

Instructivo para desarrollar ejercicios de analítica de datos para la gestión preventiva de conflictos de intereses



FUNCIÓN PÚBLICA

**DEPARTAMENTO ADMINISTRATIVO DE
LA FUNCIÓN PÚBLICA**

**DIRECCIÓN DE PARTICIPACIÓN
TRANSPARENCIA Y SERVICIO AL CIUDADANO**

2022



Presentación

El Departamento Administrativo de la Función Pública (DAFP) presenta el instructivo para facilitar tanto a las entidades, como a la ciudadanía el proceso de análisis de datos para la gestión preventiva de conflictos de intereses, entendiendo el concepto de análisis de datos como una disciplina que se centra en la generación de conocimiento, a partir de grandes volúmenes de datos que incluye masivos y complejos procedentes de diversas fuentes, dentro de un proceso de resolución de algún problema o cuestionamiento.

La analítica de datos, un área explotada principalmente por el sector privado para sus modelos de marketing y crecimiento empresarial, hace referencia a «un proceso que permite transformar conjuntos de datos en información para la toma de decisiones basada en evidencias» (Función Pública, 2021) y requiere del uso de modelos matemáticos y estadísticos para su procesamiento y análisis, así como de metodologías y tecnologías para el procesamiento y análisis de los datos masivos mediante lo que se ha denominado la Ciencia de Datos (BID, 2017, p 7)¹.

En este contexto, el despliegue de la transparencia activa y los datos abiertos en el sector público han permitido comenzar a utilizar la analítica de datos para la toma de decisiones estratégicas en política pública, así como en el control y prevención de los actos que atentan contra la integridad pública.

El avance que ha tenido Colombia en materia de publicación y acceso a la información pública ha permitido promover un enfoque de Estado Abierto, una participación ciudadana informada, así como una rendición de cuentas constante. Sin embargo, uno de los grandes desafíos ha consistido en el uso de esta gran cantidad de información que está a disposición de la ciudadanía y de las mismas entidades, para el proceso de toma de decisiones estratégicas de política pública y la mejora de sus proyectos y políticas.

El Departamento Administrativo de la Función Pública (DAFP), como entidad orientada a promover una gestión pública moderna, eficiente y transparente pone a disposición de las entidades públicas del orden nacional y territorial este instructivo para la analítica de datos sobre las declaraciones de bienes, renta y el registro de conflictos de intereses de los sujetos obligados de la Ley 2013 de 2019 y que es alojada en el Aplicativo por la Integridad Pública.

El Aplicativo por la Integridad Pública es la herramienta del Sistema de Información y Gestión del Empleo Público (SIGEP), desarrollada por Función Pública para que los servidores públicos de altos cargos del Estado, contratistas y demás sujetos obligados cumplan lo estipulado en la Ley 2013 de 2019 y el Decreto 830 de 2021, relacionado con las Personas Expuestas Políticamente (PEP), y la ciudadanía, antes de control, medios de comunicación y organizaciones interesadas consulten y puedan usar esta información en cumplimiento de los principios de integridad, transparencia y el ejercicio del control social.

¹ La Ciencia de Datos combina un conjunto amplio de técnicas provenientes de disciplinas tales como Ciencias de la Computación, Matemáticas, Estadística, Econometría e Investigación operativa. Ver en o El uso de datos masivos y sus técnicas analíticas para el diseño e implementación de políticas públicas en Latinoamérica y el Caribe, 2017, p. 7.



El objetivo de este instructivo es orientar a las entidades públicas del orden nacional y/o territorial en el desarrollo de ejercicios de análisis de datos de las declaraciones de bienes, rentas y el registro de conflictos de intereses de los sujetos obligados conforme a lo establecido en la Ley 2013 de 2019 y la declaración de información financiera complementaria de las Personas Expuestas Políticamente - PEP (Decreto 830 de 2021) con base en la información registrada en el Aplicativo por la Integridad.

Asimismo, la finalidad se centra en el cumplimiento de los principios de integridad, transparencia y el ejercicio del control social, orientándose en la promoción de una gestión pública moderna, eficiente y transparente.

Este instructivo contiene seis (6) pasos que se encuentran acompañados de ejemplos ilustrativos para la generación de nuevos aprendizajes para el análisis de datos que contribuyan a la formulación de acciones para la gestión preventiva de conflictos de intereses.

Pasos para adelantar ejercicios de analítica de datos sobre las declaraciones de bienes, rentas y conflictos de intereses en las entidades públicas

Con el fin de que su entidad pueda mejorar la gestión preventiva de conflictos de intereses con base en la identificación de posibles riesgos de corrupción, la generación de un mapeo de los conflictos de intereses más reiterados, la precisión de las áreas en las cuales existen mayor cantidad de declaraciones o recusaciones o que tienen mayor probabilidad de que se generen, se propone desarrollar ejercicios de analítica de datos sobre las declaraciones de conflictos de intereses y recusaciones de sus servidores y contratistas mediante los siguientes 6 pasos:

Ilustración 1. Pasos para adelantar el ejercicio de analítica de datos sobre potenciales conflictos de intereses en las entidades públicas



Fuente: Departamento Administrativo de la Función Pública (2022)

Estos pasos le permitirán identificar sus necesidades de información, determinar las principales fuentes de datos, precisar los análisis que requiere y elaborar unas conclusiones que sirvan de insumo para actualizar su estrategia anual de Integridad Pública e implementar acciones de mejora en las actividades correspondientes.

A continuación, se explican un resumen de los diferentes tipos de análisis de datos con sus principales características y se presentan cada uno de los pasos propuestos y las acciones por desarrollar para su ejecución en el marco de los ejercicios de analítica de datos.




Es importante tener en cuenta que existen diferentes tipos de análisis de datos, como el análisis descriptivo, el análisis predictivo y el análisis prescriptivo, los cuales tienen las siguientes características:

Ilustración 2. Tipos de análisis de datos

Descriptivo (Qué sucede o qué sucedió)	Predictivo (Qué pasará o puede pasar)	Prescriptivo (Qué se debería hacer)
<ul style="list-style-type: none">• Analiza los cambios de forma detallada (dónde/cuáles).• Se puede dividir entre descriptiva y diagnóstica.• Busca entender el comportamiento en las variables para encontrar una mejor solución, alineándose al propósito.• Datos + Procesamiento = Qué pasó y qué está pasando + Por qué pasó y está pasando.	<ul style="list-style-type: none">• Son los datos históricos al servicio del futuro probable. Pregunta ¿qué podría pasar en el futuro?, ¿qué sucedería si...?• Es un proceso de análisis que da respuesta de futuro, para lo cual se emplean técnicas estadísticas, matemáticas y de inteligencia artificial.• Datos + modelación = futuro probable.	<ul style="list-style-type: none">• Se caracteriza por la optimización de escenarios probables para seleccionar la «mejor» decisión.• Cuando se realiza el proceso de toma de decisión, la mejor solución se centra en estrategias de simulación y optimización.• Datos + modelación = opción óptima.

Fuente: Adaptación. Departamento Administrativo de la Función Pública (2022)

Ilustración 3. Ejemplos para los diferentes tipos de análisis de datos

 Descriptivo	 Predictivo	 Prescriptivo
<p>¿Cuántos declarantes tienen conflictos de intereses en la entidad?</p> <p>Cuáles son las áreas o dependencias con más conflictos de intereses?</p> <p>¿Cuáles son los factores de conflictos de intereses en la entidad?</p>	<p>¿Qué áreas o dependencias tendrán mayor riesgo de conflictos de intereses?</p> <p>¿Qué cargos tendrán mayor riesgo de conflictos de intereses?</p> <p>¿Cuál es la probabilidad de que ocurran conflictos de intereses?</p>	<p>¿Qué sucedería si aumentan o disminuyen las declaraciones?</p> <p>¿Qué cambiaría frente a una acción específica de la entidad al respecto?</p>

Fuente: Adaptación. Departamento Administrativo de la Función Pública (2022)



Como paso previo, conforme equipos de trabajo e involucre en ellos personas innovadoras, que formen parte de las áreas involucradas en los procesos asociados a la gestión preventiva de conflictos de intereses, que puedan contribuir en el análisis y que conozcan los objetivos estratégicos de la entidad.

Cabe precisar que, con el fin de facilitar su desarrollo, los pasos propuestos en este instructivo para desarrollar los ejercicios de analítica de datos son iguales para cada tipo de análisis sugerido.

Si la entidad no ha realizado ejercicios de analítica de datos sobre la información disponible acerca de las declaraciones de bienes, renta y conflictos de intereses, se recomienda que los primeros ejercicios se enfoquen en análisis descriptivos que permitan tener un diagnóstico de la situación de la gestión de conflictos de intereses en la entidad para, posteriormente, si se desea, realizar análisis más complejos que deriven en modelos predictivos y prescriptivos.

Paso 1. Establezca el propósito de su ejercicio de analítica de datos

Identifiquen las necesidades de su entidad frente al tema y establezcan el propósito u objetivo de su ejercicio con base en ellas, definiendo el tipo de análisis más acorde, teniendo en cuenta los tipos señalados en la ilustración anterior.

Paso 2. Defina las preguntas para orientar el proceso de análisis

Las preguntas orientadoras le permitirán entender al equipo el alcance del ejercicio de analítica por desarrollar y cuál será su utilidad en la estrategia de gestión preventiva de conflictos de intereses de la entidad, por consiguiente, tengan en cuenta los siguientes puntos para establecer dicha(s) pregunta(s) para su ejercicio de analítica de datos:

1. Asegúrense de que todo el equipo comprenda el objetivo del ejercicio, los tipos de análisis y el contexto.
2. En equipo, definan el tipo de análisis que mejor se adecúe a su objetivo y generen varias preguntas que podrían resolver las necesidades identificadas. Se sugiere que sean máximo cinco (5) preguntas.
3. Revisen y ajusten la redacción de estas preguntas para que sean claras, concretas y efectivas, y para que permitan dar respuesta a las necesidades identificadas inicialmente.
4. Seleccionen la(s) pregunta(s) que consideren más efectivas para el ejercicio y que le brinden mayor valor, de acuerdo con el propósito definido. Se sugiere que sean máximo dos (2) preguntas por ejercicio.



Paso 3. Defina las variables, fuentes de información y resultados esperados

Ahora, de acuerdo con el propósito definido para el ejercicio y su materialización en las preguntas orientadoras, definan de forma simultánea las variables que se requieren para desarrollar el ejercicio, las principales fuentes de información de las cuales pueden obtenerse y los resultados esperados.

Las principales fuentes de información para estos ejercicios son:

- ✚ las declaraciones de conflictos de intereses y recusaciones recibidas por la entidad, para cuyo seguimiento y control desarrollamos la *Matriz de seguimiento anual a la gestión de conflictos de intereses y recusaciones*, uno de los instrumentos disponibles en la Caja de herramientas de nuestro [micrositio de conflictos de intereses](#).
- ✚ y las declaraciones de bienes y rentas y conflictos de intereses registradas en el **Aplicativo por la Integridad Pública** en cumplimiento de la Ley 2013 de 2019, de las cuales pueden generar el reporte correspondiente a su entidad, actualizado en tiempo real, en el momento en que lo requieran, a través del módulo de [Consulta Ciudadana](#) de este aplicativo.

Dado lo anterior, consideren las siguientes orientaciones:

1. Establezcan las variables² que formarán parte del ejercicio y, de manera paralela, sus posibles fuentes de información.
2. Contemplan también fuentes de información complementarias, como el [Sistema de Información y Gestión del Empleo Público](#) (SIGEP), [el Portal Anticorrupción de Colombia](#) (PACO), sus matrices de control internas, análisis de gestión de riesgos e informes de control interno, entre otras.
3. Procuren acudir a fuentes de información oficiales, que les permitan obtener datos fiables de manera libre (p.e. plataforma de [Datos Abiertos](#)) y en formato reutilizable.

² Glosario Modelo Integrado de Planeación y Gestión (MIPG) ver. 7. **Variable:** Característica de las unidades estadísticas, que puede asumir un conjunto de valores los cuales pueden ser cuantitativos o cualitativos. (https://www.dane.gov.co/files/sen/normatividad/NTC_Proceso_Estadistico.pdf). Las variables utilizadas para medir el nivel de desarrollo de las políticas de gestión y el desempeño de las entidades nacionales y territoriales objeto de aplicación son variables categóricas ordinales, es decir, variables cuyas categorías pueden ser ordenadas ascendentemente (o descendentemente) a partir de un criterio lógico que en este caso corresponde al nivel de cumplimiento. Consultado en www.funcionpublica.gov.co/documents/28587410/38054865/2021-10-25_Glosario_mipg_v7.pdf/0b0d6867-1805-e77e-a95c-ad04ca533d9a?t=1635182340713 el 28 de octubre de 2022.



-
4. Establezcan el periodo de referencia de la información sobre la cual se desarrollará el análisis.
 5. Definan los resultados esperados de forma que concuerden con el objetivo del ejercicio, las preguntas orientadoras y las variables y fuentes definidas. Estos resultados esperados deben corresponderse con los reportes que surgirán del procesamiento de la información obtenida a partir de las variables definidas.



Para el desarrollo de estos primeros pasos se propone la siguiente matriz, para la cual se generan algunos ejemplos:

No.	Preguntas por resolver / Objetos de análisis <i>Máximo 5 preguntas</i>	Variables por incorporar en el análisis	Fuentes complementarias <i>Adicionales al Aplicativo por la Integridad Pública y las declaraciones y recusaciones registradas por la entidad</i>	Resultados esperados	Priorización <i>Orden en el cual desarrollará los ejercicios de analítica, según las preguntas planteadas</i>	Observaciones <i>Registre las principales observaciones que considere importante considerar en cada ejercicio</i>
1	¿Cuáles son las áreas con mayor número de conflictos de intereses declarados?	- Cargo - Área - Tipo de conflicto de intereses - Otros	- Sigep - Procuraduría - PACO - Otros sistemas de información	Frecuencia de posibles conflictos de intereses por área de la entidad.	4	Abcde fgh ijklmn opqrs...
2	¿Qué tanto variaron las declaraciones de bienes y rentas de nuestros directivos?	- Cargo - Área - Ingresos - Bienes patrimoniales - Otros	- Sigep - Procuraduría - PACO - Otros sistemas de información	Total de personas por área con variaciones significativas en ingresos declarados y/o bienes patrimoniales	1	Abcde fgh ijklmn opqrs...
3	¿Qué tanto inciden los cónyuges/compañeros permanentes y familiares en la presencia de conflictos de interés?	- ... - ... - ...	- ... - ... - ...	Abcde fgh ijklmn opqrs...	5	Abcde fgh ijklmn opqrs...
4	¿Qué causal de conflictos de intereses es más recurrente en cada área?	- ... -	- ... - ... - ...	Abcde fgh ijklmn opqrs...	3	Abcde fgh ijklmn opqrs...
5	¿Qué diferencias se presentan entre las declaraciones de bienes y rentas en SIGEP y en el Aplicativo por la Integridad Pública de nuestros directivos y ordenadores del gasto?	- ... -	- ... - ... - ...	Abcde fgh ijklmn opqrs...	2	Abcde fgh ijklmn opqrs...

Fuente: Función Pública (2022)



Paso 4. Recopile, prepare y procese la información que analizará

El documento *El uso de datos masivos y sus técnicas analíticas para el diseño e implementación de políticas públicas en Latinoamérica y el Caribe*³ plantea que la gestión de los datos se desarrolla a partir de tres aspectos: «(1) adquisición y almacenamiento de los datos, (2) limpieza y depuración de los datos y, (3) la preparación para su análisis».

En ese sentido, se sugiere desarrollar las siguientes acciones correspondientes a cada uno de estos aspectos:

Adquisición y almacenamiento de los datos

A partir de las fuentes de información definidas, recopilen los datos correspondientes y agrúpenlos en una o varias bases de datos, de acuerdo con las preguntas orientadoras definidas para el ejercicio, verificando en cada caso que la información que proviene de diferentes fuentes corresponda al mismo concepto y periodo y responda a las variables establecidas.

Por ejemplo, al unificar los datos en el campo **Fecha**, todos aquellos que se agrupen en esta columna deben corresponder a la misma categoría (fecha de ingreso, fecha de declaración, fecha del reporte u otro similar).

Limpieza y depuración de los datos

Una vez cuenten con la base de datos unificada, verifiquen que los datos requeridos estén completos y tengan consistencia y concordancia. Para ello, comprueben que:

- ✚ no haya campos vacíos.
- ✚ no haya caracteres que no correspondan al campo. Por ejemplo, signos o letras en un campo de números o viceversa.
- ✚ haya unidad en el registro de la información de cada campo en cuanto al uso de mayúsculas y minúsculas, tildes, caracteres especiales, etc., ya que datos sencillos pueden estar escritos de forma diferente en distintos campos, como una fecha (2022-10-01, 01/10/22, 2022-oct.-01, 01/10/22, etc.) o el nombre de un municipio (chigorodo, CHIGORODÓ, Chigorodo, CHIGORODO, Chigorodó, chigorodó, etc.).

Así, al depurar los datos se debe unificar la estructura de cada variable para que cada dato esté escrito de la misma forma, con el fin de que no se generen errores o diferencias en el momento del procesamiento de la información.

³ Banco Interamericano de Desarrollo (BID), *El uso de datos masivos y sus técnicas analíticas para el diseño e implementación de políticas públicas en Latinoamérica y el Caribe*. 2017.



- ✚ no exista duplicidad en los registros de correspondientes a las declaraciones de bienes, rentas y registro de conflictos de interés provenientes del Aplicativo por la Integridad Pública; es decir, que para un mismo sujeto obligado durante una vigencia específica no existan múltiples declaraciones del mismo tipo, ya que puede suceder que se hayan realizado correcciones sobre una declaración, y si se toman para el procesamiento todos los registros, los resultados generados podrían estar distorsionados. En este caso se recomienda dejar el registro que tenga fecha más reciente dentro de la misma vigencia.

La preparación para su análisis

Generen una o varias tablas dinámicas, según requieran, a partir de la información correspondiente a las variables definidas que les permitan agruparlas y establecer relaciones entre ellas que ayuden en el análisis.

Estas tablas dinámicas son una herramienta que permite resumir, analizar y agrupar datos con el objetivo de realizar comparaciones, evidenciar patrones e identificar tendencias.

En ese orden de ideas, su objetivo principal es facilitar el análisis de grandes volúmenes de datos y hacer así más sencilla la tarea de consultar, profundizar, ordenar, filtrar para obtener los resultados esperados de la información analizada.

Igualmente, esta es una de las opciones más sencillas y prácticas, en la medida en que, al no tener una estructura fija, permite organizar la información de múltiples formas hasta encontrar la que mejor se adapte para responder la pregunta de estudio.

De acuerdo con el análisis preliminar de estas tablas dinámicas, generen las gráficas, tablas y cuadros que les permitan ilustrar los resultados hallados y facilitar su descripción.

Paso 5. Analice la información y genere el informe correspondiente

Al igual que en investigaciones académicas u otro tipo de trabajos investigativos, el informe final que se presenta, contiene las principales conclusiones y recomendaciones generadas a partir del análisis de los resultados del ejercicio.

Para desarrollar el informe correspondiente al ejercicio de analítica de datos planteado tengan en cuenta las siguientes recomendaciones:

- ✚ Una vez cuenten con el reporte que incluye la base de datos, tablas dinámicas y gráficas definido como objeto de estudio analicen la información general recopilada y depurada, junto con los gráficos, tablas y cuadros que surgieron de su procesamiento.



-
- ✚ Analicen esta información comparando resultados entre sí y contrastando los resultados obtenidos contra información relevante asociada, como resultados de vigencias anteriores, universo o muestra definidos, entre otros.
 - ✚ Este análisis debe estar enfocado en responder las preguntas de análisis planteadas, a partir de las cuales deberán generar las conclusiones del ejercicio. Este análisis deberá estar en armonía con el tipo de análisis propuesto en el ejercicio bien sea descriptivo, predictivo o prescriptivo.
 - ✚ Presenten las conclusiones correspondientes a partir de la descripción de los resultados esperados.
 - ✚ Una vez descritos dichos resultados, generen las recomendaciones correspondientes para su entidad con el fin de mejorar procesos, identificar debilidades y fortalezas sobre el objeto de estudio, entre otros, según corresponda.
 - ✚ Presenten el informe final al personal directivo o tomadores de decisión de las áreas relacionadas con el fin de que conozcan y tengan en cuenta las conclusiones y recomendaciones para la actualización y mejora de sus procesos y planeación institucional.



Recomendaciones para lograr la sostenibilidad de ejercicios de analítica en las diferentes áreas de la entidad

- ✚ Conformar equipos de trabajo interdisciplinarios que definan ejercicios de analítica que resuelvan preguntas de interés de la entidad.
- ✚ Identificar las necesidades de información de la entidad y establecer el desarrollo periódico de los ejercicios de analítica requeridos.
- ✚ Documentar el proceso, los resultados, las buenas prácticas, guías y los demás elementos que puedan surgir por el desarrollo de los ejercicios de analítica de datos, para que estos puedan ser reutilizados por otras áreas y funcionarios dentro de la entidad.
- ✚ Capacitar a los tomadores de decisión de la entidad en el uso de los productos generados en los ejercicios (tableros, visualizaciones, informes, gráficos, etc.) con el fin de que ajusten o complementen las conclusiones y recomendaciones presentadas.
- ✚ Sensibilizar a los tomadores de decisión de la entidad sobre la importancia de tener en cuenta las recomendaciones y conclusiones generadas en sus procesos institucionales.
- ✚ Complementar la información presentada en este instructivo con la brindada en la [Guía para la analítica de datos y su uso en la planificación y ejecución de auditorías internas basadas en riesgos](#).
- ✚ Dinamizar la formación y autoformación de los contratistas y funcionarios de la entidad en temas de analítica de datos; herramientas para su manejo, como PowerBI; y lenguajes de ciencia de datos como R, Python, entre otros.