

Procedimiento para documentar errores en datos publicados

OPERACIÓN ESTADÍSTICA DE TRÁFICO Y RECAUDO DEL MODO CARRETERO

Agencia Nacional de Infraestructura

Introducción

Los registros administrativos de tráfico y recaudo del modo carretero se caracterizan por conformar una serie de datos sujetos a regulación y control de la función administrativa de la entidad, captando de manera continua la información de acuerdo con procesos de recolección y aprovechamiento estadístico de la información generada para la toma de decisión interna y de consulta a nivel sectorial.

Principios de calidad

De acuerdo a la utilidad que representa para los usuarios la información publicada en la página web de la Agencia, se requiere garantizar la calidad de la información y la transparencia en la publicación de cifras y los protocolos tanto de acceso como de posibles errores en la recolección, sistematización y/o publicación de los resultados.

Como parte de la adopción del código de Buenas prácticas para las estadísticas oficiales del Departamento Administrativo Nacional de Estadística – DANE, como ente rector del Sistema Estadístico Nacional, este protocolo aborda los criterios de calidad relacionados con el Principio No 4: Imparcialidad y normas en la difusión de procedimientos que garanticen la calidad técnica del proceso estadístico (DANE, 2011)

Específicamente los principios:

4.10. Corregir y comunicar oportunamente los errores identificados en las estadísticas publicadas.

4.11. Formular precisiones frente a las interpretaciones erróneas de las estadísticas oficiales.

Tipos de errores:

Teniendo en cuenta los Lineamientos básicos de una investigación estadística (DANE, 2012), para lo correspondiente a las operaciones estadísticas basadas en registros administrativos, los tipos de errores (errores no muestrales) más comunes son los siguientes:

- **Errores en la identificación de las unidades estadísticas:** Se presenta dado el caso que algún peaje no reporte información de tráfico y recaudo mes a mes, las principales causas son rechazo de la unidad informante por inaccesibilidad operativa.
- **Errores de observación:** Cuando el valor capturado es diferente al valor real de la variable, por deficiencias en el proceso de clasificación de vehículos o por errores en el tratamiento de la información:
 - **Errores en el formulario de recolección:** Se genera cuando el cuestionario tiene deficiencias en el diseño o no es claro su diligenciamiento, cuando no se verifica la completitud del diligenciamiento del formulario y se presentan datos faltantes.
 - **Errores en la recolección.** De responsabilidad del recolector en campo debido a que no tuvo una capacitación apropiada o a los débiles procesos de supervisión y control.
 - **Errores en el procesamiento:** cuando se presentan problemas en los procesos de crítica y captura de la información (pérdida de información luego de la integración de los formularios de recolección en el sistema).

Según los Lineamientos para el Manejo de la Imparcialidad en la Actividad Estadística (DANE, 2014), los responsables de la operación estadística deben alertar y corregir de manera oportuna los errores detectados en datos publicados, como criterio de calidad de transparencia hacia los usuarios.

Procedimientos recomendados para la identificación de errores

Realizar sondeos aleatorios cada mes para verificar que la información ingresada al sistema coincida con los reportes entregados en los informes de interventoría de cada concesión.

Diseñar mecanismos de validación y consistencia de las cifras en las cuales se analicen las variaciones del tráfico y recaudo con respecto al periodo anterior y al mes correspondiente al año anterior, identificando variaciones porcentuales negativas superiores a un porcentaje determinado por cada supervisor de acuerdo a las estimaciones de crecimiento del tráfico.

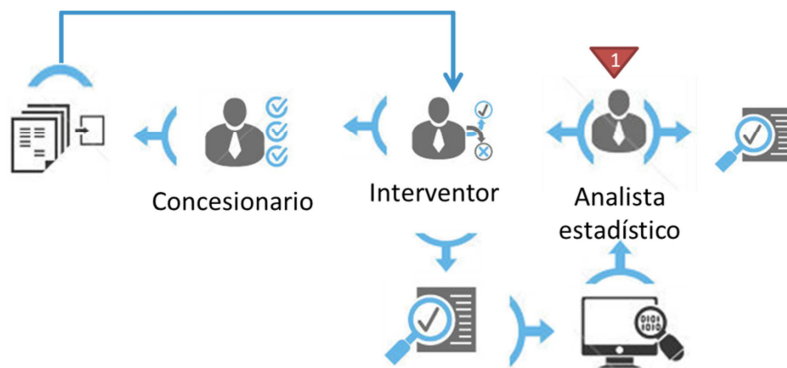
Analizar las cifras de recaudo teniendo en cuenta el volumen de tráfico y las tasas de peaje vigentes, identificando diferencias considerables en los resultados reportados para cada peaje.

Procedimientos recomendados para verificación y corrección de errores antes de la etapa de difusión

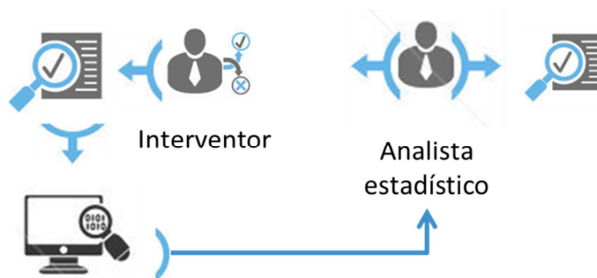
Dentro de los procedimientos recomendados para corregir los errores detectados, se cuenta con los siguientes:

- a. **Se vuelve a contactar a la fuente de información:** en caso que el error sea de digitación, se debe generar el reporte de corrección al interventor, este debe verificar la información con el concesionario e identificar si es error de medición o de digitación de acuerdo a los reportes del

concesionario. Se sugiere que el plazo máximo sea de **diez (10) días hábiles** para el ajuste de novedades.



b. Se realiza la corrección al formato de recolección: cuando es un error de digitación del interventor de acuerdo con la información suministrada por el concesionario, se procede a ingresar al sistema y realizar la corrección correspondiente. Se sugiere que el plazo máximo sea de **cinco (5) días hábiles** para el ajuste de novedades.



c. Se verifica mediante controles de calidad: Si en el proceso de compilación de los formularios para la generación del reporte mensual se elimina información o se pierde debido a alguna falla de integración de las bases de datos, el analista estadístico y el ingeniero de sistemas deben verificar los formatos de recolección de los datos faltantes. Se sugiere que el plazo máximo sea de **cinco (5) días hábiles** para el ajuste en las cifras y el cargue de las novedades con la respectiva anotación de modificación en el sistema.



Nota: Existen métodos más sofisticados para corregir errores, sin embargo de acuerdo a las características de la operación estadística no es recomendable realizar procedimientos de imputación de datos¹, ya que los resultados del volumen de tráfico son base para el recaudo, por tanto no se aconseja sustituir valores de las variables erróneas o ausentes, sin verificar con el interventor y concesionario.

Procedimientos recomendados para verificación y corrección de errores en la etapa de difusión

- a. En el enlace <http://www.ani.gov.co/carreteras> se genera una nota para comunicación de usuarios en la cual se habilita una dirección de correo electrónico para el reporte de inconsistencias o identificación de datos faltantes en los resultados.
- b. Reuniones anuales con usuarios estratégicos de la información estadística para presentar resultados y retroalimentar comportamiento de la serie y cuadros de salida generados y publicados en la página web.
- c. Se compila a lo largo del año un documento que incluya los datos que fueron verificados y modificados en la serie, bajo el nombre de “Fe de erratas” para su posterior publicación. En dicho documento se debe especificar el tipo de error, su magnitud y las posibles consecuencias en análisis inconsistente de la información para los usuarios.
- d. Al realizar las correcciones se deben cargar de nuevo los archivos Excel, es importante a pie de página incluir la siguiente leyenda ***“La Agencia Nacional de Infraestructura se reserva la modificación y/o ajustes a las cifras correspondientes, teniendo en cuenta los procedimientos de control de calidad de sus series históricas, al finalizar cada año el usuario puede encontrar una Fe de erratas con las precisiones en los datos actualizados, en caso que estos se generen”.***

¹ De acuerdo con (DANE, 2012), cuando persiste información faltante, posterior a los controles de calidad básicos para solución de datos ausentes, se recurre a aplicar procesos de imputación, los cuales se dividen en las siguientes categorías:

“Determinístico: El dato faltante recibe un único valor plausible, con el fin de que se cumplan las reglas de crítica planteadas.

Sustitución: Se hace uso de una base de datos auxiliar o de un registro administrativo que contenga información referente a las mismas variables del mismo individuo en un tiempo anterior.

En frío o de expertos: El valor faltante es imputado a criterio de un experto en la temática o con un dato histórico no necesariamente del mismo individuo.

En caliente: Consiste en la sustitución del dato faltante por uno perteneciente a otro individuo de la misma encuesta, con características similares” (DANE, 2012, pág. 50).