



Análisis estadístico del comportamiento de siniestros de tránsito por provincias del Ecuador, 2023

Statistical Analysis of Traffic Accident Behavior by Provinces of Ecuador, 2023

Iris Aracely Castillo Plaza, MSc*.

icastillo@uteq.edu.ec

Emma Yolanda Mendoza Vargas, MSc*.

emendoza@uteq.edu.ec

Eva Rosario Chávez Rojas, MSc.*

echavez@uteq.edu.ec

*Universidad Técnica Estatal de Quevedo, Ecuador

Recibido: 24/06/2024 – Aceptado: 21/09/2024

Correspondencia: icastillo@uteq.edu.ec

RESUMEN

El artículo titulado centra en el análisis de los accidentes de tránsito en las provincias ecuatorianas durante el año 2023. Utilizando métodos estadísticos multivariados como el análisis de componentes principales (PCA) y su variante robusta MacroPCA, se busca identificar patrones y correlaciones entre diferentes tipos de vehículos involucrados en los accidentes y las provincias del Ecuador. El estudio revela que las provincias de Guayas y Pichincha muestran una alta incidencia de siniestros con valores atípicos en la mayoría de los vehículos analizados, mientras que provincias como Galápagos, Esmeraldas, Carchi, Cañar y Bolívar no presentan atipicidades significativas. También se destaca que los siniestros de tránsito en Ecuador afectan mayoritariamente a automóviles, motocicletas y vehículos no identificados. Las provincias con mayor riesgo, según el análisis, son Pichincha y Guayas, estas necesitan políticas de seguridad vial específicas debido a la alta variabilidad en los tipos de vehículos involucrados en los siniestros.

Palabras claves: siniestros de tránsito, estadísticos multivariante, valores atípicos, PCA, MacroPCA.

ABSTRACT

This article focuses on the analysis of traffic accidents in Ecuadorian provinces during the year 2023. Using multivariate statistical methods such as principal component analysis (PCA) and its robust variant MacroPCA, we seek to identify patterns and correlations between different types of vehicles involved in accidents and the provinces of Ecuador. The study reveals that the provinces of Guayas and Pichincha show a high incidence of accidents with outliers in most of the vehicles analyzed, while provinces such as Galapagos, Esmeraldas, Carchi, Cañar and Bolivar do not show significant atypicalities. It is also noted that most traffic accidents in Ecuador involve automobiles, motorcycles and unidentified vehicles. The provinces with the highest risk, according to the analysis, are Pichincha and Guayas, which need specific road safety policies due to the high variability in the types of vehicles involved in crashes.

Keywords: traffic crashes, multivariate statistics, outliers, PCA, MacroPCA.

Cómo citar

Castillo Plaza, I. A., Emma Yolanda, & Chávez Rojas, E. R. (2024). Análisis estadístico del comportamiento de siniestros de tránsito por provincias del Ecuador, 2023. *GADE: Revista Científica*, 4(4), 247-259. Recuperado a partir de <https://revista.redgade.com/index.php/Gade/article/view/503>



INTRODUCCIÓN

Nuestra investigación se basa en el análisis de los accidentes de tránsito y su incidencia en la vida normal de las personas. A continuación, exponemos algunos conceptos respecto a dicha envergadura, como, por ejemplo:

Accidentes de tránsito: Son incidentes que ocurren en la vía pública y que involucran al menos un vehículo en movimiento, ya sea un automóvil, motocicleta, bicicleta, camión, entre otros. Los accidentes de tránsito pueden variar en gravedad, desde colisiones menores hasta choques graves que involucran lesiones serias o incluso la muerte. Los tipos comunes de accidentes incluyen:

Colisiones frontales: Dos vehículos chocan de frente.

Colisiones laterales: Un vehículo golpea el costado de otro, también conocido como "t-bone".

Colisiones por alcance: Un vehículo impacta a otro por detrás.

Atropellos: Un vehículo impacta a un peatón.

Salidas de carretera: Un vehículo se sale de la carretera y puede volcar o chocar con objetos en la vía.

Impacto en el ser humano:

Físico: Los accidentes de tránsito pueden causar una amplia gama de lesiones,

desde contusiones y esguinces hasta fracturas, lesiones internas y traumatismos cerebrales. La gravedad de las lesiones puede variar dependiendo de factores como la velocidad del impacto, el uso de cinturones de seguridad y la infraestructura de seguridad del vehículo.

Psicológico: Las víctimas de accidentes pueden experimentar estrés postraumático, ansiedad, depresión y otras afecciones psicológicas. Incluso los testigos de un accidente o los familiares de los involucrados pueden verse afectados emocionalmente.

Económico: Los costos asociados con los accidentes de tránsito pueden ser significativos. Estos incluyen gastos médicos, pérdida de ingresos debido a incapacidades, reparación de vehículos y costos legales. Además, las aseguradoras pueden incrementar las primas de seguro tras un accidente.

Social: Los accidentes pueden afectar las relaciones sociales y familiares, ya sea por el estrés que causa a las familias o por las posibles alteraciones en la vida cotidiana debido a lesiones o pérdidas.

Legal: Los accidentes de tránsito pueden llevar a conflictos legales, como demandas por daños y perjuicios, multas y sanciones, y problemas con el sistema



de justicia si hay responsabilidades legales involucradas.

En general, los accidentes de tránsito tienen un impacto profundo y multifacético en la vida de las personas, y es fundamental adoptar medidas preventivas para reducir su frecuencia y severidad, como el uso del cinturón de seguridad, el respeto a las señales de tránsito y la conducción responsable.

Aproximadamente 1,19 millones de personas mueren cada año como consecuencia de accidentes de tráfico, siendo las lesiones causadas por el tránsito la principal causa de muerte en niños y adultos jóvenes de entre 5 y 29 años. El 92% de las muertes en las carreteras del mundo ocurren en países de ingresos bajos y medios, a pesar de que estos países tienen alrededor del 60% de los vehículos del mundo. (Organización Mundial de la Salud, 2023).

Más de la mitad de todas las muertes por accidentes de tránsito se producen entre usuarios vulnerables de la vía, incluidos peatones, ciclistas y motociclistas. Estos accidentes cuestan a la mayoría de los países el 3% de su producto interno bruto. En respuesta a esta situación, la Asamblea General de las Naciones Unidas ha establecido el ambicioso objetivo de reducir a la mitad el número

mundial de muertes y lesiones por accidentes de tránsito para 2030. (Organización Mundial de la Salud, 2023).

En Ecuador, la Agencia Nacional de Tránsito (ANT), es una entidad pública encargada de planificar, regular y controlar la gestión del Transporte Terrestre, Tránsito y Seguridad Vial en el territorio nacional, a fin de garantizar la libre y segura movilidad terrestre, prestando servicios de calidad que satisfagan la demanda ciudadana; coadyuvando a la preservación del medio ambiente y contribuyendo al desarrollo del País, en el ámbito de su competencia (Agencia Nacional de Tránsito, s.f.).

La agencia nacional de tránsito define a los siniestros de tránsito como: “Todo suceso eventual o acción involuntaria, que como efecto de una o más causas y con independencia del grado de estas, ocurre en vías o lugares destinados al uso público o privado, ocasionando personas muertas, individuos con lesiones de diversa gravedad o naturaleza y daños materiales en vehículos, vías o infraestructura, con la participación de los usuarios de la vía, vehículo, vía y/o entorno”. (ANT, 2021).

Los accidentes de tránsito en Ecuador dejan, en promedio, seis fallecidos cada



día. Según la Agencia Nacional de Tránsito, entre el 1 de enero y el 31 de octubre de 2023, 1.942 personas perdieron la vida en 17.257 siniestros de este tipo, que también dejaron 15.357 heridos. El 73 % de los afectados son varones y las edades oscilan entre 18 y 29 años con 2,054 personas.

METODOLOGÍA

El estudio se centra en el análisis de siniestros de tránsito en las provincias de Ecuador, utilizando una metodología estadística multivariada que incluye análisis de *DetectDeviatingCells* (DDC), análisis de componentes principales (PCA), y su variante robusta, MacroPCA. El objetivo es identificar patrones de atipicidad y correlaciones entre diferentes tipos de vehículos implicados en los accidentes, así como analizar la variabilidad y representatividad de las provincias en el contexto.

La fuente de información es la base de datos de siniestros de tránsito nacional del Ecuador. Para el análisis se considerará el número de casos por provincia y por tipo de vehículo involucrado. Está conformada por 24 filas que hacen referencia a las provincias del Ecuador y 13 variables en relación con los vehículos involucrados

en los siniestros en el año 2023.

Las variables de los vehículos involucrados en los siniestros son:

Automóvil. Vehículo liviano destinado al transporte de un reducido número de personas.

Bus. Vehículo automotor diseñado para el transporte de pasajeros compuesto por un chasis y una carrocería acondicionada para el transporte de pasajeros con una capacidad desde 36 asientos incluido el conductor.

Camioneta. Vehículo a motor construido para el transporte de carga, con capacidad de hasta 3.500 Kg.

Especial. Vehículos que pertenecen a las categorías M, N u O destinados al transporte de pasajeros o mercancías.

Motocicleta. Vehículo automóvil de dos, tres o cuatro ruedas cuya masa en vacío no excede de 400 kg. de peso. Se incluyen los vehículos con una cilindrada inferior a 50cc. si no están incluidos en la definición de ciclomotor

Bicicleta. Vehículo de tracción humana de dos o más ruedas en línea.

Camión. Vehículo a motor construido especialmente para el transporte de carga, con capacidad de más de 3.500 Kg.

Emergencias. El perteneciente a la Policía Nacional o al Cuerpo de Bomberos y las ambulancias de las



instituciones públicas o privadas.

Furgoneta. Vehículo ligero diseñado para el transporte de pasajeros y mercancías, compuesto por una superestructura integral entre el chasis y la carrocería.

Vehículo deportivo utilitario. Vehículo utilitario fabricado con carrocería cerrada o abierta, con techo fijo o desmontable y rígido o flexible. Para cuatro o más asientos en por lo menos dos filas.

También, *Scooter Eléctrico, Tricimoto, No Identificado.*

DetectDeviatingCells (DDC)

La detección de celdas atípicas, este método fue desarrollado por Peter J. Rousseeuw (PhD en Matemáticas, experto en estadística robusta y métodos de regresión). El método DDC asume que las filas x_i son generadas de una distribución normal multivariante con vector de medias d -variante μ y matriz de covarianza Σ positiva semidefinida que son desconocidos, luego de lo cual se asume que algunas celdas fueron contaminadas o su información se perdió. (Rousseeuw & Leroy, 2003)

Algoritmo DetectDeviatingCells

Disponible en el paquete R cellwise en CRAN

Paso 1: estandarización, para cada

columna j de X estimamos:

$$m_j = \text{robPos}_i(x_{ij}) \text{ y } S_j =$$

$\text{robEsc}_i(x_{ij} - m_j)$ donde robPos es un estimador robusto de posición (como por ej. La mediana muestral) y robEsc es un estimador robusto de escala (como por ej. el MAD).

Luego, estandarizamos X a Z según

$$z_{ij} = (x_{ij} - m_j) / S_j$$

Paso 2: detección de outliers univariante, definimos una nueva matriz

U con entradas:

$$u_{ij} = \begin{cases} z_{ij}, & \text{si } |z_{ij}| \leq c \\ NA, & \text{si } |z_{ij}| > c \end{cases}$$

El punto crítico es $c = \sqrt{\chi_{1,p}^2}$ donde la probabilidad p es igual a 99% por defecto.

Paso 3: relaciones bivariantes, para dos columnas cualesquiera $h \neq j$ calculamos:

$$\text{cor}_{jh} = \text{robCorr}_i(u_{ij}, u_{ij})$$

donde robCorr es un estimador de correlación robusto. Ver (Rousseeuw et al., 2017).

Solo usamos la relación entre variables j y h cuando

$$|\text{cor}_{jh}| \geq \text{corrlim}$$

corrlim se establece por default en 0,5. Las variables j que satisfagan la condición para cada $h \neq j$ serán llamados conectores. El resto de variables serán autónomas.



Para los pares (j,h) que cumplan con la condición

$$|cor_{jh}| \geq corrlim \text{ se calcula } b_{jh} = robPend_i(u_{ij}|u_{ih})$$

Donde robPend calcula la pendiente de una recta de regresión robusta sin término de intercepto que predice la variable j a partir de la variable h .

Paso 4: valores de predicción, calculamos las predicciones \hat{z}_{ij} para todas las celdas. Para cada variable j consideramos el conjunto H_j que consiste de todas las variables h que cumplen con la condición $|cor_{jh}| \geq corrlim$ incluyendo la misma variable j .

Para todos $i = 1, \dots, n$ obtenemos:

$$\hat{z}_{ij} = \frac{\sum_h w_{jh} b_{jh} u_{ih}}{\sum_h w_{jh}}$$

donde $w_{jh} = |cor_{jh}|$. Note que otras opciones son posibles, como por ejemplo una mediana ponderada.

Paso 5: alargamiento, calcular \hat{z}_{ij} a menudo encoge la escala de las entradas, lo cual no es deseable. Para contrarrestar este encogimiento reemplazamos por:

$$\hat{z}_{ij} robPend_i(z_{ij} | \hat{z}_{ij})$$

para todas las variables conectadas j .

Paso 6: marcar celdas (potencialmente atípica, calcule los residuos en base a las celdas estandarizadas:

$$r_{ij} = \frac{z_{ij} - \hat{z}_{ij}}{rodEsc_i(z_{ij} - \hat{z}_{ij})}$$

En cada columna j marcamos como celdas potencialmente atípicas a todas las celdas que cumplan con la condición $|r_{ij}| > c$.

Luego, ensamblamos la matriz 'imputada' Z_{imp} dada por:

$$(Z_{imp})_{ij} = \begin{cases} \hat{z}_{ij}, & \text{si } z_{ij} \text{ fue marcada o es un NA} \\ z_{ij}, & \text{de otra manera} \end{cases}$$

Paso 7: Marcar filas (potencialmente atípicas, el método DDC también puede marcar alguna fila (potencialmente) atípica en base a los residuos por celdas estandarizados r_{ij} .

Para datos Gaussianos multivariados sin valores atípicos tenemos que $r_{ij} \approx N(0,1)$ de tal manera que la función de distribución acumulativa de r_{ij}^2 es aproximadamente la función de distribución acumulativa F de χ_1^2 . Esto nos conlleva al criterio:

$$T_i = \text{promedio}_{j=1}^d F(F_{ij}^2)$$

realizamos luego una estandarización robusta de T_i y marcamos las filas i para los cuales el T_i estandarizado excede el punto crítico $\sqrt{\chi_{1,p}^2}$

Paso 8: desestandarizar, nos deshacemos de la estandarización del



paso 1.

MacroPCA permite estimar las componentes principales en datos con valores perdidos y que pueden contener celdas y filas atípicas. Además, MacroPCA permite detectar valores atípicos con respecto a las componentes principales: por celdas y por filas en la matriz de datos multivariante, y, dentro del subespacio (con respecto a los “scores” robustos). (Cevallos-Valdiviezo, 2020)

En matrices de datos, pueden existir tanto filas como celdas atípicas, y mientras que el PCA clásico es vulnerable a estos valores atípicos, especialmente en datos de gran dimensión, se sugiere el uso de métodos robustos que sean resistentes a dichas anomalías. La comparación entre los ajustes obtenidos mediante ACP clásico y robusto es crucial, ya que diferencias significativas pueden señalar problemas en los datos, los cuales pueden ser luego investigados e interpretados. (Cevallos-Valdiviezo, 2020)

RESULTADOS

Con la matriz muestral se realiza análisis de DDC, para observar la atipicidad de las variables. Las provincias de Guayas, Pichincha muestran valores atípicos altos en 10 de los 13 vehículos

estudiados. En Santo Domingo de los Tsáchilas, se observa que los siniestros de tránsito se dan con mayor incidencia en bicicleta, bus, camión, furgoneta, motocicleta y en vehículo no identificado. Las provincias de Galápagos, Esmeraldas, Carchi, Cañar y Bolívar no evidencian valores atípicos en los siniestros de tránsito.

Los Ríos, tiende a tener siniestros automovilísticos por debajo de la media. En las provincias del oriente ecuatoriano no se observan valores atípicos.

Mapa de celdas provincias por tipo de vehículo involucrado en siniestro

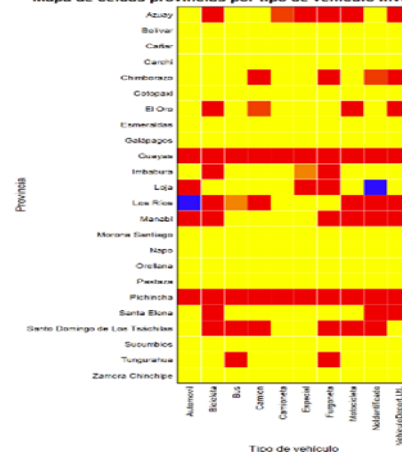


Gráfico 1

Análisis Exploratorio.

Según el tipo de vehículo, las medias más representativas de siniestros son: En Ecuador el mayor número de siniestros se dan en automóviles (media de 454.5), seguidos por motocicletas (media de 311.2) y en vehículos no identificados (media de 236.5); los agentes cuando llegan al lugar no encuentran implicados.

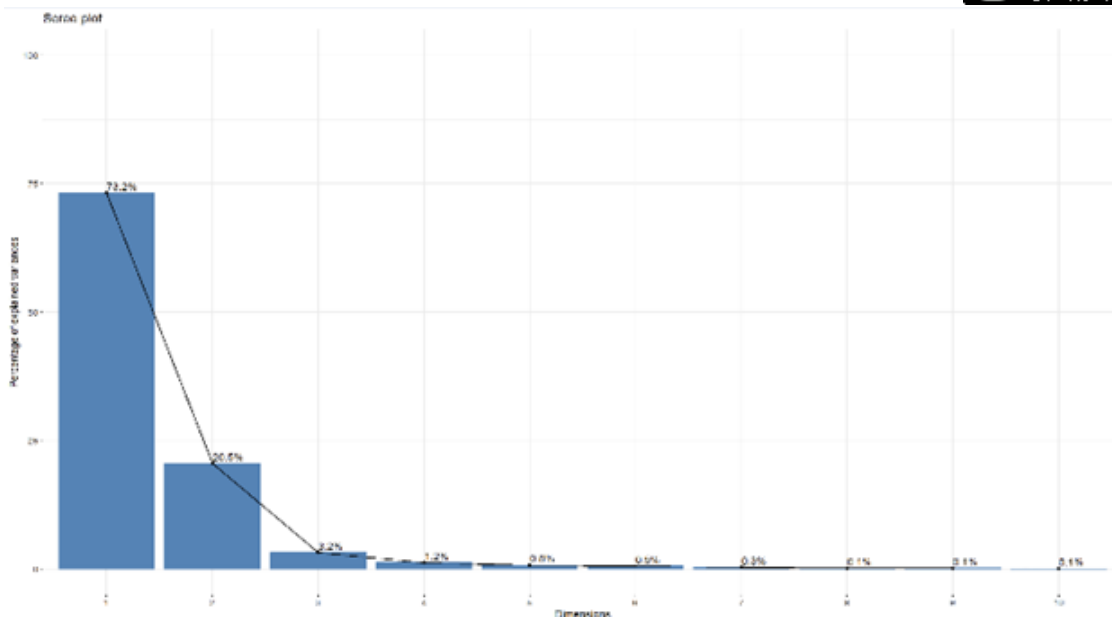


Gráfico 2

Análisis PCA

En el Gráfico 2, se ilustra que las dos primeras componentes principales explican aproximadamente el 93% de la

varianza. La primera componente principal captura el 73.19% de la varianza en los datos, la segunda componente el 20.53%.

PCA_Biplot

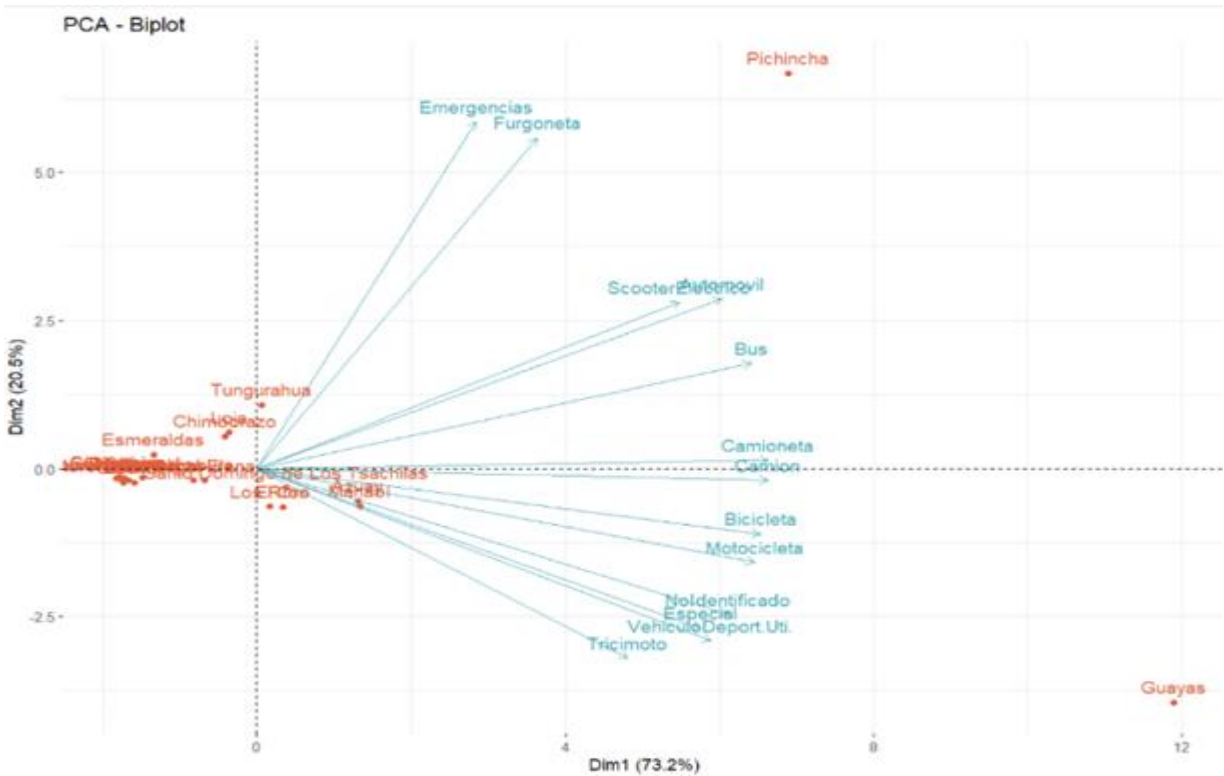


Gráfico 3



En el PCA_Biplot, Pichincha y Guayas, están más alejadas en el eje de la dimensión 1, lo que podría indicar diferencias en la composición de los siniestros de tránsito comparadas con otras provincias que se agrupan cerca del origen. Las provincias de Esmeraldas, Santa Elena, Loja, Chimborazo y algunas provincias del Oriente tienen menos siniestros de tránsito.

En la provincia de Pichincha los siniestros de tránsito por tipo de vehículo “Emergencia” y “Furgoneta” tienen una variabilidad significativa en los datos. Las variables camioneta y camión, bicicleta y motocicleta, vehículos deportivos utilitarios y especial tienen mayor correlación en el conjunto de datos analizado.

Análisis MacroPCA

Aplicando MacroPCA la matriz reducida muestra 23 provincias y 10 variables. Haciendo el análisis de celdas con valores atípicos con respecto al ajuste de los componentes principales, la provincia de Loja muestra que los siniestros por automóvil y vehículos especiales están por encima de la media; bicicleta tiene valores atípicos con tendencia a la baja, así como los siniestros en vehículos no identificados. En Pichincha, atipicidad se da en bicicleta y furgonetas con valores altos.

Los siniestros en furgonetas tienen registros bajos en las provincias de Cotopaxi, Esmeraldas y Sucumbios. En algunas provincias como Bolívar, Cañar, Carchi, Esmeraldas tienen valores perdidos en los siniestros en bicicleta. En las provincias del Oriente existen valores perdidos en los siniestros en furgoneta, a excepción de Sucumbios.

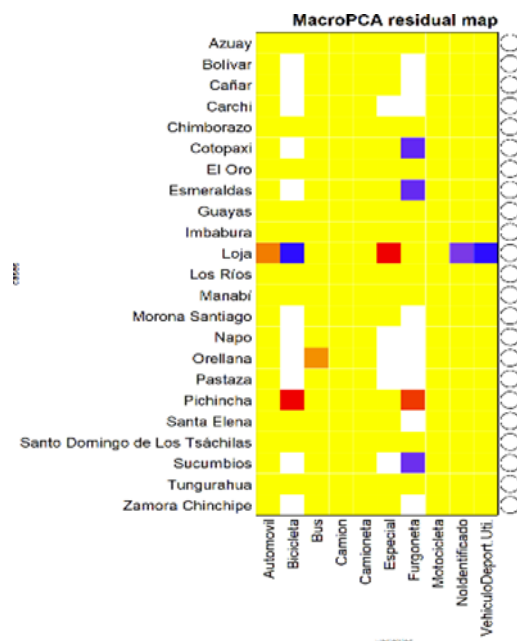


Gráfico 4

MacroPCA, outlier map

En el MacroPCA outlier map se observa que la provincia de Pichincha es atípica de manera ortogonal y Score distance, presentando características muy distintas en los siniestros de tránsito con relación a otras provincias. Ortogonalmente, son atípicos Los Ríos, Loja, Sucumbios, Cotopaxi, Esmeraldas y El Oro. Guayas muestra ser atípica de manera Score distance, está alejada significativamente.

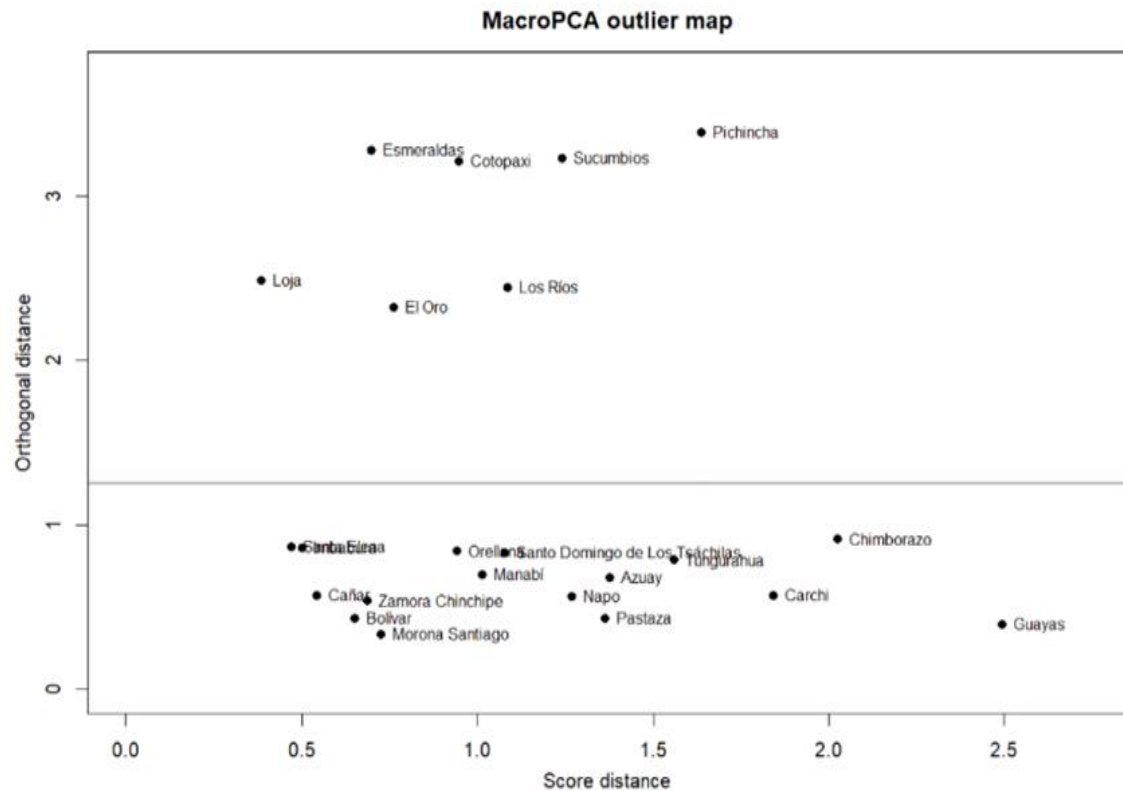


Gráfico 5

PCA_Biplot robusto

El PCA_Biplot Robusto, en el eje horizontal (Dim1) captura el 80.3% de la varianza y el eje vertical (Dim2) el 10.3%. Existe correlación entre las variables analizadas, a excepción de furgoneta. Automóvil está cercano a Camioneta, lo que implica que estos dos

tipos de vehículos comparten características similares en los datos de siniestros.

La provincia de Pichincha indica una alta cantidad de siniestros a causa del tipo de vehículo “Bicicleta”; Guayas, los siniestros en su mayoría son por: automóvil, bus, camioneta, camión.

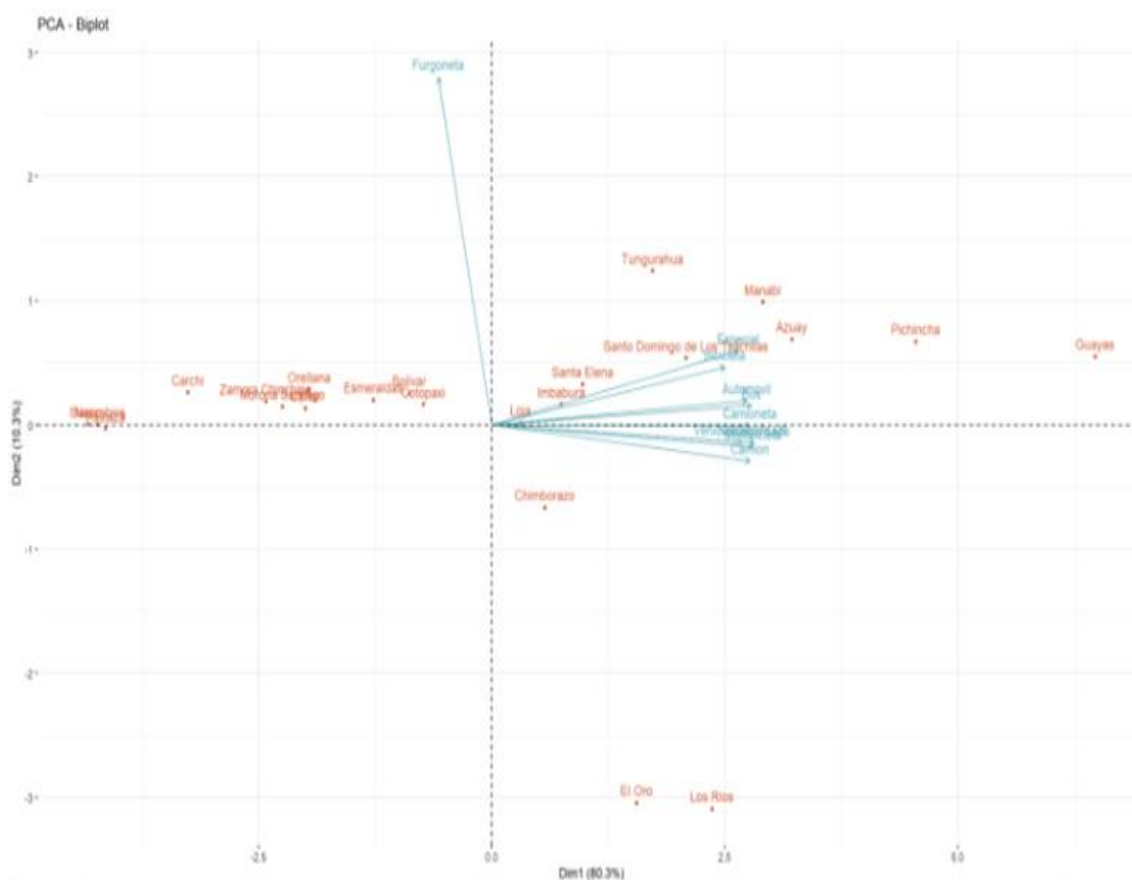


Gráfico 6

Análisis comparativo PCA_Biplot método clásico Vs PCA_Biplot Robusto.

El PCA clásico, la primera dimensión (Dim.1) explica el 73.19% de la varianza, lo que es bastante alto. El Método MacroPCA sugiere seleccionar la primera componente principal ($k = 1$), ya que incluso la primera componente ya explica más del 80% de la varianza por sí sola.

Análisis comparativo del DDC matriz muestral Vs MacroPCA residual.

En MacroPCA, reduce las provincias a un número de 23 y ambos análisis

mantienen el mismo número de variables, es decir, 10 tipos de vehículos donde suceden los siniestros.

CONCLUSIONES

Una de las limitaciones de este estudio es que la base proporcionada por la Dirección Nacional de Control de Tránsito y Seguridad Vial de la Policía Nacional del Ecuador no cuenta con información más completa de las personas accidentadas y las condiciones del accidente, así como del tipo de lesiones que tuvieron los involucrados, el nivel de educación de los



accidentados, la calidad de la vía, el tipo de unión de la vía, entre otros. El contar con esta información permitiría realizar análisis más completos de los factores que intervienen en un siniestro. De esta forma, se recomienda que estos datos sean recolectados con el fin de realizar análisis más completos. (Castillo, Coral, & Salazar Méndez, 2020)

Los accidentes de tránsito son la principal causa de muerte en niños y adolescentes en Ecuador. Un informe del Instituto Ecuatoriano de Estadística y Censos (INEC) detalla que en 2022 estos siniestros provocaron la muerte de 195 personas de entre 5 y 17 años. (Primicias, 2023)

En el estudio, el análisis de DDC identifica que las provincias de Guayas y Pichincha presentan alta atipicidad en la mayoría de los medios de transporte, señalando una mayor incidencia de siniestros de tránsito. Santo Domingo de los Tsáchilas también muestra un riesgo elevado en varios tipos de vehículos. En contraste, las provincias del Oriente ecuatoriano, Galápagos, Esmeraldas, Carchi, Cañar y Bolívar no presentan valores atípicos significativos, mientras que Los Ríos registra siniestros automovilísticos por debajo de la media. Estos hallazgos subrayan la necesidad de medidas de seguridad vial específicas.

Los datos revelan una relación entre el tipo de vehículo involucrado en los siniestros y la provincia donde estos ocurren, resaltando patrones distintos en las provincias más pobladas, como Pichincha y Guayas. En particular, Pichincha muestra una prevalencia de incidentes con automóviles y camionetas, mientras que en Guayas se destacan las motocicletas y vehículos no identificados en los accidentes. El método MacroPCA, efectivo en la gestión de datos atípicos y perdidos; resalta la provincia de Loja por siniestros de automóviles y vehículos especiales sobre la media, y a Pichincha con atipicidades en bicicletas y furgonetas. Estos resultados sugieren la necesidad de políticas de seguridad vial específicas para las provincias con mayor riesgo.

REFERENCIAS

- ANT. (2021). Reportes Nacionales, Glosario de términos. Siniestros de Tránsito. Obtenido de <https://www.ant.gob.ec/estadisticas-siniestrosde-transito/>.
- Agencia Nacional de Tránsito - Dirección de Estudios y Proyectos (2011). Manual para el llenado correcto de la matriz mensual de siniestros de tránsito. Quito, Pichincha, Ecuador.



- ANT. (2021). Reportes Nacionales, Glosario de términos. Siniestros de Tránsito. Obtenido de <https://www.ant.gob.ec/estadisticas-siniestrosde-transito/>.
- Castillo, D., Coral, C., & Salazar Méndez, Y. (2020). Modelización Econométrica de los Accidentes de Tránsito en el Ecuador.
- Cevallos-Valdiviezo, H. (2020). Estimación de componentes principales resistente a valores atípicos: una revisión de la literatura. *LACCEI International Multi-Conference for Engineering, Education, and Technology*.
- Organización Mundial de la Salud. (diciembre de 2023). Organización, Lesiones causas por el tránsito. Obtenido de <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>.
- Primicias. (2023). Accidentes, asesinatos y suicidios, las principales causas de muerte de niños y adolescentes. www.primicias.ec/noticias/sociad/accidentes
- Rousseeuw et al. (2017). Detecting Deviating Data Cells. Belgium: Department of Mathematics, KU Leuven.
- Rousseeuw, P., & Leroy, A. (2003). *Regresión robusta y detección de valores atípicos*. Nueva York.
- Raymaekers, J., Rousseeuw, P.J., and Van den Bossche, W. (2018), “cellWise,” CRAN, <https://CRAN.R-project.org/package=cellWise>.