



Reporte de validaciones de BDD,
incluir los reportes de imputación-
ESTRA 2024 realizado



ESTRA, 2024

Julio · 2025

Reporte de validaciones de BDD, incluir los reportes de imputación

ESTRA 2024

Contenido

| | |
|---|----|
| Contenido | 3 |
| Índice de tablas..... | 3 |
| Índice de Tablas | 3 |
| Índice de Anexos | 4 |
| Antecedentes..... | 5 |
| Objetivo | 5 |
| Desarrollo..... | 5 |
| 1. Validación e imputación de datos | 6 |
| 1.1 Vehículos Matriculados..... | 6 |
| 1.2 Siniestros de Tránsito | 13 |
| 1.3 Transporte Aéreo | 16 |
| 1.4 Transporte Marítimo | 16 |
| Conclusiones..... | 21 |
| ANEXOS..... | 23 |

Índice de tablas

| | |
|--|----|
| Tabla 1. Detalle de la cobertura - ESTRA | 6 |
| Tabla 2. Consistencia de la data..... | 7 |
| Tabla 3. Serie histórica variable "Servicio" | 8 |
| Tabla 4. Clasificación de vehículos matriculados por clase | 9 |
| Tabla 5. Cifras de vehículos matriculados de años anteriores..... | 10 |
| Tabla 6. Capacidad de pasajeros | 11 |
| Tabla 7. Capacidad de carga | 11 |
| Tabla 8. Capacidad de pasajeros imputado..... | 12 |
| Tabla 9. Capacidad de carga imputado..... | 13 |
| Tabla 10. Clase del siniestro INEC y ANT | 14 |
| Tabla 11. Causa del siniestro INEC y ANT..... | 15 |
| Tabla 12. Resumen de datos obtenidos de la información recibida, año 2024, antes de sustento de las novedades por parte de puertos. | 17 |
| Tabla 13. Homologación Banderas..... | 18 |
| Tabla 14. Resumen de datos obtenidos de la información recibida, año 2024. 20 | |
| Tabla 14. Resumen de datos obtenidos de la información recibida, año 2024. 20 | |
| Tabla 15. Homologación de Cifras INEC vs SPTMF | 21 |

Índice de Tablas

| | |
|--|----|
| Ilustración 2. Proceso de completitud de datos faltantes- INEC | 8 |
| Ilustración 3. Marine Traffic | 19 |



Índice de Anexos

| | |
|---|----|
| Anexo 1. Informe de mallas de validación..... | 23 |
| Anexo 2. Sintaxis de validación de los componentes de ESTRA..... | 23 |
| Anexo 3. Sintaxis y reporte de imputación de vehículos matriculados | 23 |

Reporte de validaciones de BDD, incluir los reportes de imputación - ESTRA 2024 realizado - INNOVA_EC_C3.

Antecedentes

Las Estadísticas de Transporte - ESTRA, es una operación estadística con periodicidad anual que recoge información de los diferentes tipos de transporte del país mediante el procesamiento de registros administrativos (RA), mismos que se remiten a través de diversas instituciones gubernamentales como: La Agencia Nacional de Tránsito - ANT, Dirección General de Aviación Civil - DGAC, Ministerio de Transporte y Obras Públicas – MTOP, a través de la Subsecretaría de Puertos y Transporte Marítimo y Fluvial – SPTMF.

Objetivo

En el presente informe se detalla el proceso de captación, validación e imputación ejecutadas para las bases de datos de: Vehículos Matriculados, Siniestros de Tránsito, Transporte Marítimo; y, para la información del Boletín Estadístico de Tráfico Aéreo. Esto como parte de la construcción del Anuario de Estadísticas de Transporte, periodo de referencia 2024.

Desarrollo

La etapa de recolección inició en el mes de enero y febrero, una vez se envió el requerimiento de información a las diferentes fuentes de información del anuario; en este contexto a continuación, se detalla el control de cobertura y novedades presentadas para el proceso de recolección de cada una de las secciones de ESTRA, periodo de referencia 2024.

El control de cobertura permite evaluar el grado de cumplimiento de los objetivos en la recolección, tanto a nivel cualitativo como cuantitativo. La meta a conseguir con este control fue recolectar 4 bases del periodo 2024, información que es necesaria para la construcción del anuario de transportes del periodo en referencia. Para llevar a cabo este proceso se considera lo establecidos en las mallas de validación de la OE.

La planificación de recolección contempló el uso de herramientas como: oficios, llamadas telefónicas, correos y, de ser necesario, reuniones con los diferentes equipos técnicos de las instituciones.

Tabla 1. Detalle de la cobertura - ESTRA

| Detalles | | Estrategias de recolección | Ene | Feb | Mar | Abr | May | Jun | Fecha de recolección |
|----------|------------------------|---|----------|-----|-----|-----|-----|-----|----------------------|
| | | | Acciones | | | | | | |
| ANT | Vehículos matriculados | Oficio de solicitud de información | x | | | | | | 28/02/25 |
| | | Llamada | | x | | | | | |
| | | Correo electrónico | | x | | | | | |
| | | Otro (reuniones, conversaciones extraoficiales) | | x | | | | | |
| | Sinistros de Tránsito | Oficio de solicitud de información | x | | | | | | 31/01/25 |
| | | Llamada | | x | | | | | |
| | | Correo electrónico | x | x | | | | | |
| | | Otro (reuniones, conversaciones extraoficiales) | | | | | | | |
| DGAC | Transporte Aéreo | Oficio de solicitud de información | x | | | | | | 28/02/25 |
| | | Correo electrónico | | | | | | | |
| | | Otro (reuniones, etc.) | | | | | | | |
| SPTMF | Transporte marítimo | Oficio de solicitud de información | x | | | | | | 28/02/25 |
| | | Llamada | | x | x | x | | | |
| | | Correo electrónico | | x | x | x | | | |
| | | Otro (reuniones, etc.) | | | | | x | | |

Elaboración: Gestión de Estadísticas Estructurales - GESE

Esto nos permitió evaluar:

- El tiempo transcurrido desde la última llamada/correo a la fuente.
- Determinar tiempo estimado para recolectar la información.
- Complicaciones ocasionadas a la fuente.
- Identificar de forma temprana problemas con la data.

1. Validación e imputación de datos

Una vez recolectadas las bases de datos y los datos del Boletín de Tráfico Aéreo se inició la revisión de los registros administrativos y, para aquellas bases que los necesitan: la imputación de datos mediante el uso de softwares como: SPSS, Stata, R Studio, entre otros.

A continuación, se detalla el proceso de validación e imputación para las BDD por fuente de información:

1.1 Vehículos Matriculados

La información referente a vehículos matriculados es remitida a la institución por la Agencia Nacional de Tránsito ANT el 12 de mayo de 2025 mediante oficio Nro. ANT-DTIC-2025-1747-M.

Como primer paso, se realizó una primera validación de la base para detectar inconsistencias en los datos y si es que existen todas las variables necesarias para continuar con la validación. Cabe resaltar que, previo al envío de la base de datos del 2024, durante el año 2024 se solicitó una base con corte trimestral sobre la cual se remitieron novedades que, para el envío de la base anual, estuvieron sustentas obteniéndose así una base más consistente para trabajar.

Tabla 2. Consistencia de la data

| Variable | Observación primera entrega abril 2025 |
|-------------------------------------|--|
| Placa | Placas duplicadas. |
| Provincia y cantón de domicilio | Variable con datos vacíos, NA, 'Migración', así como datos de otros países como España, entre otros. |
| Provincia y cantón de matriculación | Variable con datos vacíos, NA, 'Migración', así como datos de otros países como España, entre otros. |
| Año de fabricación vehículo | Casos vacíos. |
| Clase | La clasificación de la variable omite furgonetas. |
| Matriculado por primera vez | No reportada en la base fue remitida en archivo aparte. |

Fuente: Base primaria - ANT

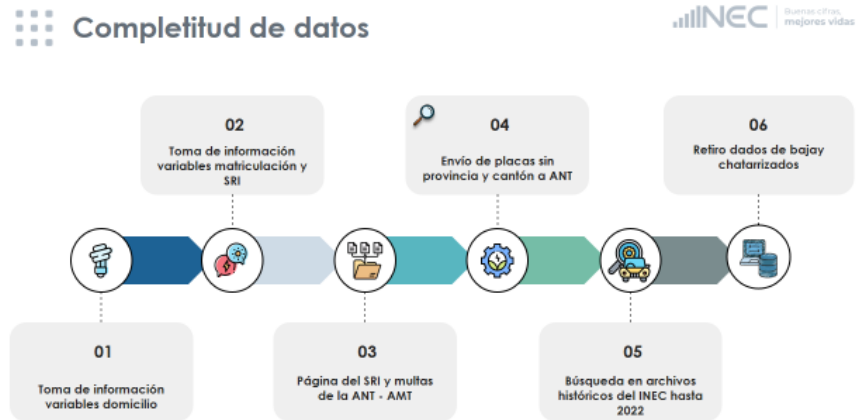
Elaboración: Gestión de Estadísticas Estructurales - GESE

Mediante correos electrónicos, se remite a la ANT, como fuente productora de la información, los ID de casos faltantes de información para que se realice una búsqueda en sus sistemas.

Una vez realizados los análisis de tendencia entre la base de 2024 versus la base de 2023, se identificó variaciones considerables, por lo que se contactó al equipo técnico de la ANT y mediante oficio Nro. INEC-INEC-2025-0240-O del 01 de mayo de 2025 se solicitó nuevas bases de los años 2023 y 2024, especificando que deben ser las completas obtenidas al final del año con el objetivo de que recabe toda la información del año y que recoja posibles casos no considerados en las bases mensualizadas que se remitieron al transcurso del año. Es así que, mediante oficio Memorando Nro. ANT-DTIC-2025-1747-M con fecha 12 de mayo de 2025 la ANT entregó al equipo técnico de INEC las bases completas del año 2023 y 2024.

Por otro lado, se resalta que el equipo técnico del INEC realizó la validación y procesamiento de los datos conforme llegaba la información desde la ANT en las bases mensuales, lo que también ayudó a reducir la carga de trabajo. El procesamiento consistió en completar los datos faltantes con búsquedas específicas realizadas por la ANT, además de los datos históricos internos, es decir, buscar en las bases de años anteriores, así como la búsqueda en páginas web como de la propia ANT y del SRI. Una ayuda valiosa fue la entrega de IDs de la ANT, sobre todo para aquellas nuevas.

Ilustración 1. Proceso de completitud de datos faltantes- INEC



Fuente: Base primaria- ANT
Elaboración: GESE

Como se observa en la ilustración anterior, internamente se llevó a cabo un proceso de búsqueda de datos de aquellos ID de placas que se encontraban faltantes. Cabe mencionar que, fue importante trabajar con bases separadas, una para arreglo de provincia y otra para arreglo de clase, esto por el tamaño de los datos y para evitar confusiones al momento de unificar la base en una sola previo imputación de datos.

Este proceso se replicó en la base nueva del año 2023, sólo para aquellos casos que aparecían como faltantes. Por lo demás no fue necesario realizar dobles procesamientos ya se contaba con lo trabajado el año pasado y fue necesario una migración de resultados. Para aquellos casos nuevos sí fue necesario realizar procesamiento de validación y completitud de la misma. Finalmente, se realizó su imputación de capacidad de pasajeros y peso, paso necesario para obtener resultados limpios.

Para el año 2024, la variable servicio, mantuvo las categorías del año 2023, como se puede apreciar en la serie histórica de la variable:

Tabla 3. Serie histórica variable “Servicio”

| SERVICIO | 2020 | 2021 | 2022 | 2023 | 2024 |
|-----------------------------------|------------------|------------------|------------------|------------------|------------------|
| TOTAL | 2.361.175 | 2.535.853 | 2.880.910 | 3.078.364 | 3.138.562 |
| ESTADO | 32.548 | 284 | 29.980 | 27.427 | 28.211 |
| USO PÚBLICO, ALQUILER O COMERCIAL | 168.862 | 177.297 | 196.426 | 208.359 | 209.615 |
| PARTICULAR | 2.158.933 | 2.358.230 | 2.653.255 | 2.841.514 | 2.899.360 |
| OTROS | 832 | 42 | 1.249 | 1.064 | 1.376 |

Fuente: Base primaria- ANT
Elaboración: GESE

Se evidencia un incremento general en cada categoría de 'servicio' dando continuidad a la serie histórica.

- **Clase de vehículo**

Clasificar los vehículos por su Clase, es la validación más compleja que se lleva a cabo en la base de vehículos matriculados. Esto debido a que la variable de vehículo presenta inconsistencias en los datos ocasionando que 40 mil automotores aproximadamente se clasifiquen como "Otra Clase" siendo incomparable la información con periodos anteriores.

A continuación, el detalle:

Tabla 4. Clasificación de vehículos matriculados por clase

| Clase Agrupada- INEC | Clase Original (ANT) | Nro. Vehículos |
|----------------------|----------------------------|----------------|
| Automóvil | Automóvil | 840.653 |
| Autobús | Autobús | 25.728 |
| | Ómnibus | 1.334 |
| Camión | Camión | 135.823 |
| Camioneta | Camioneta | 553.317 |
| SUV | Jeep | 62.607 |
| | Veh Utilitario | 3.611 |
| | Veh. Utilitario | 11 |
| | Vehículo Utilitario | 536.006 |
| Motocicleta | Motocicleta | 913.933 |
| Tanquero | Tanquero | 81 |
| Tráiler | Tráiler | 1.181 |
| | Uni Carga Y Rem | 67 |
| | Unidad De Carga Y Remolque | 15.699 |
| Volqueta | Volqueta | 673 |
| Otros | Especial | 453 |
| | Otros Tipos De Vehiculos | 6 |
| | Vehiculo Esp | 226 |
| | Vehiculo Especial | 46.819 |
| No registrado | (Vacío) | 333 |
| | Pub | 1 |

Fuente: Base primaria - ANT

Elaboración: Gestión de Estadísticas Estructurales - GESE

Se observa que, dentro de la categoría de clase original de la ANT, no existe la clase 'Furgoneta'. Ante esta situación, se utilizó la variable 'TIPO_VEHICULO' enviada por al ANT realizar combinaciones entre la Clase y Tipo de automotor a fin de determinar con mayor precisión la clase a la que el vehículo corresponde.

Desde años anteriores se ha venido trabajando en la clasificación de vehículos por lo que se cuenta con bases trabajadas desde donde se extrae, por ID, la clase de vehículo para que facilite la validación de los datos. Cabe destacar que, además

de ese procedimiento, se realizaron cruces de Marca, Modelo y tipo de vehículo para comprobar la correcta clasificación, este procedimiento se lo realiza varias veces hasta que se logra reducir las inconsistencias y tener una base limpia para imputación.

Una vez realizado estos procesos, la clasificación vehicular mejora y permite la comparación con periodos anteriores. A continuación, se presenta los datos antes y después de realizar las validaciones:

Tabla 5. Cifras de vehículos matriculados de años anteriores

| CLASE | 2021 (Inicial ANT) | 2021 (Final INEC) | 2022 (Inicial ANT) | 2022 (Final INEC) | 2023 (Inicial ANT) | 2023 (Final INEC) | 2024 (Inicial ANT) | 2024 (Final INEC) |
|---------------|--------------------------|----------------------|-----------------------|----------------------|-----------------------|----------------------|-----------------------|----------------------|
| Total | 2.535.853 | 2.535.853 | 2.880.910 | 2.880.910 | 3.078.354 | 3.078.364 | 3.138.562 | 3.138.562 |
| Automóvil | 747.148 | 746.983 | 754.708 | 812.193 | 853.697 | 857.577 | 840.653 | 846.052 |
| Autobús | 22.161 | 22.182 | 24.659 | 26.853 | 27.018 | 27.312 | 27.062 | 27.684 |
| Camión | 103.305 | 102.061 | 113.465 | 128.492 | 135.821 | 136.763 | 135.823 | 136.670 |
| Camioneta | 467.717 | 423.444 | 452.996 | 487.378 | 555.504 | 500.926 | 553.317 | 501.253 |
| Furgoneta | - | 51.832 | - | 51.913 | - | 60.340 | - | 59.998 |
| Jeep | 441.301 | 436.058 | 440.694 | 516.588 | 573.682 | 578.265 | 602.235 | 607.960 |
| Motocicleta | 722.272 | 722.381 | 606.333 | 815.518 | 871.608 | 872.273 | 913.933 | 914.464 |
| Tanquero | 47 | 2.441 | 22 | 2.801 | 71 | 2.811 | 81 | 2.739 |
| Tráiler | 11.944 | 12.013 | 13.418 | 16.046 | 16.856 | 17.331 | 16.947 | 17.308 |
| Volqueta | 339 | 10.715 | 259 | 13.496 | 583 | 13.154 | 673 | 12.513 |
| Otra clase | 19.619 | 5.743 | 8.152 | 9.632 | 41.488 | 11.612 | 47.504 | 11.921 |
| Sin categoría | - | - | 466.204 | - | 2.026 | - | 334 | - |

Fuente: Base primaria - ANT

Elaboración: Gestión de Estadísticas Estructurales - GESE

• Capacidad de pasajeros y carga

Las variables que registran: capacidad de carga en toneladas y capacidad de pasajeros, registraron datos atípicos; valores en cero, datos faltantes y valores con letras o descripciones varias. Lo cual se vuelve evidente en los cuadros al momento de agrupar los vehículos por grupos de capacidad de carga y de pasajeros y desglosarlos por Clase.

Además, existieron observaciones sobre vehículos tipo Automóviles con capacidad de carga de más de 3.25 ton., Autobuses con menos de 3 ton., así también camionetas, furgonetas y SUV con más de 6 toneladas.

A continuación, se presenta los tabulados en los que se puede observar los datos iniciales obtenidos y los problemas encontrados resaltados con color celeste:

Tabla 6. Capacidad de pasajeros

| Clase | Vehículos matriculados por clase según grupo de capacidad de pasajeros | | | | | |
|--------------|--|----------------|------------------|---------------|---------------|------------------|
| | Total | 1 A 10 | 11 A 20 | 21 A 30 | 31 A 40 | 41 Y MÁS |
| Total | 3.138.123 | 214.892 | 1.003.203 | 95.164 | 22.101 | 1.802.763 |
| Automóvil | 846.015 | 118 | 887 | 274 | 4.255 | 840.481 |
| Autobús | 27.679 | 51 | 1.077 | 3.309 | 6.754 | 16.488 |
| Camión | 136.653 | 52 | 84.682 | 50.000 | 118 | 1.801 |
| Camioneta | 501.225 | 190 | 144.272 | 31.613 | 841 | 324.309 |
| Furgoneta | 59.996 | 381 | 46.219 | 646 | 88 | 12.662 |
| SUV | 607.936 | 77 | 1.135 | 443 | 2.726 | 603.555 |
| Motocicleta | 914.152 | 213.987 | 692.591 | 289 | 6.855 | 430 |
| Tanquero | 2.739 | - | 1.947 | 772 | 5 | 15 |
| Tráiler | 17.305 | 13 | 14.060 | 3.181 | 6 | 45 |
| Volqueta | 12.510 | 4 | 10.065 | 2.397 | 8 | 36 |
| Otra clase | 11.913 | 19 | 6.268 | 2.240 | 445 | 2.941 |

Fuente: Base primaria - ANT

Elaboración: Gestión de Estadísticas Estructurales - GESE

Tabla 7. Capacidad de carga

| Clase | Vehículos matriculados por clase según grupos de capacidad de carga | | | | | |
|--------------|---|------------------|---------------|---------------|---------------|---------------|
| | Total | 1/4 A 3 | 3 1/4 A 6 | 6 1/4 A 10 | 10 1/4 A 15 | 15 1/4 Y MÁS |
| Total | 3.138.221 | 2.963.758 | 74.745 | 36.453 | 29.055 | 34.210 |
| Automóvil | 846.046 | 845.534 | 90 | 11 | 10 | 401 |
| Autobús | 27.684 | 884 | 3.222 | 9.866 | 9.601 | 4.111 |
| Camión | 136.667 | 41.631 | 57.091 | 20.658 | 12.136 | 5.151 |
| Camioneta | 501.243 | 499.066 | 2.014 | 13 | 16 | 134 |
| Furgoneta | 59.997 | 51.741 | 8.219 | 7 | 14 | 16 |
| SUV | 607.945 | 607.383 | 430 | 14 | 3 | 115 |
| Motocicleta | 914.168 | 913.976 | 8 | 1 | 18 | 165 |
| Tanquero | 2.739 | 11 | 663 | 686 | 820 | 559 |
| Tráiler | 17.306 | 20 | 73 | 1.224 | 1.335 | 14.654 |
| Volqueta | 12.512 | 41 | 1.014 | 2.310 | 3.030 | 6.117 |
| Otra clase | 11.914 | 3.471 | 1.921 | 1.663 | 2.072 | 2.787 |

Fuente: Base primaria - ANT

Elaboración: Gestión de Estadísticas Estructurales - GESE

Cabe resaltar que, el equipo técnico del INEC recibió datos desde la ANT de forma periódica, sin embargo, al momento de imputar los datos, estos no eran tan precisos como al realizar una imputación de la base completa. Por tal motivo, las imputaciones se realizaron con una base unificada tal como se describe en el siguiente punto:

Imputación para datos atípicos

Ante esto, el equipo técnico del INEC realizó un ejercicio de imputación para los datos atípicos, valores cero y datos perdidos que existían en la Base de Datos, tanto para la capacidad de pasajeros como para la capacidad de carga, lo que permitió mejorar la clasificación de los vehículos de acuerdo a estas dos variables.

El ejercicio se estableció principalmente en una imputación aleatoria de los datos "missing", basados en el modelo "Hot-Deck", cuyo procedimiento consiste en el reemplazo de los valores perdidos por otro registro ya existente en la muestra para reemplazarlo. Para ello se categorizaron los vehículos por estratos por su "Marca" y "Clase" final.

La finalidad de realizar este proceso se fundamentó en que los valores de las imputaciones, para las variables de interés, sean muy similares a los vehículos dentro de su misma marca y clase lo más homogénea posible.

En este ejercicio se realizaron 5 imputaciones para cada "missing", cero o atípico; para luego calcular la media de las 5 imputaciones generadas; y para el caso de la variable capacidad de pasajeros tomar el redondeo del valor obtenido, puesto que se considera como números enteros para pasajeros y números continuos para la capacidad de carga. Para esta acción se utilizó el software estadístico STATA, basados en el comando "hotdeck".

Una vez realizado el procedimiento antes detallado se obtuvieron los siguientes resultados:

Tabla 8. Capacidad de pasajeros imputado

| Clase | Vehículos matriculados por clase según grupo de capacidad de pasajeros | | | | | |
|--------------|--|------------------|---------------|--------------|--------------|---------------|
| | Total | 1 A 10 | 11 A 20 | 21 A 30 | 31 A 40 | 41 Y MÁS |
| Total | 3.138.562 | 3.068.785 | 42.693 | 2.944 | 6.932 | 17.208 |
| Automóvil | 846.052 | 846.052 | - | - | - | - |
| Autobús | 27.684 | - | 614 | 2.938 | 6.930 | 17.202 |
| Camión | 136.670 | 136.670 | - | - | - | - |
| Camioneta | 501.253 | 501.163 | 90 | - | - | - |
| Furgoneta | 59.998 | 18.166 | 41.832 | - | - | - |
| SUV | 607.960 | 607.927 | 33 | - | - | - |
| Motocicleta | 914.464 | 914.464 | - | - | - | - |
| Tanquero | 2.739 | 2.739 | - | - | - | - |
| Tráiler | 17.308 | 17.308 | - | - | - | - |
| Volqueta | 12.513 | 12.513 | - | - | - | - |
| Otra clase | 11.921 | 11.783 | 124 | 6 | 2 | 6 |

Fuente: Base primaria - ANT

Elaboración: Gestión de Estadísticas Estructurales - GESE

Tabla 9. Capacidad de carga imputado

| Clase | Vehículos matriculados por clase según grupos de capacidad de carga | | | | | |
|--------------|---|-----------|-----------|------------|-------------|--------------|
| | Total | 1/4 A 3 | 3 1/4 A 6 | 6 1/4 A 10 | 10 1/4 A 15 | 15 1/4 Y MÁS |
| Total | 3.138.562 | 2.964.314 | 75.272 | 36.654 | 29.120 | 33.202 |
| Automóvil | 846.052 | 846.052 | - | - | - | - |
| Autobús | 27.684 | 889 | 3.222 | 9.874 | 9.632 | 4.067 |
| Camión | 136.670 | 40.851 | 57.706 | 20.885 | 12.213 | 5.015 |
| Camioneta | 501.253 | 499.239 | 2.014 | - | - | - |
| Furgoneta | 59.998 | 51.776 | 8.222 | - | - | - |
| SUV | 607.960 | 607.526 | 434 | - | - | - |
| Motocicleta | 914.464 | 914.464 | - | - | - | - |
| Tanquero | 2.739 | 11 | 664 | 690 | 821 | 553 |
| Tráiler | 17.308 | - | 73 | 1.225 | 1.339 | 14.671 |
| Volqueta | 12.513 | 31 | 1.014 | 2.312 | 3.042 | 6.114 |
| Otra clase | 11.921 | 3.475 | 1.923 | 1.668 | 2.073 | 2.782 |

Fuente: Base primaria - ANT

Elaboración: Gestión de Estadísticas Estructurales - GESE

Al final del proceso de validación para la Base de Datos de Vehículos Matriculados se obtuvo las variables validadas de acuerdo a lo planificado, cumpliendo así con la meta establecida.

1.2 Siniestros de Tránsito

La información de siniestros de tránsito se solicitó mediante oficio Nro. INEC-INEC-2025-0019-O del 17 de enero de 2025 a la Agencia Nacional de Tránsito. La entidad entregó la información 2024 mediante correo electrónico el 15 de enero de 2025.

Al revisar la base de datos, se presentó dudas en la variable edad de los conductores involucrados, una vez identificado esas novedades, se procede a la validación normal de la base de datos, generalmente es una base que desde ANT es muy bien trabajada y completa.

• Hora

Esta variable, debe ser codificada para generar rangos que permita la rápida interpretación del dato, es así que en lugar de segundos y minutos se cuenta con rangos separados por las 24 horas del día quedando:

- 0 "00:00 A 00:59"
- 1 "01:00 A 01:59"
- 2 "02:00 A 02:59"
- ...
- 23 "23:00 A 23:59".

- **Provincia y cantones**

Las variables referentes a provincia y cantón, son validadas y posteriormente categorizadas de acuerdo al Clasificador Geográfico Estadístico – DPA.

- **Total víctimas**

Esta variable se calcula a partir del número de lesionados y fallecidos en el sitio, la sumatoria de estas dos variables permite la generación del total de víctimas.

- **Clase**

ANT clasifica los siniestros en 13 clases, mientras que el equipo técnico del INEC, para homologar la información presenta 8 clases de siniestros:

Tabla 10. Clase del siniestro INEC y ANT

| Clase INEC | CLASE ANT |
|--------------------|-------------------|
| Estrellamientos | Estrellamientos |
| Otros | Otros |
| Rozamientos | Rozamientos |
| Volcamientos | Volcamientos |
| Atropellos | Arrollamientos |
| | Atropellos |
| Choques | Choque frontal |
| | Choque lateral |
| | Choque posterior |
| | Colisión |
| Pérdida de pista | Pérdida de carril |
| | Pérdida de pista |
| Caída de pasajeros | Caída de pasajero |

Fuente: Base primaria- ANT

Elaboración: Gestión de Estadísticas Estructurales - GESE

- **Causa**

Esta variable describe las causas probables del siniestro, ANT tiene entre 27 y 28 causas las cuales se reportan de acuerdo a lo reportado por los entes de control. La validación de esta variable consiste en organizar esas causas dentro de las 10 categorías que maneja el INEC a fin de tener información homologada.

A continuación, se detalla la clasificación realizada:

Tabla 11. Causa del siniestro INEC y ANT

| CAUSA PROBABLE DEL SINIESTRO - INEC | CAUSA PROBABLE DEL SINIESTRO - ANT |
|--|---|
| Embraguez o droga | Conduce bajo la influencia de alcohol, sustancias estupefacientes o psicotrópicas y/o medicamentos. Peatón transita bajo influencia de alcohol, sustancias estupefacientes o psicotrópicas y/o medicamentos. |
| Mal rebasamiento invadir carril | Adelantar o rebasar a otro vehículo en movimiento en zonas o sitios peligrosos tales como: curvas, puentes, túneles, pendientes, etc. Realizar cambio brusco o indebido de carril. |
| Exceso velocidad | Conducir vehículo superando los límites máximos de velocidad. |
| Impericia e imprudencia del conductor | Conducir desatento a las condiciones de tránsito (celular, pantallas de video, comida, maquillaje o cualquier otro elemento distractor). Conducir en estado de somnolencia o malas condiciones físicas (sueño, cansancio y fatiga). Dejar o recoger pasajeros en lugares no permitidos. Mal estacionado - el conductor que detenga o estacione vehículos en sitios o zonas que entrañen peligro, tales como zona de seguridad, curvas, puentes, túneles, pendientes. No ceder el derecho de vía o preferencia de paso a vehículos. No ceder el derecho de vía o preferencia de paso al peatón. No guardar la distancia lateral mínima de seguridad entre vehículos. No mantener la distancia prudencial con respecto al vehículo que le antecede. Peso y volumen - no cumplir con las normas de seguridad necesarias al transportar cargas. |
| Imprudencia del peatón | Bajarse o subirse de vehículos en movimiento sin tomar las precauciones debidas. No transitar por las aceras o zonas de seguridad destinadas para el efecto. Peatón que cruza la calzada sin respetar la señalización existente (semáforos o señales manuales). |
| Daños mecánicos | Daños mecánicos previsible. Falla mecánica en los sistemas y/o neumáticos (sistema de frenos, dirección, electrónico o mecánico). |
| No respeta las señales de tránsito | Conducir en sentido contrario a la vía normal de circulación. No respetar las señales manuales del agente de tránsito. No respetar las señales reglamentarias de tránsito. (Pare, ceda el paso, luz roja del semáforo, etc.). |
| Factores climáticos | Condiciones ambientales y/o atmosféricas (niebla, neblina, granizo, lluvia). |
| Mal estado de la vía | Malas condiciones de la vía y/o configuración. (Iluminación y diseño). Presencia de agentes externos en la vía (agua, aceite, piedra, lastre, escombros, maderos, etc.). |
| Otras causas | Caso fortuito o fuerza mayor (explosión de neumático nuevo, derrumbe, inundación, caída de puente, árbol, presencia intempestiva e imprevista de semovientes en la vía, etc.). |

Fuente: Base primaria- ANT

Elaboración: Gestión de Estadísticas Estructurales - GESE

En la revisión de la base se identificaron inquietudes sobre la edad de los conductores involucrados, las cuales fueron remitidas a la fuente para su verificación y aclaración de novedades; las novedades fueron solventadas por la institución mediante correo electrónico con fecha 17 de enero de 2025, en el cual se corrige la variable de edad en los casos identificados en la base de datos.

1.3 Transporte Aéreo

Mediante Oficio Nro. INEC-INEC-2025-0021-O del 17 de enero del 2025, se solicitó a la Dirección General de Aviación Civil del Ecuador (DGAC), en calidad de fuente, la base de datos estructurada sobre transporte aéreo nacional e internacional de pasajeros y carga, correspondiente al periodo 01/01/2024 al 31/12/2024.

Mediante Oficio Nro. DGAC-DGAC-2025-0218-O del 26 de febrero del 2025, la DGAC remitió la información relacionada con transporte aéreo correspondiente al periodo solicitado por el INEC; como parte de la política de transparencia y acceso a la información pública, la DGAC presentó al INEC el Boletín Estadístico Aerocomercial de Tráfico Aéreo 2024, el cual refleja el comportamiento del tráfico aéreo nacional e internacional en los servicios de transporte aéreo público regular y no regular durante el año 2024 y la variación con el año 2023; así también, contiene la relación de variables de pasajeros, carga vuelos y capacidad ofrecida de las operaciones realizadas en cada uno de los aeropuertos del Ecuador y proyecciones de pasajeros de carga en rutas de mayor demanda.

El documento fue cargado en la plataforma digital (PIMCORE), cuyo acceso requirió credenciales de seguridad, las cuales fueron proporcionadas por la entidad fuente a la responsable de la operación estadística.

El proceso de validación de la información sobre transporte aéreo 2024, inició con la recepción de los registros administrativos remitidos oficialmente por parte de la Dirección General de Aviación Civil del Ecuador (DGAC), la cual se encuentra registrada en formato Excel, clasificada en tablas desde las cuales se generan los datos comparativos para verificar las correctas sumas de datos.

Los últimos dos años ha sido posible el traslado de las mismas a un software de procesamiento que es R, desde donde se realiza la clasificación de las tablas y se las descompone para generar datos internacionales y nacionales.

1.4 Transporte Marítimo

Mediante oficio Nro. INEC-INEC-2025-0020-O con fecha 17 de enero de 2025, el INEC realiza la solicitud de información 2024 a la Subsecretaría de Puertos – SPTMF. Esta entidad debe recolectar la información de 4 Autoridades Portuarias, 3 Superintendencias Petroleras y de 9 puertos privados del país.

La información se recibe mediante oficio Nro. MTOP-SPTM-25-227-OF del 06 de marzo de 2025 conforme el formato de matriz de recolección solicitada por el INEC. A esta respuesta se suman las matrices de información de los puertos privados participantes de la Asociación de Terminales Portuarios Privados del Ecuador-ASOTEP correspondiente a Terminales Portuarios habilitados. Las matrices fueron remitidas mediante enlace de acceso: <https://almacenamiento.mtop.gob.ec/index.php/s/HiS88Cn5SRbZ23X>.

En un primer análisis para determinar la consistencia de la información se genera un cuadro resumen de los datos para ver el panorama. Esto va de la mano de una

revisión exhaustiva de cada matriz detectando inconsistencias para ser remitidas a la entidad encargada para que ésta remita a cada puerto. A continuación, se detalla el resumen de los primeros datos obtenidos:

Tabla 12. Resumen de datos obtenidos de la información recibida, año 2024, antes del sustento de las novedades por parte de puertos.

| INEC - TRÁFICO INTERNACIONAL DE CARGA (EN TONELADAS MÉTRICAS) | | | | | | |
|--|---------------|-----------------|----------------------|----------------------|------------------------------|-----------------------------|
| Nro. | Descripción | Número de naves | Importación | Exportación | Número de pasajeros entrados | Número de pasajeros salidos |
| 1 | APE | 26 | 43.327,00 | 65.314,00 | | |
| 2 | APG | 1.061 | 4.827.801,90 | 5.248.645,02 | 1.993 | 1.993 |
| 3 | APM | 515 | 1.028.648,66 | 6.599,38 | 12.056 | 12.056 |
| 4 | APPB | 315 | 82.089,36 | 2.200.958,19 | | |
| 5 | SUINBA | 369 | 2.353.248,00 | 23.075.400,00 | | |
| 6 | SUINLI | 298 | 5.165.395,92 | | | |
| 7 | SUINSA | 54 | 972.897,63 | | | |
| 8 | ASTIESMAR C.A | 12 | 2.109,24 | | | |
| 9 | BANANAPUERTO | 322 | 1.051.530,24 | 2.920.505,77 | | |
| 10 | ECUABULK | 50 | 594.056,41 | | | |
| 11 | FERTIGRAM | 336 | 3.265.767,48 | 3.413.330,70 | | |
| 12 | FERTISA | 129 | 1.638.967,94 | 308.989,02 | | |
| 13 | INDUMOL | 2 | 3134701 | | | |
| 14 | NIRSA S.A. | 2 | 7.109,48 | | | |
| 15 | QCTE | 58 | 310.606,00 | 23.885,00 | | |
| 16 | STOREOCEAN | 27 | 269.043,87 | | | |
| Total | | 3.489 | 21.261.225,58 | 37.208.265,84 | 14.049 | 14.049 |

Fuente: Base primaria

Elaboración: Gestión de Estadísticas Estructurales - GESE

Una vez se tuvo la respuesta a las novedades remitidas por puerto, se inició la validación de los registros administrativos para convertirlos en datos estadísticos. Entre las principales validaciones tenemos:

- **Puerto**

Se homologan los nombres de los puertos ya que en ocasiones se colocan solamente la abreviación del puerto y otros el nombre completo, de esa forma la validación se vuelve más llevadera ya que la variable se utiliza para filtros más adelante en el proceso.

- 'APE' 'AUTORIDAD PORTUARIA DE ESMERALDAS'
- 'APG' 'AUTORIDAD PORTUARIA DE GUAYAQUIL'
- 'APM' 'AUTORIDAD PORTUARIA DE MANTA'
- 'APPB' 'AUTORIDAD PORTUARIA DE PUERTO BOLIVAR'
- 'SUINBA' 'SUPERINTENDENCIA DEL TERMINAL PETROLERO DE BALAO'
- 'SUINLI' 'SUPERINTENDENCIA DEL TERMINAL PETROLERO DE LA LIBERTAD'
- 'SUINSA' 'SUPERINTENDENCIA DEL TERMINAL PETROLERO DE EL SALITRAL'.

- 'ASTIESMAR C.A'
- 'BANANAPUERTO'
- TPG-FERTIGRAN'
- FERTISA'
- 'QCTE'
- 'STOREOCEAN S.A'
- 'ECUABULK S.A.'
- 'INDUMOL'
- 'NIRSA S.A.'

Cabe resaltar que, en la base de datos, se consideran los nombres de los puertos privados, sin embargo, en tabulados se homologan a TPH 'TERMINALES PORTUARIOS HABILITADOS'

- **Tráfico**

ESTRA publica cifras de transporte marítimo de los puertos respecto al tráfico internacional. En ese sentido, se valida el tráfico en función de si es nacional o de cabotaje o internacional y se homologa en dos categorías: Nacional e Internacional.

Una vez homologado, se procede a crear una base donde sólo se disponga de datos internacionales para continuar con la validación de esos datos, sin embargo, esto no significa que los datos nacionales no sean revisados, al contrario, las novedades son remitidas sobre toda la base sin distinción de tráfico.

- **Bandera**

La bandera de procedencia de la nave es descrita generalmente con un gentilicio como alemana o cubana. Para este paso se realiza una investigación de países y zonas:

Tabla 13. Homologación Banderas

| ENVIADO POR PUERTOS | CATEGORIZADO INEC |
|----------------------|---------------------|
| Holandesa | Países bajos |
| Inglesa | Inglaterra |
| Italiana | Italia |
| Japonesa | Japón |
| Netherlands | Países bajos |
| Netherlands antilles | Antillas holandesas |
| Países bajos | Países bajos |
| Ecuatoriana | Ecuador |

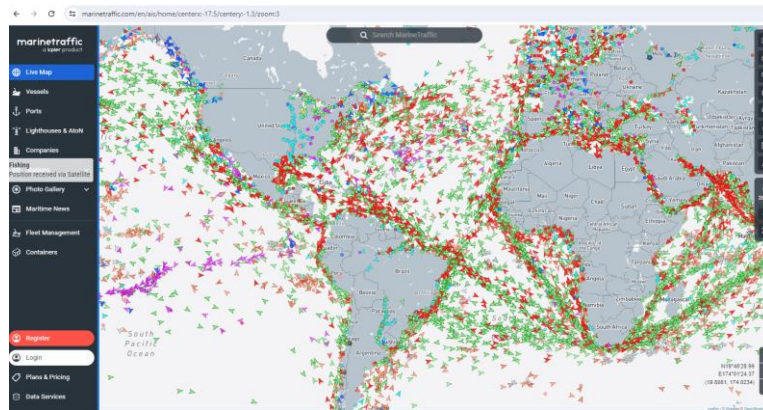
Fuente: Base primaria

Elaboración: Gestión de Estadísticas Estructurales - GESE

Como se observa, se coloca el país de procedencia, así mismo se trata, en lo posible de categorizar de acuerdo con la vigencia actual, por ejemplo Holanda, que en realidad son Países Bajos. A pesar de ello categoriza banderas como británica, inglesa o Reino Unido, es complicado ya que Inglaterra es un país que forma parte de UK.

En otros casos, se presentaron datos en los cuales se consideraba el nombre de la nave en lugar de asignar la bandera correspondiente, en este sentido, se empleó el portal web de Marine Traffic, cuyo enlace de acceso utilizado es el siguiente: <https://www.marinetraffic.com/en/ais/home/centerx:-17.5/centery:-1.3/zoom:3>

Ilustración 2. Marine Traffic



Fuente: www.marinetraffic.com

- **Pasajeros**

Los pasajeros entrados y salidos deben coincidir con el dato generado en cada caso, es decir los entrados deben ser igual a los salidos, por tanto, en el caso de la información correspondiente al 2024, se verifica que existe un solo dato en el cual se registra número de pasajeros, tanto entrados como salidos, dato que concuerda con la información proporcionada por la fuente. Es el caso de APM, que registra 12.056 y APG con 1.993 pasajeros, dato validado conforme la comparación de ambas variables, las cuales coinciden con la información preliminar y validada.

- **Verificación de sumas**

Se verificó que las variables que asumen la totalidad de carga tanto entrada como salida, coincidan con los datos que componen dicha variable, para ello se verifica la variable que crea el total de carga entrada y salida, dado que en algunos casos se considera la carga de transbordo, la cual no conforma la sumatoria de la carga total.

- **Otros**

En el caso de omisión de información como: IDBUQUE, IDREGISTRO, NÚMERO DE MATÍCULA, TRÁFICO, BANDERA, AGENCIA NAVIERA, se procedió con la actualización de sintaxis en la herramienta de validación con la finalidad de actualizar la información preliminar conforme las novedades presentadas.

Las novedades de todos los puertos, autoridades portuarias y superintendencias se remitieron vía correo electrónico a la Subsecretaría de Puertos, quienes remitieron las observaciones a los responsables durante el mes de marzo, abril y mayo, a través de correo electrónico.

Tabla 14. Resumen de datos obtenidos de la información recibida, año 2024.

| INEC - TRÁFICO INTERNACIONAL DE CARGA (EN TONELADAS MÉTRICAS) | | | | | | |
|--|---------------|--------------|----------------------|----------------------|------------------------------|-----------------------------|
| Nº | Descripción | Nº naves | Importación | Exportación | Número de pasajeros entrados | Número de pasajeros salidos |
| 1 | APE | 26 | 43.327,00 | 65.314,00 | | |
| 2 | APG | 1.061 | 4.827.801,90 | 5.248.645,02 | 1.993 | 1.993 |
| 3 | APM | 515 | 1.028.648,66 | 6.599,38 | 12.056 | 12.056 |
| 4 | APPB | 315 | 78.782,36 | 2.200.958,19 | | |
| 5 | SUINBA | 369 | 2.353.248,36 | 23.075.400,43 | | |
| 6 | SUINLI | 298 | 5.165.395,92 | | | |
| 7 | SUINSA | 54 | 972.897,63 | | | |
| 8 | ASTIESMAR C.A | 12 | 571,27 | | | |
| 9 | BANANAPUERTO | 322 | 673.646,39 | 2.865.125,39 | | |
| 10 | ECUABULK | 50 | 594.056,41 | | | |
| 11 | FERTIGRAM | 336 | 3.265.767,48 | 3.413.330,70 | | |
| 12 | FERTISA | 129 | 1.638.967,95 | 308.989,02 | | |
| 13 | INDUMOL | 2 | 31.347,01 | | | |
| 14 | NIRSA S.A | 2 | 7.109,48 | | | |
| 15 | QCTE | 58 | 310.613,89 | 23.885,71 | | |
| 16 | STOREOCEAN | 27 | 269.043,87 | | | |
| Total | | 3.489 | 21.261.225,58 | 37.208.265,84 | 14.049 | 14.049 |

Fuente: Base primaria
Elaboración: GESE

Comparación de datos: INEC vs Subsecretaría de Puertos

Mediante correo electrónico, con fecha 29 de abril del presente año, se remite el Boletín de Estadísticas 2024 por parte de la Subsecretaría de Puertos y Transporte Marítimo y Fluvial.

Una vez la Subsecretaría de Puertos remitió al INEC las cifras presentadas a su institución por parte de los puertos para el año 2024, se procedió con la revisión y comparación de las mismas.

Esto arrojó algunas novedades en los puertos, las cuales se resumen en la siguiente tabla:

Tabla 15. Resumen de datos obtenidos de la información recibida, año 2024.

| Puerto | Número de naves | | | Carga Entrada | | | Carga Salida | | |
|---------------------|-----------------|-------|------|---------------|--------------|------------|---------------|---------------|-----------|
| | INEC | SPTMF | DIF. | INEC | SPTMF | DIF. | INEC | SPTMF | DIF. |
| SUINBA | 369 | 369 | - | 2.353.248,00 | 2.353.248,36 | 0,36 | 23.075.400,00 | 23.075.400,43 | 0,43 |
| ASTIESMAR | 10 | 10 | - | 2.109,24 | 571,27 | 1.537,97 | - | - | - |
| BANANAPUERTO | 322 | 322 | - | 1.051.530,12 | 673.646,39 | 377.883,85 | 2.920.505,77 | 2.865.125,39 | 55.308,38 |
| QCTE | 58 | 58 | - | 310.606,00 | 310.613,89 | 7,89 | 23.885,00 | 23.885,71 | 0,71 |

| Puerto | Número de naves | | | Carga Entrada | | | Carga Salida | | |
|--------|-----------------|-------|------|---------------|--------------|----------|--------------|--------------|-------------|
| | INEC | SPTMF | DIF. | INEC | SPTMF | DIF. | INEC | SPTMF | DIF. |
| APM | 440 | 440 | - | 1.028.197,83 | 1.028.648,66 | -450,83 | - | - | - |
| APPB | 315 | 315 | - | - | - | - | 4.8572,41 | 1.031.896,64 | -983.324,23 |
| SUINSA | 54 | 54 | - | 972.897,63 | 969.897,63 | 3.000,00 | | | |

Fuente: Base primaria

Elaboración: GESE

Homologación de cifras entre el INEC y la SPTMF

Las novedades de todos los puertos, autoridades portuarias y superintendencias se remitieron vía correo electrónico a la Subsecretaría de Puertos, quienes remitieron las observaciones a los responsables durante el mes de abril y mayo.

Con fecha 7 de mayo de 2025, se realiza la reunión con la Subsecretaría de Puertos y Transporte Marítimo y Fluvial, en la cual, la información presentada no se identifica diferencias, por lo que se acepta las cifras homologadas entre las instituciones obteniendo las siguientes cifras finales.

Tabla 16. Homologación de Cifras INEC vs SPTMF

| Descripción | | INEC | SPTMF | DIFERENCIA |
|-------------------------------|---------------|----------------------|----------------------|------------|
| Número de naves | | 3.489 | 3.489 | |
| Carga Entrada | Contenerizada | 5.909.556,21 | 12.149.246,53 | - |
| | Carga Sólida | 5.351.719,20 | 826.149,01 | - |
| | Carga Líquida | 8.767.866,50 | 23.144.747,21 | - |
| | Carga General | 1.232.083,67 | 1.088.123,10 | - |
| Total de carga Entrada | | 21.261.225,58 | 37.208.265,84 | - |
| Carga Salida | Contenerizada | 5.909.556,21 | 12.149.246,53 | - |
| | Carga Sólida | 5.351.719,21 | 826.149,01 | - |
| | Carga Líquida | 8.767.866,50 | 23.144.747,21 | - |
| | Carga General | 1.232.083,67 | 1.088.123,10 | - |
| Total de carga Salida | | 21.261.225,58 | 37.208.265,84 | - |

Fuente: Base preliminar-SPTMF

Elaboración: GESE

Conclusiones

- El proceso de validación ha permitido verificar la consistencia de la información para la generación de resultados preliminares, mediante el uso de herramientas estadísticas tales como EXCEL, SPSS, STATA y R Studio.
- Se identificó novedades en transporte marítimo y se realizaron las debidas notificaciones a los puertos, quienes presentaron información con inconsistencias referentes a fechas incorrectas, registros sin datos en algunas variables y errores de digitación. Además, a través de la corrección de novedades, se homologó los datos de entre las dos instituciones: INEC y SPTMF.

- El resultado de la validación generó observaciones y novedades relacionadas con los datos enviados por las diferentes fuentes; en cada caso se establecieron comunicaciones permanentes con la fuente con la finalidad de realizar las actualizaciones correspondientes a los datos.
- Las observaciones fueron oportunamente solventadas por la fuente y actualizadas en los formatos establecidos y requeridos por el INEC. Sin embargo, por su complejidad tanto vehículos matriculados como transporte marítimo requirieron mayor tiempo para responder y justificar las observaciones.

Quito DM., 22de julio de 2025

| Elaborado por: | Revisado por: | | Aprobado por: |
|--------------------------------------|---|--------------------------------------|-----------------------------------|
| | | | |
| Kevin Cordova (Miembro de Equipo) | Gabriela Hidalgo (Responsable ESTRA) | Roberto Chaves (Responsable GESE) | Diana Barco (Director de Área) |

ANEXOS

Anexo 1. Informe de mallas de validación



Anexo 2. Sintaxis de validación de los componentes de ESTRA



SINTAXIS_SINIESTROS VALIDACION.R



Validación_Aereo.r



Validacion_Maritimo.R



sintaxis de unificación y trabajo nueva 2024.R



Validacion sintaxis 1.R

Anexo 3. Sintaxis y reporte de imputación de vehículos matriculados



Sint_vm24.do

```

# R Script for Vehicle Registration Data Processing
# Author: INEC
# Date: 2024-06-01
# Version: 1.0

# Load required libraries
library(tidyverse)
library(readr)
library(dplyr)
library(stringr)

# Define file paths
input_file = "data/vehiculos_2024.csv"
output_file = "reporte_vehiculos_2024.csv"

# Read input data
vehiculos = read_csv(input_file)

# Data cleaning and processing
vehiculos = vehiculos %>%
  # Remove rows with missing values in key columns
  filter(!is.na(Placa), !is.na(Provincia), !is.na(Ciudad)) %>%
  # Standardize province names
  mutate(Provincia = tolower(Provincia)) %>%
  # Add a unique identifier for each vehicle
  mutate(Identificador = paste0(Provincia, Placa))

# Write output data
write_csv(vehiculos, output_file)

# Print a message
cat("Reporte de vehículos matriculados generado exitosamente.\n")

```




Buenas cifras,
mejores vidas



@InecEcuador



@ecuadorencifras



@ecuadorencifras



INECEcuador