República del Ecuador

Proceso y metodología de validación del Anuario de Estadísticas de Transporte - ANET

Diciembre, 2021

www.ecuadorencifras.gob.ec

INEC Buenas cifras, mejores vidas

Proceso y metodología de validación del Anuario de Estadísticas de Transporte – ANET

1. Introducción

En vista de la necesidad de información básica, actualizada y completa que le permita al gobierno de Ecuador planear una adecuada política de desarrollo y organizar adecuadamente el sistema de transportación a nivel nacional; el Anuario de Estadísticas de Transporte - ANET, se ha realizado desde el año de 1966 al año de 1974 y de 1976 hasta la actualidad; su primera edición se publicó bajo el nombre Estadística de Transporte 1965 – 1966.

El ANET tiene como objetivo presentar información a partir de registros administrativos (RA) de diferentes instituciones, a fin de ofrecer una amplia descripción del sector del transporte en el país en ámbitos como: número de vehículos matriculados, pasajeros transportados por ferrocarril, carga y pasajeros transportados por vía aérea y marítima, así como, siniestros de tránsito a nivel nacional cuyas fuentes de información son las siguientes entidades gubernamentales:

- La Agencia Nacional de Tránsito - ANT
- Empresa Ferrocarriles del Ecuador Empresa Pública - FEEP
- Dirección General de Aviación Civil – DGAC
- Subsecretaría de Puertos y Transporte Marítimo y Fluvial - SPTMF

En este contexto, existe la necesidad de definir efectivamente los pasos para el procesamiento de información de los diferentes componentes del anuario para la optimización de tiempos y solucionar posibles problemas que se presenten. Esto mediante un proceso de validación y verificación de la información; y obtener bases de datos consistentes para la generación de resultados.

2. Objetivo

Diseñar un documento de procesos y establecer la metodología de las distintas etapas de validación a las cuales debe ser sometida la información recibida para la construcción del ANET.

3. Etapas de validación

Las etapas de validación y verificación a las cuales debe ser sometida la información recibida, se encuentran detallados en la fase de recolección, procesamiento y análisis de los datos de a las actividades del cronograma. Una vez que la información es recibida se realiza la planificación de la validación de acuerdo al orden de llegada de la data de los diferentes componentes.

4. Recolección

La actividad previa al procesamiento estadístico es la transferencia y recolección de datos desde las fuentes de registros administrativos. Para esto, se establecen escenarios de transferencia (identificación de herramientas para entrega -

www.ecuadorencifras.gob.ec

recepción de información y actores que intervienen) que permiten el acceso oportuno a los datos y se preserve los principios de confidencialidad, integridad y disponibilidad de la información.

Tabla 1. Calendarización de solicitud y entrega de información por fuentes

Fuentes de información		Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Vehículos Matriculados	Solicita												
	Recibe												
Transporte Ferroviario	Solicita												
	Recibe												
Transporte Aéreo	Solicita												
	Recibe												
Transporte Marítimo y fluvial	Solicita												
	Recibe												
Siniestros de tránsito	Solicita												
	Recibe												

Fuente: Instituto Nacional de Estadística y Censos (INEC)

4.1 Crítica e integración de la base de datos

Para la crítica del ANET 2020 se considera las siguientes acciones:

- Revisión de formatos establecidos para la entrega de la información.
- Revisión del número de variables requeridas.
- Revisión de datos omitidos.

La integración de las bases de datos se realiza de manera homologada en función de la información entregada por cada una de las fuentes de información:

Tabla 2. Proceso de integración de bases de datos ANET

Fuentes de Información	Proceso de integración de bases de datos[1]
Vehículos Matriculados	Es la integración de las bases de datos entregadas en formato XLS y CSV para luego exportar en SPSS.
Transporte Ferroviario	Es la compilación de tabulados entregados en formato XLS y CSV para luego consolidar la información y exportar en SPSS.
Transporte Marítimo y fluvial	Es la integración de tabulados entregados en formato XLS para luego exportar en SPSS.
Siniestros de tránsito	Es la integración de bases de datos entregados en formato XLS y CSV para luego consolidar la información y exportar en SPSS.

Fuente: Instituto Nacional de Estadística y Censos (INEC)

Cabe señalar que la información presentada en el ANET sobre Transporte Aéreo no dispone de bases de datos como tal, sino que se construye en base al Boletín Estadístico de Tráfico Aéreo de la DGAC.

[1] Las estadísticas de vía aérea provienen del boletín estadístico que publica la DGAC más no entregan bases de datos.

4.2 Clasificaciones y/o codificaciones

Para la clasificación y codificación de las variables suministradas en los distintos registros administrativos, se realiza y/o actualiza las sintaxis de codificación (en SPSS o STATA), a fin de garantizar homologación y consistencia (con las publicaciones de años anteriores) en la asignación de las categorías requeridas para la presentación de resultados, por ejemplo: clase de vehículos, tipo de cargas, otras.

En esta etapa se utiliza el **Clasificador Geográfico Estadístico 2020**, especialmente para las variables de tipo geográfico como: provincia y cantón.

A continuación, se describen las actividades realizadas para la generación y revisión de las bases de datos por componente.

Buenas cifras, mejores vidas

01.

Vehículos Matriculados

Proceso y metodología de las distintas etapas de validación de Vehículos Matriculados

1. **Ordenar la BDD e identificar casos duplicados**

Ordenar los casos por placa de forma ascendente.

PLACA	IDENTIFICACION	NOMBRE_PROPIE
A001669	702344565	GONZALEZ VASQUEZ PRO
A001794	105280309	BARRERA SUAREZ MAURI
A002382	1707335244	ESPINOZA HEREDIA JOSE
A003086	917920019	GARCIA VERA JOSE VITER
A004070	301156261	ZHANGALLIMBAY CHICAIZA
A004111	105813661	PADILLA BALAREZO LUIS A
A004370	925091514	JARA PESANTES CRISTHIA
A004474	927344499	SALGADO CAMEJO JEFFE

Una vez ordenada la base de datos se procede a identificar casos duplicados en la base mediante la aplicación de sintaxis; en la cual se genera el reporte.

1 Duplicado

0 Registros únicos.

Estadísticos

testd Registros duplicados

N	Válido	2361175
	Perdidos	0

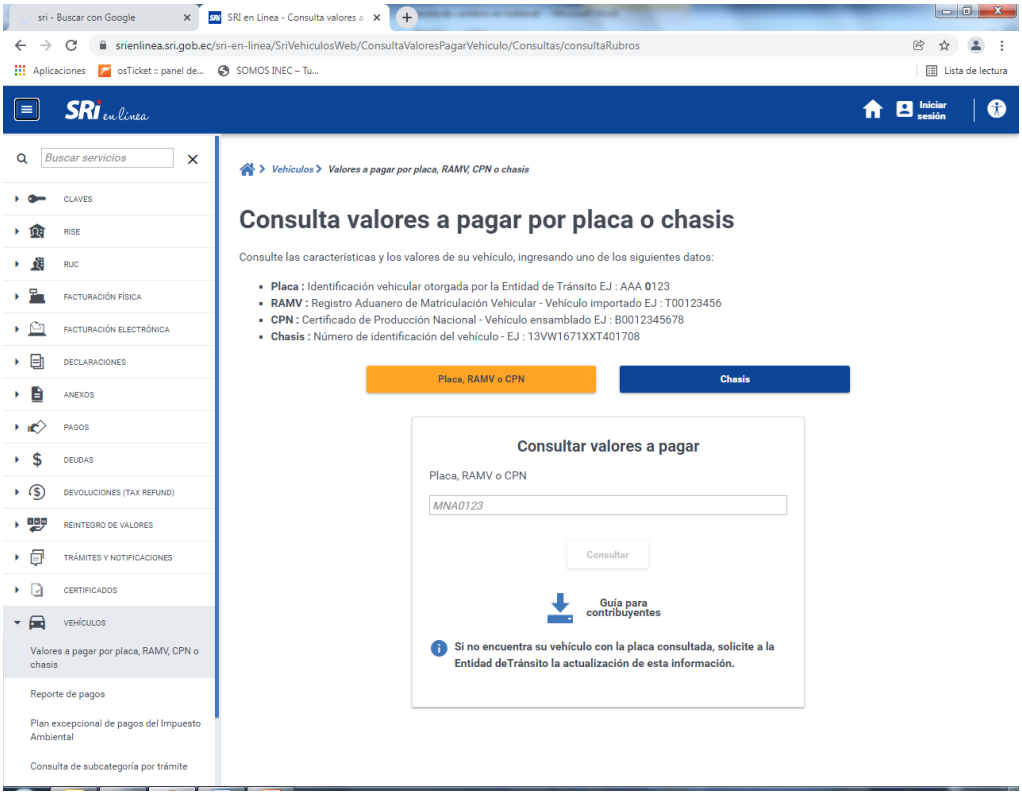
-
- En caso de presentar casos de duplicados se procede a verificar en la base, y si, la base contiene la misma información en todas las variables con el mismo identificador; una vez verificado, se realiza la depuración de los datos.

2. **Para la bdd 2020, no existieron casos duplicados, por lo que se considera una base apta para iniciar con el proceso de validación. Creación de variables.**

- Se crea la variable ***aniofab*** basándose en la variable original **año**. Algunos casos no tienen la información sobre el año de fabricación del vehículo. Para corregir esos casos y completar faltantes, se debe utilizar la matrícula e investigar en el sistema del SRI para saber el año correcto, así como la página de la ANT:

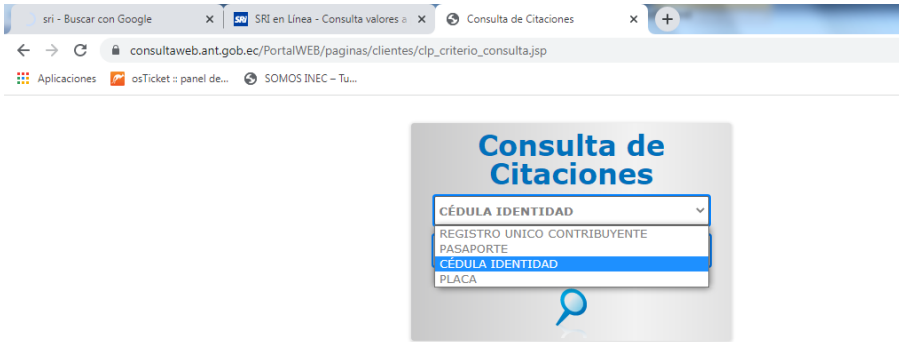
  ✓ https://srienlinea.sri.gob.ec/sri-en-linea/SriVehiculosWeb/ConsultaValoresPagarVehiculo/Consultas/consultaRubros

En la página se accede a información por Placa, RAMV o CPN:



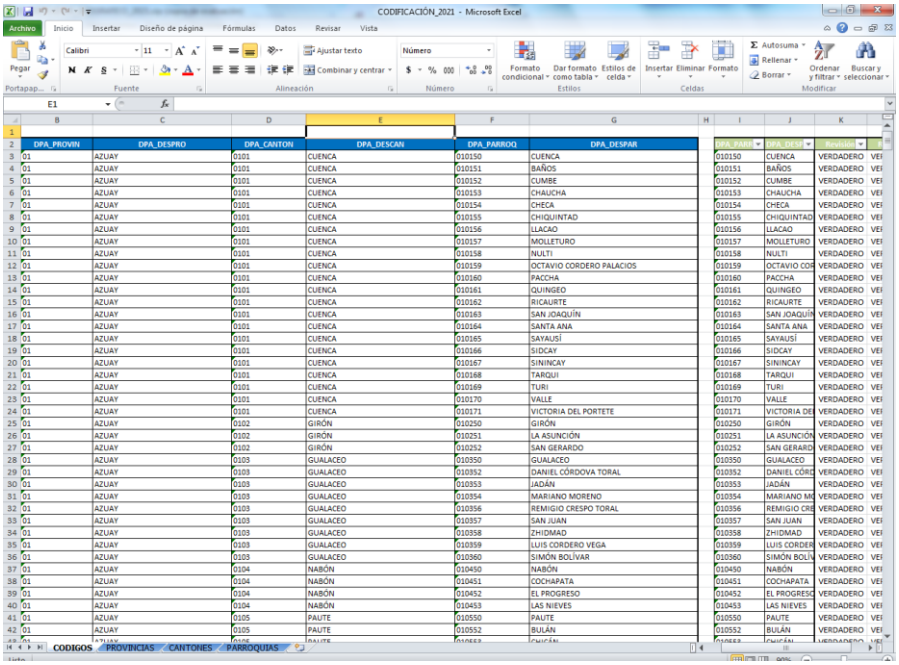
✓ https://consultaweb.ant.gob.ec/PortalWEB/paginas/clientes/clp_criterio_consulta.jsp

En la página de consulta de multas se puede acceder a los datos del vehículo por RUC, pasaporte, cedula de identidad o placa.



- Crear la variable ***prov_dom*** recodificando la variable original usando el codificador geográfico. Para datos missing colocamos 99 SIN REGISTRO.
- Crear la variable ***canton_dom*** recodificando la variable original usando el codificador geográfico. Para datos missing colocamos 99 SIN REGISTRO.
- Crear la variable ***prov_matricu*** recodificando la variable original usando el codificador geográfico. Para datos missing colocamos 99 SIN REGISTRO.
- Crear la variable ***canton_matricu*** recodificando la variable original usando el codificador geográfico. Para datos missing colocamos 99 SIN REGISTRO.

✓ Para la creación y codificación de esas variables se trabaja con el Clasificador Geográfico Estadístico – DPA



DPA_PROVIN	DPA_DESPRO	DPA_CANTON	DPA_DESCAN	DPA_PARROQ	DPA_DESPAR
01	AZUAY	0101	CUENCA	010150	CUENCA
01	AZUAY	0101	CUENCA	010151	BAÑOS
01	AZUAY	0101	CUENCA	010152	CUMBE
01	AZUAY	0101	CUENCA	010153	CHAUCHA
01	AZUAY	0101	CUENCA	010154	CHECA
01	AZUAY	0101	CUENCA	010155	CHIQUINTAD
01	AZUAY	0101	CUENCA	010156	LLACAO
01	AZUAY	0101	CUENCA	010157	MOLLETURO
01	AZUAY	0101	CUENCA	010158	NULTI
01	AZUAY	0101	CUENCA	010159	OCTAVIO CORDERO PALACIOS
01	AZUAY	0101	CUENCA	010160	PACCHA
01	AZUAY	0101	CUENCA	010161	QUINGEO
01	AZUAY	0101	CUENCA	010162	RICAURTE
01	AZUAY	0101	CUENCA	010163	SAN JOAQUÍN
01	AZUAY	0101	CUENCA	010164	SANTA ANA
01	AZUAY	0101	CUENCA	010165	SAYAUSÍ
01	AZUAY	0101	CUENCA	010166	SIDCAY
01	AZUAY	0101	CUENCA	010167	SININCAY
01	AZUAY	0101	CUENCA	010168	TARQUI
01	AZUAY	0101	CUENCA	010169	TURI
01	AZUAY	0101	CUENCA	010170	VALLE
01	AZUAY	0101	CUENCA	010171	VICTORIA DEL PORTETE
01	AZUAY	0102	GIRÓN	010250	GIRÓN
01	AZUAY	0102	GIRÓN	010251	LA ASUNCIÓN
01	AZUAY	0102	GIRÓN	010252	SAN GERARDO
01	AZUAY	0103	GUALACEO	010350	GUALACEO
01	AZUAY	0103	GUALACEO	010352	DANIEL CÓRDOVA TORAL
01	AZUAY	0103	GUALACEO	010353	JADÁN
01	AZUAY	0103	GUALACEO	010354	MARIANO MORENO
01	AZUAY	0103	GUALACEO	010356	REMIGIO CRESPO TORAL
01	AZUAY	0103	GUALACEO	010357	SAN JUAN
01	AZUAY	0103	GUALACEO	010358	ZHIDMAD
01	AZUAY	0103	GUALACEO	010359	LUIS CORDERO VEGA
01	AZUAY	0103	GUALACEO	010360	SIMÓN BOLÍVAR
01	AZUAY	0104	NABÓN	010450	NABÓN
01	AZUAY	0104	NABÓN	010451	COCHAPATA
01	AZUAY	0104	NABÓN	010452	EL PROGRESO
01	AZUAY	0104	NABÓN	010453	LAS NIEVES
01	AZUAY	0105	PAUTE	010550	PAUTE
01	AZUAY	0105	PAUTE	010552	BULÁN

La codificación utilizada por el INEC, se encuentra estandarizada con el Registro Civil y demás entidades del Sector Público que manejan un sistema con las siguientes características:

- Se asignan los dos primeros códigos de la izquierda al código de provincia.
- El código 01 es asignado al cantón, cuya cabecera es también capital provincial.
- Dentro de cada cantón se codifican las parroquias urbanas en orden alfabético desde 01 hasta 49, asignando el 50 a la cabecera cantonal.
- A las parroquias rurales se les ordena alfabéticamente y se asigna el código respectivo del 51 al 99.

Las provincias permanecen codificadas en orden alfabético, a excepción de Galápagos, Sucumbíos, Orellana, Santo Domingo de los Tsáchilas y Santa Elena, creadas en fechas posteriores al establecimiento de este criterio de codificación; a éstas se les asigna el código inmediato superior al que tiene la última provincia, de acuerdo a la fecha de su respectiva creación.

- Crear la variable ***servicio*** recodificando la variable original. Se consolidan en las categorías:

  1 ESTADO
  2 USO PÚBLICO, ALQUILER O COMERCIAL
  3 PARTICULAR
  4 OTROS.

- Para los valores missings, se investiga el modelo del vehículo en las páginas mencionadas y se asigna un uso. Para casos, donde no se pueda realizar por ese medio, se toma la información de bases de años anteriores sí se dispone.

- Crear la variable ***tipo_combustible*** recodificando los datos de la variable. De tal manera que se dispondrá de las siguientes categorías:

  1 DIESEL
  2 GASOLINA
  3 HÍBRIDO
  4 ELÉCTRICO
  5 GAS LICUADO DE PETROLEO
  6 OTRO
  99 NO REPORTADO.

3. Validación de la variable clase de vehículo

Para la creación de la clase de vehículo, partir de las variables validadas en años anteriores. En base a la placa, se toma la clase final del 2019, luego la de 2018 teniendo así: clase2019_3 y case 2018_3.

NOTA: Considerar; desde este punto, trabajar con una base exclusiva para la clase, que incluya variables de validación como placa, marca, modelo, identificación, vin, motor, color y carga y pasajeros.

Todo esto se trabaja con las sintaxis en SPSS donde también existen notas.

Crear la variable clase_2020 partiendo de la clase original de la base. Teniendo estas 3 variables, procedemos a identificar los casos que deberían entrar a revisión y ser considerados para la reclasificación.

Crear variable clase2020_2, tomando los valores de la variable clase2019_3, luego los caos que queden vacíos se toman los valores de la clase2018_3, y posterior a esto, para los casos restantes que quedaron sin categoría, tomar los valores de la variable clase2020. Esto ayuda a disminuir los casos sobre los cuales se debe trabajar y categorizar.

Así habrán quedado pocos casos para trabajar en ellos y asignar categorías.

Al momento de categorizar adecuadamente los vehículos, en lugar de ***DO IF***, que toma a todos los vehículos y los categoría automáticamente, prefiérase crear un filtro para verificar que, por ejemplo, en automóvil que se encuentre dentro de la: clase2020_3=1 & pasajerof>10 esté categorizándose de forma correcta.

Para ello es preferible generar un documento de Excel y revisar caso por caso si es correcto. Sí suele suceder que uno o dos casos están mal categorizados, asignados de marca errónea, y por ejemplo, el automóvil en realidad era una moto o camioneta. Lo mismo para buses y otras clases de vehículos.

NOTA: Si clase2018_3, clase2019_3, y clase_2020 coinciden la categoría está correcta, sin embargo sucede que por x motivo la categorización pudo ser errónea. Por lo que si ninguno de los años coincide, ese caso también debe ser tomado en cuenta para revisar.

El siguiente paso es generar segmentos de los casos en los que se trabajarán para realizar validaciones por caso o casos, dependiendo de la marca y el modelo.

Identificar los casos para asignar clase y abrir un documento auxiliar de sintaxis en Excel.

PLACA	marca2020	TIPO VEHICULO	MODELO_COMPLETO	clase2020_3	pasajerof Capacidad de pasajero	tonelajefn	nueva clase	sintaxis
AAU0240	MERCEDES BE	CAMION	OF 1721/52	3	40,00	7,00	2	IF (PLACA='AAU0240' AND marca2020='marca2020' AND clase2020_3=3) clase2020
CAE0007	HINO	CAMION	GH1JMSA	3	42,00	12,00	2	IF (PLACA='CAE0007' AND marca2020='MERCEDES BENZ' AND clase2020_3=3) cl
CAE0140	CHEVROLET	BUS	FTR32M CHASIS TORPEDO	3	40,00	8,00	2	IF (PLACA='CAE0140' AND marca2020='HINO' AND clase2020_3=3) clase2020_3=2.
EAA2321	HINO	BUS	AK8JRSA 7.7 3P 4X2 TM DIESEL	3	40,00	7,00	2	IF (PLACA='EAA2321' AND marca2020='CHEVROLET' AND clase2020_3=3) clase20
EAG0574	CHEVROLET	BUS	FTR32M CHASIS TORPEDO	3	40,00	8,00	2	IF (PLACA='EAG0574' AND marca2020='HINO' AND clase2020_3=3) clase2020_3=2

Considerar validar los casos usando: PLACA, MARCA, TIPO DE VEHICULO, MODELO CLASE2020_3, PASAJEROS Y CARGA. Al momento de asignar el nuevo valor de clase, estas variables ayudan a determinar de mejor manera el vehículo.

Para sintaxis, preferir realizar el cambio desde PLACA y aplicando filtros sobre estos, ya que usando marca modelo y clase, tienden a realizar cambios sobre los vehículos no considerados para el cambio y que están asignados correctamente.

Cada vez que se validen la clase del vehículo y se corra la sintaxis de cambios, se debe generar varias veces el tabulado 8 que es marca de vehículos por clase, este sirve para identificar los casos donde se debe poner especial cuidado, porque existen marcas que no fabrican cierto tipo de vehículos.

	clase2020_3										
MARCA2020	1	2	3	4	5	6	7	8	9	10	11
HINO	-	11.817	31.820	-	-	-	-	1.110	1.150	4.244	1.502
JMC	-	4	1.354	273	-	16	-	7	-	11	12
KING LONG	-	74	2	-	351	-	-	-	-	-	-

Sin embargo, hay casos de vehículos muy antiguos que cierta marca dejo de producir pero que en el país siguen circulando.

También existen marca que están iniciando en la fabricación, de por ejemplo, motocicletas y antes no lo hacía. O de SUV cuando antes sólo tenían camiones. En esos casos, las páginas oficiales de las marcas son de gran ayuda para validar que el vehículo está correcto.

La validación va de la mano con la constante generación de este tabulado y cuando se tengan estos resultados se pasa a la imputación de datos que se basa en la clase del vehículo.

Imputación

4. CAPACIDAD_PASAJEROS Capacidad de pasajeros.

Crear la variable pasajerof partiendo de la variable original.

Antes de reemplazar caracteres extras como comas o espacios, se sugiere retirar mediante la función reemplazar en Excel, letras y espacios, doble comas y otros, ya que reemplazar directamente con sintaxis crea muchos errores que afectan la imputación.

Por ejemplo: 4-5p puede generar una capacidad de pasajeros de 45, cuando en realidad era de entre 4 a 5 pasajeros.

Posterior a estos reemplazos se pueden retirar de forma global los otros errores de digitación.

5. CAPACIDAD_DE_CARGA Capacidad de carga

Crear la variable PESOT partiendo de la variable original.

Antes de reemplazar caracteres extra como comas o espacios, se sugiere retirar mediante la función reemplazar en Excel, letras y espacios, doble comas y otros, ya que reemplazar directamente con sintaxis crea muchos errores que afectan la imputación.

Por ejemplo: 5 0 puede generar un peso de 50 una vez retirado el espacio, lo que en realidad corresponde a 5,0.

CAPACIDAD_PASAJEROS_COMPLETO	CAPACIDAD_DE_CAR...
5,00	,5
5,00	0,375P
5,00	,75
5,00	.75

Posterior a estos reemplazos se pueden retirar de forma global los otros errores de digitación

Para llevar a cabo la imputación, se debe realizar mediante STATA con la sintaxis ya establecida del método de imputación que es una variante del modelo Hot-Deck

En la imputación puede ocurrir que no se corra la sintaxis sobre algunos vehículos, esos deben ser analizados nuevamente y de ser el caso arreglados manualmente para que la imputación no arroje errores.

6. Imputación en STATA

La imputación en STATA es una sintaxis construida por Franklin Tenesaca basándose en el modelo Hot Deck. Lo que se realiza es correr la sintaxis una vez se tiene trabajada la variable clase.

La imputación puede ser realizada varias veces ya que suele arrojar errores y requiere varias revisiones.

Una vez imputados los datos en STATA, se traslada a formato SPSS para ser incorporados en la base donde se trabajó la clase. Unificar estas variables a la base original y generar la base preliminar para correr sintaxis y generar los tabulados.

Este proceso garantiza que tanto la base trabajada como la base original tengan las variables de micro dato con las variables trabajadas finales.

Buenas cifras, mejores vidas

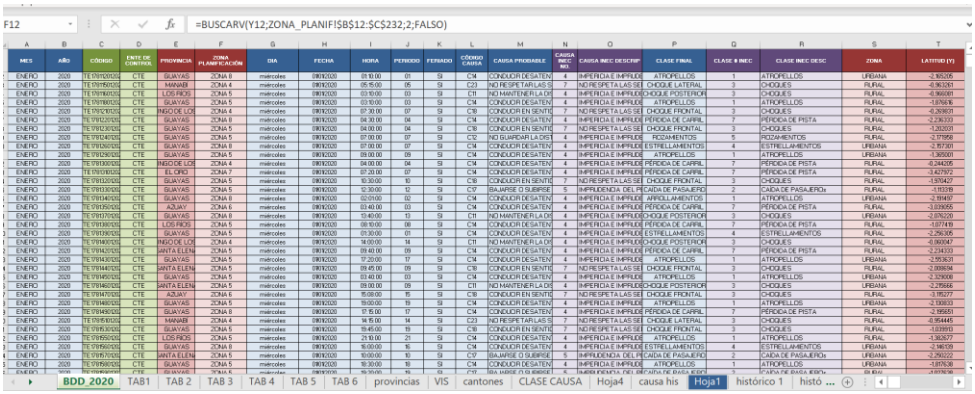
02.

Siniestros de Tránsito

Proceso y metodología de las distintas etapas de validación de Siniestros de Tránsito.

Previo al traslado de la base de Excel a SPSS, se debe realizar algunos trabajos de la información en ese estado. Como por ejemplo, verificar que los totales y casos sean correctos y que coincidan con lo publicado en la página de ANT.

Es importante señalar que se debe replicar los tabulados de la ANT y verificar coincidencia de sumas y totales.



Con las validaciones realizadas se crea el archivo que se enviará a ANT sí existen diferencias en algún dato para ser aclarado y/o corregido.

1. **Creación de variables (SPSS)**

Utilizando la variable MES como base, recodificar dicha variable en una nueva denominada: MES_T

Codificar los entes de control que reportan os casos a la ANT de 1 a 15, esto crea la variable ENTCONTROL. Seguido, asignar etiqueta a la variable.

Recodificar la variable original PROVINCIA para construir la nueva variable: PROVINCIA_T. Esto basado en el Clasificador Geográfico Estadístico – DPA. Seguido, asignar etiqueta a la variable

Se recodifica la variable DIA para generar la nueva variable DIA_T. Seguido, asignar etiqueta a la variable

Usando la variable HORA, se genera la variable HORA_T codificando las horas en rangos: 0 00:00 A 00:59 hasta 23 23:00 A 23:59. Seguido, asignar etiqueta a la variable

Generar la variable FERIADO_T donde se recodifique la variable original FERIADO donde: (SI=1) (NO=2). Seguido, asignar etiqueta a la variable

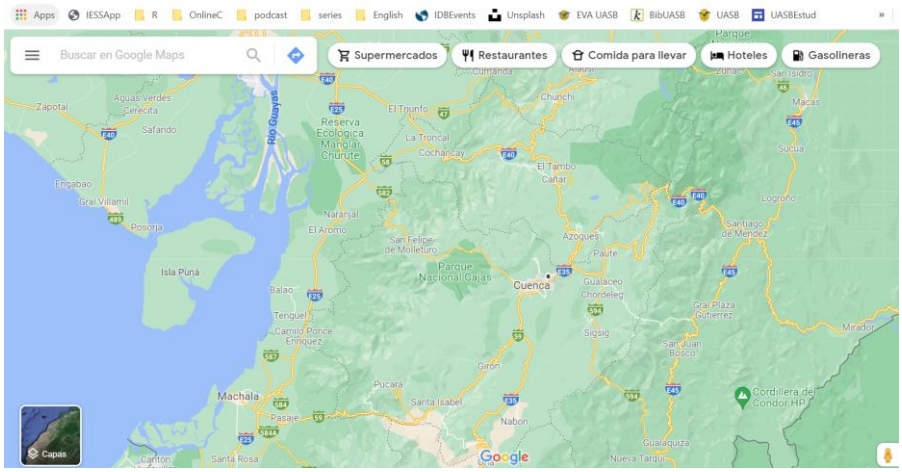
CAUSAPROBABLE: Esta variable posee varias causas probables que reporta la ANT. Se deben categorizar dentro de las causas que presenta INEC de forma homologada. La nueva variable es: CAUSA_T y presenta 10 categorías de las cerca de 27 causas que presenta ANT.

La variable CLASEFINAL tiene diferentes causas que se agrupan en las categorías que se presentan en INEC que son 8 categorías. De ese modo se crea la variable CLASE_T con las siguientes categorías:

1 ATROPELLOS
2 CAÍDA DE PASAJEROS
3 CHOQUES
4 ESTRELLAMIENTOS
5 ROZAMIENTOS
6 VOLCAMIENTOS
7 PÉRDIDA DE PISTA
8 OTROS

Recodificar la variable original ZONA en ZONA_T (URBANA=1) (RURAL=2). Seguido, asignar etiqueta a la variable

La variable LATITUDY es una variable de validación. Sí por alguna circunstancia se necesita determinar la ubicación exacta del siniestro, al igual que LONGITUDX, se pueden usar estas coordenadas en Google Maps para determinar el lugar exacto y realizar la corrección respectiva.



La variable CANTÓN se recodifica para construir la nueva variable: CANTON_T. Esto basado en el Clasificador Geográfico Estadístico – DPA.

Crear la variable: NUM_FALLECIDO_T que es igual a NUM_FALLECIDO.

Crear la variable: NUM_LESIONADO_T que es igual a NUM_LESIONADO. Para crear la variable: TOTAL_VICTIMAS, esta se genera como la suma de NUM_FALLECIDO_T y NUM_LESIONADO_T.

Posterior a la creación de variables se procede asignar etiquetas a las variables.

www.ecuadorencifras.gob.ec

Buenas cifras, mejores vidas

03.

Transporte Ferroviario

Proceso y metodología de las distintas etapas de validación de Transporte Ferroviario

1. **Cambios y creación de variables**

Usando como base la variable NombreFilial, categorizar los nombres de las fialiales por las categorías:

- FILIAL 1 Centro
- FILIAL 2 Litoral
- FILIAL 3 Norte
- FILIAL 4 Sur.

La variable NombreViaje sirve para calcular el mes de la salida en ferrocarril. Generar una nueva variable denominada MES_SALIDA.

Seguido, crear la variable FECHA una vez se ha revisado las variables EDAD y FECHA DE NACIMIENTO.

Nota: En este punto se debe trabajar de cerca con la variable **FechaNacimiento** para calcular la edad exacta de la persona al momento del viaje.

Usando la variable: **TipoIdentificación** se puede corregir posibles errores al momento de calcular la edad, ya que suele pasar que existen problemas de digitación.

Así también, se puede usar la variable NombreCompleto, sólo para los pasajeros nacionales, ya que en otros casos no es posible dar con la fecha correcta.

Se extrae los datos de mes y día de la variable NombreViaje y se concatena con el año de referencia, ejem 01/02/2020

Para la variable País, se requiere recodificar la variable y homologar nombre de países.

Generar una nueva variable denominada SEXO, con las categorías:

1 Hombre
2 Mujer
9 No registrado.

www.ecuadorencifras.gob.ec

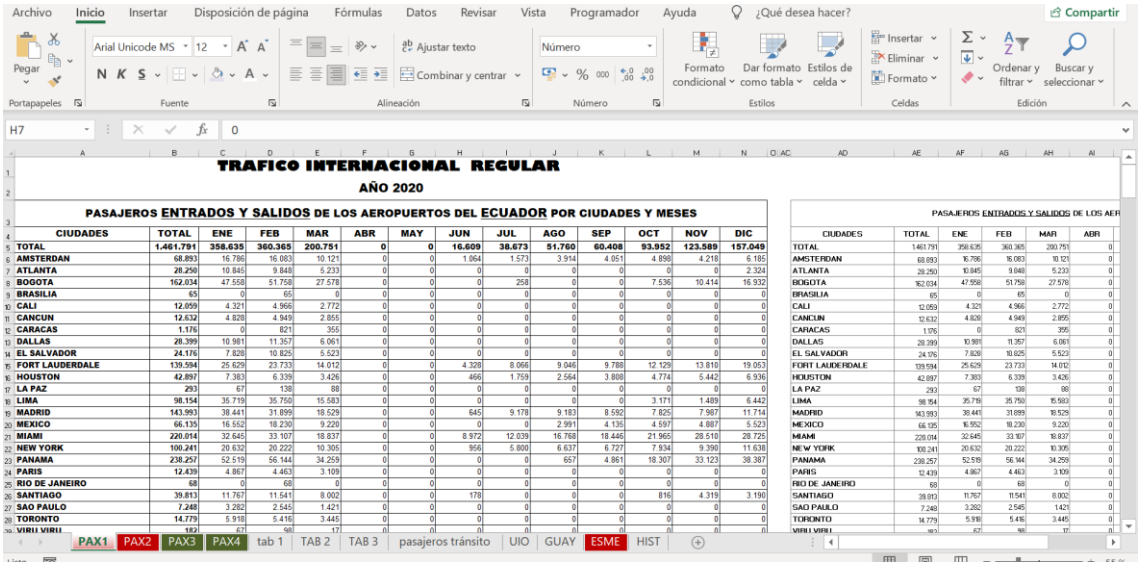
Buenas cifras, mejores vidas

04

Transporte Aéreo

Proceso y metodología de las distintas etapas de validación de Transporte Aéreo

Para este componente no existe una base de datos como tal. En su lugar se recibe una serie de tablas de transporte aéreo internacional y doméstico para pasajeros y carga. Con esta información se trabaja en replicar los cuadros de Excel que realiza la Dirección General de Aviación Civil.



TRAFICO INTERNACIONAL REGULAR

AÑO 2020

PASAJEROS ENTRADOS Y SALIDOS DE LOS AEROPUERTOS DEL ECUADOR POR CIUDADES Y MESES

CIUDADES	TOTAL	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
TOTAL	1.461.791	358.635	360.365	200.751	0	0	16.609	38.673	51.760	60.408	93.952	123.589	157.049
AMSTERDAN	68.893	16.786	16.083	10.121	0	0	1.064	1.573	3.914	4.051	4.898	4.218	6.185
ATLANTA	28.250	10.845	9.848	5.233	0	0	0	0	0	0	0	0	2.324
BOGOTA	162.034	47.558	51.758	27.578	0	0	0	258	0	0	7.536	10.414	16.932
BRASILIA	65	0	65	0	0	0	0	0	0	0	0	0	0
CALI	12.059	4.321	4.966	2.772	0	0	0	0	0	0	0	0	0
CANCUN	12.632	4.828	4.949	2.855	0	0	0	0	0	0	0	0	0
CARACAS	1.176	0	821	355	0	0	0	0	0	0	0	0	0
DALLAS	28.399	10.981	11.357	6.061	0	0	0	0	0	0	0	0	0
EL SALVADOR	24.176	7.828	10.825	5.523	0	0	0	0	0	0	0	0	0
FORT LAUDERDALE	139.594	25.629	23.733	14.012	0	0	4.328	8.066	9.046	9.788	12.129	13.810	19.053
HOUSTON	42.897	7.383	6.339	3.426	0	0	466	1.759	2.564	3.808	4.774	5.442	6.936
LA PAZ	293	67	138	88	0	0	0	0	0	0	0	0	0
LIMA	98.154	35.719	35.750	15.583	0	0	0	0	0	0	3.171	1.489	6.442
MADRID	143.993	38.441	31.899	18.529	0	0	645	9.178	9.183	8.592	7.825	7.987	11.714
MEXICO	66.135	16.552	18.230	9.220	0	0	0	0	2.991	4.135	4.597	4.887	5.523
MIAMI	220.014	32.645	33.107	18.837	0	0	8.972	12.039	16.768	18.446	21.965	28.510	28.725
NEW YORK	100.241	20.632	20.222	10.305	0	0	956	5.800	6.637	6.727	7.934	9.390	11.638
PANAMA	238.257	52.519	56.144	34.259	0	0	0	0	657	4.861	18.307	33.123	38.387
PARIS	12.439	4.867	4.463	3.109	0	0	0	0	0	0	0	0	0
RIO DE JANEIRO	68	0	68	0	0	0	0	0	0	0	0	0	0
SANTIAGO	39.813	11.767	11.541	8.002	0	0	178	0	0	0	816	4.319	3.190
SAO PAULO	7.248	3.282	2.545	1.421	0	0	0	0	0	0	0	0	0
TORONTO	14.779	5.918	5.416	3.445	0	0	0	0	0	0	0	0	0

Una vez que detectan los datos que se encuentran erróneos, se realiza una tabla donde se evidencie el error u omisión para enviar a la DGAC para su corrección o aclaración respectiva.

De lo contrario, se procede a generar los tabulados de este componente.

Buenas cifras,
mejores vidas

05.

Transporte Marítimo

Proceso y metodología de las distintas etapas de validación de Transporte Marítimo

Las bases de transporte marítimo se reciben en formato Excel. Cada uno debe ser sometida a una revisión por separado, posteriormente se realiza la consolidación de las bases para ver todo en su conjunto y realizar validaciones.

La homologación de la información recibida por los diferentes puertos se debe realizar de acuerdo a las variables que disponga en las bases ya que algunos puertos, como las superintendencias petroleras, simplemente omiten o borran algunas variables. O, por otro lado, como APE incrementan variables. Una vez homologados, sobre todo en número de variables, se traslada a software con el que se trabajará.

1. Creación y validación de variables

Una vez se realiza la unión de bases en SPSS se puede iniciar la validación. No se realizan cambio y/o arreglos en esa unión de bases, simplemente se colocan de acuerdo al formato establecido ya que muchos puertos sólo envían las variables que ellos generan información.

NOMBREDELPUERTO Homologar los nombres de los puertos:

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

TIPODEPUERTO: Variable de validación

IDREGISTRO: Variable de validación que sirve para completar datos del buque en caso de que falte alguna información. Se lo realiza en la página: https://www.marinetraffic.com/

IDBUQUE: Variable de validación que sirve para completar datos del buque en caso de que falte alguna información. Se lo realiza en la página: https://www.marinetraffic.com/

NRO.MATRÍCULA: Variable de validación que sirve para completar datos del buque en caso de que falte alguna información. Se lo realiza en la página: https://www.marinetraffic.com/

NOMBREDELANAVE: Variable de validación que sirve para completar datos del buque en caso de que falte alguna información. Se lo realiza en la página: https://www.marinetraffic.com/

TIPODELANAVE: Variable de validación

TRÁFICO: Se homologa el tipo de puerto:

string Codtraf(A15).

CABOTAJE o NACIONAL) = NACIONAL.

INTERNAC o INTERNACIONAL o Internacional = INTERNACIONAL: Se crea a una segunda base sólo de tráfico internacional para trabajar a partir de esa. Eso no quiere decir que la anterior no sea trabajada, ya que, se solicita las novedades sobre toda la información nacional o internacional

FECHAENTRADA: Se extrae sólo la fecha formato: DD/MM/YYYY sin horas. Se genera una variable auxiliar para comparación de fechas, que verifique que la fecha de entrada no sea mayor a la de salida

FECHASALIDA: Se extrae sólo la fecha formato: DD/MM/YYYY sin horas. Se genera una variable auxiliar para comparación de fechas, que verifique que la fecha de entrada no sea mayor a la de salida

BANDERA: Homologación de nombres de países. Se genera nueva variable Bandera con nombres de países homologados.

TONELAJEDEREGISTROBRUTOTRB: No se realizan cambios, sin embargo se revisa la información.

TONELAJEDEREGISTRONETOTRN: No se realizan cambios, sin embargo se revisa la información.

NÚMEROPASAJEROSENTRADOS: Se corroborar la información con lo remitido por Puertos. Generalmente APG y APM, son puertos de pasajeros, especialmente de cruceros, así que los pasajeros entrados deben ser los pasajeros que salieron. Se homologa pasajeros entrados y salidos.

NÚMEROPASAJEROSSALIDOS: Se corroborar la información con lo remitido por Puertos. Generalmente APG y APM, son puertos de pasajeros, especialmente de cruceros, así que los pasajeros entrados deben ser los pasajeros que salieron. Se homologa pasajeros entrados y salidos.

AGENCIANAVIERA Se homologa los nombres de las agencias navieras. Surge dificultad ya que una sola empresa suele manejar varias agencias en Ecuador.

NÚMERODECONTENEDORESVACIOS20ENTVariable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

NÚMERODECONTENEDORESLLENOS20ENT Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

CARGACONTENEDORES20™ENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

NÚMERODECONTENEDORESVACIOS40ENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

NÚMERODECONTENEDORESLLENOS40ENT: Variable de publicación - se debe confirmar con suma total: Se crea variable auxiliar de comparación de sumas y totales

CARGACONTENEDORES40™ENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

TOTALDECARGACONTENERIZADA™ENT: Variable de publicación - se debe confirmar con suma total: Se crea variable auxiliar de comparación de sumas y totales.

CARGAGRANELSÓLIDA™ENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGAGRANELLÍQUIDA™ENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGAGENERAL™ENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGADETRANSBORDO™ENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

NÚMERODEVEHÍCULOSENT Variable de publicación - se debe confirmar con suma total incorporando a carga general ya que un solo puerto realiza esta diferenciación. Se crea variable auxiliar de comparación de sumas y totales.

VEHÍCULOS™ENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

TOTALCARGA™ENT Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

NÚMERODECONTENEDORESVACIOS20SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

NÚMERODECONTENEDORESLLENOS20SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

CARGACONTENEDORES20™SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

NÚMERODECONTENEDORESVACIOS40SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

NÚMERODECONTENEDORESLLENOS40SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGACONTENEDORES40™SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

TOTALDECARGACONTENERIZADA™SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGAGRANELSÓLIDA™SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGAGRANELLÍQUIDA™SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

CARGAGENERAL™SAL: Variable de publicación - se deben confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGADETRANSBORDO™SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

NÚMERODEVEHÍCULOSSAL Variable de publicación - se debe confirmar con suma total incorporando a carga general ya que un solo puerto realiza esta diferenciación.

VEHÍCULOS™SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

TOTALCARGA™SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

TOTLACARGAEXPORIMPOR: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

Zona: Esta variable tiene información para APG, permite diferenciar de los puertos públicos de privados. (Contecon, Andipuerto y DP World).

Observaciones generales

APE presenta una variable denominada: vehículos en Tm, variable que no estuvo en la matriz del año pasado enviada por el puerto, tanto en entrada como salida.

De acuerdo a la técnica del puerto, es para mayor detalle. Sin embargo, no se usa ese dato en validaciones

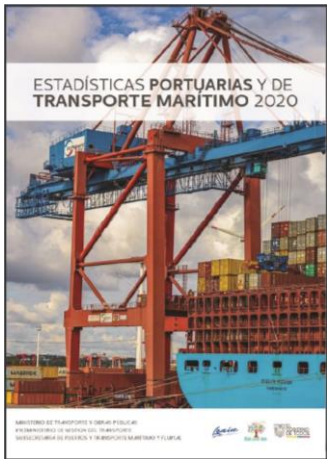
APG presenta información de todos los puertos públicos, privados y bajo concesión que posee, lo que impide una depuración de la data al no poder identificar cuales conforman la data regular. Se espera las cifras del boletín de transporte marítimo para comparar y depurar de forma correcta

5. Comparación de la información

Es importante comparar la información de cada componente con la fuente que los crea. Esto permite identificar errores y someterlos a un profundo análisis de los datos. Para esto, se contacta a las fuentes con las novedades presentadas y se solicita su revisión.

Para aquellos que disponen datos ya elaborados y publicados se realiza una comparación con estos:

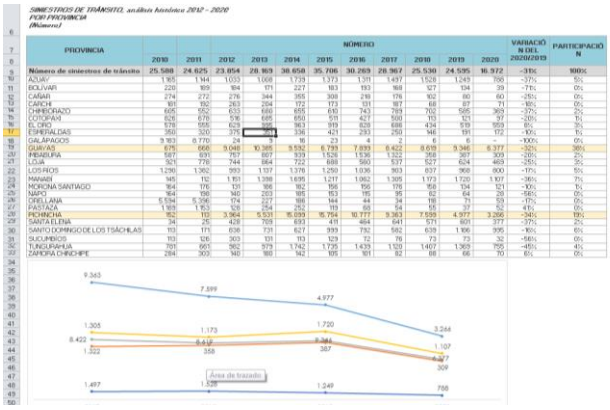
Ilustración 1 Documentos comparativos para las estadísticas de transporte



Fuente: Varios

Así también es necesario revisar la consistencia de los datos a través del tiempo.

Ilustración 2 Consistencia de las estadísticas histórica



SINIESTROS DE TRÁNSITO, análisis histórico 2012 - 2020
POR PROVINCIA
(Número)

PROVINCIA	NÚMERO											VARIACIÓN DEL 2020/2019	PARTICIPACIÓN
	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020		
Número de siniestros de tránsito	25.588	24.625	23.854	28.169	30.658	35.706	30.269	28.967	25.530	24.595	16.572	-33%	100%
AZUAY	1.165	1.144	1.033	1.008	1.739	1.373	1.171	1.437	1.528	1.249	786	-37%	5%
BOLÍVAR	220	199	184	171	227	183	193	168	127	134	39	-71%	0%
CAÑAR	274	272	276	344	355	309	219	176	102	80	60	-25%	0%
CARCHI	181	192	263	204	172	173	131	187	68	87	71	-18%	0%
CHIMBORAZO	605	552	653	686	655	610	743	789	702	585	369	-37%	2%
COTOPAXI	628	676	576	685	650	571	427	500	113	121	97	-20%	1%
EL ORO	578	555	629	395	963	919	639	686	434	519	559	8%	3%
ESMERALDAS	350	320	375	[illegible]	336	421	233	250	146	191	172	-10%	1%
GALÁPAGOS	9.383	8.770	24	3	16	23	4	2	6	6	1	-100%	0%
GUAYAS	675	666	5.048	10.385	9.532	6.793	7.893	8.422	8.619	3.346	6.377	-90%	38%
IMBABURA	587	630	767	807	533	1.526	1.530	1.322	368	387	309	-20%	2%
LOJA	761	778	744	864	722	688	560	537	527	624	469	-25%	3%
LOS RÍOS	1.290	1.362	993	1.137	1.376	1.250	1.036	903	837	968	800	-17%	5%
MANABÍ	145	112	1.161	1.398	1.695	1.017	1.062	1.305	1.173	1.120	1.107	-36%	7%
MORONA SANTIAGO	164	176	131	186	182	156	156	176	158	134	121	-10%	1%
NAPO	164	198	140	203	185	153	115	35	62	64	28	-56%	0%
ORELLANA	5.134	5.356	174	227	186	144	44	34	118	71	59	-17%	0%
PASTAZA	1.169	1.153	126	254	252	119	69	54	45	37	52	41%	0%
PICHINCHA	152	113	3.964	5.531	15.039	15.754	10.777	9.363	7.599	4.977	3.266	-34%	19%
SANTA ELENA	34	25	418	709	693	411	464	641	571	501	377	-27%	2%
SANTO DOMINGO DE LOS TSÁCHILAS	113	171	638	731	627	993	732	582	639	1.166	995	-16%	6%
SUCUMBÍOS	113	126	303	131	113	123	72	78	71	73	32	-56%	0%
TUNGURAHUA	791	661	982	979	1.742	1.725	1.439	1.120	1.407	1.363	755	-45%	4%
ZAMORA CHINCHIPE	284	303	140	180	142	105	101	82	68	66	70	6%	0%

Fuente: GESE

6. Conclusiones

- Este documento es importante ya que, detalla las diferentes etapas de validación a los que es sometida la información recibida desde las diferentes fuentes del Anuario y, dicta las pautas para procedimientos futuros así como la mejora de estos.

- Es importante generar procesos de validación para prevenir errores en la información recibida. Es así que, la planificación de la recolección es clave para tener datos oportunos y que estos puedan ser sometidos a las validaciones para identificar errores y tener retroalimentaciones con las fuentes de los registros administrativos para obtener datos de calidad.

- El trabajo conjunto entre el equipo técnico y los técnicos de las diferentes instituciones fuente, ha favorecido al momento de identificar y explicar datos que pueden estar inconsistentes. De esta manera, se obtienen cifras homologadas entre las instituciones que, sobre todo ayudan al usuario.

- Después de realizadas las validaciones, se generan resultados oportunos y de calidad, así como otros productos para la revisión ya probación de las diferentes instancias del INEC. Finalmente, estos son presentados a los usuarios del anuario de acuerdo al calendario estadístico.

Elaborado por:	**Revisado y Aprobado por:**
Gabriela Hidalgo	
Gabriela Hidalgo **Miembro del Equipo**	Roberto Chaves **Jefe de área**

www.ecuadorencifras.gob.ec

INEC
Buenas cifras,
mejores vidas
@ecuadorencifras
@ecuadorencifras
@InecEcuador
t.me/ecuadorencifras
INEC/Ecuador
INECEcuador
INEC Ecuador