedestrians: a comparison with other unintentional trauma patients. *Accident Analysis & Prevention*. 39 (2), 313-318.
- Shen, L., Lu, J., Long, M. y Chen, T. (2019). Identification of accident blackspots on rural roads using grid clustering and principal component clustering. *Mathematical Problems in Engineering*, 2019. 10, 75-82.
- Torres-Barragan, C. A., Cottrill, C. D. y Beecroft, M. (2020). Spatial inequalities and media representation of cycling safety in Bogotá, Colombia. *Transportation research interdisciplinary perspectives*. 7, 100-108
- Vargas Aguilar, D. y Agüero Valverde, J. (2020). Análisis estadístico de choques viales a nivel distrital para la gran área metropolitana. *Infraestructura vial*. 40, 1-10.
- Yang, Y., Chung, H. y Kim, J. S. (2021). Local or neighborhood? Examining the relationship between traffic accidents and land use using a gradient boosting machine learning method: the case of Suzhou industrial park, China. *Journal of advanced transportation*, 2021. 2, 1-30

