Aporte de las estadísticas de accidentes de tránsito no laborales en la posible disminución de afectaciones empresariales en el departamento del Putumayo.

Contribution of non-work traffic accident statistics in the possible reduction of business impacts in the department of Putumayo.

Estudiante:
Lize Sambony Perez
lsambony@libertadores.edu.co
Facultad de Ingeniería y Ciencias Básicas
Fundación Universitaria Los Libertadores
Bogotá D.C

Asesor:
Manuel Francisco Romero Ospina, M.Sc
mfromero@libertadores.edu.co
Líder Servicios Transversales de Estadística
Facultad de Ingeniería y Ciencias Básicas
Docente

Receiver, December, 2021

RESUMEN. El siguiente trabajo busca comparar los individuos o unidades estadísticas mediante el uso de técnicas de técnicas de Estadística Descriptiva Multivariada en su versión descriptiva. Es sabido que estas técnicas y métodos hacen parte del aprendizaje no supervisado, lo que hoy en día esta en continua evolución a la ciencia de Datos. Para ello se parte de una base o conjunto de datos, por medio del cual se pueda estudiar la realidad de los accidentes de tránsito con consecuencias fatales en el departamento de Putumayo. Los datos utilizados corresponden a las bases de datos contenidas en la secretaria de Salud departamental de Putumayo cargada en el Registro Único de Afiliados RUAF-ND, para lo cual se cuenta con información de los periodos comprendidos entre los años 2019 y 2020, a partir de 8 variables (Municipio, Área de Defunción, Sitio de Defunción, Sexo del Fallecido, Estado Civil, Edad, Nivel Educativo y Régimen de Seguridad) y 854 datos por cada variable para un total de 195.239 datos. El procedimiento estadístico se realiza mediante el estudio

descriptivo de las variables, donde se incluye la selección, depuración, concatenación y transformación. La base de datos cuenta con variables de tipo Categórico (Nominales, Ordinales, de intervalo y razón) y la variable Edad cómo numérica. Por esta razón, se parte de la base que algunas variables del caso serán sometidas a análisis siguiendo los niveles de investigación cómo son: Exploratorio y Descriptivo multivariados. En algunas oportunidades se realiza descripciones de tipo bivariados y en otros casos multivariadas para establecer la relación entre variables.

Palabras clave: Estadística Descriptiva Multivariada, variables Categóricas y variables numéricas, depuración, concatenación, transformación de variables y Accidentes de Tránsito.

ABSTRACT. The following work seeks to compare the individuals or statistical units by using the techniques of multivariate descriptive statistics techniques in their descriptive version. It is known that these techniques and methods are part of unsupervised learning, which today is in continuous evolution to data science. To do this, we start from a database or set of data, through which the reality of traffic accidents with fatal consequences in the department of Putumayo can be studied. The data used correspond to the databases contained in the Putumayo departmental health secretary uploaded in the RUAF-ND single register of affiliates, for which there is information on the periods between 2019 and 2020, as of 7 variables (area of death, place of death, sex of the deceased, marital status, age, educational level and security regime) and 854 data for each variable for a total of 195,239 data. The statistical procedure is carried out by means of the descriptive study of the variables, which includes the selection, debugging, concatenation and transformation. The database has variables of the categorical type (nominal, ordinal, interval and ratio) and the age variable as numerical. For this reason, it is based on the basis that some variables of the case will be subjected to analysis following the levels of investigation such as: exploratory and descriptive multivariate. On some occasions bivariate descriptions are made and in other cases multivariate to establish the relationship between variables.

Keywords: Multivariate Descriptive Statistics, Categorical variables and numerical variables, debugging, concatenation, transformation of variables and traffic accidents.

1. INTRODUCCIÓN

Durante los últimos años el número de accidentes de tránsito se ha incrementado, exceptuando los acontecidos durante el año 2020, que por los meses de aislamiento preventivo obligatorio se vieron un poco disminuidos. Para el año 2019 se reportaron cerca de 7000 muertes, mientras que otras 35000 personas sufrieron lesiones graves por este tipo de accidentes. Para el año 2020, durante los meses de enero a octubre se reportaron en total 4156 personas fallecidas (Portafolio, 2020).

Estas pérdidas humanas, además, de afectar a las familias y generar costos de aproximadamente 23,9 billones de pesos al país, afectan a las empresas, que no se encontraban preparadas para estás pérdidas en el caso de su personal, razón por la cual muy pocas veces se preocupan por estas cifras y por ende tienen poca preparación para afrontar posibles pérdidas de miembros de sus equipos que utilizan algún tipo de transporte vehicular o se movilizan por las diferentes calles o vías.

Es por esto que se evalúa las estadísticas de accidentalidad vial en el departamento del Putumayo con relación a las variables de contempladas en el estudio, este documento puede convertirse en apoyo a los diversos instrumentos técnicos claves a la hora de formular estrategias que permitan al departamento del Putumayo y a los gobiernos locales contar con una herramienta decisoria efectiva para apoyar las situaciones que pueden alterar la seguridad vial de los conductores, transeúntes y habitantes de las regiones del departamento, a su vez, se convierte en un instrumento de gestión empresarial por medio del cual los empresarios y/o empleadores pueden determinar las posibles afectaciones que podrían sufrir las empresas por pérdidas de su personal, y así proponer los mecanismos de prevención que deberían hacer parte de los sistemas de gestión de seguridad y salud en el trabajo, para contribuir a prevenirlos a través de la concientización, educación vial y legal que debería asumir los empleados cuando utilicen transportes particulares como autos, motocicletas o bicicletas.

De acuerdo al informe municipal de seguridad y convivencia ciudadana, presentado por el Departamento Nacional de Planeación en abril de 2020, en Mocoa, Putumayo, la tasa de lesiones en accidentes de tránsito por cada 100.000 habitantes a nivel departamental, municipal y ciudades intermedias es de 39% para el departamento, 125% es la tasa a nivel de municipios y 53.6% a nivel de ciudades intermedias. Asocapitales. (Abril, 2020). https://www.asocapitales.co/nueva/wp-content/uploads/2020/06/Putumayo-Mocoa.pdf.

En Mocoa, se registraron 362 casos de accidentes de transito entre los periodos 2014 y 2019, incrementándose un 7.3% entre ambos periodos, representando el 0.149% de los casos registrados a nivel nacional durante el 2019 y concentrando este municipio el 40.14% de los casos registrados a nivel departamental.

Evitar riesgos podría considerarse como una contribución a la disminución de accidentalidad vial y así mantener estable en la prestación de servicios de salud que se vio afectada sin duda alguna en época de cuarentena por SARS-COVID 19 y evitando el riesgo de pérdida de la capacidad laboral y productividad empresarial, que es altamente afectada por la pérdida de colaboradores de manera inesperada.

Es por ello, que con el presente trabajo se pretende desarrollar un documento base que permita identificar las relaciones existentes entre variables interés y establecer prioridades al momento de tomar decisiones gubernamentales en pro de la seguridad vial del departamento del Putumayo.

2. REFERENTES TEORICOS

Se han utilizado las técnicas del análisis estadístico multivariado (EDM) para encontrar o identificar patrones que no se pueden observar fácilmente en bases de datos, asociadas a conjuntos de individuos u objetos. Es sabido que el conjunto de individuos u objetos pueden disponerse en arreglo matricial, es decir, tablas de *n* filas y *p* columnas, como nubes de puntos. Las filas como *n* vectores en $\mathbb{R}^p$ y las columnas como *p* en vectores en $\mathbb{R}^n$. Estas representaciones pueden sacar a flote contenido oculto dentro de la base de datos. (Lebart et al., 2006).

A continuación, se presentan algunos antecedentes teóricos donde se utilizan técnicas de Estadística Descriptiva Multivariada para este tipo de estudios:

De los accidentes de Tránsito ocurridos y registrados en el Municipio de Pereira durante los años 2007 y 2008, se tomaron los ocurridos en la zona urbana, específicamente en los que el vehículo causante (implicado) es de tipo liviano (automóvil, microbús, etc.) ó motocicletas; según información existente en el Instituto Municipal de Tránsito y Transporte del Municipio de Pereira y en el Instituto Nacional de Medicina Legal. Inicialmente se realizó el análisis de Correspondencia en el programa SPSS 13.0 a las variables de manera individual, así como a la totalidad de las variables existentes, encontrando como resultado en ambos casos que varias de ellas no son significativas, por lo tanto, de acuerdo a este análisis se aplicó nuevamente la Regresión Logística con las variables que tenían una significancia menor al 5%. (González y Ríos, 2010).

(Tomalá y Ribadeneira, 2017), en el trabajo "Análisis Estadístico de las causas de Accidentes de Tránsito en la Ciudad de Guayaquil", toman como base las publicaciones que realiza la Comisión de Transito del Guayas, por medio del cual buscaron determinar el comportamiento de las causas y sus consecuencias en la sociedad de los accidentes de tránsito, mediante el uso de técnicas de Estadística Descriptiva Multivariada. Para el estudio se trabajó por capítulos; teniendo en cuenta información relacionada con el histórico del transporte y transito de Guayaquil; el estudio de accidentes en la ciudad y sus consecuencias; el tercer capitulo conformado por el marco teórico; y el cuarto y último capitulo conformado por uso de técnicas de estadística descriptiva multivariada de las causas de transito en la ciudad de Guayaquil. Se utilizaron técnicas como Análisis de componentes principales (PCA) bivariados, analizando variables como el sexo de las victimas y la condición social.

En Ecuador y específicamente en el Cantón Cuenca, se realizó un análisis estadístico en el cual se analizaron las principales causas que provocan con mayor frecuencia los accidentes de transito en las vías, para ello, se tomaron en consideración todos los vehículos que circulan en el Cantón, para en base a ellos, identificar propuestas y soluciones con el fin de reducir las cifras negativas de esta problemática social. Para el desarrollo de este estudio se utilizó un modelo ARIMA, ya que el objeto de dicho trabajo consistió en predecir si los accidentes se van a mantener o si va haber un crecimiento o reducción en los mismos. El modelo fue univariante por lo cual se trabajó sobre el número de accidentes por cada mes. (Cabrera y Vele, 2017).

(Moya, P. 2019), llevo a cabo un trabajo en el cual se analizó los choques en la ruta 27 usando modelos multivariados, para ello se realizó la recopilación de datos del TPDA (Tránsito promedio diario anual), teniendo en cuenta los choques viales y la geometría de la ruta. La información fue procesada y organizada utilizando Sistemas de Información Geográfica (SIG) para ubicar espacialmente las estaciones de recuento, los choques y la geometría de la carretera. Se procedió a utilizar a correr los modelos Binomial Negativo o Poisson, según el mejor ajuste de cada caso particular, posteriormente aplicaron funciones de desempeño, mediante el modelo Bayes completo (full Bayes) para todas las severidades identificadas. El uso de métodos multivariados permitió hacer un análisis con varias variables de manera simultánea, de manera que se toma en cuenta el efecto de todas y cada una en presencia de las demás.

García, G. (2001). Llevo a cabo un análisis de la ocurrencia del factor humano en la accidentalidad vial en Colombia durante la última década de los años 90 y los primeros años del siglo actual, donde se establecieron como principales víctimas los peatones seguidos por pasajeros y en tercer lugar se encontraron los motociclistas; también determina la edad como un factor relevante en este tipo de muertes.

Otra investigación desarrollada en Colombia, permitió conocer las dinámicas sociales de la accidentalidad vial, formas de ocurrencia, tipos de consecuencias, tipos de víctimas y su relación con variables sociodemográficas. En esta oportunidad se analizaron datos procedentes de dos fuentes: una encuesta de victimización vial (accidentalidad subjetiva) a una muestra no aleatoria de estudiantes de ciclos superiores de 20 ciudades capitales (n=2292) de Colombia y estadísticas oficiales de heridos en tránsito (accidentalidad objetiva) en Bogotá en el año 2009 (n=2300). Mediante un análisis de clasificación jerárquica y combinando características del vehículo, del daño causado y del conductor teniendo variables como la edad, el sexo y otros actores de la vía, los resultados mostraron cuatro tipologías de accidentes: 1) conductor particular, atropello, víctima peatón de 61 años o más, 2) víctima peatón de 0 a 10 años por atropello y lesión por trauma facial, 3) víctima pasajero o mujer, por caída desde el automóvil o bus de transporte público, y lesiones en el cuello, 4) víctima conductor de moto o cicla, entre 21 y 30 años, varón, por colisión, y daños en las extremidades. Las implicaciones de los resultados se discuten con relación a políticas de prevención de la accidentalidad vial. Mediante este trabajo se llevo a concluir los siguiente: La combinación de múltiples variables de los datos oficiales de accidentalidad en Bogotá, permite identificar unas tipologías de accidentes, que indican la mayor vulnerabilidad de ciertos usuarios de la vía –niños, adultos mayores, peatones, conductores de motos o de bicicletas- a los accidentes, en las interacciones con los vehículos de mayor potencia (automóviles, buses). Es por ello que es necesario volver a señalar, siguiendo a Montoro et al. (2000), la responsabilidad que implica conducir un vehículo a motor. De otro lado, también la formación de los conductores de transporte público y su estilo de conducción se hacen visibles, al ligarse a la accidentalidad de los pasajeros, especialmente en el caso de las mujeres. Otro hallazgo relevante es que el tipo de accidente vial se asocia con mayor probabilidad a un tipo de lesiones, y ello puede ser tenido en cuenta para la investigación que deben adelantar las autoridades judiciales para evaluar los daños (Vargas, Cuervo, & Malaver, en prensa). Por su parte, el tipo de usuario de la vía (peatón, motociclista, chófer) también se asocia con una dinámica de accidentalidad específica. (Ruiz y Herrera, 2016).

Universidad de Jáen. (2015). Factores Causantes de los accidentes. Asignatura Prevención de riesgos laborales, 7-52. En este documento se establecen los principales factores asociados a los accidentes de tránsito entre los que se plantean la responsabilidad humana con todas sus características, estado mecánico general de los vehículos, estado de las vías y señales de tránsito. Así mismo, determina la responsabilidad de los diferentes actores involucrados y responsables en los accidentes de tránsito. Las conclusiones más relevantes descritas en el documento de la universidad hacen referencia al uso de altas velocidades, a la conducción en estado de alicoramiento, factores de distracción mientras se conduce, mal estado de la vía y falta de señalización.

El grupo de educación sanitaria y promoción de la salud del PAPPS, de España, realizó un estudio con el fin de establecer los principales factores asociados a los accidentes de tránsito entre los que se plantean la responsabilidad humana con todas sus características, estado mecánico general de los vehículos, estado de las vías y señales de tránsito. Así mismo, determinar la responsabilidad de los diferentes actores involucrados y responsables en los accidentes de tránsito. Dicho trabajo genero conclusiones diferenciadas de otros trabajos como accidentes generados por factores de tipo patológico como el tabaquismo, la apnea del sueño, por lo cual se concluyó que un factor determinante para disminuir los índices de accidentalidad es incluir exámenes médicos a los conductores sobre diagnóstico y tratamiento de los trastornos del sueño, e investigar los síntomas

de somnolencia y fatiga en los conductores viales. https://www.mscbs.gob.es/ciudadanos/accidentes/docs/GUIA_PREV_ACC_TR_AFICO.pdf

Matamoros, H., (2018). En el trabajo denominado "Elementos que participan en la incidencia de accidentes de tránsito", se determinó que el campo de la seguridad vial es uno de los aspectos de especial importancia en el desarrollo de España. Día a día, la prensa se ve colmada de noticias sobre accidentes de tránsito, los cuales oscilan desde la colisión más simple hasta situaciones complicadas y violentas en las que seres humanos pierden su vida o quedan irremediablemente lisiados para el resto de sus días. En un país como España, cuyos recursos económicos son limitados, es preciso establecer las causas y situaciones comunes que presentan los percances, de tal manera que sea posible orientar las políticas correctivas en forma eficiente. Finalmente, el autor concluye que es determinante conocer la verdadera causa de ocurrencia de accidente de tránsito, es importante para plantear estrategias que conlleven a la planificación de acciones o medidas correctivas relacionadas. http://creandoconciencia.org.ar/enciclopedia/accidentologia/relevamiento-de-rastros/ELEMENTOS-QUE-PARTICIPAN-EN-LA-INCIDENCIA-DE-ACCIDENTES-DE-TRANSITO.pdf

El Instituto Nacional de Medicina Legal y Ciencias forenses de Colombia, determino en un informe elaborado en el año 2004 que los accidentes de tránsito son la segunda causa de muerte violenta en nuestro país (después de los homicidios). Aunque este es un fenómeno que en los últimos años ha descendido considerablemente, en parte por las innovaciones tecnológicas, mecánicas y de seguridad incluidas actualmente en los vehículos, la reducción es inferior a la esperada en parte por el comportamiento tanto de conductores y peatones al momento de asumir su rol como usuarios activos de la vía. La prevención en los accidentes de tránsito, si bien está enfocada hacia las modificaciones vehiculares haciendo día tras día automotores y vías más seguras, un aspecto un tanto descuidado es la modificación de la conducta y el comportamiento riesgoso de los actores en las vías. Además del simple hecho de circular, según algunos autores, hasta un 95% de los accidentes de tránsito se deben a factores humanos. Estos factores podrían dividirse en dos tipos: errores y violaciones intencionales. Las mujeres tienden a accidentarse más por errores de percepción o juicio mientras los hombres se accidentan más por violaciones. (Hernández et al., 2004).

(Herrera y Pulido, 2005). Llevaron a cabo un trabajo de grado en el cual tenía como objetivo determinar y analizar los índices de Accidentalidad en el Departamento de Boyacá en los últimos tres años y medio, analizar sus posibles causas, identificar puntos críticos, los elementos que intervienen y proporcionar información a través de una base de datos, a las autoridades y empresas para ayudarles en la formulación de estrategias cuyo objetivo sea buscar soluciones que disminuyan este flagelo, proporcionando herramientas para buscar soluciones que ayuden a mitigar o reducir el gran fenómeno de accidentalidad vial en Colombia el cual se ha convertido en la segunda causa de mortalidad luego de los homicidios violentos. En dicho estudio se concluyo que Las causas de accidentalidad que se repiten con mayor frecuencia en el departamento de Boyacá son; con código 110 Distraerse (cuando el conductor no está pendiente de la vía y de las acciones de los demás conductores), código 134 Impericia en el manejo (cuando el conductor no sabe maniobrar ante una situación de peligro) y código 140 Falta de precaución por niebla, lluvia o humo (conducir en estas circunstancias sin disminuir la velocidad y/o sin utilizar luces).

En el trabajo realizado por parte de la Secretaria Distrital de Movilidad de Boyacá, se analizó la incidencia del factor humano en la accidentalidad vial en el territorio colombiano, y se identificó evidencia empírica para la formulación de política pública del tránsito. Se desarrolló el diseño es descriptivo-correlacional. Instrumentos: cuestionario de comportamiento para conductores y motociclistas (D.B.Q.) y encuesta tipo Likert. Muestra no probabilística intencional: 16.322 personas (8.631 conductores de automotores, 5.133 motociclistas y 2.558 peatones, pasajeros y acompañantes). Estableciéndose: a) conductores con nivel educativo superior inciden en menos accidentes; b) peatones con menor nivel educativo inciden en mayores conductas riesgosas; c) las mujeres tienen actitud positiva y perciben eficacia de campañas en prevención; d) estilos de conducción iracundo, ansioso, riesgoso y de alta velocidad cometen más infracciones y accidentes; e) desobedecer señales de tránsito, la principal causa de accidentalidad; f) agresividad, hostilidad y estrés en el tránsito son factores que aumentan la probabilidad de accidente; g) campañas de prevención no están siendo captadas por la población más afectada. En dicho trabajo se concluyó que los factores de accidentalidad vial guardan correspondencia con el Triángulo de Seguridad Vial establecido por la Organización de las Naciones Unidas: factor humano preponderante en la accidentalidad. Lineamientos de política pública fundamentados en procesos educativos y corporativos tendientes a disminuir la accidentalidad por lo general no se concibe aplicada a la seguridad pública. Por ello, se presenta la inteligencia criminal como una disciplina que se ocupa de anticipar los riesgos criminales contra la seguridad pública. Esto permitiría disminuir la incertidumbre durante la toma de decisiones y calcular los daños contra la seguridad pública, que se pretende prevenir. (Norza et al., 2004).

Coy Jiménez, E. (2014). Mortalidad prematura por accidentes de transito en Colombia en el periodo 2005-2013. Plantea definiciones claras sobre accidentalidad y mortalidad en accidentes de tránsito, ocurridos en los diferentes departamentos de Colombia, estableciendo rango de edades, características de los actores lesionados y victimas junto a la tasa de mortalidad que se ha presentado durante los últimos años. En dicho trabajo se concluyó que la utilización de fuentes de información de dos sectores diferentes como las bases de datos sobre accidentes de tránsito y las estadísticas vitales, evidenció la necesidad de un trabajo articulado que permita concatenar las bases de datos para hacer y potenciar un mejor uso de las fuentes de información y la elevada mortalidad prematura por los accidentes de tránsito en Colombia coloca el tema nuevamente en la agenda nacional como prioridad, dado que afecta principalmente el grupo de 20 a 34 años, lo que incide en la productividad, afectando varias esferas de la vida nacional. Es de resaltar como el peatón hombre y los motociclistas hombres son mayoritariamente vulnerables a una muerte temprana, dato que mantienen similitud con los que se pueden observar en las otras muertes por causa de lesión externa.

Por su parte el Departamento Nacional de Estadística DANE. (2021, 25 junio). Genero un informe detalla claramente las estadísticas correspondientes a la natalidad registrada en Colombia durante el año 2020 y el primer trimestre del año 2021. Además, establece los datos de número de muertes y tasa de mortalidad por las diferentes causas registrados en Colombia en los que se incluyen los datos correspondientes a las muertes por accidentes de tránsito y su comportamiento durante los últimos 6 años. En dicho informe se concluyó que para el año 2020pr los departamentos con mayores tasas de muertes por accidentes de tránsito son: Casanare, Arauca, San Andrés, Meta y

Cesar. En departamentos como Bogotá, Atlántico, Antioquia y Valle del Cauca la tasa de muertes por accidentes de tránsito ha disminuido gradualmente desde 2015 hasta 2020. Para el año 2020 se observa una reducción del 18,2% en el número de muertes por accidentes de tránsito en el país. Se observa cómo la tasa de muertes por accidentes de tránsito por 100.000 habitantes cae considerablemente entre 2019 y 2020pr, probablemente por las medidas de confinamiento. Según grupos de edad, las defunciones por accidentes de tránsito se concentran entre los 20 y 24 años para todos los años, seguidos de los grupos de 25 a 29 años. https://www.dane.gov.co/files/investigaciones/poblacion/pre_estadisticasvitales_Itrim_2021pr.pdf.

3. METODOLOGÍA

Para llevar a cabo el presente trabajo, se abordaron dos métodos de estadística descriptiva multivariada denominados: Factoriales o en ejes principales y de clasificación o agrupamiento.

Para entender la metodología utilizada, podemos pensar en la base de datos utilizada, en la cual denominamos los municipios en *n* filas que denominaremos "individuos", y las variables de interés *p*, que se tomaran como medidas continuas. Una tabla de estas características es una matriz numérica de *n* filas y *p* columnas y tiene representaciones geométricas, así: *n* vectores fila en $\mathbb{R}^p$ y las columnas como *p* en vectores en $\mathbb{R}^n$.

En el primer caso, la representación geométrica de los municipios denominados "individuos", es un conjunto de *n* puntos en un sistema de *p* ejes ortogonales, cada eje asociado a una variable, y las coordenadas de un individuo son *p* valores que toma para la variable. Aunque esta es una representación abstracta, tiene las mismas propiedades de un diagrama de dispersión.

Por ejemplo, si dos individuos de edad diferente, están cerca a $\mathbb{R}^p$, tienen coordenadas similares o mas o menos las mismas coordenadas, es decir, presentan valores similares para las *p* variables. Por esta razón, resulta que la representación geométrica es importante a la hora de comparar a los individuos entre si y visualizar la estructura de la nube de individuos, en el sentido de identificar su forma y patrones que se pueden manifestar como grupos e individuos. Para detectar dichos patrones, se utilizan los métodos factoriales o en ejes principales y de clasificación o agrupamiento. Los primeros buscan los mejores ejes y planos de proyección, para si observar nubes de puntos de forma aproximada, y los de agrupamiento o clasificación, buscan descubrir puntos en los datos, con el fin de que los individuos se parezcan lo más posible cuando pertenecen a un mismo grupo y se diferencien al máximo cuando pertenecen a grupos diferentes.

La nube de variables esta representada por el conjunto de *p* vectores en $\mathbb{R}^n$ y su representación imaginaria pueden ser flechas que empiezan en el origen de los *n* ejes y finalizan en el punto cuyas coordenadas son los *n* valores que toma la respectiva variable. Se debe tener claridad que cada uno de los *n* puntos se asocia con un individuo y la longitud de las flechas de los vectores variables representan sus desviaciones estándar y el ángulo entre parejas de ellos su correlación. (Pardo, E. 2020).

Según (Pardo, E. 2020), el Análisis de Componentes Principales (ACP) generalizado o también denominado *Análisis Factorial General,* es el marco de referencia para este tipo de métodos. El análisis de componentes principales (ACP), se utiliza principalmente para describir tablas que

representan las unidades estadísticas en las filas, se conocen como "Individuos", y en las columnas se ubican las variables de tipo continuo que se han medido sobre los individuos. Los objetivos del análisis de componentes principales (ACP) son:

- Contrastar los individuos entre si. Las nubes de puntos observadas en las gráficas permiten detectar patrones en los datos,
- Describir las relaciones entre las variables,
- Disminuir la dimensión de la representación. A mayor relación entre las variables mayor será la capacidad de síntesis del análisis de componentes principales (ACP).

En síntesis, un análisis de componentes principales (ACP) se constituye en un método estadístico que permite compendiar la complejidad de espacios muestrales con muchas dimensiones a la vez que conserva su información. Finalmente, el ACP permite por lo tanto "condensar" la información contenida por múltiples variables en solo unas pocas componentes. Esto lo convierte en un método muy útil de aplicar previa utilización de otras técnicas estadísticas tales como regresión, clustering, entre otras. Aun así, no hay que olvidar que sigue siendo necesario disponer del valor de las variables originales para calcular las componentes.

Para utilizar la metodología de análisis de componentes principales (ACP), se deben tener en cuenta algunos supuestos:

- Se tienen múltiples variables que deben medirse a nivel continuo (aunque las variables ordinales se usan con mucha frecuencia). Los ejemplos de variables continuas (es decir, variables de razón o intervalo) incluyen el tiempo de revisión (medido en horas), la inteligencia (medido con la puntuación de CI), el rendimiento del examen (medido de 0 a 100), el peso (medido en kg), etc. Ejemplos de variables ordinales comúnmente utilizadas en ACP pueden incluir una amplia gama de escalas Likert (por ejemplo, una escala de 5 puntos desde 'vía pública', 'Otro sitio', 'Hospital/Clínica', 'Centro/Puesto de salud' y 'Casa/Domicilio', como la utilizada para la variable "Sitio de Defunción" en el presente trabajo,
- Debe haber una relación lineal entre todas las variables. La razón de esta suposición es que un ACP se basa en los coeficientes de correlación de Pearson y, como tal, debe haber una relación lineal entre las variables. En la práctica, este supuesto es algo relajado (incluso si no debería serlo) con el uso de datos ordinales para las variables. Aunque la linealidad se puede probar usando un diagrama de dispersión matricial, esto a menudo se considera excesivo porque el diagrama de dispersión a veces puede tener más de 500 relaciones lineales,
- Debe tener la adecuación del muestreo, lo que simplemente significa que para que el ACP produzca un resultado confiable, se requieren tamaños de muestra suficientemente grandes. Se han propuesto muchas reglas prácticas diferentes. Estos difieren principalmente dependiendo de si se propone un tamaño de muestra absoluto o si se utiliza un múltiplo del número de variables en su muestra. En términos generales, se ha recomendado un mínimo de 150 casos, o de 5 a 10 casos por variable, como tamaño mínimo de muestra. Hay algunos métodos para detectar la adecuación del muestreo: (1) la Medida de adecuación del muestreo de Kaiser-Meyer-Olkin (KMO) para el conjunto de datos general; y (2) la medida KMO para cada variable individual,

- Los datos deben ser adecuados para la reducción de dimensionalidad. Efectivamente, debe tener correlaciones adecuadas entre las variables para que las variables se reduzcan a un número menor de componentes. Típicamente esto se puede evaluar visualizando la matriz de correlaciones, realizando una prueba de esfericidad de Bartlett o haciendo figuras por pares (pair plots),
- No debe haber valores atípicos significativos. Los valores atípicos son importantes porque pueden tener una influencia desproporcionada en los resultados. Se recomienda determinar valores atípicos como puntuaciones de componentes superiores a 3 desviaciones estándar de la media,

Las ideas generales de la metodología utilizada se describen a continuación:

- **Vector de Medias, Matriz de Varianzas y Covarianzas y Matriz de Correlaciones**

Los parámetros poblaciones, sea $(X_1, X_2,\ldots,X_p)$ un vector de variables.

Se define el vector de medias como $\mu = (\mu 1, \mu 2,\cdots, \mu p)$, donde $\mu i = E(Xi)$, y la matriz de Varianzas y Covarianzas definida como:

$$\sum = \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 \ldots & \sigma_{1p}^2 \\ \sigma_{21}^2 & \sigma_2^2 \ldots & \sigma_{2p}^2 \\ . & . & . \\ . & . & . \\ \sigma_{p1}^2 & \sigma_{p2}^2 \ldots & \sigma_p^2 \end{pmatrix}$$

donde, $\boldsymbol{\sigma_{ij}^2} = \mathbf{Cov}(X_i,X_j) = \mathbf{E}((X_i - \overline{X_i})(X_j - \overline{X}_j))$ **y** $\boldsymbol{\sigma_i^2} = \mathbf{Var}(X_i)$**,**

Finalmente, la matriz de correlaciones es obtenida como:

$$\rho = \begin{pmatrix} 1 & \rho_{12} \ldots & \rho_{1p} \\ \rho_{21} & 1 \ldots & \rho_{2p} \\ . & . & . \\ . & . & . \\ \rho_{p1} & \rho_{p2} \ldots & 1 \end{pmatrix}$$

donde, $\rho_{ij} = \dfrac{\sigma_{ij}^2}{\sigma_i\,\sigma_j}$

- **Parámetros muestrales:**

Los vectores de medias definidos como:

$\bar{\boldsymbol{X}} = (\overline{\boldsymbol{X_1}}, \bar{\boldsymbol{X}}_{\boldsymbol{2}}, \dots\dots\dots, \bar{\boldsymbol{X}}_{\boldsymbol{p}}), \boldsymbol{donde}\ \bar{\boldsymbol{X}}_{\boldsymbol{i}} = \boldsymbol{n^{-1}} \sum \boldsymbol{X_i}$, siendo *n* el tamaño de la muestra.

Ahora, se define la matriz de varianzas y covarianzas como:

$$\boldsymbol{S} = \frac{\boldsymbol{1}}{\boldsymbol{n-1}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{p}} (\boldsymbol{Xi} - \bar{\boldsymbol{X}})(\boldsymbol{Xi} - \bar{\boldsymbol{X}})^{\boldsymbol{t}}$$

Seguidamente, se define la matriz de correlaciones como:

$$\boldsymbol{R} = \boldsymbol{D}^{-\frac{1}{2}} \boldsymbol{S} \boldsymbol{D}^{-\frac{1}{2}}, \qquad donde:$$

$$\boldsymbol{D}^{-\frac{1}{2}} = \boldsymbol{diag}\left(\frac{\boldsymbol{1}}{\boldsymbol{s_1}}, \frac{\boldsymbol{1}}{\boldsymbol{s_2}}, \dots\dots, \frac{\boldsymbol{1}}{\boldsymbol{s_p}}\right), y\ s_p\ es\ el\ elemento\ en\ la\ posición\ (p,p) de\ la\ matriz\ S$$

- **Distancias**

Una de las principales características del análisis de componentes principales (ACP) es la reducción del número de dimensiones. Naturalmente, esto conlleva una pérdida en la información contenida en la base de datos como tal. Lo que se busca es perder la menor cantidad de información, en términos espaciales, se quiere que la distancia entre dos puntos, sea lo más preservada posible al considerar su proyección con menos variables.

La idea es perder la menor cantidad de información posible. En términos espaciales, la idea es que la distancia entre dos puntos, sea preservada lo máximo posible al considerar (proyectar en) menos variables. La idea de distancia puede ser entendida como Euclidiana:

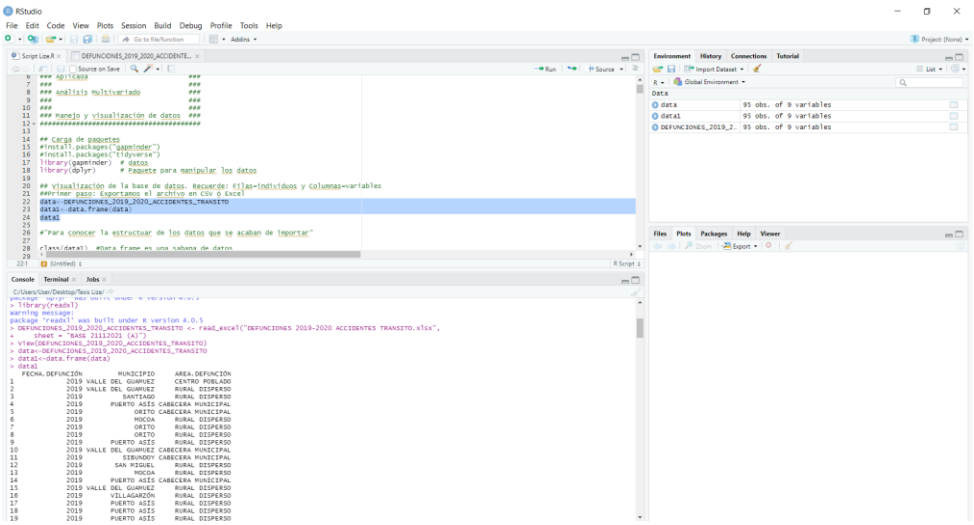
$$\boldsymbol{d_j^i} = \sqrt{\sum_{\boldsymbol{k=1}}^{\boldsymbol{p}} (\boldsymbol{X_{ik}} - \boldsymbol{X_{jk}})^{\boldsymbol{2}}}$$

En el caso del presente trabajo, la metodología esta determinada por un conjunto de variables categóricas (Nominales y Ordinales) y una variable numérica (EDAD), para lo cual se utilizó una base de datos con un total de 7 variables y 96 individuos para un total de 854 datos para cada variable y un total de 195.239 datos. Para esto se utilizo el entorno de ejecución de R-Studio.

Como primera fase se procedió a cargar las librerías a utilizar como son: gapminder, dplyr, pca3d, moments, ggplot2, FactoMineR, ade4, FactoClass, Factoextra, missMDA, devtools, ggbiplot y ggforce, todas utilizadas para el manejo y manipulación de los datos, obtener las medidas de forma y las necesarias para la elaboración de gráficas propias del ejercicio.

Seguido se procede a cargar la base de datos en R-Studio, para lo cual se cuenta con un archivo en Excel:

Figura 1
Base de datos utilizada para el trabajo



Luego de cargar la base de datos, se llevo a cabo la visualización y análisis descriptivo de las variables (Municipio, Área de Defunción, Sitio de Defunción, Sexo del Fallecido, Estado Civil, Edad, Nivel Educativo y Régimen Seguridad).

Seguido se procede a conocer la estructura de los datos que se acaban de importar, verificando inicialmente que no se presenten observaciones ausentes, a continuación, se verifico el orden de las columnas en el dataframe para el desarrollo de las operaciones siguientes. En este caso se considero importante definir al programa las variables que están en escala nominal para lo cual se requiere la librería FactoClass, recordemos en este punto que la única variable numérica en EDAD, y determinando que las demás variables se consideran nominales se procede a definirlas en R-Studio, igualmente es importante definir al programa la variable cuantitativa, para lo cual se realiza la verificación de las variables y garantizar la definición realizada a cada variable.

En la segunda fase, se generaron las tablas de frecuencias y gráficos para las variables cualitativas, a continuación, se presenta las tablas para la variable “MUNICIPIO”, para el total de las 96 observaciones:

Tabla 1
Tabla de Frecuencia para la variable MUNICIPIO

MUNICIPIO	FRECUENCIA	FRECUENCIA ACUMULADO
COLÓN	7	7.07
MOCOA	14	14.5
ORITO	12	12.13
PUERTO ASIS	22	22.23
PUERTO CAICEDO	4	4.04
PUERTO GUZMÁN	2	2.02
PUERTO LEGUÍZAMO	1	1.01
SAN MIGUEL	6	6.06
SANTIAGO	2	2.02
SIBUNDOY	5	5.05
VALLE DEL GUAMEZ	15	15.16
VILLA GARZÓN	5	5.05

Tabla 2
Tabla de Frecuencia para la Variable SITIO DE DEFUNCIÓN

SITIO DE DEFUNCIÓN	FRECUENCIA	FRECUENCIA ACUMULADA
CABECERA MUNICIPAL	39	39.41
CENTRO POBLADO	2	2.02
RURAL DISPERSO	54	54.57

Como se puede observar en la tabla 1, la mayor frecuencia de accidentes de transito con fatalidad se presenta en el municipio de Puerto Asís con 22 casos, seguido de Valle del Guamez con un total de 15 casos y siendo los municipios de Puerto Guzmán y Santiago los de menos casos de fatalidades por accidentes de transito con 2 casos cada uno. Por ejemplo, la tabla 2, permite identificar que la mayoría de las fatalidades se presentan en zona Rural Dispersa con un total de 54 casos y Cabecera Municipal con 39 casos presentados, esto se puede considerar debido a las altas velocidades de los conductores en vías rurales, al estado de las vías terciarias, a la falta de iluminación, señalización, entre otras, igualmente las zonas circundantes a la Cabecera Municipal generalmente son vías de tránsito rápido como vías secundarias o autopistas, lo que de la mano de posibles imprudencias de los conductores puede desencadenar los fatídicos casos en el departamento de Putumayo.

Un aspecto interesante a tener en cuenta en el Nivel Educativo de los fallecidos, la tabla 3, permite identificar los niveles de escolaridad en los cuales se presentan mayor número de fatalidades:

Tabla 3
Tabla de Frecuencia para la variable Nivel Educativo

NIVEL EDUCATIVO	FRECUENCIA	FRECUENCIA ACUMULADA
NINGUNO	2	2.02
PREESCOLAR	1	1.01
BÁSICA PRIMARIA	27	27.28
BÁSICA SECUNDARIA	17	17.18
MEDIA ACADÉMICA O CLÁSICA	26	26.27
TÉCNICO PROFESIONAL	3	3.03

PROFESIONAL	6	6.06
SIN INFORMACIÓN	13	13.14

De acuerdo a la tabla 3, los casos de mayor frecuencia se presentan para el nivel educativo de Básica Primaria con 27 casos, seguido de Media Académica o Clásica (corresponde a grados de decimo y undécimo) con un total de 26 fatalidades, también se considera importante el nivel de Básica Secundaria ya que en este nivel se presentaron 17 accidentes fatales. Lo anterior puede generar la hipótesis de si la mayoría de los casos de fatalidades identificados se debe al nivel de escolaridad, esto debido a posibles falencias en los cursos de conducción en los cuales se instruye sobre las velocidades permitidas de acuerdo al tipo de vía o zona, el significado de las señalizaciones viales y muy importante el acceso a cursos de manejo preventivo.

Posteriormente, se obtuvieron las tablas de doble entrada para variables a escala de medida nominal, para el ejemplo se presenta los resultados generados para las variables Sexo del Fallecido y Estado Civil:

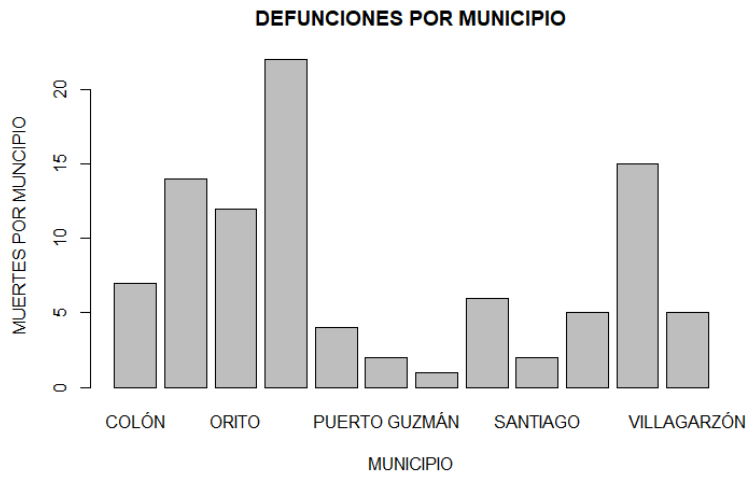
Tabla 4
Tabla de doble entrada para las variables Sexo Fallecido Vs Estado Civil

	ESTADO CIVIL					
SEXO FALLECIDO	CASADO	CONVIVIENTE	SEPARADO	SIN INFORMACIÓN	SOLTERO	VIUDO
FEMENINO	2	0	0	4	5	1
MASCULINO	9	20	1	10	46	1
Sum	11	20	1	10	51	2

Como se puede observar en la tabla 4, la mayor cantidad de personas fallecidas en accidentes de transito en el departamento de Putumayo, en su mayoría son hombres solteros, seguido hombres en estado de convivencia y siendo los casos presentados superiores a las fatalidades presentadas en mujeres, uno de los aspectos a considerar y siendo conocido es el uso de altas velocidades por parte del genero masculino, en este caso sería interesante determinar si estos accidentes se presentan en su mayoría en motocicletas o automóviles, pero dicha información no se encontró disponible. Dentro de los cálculos realizados se pudo establecer que la mayor cantidad de hombres fallecidos se encuentran entre los 19, 20, 21, 22 y 23 años y una media de 3.5 casos. Del total de casos de fatalidades presentadas en el departamento se pudo determinar que la mayoría de las personas involucradas se encontraban en Régimen de Seguridad Subsidiado con un total de 52 casos, otros 25 casos no presentaron información y 16 casos pertenecientes a régimen contributivo.

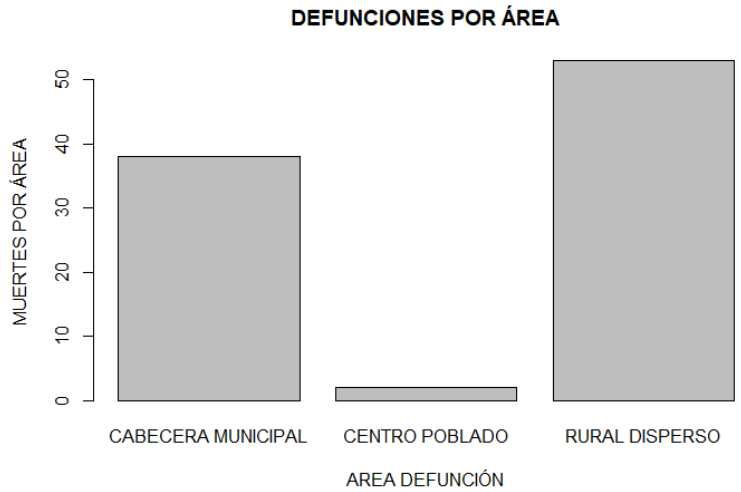
Seguido de los cálculos anteriores, se llevó a cabo la tercera fase correspondiente al desarrollo de gráficos para variables cualitativas, a continuación, se presenta el gráfico correspondiente a las muertes por municipio:

Gráfica 1
Defunciones por Municipio



Efectivamente la gráfica 1, concuerda con la información arrojada por la tabla 1, siendo el municipio de Puerto Asís, el que mayor número de casos fatales presenta.

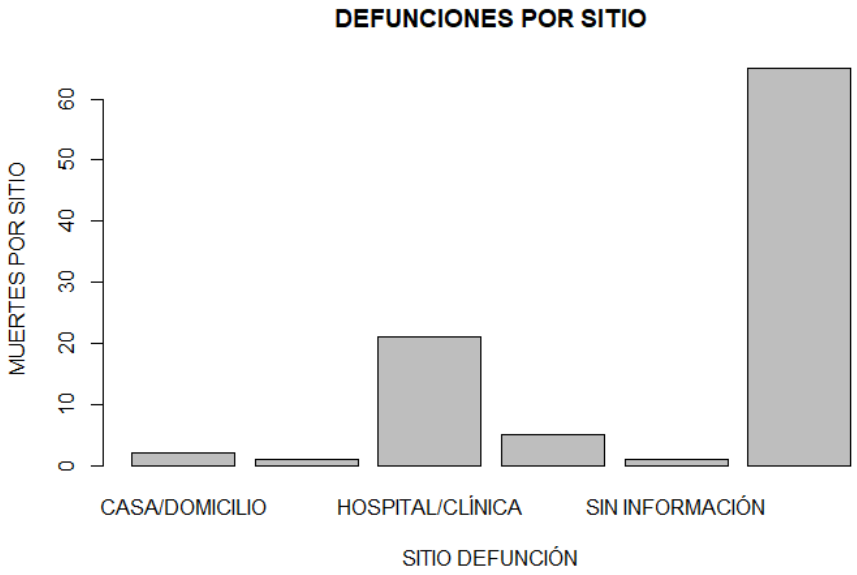
Gráfica 2
Defunciones por Área



Ocurre lo mismo con la gráfica 2, en la cual se asegura que es en el sector rural disperso donde se generan la mayor cantidad de casos fatales por accidentes de tránsito.

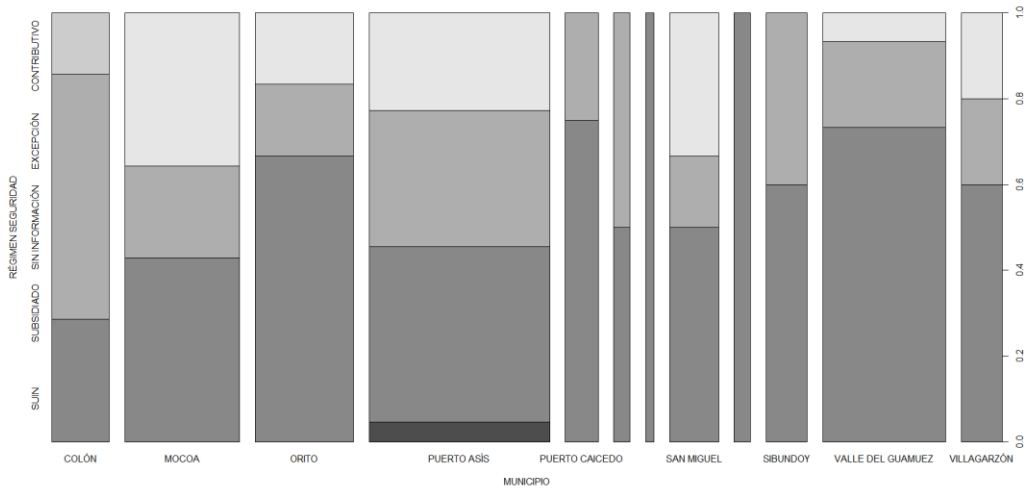
La gráfica 3, por su parte deja en evidencia que las muertes se generan de manera inmediata o fulminante o no alcanzan a ser trasladados a centros asistenciales ya que la mayoría de registros se encuentran en vía pública, seguido de Hospital/Clínica.

Gráfica 3
Defunciones por Sitio



Otro gráfico interesante que permite determinar el régimen de seguridad por municipio se presenta a continuación:

Gráfica 4
Régimen de Seguridad Vs Municipio



Por medio del gráfico 4, se puede determinar por ejemplo que en el municipio de Puerto Asís en el cual se presenta el mayor número de casos la mayoría de los fallecidos pertenecen a régimen subsidiado, seguido de personas sin información y finalmente pertenecientes a régimen contributivo.

Dentro de los análisis realizados se utilizó la función "*filter",* la cual permite filtrar por valores de una variable, por ejemplo, mediante dicha función se estableció que para el año 2019 se presentaron 42 accidentes fatales y para el año 2020 se presentaron 53 fatalidades, siendo una situación extraña debido a los extensos periodos de cuarentena establecidos por el Gobierno Nacional a causa de COVID – 19.

Uno de los aspectos relevantes calculados en el trabajo es el promedio de expectativa de vida para los periodos 2019 y 2020, así como la edad máxima de las personas fallecidas:

Gráfica 5
Expectativa de vida Vs Máxima edad de los fallecidos

AÑO DE DEFUNCIÓN	EXPECTATIVA DE VIDA	MAXIMA EDAD
2019	25	85
2020	33	87

Para finalizar la tercera fase se procedió a realizar el resumen de cada una de las variables en sus términos descriptivos, tal y como se presenta en la siguiente imagen:

Imagen 1
Resúmenes estadísticos descriptivos para cada variable

```
              SITIO.DEFUNCIÓN   SEXO.FALLECIDO          ESTADO.CIVIL
CASA/DOMICILIO        : 2     MASCULINO: 0     CASADO         :11
CENTRO/PUESTO DE SALUD: 1     FEMENINO : 0     CONVIVIENTE    :20
HOSPITAL/CLÍNICA      :21     NA's     :95     SEPARADO       : 1
OTRO SITIO            : 5                      SIN INFORMACIÓN:10
SIN INFORMACIÓN       : 1                      SOLTERO        :51
VÍA PÚBLICA           :65                      VIUDO          : 2

     EDAD                        NIVEL.EDUCATIVO
Min.   : 2.00   BÁSICA PRIMARIA          :27
1st Qu.:20.00   MEDIA ACADÉMICA O CLÁSICA:26
Median :29.00   BÁSICA SECUNDARIA        :17
Mean   :35.59   SIN INFORMACIÓN          :13
3rd Qu.:51.00   PROFESIONAL              : 6
Max.   :87.00   TÉCNICA PROFESIONAL      : 3
                (Other)                  : 3
      RÉGIMEN.SEGURIDAD
CONTRIBUTIVO   :16
EXCEPCIÓN      : 1
SIN INFORMACIÓN:25
SUBSIDIADO     :52
SUIN           : 1
```

Ya en la cuarta fase se llevó a cabo el Análisis de Componentes Principales (ACP), para esta fase del trabajo, se realizaron las transformaciones a las variables categóricas (Nominales u Ordinales) a numéricas, para lo cual se recodificaron las variables para que estas sean numéricas, así:

- ÁREA DE DEFUNCIÓN (1; CENTRO POBLADO, 2: RURAL DISPERSO, 3; CABECERA MUNICIPAL),
- SITIO DE DEFUNCIÓN (1: VÍA PÚBLICA, 2: OTRO SITIO, 3: HOSPITAL/CLÍNICA, 4: CENTRO/PUESTO DE SALUD, 5: CASA/DOMICILIO),
- SEXO FALLECIDO (1: MASCULINO, 2: FEMENINO),
- ESTADO CIVIL = (1: SOLTERO, 2: CONVIVIENTE, 3: CASADO, 4: SIN INFORMACIÓN, 5: VIUDO, 6: SEPARADO),
- EDAD (ES NUMÉRICA),
- NIVEL EDUCATIVO (1: BÁSICA PRIMARIA, 2: BÁSICA SECUNDARIA, 3: MEDIA ACADEMICA O CLÁSICA, 4: NINGUNO, 5: TECNICA PROFESIONAL, 6: NO DETERMINADA, 7: PROFESIONAL, 8: PREESCOLAR), y
- RÉGIMEN SEGURIDAD (1: SUBSIDIADO, 2: CONTRIBUTIVO, 3: NO SE SABE, 4: EXCEPCIÓN)

A su vez, se codifican los municipios de acuerdo a la codificación Nacional, así: Valle del Guamuez: 865, Villa Garzón: 885, Mocoa: 1, Colón: 219, Orito: 320, Puerto Asís: 568, Puerto Caicedo: 569, Puerto Guzmán: 571, Sibundoy: 759, San Miguel: 757, Santiago: 760 y Puerto Leguizamo: 573.

Una vez se realizó la recodificación de variables, se procedió a calcular la independencia entre las variables mediante la prueba de Chi-Cuadrado, utilizando la variable MUNICIPIO con cada una de las variables de interés:

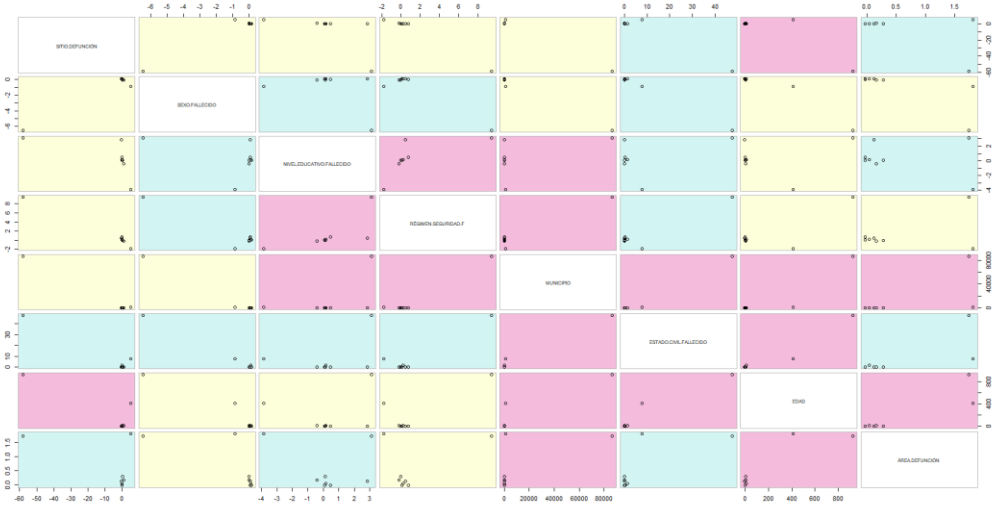
Imagen 2
Chi-Cuadrado Municipio vs Área de defunción

```
> chisq.test(MUNICIPIO, data2$ÁREA.DEFUNCIÓN)

        Pearson's Chi-squared test

data:  MUNICIPIO and data2$ÁREA.DEFUNCIÓN
X-squared = 17.193, df = 22, p-value = 0.7526
```

Con un 95% de confianza, existe evidencia suficiente para rechazar la hipótesis nula de independencia estadística entre las variables MUNICIPIO y ÁREA DE DEFUNCIÓN percibido (χ^2=17.193, df=22, p>0.05). Por lo tanto, es posible sostener que, el área de defunción según el municipio no tiene una variación significativa. Lo anterior es totalmente acertado teniendo presente que en el análisis descriptivo la mayor cantidad de fatalidades se presentaron en zona Rural Dispersa.

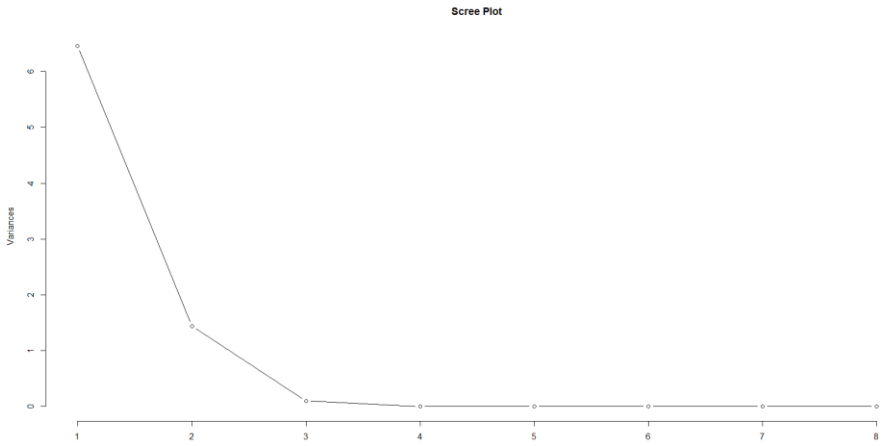
La gráfica 6, es importante a la hora de verificar las relaciones lineales entre cada una de las variables de interés, por medio de esta gráfica podemos definir que no existe correlación entre la mayoría de las variables que componen la base de estudio.

Gráfica 6
Correlaciones entre las Variables



La gráfica 7, permite determinar que las componentes 1, 2 y 3 pueden ser consideradas para el desarrollo del trabajo, aún así, se tendrán en cuenta las componentes 1 y 2:

Gráfica 7
Scree Plot



Para visualizar los eigenvalues (autovalores): Número componentes se escoge en función a los eigenvalues: Componentes = eigenvalues mayores a 1, si el valor es mayor a 1 se sugiere considerar ese componente principal en el análisis del presente estudio, por esta razón se seleccionaron los componentes 1 y 2 respectivamente: Comp. 1 = 1.83 y Comp. 2 = 1.52.

Para ver los eigenvalues (autovalores): Número componentes se escoge en función a los eigenvalues: Componentes = eigenvalues mayores a 1:

Imagen 3
Valores de los compontes principales

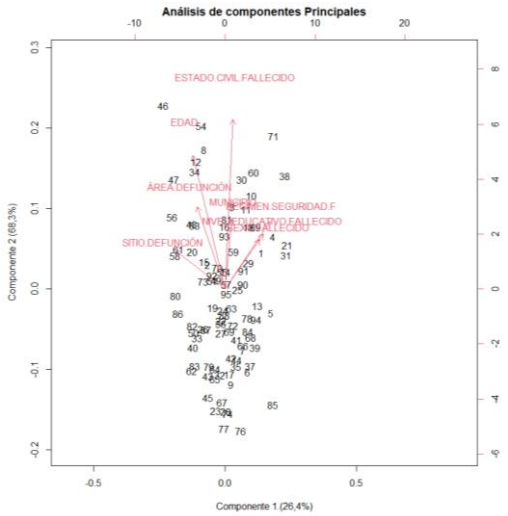
```
Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8
 1.834  1.520  1.182  1.123  0.719  0.655  0.524  0.443
```

Posteriormente se realiza el calculo de la proporción de varianza entre las componentes, verificando una importante proporción de varianza entre las componentes 1 y 2: Comp. 1: 26,4% y Comp. 2: 68.3%

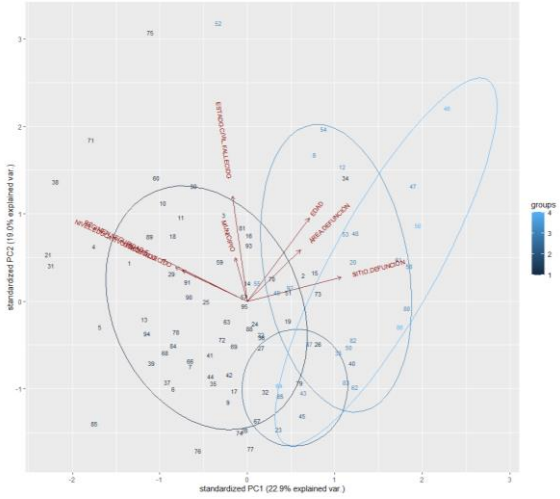
Imagen 4
Proporción de varianza entre las componentes

```
Loadings:
                          Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6
MUNICIPIO                         0.264  0.683  0.274  0.465  0.288
EDAD                      -0.342  0.507  0.226        -0.393 -0.240
ÁREA.DEFUNCIÓN            -0.295  0.313 -0.390  0.506  0.360  0.150
SITIO.DEFUNCIÓN           -0.521  0.146 -0.382 -0.182         0.309
SEXO.FALLECIDO             0.361  0.188 -0.247 -0.542  0.544
ESTADO.CIVIL.FALLECIDO            0.641        -0.298        -0.180
NIVEL.EDUCATIVO.FALLECIDO  0.399  0.207 -0.305  0.492        -0.486
RÉGIMEN.SEGURIDAD.F        0.471  0.254 -0.163        -0.443  0.687
```

Gráfica 8
Análisis de Componentes Principales



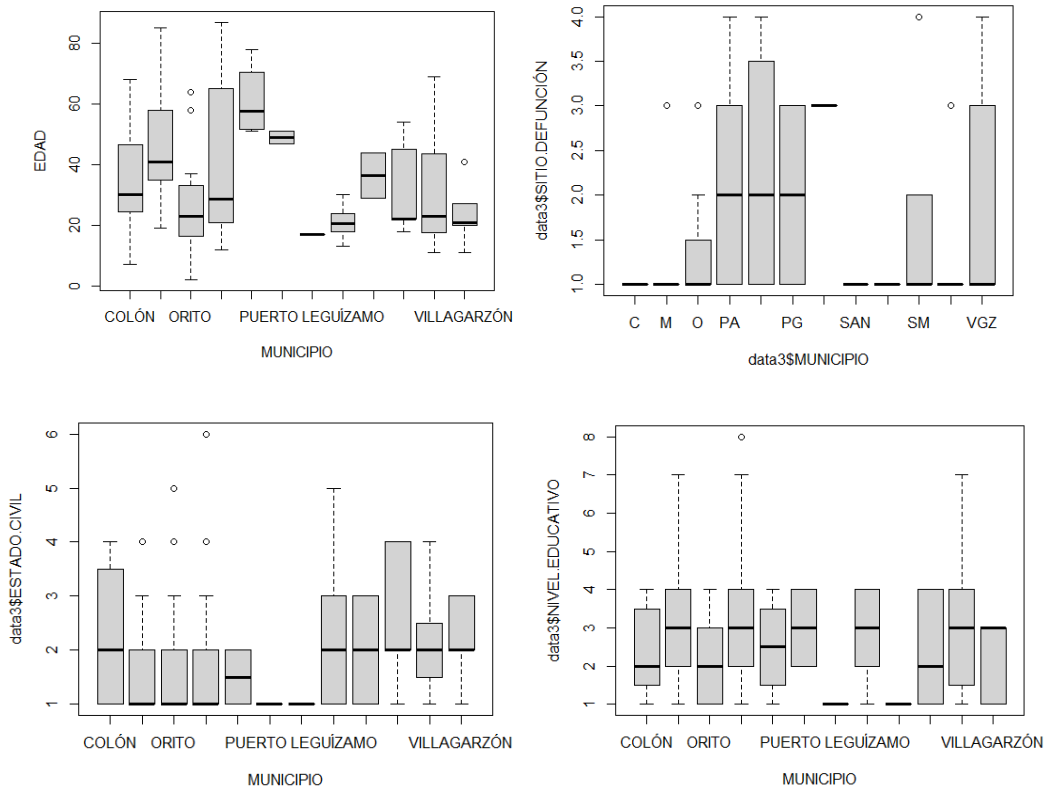
Gráfica 9
Componentes Principales Estandarizados para La variable Sitio de defunción



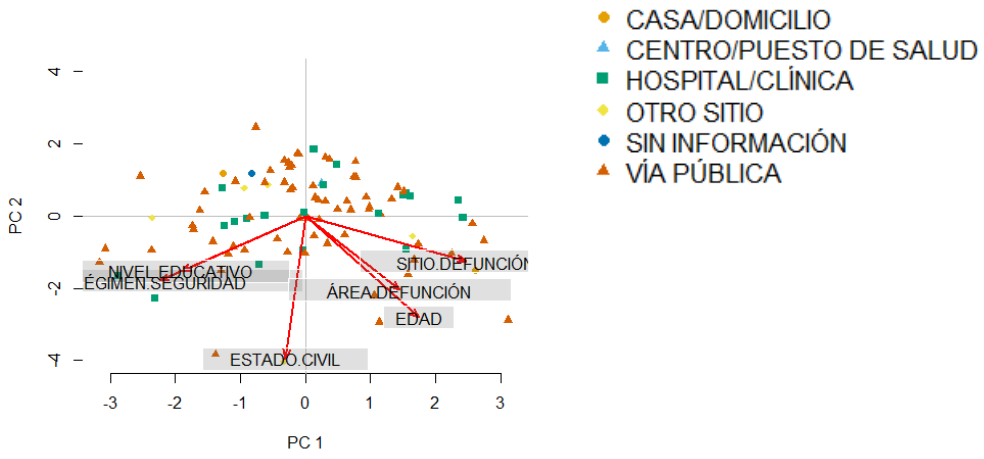
Mediante la gráfica 8, análisis de componentes principales, se puede concluir algunos patrones de la base de datos, por ejemplo: las variables incluidas en el estudio área de defunción vs EDAD, tienen varianzas diferentes, pero posiblemente presentan algo de correlación entre las variables. La nube de puntos conformada por la edad de los individuos fallecidos permite establecer que se tienen edades que no se encuentran dentro las variables, esto puede deberse a que la mayoría de las fatalidades ocurrieron en un rango de edad previamente identificado en el análisis descriptivo.

A continuación, se presentan algunos boxplot por medio de los cuales es posible verificar como en Puerto Asís se presentan mayor número de personas fallecidas con edad entre los 20 y 30 años con media aproximada a los 25 años, edad muy similar a la calculada en la Imagen 5 Resúmenes estadísticos descriptivos para cada variable.

Gráfica 10
Boxplot para las variables edad, sitio de defunción, estado civil y nivel educativo



Gráfica 11
Componentes principales en 2d



4. RESULTADOS

Durante el desarrollo del presente trabajo, se realizó el análisis descriptivo de las variables de interés, información que permitió identificar algunos patrones de la base de datos como por ejemplo que la mayoría de los accidentes de transito se generan en el municipio de Puerto Asís con un total de 22 fallecimientos la mayoría en dados en zona rural dispersa, esto es importante teniendo presente que Puerto asís es el municipio de Putumayo con mayor densidad población (67.211 habitantes según censo 2020), seguido por el municipio de Mocoa que cuenta con un total de 14 fatalidades siendo el segundo municipio de mayor densidad poblacional con un total de 58.938 habitantes. Los municipios que presentaron menor número de fatalidades cada uno con 2 fallecidos son Puerto Guzmán y Santiago, el primero con un censo poblacional para el 2020 de 36.665 habitantes y el municipio de Santiago con un total de 10.776 habitantes, lo que refleja una menor cantidad de accidentes fatídicos vs el número de habitantes por municipio.

Mediante el análisis descriptivo realizado a las variables fue posible visualizar que la mayoría de los accidentes se generaron en zona rural dispersa con un total de 54 fallecimientos, esto puede deberse a varios factores como son: mal estado de las vías, falta de señalización, iluminación deficiente y transito a altas horas de la noche por parte de los conductores de vehículos sin el uso de equipos de protección EPP o posiblemente algunos en estado de embriaguez, así mismo ocurre para la zona centro poblado la cual cuenta con un total de 39 fallecidos.

Es importante resaltar que la mayoría de casos de muertes en accidentes de transito son generados por el género masculino con un total de 87 muertes generadas por accidentes de transito vs 12 muertes ocasionadas por los mismos hechos en mujeres. La mayoría de los accidentes fatales fueron generados por hombres solteros con un total de 46 casos, un valor alto si es comparado con las 5 mujeres casadas que perecieron en accidentes de transito en el departamento de Putumayo.

La mayoría de las muertes se presentaron en personas de edades entre los 19, 20, 21, 22 y 23 años respectivamente, siendo la edad promedio de vida de 25 años y la expectativa de vida máxima 85, algo particular que se logró identificar en el estudio realizado es que la mayor cantidad de fatalidades se presentaron en el año 2020, con un total de 53 fatalidades vs 42 casos presentados en el año 2019. Lo anterior es algo importante debido a que, durante el año 2020, se llevo a cabo el aislamiento preventivo obligatorio por parte del Gobierno Nacional debido al SARS COVID -19, razón por la cual es importante considerar porque el mayor número de casos fatales se presentó durante este periodo.

Otro patrón identificado en el estudio es que la mayoría de los casos de accidentes fatales fueron generados por personas de régimen contributivo subsidiado con un total de 52 personas, y un nivel de escolaridad bajo principalmente básica primaria con un total de 27 personas, 27 personas lograron realizar decimo y undécimo y solo 17 culminaron la totalidad de básica secundaria.

Posteriormente se realizó el análisis estadístico multivariado para lo cual se utilizó el método de Análisis de Componentes Principales (ACP), en esta etapa fue necesario ajustar variables categóricas a numéricas. El primer resultado generado es no existe asociaciones a simple vista entre las variables de interés, lo que se puede verificar mediante la gráfica 6, una vez, se determinan la existencia o no de relación entre las variables se procedió a gráficas el scree plot (ver gráfica 7) por medio del cual se determinaron el número de componentes que mediante análisis de componentes principales

(ACP) se deben considerar dentro del análisis, pare ello es importante considerar los autovalores y los autovalores, que sean mayores a 1, ya que como regla general si estos valores arrojan resultados mayores a 1 es recomendable considerar ese componente principal dentro del análisis, recordemos en este punto los valores de los componentes determinados en el análisis de componentes principales (ACP): Comp. 1 = 1.83 y Comp. 2 = 1.52.

Como resultado final mediante el análisis de componentes principales (ACP) se puede definir el porcentaje que alcanza hacer explicado por las componentes seleccionadas: Comp.1 = 26.4% y Comp.2 = 68.3%

5. CONCLUSIONES

A partir del estudio realizado y los resultados obtenidos en el desarrollo del trabajo se generaron las siguientes conclusiones:

- Se identifico que la mayoría de fatalidades se están generando en personas jóvenes, principalmente de genero masculino entres los 19 a 23 años respectivamente, siendo la expectativa de vida de 25 años para el periodo de 2019 y de 33 años para el periodo de 2020,
- El mayor número de accidentes de transito se presentan en los municipios con mayor número de habitantes y principalmente en zona rural dispersa con un total de 54 fatalidades, seguido por accidentes ocurridos en Cabecera Municipal con un total de 39 fatalidades,
- El municipio de Puerto Asís el de mayor número de accidentes fatales con una frecuencia de 22 accidentes fatales, seguido por Valle del Guamez con 15 fatalidades,
- Los casos de fatalidades se generaron en bajos niveles de escolaridad, un total de 27 fallecidos en accidentes de tránsito contaban con Básica Primaria, un caso interesante a tener en cuenta es establecer el tipo de vehículo involucrados en los accidentes y si las personas fallecidas contaban con licencias de conducción vigentes o no al día de accidente, esto para verificar el nivel de preparación para acceder a licencia de conducción de los implicados en los accidentes fatales,
- El modelo de análisis de componentes principales (ACP), genero dos componentes a tener en cuenta, aún así, se considera interesante realizar otros análisis adicionales como clúster o de agrupamiento para identificar otros patrones ocultos en la base de datos,
- En definitiva, el género masculino genera más accidentes de transito que el genero femenino, esto se evidencia en el número de muertes en hombres vs el número de muertes presentado en mujeres el cual presento una diferencia entre géneros de 41,
- La mayoría de las fatalidades fueron fulminantes o se dieron en vías públicas,
- La base de datos carece de otra información que podría ser importante, como es: el tipo de vehículo, el estado de alicoramiento, conductor o acompañante, entre otras.
- Los accidentes de transito en definitiva se constituyen como una de los problemas más relevantes llegando a ser una de las principales causas de muerte, la falta de sensibilización, educación en torno a temas de manejo preventivo, entendimiento de las señales de transito y normas de seguridad, es preocupante para el departamento de Putumayo y principalmente para las ciudades de mayor densidad poblacional.

- Es oportuno utilizar otras metodologías de análisis descriptivo multivariado como análisis de conglomerados o clúster y realizar comparaciones entre los métodos estadísticos para la toma de conclusiones.

6. REFERENCIAS BIBLIOGRÁFICAS

- Marcos E. Salinas Cabrera, Luis H. Vele Figueroa. (2014). Estudio científico de la accidentalidad de tránsito en el Cantón Cuenca. Universidad Politécnica Salesiana, Ecuador.
- Lil M. Moya Fernández. (2019). Análisis de Choques en la ruta 27 usando modelos multivariados. Universidad de Costa Rica, Costa Rica.
- Christian Tomalá Vergara, Eduardo Ribadeneira Molina. (2017). Análisis estadístico de las causas de accidentes de transito en la ciudad de Guayaquil. Universidad Amistad de los Pueblos en Moscú.
- Ministerio de Salud de Colombia. https://www.minsalud.gov.co/proteccionsocial/Paginas/Aplicaci%C3%B3nNacimientosyDefunciones-RUAF.aspx.
- Campo Elías Pardo. (2020). Estadística Descriptiva Multivariada. Universidad Nacional Sede Bogotá.
- Fredes, D. (2020). Prueba de independencia estadística Chi-Cuadrado. https://rpubs.com/fredesdanyel/646504.
- Margeé García Gutierrez. (2001). Lesiones no intencionales. Centro de Referencia Nacional sobre Violencia. Colombia.
- Ruiz. J, Herrera N. (2016). Accidentes de tránsito con heridos en Colombia según fuentes de información: Caracterización general y tipologías de accidentes. Colombia.