

Guía de uso de base de datos de las Estadísticas de Transporte (ESTRA)

Septiembre, 2022

Dirección/Departamento

Dirección de Estadísticas Económicas (DECON)

Unidad

Gestión de Estadísticas Estructurales (GESE)

Elaborado por:

Gabriela Hidalgo

Revisado por:

Roberto Chaves

Aprobado por:

Diana Barco

Contacto:

inec@inec.gob.ec

www.ecuadorencifras.gob.ec

(02) 2234 164 · (02) 2235 890 · (02) 2526 072

ÍNDICE DE CONTENIDO

Guía de uso de base de datos de las Estadísticas de Transporte (ESTRA)- 2021 .	4
1. Introducción	4
2. Datos generales de la operación estadística	4
3. Descripción de las bases de datos	5
4. Cuestionarios o instrumentos de captura	6
5. Anonimización de la Base de datos	6
6. Variables.....	6
6.1 Variables derivadas.....	8
7. Tratamiento de valores perdidos, validaciones e imputaciones	9
8. Clasificaciones y Nomenclaturas	10
9. Interpretación de indicadores principales.....	11
10. Recomendaciones para el usuario.....	12

ÍNDICE DE TABLAS

Tabla 1 Información de la operación estadística	4
Tabla 2 Descripción de las bases de datos	5
Tabla 3. Lista de principales variables.....	7
Tabla 4. Variables derivadas	9

ÍNDICE DE FIGURAS

Figura 1 Formato de ficha metodológica de indicadores	11
Figura 2 Sección de interpretación del indicador en la ficha metodológica....	12

Guía de uso de base de datos de las Estadísticas de Transporte (ESTRA)- 2021

1. Introducción

El Instituto Nacional de Estadística y Censos (INEC), desarrolla las Estadísticas de Transporte (ESTRA), que es una operación estadística que procesa registros administrativos (RA) para mostrar información relevante sobre: matriculación vehicular, siniestros de tránsito, transporte por vía aérea y marítimo a nivel nacional; información que es importante a la hora de planificar y organizar los sistemas de transporte del país, por parte de las autoridades competentes y de los diferentes actores públicos y privados.

En este marco, el Anuario presenta información de las siguientes entidades gubernamentales:

- Agencia Nacional de Tránsito (ANT), que proporciona información sobre vehículos motorizados matriculados y siniestros de tránsito;
- Dirección General de Aviación Civil (DGAC), que remite la información de tráfico aéreo (Pasajeros y Carga en TM); y,
- Subsecretaría de Puertos y Transporte Marítimo y Fluvial (SPTMF), que entregan información de transporte marítimo y fluvial (Pasajeros y Carga en TM).

En este documento se detallan aspectos a considerar al momento de usar Bases de Datos de las Estadísticas de Transporte (ESTRA), con la finalidad de que el usuario pueda usar de manera óptima la información. El documento hace referencia a datos generales de la operación estadística, descripción de base de datos, instrumentos de recolección, entre otros temas de interés.

Para la descarga de la base de datos e información relevante de las Estadísticas de Transporte (ESTRA), ingresar al siguiente enlace: <https://www.ecuadorencifras.gob.ec/transporte/>

2. Datos generales de la operación estadística

A continuación, se presenta los datos generales del Anuario de Estadísticas de Transporte 2020:

Tabla 1 Información de la operación estadística

Criterio	Descripción
Nombre de la operación	Estadísticas de Transporte (ESTRA)
Población Objetivo	<ul style="list-style-type: none"> • Transporte Terrestre: Vehículos motorizados matriculados y Siniestros de tránsito. • Transporte Aéreo: Pasajeros y carga transportada por vía aérea.

Criterio	Descripción
	<ul style="list-style-type: none"> Transporte Marítimo: Naves arribadas, pasajeros transportados por vía marítima y carga transportada por vía marítima.
Unidad de observación	<ul style="list-style-type: none"> Vehículos motorizados matriculados: vehículos. Siniestros de tránsito: personas. Transporte aéreo: personas y carga. Transporte marítimo y fluvial: Naves, personas y carga.
Cobertura geográfica	<ul style="list-style-type: none"> Vehículos motorizados matriculados: Nacional, regional y provincial. Transporte aéreo: Por principales aeropuertos a nivel nacional (Quito, Guayaquil, Latacunga, Manta). Transporte marítimo y fluvial: Por puertos públicos y privados nacionales (Esmeraldas, Guayaquil, Manta, Bolívar, Balao, La Libertad, El Salitral, Asotep, Astiesmar). Siniestros de Tránsito: Nacional, regional, provincial
Tipo de muestreo	N/A
Tamaño de la muestra	N/A
Desagregación de la información	Las Estadísticas de Transporte ofrecen información sobre el número de vehículos matriculados, carga y pasajeros transportados por vía aérea y marítima, así como, siniestros de tránsito a nivel nacional.
Comparabilidad de la serie histórica	Representación de los resultados en formato tabla, mediante la combinación de variables del último período y datos históricos periodo 2008 - 2021

Fuente: Instituto Nacional de Estadística y Censos

3. Descripción de las bases de datos¹

La operación estadística ESTRA, dispone de 3 bases de datos, en la siguiente tabla, se detalla las principales características de las bases de datos:

Tabla 2 Descripción de las bases de datos

Nombre de la Base de Datos	Criterio	Descripción
Vehículos Matriculados	Número de BDD	1
	Formato de la BDD	La base de datos se encuentra en formato EXCEL (.csv) y Spss (.sav).
	Número de registros	2.535.853
	Número de variables (columnas)	10
	Unidad de análisis de la BDD	Vehículos Matriculados
	Bases de datos con las que se puede realizar combinación o cruce de variables (match)	Siniestros de tránsito
	Variables necesarias para realizar combinación o cruce de variables (match)	N/A
Siniestros de Tránsito	Número de BDD	1
	Formato de la BDD	La base de datos se encuentra en formato EXCEL (.csv) y Spss (.sav).
	Número de registros	21.352

¹ No incluye base de Transporte Aéreo, los datos son remitidos a través del Bolefín de

Nombre de la Base de Datos	Criterio	Descripción	
	Número de variables (columnas)	11	
	Unidad de análisis de la BDD	Siniestros de Tránsito, personas lesionadas y fallecidas.	
	Bases de datos con las que se puede realizar combinación o cruce de variables (match)	Vehículos Matriculados	
	Variables necesarias para realizar combinación o cruce de variables (match)	N/A	
	Número de BDD	1	
	Formato de la BDD	La base de datos se encuentra en formato EXCEL (.csv) y Spss (.sav).	
	Número de registros	3.580	
	Número de variables (columnas)	32	
	Transporte Marítimo	Unidad de análisis de la BDD	Naves, pasajeros y carga transportada
		Bases de datos con las que se puede realizar combinación o cruce de variables (match)	N/A
Variables necesarias para realizar combinación o cruce de variables (match)		N/A	

Fuente: Instituto Nacional de Estadística y Censos

4. Cuestionarios o instrumentos de captura

ESTRA es una operación estadística basada en registros administrativos entregados por varias instituciones gubernamentales de acuerdo a sus competencias, como: ANT, DGAC y SPTMF.

5. Anonimización de la Base de datos

De acuerdo al principio de confidencialidad estadística o secreto estadístico establecido en el Código de Buenas Prácticas Estadísticas del Ecuador, se garantiza la confidencialidad de la información estadística y resguardo de datos de carácter personal de las unidades de observación.

Las bases de datos que conforman las Estadísticas de Transporte son sometidas al proceso de anonimización, donde se procede con el reemplazo de la variable de identificación directa, así como la eliminación de las variables de identificación auxiliares como: nombres y apellidos, razón social, cédula de identidad, RUC, pasaporte, entre otras. Este proceso permite ocultar la identificación directa del individuo y garantiza la vinculación con otros registros administrativos.

6. Variables

Para el diseño y construcción de las variables, se considera las categorías utilizadas en años anteriores, con la finalidad de estandarizar la información y por ende

realizar análisis comparativos y evolutivos de la información procesada, para mayor información puede visitar la metodología de la operación estadística.

En la siguiente tabla se señalan las variables según base de datos:

Tabla 3. Lista de principales variables

Código de la variable	Nombre de la variable	Definición de la variable	Formato del dato de la variable	Categorías de la variable
Vehículos Motorizados Matriculados				
PROVINCIA	Provincia	Indica la provincia de matriculación de los vehículos que existe en el país	Categórica	1 = "Azuay" 2 = "Bolívar" 3 = "Cañar" ... Resto provincias, según DPA.
MODELO	Modelo/ Año de fabricación	Indica el modelo (año de fabricación) de los vehículos matriculados; es decir, es el año en el que se fabricó el vehículo y salió a la venta.	Categórica	1111= "2003 y anteriores" 2004= "2004" 2005= "2005" 2006= "2006" ... 2020="2020" 2021="2021" 2022="2022" ²
MARCA	Marca del vehículo	Determina la marca que predomina en el parque automotor del país.	Cadena	-
CLASE	Clase del vehículo	Identifica la clase del vehículo que transita en el país.	Categórica	1= "Automóvil", 2= "Autobús", 3= "Camión", 4= "Camioneta", 5= "Furgoneta", 6= "SUV", 7= "Motocicleta", 8= "Tanquero", 9= "Tráiler", 10= "Volqueta", 11= "Otra Clase"
Transporte Marítimo y Fluvial				
PUERTO	Nombre del puerto	Indica el nombre del puerto en el que las naves entraron y salieron.	Cadena	APE= "Autoridad Portuaria Esmeraldas" APG= "Autoridad Portuaria Guayaquil" APM= "Autoridad Portuaria de Manta" APPB= "Autoridad Portuaria de Puerto Bolívar" SUINBA= "Superintendencia Petrolera de Balao" SUINLI= "Superintendencia Petrolera La Libertad" SUINSA= "Superintendencia Petrolera el Salitral" TPH= "Terminales Portuarios Habilitados"
TRB	Tonelaje de registro bruto	Indica el tonelaje bruto de las naves entradas y salidas.	Numérico	-
TRN	Tonelaje de registro neto	Indica el tonelaje neto de las naves entradas y salidas.	Numérico	-
PASAJ_ENT	Número de pasajeros que entran	Indica los pasajeros entrantes en los diferentes puertos.	Numérico	-

² La categoría de esta variable corresponde al año del modelo de vehículo (versiones) que fue matriculado en el 2021.

Código de la variable	Nombre de la variable	Definición de la variable	Formato del dato de la variable	Categorías de la variable
PASAJ_SAL	Número de pasajeros que salen	Indica los pasajeros salidos en los diferentes puertos.	Numérico	-
STC	Salida total de carga	Indica la salida total de la carga en tonelada métrica	Numérico	-
ETC	Entrada total de carga	Indica la entrada total de la carga en tonelada métrica	Numérico	-
TC	Total de carga	Indica el total de la carga en tonelada métrica	Numérico	-
TRAFICO_C	Tráfico	Indica el tráfico Internacional de las naves	Cadena	-
AGENCIA_NAVI ERA_C	Agencia Naviera	Indica el nombre de la agencia naviera con él se transporta la mercancía.	Cadena	-
Siniestros de Tránsito				
PROVINCIA	Provincia	Indica la provincia en la que ocurrió el siniestro de tránsito	Categórica	1 = "Azúay" 2 = "Bolívar" 3 = "Cañar" ... Resto provincias, según DPA.
CLASE	Clases de siniestros de tránsito	Indica las principales clases de siniestros de tránsito en el país.	Categórica	1= "Atropellos" 2= "Caída Pasajeros" 3= "Choques" 4= "Estrellamientos" 5= "Rozamientos" 6= "Volcamientos" 7= "Pérdida De Pista" 8= "Otros "
CAUSA	Causas de siniestros de tránsito	Indica las causas frecuentes de siniestros de tránsito en el país.	Categórica	1= "Embriaguez O Droga" 2= "Mal rebasamiento e Invadir carril" 3= "Exceso velocidad" 4= "Impericia e Imprudencia del conductor" 5= "Imprudencia del peatón" 6= "Daños mecánicos" 7= "No respeta las señales de tránsito" 8= "Factores climáticos" 9= "Mal estado de la vía" 10= "Otras Causas"
NUM_LESIONADO	Número de lesionados	Indica el número de lesionados en los siniestros de tránsito.	Numérico	-
NUM_FALLECIDO	Número de fallecidos	Indica el número de fallecidos en los siniestros de tránsito.	Numérico	-
TOTAL_VICTIMAS	Víctimas de siniestros de tránsito	Indica el total de víctimas en los siniestros de tránsito.	Numérico	-

Fuente: Instituto Nacional de Estadística y Censos (INEC)

6.1 Variables derivadas

Al generar resultados, se construye variables derivadas a partir de la información de la base proporcionada en los registros administrativos del MTOP y la ANT; estas derivaciones se realizan con la finalidad de dar una fácil interpretación y comprensión de los resultados y cumplir los objetivos del anuario. Las variables derivadas que se procesan son las siguientes:

Tabla 4. Variables derivadas

Nº	Nombre de la BDD	Variables derivadas	Descripción
1	2021_VEHICULOS_MATRICULADOS_BDD	ESTRATONE	Es la agrupación simplificada según rangos de capacidad de carga en toneladas que transporta los vehículos motorizados por vía terrestre.
		ESTRAPASAJE	Comprende la agrupación simplificada según rangos de capacidad de pasajeros que se transportan en vehículos motorizados por vía terrestre.
2	2021_TRANSPORTE MARITIMO_BDD	TCONT	Es la suma de los contenedores con capacidad de carga de 20 y 40 pies.
		STCC	Es la suma de los contenedores de capacidad de carga 20 y 40 pies que registran salidas de carga en el país.
		ETCC	Es la suma de los contenedores de capacidad de carga 20 y 40 pies que registran entradas de carga en el país.
		STC	Es la suma de las salidas del tipo de carga (sólida, líquida, general, contenerizada)
		ETC	Es la suma de las entradas del tipo de carga (sólida, líquida, general, contenerizada)
		TC	Es la suma de las entradas y salidas de carga en los diferentes puertos públicos del país.
		TCONT_20	Es el total de los contenedores tipo 20 (pies) los cuales registran entradas y salidas en los puertos públicos del país.
3	2021_SINIESTROS_DE_TRANSITO_BDD	TCONT_40	Es el total de los contenedores tipo 40 (pies) los cuales registran entradas y salidas en los puertos públicos del país.
		HORA	Son los rangos establecidos en cuanto a la hora del siniestro de tránsito suscitado en el sitio.
		TOTAL_VICTIMAS	Es la suma de personas que se lesionaron y fallecieron en un siniestro de tránsito.

Fuente: Instituto Nacional de Estadística y Censos (INEC)

7. Tratamiento de valores perdidos, validaciones e imputaciones

Validación

Una vez criticada la información de vehículos matriculados, transporte ferroviario, transporte aéreo, transporte marítimo y fluvial y siniestro de tránsito, es sometida a análisis y validación en su totalidad de acuerdo con los siguientes pasos:

- Revisión de consistencia de información
- Derivación de nuevas variables
- Generación y revisión de tabulados preliminares
- Generación y revisión de indicadores
- Generación de archivos preliminares

Imputación

La imputación se considera parte importante del proceso de validación, con el objetivo de presentar un conjunto de datos limpios, completos y coherentes de los RA que conforman esta operación estadística.

El proceso de imputación se aplica específicamente para la sección de Vehículos Matriculados sobre la información proveniente de la ANT, en variables utilizadas para la construcción de los principales resultados, como lo son la capacidad de carga y pasajeros de los automotores. Para ello se observa los registros que durante la validación y procesamiento reflejaron datos faltantes “missing”, valores en cero y datos atípicos; estos, al ser analizados al detalle por cada una de las clases de vehículo, se presume podría haberse tratado a errores en la digitación, al ingresar la información al sistema de matriculación.

En el análisis de resultados de vehículos según su estratificación de capacidad de carga o de pasajeros, existieron inconsistencias que los usuarios podían interpretar, como una mala validación de los registros, poniendo en duda la calidad de la información.

Para esto se adoptó como método de imputación una variante del modelo “Hot-Deck” el cual tiene como objetivo “llenar los registros vacíos (receptores) con información de campos con información completa (donantes), y los datos faltantes se reemplazan a partir de una selección aleatoria de valores observados, lo cual no introduce sesgos en la varianza del estimador... El algoritmo consiste en ubicar registros completos e incompletos, identificar características comunes de donantes y receptores, y decidir los valores que se utilizarán para imputar los datos omitidos. Para la aplicación del procedimiento es fundamental generar agrupaciones que garanticen que la imputación se llevará a cabo entre observaciones con características comunes, y la selección de los donantes se realiza en forma aleatoria evitando que se introduzcan sesgos en el estimador de la varianza.” (CEPAL; Fernando Medina, Marco Galván, 2007)

En síntesis, para este ejercicio de imputación en la data de matriculación, el modelo utiliza como variables de clasificación para construir los estratos, la “marca” y la “clase” de cada vehículo. Al imputar los datos faltantes, se toman otros de similares características en cuanto a la marca y clase del automotor.

Para realizar el proceso de imputación, se utilizó el software estadístico Stata, ya que facilita la aplicación del modelo, al disponer de un paquete específico para este ejercicio; la sintaxis simula cinco imputaciones diferentes para cada una de las variables tratadas, en este caso capacidad de carga y pasajeros.

Como resultado se obtuvo cifras comparables con respecto a años anteriores, con relación a la clasificación de sus vehículos por capacidad de carga y pasajeros.

8. Clasificaciones y Nomenclaturas

Esta operación estadística utiliza el siguiente Clasificador:

División Política Administrativa (DPA 2021): Es un documento que permite clasificar y homologar la nomenclatura correspondiente a las provincias, cantones y parroquias (urbanas y rurales) del país.

Nomenclaturas y clasificaciones usadas	Versión/ revisión	Variables que aplican el clasificador	Nombre de la BDD
Clasificador Geográfico (DPA)	2021	Provincia	2021_VEHICULOS_MATRICULADOS_BDD
		Provincia Cantón	2021_SINIESTROS_DE_TRANSITO_BDD

9. Interpretación de indicadores principales

Los principales indicadores de la operación estadística han sido homologados, tomando en cuenta estándares internacionales y criterios expertos. Las fichas metodológicas permiten generar y replicar indicadores con una misma metodología y sirven de insumo al usuario para su análisis³:

Figura 1 Formato de ficha metodológica de indicadores

FICHA METODOLÓGICA	
NOMBRE DEL INDICADOR	Proporción de carga importada por vía aérea en relación al total de importaciones nacionales no petroleras.
DEFINICIÓN/OBJETIVO	Conocer la proporción de carga importada por vía aérea en relación al total de importaciones nacionales de productos no petroleros.
FÓRMULA DE CÁLCULO	
$PCMA = \left[\frac{CMA}{CMPNP} \right] \cdot 100$	
Dónde:	
PCMA = Proporción de carga importada por vía aérea. CMA = Peso de la carga entrada por vía aérea. CMPNP = Peso de la carga del total de importaciones nacionales de los productos no petroleros.	
DEFINICIÓN DE LAS VARIABLES RELACIONADAS	
Peso de la carga entrada (importada) por vía aérea: Corresponde a la carga que es ingresada en los diferentes aeropuertos del país. Peso de la carga importada de los productos no petroleros: Hace referencia al peso del total de importaciones nacionales no petroleras cuya fuente es el Banco Central del Ecuador.	
METODOLOGÍA DE CÁLCULO	
Para obtener la proporción de la carga importada por vía aérea, se considera la información del peso de carga entrada por vía aérea de enero a diciembre de 2020, dividido para el peso total de las importaciones de productos no petroleros en el mismo periodo, multiplicado por 100.	
LIMITACIONES TÉCNICAS	
El total de las toneladas importadas de los productos no petroleros está registrado como dato provisional según la fuente de información.	
UNIDAD DE MEDIDA O EXPRESIÓN DEL INDICADOR	La unidad de medida es: Porcentaje de carga importada por vía aérea.

INTERPRETACIÓN DEL INDICADOR	Del total de importaciones no petroleras del país, el (x) porcentaje que corresponde a la carga de entrada de productos no petroleros por vía aérea reportada por las autoridades competentes.	
FUENTE DE DATOS	Banco Central del Ecuador (BCE); Dirección General de Aviación Civil (DGAC).	
PERIODICIDAD DEL INDICADOR	Anual.	
DISPONIBILIDAD DE LOS DATOS	2017-2020.	
NIVEL DE DESAGREGACIÓN	GEOGRÁFICO	Nacional.
	GENERAL	Peso de carga entrada.
	OTROS ÁMBITOS	Mensual.
INFORMACIÓN GEO - REFERENCIADA	No aplica.	

Fuente: Instituto Nacional de Estadística y Censos

Las fichas contienen información detallada sobre el indicador, así como su interpretación, por tanto se recomienda revisar el siguiente punto dentro de la ficha del indicador:

³ Ficha del indicador en formato establecido por el INEC.

Figura 2 Sección de interpretación del indicador en la ficha metodológica

INTERPRETACIÓN DEL INDICADOR	Del total de importaciones no petroleras del país, el (x) porcentaje que corresponde a la carga de entrada de productos no petroleros por vía aérea reportada por las autoridades competentes.
-------------------------------------	--

Fuente: Instituto Nacional de Estadística y Censos

Nota: Para mayor información de los indicadores y su forma de cálculo refiérase a los documentos y su sintaxis en el enlace:

<https://www.ecuadorencifras.gob.ec/transporte/>

10. Recomendaciones para el usuario

Para un adecuado manejo de la base de datos, se recomienda:

- Revisar la presente guía y los documentos adjuntos, como la metodología de la operación estadística.
- Descargar la o las bases de datos en el formato que mejor se ajuste a sus herramientas de trabajo.
- Manipular las bases de datos con el software estadístico de su preferencia.
- Identificar preguntas o indicadores que se desea analizar en la base de datos.
- En caso de tener dificultad del manejo de la base de datos mediante un software estadístico, descargue la base en formato CSV y extraer el archivo de texto a un archivo Excel.



INEC | Buenas cifras,
mejores vidas



@ecuadorencifras



@ecuadorencifras



@InecEcuador



t.me/euadorencifras



INEC/Ecuador



INECEcuador



INEC Ecuador