



UNIVERSIDAD NACIONAL DE CHIMBORAZO
VICERRECTORADO DE INVESTIGACIÓN,
VINCULACIÓN Y POSGRADO
DIRECCIÓN DE POSGRADO

“APLICACIÓN DE UN MODELO DE REGRESIÓN ESTADÍSTICO PARA
ANALIZAR LA PROBLEMÁTICA DE MIGRACIÓN POR PUERTOS Y
AEROPUERTOS DEL ECUADOR”

TESIS PREVIA A LA OBTENCIÓN DEL GRADO DE:
MAGÍSTER EN MATEMÁTICA APLICADA CON MENCIÓN EN
MATEMÁTICA COMPUTACIONAL

AUTOR:

Ing. Gladys Susana Morocho Yupanqui

TUTOR:

Dr. Manuel Antonio Meneses Freire, Mgs/PhD.

Riobamba, Ecuador. 2023

Certificación del Tutor

Certifico que el presente trabajo de titulación denominado: **“Aplicación de un modelo de regresión estadístico para analizar la problemática de migración por puertos y aeropuertos del Ecuador”**, ha sido elaborado por la Ingeniera Gladys Susana Morocho Yupanqui el mismo que ha sido orientado y revisado con el asesoramiento permanente de mi persona en calidad de Tutor. Así mismo, refrendo que dicho trabajo de titulación ha sido revisado por la herramienta anti plagio institucional; por lo que certifico que se encuentra apto para su presentación y defensa respectiva.

Es todo cuanto puedo informar en honor a la verdad.

Riobamba, 30 de septiembre de 2023



Dr. Manuel Antonio Meneses Freire, Mgs/PhD.

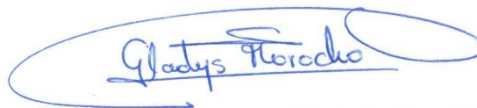
TUTOR

Declaración de Autoría y Cesión de Derechos

Yo, Gladys Susana Morocho Yupanqui con número único de identificación **060338284-7**, declaro y acepto ser responsable de las ideas, doctrinas, resultados y lineamientos alternativos realizados en el presente trabajo de titulación denominado: “Aplicación de un modelo de regresión estadístico para analizar la problemática de migración por puertos y aeropuertos del Ecuador” previo a la obtención del grado de Magíster en Matemática Aplicada con mención en Matemática Computacional.

- Declaro que mi trabajo investigativo pertenece al patrimonio de la Universidad Nacional de Chimborazo de conformidad con lo establecido en el artículo 20 literal j) de la Ley Orgánica de Educación Superior LOES.
- Autorizo a la Universidad Nacional de Chimborazo que pueda hacer uso del referido trabajo de titulación y a difundirlo como estime conveniente por cualquier medio conocido, y para que sea integrado en formato digital al Sistema de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor, dando cumplimiento de esta manera a lo estipulado en el artículo 144 de la Ley Orgánica de Educación Superior LOES.

Riobamba, 30 de septiembre de 2023



Ing. Gladys Susana Morocho Yupanqui

N.U.I. 0603382847



Riobamba, 15 de septiembre de 2023

ACTA DE SUPERACIÓN DE OBSERVACIONES

En calidad de miembro del Tribunal designado por la Comisión de Posgrado, CERTIFICO que una vez revisado el Proyecto de Investigación y/o desarrollo denominado "APLICACIÓN DE UN MODELO DE REGRESIÓN ESTADÍSTICO PARA ANALIZAR LA PROBLEMÁTICA DE MIGRACIÓN POR PUERTOS Y AEROPUERTOS DEL ECUADOR", dentro de la línea de investigación de Ingeniería Informática, presentado por el maestrante Morocho Yupanqui Gladys Susana, portador de la CI. 0603382847, del programa de Maestría en Matemática aplicada con mención en Matemática Computacional, cumple al 100% con los parámetros establecidos por la Dirección de Posgrado de la Universidad Nacional de Chimborazo.

Es todo lo que podemos certificar en honor a la verdad.

Atentamente,



JESSICA ALEXANDRA
MARCATOMA TIXI

Alexandra Marcatoma
MIEMBRO DEL TRIBUNAL



Riobamba, 15 de septiembre de 2023

ACTA DE SUPERACIÓN DE OBSERVACIONES

En calidad de miembro del Tribunal designado por la Comisión de Posgrado, CERTIFICO que una vez revisado el Proyecto de Investigación y/o desarrollo denominado "Aplicación de un modelo de regresión estadístico para analizar la problemática de migración por puertos y aeropuertos del Ecuador", dentro de la línea de investigación de Ingeniería Informática, presentado por la maestrante Gladys Susana Morocho Yupanqui, portador de la CI. 0603382847, del programa de Maestría en Matemática aplicada con mención en Matemática Computacional, cumple al 100% con los parámetros establecidos por la Dirección de Posgrado de la Universidad Nacional de Chimborazo.

Es todo lo que podemos certificar en honor a la verdad.

Atentamente,



Antonio Meneses
TUTOR DE TESIS



Riobamba, 15 de septiembre de 2023

ACTA DE SUPERACIÓN DE OBSERVACIONES

En calidad de miembro del Tribunal designado por la Comisión de Posgrado, CERTIFICO que una vez revisado el Proyecto de Investigación y/o desarrollo denominado "**Aplicación de un Modelo de Regresión Estadístico para analizar la problemática de migración por puertos y aeropuertos del Ecuador**", dentro de la línea de investigación de Ingeniería Informática (**Modelos Matemáticos**), **presentado por el maestrante Morocho Yupanqui Gladys Susana**, portador de la CI. 0603382847, del programa de **Maestría en Matemática Aplicada con Mención en Matemática Computacional**, cumple al 100% con los parámetros establecidos por la Dirección de Posgrado de la Universidad Nacional de Chimborazo.

Es todo lo que podemos certificar en honor a la verdad.

Atentamente,



SILVIA MARIANA HARO
RIVERA

Silvia Haro
MIEMBRO DEL TRIBUNAL



Riobamba 30 de septiembre de 2023

CERTIFICADO

Yo, Dr. Manuel Antonio Meneses Freire, Mgs/PhD., certifico que Gladys Susana Morocho Yupanqui con cedula de identidad N° 060338284-7 estudiante del programa de Maestría en Matemática aplicada con mención en Matemática Computacional cohorte Primera (2021), presentó su trabajo de titulación bajo la modalidad de Proyecto de titulación con componente de investigación aplicada/desarrollo denominado: **“APLICACIÓN DE UN MODELO DE REGRESIÓN ESTADÍSTICO PARA ANALIZAR LA PROBLEMÁTICA DE MIGRACIÓN POR PUERTOS Y AEROPUERTOS DEL ECUADOR”** el mismo que fue sometido al sistema de verificación de similitud de contenido **URKUND** identificando el porcentaje de similitud del 6% .

Es todo cuanto puedo certificar en honor a la verdad

Atentamente;



Dr. Manuel Antonio Meneses Freire, Mgs/PhD.

C.I. 180251584-9

Ad. Resultado de análisis

Agradecimiento

Al culminar esta etapa maravillosa de mi vida quiero extender un profundo agradecimiento, a quienes hicieron posible este sueño, aquellos que caminaron junto a mí en todo momento.

Siendo inspiración, apoyo y fortaleza. Esta mención en especial a Dios todopoderoso, a mi esposo, mis hijos y mi madre. Muchas gracias a ustedes por demostrarme que “El verdadero amor no es otra cosa que el deseo inevitable de ayuda al otro para que este se supere”

Un eterno agradecimiento, también a la Universidad Nacional de Chimborazo, a los docentes que nos brindaron sus conocimientos, consejos y estímulos para el logro de esta meta. A mis compañeros que siempre fueron apoyo para seguir adelante.

Dedicatoria

A Dios Padre Todopoderoso por darme la sabiduría y fortaleza para desarrollar mis capacidades, a la memoria de mi padre Ramón Morocho, a mi madre Fidelia Yupanqui por sus sabios consejos que han servido para hacer de mí una persona con buenos valores. En especial a mi esposo Alberto Paguay con quien día a día he luchado y compartido sudor, lágrimas y mil alegrías; a mis hijos Christian y Kevin que me han enseñado que a veces la vida es difícil, pero que no hay que rendirse jamás.

Índice General

Certificación del Tutor	
Declaración de Autoría y Cesión de Derechos	
Acta de Superación de Observaciones	
Certificado Antiplagio	
Agradecimiento	
Dedicatoria	
Índice General	
Índice de Tablas	
Índice de Figuras	
Resumen	
Abstract	
Introducción	18
Capítulo 1 Generalidades	21
1.1 Planteamiento del problema.....	21
1.2 Justificación de la investigación	22
1.3 Objetivos	24
1.3.1 Objetivo general	24
1.3.2 Objetivos específicos.....	24
Capítulo 2 Estado del arte y la práctica	25
2.1 Antecedentes investigativos.....	25
2.2 Fundamentación legal	27
2.3 Fundamentación teórica	32
2.3.1 Regresión estadística	32
2.3.2 Migración ecuatoriana.....	41
2.3.3 Diseño de modelos matemáticos en R.....	47
Capítulo 3 Diseño Metodológico	50

3.1	Enfoque de la investigación	50
3.2	Diseño de la investigación	50
3.3	Tipo de investigación	50
3.3.1	Por su finalidad.....	51
3.3.2	Por las fuentes de información	51
3.3.3	Por la manipulación de las variables.....	51
3.4	Nivel de investigación	51
3.5	Técnicas e instrumentos de recolección de datos	51
3.6	Técnicas para el procesamiento e interpretación de datos	52
3.7	Población y muestra.....	54
3.7.1	Población.....	54
3.7.2	Tamaño de la muestra	54
Capítulo 4 Análisis y Discusión de los Resultados		55
4.1	Análisis descriptivo de los resultados	55
4.1.1	Identificación de las variables	61
4.1.2	Formulación del modelo.....	61
4.1.3	Modelo de regresión logística en R.....	68
4.1.4	Cargar las librerías	69
4.1.5	Lectura de las bases de datos.....	70
4.1.6	Selección de variables y recodificación a nivel o factores	70
4.1.7	Determinación de variables para la regresión	72
4.1.8	Selección de tipos de variables.....	72
4.1.9	Transformación de la variable dependiente a numérico.....	72
4.1.10	Número de observaciones o personas en el estudio	73
4.2	Discusión de los resultados	73
4.2.1	Número de personas por vía de transporte	73
4.2.2	Correlación	74

4.2.3	Modelo de regresión logística	76
4.2.4	Predecir los valores para el conjunto de datos del test o prueba	76
4.2.5	Precisión de cálculo (Accuracy):.....	77
4.2.6	Evaluar la confianza del modelo	77
	Capítulo 5 Marco Propositivo	79
5.1	Propuesta de predicción, control y optimización de flujo migratorio en el Ecuador	79
	CONCLUSIONES	86
	RECOMENDACIONES	87
	ANEXOS	88
	Anexo A. Código de programación en R	88
	Anexo B. Lectura de base de datos	91
	Anexo C. Número de personas por vía de transporte.....	91
	Anexo D. Correlación.	91
	Anexo E. Código para resultados de la correlación.	92
	Anexo F. Código para cargar librerías necesarias para el modelo.	92
	Anexo G. Código para aplicación del modelo de regresión.....	92
	Anexo H. Código para llamar a la función del modelo de regresión multinomial	93
	Anexo I. Código para la predicción con los datos de prueba.	93
	Anexo J. Código para la predicción con los datos de prueba.....	94
	Anexo K. Código para la predicción con los datos de prueba.	94
	Anexo L. Código de programación en R. Modelo de regresión logística.....	94
	Anexo M. Clasificación de salida de ecuatorianos y extranjeros, de acuerdo con su ocupación y porcentaje correspondiente.	102
	Anexo N Coeficientes completos del modelo de regresión logística.	106
7	Referencias.....	114

Índice de Tablas

Tabla 1 <i>Cronología de cambios en los registros estadísticos de migración</i>	29
Tabla 2 <i>Tipos de regresiones</i>	32
Tabla 3 <i>Proceso de análisis de regresión</i>	35
Tabla 4 <i>Uso de R en empresas</i>	48
Tabla 5 <i>Salida de ecuatorianos y extranjeros, según tipo de vía, cantidad de personas y porcentajes</i>	56
Tabla 6 <i>Salida de ecuatorianos y extranjeros según sexo</i>	56
Tabla 7 <i>Salida de ecuatorianos y extranjeros según nacionalidad</i>	57
Tabla 8 <i>Salida de ecuatorianos y extranjeros según motivo</i>	58
Tabla 9 <i>Salida de ecuatorianos y extranjeros según motivo top 10 de ocupaciones registradas</i>	59
Tabla 10 <i>Salida de ecuatorianos y extranjeros según el mes durante el 2021</i>	60
Tabla 11 <i>Determinación de variables para la base de datos</i>	71
Tabla 12 <i>Número de personas por vía de transporte</i>	73
Tabla 13 <i>Resultados del entrenamiento</i>	75
Tabla 14 <i>Resultado función Predict</i>	76
Tabla 15 <i>Flujo migratorio en Ecuador 1997 - 2021</i>	80
Tabla 16 <i>Variables y coeficientes del modelo A</i>	83
Tabla 17 <i>Variables y coeficientes del modelo B</i>	84
Tabla 18 <i>Variables y coeficientes del modelo C</i>	85

Índice de Figuras

Figura 1 <i>Un PCS puede ser considerado una «mesa virtual» formada por la comunidad portuaria</i>	19
Figura 2 <i>Salidas de ecuatorianos según motivo de viaje</i>	23
Figura 3 <i>Componentes del código de buenas prácticas estadísticas</i>	31
Figura 4 <i>Tipos de análisis de regresión logística</i>	33
Figura 5 <i>Función de curva sigmoïdal</i>	35
Figura 6 <i>Curvas del modelo de regresión logística</i>	36
Figura 7 <i>Flujo migratorio 2011-2021</i>	42
Figura 8 <i>Entradas y salidas de ecuatorianos y extranjeros 2017 - 2021</i>	43
Figura 9 <i>Entradas y salidas de ecuatorianos por sexo</i>	43
Figura 10 <i>Entradas y salidas de ecuatorianos según medio de transporte</i>	44
Figura 11 <i>Entradas y salidas de ecuatorianos según motivo de viaje</i>	44
Figura 12 <i>Entradas y salidas de ecuatorianos por grupo etario</i>	45
Figura 13 <i>Entradas de extranjeros según motivo de viaje</i>	45
Figura 14 <i>Entradas y salidas de extranjeros por grupo etario</i>	46
Figura 15 <i>Salidas de ecuatorianos según país de destino</i>	46
Figura 16 <i>Entrada de extranjeros según nacionalidad</i>	47
Figura 17 <i>Partes del paquete R</i>	49
Figura 18 <i>Fases y paso metodológicos</i>	53
Figura 19 <i>Salida de ecuatorianos y extranjeros según tipo de vía</i>	55
Figura 20 <i>Salida de ecuatorianos y extranjeros según sexo</i>	56
Figura 21 <i>Salida de ecuatorianos y extranjeros según tipo de nacionalidad</i>	57
Figura 22 <i>Salida de ecuatorianos y extranjeros según motivo</i>	58

Figura 23 <i>Salida de ecuatorianos y extranjeros según top 10 de ocupaciones registradas.</i>	59
Figura 24 <i>Salida de ecuatorianos y extranjeros según el mes durante el 2021.</i>	60
Figura 25 <i>Detección de k cluster mediante método de codo (Elbow).</i>	63
Figura 26 <i>Detección de k cluster mediante método de Silhouette</i>	63
Figura 27 <i>Comparativa de métodos</i>	64
Figura 28 <i>Agrupación de los individuos del mejor algoritmo</i>	65
Figura 29 <i>Propuesta de modelo con clúster para regresión logística</i>	66
Figura 30 <i>Propuesta de modelo aplicado de regresión logística</i>	67
Figura 31 <i>Carga de paquetes de librerías en RStudio</i>	69
Figura 32 <i>Matriz de confusión</i>	74
Figura 33 <i>Provincias de puntos migratorios en el Ecuador</i>	79
Figura 34 <i>Cantidad de puntos migratorios del Ecuador por tipo de transporte</i>	80

Resumen

En la presente investigación se realizó un análisis estadístico utilizando la data migratoria obtenida del Instituto Nacional de Estadísticas y Censo INEC respecto al flujo de salida del año 2021, cuyo principal propósito fue aplicar un modelo de regresión estadístico que permita el análisis de la problemática migratoria por puertos y aeropuertos del Ecuador para predecir, controlar u optimizar los flujos migratorios. El estudio se elaboró debido al creciente flujo migratorio que se evidencia en el país desde el año 2010, con un pico de 7653 entradas y salidas de personas nacionales y extranjeras en el 2018, luego se dio un decremento a causa de la pandemia por COVID, retomando el crecimiento desde 2021, lo que derivó en la des habilitación de dos puntos de control Migratorio Puerto el Carmen y Puerto Nuevo Rocafuerte. Esta investigación metodológica aplicó un diseño no experimental transeccional, con base en los datos del INEC en el que constan 1468567 a salidas internacionales de ecuatorianos y extranjeros. Mediante el uso del software estadístico R se procedió a plantear un modelo estadístico de regresión Logit multinomial en la que se utilizó la variable dependiente Y, vía de transporte (terrestre, marítimo, aéreo y fluvial) y se tomaron 5 de las 24 variables independientes X, existentes (Tipo de nacionalidad, mes de movilidad, sexo de migrante, motivación del viaje y ocupación del migrante). De la formulación y ajuste del modelo planteado se obtuvo una precisión del 97.74 % y coeficientes de determinación de McFadden del 61.21%, y de Nagelkerke del 65.35%. Es decir, con estos resultados se pudo concluir que el modelo de regresión logística multinomial propuesto es excelente, y sirve para la predicción de las vías de transporte de salida desde Ecuador, considerando la variable descrita para elaborar planes de mejora respecto al control y optimización de los flujos migratorios.

Palabras claves: Modelo, regresión, logística, vías de transporte, flujo migratorio.

ABSTRACT

In this research, a statistical analysis was conducted using migration data obtained from the National Institute of Statistics and Census INEC regarding the outflow of the year 2021, whose primary purpose was to apply a statistical regression model that allows the analysis of the migration problem by ports and airports of Ecuador to predict, control or optimize migration flows. The study was developed due to the increasing migratory flow evident in the country since 2010, with a peak of 7653 entries and exits of national and foreign people in 2018. Then, there was a decrease because of the pandemic by COVID, resuming growth in 2021, which resulted in turning off two Migratory checkpoints, Puerto del Carmen and Puerto Nuevo Rocafuerte. This methodological research applied a non-experimental, cross-sectional design based on INEC data, which includes 1468567 international departures of Ecuadorians and foreigners. Using the statistical software R, a multinomial Logit regression statistical model was used in which the dependent variable Y, transport route (land, sea, air, and river) was used and 5 of the 24 existing independent variables X (type of nationality, month of mobility, sex of migrant, motivation for the trip and occupation of the migrant) were taken. From the formulation and adjustment of the proposed model, an accuracy of 97.74% and, McFadden's determination coefficients of 61.21% and Nagelkerke's of 65.35% were obtained. In other words, with these results, it was possible to conclude that the proposed multinomial logistic regression model is excellent and helps predict outbound transportation routes from Ecuador, considering the described variable to elaborate improvement plans regarding the control and optimization of migratory flows.

Keywords: Model, regression, logistics, transport routes, migratory flow.



Reviewed by:
Danilo Yépez Oviedo
English professor UNACH
0601574692

Introducción

De acuerdo con Vargas (2022) en la actualidad los puertos y aeropuertos son espacios estratégicos para cada país, debido a sus características naturales por ubicación geográfica. Una correcta regulación y gestión permiten un crecimiento económico local y regional, motivando la inversión. Si bien, estas concepciones están claras para los gobiernos, se puede distinguir que aquellos que logran potencializar sus puertos y aeropuertos han sido en su mayoría países desarrollados. En la zona de América del Sur, existe aún mucho por hacer respecto a la explotación de estas infraestructuras de entrada y salida de ciudadanos.

El presente estudio se justifica considerando diversos criterios, Lucenti y Méndes (2019) señalan que, los países de la región que han incorporado o están en el proceso de implementar sistemas portuarios avanzados y conectados digitalmente son Brasil, Chile, Jamaica y Panamá, a diferencia de los países del continente europeo, los cuales incorporaron sistemas portuarios avanzados en los años setenta.

Los países de América Latina y el Caribe (ALC) podrían aprovechar la transformación digital para tener puertos más eficientes. Estos iniciarían con la generación de una comunidad portuaria que incluya todos los interesados en el transporte marítimo, terrestre y aéreo. Es así que, esta investigación pretende proveer de información estadística útil para el desarrollo y optimización de los puertos y aeropuertos en el Ecuador.

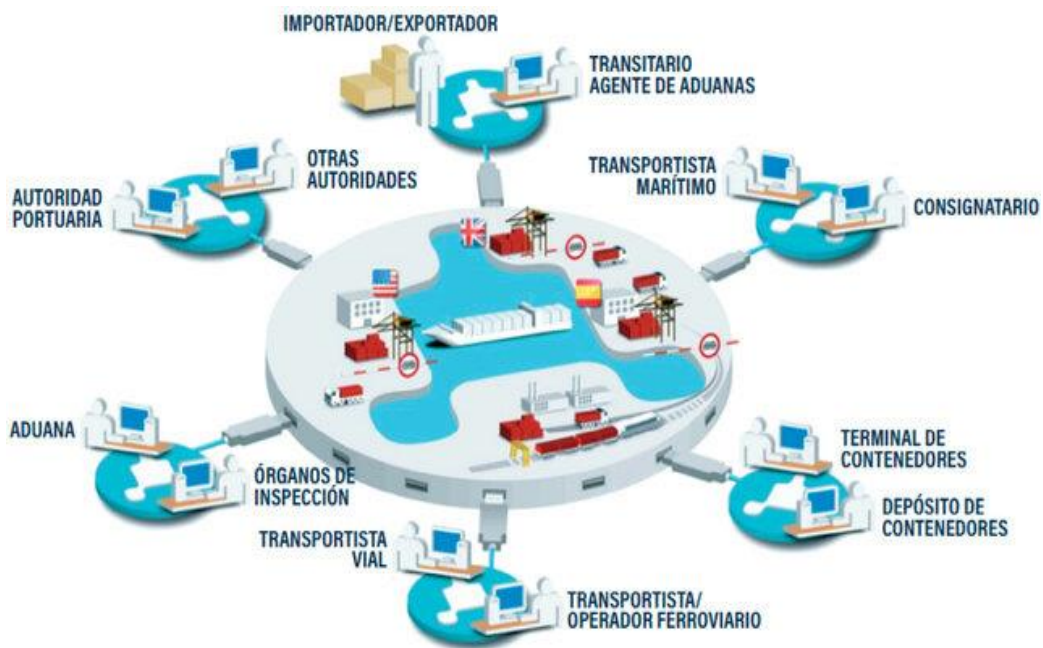
En el aspecto académico se aporta con la incorporación de modelos matemáticos que brinden una proyección del flujo migratorio, y que a futuro se consoliden en proyectos generadores de beneficios para todos los involucrados en las actividades relacionadas con los respectivos puertos o aeropuertos, beneficiando a los gremios empresariales cuyo giro de negocio se vincula entorno a estas infraestructuras. En el aspecto social, se reflejan las ventajas comerciales y desarrollo de la población, que se derivan de los proyectos aplicados.

Para ejecutar los análisis mencionados, se aplicará una metodología con enfoque cuantitativo, la misma que aborda los datos históricos o estadísticos de flujos migratorios del Ecuador y que se los utiliza como fuente para el tratamiento de modelos de regresión estadístico, que orienten sobre la incorporación de proyectos, planes, políticas públicas, entre otros que beneficien al sector.

Como resultado, se pretende alcanzar modelos de regresión que permitan una determinación de los factores que inciden en los puertos y aeropuertos, para posteriormente aplicar alguna de las alternativas válidas y comprobadas para un mejoramiento como por ejemplo de Sistemas de Comunidad Portuaria (PSC por sus siglas en inglés). En la Figura 1 se aprecia la explicación gráfica de un PSC.

Figura 1

Un PCS puede ser considerado una «mesa virtual» formada por la comunidad portuaria



Nota: Tomado de ValenciaportPCS.net (2023)

Esta investigación analiza en su primer capítulo las generalidades, planteamiento del problema, justificación y objetivos del mismo. El segundo capítulo aborda el estado del arte y los fundamentos que dan soporte al estudio, como la regresión estadística, la data de

migración ecuatoriana y el software R de diseño de modelos matemáticos. El tercer capítulo detalla la metodología, en la que constan enfoque, diseño, tipo, muestra, recolección de datos entre otros. El cuarto capítulo desglosa la obtención de resultados del programa R, mediante el que se han probado modelos matemáticos y se ha desarrollado el modelo de regresión estadístico mediante el método Logit, con sus correspondientes valores y variables.

Finalmente, se plantea en el capítulo quinto, un marco propositivo de acción en el que indique la manera de aplicar la información del modelo de regresión estadístico para proyectos o planes futuros de intervención en puertos y aeropuertos.

Capítulo 1

Generalidades

1.1 Planteamiento del problema

En Ecuador el fenómeno migratorio se presentó desde el año 1990, cuando el país sufrió una fuerte crisis económica, el principal destino de los migrantes fue España, como señalan Herrera, Carrillo, y Torres (2005):

En una primera etapa, a partir de 1998, se produce la llegada de un contingente de familias ecuatorianas al mercado de trabajo agrícola en la provincia de Murcia; posteriormente, el movimiento migratorio es masivo y las trayectorias socio espaciales, también, se encaminan a las grandes ciudades como Madrid y Barcelona.
(p. 105)

Posteriormente, la migración a continuado a diversos países, pero siendo Estados Unidos, España e Italia los de mayor atractivo por parte de los migrantes, como se identifica en Datos Macro (s.f.) “La emigración de Ecuador se ha dirigido especialmente a Estados Unidos, donde van el 43,61%, seguido por España, el 35,09% y por Italia, el 7,20%” (párr.1), actualmente las olas migratorias de ecuatorianos son latentes, sin embargo, la afluencia es menor que la de aquella época.

Efectivamente al ser la migración un aspecto cotidiano, las formas de salida de los ciudadanos son diversas, en esta investigación se revisarán las que se realizan por puertos y aeropuertos, entendiéndose además que las vías ilegales de migración también competen una problemática social como indica el estudio de Troya (2007):

En el Ecuador una de las principales formas de migrar es mediante el tráfico ilegal, en gran parte a través de la ruta marítima y en menor grado por tierra triangulando la salida por terceros países, para a su vez continuar por tierra o por vía aérea al lugar de destino. (p. 225)

Por lo expuesto, esta investigación pretende: Analizar la migración que realizan los ciudadanos ecuatorianos por vías terrestre, marítimas y aéreas, identificando salidas por puertos o aeropuertos, y aplicando un modelo de regresión estadístico que permita predicción, control u optimización del fenómeno migratorio ecuatoriano.

De acuerdo con Gutiérrez y Vladimirovna (2016) el análisis de regresión se define como una “técnica estadística utilizada para la estimación de las relaciones entre las variables, que incluye muchas técnicas para modelar y analizar varias variables, cuando la atención se centra en la relación entre una variable dependiente y una o más variables independientes” (p. 272). Por lo tanto, para la investigación propuesta se aplica esta mencionada técnica en las variables de migración en los últimos años, así como la frecuencia de salidas por puertos y aeropuertos y el modelo de regresión estadístico.

Entonces, el estudio contempla resolver los siguientes interrogantes: ¿Qué relación existe entre la migración ecuatoriana y los puertos/aeropuertos de salida? ¿Qué resultados de predicción, control u optimización de flujos migratorios por puertos y aeropuertos del Ecuador se pueden obtener utilizando un modelo de regresión estadístico?

1.2 Justificación de la investigación

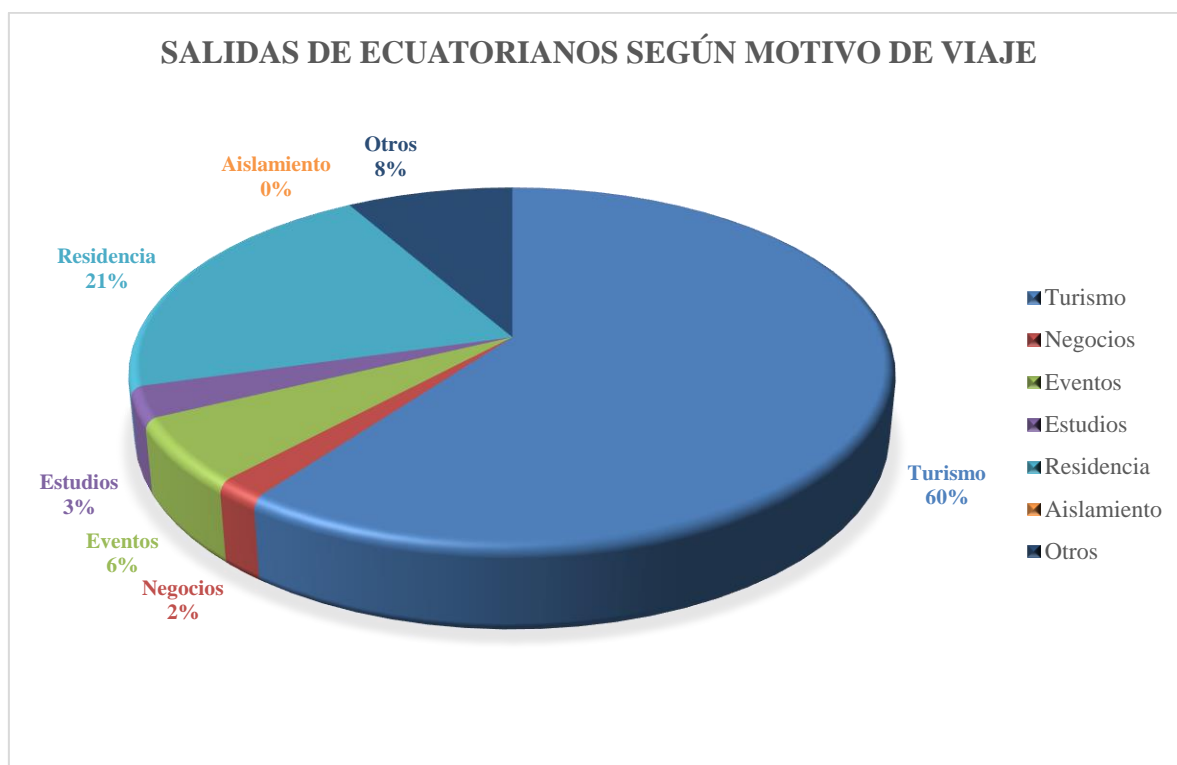
Siendo el fenómeno migratorio ecuatoriano una problemática social, económica, que de acuerdo con información de Datos Macro (s.f.) “Ecuador tiene, según los últimos datos publicados por la ONU, 1.183.685 emigrantes, lo que supone un 6,85% de la población de Ecuador” (párr.1) además “en los últimos años, el número de emigrantes ecuatorianos ha aumentado en 52.258 personas, un 4,62%”, entonces la presente investigación se realiza porque es importante conocer las vías por las cuales los migrantes realizan su desplazamiento, para que se pueda inferir condiciones de control u optimización de puertos o aeropuertos, así como establecer procesos que vigilen una migración legal, evitando así la problemática de migración ilegal, que es un conflicto social al que muchos ecuatorianos han

recorrido y que en ocasiones los ha llevado a perder la vida en la búsqueda de mejores días para su entorno.

Conforme a la información obtenida por el INEC (2020) respecto a las salidas de los ecuatorianos se detalla en la Figura 2.

Figura 2

Salidas de ecuatorianos según motivo de viaje



Nota: Adaptado de INEC (2020)

Identificando la definición de residencia: “corresponde a los movimientos de entradas o salidas internacionales cuya motivación principal se da por fines de residencia habitual en el país de destino.” (INEC, 2020, p. 9), se puede evidenciar que existe un 21.1% de viajes por esta motivación.

Por lo expuesto, para realizar la investigación se revisarán diferentes fuentes como libros, artículos científicos y tesis relacionadas que aborden el tema propuesto, además, se tomará como referencia los datos oficiales del Instituto Nacional de estadísticas y Censos (INEC),

que aporten al relacionamiento de las variables dependiente e independiente, que según Sdelsol (2020), las define como:

- Variables dependientes: Son aquellas que buscamos estudiar mediante la regresión estadística para comprender cómo se adapta al modificar las variables independientes.
- Variables independientes: Son los factores que consideramos que influyen y que afectan directamente a las variables dependientes que están bajo estudio.

Siendo fundamentalmente para este estudio la identificación y aplicación del modelo de regresión estadístico, que mejor se ajuste a las variables.

1.3 Objetivos

1.3.1 Objetivo general

Aplicar un modelo de regresión estadístico que permita el análisis de la problemática migratoria por puertos y aeropuertos del Ecuador para predecir, controlar u optimizar los flujos migratorios.

1.3.2 Objetivos específicos

- Recopilar información sobre el flujo de salida de migrantes por los puertos y aeropuertos del Ecuador.
- Seleccionar un modelo de regresión estadístico que permitan obtener relaciones entre el flujo migratorio y las salidas.
- Validar el modelo de regresión estadístico referente a las relaciones entre el flujo migratorio y las salidas.
- Aplicar el modelo de regresión estadístico seleccionado para obtener relaciones del flujo migratorio con los puntos de salida puertos o aeropuertos.

Capítulo 2

Estado del arte y la práctica

2.1 Antecedentes investigativos

De acuerdo con el artículo de López y Lorenzen denominado “Tablas de contingencia y modelos de regresión logística aplicados a cuatro rutas migratorias a partir del medio rural en Cuba” cuyo propósito es analizar las migraciones internas desde el medio rural, en Cuba; identificar las características de los migrantes, en relación con los no migrantes, y aplicando una metodología en la que se utilizan modelos de regresión logística para los cuatro movimientos migratorios: Rural-Rural-Rural, Rural-Rural-Urbano, Rural-Urbano-Rural y Rural-Urbano-Urbano, con sus pruebas de bondad de ajuste y diagnósticos.

Se pudo encontrar que cuando el destino del migrante contiene una zona urbana, el sexo influye. El movimiento migratorio se realiza en familia, para todas las rutas migratorias migran más los jóvenes; en la ruta RUR migran más las con vínculo laboral, y en la ruta RUU las más calificadas. Entre lo esencial que se concluye consta que la aplicación del método estadístico de la Regresión Logística que se presenta muestra sin lugar a dudas, resultados muy interesantes que contribuyen a ampliar el conocimiento en el tema de las migraciones. (López C. , 2014).

La investigación analizada aporta al conocimiento sobre los movimientos migratorios, se diferencia con la presente en la problemática migratoria del Ecuador, diferenciada por puertos y aeropuertos, además de otras variables relacionadas con la toma de decisión de migración del país.

En la investigación desarrollada por Castañeda, denominada “Determinantes de la migración interna juvenil por acceso a educación superior pública hacia la provincia de Pichincha” cuyo principal objetivo es analizar los factores que influyen en la migración interna de jóvenes que aceptaron un cupo en una Universidad Pública de la Provincia de Pichincha, en

el periodo 2013-2017, en el que se aplicó una metodología mediante el modelo de regresión lineal múltiple y Beta, empleados para identificar los factores que influyen en la migración. Los resultados permitieron conocer que los valores estimados muestran las significancias positivas y negativas de los factores determinantes. Inicialmente se pudo identificar que existen cinco principales razones por la que los jóvenes deciden migrar, siendo una de ellas la educación. (Castañeda, 2020). La mencionada investigación ayuda a entender factores de influencia en la migración, sin embargo, ocupa una regresión distinta a la propuesta para el presente estudio.

En el estudio de Fonseca (2002) denominado “Asociación de variables culturales y económicas con la Migración” cuyo objetivo es determinar la relación entre variables culturales y economía con la migración. Por medio de una metodología en la que aplica un modelo de regresión Logit, en donde sus variables son la migración y la ocupación, además se compararon varios modelos de regresión Logit y se usó el procedimiento de selección por eliminación hacia atrás.

Se aplicó en cada paso el test condicional por el cual la diferencia de desviación considerada con distribución chi-cuadrado con los grados de libertad correspondientes a la diferencia entre los grados de libertad de los dos modelos, uno el modelo corriente y el otro sin un término de mayor orden, permite eliminar de a un término por paso hasta obtener un modelo en el cual todos los términos de mayor orden resultan ser significativos. El estudio anterior sirvió para comparar modelos de distribución, y a diferencia del presente trabajo investigativo, se trabaja con otras variables.

De acuerdo con un análisis de regresión no lineal para el estado de Veracruz 2005-2010, denominado “El cambio climático como una causa de migración: un análisis de regresión no lineal para el estado de Veracruz 2005-2010” desarrollado por Bernabé (2018) cuyo principal propósito es explicar y observar mediante la estimación de un modelo

econométrico, si las variables climatológicas influyen en la decisión de las personas para moverse de un lugar a otro lugar, y determinar cuál es la probabilidad de que un individuo promedio del estado de Veracruz, incentivado por estos fenómenos hidrometeorológicos (efectos del cambio climático) decida o no migrar, metodológicamente señala que la regresión es observar el comportamiento de la migración (variable dependiente) con respecto a las variables explicativas de ingreso, edad, sexo, estado civil, nivel educativo, índice de desarrollo humano y de las variables hidrometeorológicas (explicadas con anticipación), obteniendo (estimando) así un valor promedio de migración en los valores de las variables independientes.

Otro de los estudios con enfoque descriptivo que determina la creciente migración venezolana en el periodo (2012- 2019), influyó sobre el comportamiento de la tasa de desempleo en Colombia durante el mismo periodo. Utilizó la técnica de regresión múltiple como metodología para realizar un análisis estadístico de datos obtenidos de fuentes secundarias, en el cual se evidenció que la migración venezolana si tuvo efectos sobre la tasa de desempleo (Bustos , 2020).

De acuerdo con las fuentes consultadas se puede evidenciar que existen investigaciones relacionadas con la migración y otros factores en los que fueron aplicados modelos de regresión estadística, sin embargo, existe una oportunidad de explorar este ámbito de investigación en virtud de la escasa información referente que relacione las variables de la presente investigación siendo la migración por las vías marítima o aérea y los modelos de regresión estadística.

2.2 Fundamentación legal

En el año 2017 la Asamblea Nacional de la República del Ecuador, promulga la Ley Orgánica de Movilidad Humana, cuyo objeto y ámbito se definen en el artículo 1 que señala:

La presente Ley tiene por objeto regular el ejercicio de derechos, obligaciones, institucionalidad y mecanismos vinculados a las personas en movilidad humana, que comprende emigrantes, inmigrantes, personas en tránsito, personas ecuatorianas retornadas, quienes requieran de protección internacional, víctimas de los delitos de trata de personas y de tráfico ilícito de migrantes; y, sus familiares. (ACNUR, 2017, p. 3)

También se revisa una definición importante que tiene que se menciona en el artículo 7. Persona en movilidad humana: “La persona que se moviliza de un Estado a otro con el ánimo de residir o establecerse de manera temporal o definitiva en él” (ACNUR, 2017, p. 4).

Además, en su Artículo 168.- Sistema Nacional Integrado de Información Sobre la Movilidad Humana, se indica que:

Se crea el Sistema Nacional Integrado de Información sobre la Movilidad Humana que al menos deberá contener datos de identidad, condición migratoria, movimientos migratorios y lugar de residencia de las personas en movilidad humana. En este sistema se consolidará la información que mantengan las entidades públicas relacionadas con la movilidad humana, la cual deberá considerar las disposiciones sobre confidencialidad. El Sistema será administrado por la autoridad de movilidad humana en coordinación con la autoridad de control migratorio y la Dirección General de Registro Civil, Identificación y Cedulación y otras entidades vinculadas con la movilidad humana, de conformidad con la legislación vigente y el reglamento de la presente Ley. (ACNUR, 2017, p. 30)

Bajo la premisa de los artículos antes mencionados se prevé que la información necesaria para realizar la presente investigación, podrá ser recopilada de forma lícita y en concordancia con la ley de movilidad humana.

A partir del año de 1985, se incorporaron todos los movimientos migratorios mediante un sistema, en la actualidad es el denominado Sistema de Migración Ecuatoriano (SIMIEC), el mismo que realiza el registro migratorio diario en las jefaturas de migración, el organismo que regula estas actividades es las Subsecretaria de Migración del Ministerio de Gobierno. Cabe hacer mención a la Jefatura de Control Migratorio que son los “lugares de ingreso y salida de pasajeros, contemplan a los pasos de frontera, puertos y aeropuertos” (Ecuador en cifras, 2022, p.11).

La cronología de los cambios en el registro estadístico de entradas y salidas internacionales, se muestran en la siguiente tabla:

Tabla 1

Cronología de cambios en los registros estadísticos de migración

Año	Principales hitos
1976	El INEC presenta por primera vez los resultados de la Encuesta de Migración Internacional. Únicamente se investigan los movimientos de las jefaturas de migración de Quito, Guayaquil, Tulcán y Huaquillas.
1979	La información recolectada comprende una muestra aleatoria simple de proporciones para Quito, Guayaquil, Tulcán y Huaquillas
1981	La información recolectada comprende una muestra aleatoria simple de proporciones para los 4 puestos y se incluyen a la investigación las jefaturas de: Macará, Machala, Esmeraldas, Salinas, Manta y Lago Agrio.
1985	La jefatura de migración de Quito comenzó a ingresar en un sistema informático los datos contenidos en la Tarjeta Internacional.
1987	Se incorpora una nueva variable en el anuario, la cual tiene que ver con el motivo de viaje tanto para ingresos y salidas de ecuatorianos, como de extranjeros
1988	Se elimina la variable de motivo de viaje para el ingreso y salida de extranjeros.
1989	Se incorpora información de entradas y salidas internacionales de ecuatorianos y extranjeros recolectada en la jefatura de migración de Ibarra.
1993	Por primera vez se presenta un informe analítico y gráfico con datos históricos
1996	Se toma en cuenta como instrumento de recolección de entradas y salidas internacionales de ecuatorianos y extranjeros, la Tarjeta Andina de Migración (TAM).
1997	Se obtiene la primera base de datos de estas estadísticas, además de cambiar el nombre de registro de estadísticas de migración internacional a Anuario Estadístico de Entradas y Salidas Internacionales.
2000	Se firma el Convenio Interinstitucional entre INEC y Ministerio del Interior (actual Ministerio de Gobierno) con el objetivo de obtener y producir estadísticas de calidad y cobertura satisfactoria. Se incorpora nuevamente la variable de motivo de viaje para entradas y salidas de extranjeros.
2007	Se incorpora una nueva jefatura de control migratorio en el cantón San Lorenzo.
2009	Se incorporan nuevas jefaturas de control migratorio en Nuevo Rocafuerte, La Balsa y Latacunga.
2014	Se incorpora una nueva jefatura de control migratorio en Santa Elena, cantón Salinas

2016	Incorporación del Sistema de Migración Ecuatoriana (SIMIEC)
2019	Se activan 3 jefaturas de control migratorio en los cantones Isabela, Santa Rosa y Lago Agrío.
2020	Se incorpora la categoría de aislamiento en la variable motivo de viaje, en el marco de la pandemia de la COVID-19
2021	Se deshabilitan dos puntos de control Migratorio Puerto el Carmen y Puerto Nuevo Rocafuerte.

Nota: Adaptado de Ecuador en cifras (2022)

La ley de estadística expedido en mayo de 1976 establece entre sus principales acápites, el artículo 10. Al Instituto Nacional de Estadística y Censo (INEC) le corresponde:

- a) Elevar a consideración del Consejo Nacional de Estadística y Censos el Programa Nacional de Estadística;
- b) Realizar las labores que le sean asignadas en el Programa Nacional de Estadística;
- c) Coordinar y supervisar la ejecución de los programas y planes de trabajo que deben realizar las demás instituciones del Sistema Estadístico Nacional;
- d) Operar como centro oficial general de información de datos estadísticos del país;
- e) Hacer inventarios estadísticos y mantener un archivo centralizado de todos los formularios, boletas, cuestionarios, instrucciones, y más instrumentos de registro que utilice el Sistema Nacional, para obtención de sus estadísticas;
- f) Canalizar la actualización y la preparación, a través del Instituto Geográfico Militar, de la cartografía necesaria para la ejecución de las investigaciones que realicen las dependencias que conforman el Sistema Estadístico Nacional;
- g) Realizar los censos de población y vivienda, agropecuarios, económicos y otros, y publicar y distribuir sus resultados, previo conocimiento de la Secretaría General del Consejo de Seguridad Nacional y del Estado Mayor Conjunto de las Fuerzas Armadas;
- h) Coordinar el funcionamiento de las comisiones especiales;

- i) Proporcionar a la Dirección General de Movilización Nacional y a la Dirección de Movilización del Estado Mayor Conjunto, los datos estadísticos que le soliciten;
- j) Las demás señaladas en esta Ley y sus Reglamentos. (Banco Central del Ecuador, 2022, p. 1)

Evidenciando la gran responsabilidad en términos de generación de datos e información de la población ecuatoriana, así también de acuerdo al artículo 13 de la Ley de Estadística, el (INEC) “tiene la facultad de crear “Comisiones Especiales de Estadística” que funcionaran como organismos auxiliares y asesores del INEC y son conformados por representantes de las instituciones productoras y usuarias de la información estadística de determinado sector”. Entonces mediante la referenciada facultad asignada al INEC, se creó la Comisión Especial de Estadísticas de Migración, en la cual se definió y estandarizó un conjunto de indicadores sobre migración internacional, necesarios para el diseño, ejecución y evaluación de las políticas públicas, según menciona la Comisión Especial de Estadísticas de Migración citada en (Ecuador en cifras, 2020).

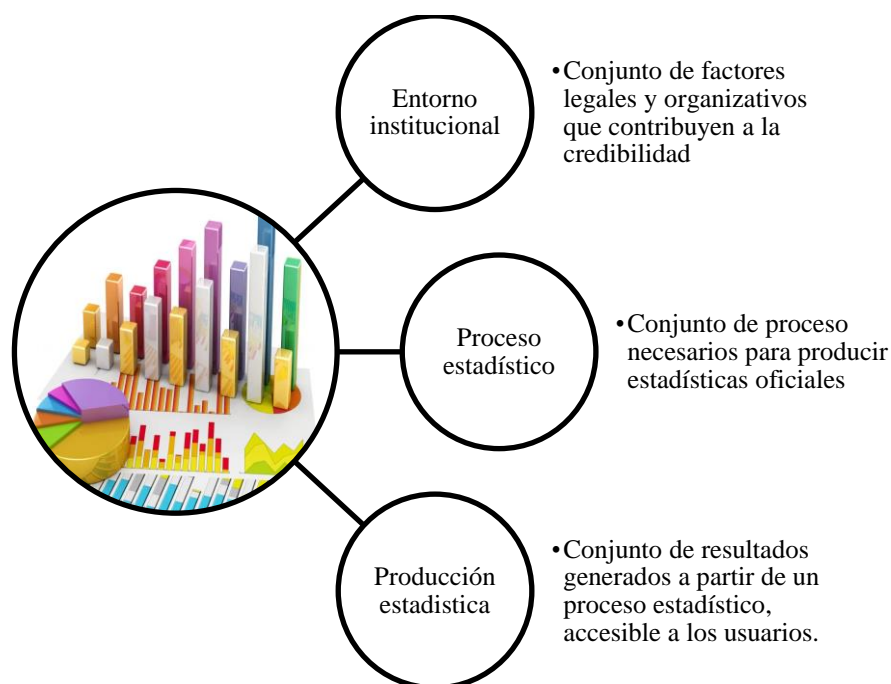
Finalmente en este estudio se cita el código de buenas prácticas estadísticas, cuyo objetivo es:

Establecer los principios y buenas prácticas que norman y regulan la producción de estadísticas oficiales de las entidades que conforman el Sistema Estadístico Nacional, estadísticas que son construidas en base a censos, encuestas, estadísticas basadas en registros administrativos y estadísticas de síntesis. (INEC, 2022, p. 1)

La norma técnica de buenas prácticas estadísticas contiene 17 principios y sus componentes se dividen en: entorno institucional, proceso estadístico y producción estadística, que se describen en la Figura 3.

Figura 3

Componentes del código de buenas prácticas estadísticas



Nota: Adaptado de INEC (2022)

2.3 Fundamentación teórica

2.3.1 Regresión estadística

De acuerdo con Estamática (2020) define al modelo de regresión estadística como “un procedimiento que puede ser usado siempre que una variable dependiente de respuesta cuantitativa pueda ser expresada como función de otra variable también cuantitativa, o de una combinación de variables independientes (cuantitativas y/o dicotómicas)” (p. 1).

Conforme a la definición de regresión estadística, es preciso conocer de forma breve cuáles son los tipos de modelos que se utilizan, es por ello que se realiza una clasificación de los modelos de regresión estadística, que se resumen en la Tabla 2.

Tabla 2

Tipos de regresiones

Tipo de regresión		Descripción
Regresión Simple	Lineal	modelo de ajuste para predecir la relación lineal existente entre la variable dependiente Y, y la independiente X, conviene estudiar los índices para cuantificar dicha relación lineal

Regresión Múltiple	Detectar si varias covariables independientes afectan significativamente a la variable de respuesta dependiente, y si efectivamente el modelo predictivo puede resultar potente de cara al ajuste por regresión.
Regresión logística binaria	La variable dependiente es la probabilidad de ocurrencia de un evento dicotómico, como suele ser la probabilidad de éxito de reintervención quirúrgica.
Regresión logística	Herramienta útil para estudiar el efecto de una o más variables independientes, sobre una variable respuesta binaria.
Regresión logística Multinomial	Detectar la influencia de los predictores continuos o dicotómicos en la probabilidad de ocurrencia de la variable de respuesta de más de 2 categorías
Regresión de COX	La regresión de Cox genera un modelo predictivo para datos de tiempo de espera hasta el evento. El modelo genera una función de supervivencia que pronostica la probabilidad de que se haya producido el evento de interés en un momento dado t para determinados valores de las variables predictoras. (IBM, 2022)

Nota: Adaptado de Estamática (2020)

Dentro de esta clasificación de los tipos de regresiones se ha realizado un análisis de cuál de estos sería el modelo que mejor aplicación tendría en la investigación, considerando las variables que se utilizan para el estudio y el tipo de datos que les corresponden.

López y Fachelli (2015) definen al análisis de regresión logística como “una técnica estadística multivariable destinada al análisis de una relación de dependencia entre una variable dependiente y un conjunto de variables independientes, de forma similar a como actúa el análisis de regresión lineal clásico” (p. 5). Siendo entonces su propósito el de obtener predicciones del comportamiento, a partir de un suceso que lo define la variable dependiente y que está en relación con la variable predictora.

Esta técnica de regresión logística apareció en los años sesenta y su utilización se generalizó en la década de los ochenta, gracias a la informática aplicada. Se caracteriza por mezclar dos prácticas del análisis estadístico.

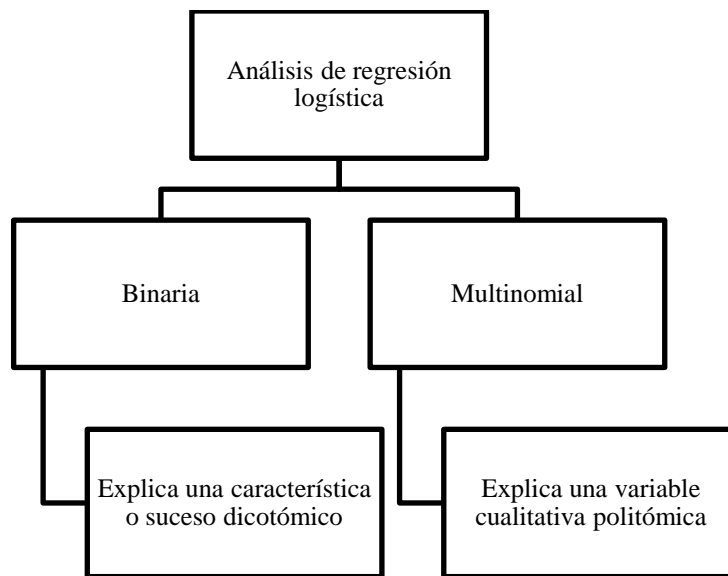
- Análisis de tablas de contingencia con el tratamiento de modelos log-lineales
- Análisis de regresión por mínimos cuadrados.

Así también las modalidades del análisis de regresión logística se pueden identificar en la

Figura 4:

Figura 4

Tipos de análisis de regresión logística

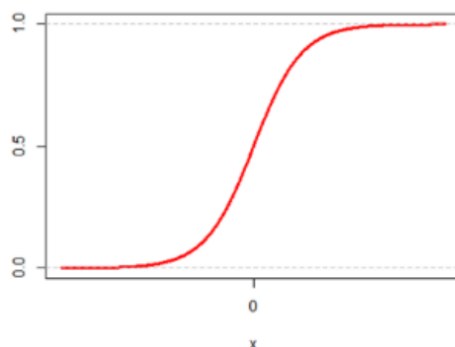


Nota: Adaptado de López y Fachelli (2015)

Entonces respecto a las variables independientes, éstas pueden ser cualitativas, tanto dicotómicas como politómicas, o cuantitativas, y se puede considerar tanto el efecto individual de cada una como el efecto de la interacción (López y Fachelli, 2015, p. 6). Por lo tanto, para la presente investigación, se podrán realizar pruebas con este tipo de variables. La forma de la función de regresión binomial es una curva sigmoideal como lo indica la Figura 5.

Figura 5

Función de curva sigmoideal



Nota: Tomado de López y Fachelli (2015)

El modelo de regresión logística binaria considera dos sucesos de un fenómeno o variable Y , excluyentes y exhaustivos, que se codifican con valores 0 y 1. Si la probabilidad de que suceda uno de ellos es P , la probabilidad de que suceda la otro es igual a 1 menos la probabilidad P :

$$\Pr(Y = 1) = P$$

$$\Pr(Y = 0) = 1 - P$$

Esta conceptualización del modelo es clara y sirve para poder determinar que en el presente estudio el evento de migración será codificado con valores entre 0 y 1. Sin embargo, de forma general para el proceso de análisis de una regresión logística se desarrolla mediante las etapas detalladas en la Tabla 3:

Tabla 3

Proceso de análisis de regresión

Etapa	Descripción
Selección de las variables del modelo	Una primera tarea fundamental es que las variables se justifiquen en el contexto de unos objetivos de investigación y a partir de criterios teóricos que fundamenten la relación de dependencia. El modelo teórico puede ser más o menos elaborado.
Estimación de los coeficientes de las variables independientes	La estimación de los coeficientes o pesos de la ecuación de regresión que determinan la importancia de cada variable independiente en la explicación de la dependiente se realiza mediante un algoritmo iterativo de máxima verosimilitud propio del modelo de la regresión logística

Clasificación de los casos	En función de la ecuación de regresión logística estimada se procede a la clasificación de los individuos según la variable dependiente pronosticada
Análisis de los residuos	Con el modelo seleccionado, se puede proceder a realizar un análisis más detallado de los residuos con el fin de detectar la existencia de casos extremos, casos que difieren notablemente entre la probabilidad observada y la probabilidad pronosticada por el modelo, y cuya eliminación puede mejorar el ajuste del modelo.

Nota: Adaptado de López y Fachelli (2015)

Además, se debe considerar los factores de confusión que son aquellos que:

Durante el proceso de selección del modelo de regresión más adecuado, el que mejor se ajusta a los datos disponibles, hay que considerar un último aspecto adicional, especialmente si el proceso de selección de variables se hace mediante el método manual de obligar a que todas las variables entren en el modelo y es el propio investigador el que paso a paso va construyendo el modelo de regresión más conveniente. (De la fuente, 2022, p. 1)

Un concepto importante al construir un modelo de regresión es que pueden introducirse términos independientes únicos (una sola variable, por ejemplo, efecto del tabaco) y además las interacciones entre variables de cualquier orden (efecto del tabaco según género), si se considera que pueden ser de interés o afectar a los resultados.

En el modelo de regresión logística binaria simple, se utilizan las siguientes ecuaciones:

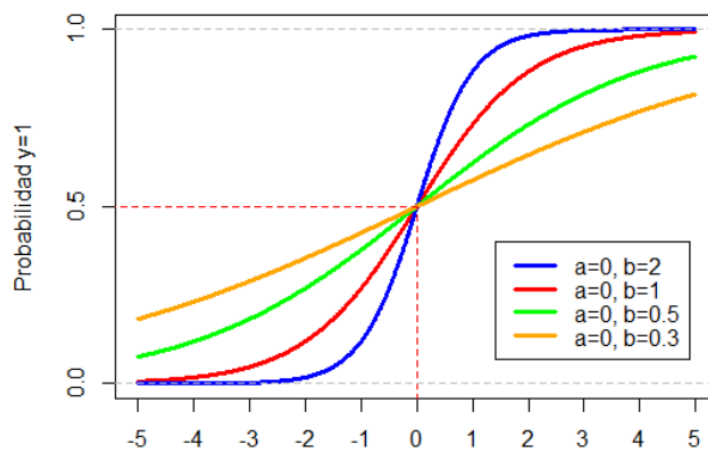
$$\Pr(Y = 1) = \frac{1}{1 + e^{-(a+bx)}} = P$$

$$\Pr(Y = 0) = 1 - \left(\frac{1}{1 + e^{-(a+bx)}} \right) = 1 - P$$

El coeficiente a representa la posición de la curva sobre el eje horizontal o de abscisas, y sitúa la curva más hacia la derecha o hacia la izquierda. El coeficiente b representa la pendiente de la curva en su punto de inflexión, en función de su valor más alto o más bajo tendremos una pendiente de la curva más inclinada o menos.

Figura 6

Curvas del modelo de regresión logística



Nota: Tomado de López y Fachelli (2015)

2.3.1.1 Regresión logística multinomial. La regresión logística multinomial es una técnica estadística utilizada para modelar y predecir variables categóricas con tres o más categorías mutuamente excluyentes. A diferencia de la regresión logística binomial, que se utiliza para variables categóricas con dos categorías, la regresión logística multinomial permite manejar múltiples categorías (Scott, 2016).

En la regresión logística multinomial, se ajusta un modelo de regresión logística separado para cada categoría en comparación con una categoría de referencia. El modelo estima la probabilidad de que una observación pertenezca a cada una de las categorías, utilizando una función Logit para modelar la relación entre las variables predictoras y las probabilidades de pertenecer a cada categoría (Hosmer, Lemeshow, & Sturdivant, 2013).

2.3.1.2 Regresión tipo Logit. Sea Y una variable dependiente binaria (con dos posibles valores: 0 y 1). Sean un conjunto de k variables independientes, (X_1, X_2, \dots, X_k) , observadas con el fin de predecir/explicar el valor de Y .

El modelo logístico establece la siguiente relación entre la probabilidad de que ocurra el suceso, dado que el individuo presenta los valores $X=X_1, X=X_2, \dots$

El objetivo es hallar los coeficientes $(\beta_0, \beta_1, \dots, \beta_k)$ que mejor se ajusten a la expresión funcional. El cociente de probabilidades se conoce como *odds* o ratio del riesgo.

2.3.1.3 Modelo de regresión logística multinomial. La ecuación de la regresión

logística multinomial es similar a la de la regresión logística binomial, pero se extiende para abarcar múltiples categorías. La ecuación general para la regresión logística multinomial se puede expresar de la siguiente manera (Hosmer, Lemeshow, & Sturdivant, 2013):

Para cada categoría j ($j = 1, 2, \dots, J - 1$), siendo J el número total de categorías, la ecuación Logit se define como:

$$\text{Logit}(P(Y = j|X)) = \beta_{0j} + \beta_{1j}X_1 + \beta_{2j}X_2 + \dots + \beta_{pj}X_p$$

Donde:

$P(Y = j|X)$ es la probabilidad condicional de que la observación pertenezca a la categoría j dado un conjunto de variables predictoras X .

$\beta_{0j}, \beta_{1j}, \beta_{2j}, \dots, \beta_{pj}$ son los coeficientes del modelo correspondientes a la categoría j .

X_1, X_2, \dots, X_p son las variables predictoras.

En la regresión logística multinomial, se ajustan $J-1$ modelos separados para las J categorías, utilizando una categoría de referencia. La categoría de referencia es aquella para la cual los coeficientes son iguales a cero. La probabilidad de pertenecer a la categoría de referencia se puede obtener mediante el complemento de las probabilidades de las categorías restantes (Hosmer, Lemeshow, & Sturdivant, 2013).

Es importante destacar que en la regresión logística multinomial, se utiliza la función softmax para transformar los logit en probabilidades. La función softmax asigna probabilidades a cada categoría y garantiza que la suma de las probabilidades para todas las categorías sea igual a 1 (Scott, 2016).

Es posible encontrar implementaciones específicas y variantes de la ecuación de la regresión logística multinomial en diferentes paquetes y software estadístico, adaptadas a las necesidades y características de cada problema en particular (Scott, 2016).

2.3.1.4 Evaluación del modelo de regresión Logit. Para evaluar un modelo de regresión logística, existen varias métricas y técnicas que se puede utilizar. Se presentan algunas a continuación:

1. Precisión (Accuracy): Es la proporción de predicciones correctas realizadas por el modelo. Se calcula dividiendo el número total de predicciones correctas entre el número total de observaciones. Sin embargo, la precisión puede ser engañosa si las clases están desbalanceadas (Gutiérrez & Vladimirovna, 2016).

2. Matriz de confusión: Es una tabla que muestra el número de predicciones correctas e incorrectas para cada clase del modelo. Permite evaluar el desempeño del modelo en términos de verdaderos positivos (TP), falsos positivos (FP), verdaderos negativos (TN) y falsos negativos (FN) (Gutiérrez & Vladimirovna, 2016).

3. Precisión Recall y F1-Score: Estas métricas se calculan a partir de la matriz de confusión y se utilizan cuando las clases están desbalanceadas. La precisión es la proporción de verdaderos positivos entre todas las predicciones positivas, el recall es la proporción de verdaderos positivos entre todos los casos positivos reales, y el F1-score es una medida combinada que tiene en cuenta tanto la precisión como el recall (Hosmer, Lemeshow, & Sturdivant, 2013).

4. Curva ROC (Receiver Operating Characteristic): La curva ROC muestra la relación entre la tasa de verdaderos positivos (recall) y la tasa de falsos positivos (1 - especificidad) para diferentes umbrales de clasificación. A partir de la curva ROC, puedes calcular el área bajo la curva (AUC-ROC), que es una

medida de qué tan bien se desempeña el modelo en general. Un valor de AUC-ROC cercano a 1 indica un buen rendimiento del modelo (Sdelsol, 2020).

5. Validación cruzada: La validación cruzada es una técnica que se utiliza para evaluar el rendimiento del modelo utilizando diferentes particiones de los datos de entrenamiento y prueba. Puedes aplicar la validación cruzada para obtener una estimación más robusta del desempeño del modelo y verificar su capacidad de generalización (Hosmer, Lemeshow, & Sturdivant, 2013).

2.3.1.5. Clustering. Los modelos Clustering son un conjunto de algoritmos que permiten clasificar un conjunto de individuos u observaciones en grupos homogéneos. Los elementos dentro de los clústers son parecidos, mientras que los elementos de diferentes clústers son diferentes (Jaramillo, 2022).

Existen diferentes algoritmos y métodos de clustering, cada uno con sus propias características y enfoques. Algunos de los algoritmos de clustering más comunes incluyen:

1. K-means: Este algoritmo busca particionar los datos en K clústeres, donde cada clúster se representa mediante su centroide. El objetivo es minimizar la suma de las distancias cuadradas entre cada punto y el centroide de su clúster asignado (Scott, 2016).

2. DBSCAN: El algoritmo DBSCAN (Density-Based Spatial Clustering of Applications with Noise) define los clústeres como regiones de alta densidad de puntos, separadas por regiones de baja densidad. No requiere especificar el número de clústeres de antemano y puede identificar puntos ruidosos o anomalías (Jaramillo, 2022).

3. Agglomerative Hierarchical Clustering: Este método comienza considerando cada punto como un clúster individual y luego fusiona iterativamente los clústeres más cercanos hasta que todos los puntos estén en un

solo clúster o se cumpla un criterio de parada (Hosmer, Lemeshow, & Sturdivant, 2013).

4. Gaussian Mixture Models (GMM): Los modelos de mezcla gaussiana asumen que los datos provienen de una combinación de varias distribuciones gaussianas. El algoritmo estima los parámetros de estas distribuciones para modelar y asignar puntos a diferentes clústeres (Jaramillo, 2022).

5. Spectral Clustering: Este enfoque utiliza técnicas de álgebra lineal y grafos para realizar el clustering. Transforma los datos en un espacio de menor dimensión y luego aplica métodos de particionamiento o corte de grafos para obtener los clústeres (Jaramillo, 2022).

2.3.2 Migración ecuatoriana

Conforme al boletín técnico N°01-2022-REESI, de Registro Estadístico de Entradas y Salidas Internacionales 2021, elaborado por él (INEC Boletín REESI, 2022) se afirma que:

El Flujo migratorio general está compuesto por la suma de entradas y salidas tanto de ecuatorianos como de extranjeros. La evolución del flujo migratorio general ha sido creciente en los últimos 20 años. Tanto las entradas como salidas internacionales aumentaron en el año 2021. (p.7)

Es importante mencionar que mediante el Sistema de Migración Ecuatoriano (SIMIEC):

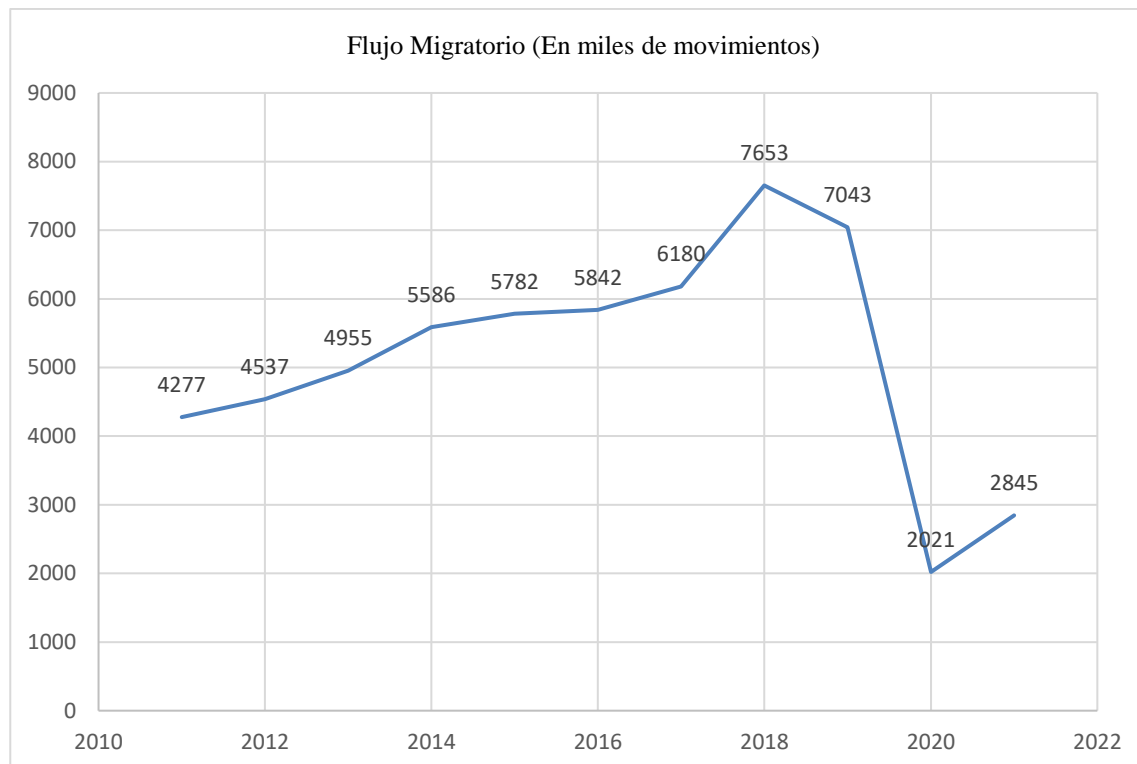
En cada jefatura de migración la información de entradas y salidas internacionales de ecuatorianos y extranjeros es captada a través del sistema SIMIEC, mismo que cuenta con parámetros de validación, esta información es validada con información de Registro Civil, Policía Judicial, aerolíneas, entre otros, además de contar con campos obligatorios de llenado. (INEC Boletín REESI, 2022, p. 6)

Mediante este sistema SIMIEC se obtienen data sobre la migración como las que se presentan a continuación.

El historial de flujo migratorio de 2011 a 2021 se muestra en la Figura 7.

Figura 7

Flujo migratorio 2011-2021

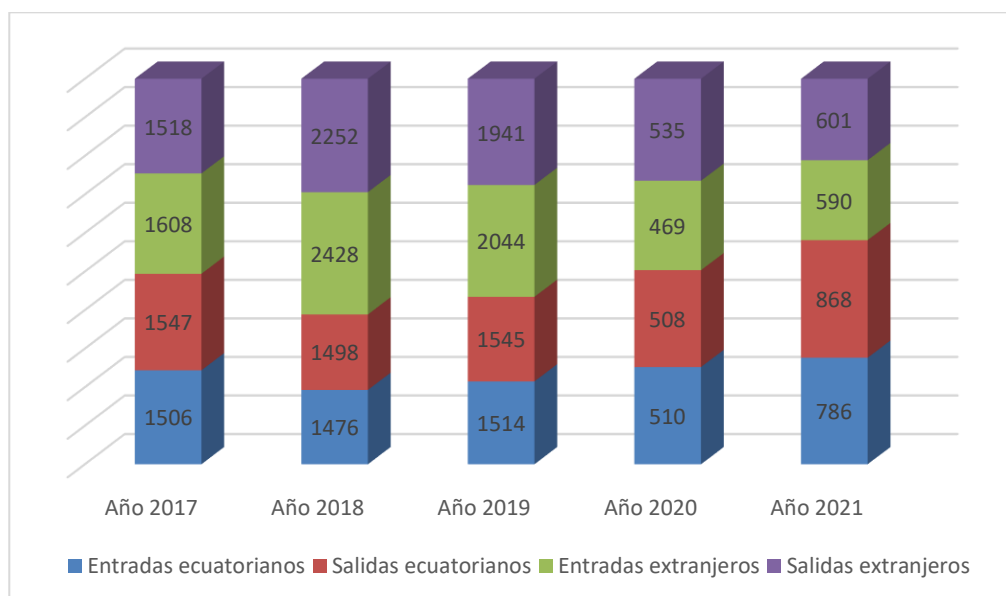


Nota: Tomado de INEC (2021)

Entradas y salidas internacionales de ecuatorianos y extranjeros entre año 2017 y 2021, datos en miles de movimientos.

Figura 8

Entradas y salidas de ecuatorianos y extranjeros 2017 - 2021

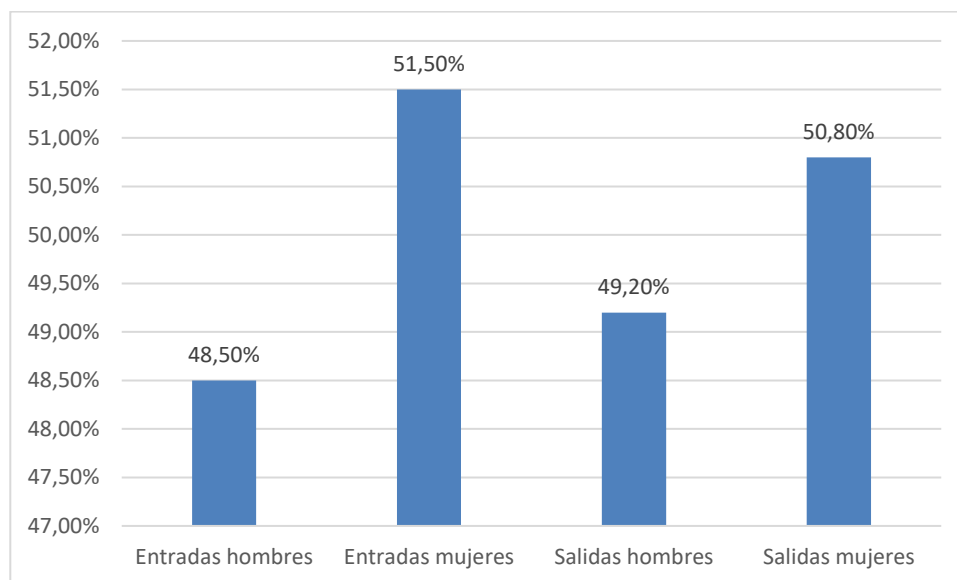


Nota: Tomado de INEC (2021)

Entradas y salidas de ecuatorianos por sexo

Figura 9

Entradas y salidas de ecuatorianos por sexo

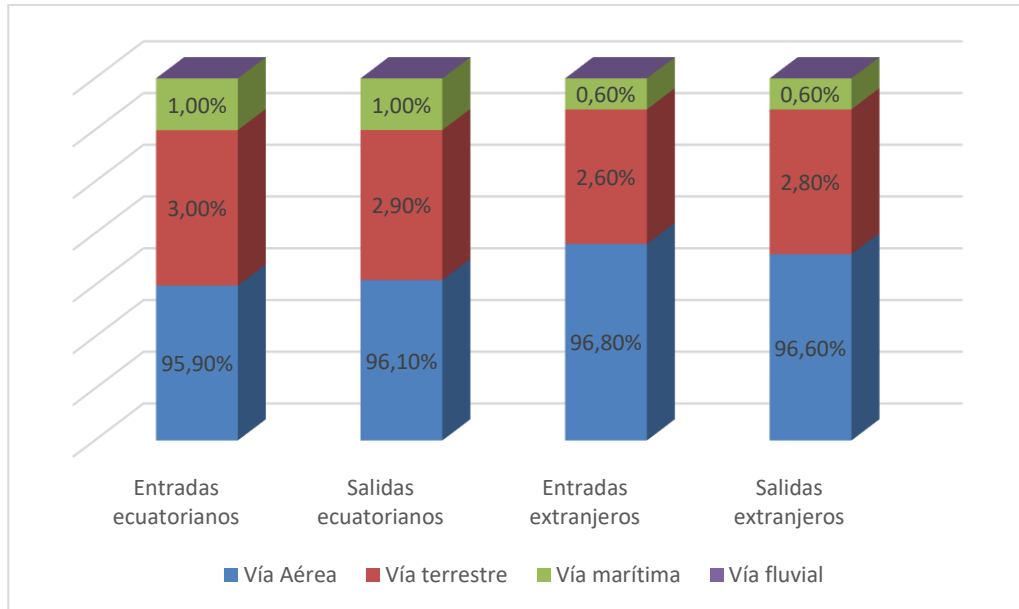


Nota: Tomado de INEC (2021)

Entradas y salidas de ecuatorianos de acuerdo al medio de transporte

Figura 10

Entradas y salidas de ecuatorianos según medio de transporte

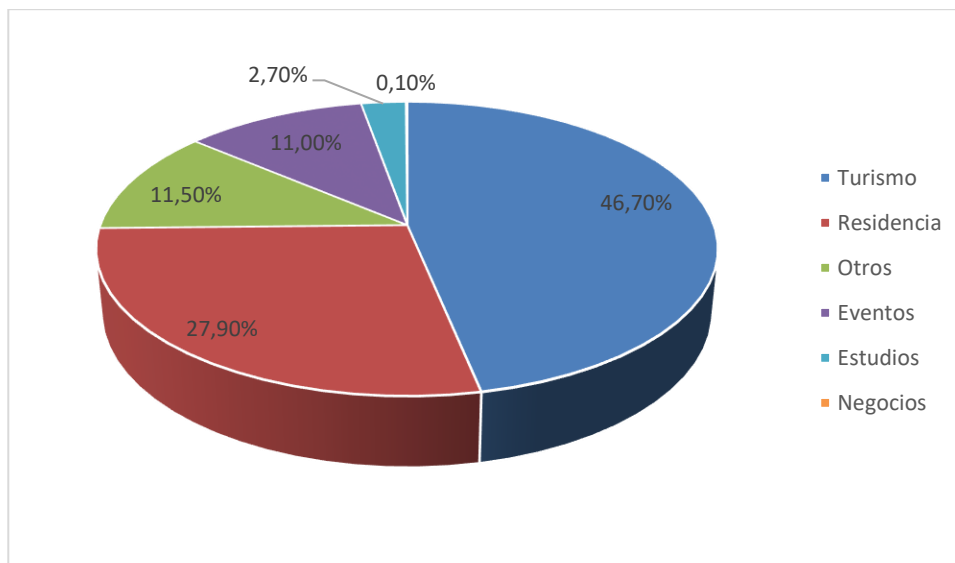


Nota: Tomado de INEC (2021)

Salida de ecuatorianos según motivo de viaje

Figura 11

Entradas y salidas de ecuatorianos según motivo de viaje

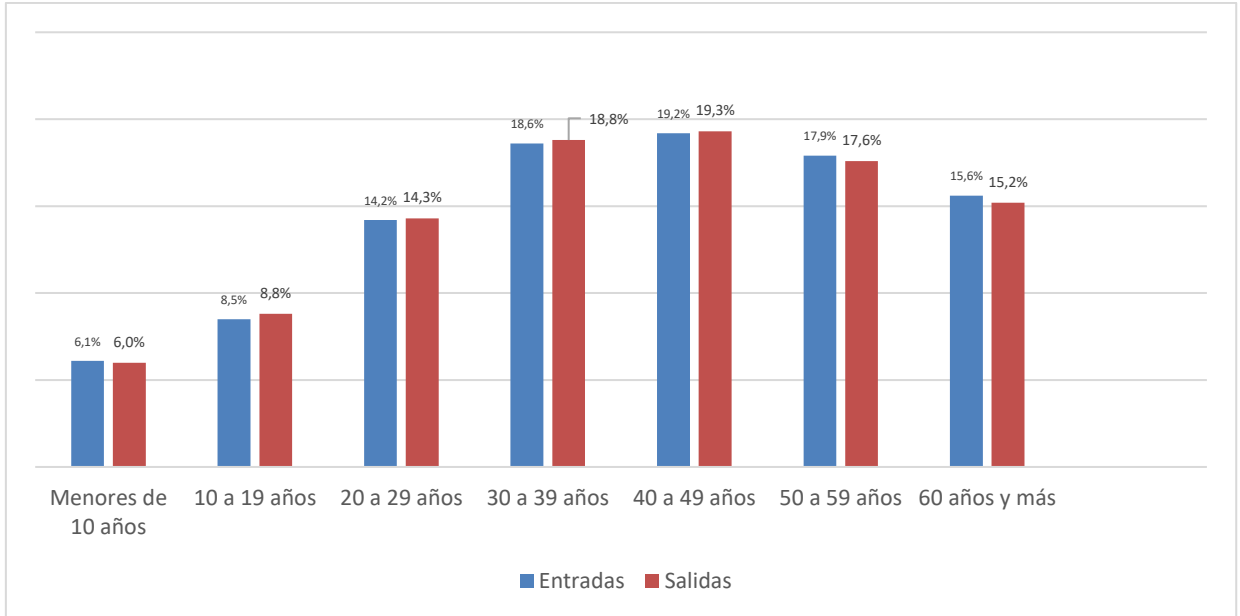


Nota: Tomado de INEC (2021)

Entradas y salidas de ecuatorianos de acuerdo a la edad.

Figura 12

Entradas y salidas de ecuatorianos por grupo etario

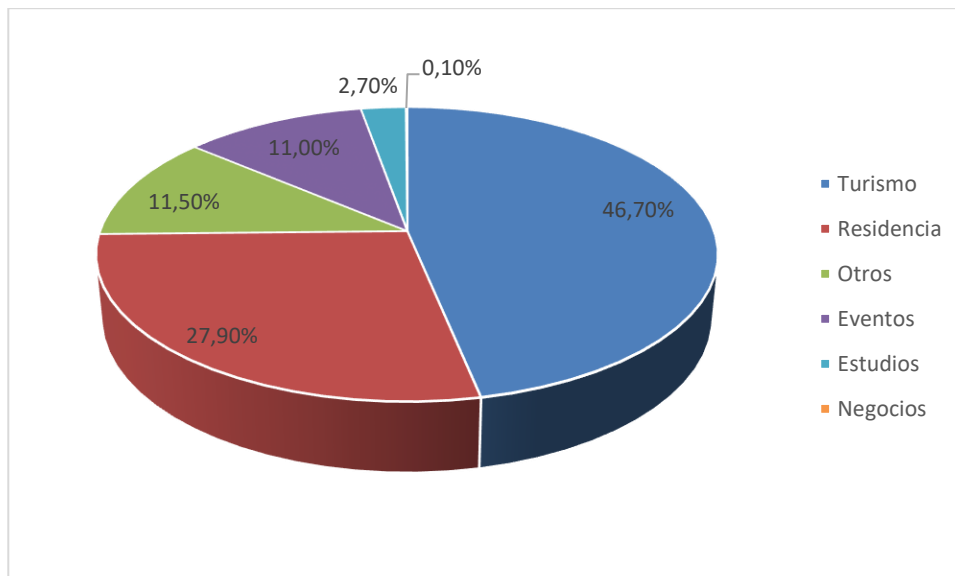


Nota: Tomado de INEC (2021)

Entrada de extranjeros según motivo de viaje

Figura 13

Entradas de extranjeros según motivo de viaje

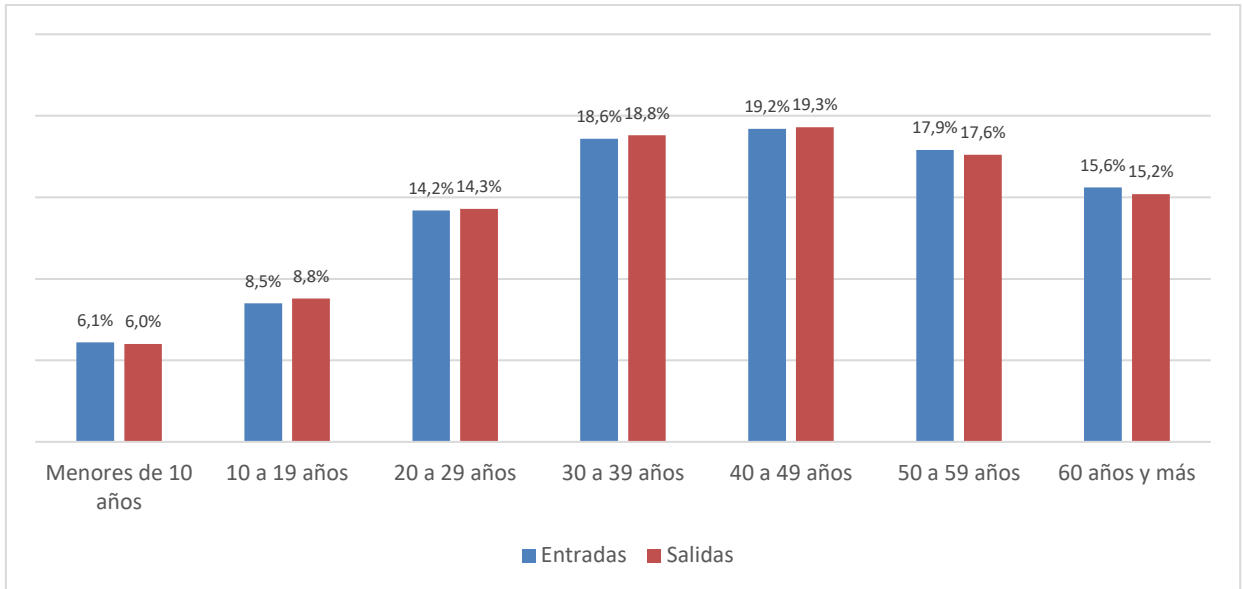


Nota: Tomado de INEC (2021)

Entrada y salida de extranjeros según grupo de edades.

Figura 14

Entradas y salidas de extranjeros por grupo etario

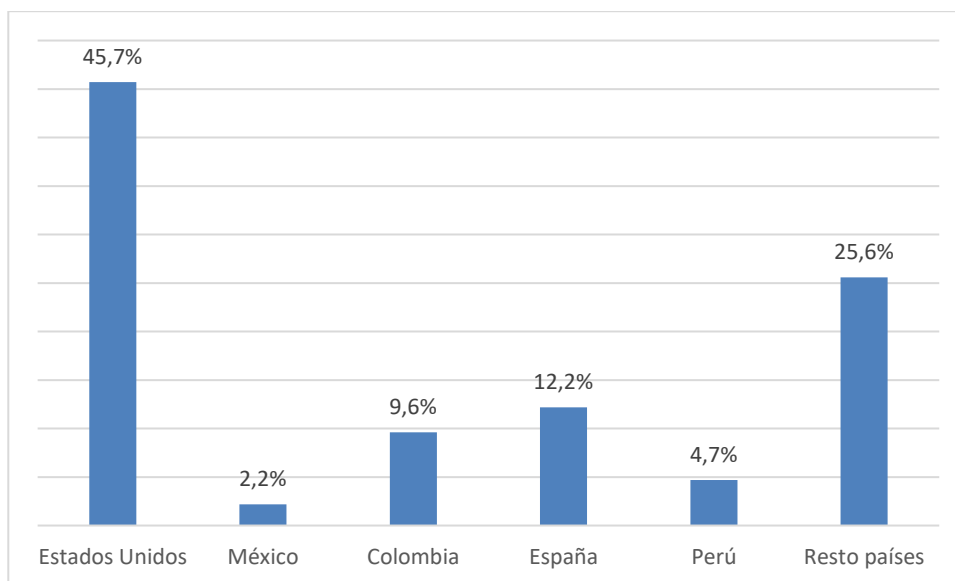


Nota: Tomado de INEC (2021)

Salida de ecuatorianos según país de destino

Figura 15

Salidas de ecuatorianos según país de destino

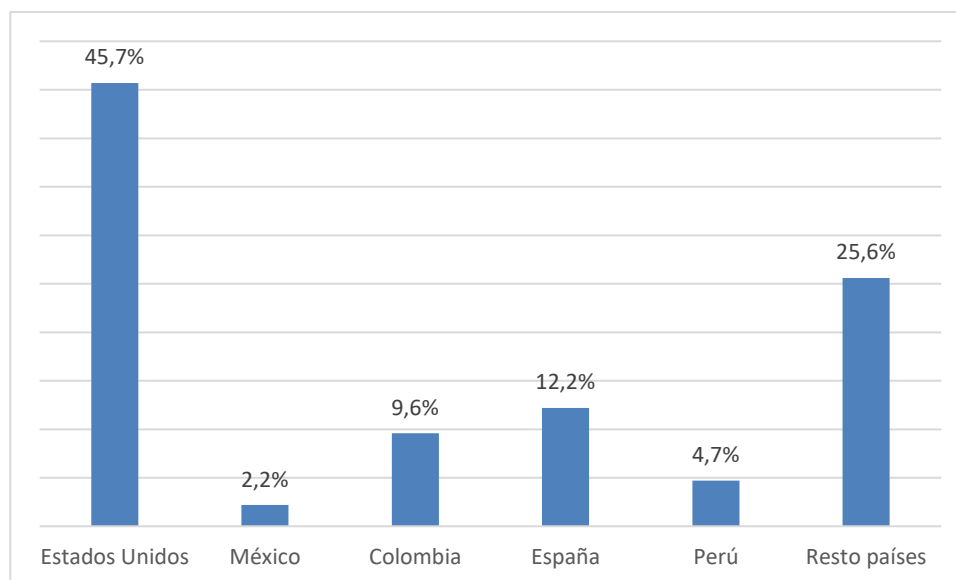


Nota: Tomado de INEC (2021)

Entrada de extranjeros según nacionalidad

Figura 16

Entrada de extranjeros según nacionalidad



Nota: Tomado de INEC (2021)

Estas fuentes de datos confiables generadas por el INEC, permiten tener la posibilidad de proyectar los flujos migratorios en el Ecuador, a fin de determinar un comportamiento y las optimizaciones en las respectivas jefaturas migratorias

2.3.3 Diseño de modelos matemáticos en R

El lenguaje de programación R fue diseñado particularmente para fines estadísticos, logrando ser una herramienta muy útil para los trabajos de modelamiento matemático.

De acuerdo con Booksdown (2022), “la adopción de R se debe en gran medida a que permite responder preguntas mediante el uso de datos de forma efectiva, y como es un lenguaje abierto y gratuito, se facilita compartir código, crear herramientas para solucionar problemas comunes” (p.1). Algunas de las empresas que utilizan R y sus aplicaciones se detallan en la Tabla 4

Tabla 4

Uso de R en empresas

Empresa	Aplicación
Facebook	Analizar la manera en que sus usuarios interactúan con sus muros de publicaciones para así determinar qué contenido mostrarles
Google	Analizar la efectividad las campañas de publicidad implementadas en sus servicios, por ejemplo, los anuncios pagados que te aparecen cuando “googleas” algo
Microsoft	Adquirió y ahora desarrolla una versión propia de R llamada OpenR, que ha hecho disponible para uso general del público. OpenR es empleada para realizar todo tipo de análisis estadísticos,

Nota: Adaptado Booksdown (2022)

Así también, cabe mencionar que, aunque R está diseñado para análisis estadístico, con el paso del tiempo los usuarios de este lenguaje han creado extensiones a R, llamadas paquetes, que han ampliado su funcionalidad. (Booksdown, 2022, p. 1).

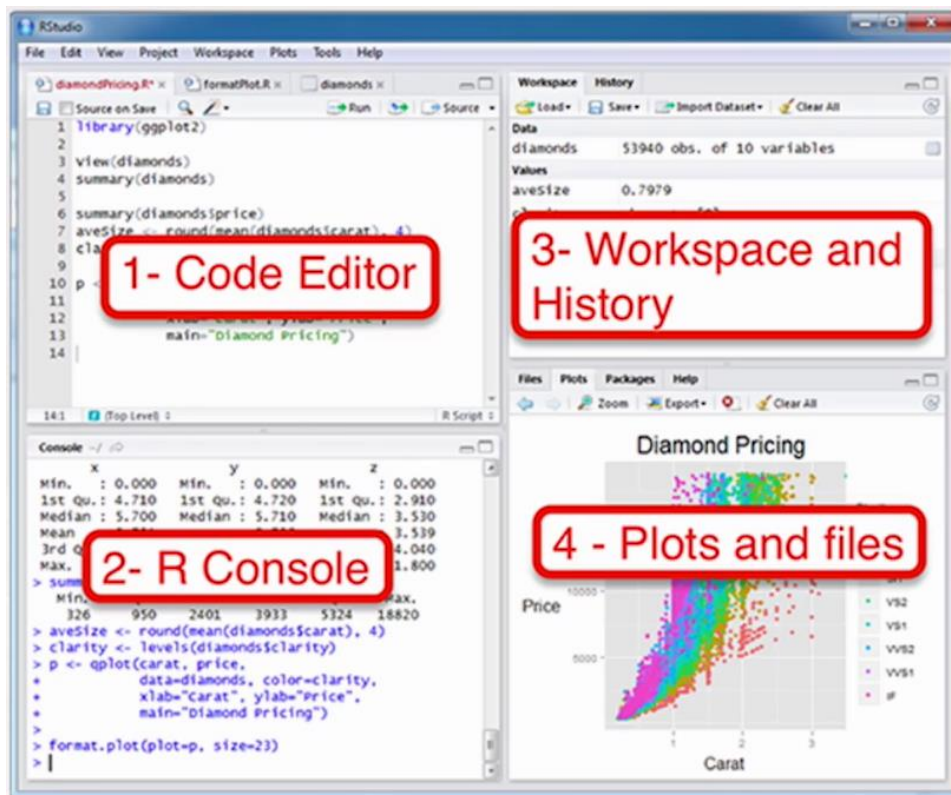
Cuando ya se ha estimado el modelo, se procede con la inferencia para extrapolar los resultados. Para ajustar un modelo en R se utiliza la función `glm` que es aquella que utiliza la estimación por mínimos cuadrados iterativamente para obtener los parámetros de los modelos.

Para poder trasladar los resultados al total de la población se debe realizar la inferencia del modelo estimado, para esto se puede utilizar los modelos de: contrastes basados en el test de Wald, test condicional de la razón o teste multiplicador de Lagrange, es recomendable aplicar también los intervalos de confianza de los parámetros, en R se utiliza la función `confint.default()`, además los valores estimados y predicciones del modelo, en R se utiliza la función `predict()`.

En la Figura 17 se muestran las partes de las que consta el paquete RStudio, utilizado para análisis masivo de modelos predictivos de regresión logística multivariante.

Figura 17

Partes del paquete R



Capítulo 3

Diseño Metodológico

3.1 Enfoque de la investigación

La metodología que permite cumplir con los propósitos del estudio, se basa en la selección y aplicación de un modelo estadístico que permita analizar la relación entre el flujo migratorio ecuatoriano y los puntos de salida puertos o aeropuertos.

La presente investigación realiza un análisis del fenómeno migratorio y relaciona con el flujo de salida, se toma la información con datos existentes del año 2021, considerando que se regularizaron las salidas por puertos y aeropuertos del Ecuador.

El enfoque de la investigación es cuantitativo, en virtud que se realiza una recopilación de datos de salidas por vías terrestres, marítimas y aéreas, así como la relación con las motivaciones de los viajes, orientando el análisis hacia aquellos que son por causas migratorias.

3.2 Diseño de la investigación

El diseño de la investigación es no experimental transeccional, debido a que se obtendrán datos oficiales respecto a las variables de estudio y las mismas no sufrirán alteraciones.

Dentro de este tipo de estudio se realizará en un tiempo determinado o único momento, por lo que es transeccional.

3.3 Tipo de investigación

La tipología de la investigación se determina por varios factores tales como el control de las variables, finalidad y fuentes de información, de manera que se explica cada una de las tipologías.

3.3.1 Por su finalidad

Al existir investigaciones previas sobre el tema que se pretende investigar, estas sirven de base para enriquecer el presente estudio, permitiendo reestructurar y mejorar las investigaciones anteriores, constituyéndose en un estudio aplicado (Niño V. , 2011).

3.3.2 Por las fuentes de información

Las fuentes secundarias utilizadas en la descripción teórica, legal y referencial de la presente investigación permiten que sea de tipo documental a fin de analizar y utilizar como base de estudio la información encontrada en la revisión bibliográfica (Bernal, 2010).

3.3.3. Por la manipulación de las variables

Las variables en estudio no serán intervenidas, se estudiará el problema migratorio en sus condiciones normales, es decir, sin manipular ninguno de los factores que causan la migración de ciudadanos ecuatorianos, por ello el estudio es de tipo no experimental (Bernal, 2010).

3.4 Nivel de investigación

Para esta investigación se propone realizar un nivel de estudio explicativo, que va más allá de la descripción de conceptos o fenómenos o del establecimiento de relaciones entre conceptos; es decir, están dirigidos a responder por las causas de los eventos y fenómenos físicos o sociales. Como su nombre lo indica, su interés se centra en explicar por qué ocurre un fenómeno y en qué condiciones se manifiesta o por qué se relacionan dos o más variables. (Hernández, et. al, 2014, p. 95).

3.5 Técnicas e instrumentos de recolección de datos

Dentro de los métodos a utilizar, se procederá con el análisis de contenidos basados en la revisión de datos de manera objetiva y sistemática, para cuantificar y proceder con tratamientos estadísticos (Hernández, et al., 2014), además, se realizará la búsqueda y contraste de otras investigaciones en el que se encuentre inmersa alguna de las variables

del estudio, por ejemplo, artículos científicos, investigaciones estadísticas relacionadas con las regresiones.

Finalmente se usará el método de datos secundarios recolectados por otros investigadores, registros documentales y públicos, se tomarán bases de datos públicas tanto de flujo de vuelos, marítimas y terrestres, y se utilizó la data de salidas con las que se trabajó en el modelo de regresión.

3.6 Técnicas para el procesamiento e interpretación de datos

Los datos que se obtendrán a través de estos portales serán consolidados en una sola base de datos a manera de poder crear una matriz adaptada para aplicar el modelo de regresión mediante el método Logit y hacer la respectiva comparativa mediante estadística inferencial.

La investigación tiene un enfoque cuantitativo de manera que el tratamiento de los datos y aplicación del modelo de regresión se hará en un software estadístico. Primero, se procede a revisar las bases de datos obtenidas de los portales digitales para luego depurar y crear una base que contemple únicamente las variables de interés. Una vez con la base definida se procede a seleccionar el modelo de regresión óptimo para este estudio a través del método Logit multinomial.

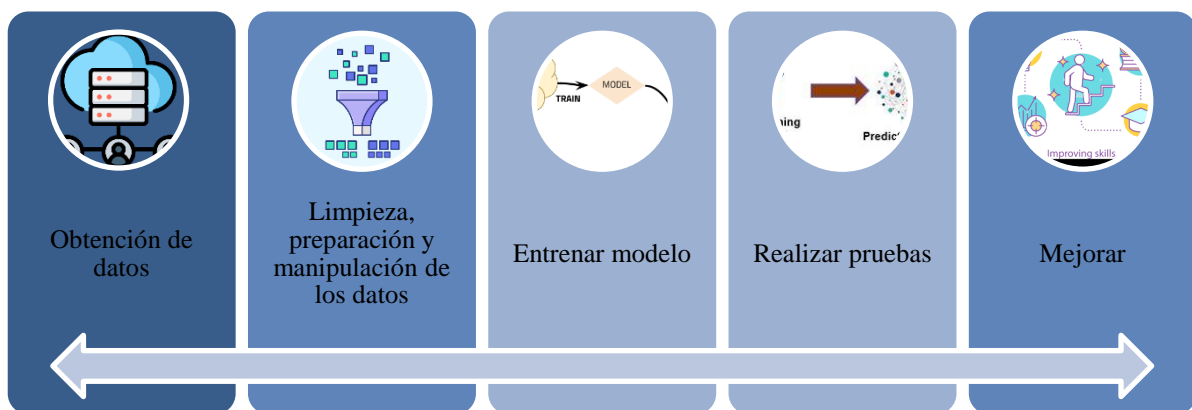
Posteriormente, se comprueba si las variables independientes son significativas mediante la prueba de Wald, tomando en cuenta que el valor de Wald menor será el p-valor y significará que las variables independientes son más significativas para el modelo de regresión.

Dentro de la metodología es importante incorporar las fases metodológicas, que particularmente para nuestra investigación se desarrolla en tres fases; en la primera se realiza la recopilación de los datos relacionados con las variables de investigación que permitan ejecutar regresiones, posteriormente en una segunda etapa se realiza el análisis

de los modelos de regresión estadística adecuados para el tipo de estudio, finalmente en una tercera etapa se lleva a cabo la aplicación del modelo y se procede con el análisis de los resultados. La Figura 18 muestra las fases y paso metodológicos para la elaboración del modelo estadístico.

Figura 18

Fases y paso metodológicos



Nota: Elaboración propia

- **Obtención de datos:** Conseguir datos de diferentes fuentes.
- **Limpieza, preparación y manipulación de los datos:** En este paso lo que se realiza principalmente es validar los datos, eliminar/corregir los datos nulos y estandarizar los datos previamente.
- **Entrenar modelo:** Elegir uno o varios modelos, configurar los hiperparámetros y realizar el entrenamiento del modelo con base a los datos.
- **Realizar pruebas:** Realizar pruebas del modelo entrenado para validar su rendimiento, usando el base de datos de entrenamiento.
- **Mejorar:** Con base a los resultados de las pruebas, seleccionar otros algoritmos y/o configurar otros hiperparámetros para comparar modelos y elegir el mejor.

3.7 Población y muestra

3.7.1 Población

Se considerará como población de estudio a los ciudadanos ecuatorianos que hayan salido del país en el año 2021 conforme al reporte realizado por el INEC Entradas y Salidas Internacionales (2021) se asevera que “el flujo migratorio general fue de 2844788 movimientos, de los cuales 1376221 corresponden a entradas internacionales y 1468567 a salidas internacionales de ecuatorianos y extranjeros”.

3.7.2 Tamaño de la muestra

La muestra que servirá de base para el desarrollo de este estudio será el número de salidas hechas en el período de interés por un puerto y un aeropuerto del país, correspondiente a 1468567 observaciones. Los mismos que se elegirán de acuerdo a la disponibilidad de datos. El estudio propone la aplicación de un modelo de regresión por lo cual se distinguen dos tipos de variables, la variable dependiente es la migración en tanto que las variables independientes corresponden a:

- Cantidad de personas que salen del país de acuerdo al género.
- Cantidad de personas que salen del país de acuerdo a la edad.
- Cantidad de personas que salen del país por puerto marítimo
- Cantidad de personas que salen del país por puerto terrestre
- Cantidad de personas que salen por aeropuerto.
- Cantidad de personas que salen del país de acuerdo al motivo de viaje, etc.

El MRE proporciona en su portal la base de datos histórica del número de salidas internacionales del año 2021, esta base será de gran importancia en la investigación, además, se utilizará bases de datos del INEC para completar las variables de interés propuestas en el estudio.

Capítulo 4

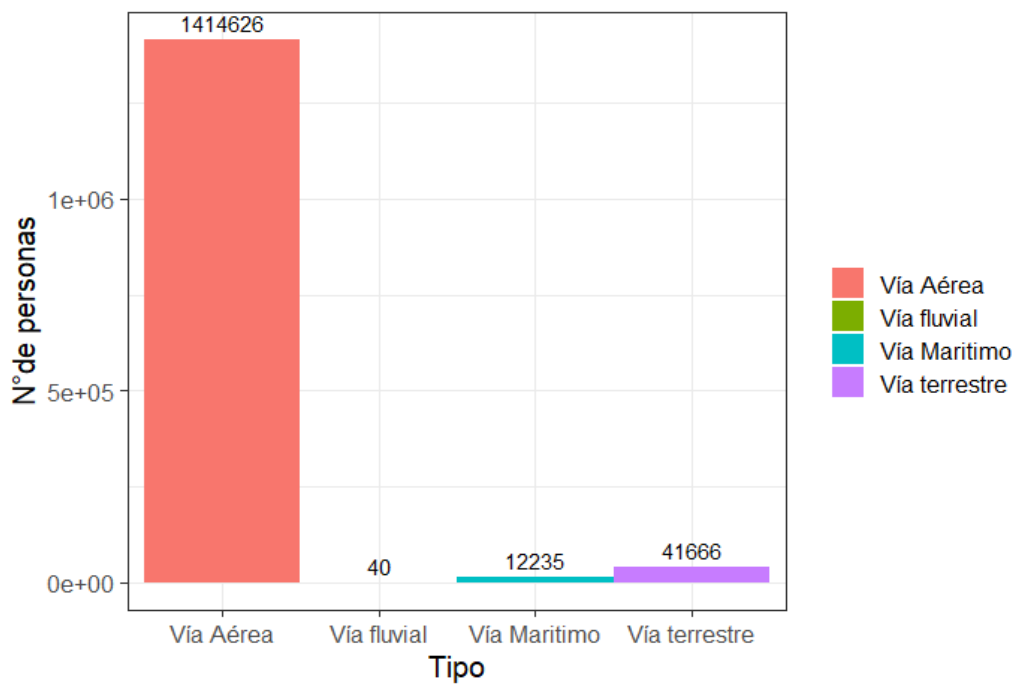
Análisis y Discusión de los Resultados

4.1 Análisis descriptivo de los resultados

De acuerdo a la Figura 19, el principal medio de transporte utilizado por los ecuatorianos y extranjeros en el año 2021 viajan mediante vía aérea con el 96,30% de todas las salidas internacionales registradas. La vía terrestre y marítima, también son utilizados en un 2,84% y 0,83% respectivamente, mientras la vía fluvial no es muy utilizado representado apenas en un 0,003%.

Figura 19

Salida de ecuatorianos y extranjeros según tipo de vía



Nota: Registro Estadístico de Entradas y Salidas Internacionales 2021

A continuación, se muestran en la Tabla 5, el resumen de salida de ecuatorianos y extranjeros organizados de acuerdo al tipo de vía, cantidades de personas, y porcentajes correspondientes.

Tabla 5

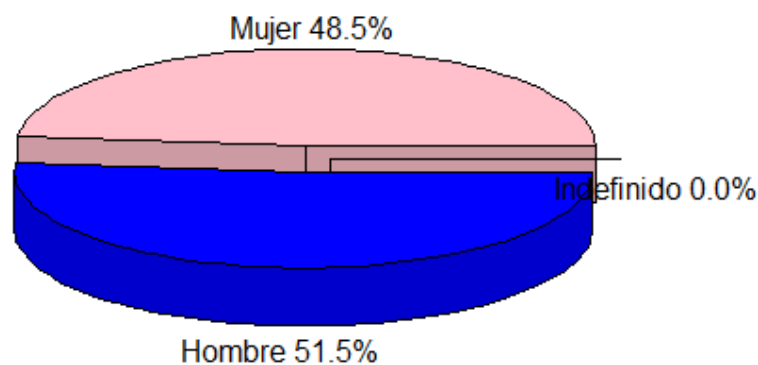
Salida de ecuatorianos y extranjeros, según tipo de vía, cantidad de personas y porcentajes

Tipo de vía	N° de personas	Porcentaje
Vía Aérea	1414626	96.30%
Marítimo	12235	0.83%
Fluvial	40	0.003%
Terrestre	41666	2.84%
Total	1468567	100%

En el año 2021, según la Figura 20 los hombres ecuatorianos y extranjeros reportan un mayor porcentaje de salidas del país en relación a las mujeres. En este contexto los hombres registran un total de 756127 salidas el cual representa el 51.5 % y las mujeres el 48.5%.

Figura 20

Salida de ecuatorianos y extranjeros según sexo



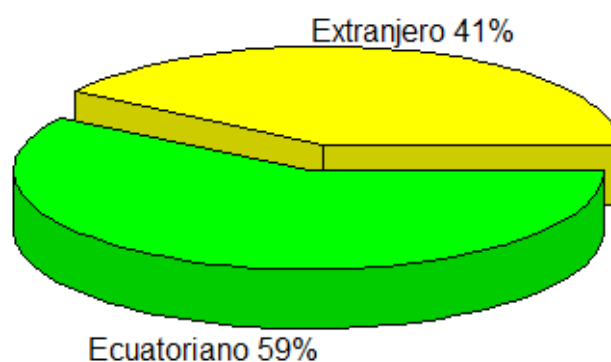
Nota: Registro Estadístico de Entradas y Salidas Internacionales 2021

La Tabla 6 muestra las relaciones entre la cantidad de personas ecuatorianos o extranjeros que han salido del país, organizados por sexo.

Tabla 6*Salida de ecuatorianos y extranjeros según sexo.*

Sexo	N° de personas	Porcentaje
Mujer	712410	48.5%
Hombre	756127	51.5%
Indefinido	30	0.0%
Total	1468567	100%

Del total de 1468567 de salidas del país, 867973 son ecuatorianos representando el 59% y el 41% corresponde a los extranjeros, tal como se evidencia en la Figura 21 y en la Tabla 7.

Figura 21*Salida de ecuatorianos y extranjeros según tipo de nacionalidad*

Nota: Registro Estadístico de Entradas y Salidas Internacionales 2021

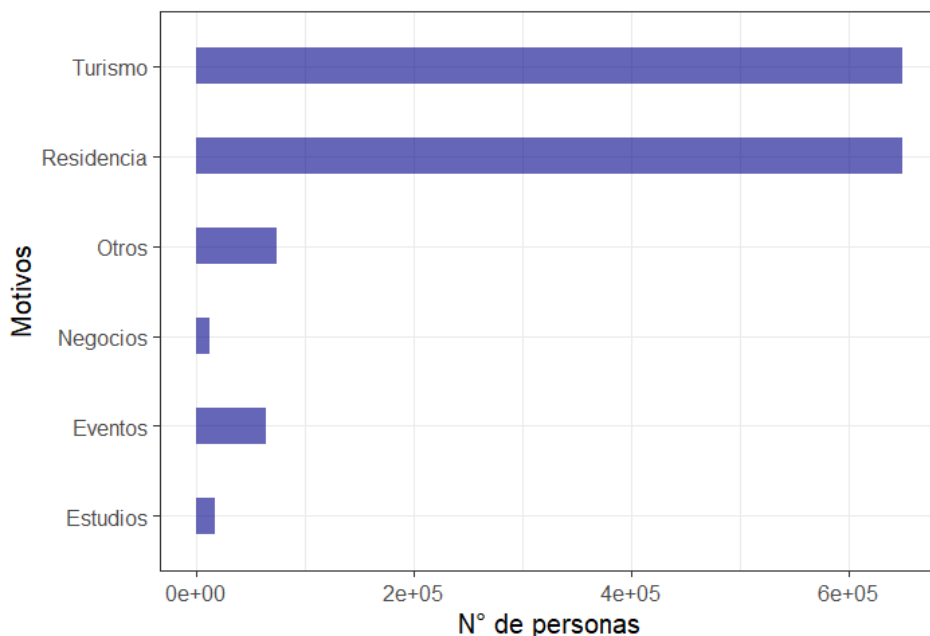
Tabla 7*Salida de ecuatorianos y extranjeros según nacionalidad.*

Nacionalidad	N° de personas	Porcentaje
Extranjero	600594	41%
Ecuatoriano	867973	59%
Total	1468567	100%

De acuerdo a la Figura 22 y su correspondiente Tabla 8, el 44.30% de salidas del país se debe por motivos de turismo, similar al de residencia el 44.20%, seguido de eventos 4,38%, estudios,1.19%, otros 5,06 y un porcentaje bajo de 0.09% por motivo de negocios.

Figura 22

Salida de ecuatorianos y extranjeros según motivo



Fuente: Registro Estadístico de Entradas y Salidas Internacionales 2021

Tabla 8

Salida de ecuatorianos y extranjeros según motivo

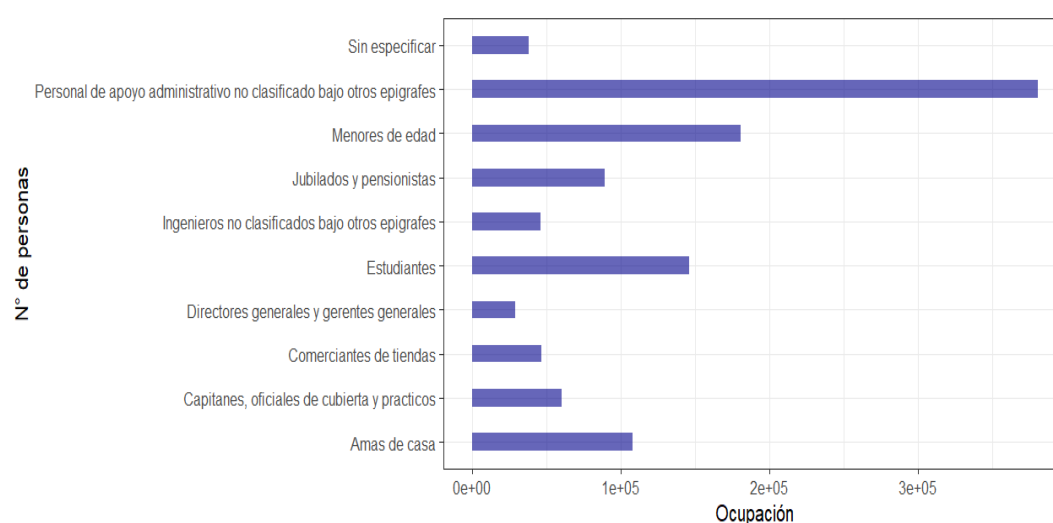
Motivo	N° de personas	Porcentaje
Turismo	650067	44,30%
Residencia	649595	44,20%
Otros	74275	5,06%
Eventos	64295	4,38%
Estudios	17533	1,19%
Negocios	12802	0,09%
Total	1468567	100%

Existen 150 ocupaciones registradas en el 2021, en la Figura 23 y la Tabla 9, se muestran que el 25,92% del total de salidas internacionales corresponden al personal de apoyo

administrativo no clasificado bajo otros epígrafes. Mientras las 9 ocupaciones restantes dentro del top 10 corresponde a Menores de edad, Estudiantes, Amas de casa, Jubilados y pensionistas, Capitanes, oficiales de cubierta y prácticos, Comerciantes de tiendas, ingenieros no clasificados bajo otros epígrafes Sin especificar y Directores y gerentes generales.

Figura 23

Salida de ecuatorianos y extranjeros según top 10 de ocupaciones registradas.



Nota: Registro Estadístico de Entradas y Salidas Internacionales 2021

Tabla 9

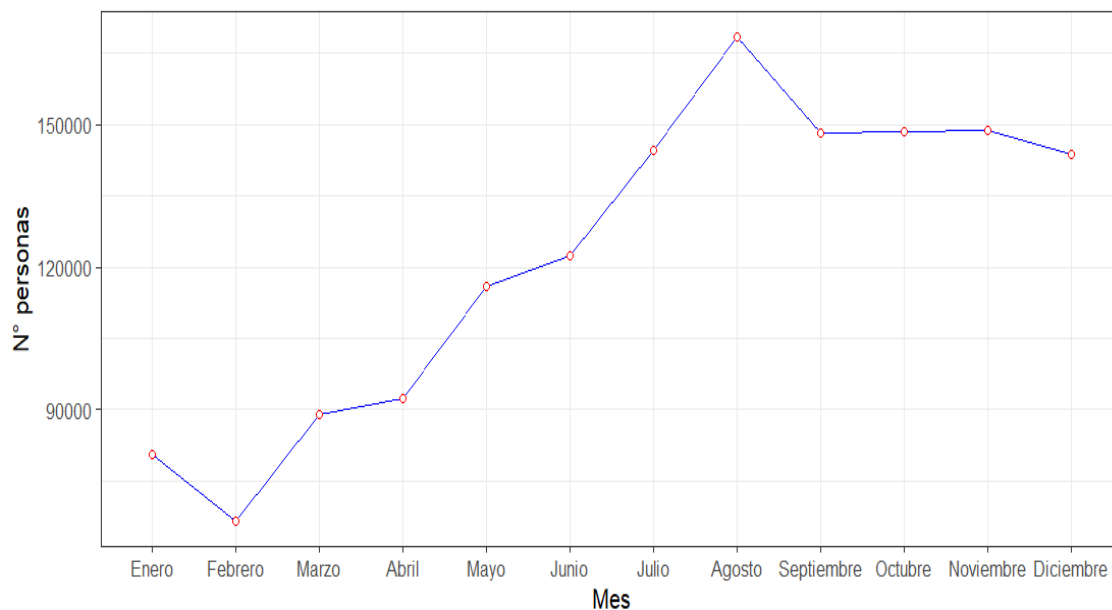
Salida de ecuatorianos y extranjeros según motivo top 10 de ocupaciones registradas.

Ocupación	Nº de personas	Porcentaje
Personal de apoyo administrativo no clasificado bajo otros epígrafes	380698	25.92%
Menores de edad	180466	12.29%
Estudiantes	146328	9.96%
Amas de casa	107935	7.35%
Jubilados y pensionistas	89242	6.08%
Capitanes, oficiales de cubierta y prácticos	60138	4.10%
Comerciantes de tiendas	46657	3.18%
Ingenieros no clasificados bajo otros epígrafes	46219	3.15%
Sin especificar	37848	2.58%
Directores generales y gerentes generales	28820	1.96%
Zapateros y afines	442	0.03%
Total	1468567	100%

La Figura 24 y la Tabla 10, muestran un aumento significativo de salidas entre los meses de julio a diciembre, mientras febrero presenta el menor porcentaje (4.52%) y en el mes de agosto se registra un mayor porcentaje (11.47%) de salidas internacionales.

Figura 24

Salida de ecuatorianos y extranjeros según el mes durante el 2021.



Nota: Registro Estadístico de Entradas y Salidas Internacionales 2021

Tabla 10

Salida de ecuatorianos y extranjeros según el mes durante el 2021

Mes	N° de personas	Porcentaje
Enero	80671	5.49%
Febrero	66444	4.52%
Marzo	88984	6.06%
Abril	92231	6.28%
Mayo	115882	7.89%
Junio	122428	8.34%
Julio	144479	9.84%
Agosto	168503	11.47%
Septiembre	148286	10.10%
Octubre	148353	10.10%
Noviembre	148719	10.13%
Diciembre	143587	9.78%
Total	1468567	100%

El código utilizado en R, para la elaboración de la estadística descriptiva, se lo puede encontrar en el Anexo A.

4.1.1 Identificación de las variables

Se inicia con la descripción de las variables:

Variable predictivas o independientes: X_i

1. Tipo de nacionalidad (Ecuatoriano – Extranjero) X_1
2. Sexo (Hombre – mujer- indefinido) X_2
3. Mes de movilidad (Enero – Diciembre) X_3
4. Motivo del viaje (Eventos – Negocios – Residencia – Turismo – Otros) X_4
4. Ocupación de la persona que migra. X_5

Variables de respuesta o dependiente: Y_j

- A. Terrestre Y_1
- B. Marítimo Y_2
- C. Aéreo Y_3

Entendiendo que la regresión logística multinomial estima las probabilidades de esta distribución para cada individuo i , teniendo en cuenta un conjunto de variables explicativas (Mc Fadden, 1974).

4.1.2 Formulación del modelo

Para proceder con la formulación del modelo resultante se ejecutaron algunas pruebas de acuerdo a la data disponible.

Debemos previamente definir los estadísticos de validación que se utilizan en análisis de agrupamientos.

APN (Average Pairwise Nearest Neighbor): Este estadístico calcula la distancia promedio entre los puntos de datos y sus vecinos más cercanos dentro del mismo grupo. Un valor

bajo de APN indica que los puntos dentro de un mismo grupo están cercanos entre sí, lo que sugiere una buena agrupación.

AD (Average Distance): El AD mide la distancia promedio entre todos los pares de puntos en el mismo grupo. Un valor bajo de AD indica que los puntos dentro del mismo grupo están cerca entre sí en términos de distancia euclidiana u otra métrica de distancia.

ADM (Average Distance to Centroid): Calcula la distancia promedio entre los puntos de datos y el centroide de su grupo asignado. Una menor ADM indica que los puntos están más cerca de su centroide, lo que es deseable en agrupamientos bien definidos.

FOM (Figure of Merit): El FOM combina información sobre la dispersión dentro de los grupos y la separación entre los grupos. Un valor alto de FOM indica que los grupos están bien separados y los puntos dentro de los grupos están cerca entre sí.

Connectivity: La conectividad mide cómo de bien están conectados los puntos dentro de cada grupo. Un valor alto de conectividad indica que los puntos están fuertemente conectados dentro de los grupos, lo que es característico de una buena agrupación.

Dunn Index: El índice de Dunn es la relación entre la mínima distancia entre los centros de dos grupos y la máxima distancia entre los puntos dentro de cada grupo. Un valor alto de Dunn indica grupos compactos y bien separados.

Silhouette Score: El coeficiente de silueta compara la distancia promedio entre un punto y los demás puntos en su mismo grupo con la distancia promedio entre el punto y los puntos en el grupo más cercano. Puede variar entre -1 y 1, donde un valor más alto sugiere una mejor separación entre los grupos (STHDA, 2023).

Se procedió aplicar Clustering, entonces a partir de la revisión de las técnicas de clúster, se aplica 3 métodos tales como: K-means clustering, K-medoids clustering PAM y Hierarchical K-means clustering, cabe mencionar que no se tomó en cuenta Fuzzy

clustering dado que cada individuo debe pertenecer a un solo grupo y Density based clustering (DBSCAN) debido a que solo se detecta un clúster mediante experimentación. Las técnicas de clúster permiten encontrar agrupaciones de personas que ha salido del país, con la finalidad de determinar para el caso de estudio, las posibles características similares. Se considera como variables de entrada tipo de nacionalidad, sexo, motivo de viaje y vía de transporte.

Previo a la creación de los modelos, se detecta vacíos (missing) y se estandariza las variables independientes, además para analizar si la data de salidas internaciones puede aplicarse técnicas de clúster, se utiliza el estadístico Hopkins cuyo valor 0.89 (cercano a 1), donde podemos concluir que el conjunto de datos es significativamente agrupable. Para establecer el numero óptimo de k clúster se aplica el método Elbow y Silhouette tal como se observa en la Figura 25 y 26, donde se recomienda agrupar en k=2 clúster.

Figura 25

Detección de k cluster mediante método de codo (Elbow).

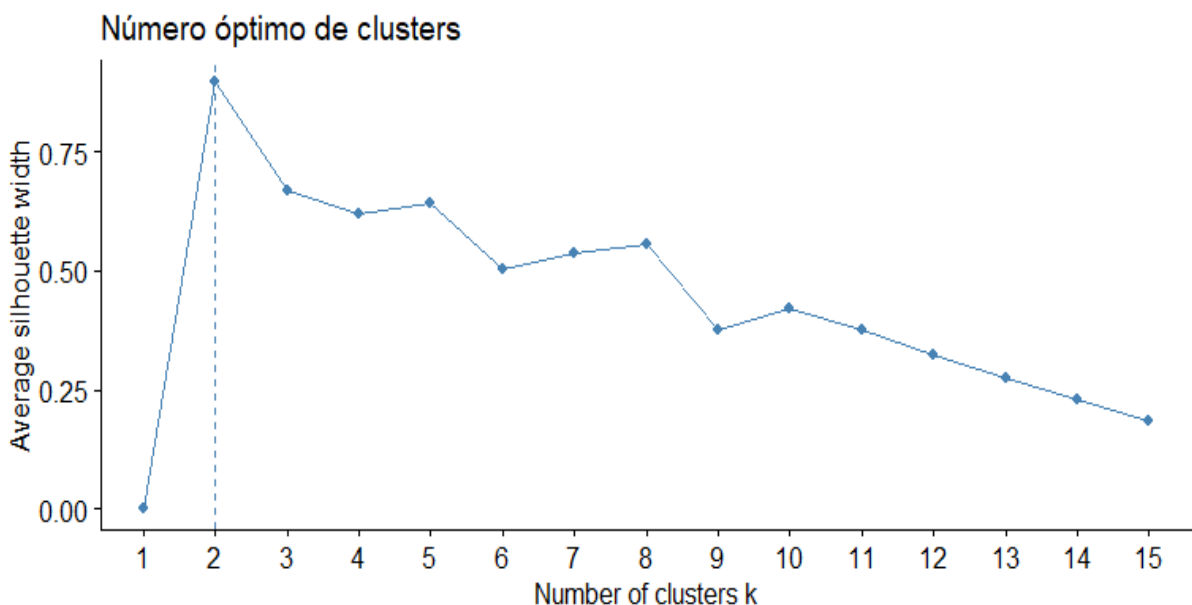
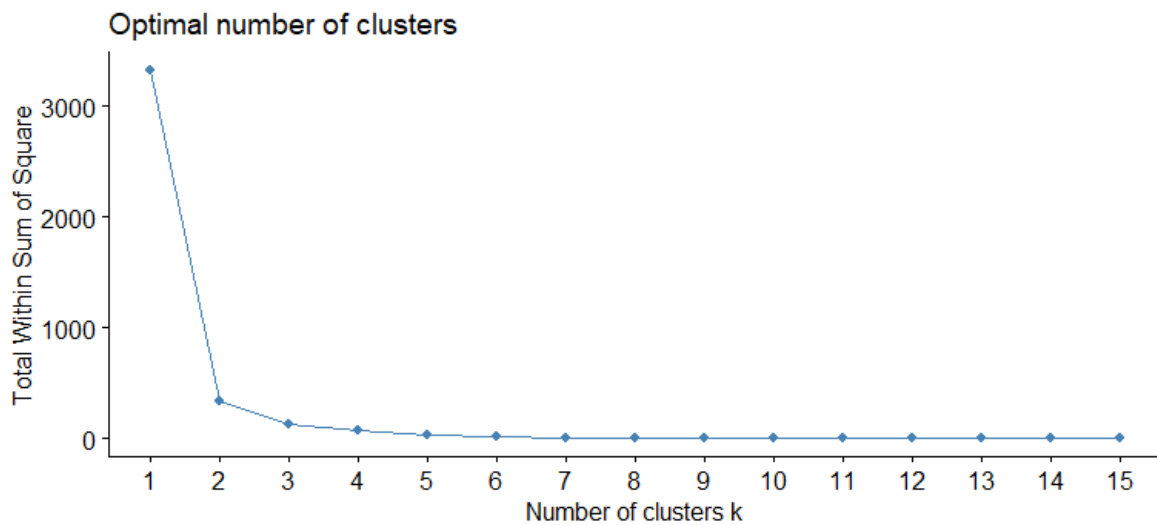


Figura 26

Detección de k cluster mediante método de Silhouette



Se procedió a comparar los resultados de los distintos métodos escogidos como se observa en la Figura 27, considerando los estadísticos de validación tales como: APN, AD, ADM, FOM, Connectivity, Dunn y Silhouette, donde el mejor método es el PAM con k= 2 clúster

Figura 27

Comparativa de métodos

validation Measures:		2	3	4	5	6	7	8
pam	APN	0.0000	0.0000	0.0016	0.0000	0.0014	0.0011	0.0025
	AD	3.2337	2.2596	1.6912	1.1196	0.7448	0.4906	0.3409
	ADM	0.0000	0.0000	0.0193	0.0000	0.0040	0.0058	0.0094
	FOM	0.2440	0.0701	0.0698	0.0410	0.0421	0.0414	0.0345
	Connectivity	3.8579	5.8579	8.7869	11.7159	14.6448	19.9317	22.8607
	Dunn	1.7540	2.0808	0.6050	0.9456	1.4321	0.4015	0.5145
	Silhouette	0.8952	0.8394	0.6209	0.6323	0.5952	0.4575	0.4808
kmeans	APN	0.0000	0.0000	0.0016	0.0000	0.0014	0.0029	0.0015
	AD	3.2337	2.2596	1.6912	1.1196	0.7448	0.4921	0.3769
	ADM	0.0000	0.0000	0.0193	0.0000	0.0040	0.0094	0.0063
	FOM	0.2440	0.0701	0.0698	0.0410	0.0421	0.0417	0.0307
	Connectivity	3.8579	5.8579	8.7869	11.7159	14.6448	19.9317	23.4317
	Dunn	1.7540	2.0808	0.6050	0.9456	1.4321	0.4015	0.6689
	Silhouette	0.8952	0.8394	0.6209	0.6323	0.5952	0.4575	0.4377
hierarchical	APN	0.0000	0.0000	0.0016	0.0000	0.0005	0.0016	0.0027
	AD	3.2337	2.2596	1.6912	1.1196	0.7448	0.5591	0.4123
	ADM	0.0000	0.0000	0.0193	0.0000	0.0033	0.0071	0.0097
	FOM	0.2440	0.0701	0.0698	0.0410	0.0422	0.0422	0.0353
	Connectivity	3.8579	5.8579	8.7869	11.7159	14.6448	17.5738	20.5028
	Dunn	1.7540	2.0808	0.6050	0.9456	1.4321	0.6689	0.7053
	Silhouette	0.8952	0.8394	0.6209	0.6323	0.5952	0.4653	0.4325

	Score	Method	Clusters
APN	0.0000	pam	2
AD	0.3409	pam	8
ADM	0.0000	pam	2
FOM	0.0307	kmeans	8
Connectivity	3.8579	pam	2
Dunn	2.0808	pam	3
Silhouette	0.8952	pam	2

Para finalizar en la Figura 28 se observa la agrupación de los individuos cuando se aplica el mejor algoritmo K-medoids clustering PAM.

El algoritmo K-medoids es una variante del algoritmo K-means que utiliza medoides en lugar de centroides para representar los grupos. PAM (Partitioning Around Medoids) es un enfoque específico dentro del algoritmo K-medoids que busca mejorar la elección de los medoides durante la optimización del agrupamiento. PAM es más robusto frente a valores atípicos y ruidosos en los datos, ya que los medoides son puntos reales en el conjunto de datos (Montoya, 2023).

Figura 28

Agrupación de los individuos del mejor algoritmo



Se procedió a realizar agrupación por sexo, en la que constaban dos grupos: uno de hombres y otro de mujeres, para luego determinar las diversas variables independientes como: edad, vía de transporte, motivo de viaje, nacionalidad, entre otras, pero lastimosamente al correr esta alternativa no se logró que funcione la regresión. En la Figura 29, se identifican los clústeres que se propusieron.

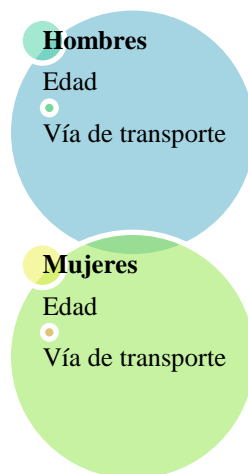
Al realizar el modelo con las siguientes variables `tip_naci` , `sex_migr` , `mes_movi` y `mot_viam`. En donde se aplicó la medición del nivel de confianza mediante El R-cuadrado de McFadden y el R-cuadrado de Nagelkerke, son medidas estadísticas que se utilizan para evaluar la bondad de ajuste de un modelo de regresión logística, permiten explicar la variación en de la variable dependiente en comparación a un modelo nulo o sin predictores. Se obtuvo una precisión (suma de elementos diagonales dividida por observaciones totales) de 96.28, y un R^2 de:

McFadden	Nagelkerke
0.2842012	0.3211000

(Estos resultados fueron bajos la diferencia del modelo que se logró posteriormente)

Figura 29

Propuesta de modelo con clúster para regresión logística

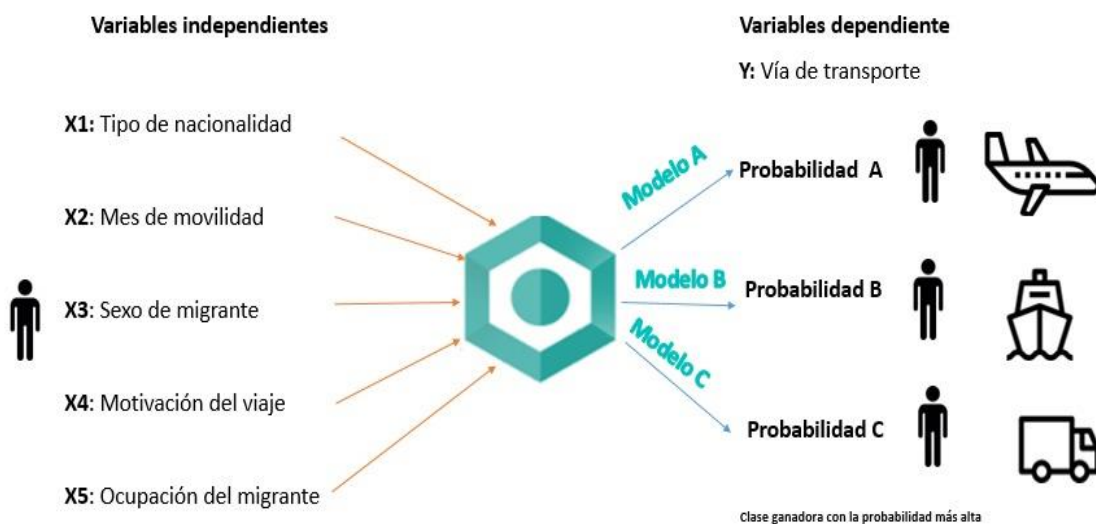


Otras alternativas en las que se aplicó las variables dependientes: Puerto terrestre, puerto marítimo, aeropuerto, tipo de movilidad, pero que no se pudieron obtener una gran representatividad en los resultados de la regresión logística multinomial.

Entonces, se probó aplicar la predicción de si una persona va a viajar por una vía de transporte terrestre, marítimo o aéreo, de acuerdo al tipo de nacionalidad, Mes de movilidad, sexo del migrante, vía de transporte y motivación del viaje, como se muestra en la representación de la Figura 30.

Figura 30

Propuesta de modelo aplicado de regresión logística



Nota: Elaboración propia

Finalmente, se adoptó la propuesta de la prueba con la se obtuvieron buenos resultados en la regresión logística, en la que se utilizaron las variables dependientes de vía de transporte (terrestre, marítimo y aéreo) y se tomaron 5 de las 24 variables independientes existentes.

Considerando una variable aleatoria dependiente Y categórica nominal politómica con Soporte(Y) = {1,2,3} y con probabilidades $p_1 = p(Y = 1)$, $p_2 = p(Y = 2)$ y $p_3 = p(Y = 3) = 1 - p_1 - p_2$.

Supongamos que se quiere analizar el efecto que ejercen dos variables explicativas continuas X_1, X_2 sobre las probabilidades p_1, p_2 que caracterizan a la variable Y . Se puede redefinir a la variable Y mediante un vector Y_1, Y_2 construido de la siguiente forma:

$$(Y_1, Y_2) = \begin{cases} (1,0) & \text{si } Y = 1 \\ (0,1) & \text{si } Y = 2 \\ (0,0) & \text{si } Y = 3 \end{cases}$$

Para realizar las estimaciones de los parámetros en el modelo logístico multinomial, se utiliza la función canónica de enlace de la distribución multinomial a la familia exponencial, la cual es llamada transformación Logit:

$$Y_i = \log \left(\frac{P_{ij}}{P_{ig}} \right)$$

Donde P_{ij} es la probabilidad del individuo i pertenezca a la categoría j , P_{ig} corresponde a la probabilidad del mismo individuo en la categoría g , la cual es definida como la categoría de referencia, de la variable con distribución multinomial Y . Además, se debe mencionar por ejemplo que al cociente $\frac{p_1}{p_3}$ se le denomina ‘odds’ de la categoría 1 respecto de la categoría 3, y que Odds ratio es la razón entre dos odds. Dando como resultado el modelo:

$$\ln (p(Y)) = \beta_{01} - \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \dots \beta_5 X_5$$

4.1.3 Modelo de regresión logística en R

Previo a la creación del modelo final, se realizaron varias pruebas de regresión que dependía de la data que se disponía, en la cual se ejecutaron cinco ideas verificadas.

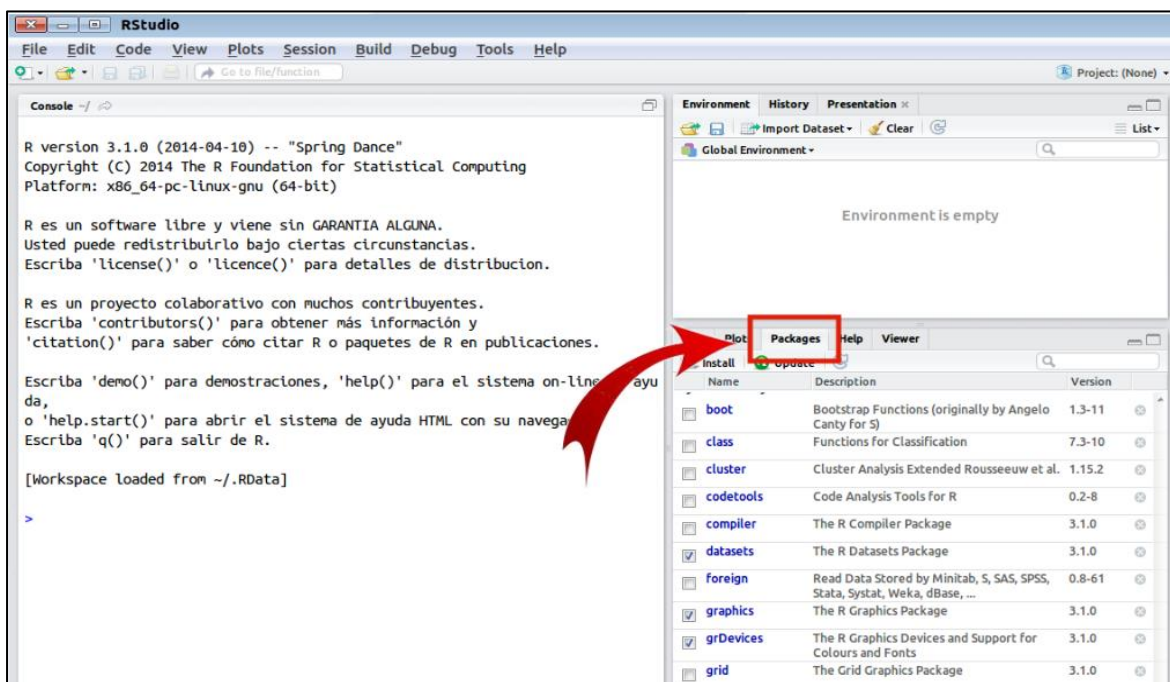
En una primera idea se probó realizar la regresión logística basándose en una agrupación de clúster considerando agrupación por sexo, luego tres por los tipos de migración por puertos terrestre, marítimo y aeropuertos, dentro de los cuales se distinguirían características de las personas que migran por estas salidas y la quinta por tipo de movilidad.

Se seleccionó la regresión logística identificada por medios de transporte como variables dependientes y las características de las personas como las variables independientes.

Entonces para poder introducir el modelo de regresión logística en R, es preciso iniciar por la carga de las librerías que se requieren utilizar. En la siguiente figura se puede apreciar donde se deben realizar la carga de librerías.

Figura 31

Carga de paquetes de librerías en RStudio



4.1.4 Cargar las librerías

A continuación, se listan las librerías que fueron requeridas para poder ejecutar el modelo de regresión logística multinomial.

- `library(dplyr)`.- el paquete `dplyr` proporciona una "gramática" (particularmente verbos) para la manipulación y operaciones con data frames.
- `library(tidyverse)`.- es un conjunto de librerías de R diseñadas para la "Ciencia de datos (Data Science)". Todas las librerías de este conjunto

comparten la misma filosofía de trabajo, la misma gramática y las mismas estructuras de datos.

- `library(tidyr)`.- es un paquete de R que permite ordenar datos “sucios” para obtener objetos de datos en R, a idea de los datos ordenados es que queden organizados del siguiente modo: Cada variable está en una columna, cada observación está en una fila.
- `library(magrittr)`.- tiene dos objetivos: disminuir el tiempo de desarrollo y mejorar la legibilidad y el mantenimiento del código.

4.1.5 Lectura de las bases de datos

Se trabajó con la base de datos migratorios del INEC del año 2021, la misma que fue cargada en RStudio mediante el código que se puede apreciar en el Anexo B.

4.1.6 Selección de variables y recodificación a level o factores

Las variables independientes fueron abreviadas para su mejor tratamiento en la propuesta del modelo.

- Tipo de movilidad = `tip_movi`
- Tipo de nacionalidad = `tip_naci`
- Año de movilidad = `anio_movi`
- Mes de movilidad = `mes_movi`
- Día de movilidad = `dia_movi`
- Sexo de migrante = `sex_migr`
- Nacionalidad de migrante = `nac_migr`
- Sub Continente de migrante = `subcont_nac`
- Continente de migrante = `cont_nac`
- Motivación de viaje = `mot_viam`

- País de procedencia = pais_prod
- Sub continente de procedencia = subcont_prod
- Continente de procedencia = cont_prod
- Lugar de procedencia = lug_prod
- País de residencia = pais_res
- Sub continente de residencia = subcont_res
- Continente de residencia = cont_res
- Jefatura migratoria = jef_migr
- Provincia de Jefatura migratoria = pro_jefm
- Cantón de Jefatura migratoria = can_jefm
- Clasificación migratoria = cla_migr
- Ocupación migratoria = ocu_migr
- Edad = edad

Para el tratamiento del modelo es preciso dejar definidas claramente las variables, por lo que se indican en la siguiente Tabla 11:

Tabla 11

Determinación de variables para la base de datos

Variable dependiente	Subcategorías	Variables independientes
Vías de transporte	Vía de transporte_puerto terrestre	tip_movi
	Vía de transporte_puerto marítimo	tip_naci
	Vía de transporte_aeropuerto	anio_movi
		mes_movi
		dia_movi
		sex_migr
		nac_migr
		subcont_nac
		cont_nac
		mot_viam
		pais_prod
		subcont_prod
		cont_prod
		lug_prod
		pais_res

	subcont_res
	cont_res
	jef_migr
	pro_jefm
	can_jefm
	cla_migr
	ocu_migr
	edad

4.1.7 Determinación de variables para la regresión

Debido a las pruebas realizadas con las combinaciones de algunas variables, se logró determinar aquellas que sirvieron para poder establecer un modelo óptimo de regresión logística multinomial, siendo las siguientes cinco variables independientes:

1. Tipo de nacionalidad.
2. Mes de movilidad.
3. Sexo de migrante.
4. Ocupación del migrante.
5. Motivación del viaje.

4.1.8 Selección de tipos de variables

Cada una de las variables seleccionadas tienen sus correspondientes tipos por ejemplo en la nacionalidad se distinguen si es ecuatoriano o si es extranjero. Para esto utilizamos la función Unique() que “en el lenguaje de programación R se usa para devolver un vector, un marco de datos o una array sin elementos / filas duplicados” (Barcelona Geeks, 2022, p. 1).

4.1.9 Transformación de la variable dependiente a numérico

Es necesario realizar la transformación de las variables dependientes convirtiéndoles en numéricos, para esto se hace uso de la función ifelse, la misma que **toma tres argumentos:**

- El primer argumento es un vector lógico de condiciones.
- El segundo es un vector que contiene los valores que serán devueltos cuando el primer vector es TRUE.

- El tercero contiene los valores que serán devueltos cuando el primer vector es FALSE. (Sánchez R. , 2023, p. 1).

4.1.10 Número de observaciones o personas en el estudio

Otro de los aspectos para la aplicación del modelo de regresión en R, es el relacionado con la data frames que son aquellos que se usan “para almacenar datos en forma de hoja de datos. Cada fila de la hoja de datos corresponde a una observación o valor de una instancia, mientras que cada columna corresponde a un vector que contiene los datos de una variable” (Sánchez R. , 2023, p. 1).

La función nrow() devuelve el número de líneas y la función str() permite visualizar la estructura interna de un objeto, esto quiere decir la cantidad de observaciones y variables.

4.2 Discusión de los resultados

4.2.1 Número de personas por vía de transporte

Utilizamos la función table para poder construir una tabla de frecuencias absolutas el nombre de la variable (via_tran) va precedido por el nombre del data.frame data_reg_2 y del símbolo \$. “Ello permite utilizar simultáneamente varios data.frames que contengan información de las mismas variables” (Estadística DMA, 2023, p.1). El número de personas se visualiza mediante el código que consta en el Anexo C y el resultado se aprecia en la Tabla 12.

Tabla 12

Número de personas por vía de transporte.

Transporte fluvial	Transporte terrestre	Transporte marítimo	Transporte aéreo
(4)	(1)	(2)	(3)
40	41666	12235	1414626

Como se puede analizar en el resultado, se tiene una matriz de confusión que es “una herramienta muy útil para valorar que tan bueno es un modelo de clasificación basado en aprendizaje automático” (Recuero, 2021, p.1). En la cual los valores de la diagonal que va

desde la parte superior izquierda a la inferior derecha, son los resultados validos para el modelo. Como se puede apreciar en la Figura 32:

Figura 32

Matriz de confusión

Matriz de confusión		Estimado por el modelo			
		Negativo (N)	Positivo (P)		
Real	Negativo	a: (TN)	b: (FP)	Precisión ("precision") Porcentaje predicciones positivas correctas:	$d/(b+d)$
	Positivo	c: (FN)	d: (TP)		
		Sensibilidad, exhaustividad ("Recall") Porcentaje casos positivos detectados	Especificidad ("Specificity") Porcentaje casos negativos detectados	Exactitud ("accuracy") Porcentaje de predicciones correctas (No sirve en datasets poco equilibrados)	
		$d/(d+c)$	$a/(a+b)$	$(a+d)/(a+b+c+d)$	

Nota: Tomado de Recuero (2021)

4.2.2 Correlación

Para poder realizar la correlación iniciamos con la inclusión de la librería (psych) debido a que contiene las funciones describe () y describe.by () que no sólo calculan la media, sino que añaden otros estadísticos de interés. Luego se incorpora el concepto de fijar una semilla en R, de acuerdo con R-coder (2023):

Cuando se genera “números aleatorios” en R, en realidad se está generando números pseudoaleatorios. Estos números se generan con un algoritmo que requiere una semilla para inicializarse. Ser pseudoaleatorio en lugar de puramente aleatorio significa que, si se conoce la semilla y el generador, se puede predecir (y reproducir) la secuencia.

Además, esto se realiza con la función set.seed, cuyo objetivo es permitirte fijar una semilla inicial y un generador.

Como indica Villalba (2017) al preparar los datos para un modelo de aprendizaje es una buena práctica realizar lo siguiente:

Dividir el conjunto de datos de origen en dos subconjuntos, uno de ellos nos servirá para el entrenamiento (train) del modelo y el otro para la comprobación -a posteriori- (test). Con este sistema de división de la muestra inicial de datos evitamos el overfitting.

Esta parte se aplica con las funciones train y test, que para este caso se han asignado con un 70% para entrenamiento y con un 30% para prueba. El código utilizado en R para este propósito se puede revisar en el Anexo D.

El desarrollo de la correlación prosigue mediante la identificación de los resultados obtenidos respecto de las variables de vía de transporte, mostrados en la Tabla 13, utilizando el código indicado en el Anexo E. Los valores asignados de variables dependientes tienen su correspondencia como se indica: (1) Terrestre, (2) marítimo, (3) aéreo y (4) fluvial.

Tabla 13

Resultados del entrenamiento.

<i>[1] 1 2 3 4</i>			
<i>Levels: 4 1 2 3</i>			
Transporte fluvial (4)	Transporte terrestre (1)	Transporte marítimo (2)	Transporte aéreo (3)
31	29132	8575	990258

Como se puede apreciar en el resultado de la tabla de entrenamiento, se han identificado una cantidad para cada tipo de variable dependiente, Terrestre (1), Marítimo (2), Aéreo (3), Fluvial (4), que suman en total la asignación de respuesta obtenidas de la correlación con las 1027996 observaciones.

Para poder continuar es necesario cargar las librerías:

VGAMS, que es la librería necesaria para la estimación del modelo. (Martínez & Delgado, 2020).

SPLINES, las funciones spline ofrecen una forma sencilla, flexible y útil para realizar un ajuste por segmentos. (Moscote, 2022). El código utilizado para la carga de estas librerías se puede apreciar en el Anexo F.

4.2.3 Modelo de regresión logística

Para la generación del modelo fue necesario hacer uso de la función multinom en la librería nnet (Venables & Ripley 2002). Es el paquete de mayor uso y se basa en redes neuronales (LLinás, 2022). El código aplicado en el modelo de regresión se puede revisar en el Anexo G.

Posteriormente a esta ejecución se realiza el llamado a la función multinom y se procede a realizar la intercepción de los coeficientes de acuerdo con las variables independientes señaladas. El código para llamar a la función del modelo de regresión se indica en el Anexo H.

4.2.4 Predecir los valores para el conjunto de datos del test o prueba

En la regresión logística múltiple el número de predictores es mayor, en comparación con la regresión simple, pudiendo ser tanto continuos como categóricos. Por lo tanto, se utiliza la función predict para poder encontrar la predicción de valores que arroja el programa, como se muestra en el siguiente código y el resultado obtenido al ejecutar. El código de predicción con los datos de prueba se indica en el Anexo I.

Tabla 14

Resultado función Predict

	1 Terrestre	2 Marítimo	3 Aéreo
1 Terrestre	6299 Vedadero positivo	5	6230
2 Marítimo	848	1464 Vedadero positivo	1348

3 Aéreo	1141	379	422848 Vedadero positivo
----------------	------	-----	------------------------------------

El modelo clasificó a las personas de la siguiente manera

- Terrestre (1): Correctamente 66299 e incorrecto 6235
- Marítimo (2): Correctamente 1464 e incorrecto 2196
- Aéreo (3): Correctamente 422848 e incorrecto 1520

4.2.5 Precisión de cálculo (Accuracy):

Para encontrar la precisión del modelo, se aplica la suma de elementos diagonales dividida por observaciones totales. El código para la precisión de cálculo se identifica en el Anexo J y el resultado obtenido fue del 97.74 % de precisión en el conjunto de datos de entrenamiento, es decir este porcentaje de la data fue clasificada correctamente respecto a los datos de entrenamiento.

4.2.6 Evaluar la confianza del modelo

Para realizar la evaluación de la calidad del modelo, se usaron indicadores relacionados con los coeficientes denominados Pseudo R^2 , siendo estos los que permiten la comparación de modelos de regresión logística multinomial con diversas variables predictoras.

Para Niño (2022) el cociente pseudo- R^2 de Mc-Fadden “Los modelos de regresión logística pueden ser ajustados mediante el método de máxima verosimilitud, es decir, las estimaciones de los parámetros son aquellos valores que maximizan la probabilidad de los datos que se han observado” (p. 27).

Por otra parte el Cociente pseudo R^2 de Nagelkerke “se puede considerar como un ajuste del cociente de Cox-Snell, para abordar el problema en el que el límite superior de la R al cuadrado de Cox-Snell no es 1” (Niño R., 2022, p.28).

Es importante evaluar la calidad del ajuste, la que fue determinada mediante los cocientes pseudo R-cuadrado de Nagelkerke y McFadden (Niño R. , 2022, p. 52)

La aplicación de la confianza del modelo se realizó mediante los descritos cocientes de McFadden y Nagelkerke, el código se adjunta en el Anexo K y los resultados obtenidos fueron:

<i>McFadden</i>	<i>Nagelkerke</i>
0.6120854	0.6535123

El resultado obtenido al correr el programa evidencia que existe un 61% de confianza del modelo de acuerdo con McFadden y un 65% de confianza con Nagelkerke. Es decir que implicaría que las variables independientes empleadas (Tipo de nacionalidad, Sexo, Mes de movilidad, Motivo del viaje y Ocupación de los migrantes.) en el modelo de regresión explican el 65 % de la variabilidad de la variable dependiente. Por lo tanto, el modelo ha cumplido con las condiciones de precisión y confianza respectivamente.

Capítulo 5

Marco Propositivo

5.1 Propuesta de predicción, control y optimización de flujo migratorio en el Ecuador

En el Ecuador existen un total de 23 puntos migratorios distribuidos en las siguientes 12 provincias:

Figura 33

Provincias de puntos migratorios en el Ecuador



La cantidad de acuerdo al tipo de transporte para el flujo migratorio en el Ecuador, de acuerdo con los datos registrados por el INEC Entradas y Salidas Internacionales (2021) se asevera que “el flujo migratorio general fue de 2844.88 movimientos, de los cuales 1376221 corresponden a entradas internacionales y 1468567 a salidas internacionales de ecuatorianos y extranjeros” (p.1) En la presente investigación se ha trabajado con la base

de datos que cuenta con las observaciones correspondientes a las salidas internacionales de ecuatorianos y extranjeros.

El resumen del flujo migratorio del Ecuador desde 1997 hasta 2021 se lo puede apreciar en la Tabla 15.

Tabla 15

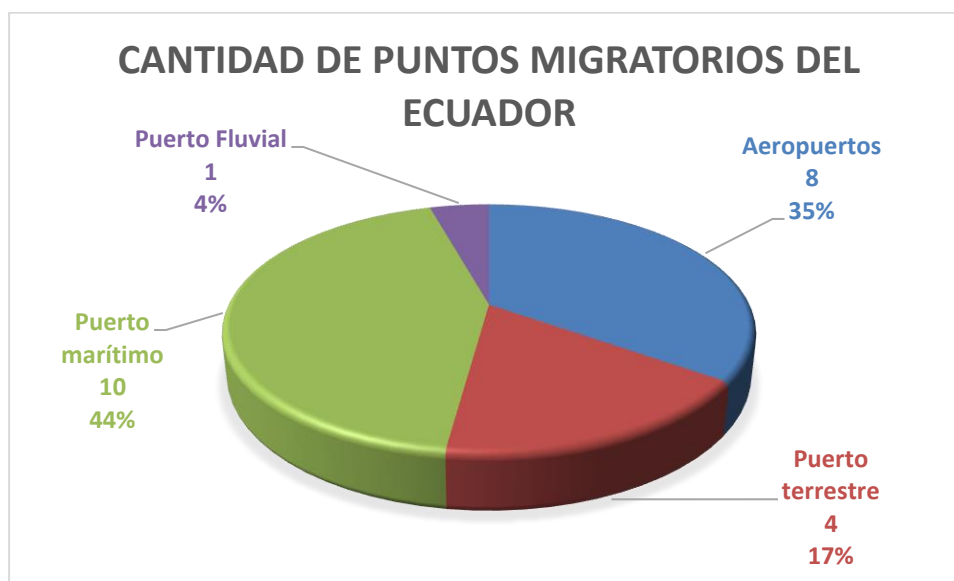
Flujo migratorio en Ecuador 1997 - 2021

Años	Entradas	Salidas	Flujo Migratorio General**	Saldo Migratorio***
1997	819.184	722.733	1.541.917	96.451
1998	705.269	624.358	1.329.627	80.911
1999	812.217	794.301	1.606.518	17.916
2000	971.142	964.900	1.936.042	6.242
2001	1.064.298	1.026.848	2.091.146	37.450
2002	1.144.358	1.114.157	2.258.515	30.201
2003	1.246.747	1.132.907	2.379.654	113.840
2004	1.347.839	1.245.697	2.593.536	102.142
2005	1.456.926	1.367.156	2.824.082	89.770
2006	1.514.822	1.512.844	3.027.666	1.978
2007	1.695.379	1.698.881	3.394.260	-3.502
2008	1.757.235	1.767.097	3.524.332	-9.862
2009	1.788.791	1.742.611	3.531.402	46.180
2010	1.940.506	1.904.307	3.844.813	36.199
2011	2.168.580	2.108.567	4.277.147	60.013
2012	2.297.211	2.240.008	4.537.219	57.203
2013	2.507.173	2.447.510	4.954.683	59.663
2014	2.826.666	2.759.821	5.586.487	66.845
2015	2.919.356	2.862.444	5.781.800	56.912
2016	2.911.927	2.929.849	5.841.776	-17.922
2017	3.114.763	3.065.412	6.180.175	49.351
2018	3.903.315	3.749.943	7.653.258	153.372
2019	3.557.506	3.485.700	7.043.206	71.806
2020	978.493	1.042.983	2.021.476	-64.490
2021	1.376.221	1.468.567	2.844.788	-92.346

Las alternativas de salidas internacionales del país pueden optarse por vías terrestre, marítima o aérea. Las mismas que se han resumido en la siguiente figura conforme a los 23 puntos existentes:

Figura 34

Cantidad de puntos migratorios del Ecuador por tipo de transporte



En la distribución se puede apreciar que el 44% de los puntos migratorios son de tipo puerto marítimo, un 35% son aeropuertos, el 17 % es por puerto terrestre y apenas el 1% corresponde a puerto fluvial. Es importante considerar que, puerto marítimo hace referencia a la salida de migrantes por mares, mientras que puerto fluvial, se refiere a la salida de migrantes por ríos, lagos y otros cuerpos de agua.

Para poder realizar una predicción, control y optimización del flujo migratorio en el Ecuador se propone la identificación de las proyecciones de decisión respecto a tres tipos de puntos de salida (terrestre, marítimo y aéreo) variables dependientes, correlacionadas con las variables independientes identificadas en la base de datos de salidas internacionales (nacionalidad, sexo, mes de movilidad, motivación del viaje, ocupación de la persona que migra)

Es importante señalar que se obtuvo el modelo aplicando las estimaciones de máxima similitud, en las que se realizaron transformación monótona, mediante logaritmos naturales.

Por tanto, con base en el establecimiento de las variables, se identificó el modelo matemático que se aplicaría y es el que se muestra a continuación.

$$\ln \left(p \left(\frac{Vía_{tran} = Puerto\ terrestre}{Vía_{tran} = Puerto\ fluvial} \right) \right) = \beta_{01} - \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \dots \beta_5 X_5$$

Este modelo se utiliza para predecir la probabilidad de que una persona viaje por vía terrestre Y_1 , dadas las cinco variables independientes X_1, X_2, \dots, X_5 (nacionalidad, sexo, mes de movilidad, motivación del viaje, ocupación de la persona que migra).

La parte izquierda de la ecuación se le conoce como logit, donde los argumentos de probabilidad están determinados en la razón de la probabilidad de vía de transporte terrestre, sobre la variable de referencia utilizada que es la vía de transporte fluvial, debido a que es la de menor representatividad de acuerdo con la base de datos. A esta se le conoce como los Odds ratio, que son precisamente la división entre dos odds, siendo esa la razón entre la probabilidad de que un evento ocurra y la probabilidad de que no ocurra.

Los parámetros a encontrar son los coeficientes $\beta_{01} - \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \dots \beta_5 X_5$ con sus respectivas alternativas de acuerdo a la variable independiente.

La identificación de los coeficientes que se han obtenido mediante este modelo logístico multinomial, nos permite una proyección de los tres tipos de salidas de migración, siendo por vía terrestre, marítima o aérea, las que se relacionan con las variables independientes analizadas.

Entonces, mediante la aplicación de los modelos matemáticos y de la cantidad de datos que posea por años, la proyección de salidas de personas de las diversas categorías que realicen una migración desde el Ecuador hacía otros parís por las tres alternativas estudiadas (terrestre, marítima, aérea). Entonces los modelos resultantes de acuerdo con cada una de las variables dependientes Y_1, Y_2 , y Y_3 son:

$$\ln \left(p \left(\frac{Vía_{tran} = Puerto\ terrestre}{Vía_{tran} = Puerto\ fluvial} \right) \right)$$

$$= \beta_{01} - \beta_{11}X_{11} + \beta_{21}X_{21} + \beta_{22}X_{22} + \dots + \beta_{31}X_{31} + \dots + \beta_{41}X_{41} + \dots + \beta_{51}X_{51}$$

Tabla 16

Variables y coeficientes del modelo A

Modelo A		
Vía transporte por puerto Terrestre (Y₁)		
Variables	Categorías	Coefficiente β
		β_{01}: 21.304
Tipo de nacionalidad X₁		
X ₁₁ :	tip_naciExtranjero	β_{11}: -5.278
Mes de movilidad X₂		
X ₂₁ :	mes_moviAgosto	β_{21}: 5.640
X ₂₂ :	mes_moviDiciembre	β_{22}: 5.790
X ₂₃ :	mes_moviEnero	β_{23}: -19.785
X ₂₄ :	mes_moviFebrero	β_{24}: 2.885
X ₂₅ :	mes_moviJulio	β_{25}: 5.207
X ₂₆ :	mes_moviJunio	β_{26}: 4.634
X ₂₇ :	mes_moviMarzo	β_{27}: 3.793
X ₂₈ :	mes_moviMayo	β_{28}: 4.604
X ₂₉ :	mes_moviNoviembre	β_{29}: 5.534
X ₂₁₀ :	mes_moviOctubre	β_{210}: 5.593
X ₂₁₁ :	mes_moviSeptiembre	β_{211}: 5.447
Sexo de migrante X₃		
X ₃₁ :	sex_migrIndefinido	β_{31}: -0.245
X ₃₂ :	sex_migrMujer	β_{32}: 18.082
Motivación del viaje X₄		
X ₄₁ :	mot_viamEventos	β_{41}: 7.874
X ₄₂ :	mot_viamNegocios	β_{42}: 10.496
X ₄₃ :	mot_viamOtros	β_{43}: 19.120
X ₄₄ :	mot_viamResidencia	β_{44}: 4.530
X ₄₅ :	mot_viamTurismo	β_{45}: 27.023
Ocupación del migrante X₅		
X ₅₁ :	ocu_migrActores	β_{51}: -0.107
X ₅₂ :	ocu_migrAgentes de servicios comerciales no clasificados bajo otros epígrafes	β_{52}: 0.826
X ₅₃ :	ocu_migrAgricultores y trabajadores calificados de huertas, invernaderos, viveros y jardines	β_{53}: 1.083
X ₅₄ :	ocu_migrAgricultores y trabajadores calificados de jardines y de cultivos para el mercado	β_{54}: 3.269
X ₅₅ :	ocu_migrAgronomos y afines	β_{55}: 0.903
X ₅₆ :	ocu_migrAlbaniles	β_{56}: -0.090

$$\ln \left(p \left(\frac{Vía_{tran} = Puerto\ marítimo}{Vía_{tran} = Puerto\ fluvial} \right) \right)$$

$$= \beta_{01} - \beta_{11}X_{11} + \beta_{21}X_{21} + \beta_{22}X_{22} + \dots + \beta_{31}X_{31} + \dots + \beta_{41}X_{41} + \dots + \beta_{51}X_{51} + \dots$$

Tabla 17

Variables y coeficientes del modelo B

Modelo B		
Vía transporte por puerto Marítimo (Y₂)		
Variables	Categorías	Coeficiente β
		β_{01} : 28.443
Tipo de nacionalidad X₁		
X ₁₁ :	tip_naciExtranjero	β_{11} : -4.539
Mes de movilidad X₂		
X ₂₁ :	mes_moviAgosto	β_{21} : 4.824
X ₂₂ :	mes_moviDiciembre	β_{22} : 4.031
X ₂₃ :	mes_moviEnero	β_{23} : -19.794
X ₂₄ :	mes_moviFebrero	β_{24} : 2.449
X ₂₅ :	mes_moviJulio	β_{25} : 4.273
X ₂₆ :	mes_moviJunio	β_{26} : 4.122
X ₂₇ :	mes_moviMarzo	β_{27} : 3.779
X ₂₈ :	mes_moviMayo	β_{28} : 4.487
X ₂₉ :	mes_moviNoviembre	β_{29} : 4.012
X ₂₁₀ :	mes_moviOctubre	β_{210} : 5.473
X ₂₁₁ :	mes_moviSeptiembre	β_{211} : 4.911
Sexo de migrante X₃		
X ₃₁ :	sex_migrIndefinido	β_{31} : -1.379
X ₃₂ :	sex_migrMujer	β_{32} : 17.848
Motivación del viaje X₄		
X ₄₁ :	mot_viamEventos	β_{41} : 2.342
X ₄₂ :	mot_viamNegocios	β_{42} : -1.912
X ₄₃ :	mot_viamOtros	β_{43} : 11.243
X ₄₄ :	mot_viamResidencia	β_{44} : -4.862
X ₄₅ :	mot_viamTurismo	β_{45} : 19.661
Ocupación del migrante X₅		
X ₅₁ :	ocu_migrActores	β_{51} : 1.627
X ₅₂ :	ocu_migrAgentes de servicios comerciales no clasificados bajo otros epígrafes	β_{52} : -1.545
X ₅₃ :	ocu_migrAgricultores y trabajadores calificados de huertas, invernaderos, viveros y jardines	β_{53} : -2.163
X ₅₄ :	ocu_migrAgricultores y trabajadores calificados de jardines y de cultivos para el mercado	β_{54} : -4.879
X ₅₅ :	ocu_migrAgronomos y afines	β_{55} : -0.726
X ₅₆ :	ocu_migrAlbaniles	β_{56} : 0.905

$$\ln \left(p \left(\frac{Vía_{tran} = Vía \text{ transporte por aeropuerto}}{Vía_{tran} = Puerto \text{ fluvial}} \right) \right)$$

$$= \beta_{01} - \beta_{11}X_{11} + \beta_{21}X_{21} + \beta_{22}X_{22} + \dots + \beta_{31}X_{31} + \dots + \beta_{41}X_{41} + \dots + \beta_{51}X_{51} + \dots$$

Tabla 18

Variables y coeficientes del modelo C

Modelo C		
Vía transporte por aeropuerto (Y₃)		
Variables	Categorías	Coefficiente β
		β_{01} : 41.874
Tipo de nacionalidad X₁		
X₁₁:	tip_naciExtranjero	β_{11} : -4.982
Mes de movilidad X₂		
X₂₁:	mes_moviAgosto	β_{21} : 6.062
X₂₂:	mes_moviDiciembre	β_{22} : 4.029
X₂₃:	mes_moviEnero	β_{23} : -19.833
X₂₄:	mes_moviFebrero	β_{24} : 2.854
X₂₅:	mes_moviJulio	β_{25} : 5.359
X₂₆:	mes_moviJunio	β_{26} : 4.656
X₂₇:	mes_moviMarzo	β_{27} : 3.788
X₂₈:	mes_moviMayo	β_{28} : 4.809
X₂₉:	mes_moviNoviembre	β_{29} : 5.362
X₂₁₀:	mes_moviOctubre	β_{210} : 5.395
X₂₁₁:	mes_moviSeptiembre	β_{211} : 5.664
Sexo de migrante X₃		
X₃₁:	sex_migrIndefinido	β_{31} : 1.628
X₃₂:	sex_migrMujer	β_{32} : 19.775
Motivación del viaje X₄		
X₄₁:	mot_viamEventos	β_{41} : -3.878
X₄₂:	mot_viamNegocios	β_{42} : -6.338
X₄₃:	mot_viamOtros	β_{43} : 2.364
X₄₄:	mot_viamResidencia	β_{44} : -10.331
X₄₅:	mot_viamTurismo	β_{45} : 12.487
Ocupación del migrante X₅		
X₅₁:	ocu_migrActores	β_{51} : -1.287
X₅₂:	ocu_migrAgentes de servicios comerciales no clasificados bajo otros epígrafes	β_{52} : 0.831
X₅₃:	ocu_migrAgricultores y trabajadores calificados de huertas, invernaderos, viveros y jardines	β_{53} : 1.306
X₅₄:	ocu_migrAgricultores y trabajadores calificados de jardines y de cultivos para el mercado	β_{54} : 2.695
X₅₅:	ocu_migrAgronomos y afines	β_{55} : 0.640
X₅₆:	ocu_migrAlbaniles	β_{56} : 1.00

CONCLUSIONES

Se recopiló la información estadística mediante una base de datos de migración del Ecuador del año 2021, la misma que cuenta con un total de 1 468 567 observaciones correspondientes a la data de salidas internacionales de ecuatorianos y extranjeros, distribuidas en 24 categorías como: tipo de movilidad (entrada/salida), tipo de nacionalidad (ecuatoriano/extranjero), año de movilidad (2021), mes de movilidad, sexo (hombre/mujer/indefinido), entre otros que constan en la base de datos.

Se planteó un modelo de regresión estadístico de tipo multinomial, el que precisa de utilizar la variable de respuesta nominal Y, en dependencia de las variables categóricas X, por lo que se identificó la variable dependiente (Y: Terrestre, marítima, aérea) y como variables independientes (nacionalidad, sexo, mes de movilidad, motivación del viaje, ocupación de persona que migra). Para llegar a la obtención del modelo:

$$\ln (p(Y)) = \beta_{01} - \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \dots \beta_5 X_5$$

Luego del ajuste, se concluye que este modelo matemático de regresión multinomial aplicado a los datos de entrenamiento (70% del conjunto de la data) da un 97,74% de precisión. Además, la variabilidad de la variable vía de transporte Y, tiene el 61% de explicación por el método McFadden y un 65% por el método Nagelkerke mediante las variables independientes, nacionalidad, sexo, mes de movilidad, motivación del viaje, ocupación de persona que migra.

Se concluye también que el modelo de regresión logística multinomial propuesto, sirve para la predicción de las vías de transporte de salida desde Ecuador, considerando las variables descritas para elaborar planes de mejora respecto al control y optimización de los flujos migratorios.

RECOMENDACIONES

Es recomendable adoptar la información estadística de la base de datos de migración del Ecuador, la misma que se considera confiable debido a que el tratamiento de este insumo está a cargo de la principal institución de estadísticas del país (INEC) y que aplican apropiadas metodologías para la recopilación de datos.

Se recomienda ejecutar varias alternativas de modelos matemáticos con diversos grupos de variables independientes, debido a que dependiendo de la data que se tenga podrán obtenerse resultados que sean oportunos y que cumplan con condiciones mínimas de confianza, que en esta investigación han sido ejecutadas por el método McFadden y Nagelkerke.

Otra recomendación está orientada al uso de data de entrenamiento y de data test, que es conveniente manejar en porcentajes como el aplicado de 70 – 30. A fin de que se puedan obtener resultados más confiables.

Finalmente se recomienda la revisión de esta investigación, a fin de utilizarla para posibles proyectos relacionados con las jefaturas migratorias, ciudades o poblaciones por donde las salidas de ciudadanos ecuatorianos o extranjeros tiene un mayor nivel de afluencia, debido a las variables independientes analizadas.

ANEXOS

Anexo A. Código de programación en R

```
#####  
# Análisis de cluster y regresion logistica  
#####  
  
#1.- Librerias  
  
library(dplyr)  
library(tidyverse)  
library(tidyr)  
library(magrittr)  
library(graphics)  
  
#2.- Lectura de base de datos  
  
data_migracion <- read.csv("C:/Users/Desktop/Tesis/esi_2021.csv", sep= ";", header=T,encoding="UTF-  
8")  
data_migracion <- filter(data_migracion, data_migracion$tip_movi == "Salida")  
str(data_migracion)  
  
# Estadisticxa descriptiva  
  
##### via_tran  
  
Tabla <- data_migracion %>% group_by(via_tran) %>% summarise(Total=n())  
Tabla <- Tabla %>% mutate(pcnt = `Total` / sum(`Total`)*100)  
Tabla  
  
sum(Tabla$Total)  
  
ggplot(Tabla, aes(x = via_tran , y=Total,fill=via_tran) ) +  
  
  geom_bar(width = 1, stat="identity",position = position_dodge() ) +  
  
  labs(x="Tipo", y= "N°de personas") +  
  
  labs(fill = "") +  
  
# ggtitle("Salida de ecuatorianos y extranjeros por tipo de via") +  
  
  geom_text(aes(label=Total), vjust=-0.5, color="black",  
    position = position_dodge(0.3), size=4.0  
  ) +  
  
  theme_bw(base_size = 15)  
  
##### sex_migr
```



```

Tabla <- data_migracion %>%
  group_by(sex_migr) %>%
  count() %>%
  ungroup() %>%
  mutate(Porcentaje = `n` / sum(`n`)) %>%
  arrange(Porcentaje) %>%
  mutate(etiquetas = scales::percent(Porcentaje)) %>%
  mutate(etiquetas_total = paste(sex_migr,etiquetas , " "))

```

```
library(plotrix)
```

```
pie3D(Tabla$Porcentaje,labels=Tabla$etiquetas_total,
  explode=0.09, labelcex = 1, col = c("blue","pink"))
```

```
legend(0,3, legend = Tabla$sex_migr, cex=0.7,yjust=0.2, xjust = -0.1,
  fill = heat.colors(length(Tabla$Porcentaje)))
```

```
##### tip_naci
```

```

Tabla <- data_migracion %>%
  group_by(tip_naci) %>%
  count() %>%
  ungroup() %>%
  mutate(Porcentaje = `n` / sum(`n`)) %>%
  arrange(Porcentaje) %>%
  mutate(etiquetas = scales::percent(Porcentaje)) %>%
  mutate(etiquetas_total = paste(tip_naci,etiquetas , " "))

```

```
Tabla
```

```
pie3D(Tabla$Porcentaje,labels=Tabla$etiquetas_total,
  explode=0.09, labelcex = 1, col = c("yellow","green"))
```

```
legend(0,3, legend = Tabla$tip_naci, cex=0.7,yjust=0.2, xjust = -0.1,
  fill = heat.colors(length(Tabla$Porcentaje)))
```

```
##### mes_movi
```

```
Tabla <- data_migracion %>% group_by(mes_movi) %>% summarise(Total=n()) %>% arrange(desc(
mes_movi))
```

```
Tabla <- Tabla %>% mutate(Mes_numero = c(9, 10, 11, 5, 3, 6, 7, 2, 1, 12, 8, 4))
```

```
Tabla <- Tabla %>% arrange(Mes_numero)
```

```
Tabla
```

```
# Gráfico de área
```

```
library(ggplot2)
```

```
ggplot(Tabla, aes(x=reorder(mes_movi, (Mes_numero)), y=Total ,group = 1)) +
  geom_line(colour="blue") +
  geom_point( size=2, shape=21, fill="white", colour="red") +
```

```
# ggtitle("Salida de ecuatorianos y extranjeros por mes") +
labs(x="Mes", y="N° personas") +

theme_bw(base_size = 15)
```

```
##### mot_viam
```

```
Tabla <- data_migracion %>%
  group_by(mot_viam) %>%
  count() %>%
  ungroup() %>%
  mutate(Porcentaje = `n` / sum(`n`)) %>%
  arrange(Porcentaje) %>%
  mutate(etiquetas = scales::percent(Porcentaje))
```

```
# En forma descendente:
```

```
ggplot( Tabla, aes(x=mot_viam, y=n)) +
  geom_bar(stat="identity", fill="darkblue", alpha=.6, width=.4) +
  coord_flip() +
  xlab("Motivos") +
  ylab("N° de personas") +
  theme_bw(base_size = 15)
```

```
##### ocu_migr
```

```
Tabla <- data_migracion %>%
  group_by(ocu_migr) %>%
  count() %>%
  ungroup() %>%
  mutate(Porcentaje = `n` / sum(`n`)) %>%
  arrange(Porcentaje) %>%
  mutate(etiquetas = scales::percent(Porcentaje))
```

```
Tabla = head(Tabla, n=10) %>% arrange(desc(n))
```

```
# En forma descendente:
```

```
ggplot( Tabla, aes(x=ocu_migr, y=n)) +
  geom_bar(stat="identity", fill="darkblue", alpha=.6, width=.4) +
  coord_flip() +
  xlab("N° de personas") +
  ylab("Ocupación") +

  theme_bw(base_size = 15)
```

Anexo B. Lectura de base de datos

```
data_migracion <- read.csv("C:/Users/Escritorio/Proyecto Tesis/Base de datos/esi_2021.csv", sep= ";",  
header=T) data_migracion <- filter(data_migracion, data_migracion$tip_movi == "Salida")  
str(data_migracion)
```

Anexo C. Número de personas por vía de transporte

```
> table(data_reg_2$via_tran)
```

Anexo D. Correlación.

```
> library(psych)  
> n<- dim(data_reg_2)[1]  
> set.seed(1234) #semillas  
> # select a random sample with  
> train <- sample(1:n , 0.7*n)  
> data.test <- data_reg_2[-train,]  
> data.train <- data_reg_2[train,]  
> ytrain <- data_reg_2$via_tran[train]  
> ytest <- data_reg_2$via_tran[-train]  
> str(data.train)  
  
'data.frame': 1027996 obs. of 8 variables:  
 $ tip_naci: chr "Extranjero" "Extranjero" "Extranjero" "Ecuatoriano" ...  
 $ mes_movi: chr "Abril" "Septiembre" "Abril" "Septiembre" ...  
 $ sex_migr: chr "Hombre" "Mujer" "Hombre" "Mujer" ...  
 $ via_tran: Factor w/ 4 levels "4","1","2","3": 4 4 4 4 4 4 4 4 4 4 ...  
 $ mot_viam: chr "Residencia" "Residencia" "Residencia" "Eventos" ...  
 $ pais_res: chr "Estados Unidos de América" "Costa Rica" "Ecuador" "Ecuador" ...  
 $ jef_migr: chr "Aeropuerto Internacional Mariscal Sucre" "Aeropuerto Internacional Mariscal  
Sucre" "Aeropuerto Internacional José Joaquín de Olmedo" "Aeropuerto Internacional José  
Joaquín de Olmedo" ...  
 $ ocu_migr: chr "Pilotos de aviacion y afines" "Comerciantes de tiendas" "Jubilados y  
pensionistas" "Jubilados y pensionistas"
```

Anexo E. Código para resultados de la correlación.

```
> unique(data_reg_2$via_tran)
> table(data.train$via_tran)
> str(data.train)

'data.frame':   1027996 obs. of  5 variables:
 $ tip_naci: chr "Extranjero" "Extranjero" "Extranjero" "Ecuatoriano" ...
 $ mes_movi: chr "Abril" "Septiembre" "Abril" "Septiembre" ...
 $ sex_migr: chr "Hombre" "Mujer" "Hombre" "Mujer" ...
 $ mot_viam: chr "Residencia" "Residencia" "Residencia" "Eventos" ...
 $ ocu_migr: chr "Pilotos de aviacion y afines" "Comerciantes de tiendas" "Jubilados y
 pensionistas" "Jubilados y pensionistas"
```

Anexo F. Código para cargar librerías necesarias para el modelo.

```
> library(VGAM)
> library(stats4)
> library(splines)
> unique(data.train$mes_movi)
```

Anexo G. Código para aplicación del modelo de regresión.

```
> library(nnet)
> multi_model <- multinom(
+ formula = via_tran ~ tip_naci+sex_migr+ mes_movi+ mot_viam + ocu_migr,
+ data = data.train + )

# weights: 680 (507 variable)
initial value 1425105.058069
iter 10 value 198370.516004
iter 20 value 170735.369092
iter 30 value 119093.042407
iter 40 value 100030.693921
iter 50 value 91670.519898
```

```
iter 60 value 82945.599396
iter 70 value 76105.943378
iter 80 value 72864.236429
iter 90 value 71235.904072
iter 100 value 70684.425248
final value 70684.425248
stopped after 100 iterations
multi_model
```

Anexo H. Código para llamar a la función del modelo de regresión multinomial

Call:

```
multinom(formula = via_tran ~ tip_naci + sex_migr + mes_movi + mot_viam + ocu_migr, data =
data.train)
```

Coefficients:

```
(Intercept) tip_naciExtranjero sex_migrIndefinido sex_migrMujer mes_moviAgosto
mes_moviDiciembre mes_moviEnero mes_moviFebrero mes_moviJulio mes_moviJunio
mes_moviMarzo mes_moviMayo mes_moviNoviembre mes_moviOctubre mes_moviSeptiembre
mot_viamEventos mot_viamNegocios mot_viam Otros mot_viamResidencia mot_viamTurismo
ocu_migrActores ocu_migrAgentes de servicios comerciales no clasificados bajo otros epigrafs
ocu_migrAgricultores y trabajadores calificados de huertas, invernaderos, viveros y jardines
ocu_migrAgricultores y trabajadores calificados de jardines y de cultivos para el mercado
ocu_migrAgronomos y afines ocu_migrAlbaniles ocu_migrAlfareros y afines (barro, arcilla y
abrasivos) ocu_migrAmas de casa ocu_migrAnalistas de gestion y organización
```

Anexo I. Código para la predicción con los datos de prueba.

```
> data.test$precticedt <- predict(multi_model, newdata = data.test, "class")
> ctable1 <- table(data.test$via_tran, data.test$precticedt)
> ctable1
```

Anexo J. Código para la predicción con los datos de prueba.

```
> round((sum(diag(ctable1))/sum(ctable 1))*100,2)
```

Anexo K. Código para la predicción con los datos de prueba.

```
> library(DescTools)
```

```
> DescTools::PseudoR2(multi_model, which = c("McFadden", "Nagelkerke"))
```

Anexo L. Código de programación en R. Modelo de regresión logística.

```
#####  
# Anàlisis de cluster y regresion logistica  
#####  
  
#1.- Librerias  
  
library(dplyr)  
library(dbplyr)  
library(tidyverse)  
library(tidyr)  
library(magrittr)  
library(graphics)  
  
#2.- Lectura de base de datos  
  
data_migracion <- read.csv("C:/Users/MICHAEL/OneDrive/Escritorio/Proyecto Tesis/Base de  
datos/esi_2021.csv", sep=";", header=T)  
data_migracion <- filter(data_migracion, data_migracion$tip_movi == "Salida")  
str(data_migracion)  
  
# =====  
# OPCION 2: Modelo  
#Variable dependiente: puerto aereo , maritimo, terrestre y fluvial  
# =====  
  
#Objetivo: Se le ha pedido que determine qué factores influyeron en la elección del tipo aereo etc  
  
#Selección de variables y recodificación a nivel o factores  
  
str(data_migracion)  
  
data_reg_2= select(data_migracion,-tip_movi, -anio_movi,-dia_movi,-subcont_nac,- cont_nac,-  
subcont_prod,-cont_prod,-lug_prod,-subcont_res,-pro_jefm , -pais_prod,-cla_migr, -cont_res, -nac_migr, -  
can_jefm, -edad)  
  
str(data_reg_2)  
  
unique(data_reg_2$tip_naci)  
levels(data_reg_2$tip_naci) <- c("Ecuatoriano","Extranjero")  
  
unique(data_reg_2$sex_migr)
```

```

levels(data_reg_2$sex_migr) <- c("Hombre", "Mujer", "Indefinido")

unique(data_reg_2$via_tran)
levels(data_reg_2$via_tran) <- c("Vía terrestre", "Vía Marítimo", "Vía Aérea", "Vía fluvial" )

unique(data_reg_2$mot_viam)
levels(data_reg_2$mot_viam) <- c("Turismo", "Otros", "Eventos", "Residencia", "Negocios", "Estudios" )

unique(data_reg_2$jef_migr)
levels(data_reg_2$jef_migr) <- c("Centro Binacional de Atención Fronteriza Huaquillas", "Unidad de
Control Migratorio Puerto Marítimo Guayaquil",
"Centro Nacional de Atención Fronteriza Rumichaca", "Unidad de Control Migratorio
Puerto Marítimo Manta",
"Aeropuerto Internacional Mariscal Sucre", "Unidad de Control Migratorio Puerto
Marítimo La Libertad",
"Aeropuerto Internacional José Joaquín de Olmedo", "Aeropuerto Internacional Eloy
Alfaro",
"Unidad de Control Migratorio Puerto Marítimo Esmeraldas", "Unidad de Control
Migratorio Puerto Marítimo Posorja",
"Aeropuerto Internacional Cotopaxi", "Unidad de Control Migratorio Puerto Marítimo
Ayora",
"Unidad de Control Migratorio Puerto Fluvial San Lorenzo", "Unidad de Control
Migratorio Puerto Marítimo Bolívar",
"Aeropuerto Mariscal Lamar", "Unidad de Control Migratorio Puerto Marítimo
Baquerizo Moreno",
"Unidad de Control Migratorio Puerto Marítimo Isabela", "Unidad de Control
Migratorio Puerto Marítimo Seymour",
"Centro Binacional de Atención Fronteriza San Miguel", "Aeropuerto Transfronterizo
Coronel Carlos Concha Torres",
"Unidad de Control Migratorio Frontera Terrestre Macara")

unique(data_reg_2$ocu_migr)
levels(data_reg_2$ocu_migr) <- c( "Capitanes, oficiales de cubierta y practicos"
,
"Marineros de cubierta y afines"
,
"Auxiliares laicos de las religiones"
,
"Directores generales y gerentes generales"
,
"Jubilados y pensionistas"
,
"Pilotos de aviacion y afines"
,
"Pescadores, cazadores y tramperos"
,
"Conductores de automoviles, taxis y camionetas"
,
"Personal de apoyo administrativo no clasificado bajo otros epigrafes"
,
"Sin especificar"
,
"Menores de edad"
,
"Estudiantes"
,
"Otros miembros de las fuerzas armadas"
,
"Comerciantes de tiendas"
,
"Amas de casa"
,
"Contables"
,
"Miembros del poder ejecutivo y de los gobiernos locales o seccionales"
,

```

"Tecnicos en ingenieria mecanica"

"Limpiadores y asistentes de oficinas, hoteles y otros establecimientos"

"Asesores financieros y en inversiones"

"Inspectores de la salud laboral, medioambiental y afines"

"Ingenieros mecanicos"

"Asistentes de venta de tiendas y almacenes"

"Ingenieros no clasificados bajo otros epigrafes"

"Chefs"

otros epigrafes" "Oficiales y operarios de la construccion (obra gruesa) y afines no clasificados bajo otros epigrafes"

"Periodistas"

"Tecnicos en ciencias fisicas y en ingenieria no clasificados bajo otros epigrafes"

"Practicantes paramedicos"

"Policias"

bajo otros epigrafes" "Oficiales, operarios y artesanos de artes mecanicas y de otros oficios no clasificados bajo otros epigrafes"

"Analistas financieros"

"Profesionales de la ensenanza no clasificados bajo otros epigrafes"

"Medicos especialistas"

"Abogados"

"Agricultores y trabajadores calificados de huertas, invernaderos, viveros y jardines"

"Matematicos, actuarios y estadisticos"

"Ingenieros civiles"

"Medicos generales"

"Agricultores y trabajadores calificados de jardines y de cultivos para el mercado"

"Disenadores y decoradores de interior"

"Auxiliares de servicios de abordo"

"Guardianes de proteccion"

"Albaniles"

"Criadores de ganado"

"Peluqueros"

"Atletas y deportistas"

"Electricistas de obras y afines"

"Profesionales del trabajo social"

"Economistas"

"Actores"

"Limpiadores y asistentes domesticos"

"Ingenieros industriales y de produccion"

"Geologos y geofisicos"

"Profesionales de enfermeria"

"Representantes comerciales"

"Desarrolladores de software"

"Lavaderos y planchadores manuales"

"Panaderos, pasteleros y confiteros"
 ,
 "Costureros, bordadores y afines"
 ,
 "Mineros y operadores de instalaciones mineras"
 ,
 "Joyereros, orfebres y plateros"
 ,
 "Pintores y empapeladores"
 ,
 "Peones de explotacion de cultivos mixtos y ganaderos"
 ,
 "Músicos, cantantes y compositores"
 ,
 "Profesores de formacion profesional"
 ,
 "Biólogos, botánicos, zoológicos y afines"
 ,
 "Fontaneros e instaladores de tuberías"
 ,
 "Dentistas"
 ,
 "Inspectores de policia y detectives"
 ,
 "Fotografos"
 ,
 "Cajeros de bancos y afines"
 ,
 "Cocineros"
 ,
 "Arquitectos"
 ,
 "Farmaceuticos"
 ,
 "Camareros de mesas"
 ,
 "Traductores, interpretes y lingüistas"
 ,
 "Ingenieros quimicos"
 ,
 "Profesionales de medicina tradicional y alternativa"
 ,
 "Psicologos"
 ,
 "Entrenadores instructores y arbitros de actividades deportivas"
 ,
 "Artesanos no clasificados bajo otros epigrafes"
 ,
 "Agronomos y afines"
 ,
 "Guías de turismo"
 ,
 "Oficinistas generales"
 ,
 "Mensajeros, mandaderos, maleteros y repartidores"
 ,
 "Productores y trabajadores calificados de explotaciones agropecuarias mixtas cuya produccion se destina al mercado" ,
 "Ingenieros electricistas"
 ,
 "Especialistas en tratamiento de belleza y afines"
 ,
 "Profesionales de la publicidad y la comercializacion"
 ,
 "Ebanistas y afines"
 ,
 "Carpinteros de armar y de obra blanca"
 ,
 "Ingenieros de minas, metalurgicos y afines"
 ,
 "Conserjes"
 ,
 "Maestros preescolares"
 ,
 "Diseñadores graficos y multimedia"

"Bibliotecarios, documentalistas y afines"

"Analistas de gestion y organizacion"

"Empacadores manuales"

"Secretarios (generales)"

"Agentes de servicios comerciales no clasificados bajo otros epigrafs"

"Alfareros y afines (barro, arcilla y abrasivos)"

"Modelos de moda, arte y publicidad"

"Ingenieros en telecomunicaciones"

"Profesionales de ventas tecnicas y medicas (excluyendo la tic)"

"Veterinarios"

"Profesionales de relaciones publicas"

"Trabajadores de explotacion de acuicultura"

"Profesionales de la proteccion medio ambiental"

"Sociologos, antropologos y afines"

"Artistas creativos e interpretativos no clasificados bajo otros epigrafs"

"Profesionales religiosos"

"Archivistas y curadores de museos"

"Bailarines y coreografos"

"Especialistas en politicas de administracion"

"Mecanicos y reparadores de vehiculos de motor"

"Zapateros y afines"

"Sastres, modistos, peleteros y sombrereros"

"Catadores y clasificadores de alimentos y bebidas"

"Artesanos de los tejidos, el cuero y materiales similares"

"Quimicos"

"Autores y otros escritores"

"Tecnicos de radiodifusion y grabacion audio visual"

"Filosofos, historiadores y especialistas en ciencias politicas"

"Bomberos"

"Profesionales en derecho no clasificados bajo otros epigrafs"

"Cobradores y afines"

"Tapiceros, colchoneros y afines"

"Jueces"

"Cuidadores de ninos"

"Artistas de artes plasticas"

"Locutores de radio, television y otros medios de comunicacion"

```

"Grabadores de datos"
"Profesionales de la salud y la higiene laboral y ambiental"
,
"Cartografos y agrimensores"
"Personal directivo de la administracion publica"
,
"Operadores de instalaciones de refinacion de petroleo y gas natural"
,
"Avicultores y trabajadores calificados de la avicultura"
,
"Apicultores y sericultores y trabajadores calificados de la apicultura y la sericultura"
,
"Peones de carga"
"Chapistas y caldereros"
"Delineantes y dibujantes Tecnicos"
,
"Economos y mayordomos domesticos"
,
"Fisicos y astronomicos"
"Meteorologos"
"Astrologos, adivinadores y afines"
,
"Otros profesores de artes"
"Tecnicos de protesis medicas y dentales"
,
"Operadores de maquinas para fabricar productos de caucho"
,
"Otros profesores de idiomas")

unique(data_reg_2$pais_res)
levels(data_reg_2$pais_res) <- c("Ecuador" , "Colombia","Estados Unidos de América" ,
"Perú","España"
,
"Alemania"
,
"Portugal", "Venezuela", "México", "Países Bajos", "Filipinas", "Federación de Rusia",
"Ucrania", "Brasil", "Argentina", "Catar", "Francia", "Costa Rica", "Canadá" ,
"Australia" , "Chile" , "Italia" , "Reino Unido" , "Sudáfrica" , "Nueva Zelanda" ,
"Croacia", "Irlanda" , "China" , "Panamá" , "Malasia", "Samoa" , "Singapur" ,
"Polonia" , "Finlandia" , "Guatemala", "Suecia" , "Israel", "República de Corea
del Sur" ,
"Myanmar", "Bélgica" , "Dinamarca" , "Turquía", "Austria", "Emiratos Árabes Unidos"
,
"Bulgaria", "Suiza" , "Bielorrusia" , "Arabia Saudita" , "Andorra", "Argelia",
"Rumania" ,
"Albania" , "Cuba", "Uruguay" , "El Salvador" , "República Dominicana" ,
"Honduras" , "Bolivia", "India" , "Curazao", "Serbia", " Hungría" , "Egipto",
"Líbano",
"Jamaica" , "República Democrática del Congo",
"Irán", "Islas Wallis y Futuna", "Etiopía" , "Islas Vírgenes Británicas" ,
"Guyana" , "Aruba", "Eslovenia", "Paraguay", "Bonaire, San Eustaquio y Saba",
"Costa de Marfil", "Eslovaquia", "República Centroafricana", "República Checa",
"Burkina Faso", "Malta", "Afganistán", "Pakistán", "Taiwán (República de China)",
"Marruecos", "Lituania", "Nigeria", "Indonesia" , "Haití", "Bahamas", "Zambia",
"Surinam",
"San Martín (Parte francesa)", "Puerto Rico", "Armenia", "Grecia", "Noruega",
"Tailandia", "Bermudas", "San Vicente y las Granadinas", "Kuwait" , "Nicaragua" ,
"Sin especificar" , "Islas Vírgenes de los Estados Unidos", "Islas Turcas y Caicos" ,
"Papúa Nueva Guinea" , "Polinesia francesa" , "Omán",
"Estonia", "Japón" , "Letonia" , "Kenia" , "Yemen", "Libia" , "Guadalupe"
, "Luxemburgo" ,

```

```

        "Somalia" , "Bangladesh" , "Lesoto" , "Senegal", "Belice" , "Micronesia (Estados
Federados de)",
        "Barein" , "Hong Kong (Región Administrativa Especial de China)" , "Cabo
Verde" , "Jordania" , "Túnez" , "Kazajistán" , "Angola" ,
        "Antigua República Yugoslava de Macedonia" , "Georgia" , "Islas Caimán" ,
        "Irak" , "Liechtenstein" , "Guinea Ecuatorial" , "Azerbaiyán"
        , "Madagascar" , "Antártida" , "Vietnam" , "Guayana Francesa" ,
        "Islandia" , "Mozambique" , "Chipre" , "Montenegro" , "República de Moldovia" ,
        "Turkmenistán" , "Trinidad y Tobago" , "Uganda" , "Ghana" , "Palestina" ,
        "Camerún" , "San Cristóbal y Nieves" , "Congo" , "Kirguistán" ,
        "Uzbekistán" , "Nueva Caledonia" , "Mónaco" ,
"Namibia" , "Vanuatu" , "República Popular Democrática de Corea" ,
        "República Árabe Siria" , "Dominica" , "Ruanda" , "Mauritania" , "Tanzania (República
Unida de)" , "Barbados" , "Fiyi" , "Tayikistán" , "Kosovo" , "Anguila" , "Níger" , "Benín" , "Brunei Darussalam"
, "Sudán" ,
        "Bosnia y Herzegovina" , "Suazilandia" , "Granada" , "Eritrea" ,
        "Gibraltar" , "Chad" , "Malí" , "Antigua y Barbuda" , "Guinea" ,
"Zimbabue" , "Nepal" , "Camboya" ,
        "Togo" , "Mongolia" , "Gabón" , "Santa Sede (Vaticano)" , "Comoras" , "Guinea-Bisau" ,
"San Marino" , "Sint Maarten (Parte Holandesa)" , "Santa Lucía" , "Isla de Man" , "Mauricio" , "Macao")

str(data_reg_2)

#Trasformación de la variable dependiente a numerico

data_reg_2$via_tran <- ifelse(data_reg_2$via_tran == "Vía terrestre", 1,
        (ifelse(data_reg_2$via_tran == "Vía Maritimo", 2,
        (ifelse(data_reg_2$via_tran == "Vía Aérea", 3, 4))))))

data_reg_2$via_tran <- relevel(as.factor(data_reg_2$via_tran), ref = 4)

#Número de observaciones o personas en el estudio
nrow(data_reg_2)
str(data_reg_2)

#Número de personas por puerto
table(data_reg_2$via_tran)

#Correlacion
library(psych)

#corPlot(data_reg_2, cex = 1.2, main = "Matriz de correlación")
#PENDIENTE

#### 2.- Creación muestras de entrenamiento y prueba
#Objetivo: Se le ha pedido que determine qué factores influyeron en la elección del tipo aereo etc

n <- dim(data_reg_2)[1]

set.seed(1234) #semillas

# select a random sample with
train <- sample(1:n , 0.7*n)

data.test <- data_reg_2[-train,]
data.train <- data_reg_2[train,]

```

```

ytrain <- data_reg_2$via_tran[train]
ytest <- data_reg_2$via_tran[-train]

str(data.train)
unique(data_reg_2$via_tran)

#### 3. Ajustar el modelo de regresión logística

table(data.train$via_tran)
str(data.train)

library(VGAM)
library(stats4)
library(splines)

#Combinacion de variables

# Opcion 1: tip_naci+sex_migr
# Opcion 2: mot_viam
# Opcion 3: tip_naci+sex_migr+mes_movi+ mot_viam (Bien)

unique(data.train$mes_moeevi)

library(nnet)

multi_model <- multinom(
  formula = via_tran ~ tip_naci+sex_migr+ mes_movi+ mot_viam + ocu_migr,
  data = data.train
)

# Descripcion de los resultados
summary(multi_model)

# Calcular z-statistics de los coeficientes

#1.-Hipotesis
#H0: variable no significativa
#H1: variable significativa aportar al modelo

#Nivel significancia 0,05 significa 95% confianza

# p < 0,05 Se va a rechazar la hipotesis nula.

z_stats <- summary(multi_model)$coefficients/
  summary(multi_model)$standard.errors

# Convertir a probabilidades
p_values <- (1 - pnorm(abs(z_stats)))*2
data.frame(t(p_values))

# El valor p nos dice que las variables son significativas.
#Ahora exploraremos todo el conjunto de datos y
# analizaremos si podemos eliminar cualquier variable que no se sume al rendimiento del modelo.

#Conclusion: Dado que el p valor para todas las variables es menor a 0.05, concluimos

```

con un 95% de confianza que las variables son significativas para el modelo.

4. Predicción

```
# Predecir los valores para el conjunto de datos del test o prueba
data.test$precticedt <- predict(multi_model, newdata = data.test, "class")
```

```
# Tabla de clasificación
# Nombre
```

```
ctable1 <- table(data.test$via_tran, data.test$precticedt)
ctable1
```

5. Validacion del modelo

```
# Precisión de cálculo (Accuracy): suma de elementos diagonales dividida por obs total
round((sum(diag(ctable1))/sum(ctable1))*100,2)
```

Conclusión: La precisión en el conjunto de datos de entrenamiento es del 97,74%.

```
#Evaluar la confianza del modelo
library(DescTools)
DescTools::PseudoR2(multi_model, which = c("McFadden", "Nagelkerke"))
```

#Implicaría que las variables independientes empleadas (tip_naci+sex_migr+ mes_movi+ mot_viam) en el modelo de regresión solamente explican el 22% de la varianza de la variable dependiente #0.61 (61%) y 0.65(65%).

Anexo M. Clasificación de salida de ecuatorianos y extranjeros, de acuerdo con su ocupación y porcentaje correspondiente.

Ocupación	N° de personas	Porcentaje
Abogados	19859	1.35%
Actores	593	0.04%
Agentes de servicios comerciales no clasificados bajo otros epigrafs	781	0.05%
Agricultores y trabajadores calificados de huertas, invernaderos, viveros y jardines	916	0.06%
Agricultores y trabajadores calificados de jardines y de cultivos para el mercado	6865	0.47%
Agronomos y afines	4941	0.34%
Albaniles	6802	0.46%
Alfareros y afines (barro, arcilla y abrasivos)	36	0.00%
Amas de casa	107935	7.35%
Analistas de gestion y organizacion	1902	0.13%
Analistas financieros	1569	0.11%

Apicultores y sericultores y trabajadores calificados de la apicultura y la sericultura	41	0.00%
Archivistas y curadores de museos	87	0.01%
Arquitectos	6303	0.43%
Artesanos de los tejidos, el cuero y materiales similares	88	0.01%
Artesanos no clasificados bajo otros epigrafs	2299	0.16%
Artistas creativos e interpretativos no clasificados bajo otros epigrafs	959	0.07%
Artistas de artes plsticas	79	0.01%
Asesores financieros y en inversiones	3081	0.21%
Asistentes de venta de tiendas y almacenes	7633	0.52%
Astrologos, adivinadores y afines	7	0.00%
Atletas y deportistas	6601	0.45%
Autores y otros escritores	383	0.03%
Auxiliares de servicios de abord	8348	0.57%
Auxiliares laicos de las religiones	4615	0.31%
Avicultores y trabajadores calificados de la avicultura	159	0.01%
Bailarines y coregrafos	312	0.02%
Bibliotecarios, documentalistas y afines	92	0.01%
Biologos, botanicos, zoologos y afines	870	0.06%
Bomberos	671	0.05%
Cajeros de bancos y afines	1586	0.11%
Camareros de mesas	2136	0.15%
Capitanes, oficiales de cubierta y prcticos	60138	4.10%
Carpinteros de armar y de obra blanca	1491	0.10%
Cartografos y agrimensores	92	0.01%
Catadores y clasificadores de alimentos y bebidas	10	0.00%
Chapistas y caldereros	126	0.01%
Chefs	2642	0.18%
Cobradores y afines	111	0.01%
Cocineros	1673	0.11%
Comerciantes de tiendas	46657	3.18%
Conductores de automoviles, taxis y camionetas	13864	0.94%
Conserjes	428	0.03%
Contables	10624	0.72%
Costureros, bordadores y afines	1375	0.09%
Criadores de ganado	1516	0.10%
Cuidadores de ninos	544	0.04%
Delineantes y dibujantes Tecnicos	58	0.00%
Dentistas	7361	0.50%
Desarrolladores de software	445	0.03%
Directores generales y gerentes generales	28820	1.96%
Disenadores graficos y multimedia	873	0.06%

Disenadores y decoradores de interior	4538	0.31%
Ebanistas y afines	443	0.03%
Economistas	7990	0.54%
Economos y mayordomos domesticos	39	0.00%
Electricistas de obras y afines	1906	0.13%
Empacadores manuales	851	0.06%
Entrenadores instructores y arbitros de actividades deportivas	1959	0.13%
Especialistas en politicas de administracion	383	0.03%
Especialistas en tratamiento de belleza y afines	1435	0.10%
Estudiantes	146328	9.96%
Farmacuticos	873	0.06%
Filosofos, historiadores y especialistas en ciencias politicas	313	0.02%
Fisicos y astronomicos	74	0.01%
Fontaneros e instaladores de tuberias	180	0.01%
Fotografos	994	0.07%
Geologos y geofisicos	660	0.04%
Grabadores de datos	69	0.00%
Guardianes de proteccion	1238	0.08%
Guias de turismo	822	0.06%
Ingenieros civiles	5626	0.38%
Ingenieros de minas, metalurgicos y afines	1007	0.07%
Ingenieros electricistas	2758	0.19%
Ingenieros en telecomunicaciones	895	0.06%
Ingenieros industriales y de produccion	3733	0.25%
Ingenieros mecanicos	3135	0.21%
Ingenieros no clasificados bajo otros epigrafes	46219	3.15%
Ingenieros quimicos	1199	0.08%
Inspectores de la salud laboral, medioambiental y afines	245	0.02%
Inspectores de policia y detectives	404	0.03%
Joyeros, orfebres y plateros	439	0.03%
Jubilados y pensionistas	89242	6.08%
Jueces	176	0.01%
Lavanderos y planchadores manuales	53	0.00%
Limpiadores y asistentes de oficinas, hoteles y otros establecimientos	6135	0.42%
Limpiadores y asistentes domesticos	6941	0.47%
Locutores de radio, television y otros medios de comunicacion	109	0.01%
Maestros preescolares	253	0.02%
Marineros de cubierta y afines	5784	0.39%
Matematicos, actuarios y estadisticos	544	0.04%
Mecanicos y reparadores de vehiculos de motor	240	0.02%
Medicos especialistas	812	0.06%

Medicos generales	22438	1.53%
Menores de edad	180466	12.29%
Mensajeros, mandaderos, maleteros y repartidores	195	0.01%
Meteorologos	16	0.00%
Miembros del poder ejecutivo y de los gobiernos locales o seccionales	2757	0.19%
Mineros y operadores de instalaciones mineras	796	0.05%
Modelos de moda, arte y publicidad	294	0.02%
Musicos, cantantes y compositores	2108	0.14%
Oficiales y operarios de la construccion (obra gruesa) y afines no clasificados bajo otros epigrafes	2129	0.14%
Oficiales, operarios y artesanos de artes mecanicas y de otros oficios no clasificados bajo otros epigrafes	8745	0.60%
Oficinistas generales	540	0.04%
Operadores de instalaciones de refinacion de petroleo y gas natural	140	0.01%
Operadores de maquinas para fabricar productos de caucho	15	0.00%
Otros miembros de las fuerzas armadas	3940	0.27%
Otros profesores de artes	32	0.00%
Otros profesores de idiomas	16	0.00%
Panaderos, pasteleros y confiteros	740	0.05%
Peluqueros	3926	0.27%
Peones de carga	84	0.01%
Peones de explotacion de cultivos mixtos y ganaderos	318	0.02%
Periodistas	4545	0.31%
Personal de apoyo administrativo no clasificado bajo otros epigrafes	380698	25.92%
Personal directivo de la administracion publica	57	0.00%
Pescadores, cazadores y tramperos	1367	0.09%
Pilotos de aviacion y afines	21106	1.44%
Pintores y empapeladores	840	0.06%
Policias	2620	0.18%
Practicantes paramedicos	199	0.01%
Productores y trabajadores calificados de explotaciones agropecuarias mixtas cuya produccion se destina al mercado	1045	0.07%
Profesionales de enfermeria	4774	0.33%
Profesionales de la ensenanza no clasificados bajo otros epigrafes	16137	1.10%
Profesionales de la proteccion medio ambiental	734	0.05%
Profesionales de la publicidad y la comercializacion	5695	0.39%
Profesionales de la salud y la higiene laboral y ambiental	106	0.01%
Profesionales de medicina tradicional y alternativa	925	0.06%
Profesionales de relaciones publicas	691	0.05%
Profesionales de ventas tecnicas y medicas (excluyendo la tic)	876	0.06%
Profesionales del trabajo social	766	0.05%
Profesionales en derecho no clasificados bajo otros epigrafes	106	0.01%
Profesionales religiosos	708	0.05%

Profesores de formacion profesional	8814	0.60%
Psicologos	3227	0.22%
Quimicos	639	0.04%
Representantes comerciales	2481	0.17%
Sastres, modistos, peleteros y sombrereros	560	0.04%
Secretarios (generales)	2867	0.20%
Sin especificar	37848	2.58%
Sociologos, antropologos y afines	700	0.05%
Tapiceros, colchoneros y afines	95	0.01%
Tecnicos de protesis medicas y dentales	6	0.00%
Tecnicos de radiodifusion y grabacion audio visual	386	0.03%
Tecnicos en ciencias fisicas y en ingenieria no clasificados bajo otros epigrafe s	5501	0.37%
Tecnicos en ingenieria mecanica	4125	0.28%
Trabajadores de explotacion de acuicultura	676	0.05%
Traductores, interpretes y linguistas	534	0.04%
Veterinarios	1535	0.10%
Zapateros y afines	442	0.03%

Anexo N Coeficientes completos del modelo de regresión logística.

Coefficients:

(Intercept) tip_naciExtranjero sex_migrIndefinido sex_migrMujer
mes_moviAgosto mes_moviDiciembre mes_moviEnero mes_moviFebrero
mes_moviJulio

1	21.30445	-5.278820	-0.2453246	18.08278
5.640465	5.790002	-19.78527	2.885485	5.207538
2	28.44352	-4.539596	-1.3797305	17.84881
4.824597	4.031072	-19.79478	2.449289	4.273021
3	41.87477	-4.982258	1.6287817	19.77552
6.062142	4.029966	-19.83336	2.854355	5.359776

mes_moviJunio mes_moviMarzo mes_moviMayo mes_moviNoviembre
mes_moviOctubre mes_moviSeptiembre mot_viamEventos mot_viamNegocios
mot_viamOtros

1	4.634730	3.793314	4.604100	5.534855
5.593431	5.447509	7.874070	10.496178	19.120968
2	4.122942	3.779063	4.487615	4.012539
5.473807	4.911353	2.342224	-1.912734	11.243680
3	4.656540	3.788691	4.809480	5.362860
5.395493	5.664050	-3.878999	-6.338072	2.364313

mot_viamResidencia mot_viamTurismo ocu_migrActores ocu_migrAgentes de servicios comerciales no clasificados bajo otros epigrafs

1 4.530929 27.02322 -0.1079299
0.8265523

2 -4.862220 19.66179 1.6278063
-1.5458087

3 -10.331351 12.48787 -1.2877571
0.8315566

ocu_migrAgricultores y trabajadores calificados de huertas, invernaderos, viveros y jardines

1 1.083667

2 -2.163437

3 1.306146

ocu_migrAgricultores y trabajadores calificados de jardines y de cultivos para el mercado ocu_migrAgronomos y afines ocu_migrAlbaniles

1 3.269131 0.9039551 -0.09011206

2 -4.879966 -0.7260273 0.90523218

3 2.695828 0.6401885 1.00468363

ocu_migrAlfareros y afines (barro, arcilla y abrasivos) ocu_migrAmas de casa ocu_migrAnalistas de gestion y organizacion

1 4.263271 -1.4339522

2 2.828423 -0.8152663 0.2982706

3 2.695235 0.2931320 1.1578593

0.8259672

ocu_migrAnalistas financieros ocu_migrApicultores y sericultores y trabajadores calificados de la apicultura y la sericultura

1 -2.846366

-1.23413612

2 1.557985

0.04455969

3 1.598421

1.20008146

ocu_migrArchivistas y curadores de museos ocu_migrArquitectos ocu_migrArtesanos de los tejidos, el cuero y materiales similares

1 1.3367370 1.6925114 0.2578200

2 -0.3621537 -0.7902912 0.4161827

3 -0.9443498 -0.8932000 0.4412965

-0.9443498

ocu_migrArtesanos no clasificados bajo otros epigrafs

ocu_migrArtistas creativos e interpretativos no clasificados bajo otros epigrafs

1 -0.1413645 1.2192305

2 1.0361645 -1.7808770

3 -0.5515222 0.8936813

-0.5515222

ocu_migrArtistas de artes plasticas ocu_migrAsesores financieros y en inversiones ocu_migrAsistentes de venta de tiendas y almacenes

1 0.5460042 -1.02026003

2 0.3496517 -0.02879184 1.0859884

3 -0.2522781 1.05990988 -0.2875916

0.5675445

ocu_migrAstrologos, adivinadores y afines ocu_migrAtletas y deportistas ocu_migrAutores y otros escritores

1				-1.08328973
2.468014				1.4084696
2				0.01220832
2.582483				0.3124745
3				1.07178980
1.134321				-1.5288597
ocu_migrAuxiliares de servicios de abord o cu_migrAuxiliares laicos de las religiones				
1				-2.5868398
1.8104342				2.9177924
2				0.4705124
-1.9160707				
3				
0.8881248				
ocu_migrAvicultores y trabajadores calificados de la avicultura				
ocu_migrBailarines y coreografos				
1				-0.8533926
1.5262199				-0.1918545
2				
-1.0272535				1.0720185
3				
-0.4476234				
ocu_migrBibliotecarios, documentalistas y afines o cu_migrBiologos, botanicos, zoologos y afines o cu_migrBomberos				
1				-0.4177522
-1.408500	3.6808434			
2				-0.2739007
2.773197	-2.9448187			
3				0.7264076
-1.153426	-0.4808274			
ocu_migrCajeros de bancos y afines o cu_migrCamareros de mesas				
ocu_migrCapitanes, oficiales de cubierta y practicos				
1		1.9017354		-0.8240271
6.4566809				
2		-0.3291608		2.9208315
6.0013180				
3		-1.2183966		-1.3835382
0.6708606				
ocu_migrCarpinteros de armar y de obra blanca o cu_migrCartografos y agrimensores o cu_migrCatadores y clasificadores de alimentos y bebidas				
1				-0.98703481
0.5040236				2.06405592
2			1.28776442	
1.1375621				-0.09853071
3			0.01887554	
-1.5432119				-1.90726462
ocu_migrChapistas y caldereros o cu_migrChefs o cu_migrCobradores y afines o cu_migrCocineros o cu_migrComerciantes de tiendas				
1		1.6553593	-0.4460159	-
0.5983786	1.8200596		3.0340019	-
2	-0.6465881	1.6826208		-
0.2825348	-0.2540912		1.8974096	
3	-0.9282358	-0.4332623		
0.8978754	-0.7904743		0.9233346	
ocu_migrConductores de automoviles, taxis y camionetas				
ocu_migrConserjes o cu_migrContables o cu_migrCostureros, bordadores y afines				
1			3.7372436	
0.3559051	1.0321373			0.99729321
2			1.6290762	-
0.8059808	0.3401492			-0.95432803
3			-0.2360302	
0.5194952	0.1935623			0.09018798
ocu_migrCriadores de ganado o cu_migrCuidadores de ninos				
ocu_migrDelineantes y dibujantes Tecnicos o cu_migrDentistas				
1		1.102501		-2.1326967
-0.4342810	0.3707083			

2	-1.973825	3.0132355	
-0.1596884	0.3304592		
3	1.094456	-0.8314544	
0.6121131	0.2284914		
ocu_migrDesarrolladores de software ocu_migrDirectores generales y gerentes generales ocu_migrDiseñadores gráficos y multimedia			
1	-2.2979218		
-0.01961188		1.436564	
2	2.1878741		
2.41162818		-2.501561	
3	0.2523524		
1.46934291		1.227108	
ocu_migrDiseñadores y decoradores de interior ocu_migrEbanistas y afines ocu_migrEconomistas ocu_migrEconomos y mayordomos domesticos			
1		0.43915885	
1.5265250	1.091206		-1.364643
2		-0.06179566	-
1.2181918	-1.690380		2.904546
3		0.21490515	-
0.1897612	1.728043		-1.495335
ocu_migrElectricistas de obras y afines ocu_migrEmpacadores manuales ocu_migrEntrenadores instructores y arbitros de actividades deportivas			
1		-0.9760555	0.09881066
1.4087638			
2		2.0292398	1.20671729
-1.5700766			
3		-0.4792726	-0.81354611
0.5702933			
ocu_migrEspecialistas en politicas de administracion ocu_migrEspecialistas en tratamiento de belleza y afines ocu_migrEstudiantes			
1		0.6341121	
1.3364196	5.116630		
2		-0.8897638	
-1.6900971	4.927836		
3		0.3193338	
0.4837896	4.292618		
ocu_migrFarmaceuticos ocu_migrFilosofos, historiadores y especialistas en ciencias politicas ocu_migrFisicos y astronomicos			
1	0.3058990		
0.6853176		-1.856840	
2	-0.9127570		
0.7498094		3.337060	
3	0.7455071		
-1.3319367		-1.427101	
ocu_migrFontaneros e instaladores de tuberias ocu_migrFotografos ocu_migrGeologos y geofisicos ocu_migrGrabadores de datos			
1		-0.9902122	0.2926483
-1.2382095		-0.62823322	
2		-0.1876129	0.4133349
1.2417250		-0.05288319	
3		1.2153122	-0.4734911
0.1184849	0.69310204		
ocu_migrGuardianes de proteccion ocu_migrGuias de turismo ocu_migrIngenieros civiles ocu_migrIngenieros de minas, metalurgicos y afines			
1	-0.32048546	0.2726359	
1.145214		0.8077334	
2	0.74795577	-1.0718859	
-1.830038		-1.0619275	
3	-0.03452399	0.9393159	
1.593115		0.4750736	
ocu_migrIngenieros electricistas ocu_migrIngenieros en telecomunicaciones ocu_migrIngenieros industriales y de produccion			
1	-0.2951637		
0.01936247		0.3305732	

2	0.6130078		-
1.09072379		-1.2950333	
3	0.3101174		
1.23881626		1.5219760	
ocu_migrIngenieros mecanicos ocu_migrIngenieros no clasificados bajo otros epigrafes ocu_migrIngenieros quimicos			
1	-1.2209247		
1.477646	0.9679592		
2	1.4643455		
2.976203	-1.0405828		
3	0.4721491		
1.967441	0.2664860		
ocu_migrInspectores de la salud laboral, medioambiental y afines ocu_migrInspectores de policia y detectives			
1		-1.4718966	
-0.3659547			
2		0.0450586	
1.0519227			
3		1.4793613	
-0.5294707			
ocu_migrJoyereros, orfebres y plateros ocu_migrJubilados y pensionistas ocu_migrJueces ocu_migrLavanderos y planchadores manuales			
1	0.1805757		0.758219
-1.359520656		2.0932221	
2	-0.4761440		4.589425
0.007506107		-0.5083256	
3	0.4054189		3.540530
1.383413319		-1.5788225	
ocu_migrLimpiadores y asistentes de oficinas, hoteles y otros establecimientos ocu_migrLimpiadores y asistentes domesticos			
1		0.02039392	
0.04742442			
2		1.69475639	
0.46138189			
3		-0.62647073	
0.24573489			
ocu_migrLocutores de radio, television y otros medios de comunicacion ocu_migrMaestros preescolares ocu_migrMarineros de cubierta y afines			
1		1.0489689	
-0.5702835	0.3124655		
2		-0.6851471	
-0.2191686	6.7723608		
3		-0.3298828	
0.8123664	-0.7904187		
ocu_migrMatematicos, actuarios y estadisticos ocu_migrMecanicos y reparadores de vehiculos de motor ocu_migrMedicos especialistas			
1	0.4160365		
0.8468999	-1.3130658		
2	0.5842541		
-0.9441878	2.2050044		
3	-0.8761795		
0.1550280	-0.7362908		
ocu_migrMedicos generales ocu_migrMenores de edad ocu_migrMensajeros, mandaderos, maleteros y repartidores ocu_migrMeteorologos			
1	0.7607602	6.126378	
0.59508255	-0.09479743		
2	0.8064123	5.347221	
-0.55149410	-0.05121877		
3	1.2354742	5.241495	
0.03770025	0.15526524		
ocu_migrMiembros del poder ejecutivo y de los gobiernos locales o seccionales ocu_migrMineros y operadores de instalaciones mineras			
1			-
3.173601		-5.2251255	
2			-
7.124574		-0.9492689	

3			-
4.414955		6.3009059	
ocu_migrModelos de moda, arte y publicidad y compositores	ocu_migrMusicos, cantantes		
1	1.5082843		
0.6943059			
2	-0.7436134		
0.5259740			
3	-0.7382652		
-0.6718284			
ocu_migrOficiales y operarios de la construccion (obra gruesa) y afines no clasificados bajo otros epigrafes			
1			
-0.3417542			
2			
1.3449267			
3			
-0.4054909			
ocu_migrOficiales, operarios y artesanos de artes mecanicas y de otros oficios no clasificados bajo otros epigrafes			
1			
1.1618596			
2			
0.7322408			
3			
0.3688340			
ocu_migrOficinistas generales ocu_migrOperadores de instalaciones de refinacion de petroleo y gas natural			
1	-0.08598103		
-3.62705420			
2	1.11452545		
0.03248199			
3	-0.81885104		
3.63955590			
ocu_migrOperadores de maquinas para fabricar productos de caucho ocu_migrOtros miembros de las fuerzas armadas			
1		-0.60239373	
-5.15059080			
2		-0.04689838	
-0.09927016			
3		0.65684376	
-5.10813824			
ocu_migrOtros profesores de artes ocu_migrOtros profesores de idiomas ocu_migrPanaderos, pasteleros y confiteros ocu_migrPeluqueros			
1	-0.3490886		-0.43688369
0.2371476	0.8441985		
2	-0.1149930		-0.08245533
0.3280774	0.4828562		
3	0.4669685		0.52439740
-0.3736688	-0.6179591		
ocu_migrPeones de carga ocu_migrPeones de explotacion de cultivos mixtos y ganaderos ocu_migrPeriodistas			
1	1.1760975		
0.8657006	-2.259182		
2	-0.7024212		
0.5916310	-2.548035		
3	-0.4136794		
-1.2080305	-2.854515		
ocu_migrPersonal de apoyo administrativo no clasificado bajo otros epigrafes ocu_migrPersonal directivo de la administracion publica			
1			-
0.1094952		-0.3805008	
2			
0.9560384		-0.1812648	
3			-
0.3316409		0.5669553	

ocu_migrPescadores, cazadores y tramperosPescadores, cazadores y tramperos		
1		-
0.6880821	1.960297	
2		
5.4197149	2.576463	
3		-
2.5559876	2.655583	
ocu_migrPintores y empapeladores ocu_migrPolicias ocu_migrPracticantes paramedicos		
1	0.43303817	-2.194669
-1.968152		
2	-0.18018143	-3.319657
-2.670303		
3	0.01520209	-2.396845
-3.612109		
ocu_migrProductores y trabajadores calificados de explotaciones agropecuarias mixtas cuya produccion se destina al mercado		
1		
0.6649546		
2		
-1.4008934		
3		
0.9136198		
ocu_migrProfesionales de enfermeria ocu_migrProfesionales de la ensenanza no clasificados bajo otros epigrafes		
1	0.7145178	
0.5408076		
2	0.9470314	
0.8430271		
3	-0.9706174	
0.6522113		
ocu_migrProfesionales de la proteccion medio ambiental ocu_migrProfesionales de la publicidad y la comercializacion		
1		-0.5220872
-0.3939923		
2		0.5106512
0.8053246		
3		0.1169488
0.3985814		
ocu_migrProfesionales de la salud y la higiene laboral y ambiental ocu_migrProfesionales de medicina tradicional y alternativa		
1		-0.2664600
0.2180636		
2		-0.1028161
0.7480719		
3		0.3789926
-0.6348705		
ocu_migrProfesionales de relaciones publicas ocu_migrProfesionales de ventas tecnicas y medicas (excluyendo la tic)		
1	0.4527037	
-4.101619		
2	-	
0.6984466		
1.262669		
3	0.3185661	
2.940232		
ocu_migrProfesionales del trabajo social ocu_migrProfesionales en derecho no clasificados bajo otros epigrafes		
1	1.7348216	
-1.07506141		
2	-0.9458374	
-0.04751654		
3	-0.6450452	
1.13858391		
ocu_migrProfesionales religiosos ocu_migrProfesores de formacion profesional ocu_migrPsicologos ocu_migrQuimicos		

1	-0.3641805	
0.74866970	1.2915814	-0.2524283
2	-1.7930331	
0.41489139	-1.2987464	0.6829246
3	2.3379114	
0.06120764	0.4206496	-0.3120035
ocu_migrRepresentantes comerciales ocu_migrSastres, modistos, peleteros y sombrereros ocu_migrSecretarios (generales)		
1	0.7520040	
1.1554521	0.78674614	
2	-1.3319901	
-1.2701187	-0.05556712	
3	0.9846061	
0.1959166	-0.45041684	
ocu_migrSin especificar ocu_migrSociologos, antropologos y afines ocu_migrTapiceros, colchoneros y afines		
1	-0.3418145	1.024156925
-0.6337848		
2	-2.3989340	-0.927415511
-0.3115262		
3	-3.1669154	0.002963529
0.9636529		
ocu_migrTecnicos de protesis medicas y dentales ocu_migrTecnicos de radiodifusion y grabacion audio visual		
1		-0.03770686
-0.4643508		
2		-0.00983658
1.6755034		
3		0.04780169
-1.0283536		
ocu_migrTecnicos en ciencias fisicas y en ingenieria no clasificados bajo otros epigrafes ocu_migrTecnicos en ingenieria mecanica		
1		0.001147509
-1.5458505		
2		1.477860831
-0.6387203		
3		-0.033153891
-2.0436754		
ocu_migrTrabajadores de explotacion de acuicultura ocu_migrTraductores, interpretes y linguistas ocu_migrVeterinarios		
1		-0.06662255
0.780347229	0.704616	
2		0.64936152
-0.663927803	-1.533891	
3		-0.37287681
-0.007730895	1.061639	
ocu_migrZapateros y afines		
1	0.6964809	
2	-1.1905048	
3	0.5729146	

REFERENCIAS

- ACNUR. (2017). <https://www.acnur.org>. Obtenido de <https://www.acnur.org/fileadmin/Documentos/BDL/2017/10973.pdf>
- Banco Central del Ecuador. (2022). <https://www.bce.fin.ec/index.php/component/k2/item/348-ley-de-estad%C3%ADstica>. Obtenido de <https://www.bce.fin.ec/index.php/component/k2/item/348-ley-de-estad%C3%ADstica>
- Barcelona Geeks. (06 de 2022). <https://barcelonageeks.com/funcion-unica-en-r/>. Obtenido de <https://barcelonageeks.com/funcion-unica-en-r/>
- Bernabé, E. (2018). El cambio climático como una causa de migración: Un análisis de regresión no lineal para el estado de Veracruz 2005-2010. [Tesis de Pregrado, Universidad Autónoma de del estado de México]. UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO. Obtenido de <http://ri.uaemex.mx/handle/20.500.11799/95347>
- Bernal, C. (2010). *Metodología de la Investigación* (Tercera Edición ed.). Bogotá: Pearson. Obtenido de <https://anyflip.com/vede/ohla/basic>
- Booksdown. (2022). <https://bookdown.org>. Obtenido de <https://bookdown.org/jboscomendoza/r-principiantes4/quien-usa-r.html>
- Bustos , A. (2020). Efecto de la migración venezolana sobre el desempleo en Combia: 2012-2019. [Tesis dde Pregrado, Fundación Universitaria Los Libertadores]. Fundación Universitaria los Libertadores. Obtenido de <https://repository.libertadores.edu.co/handle/11371/3126>
- Castañeda, J. (2020). Determinantes de la migración interna juvenil por acceso a educación superior pública hacia la provincia de Pichincha. [Tesis de Pregrado,

Escuela Politécnica Nacional]. Obtenido de <https://bibdigital.epn.edu.ec/handle/15000/21230>

Datos Macro. (s.f.). *Ecuador - Emigrantes totales*. Obtenido de Expansión/ Datos Macro:

<https://datosmacro.expansion.com/demografia/migracion/emigracion/ecuador>

De la fuente, S. (2022).

<https://www.estadistica.net/ECONOMETRIA/CUALITATIVAS/LOGISTICA/regresion-logistica.pdf>. Obtenido de

<https://www.estadistica.net/ECONOMETRIA/CUALITATIVAS/LOGISTICA/regresion-logistica.pdf>

Ecuador en cifras . (2020). https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Migracion/2020/Metodologia_ESI2020.pdf.

Obtenido de https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Migracion/2020/Metodologia_ESI2020.pdf

Ecuador en cifras. (04 de 2022).

https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Migracion/2021/Documentos_ESI_DIES/Historia_de_ESI_2021.pdf. Obtenido de

https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Migracion/2021/Documentos_ESI_DIES/Historia_de_ESI_2021.pdf

Estadística DMA. (2023). https://estadistica-dma.ulpgc.es/doctoMed/Tarea_3_-_Descripcion_de_datos.html. Obtenido de [https://estadistica-](https://estadistica-dma.ulpgc.es/doctoMed/Tarea_3_-_Descripcion_de_datos.html)

[dma.ulpgc.es/doctoMed/Tarea_3_-_Descripcion_de_datos.html](https://estadistica-dma.ulpgc.es/doctoMed/Tarea_3_-_Descripcion_de_datos.html)

Estamatica. (20 de 11 de 2020). <https://estamatica.net/tipos-de-regresion-en-estadistica/>. Obtenido de <https://estamatica.net/tipos-de-regresion-en-estadistica/>

Fonseca, C. (22-25 de junio de 2002). Asociación de variables culturales y económicas con la Migración. Una aplicación de modelos de regresión Logit. *[Resultados del Congreso]*. Obtenido de [http://editorial.unca.edu.ar/Publicacione%20on%20line/CD%20INTERACTIVO S/NOA2002/Asociacion%20Variables%20Culturales.pdf](http://editorial.unca.edu.ar/Publicacione%20on%20line/CD%20INTERACTIVO%20S/NOA2002/Asociacion%20Variables%20Culturales.pdf)

Gutiérrez, E., & Vladimirovna, O. (2016). *Estadística inferencial para ingeniería y ciencias*. Grupo Editorial Patria, S.A. de C.V. Obtenido de <http://biblioteca.univalle.edu.ni/files/original/4bee2ce5589a0b8ae82ed363b2bac6206dd28ab1.pdf>

Herrera, G., Carrillo, M., & Torres, A. (2005). *La migración ecuatoriana transnacionalismo, redes e identidad*. FLACSO. Obtenido de https://flacsoandes.edu.ec/web/imagesFTP/6417.migracion_ecuatoriana_transnacionalismo__redes_e_identidades.pdf

Hosmer, D., Lemeshow, S., & Sturdivant, R. (2013). *Applied logistic regression*. John Wiley & Sons.

IBM. (13 de 09 de 2022). <https://www.ibm.com/docs/es/spss-statistics/saas?topic=statistics-cox-regression-analysis>. Obtenido de <https://www.ibm.com/docs/es/spss-statistics/saas?topic=statistics-cox-regression-analysis>

INEC. (7 de julio de 2020). *Entradas y salidas Internacionales 2019*. Obtenido de Instituto Nacional de Estadísticas y Censos: <https://www.ecuadorencifras.gob.ec/entradas-y-salidas-internacionales-2019/>

INEC. (2022). https://www.ecuadorencifras.gob.ec/documentos/web-inec/Normativas%20Estadisticas/Normas_Tecnicas/Codigo_de_buenas_practicass_estadist.pdf. Obtenido de <https://www.ecuadorencifras.gob.ec/documentos/web->

inec/Normativas%20Estadisticas/Normas_Tecnicas/Codigo_de_buenas_practicas
_estadist.pdf

INEC Boletín REESI. (2022).

[https://www.ecuadorencifras.gob.ec/documentos/web-](https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Migracion/2021/Documentos_ESI_CGTPE/Bolet%C3%ADn%20T%C3%A9cnico%20ESI%202021.pdf)

[inec/Poblacion_y_Demografia/Migracion/2021/Documentos_ESI_CGTPE/Bolet](https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Migracion/2021/Documentos_ESI_CGTPE/Bolet%C3%ADn%20T%C3%A9cnico%20ESI%202021.pdf)

[%C3%ADn%20T%C3%A9cnico%20ESI%202021.pdf](https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Migracion/2021/Documentos_ESI_CGTPE/Bolet%C3%ADn%20T%C3%A9cnico%20ESI%202021.pdf). Obtenido de

[https://www.ecuadorencifras.gob.ec/documentos/web-](https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Migracion/2021/Documentos_ESI_CGTPE/Bolet%C3%ADn%20T%C3%A9cnico%20ESI%202021.pdf)

[inec/Poblacion_y_Demografia/Migracion/2021/Documentos_ESI_CGTPE/Bolet](https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Migracion/2021/Documentos_ESI_CGTPE/Bolet%C3%ADn%20T%C3%A9cnico%20ESI%202021.pdf)

[%C3%ADn%20T%C3%A9cnico%20ESI%202021.pdf](https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Migracion/2021/Documentos_ESI_CGTPE/Bolet%C3%ADn%20T%C3%A9cnico%20ESI%202021.pdf)

INEC Entradas y Salidas Internacionales. (2021).

<https://www.ecuadorencifras.gob.ec/entradas-y-salidas-internacionales/>.

Obtenido de [https://www.ecuadorencifras.gob.ec/entradas-y-salidas-](https://www.ecuadorencifras.gob.ec/entradas-y-salidas-internacionales/)

[internacionales/](https://www.ecuadorencifras.gob.ec/entradas-y-salidas-internacionales/)

LLinás, H. (14 de 07 de 2022). <https://rstudio-pubs-static.s3.amazonaws.com>.

Obtenido de [https://rstudio-pubs-](https://rstudio-pubs-static.s3.amazonaws.com/779662_23264cc991994522a8e0f2e1ad997cce.html)

[static.s3.amazonaws.com/779662_23264cc991994522a8e0f2e1ad997cce.html](https://rstudio-pubs-static.s3.amazonaws.com/779662_23264cc991994522a8e0f2e1ad997cce.html)

López, C. (2014). Tablas de contingencia y modelos de regresión logística.

Población y salud en Mesoamérica, 12(1), 2-31. Obtenido de

[https://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S1659-](https://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S1659-02012014000100002)

[02012014000100002](https://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S1659-02012014000100002)

López, P., & Fachelli, S. (2015). *Metodología de la investigación social*

cuantitativa. Barcelona: Universidad Autónoma de Barcelona. Obtenido de

https://ddd.uab.cat/pub/caplli/2016/163570/metinvsoccua_a2016_cap3-10.pdf

Lucenti, K., & Mendes, J. (11 de 06 de 2019). <https://blogs.iadb.org>. Obtenido de

<https://blogs.iadb.org/integracion-comercio/es/eficientes-puertos/>

Martínez, B., & Delgado, D. (09 de 10 de 2020). [*https://rstudio-pubs-static.s3.amazonaws.com/672776_b38a580b68564498b524317f6601fe9e.html#\(1\)*](https://rstudio-pubs-static.s3.amazonaws.com/672776_b38a580b68564498b524317f6601fe9e.html#(1))). Obtenido de [*https://rstudio-pubs-static.s3.amazonaws.com/672776_b38a580b68564498b524317f6601fe9e.html#\(1\)*](https://rstudio-pubs-static.s3.amazonaws.com/672776_b38a580b68564498b524317f6601fe9e.html#(1)))

Moscote, O. (18 de 02 de 2022). [*https://rpubs.com/orlandoan/868570*](https://rpubs.com/orlandoan/868570). Obtenido de [*https://rpubs.com/orlandoan/868570*](https://rpubs.com/orlandoan/868570)

Niño, R. (2022). [*https://masteres.ugr.es/estadistica-aplicada/sites/master/moea/public/inline-files/TFM_NINO%20ALFARO.pdf*](https://masteres.ugr.es/estadistica-aplicada/sites/master/moea/public/inline-files/TFM_NINO%20ALFARO.pdf). Obtenido de [*https://masteres.ugr.es/estadistica-aplicada/sites/master/moea/public/inline-files/TFM_NINO%20ALFARO.pdf*](https://masteres.ugr.es/estadistica-aplicada/sites/master/moea/public/inline-files/TFM_NINO%20ALFARO.pdf)

Niño, V. (2011). *Metodología de la Investigación* (Primera Edición ed.). Bogotá: Ediciones de la U. Obtenido de [*https://download.e-bookshelf.de/download/0003/5946/06/L-G-0003594606-0006935685.pdf*](https://download.e-bookshelf.de/download/0003/5946/06/L-G-0003594606-0006935685.pdf)

R-coder. (2023). [*https://r-coder.com/fijar-semilla-r/*](https://r-coder.com/fijar-semilla-r/). Obtenido de [*https://r-coder.com/fijar-semilla-r/*](https://r-coder.com/fijar-semilla-r/)

Recuero, P. (13 de 12 de 2021). [*https://empresas.blogthinkbig.com/como-interpretar-la-matriz-de-confusion-ejemplo-practico/*](https://empresas.blogthinkbig.com/como-interpretar-la-matriz-de-confusion-ejemplo-practico/). Obtenido de [*https://empresas.blogthinkbig.com/como-interpretar-la-matriz-de-confusion-ejemplo-practico/*](https://empresas.blogthinkbig.com/como-interpretar-la-matriz-de-confusion-ejemplo-practico/)

Sánchez, R. (2021). Programas en R. *Matemática avanzada*, 35-42.

Sánchez, R. (2023). [*https://rsanchezs.gitbooks.io/ciencia-de-datos-con-r/content/estructuras_control/condicionales/estructuras_condicionales.html*](https://rsanchezs.gitbooks.io/ciencia-de-datos-con-r/content/estructuras_control/condicionales/estructuras_condicionales.html). Obtenido de [*https://rsanchezs.gitbooks.io/ciencia-de-datos-con-r/content/estructuras_control/condicionales/estructuras_condicionales.html*](https://rsanchezs.gitbooks.io/ciencia-de-datos-con-r/content/estructuras_control/condicionales/estructuras_condicionales.html)

Scott, S. (2016). Multinomial logistic regression. *Wiley Interdisciplinary Reviews: Computational Statistics*, 74(4), 240-247. doi:10.1002/wics.1354

Sdelsol. (15 de julio de 2020). *Análisis de regresión*. Obtenido de Software DELSOL: <https://www.sdelsol.com/glosario/analisis-de-regresion/>

Troya, M. (2007). Las rutas del tráfico ilegal de migrantes. Labor de la marina ecuatoriana. *Revista de Derecho FORO UASB*, 225-250. Obtenido de <https://repositorio.uasb.edu.ec/bitstream/10644/1636/1/RF-07-TC-Troya.pdf>

Vargas, L. (17 de 11 de 2022). *Larepublica.net*. Obtenido de <https://www.larepublica.net/noticia/puertos-y-aeropuertos-son-esenciales-para-la-reactivacion-economica-y-el-desarrollo-regional>

Villalba, F. (2017). *Manual de R para el científico de datos*. Murcia. Obtenido de <https://drive.google.com/file/d/1EoLm-rqr5eikmpodb90uIGyju6E1jBjZ/view?usp=sharing>