



Universidad Internacional de La Rioja
Escuela Superior de Ingeniería y Tecnología

Máster Universitario en Análisis y Visualización de Datos
Masivos/ Visual Analytics and Big Data

“Análisis y visualización de tendencias del mercado
inmobiliario en Bogotá para mejorar la toma de decisiones”

Trabajo fin de estudio presentado por:	Miguel Alejandro Gonzalez Cardeñoza Sergio Andrés Fuentes Gomez
Tipo de trabajo:	Piloto Experimental
Director/a:	Mariana Rios Ortego
Fecha:	11/04/2025

Resumen

El presente documento plantea como objetivo general el análisis y la visualización de las tendencias del mercado inmobiliario en la ciudad de Bogotá Colombia, permitiendo destacar los cambios relevantes que garanticen la mejor toma de decisiones para algunas pequeñas y medianas empresas que están incursionando en el sector.

El trabajo toma como referencias los datos recopilados a través del uso de fuentes de datos abiertas y algunas recolecciones de datos haciendo uso de técnicas de web scraping de los portales especializados del sector inmobiliario, esto con el fin de construir un conjunto de datos heterogéneo que permita representar desde el sector comercial y desde el sector gubernamental el movimiento inmobiliario de la ciudad.

Como metodología empleada propone una exploración de datos (EDA), con el fin de identificar tendencias, dispersión y correlación entre los datos, garantizando así un planteamiento de contraste de hipótesis como metodología estadística y concluir en la construcción de un modelo de predicción analítica que permita la estimación de un avalúo comercial para la venta de inmuebles, apuntando al mejoramiento y automatización de tareas llevadas a cabo por entidades bancarias o sus terceros quienes realizan este tipo de actividad.

Así mismo incluye modelos de visualización de los datos que permiten representar fácil y comprensiblemente las conclusiones encontradas dentro del desarrollo del trabajo.

Este trabajo piloto experimental pretende ser un modelo escalable a otras ciudades y replicable a sectores tanto públicos como privados mejorando así la toma de decisiones con base en la integración de técnicas de análisis de datos en un contexto real

Palabras clave: Inmobiliario, Bogotá, Análisis, Scraping, Vivienda

Abstract

The primary objective of this document is the analysis and visualization of real estate market trends in the city of Bogotá, Colombia. This analysis aims to highlight significant changes, thereby supporting improved decision-making for small and medium-sized enterprises (SMEs) entering this sector.

The research utilizes data collected from open data sources alongside data gathered using web scraping techniques from specialized real estate portals. The goal is to construct a heterogeneous dataset that represents the city's real estate dynamics from both commercial and governmental perspectives.

The methodology employed includes Exploratory Data Analysis (EDA) to identify trends, data dispersion, and correlations. This forms the basis for applying hypothesis testing as a statistical method. The study culminates in the development of an analytical predictive model designed to estimate the commercial valuation of properties for sale. This model aims to enhance and automate valuation tasks currently undertaken by financial institutions or their third-party service providers who perform such activities.

Furthermore, the work incorporates data visualization models to present the conclusions derived from the research in a clear and comprehensible manner.

This experimental pilot study is intended as a scalable model applicable to other cities and replicable in both public and private sectors, ultimately aiming to improve decision-making through the integration of data analysis techniques within a real-world context

Keywords: Real Estate, Bogota, Analysis, Web Scraping, Housing

Índice de contenidos

1.	Introducción	1
1.1.	Motivación	2
1.2.	Planteamiento del trabajo	3
1.3.	Estructura del trabajo	3
2.	Contexto y estado del arte	4
2.1.	Contexto del problema	4
2.2.	Estado del arte	4
2.3.	C conclusiones	¡Error! Marcador no definido.
3.	Objetivos concretos y metodología de trabajo	5
3.1.	Objetivo general	5
3.2.	Objetivos específicos	5
3.3.	Metodología del trabajo	6
4.	Marco normativo	7
5.	Desarrollo específico de la contribución	8
5.1.	“Título 2” del menú de estilos	8
5.2.	“Título 2” del menú de estilos	8
5.2.1.	“Título 3” del menú de estilos	8
6.	Código fuente y datos analizados	9
6.1.	Código fuente	9
6.2.	Datos Analizados	9
7.	Conclusiones	10
8.	Limitaciones y prospectiva	11
8.1.	Limitaciones	11
8.2.	Trabajo futuro	11

Referencias bibliográficas.....	12
Anexo A. Privacidad y protección de datos	13

Índice de figuras

Figura 1. <i>Ejemplo de figura realizada para nuestro trabajo</i>	2
---	---

Índice de tablas

Tabla 1. Distribucion del trabajo en grupo	VIII
Tabla 2. Alineación del conocimiento	IX
Tabla 3. Mecanismos de coordinación	X

Organización del trabajo en grupo

Se propone el desarrollo colaborativo por los estudiantes garantizando la construcción conceptual, la definición del problema, el planteamiento de hipótesis, y la elaboración del marco metodológico. A su vez, se dividieron responsabilidades específicas en función de sus fortalezas y experiencia de acuerdo con el siguiente planteamiento así:

Partes que aborda el TFE

Análisis estadístico y modelado predictivo

Esta parte se centra en el núcleo analítico y de machine learning del proyecto. Implica la transformación avanzada de los datos brutos en información útil, la exploración estadística profunda para entender las relaciones, correlaciones y distribuciones, apuntando al desarrollo de modelos predictivos.

Visualización de datos y análisis de tendencias

Esta parte se enfoca en la comunicación efectiva de los hallazgos y en la exploración de patrones a través de recursos de visualización de datos incluyendo ejercicios de diseño de dashboards que faciliten la toma de decisiones, el análisis específico de tendencias en el sector inmobiliario.

Distribución y estructura de la memoria

Tabla 1. Distribución del trabajo en grupo

Organización del trabajo en grupo - Desarrollo de la memoria	
Apartado de la memoria	Responsables
Introducción	Miguel González y Sergio Fuentes
Contexto y estado del arte	Miguel González y Sergio Fuentes
Objetivos y metodología de trabajo	Miguel González y Sergio Fuentes
Marco normativo	Miguel González y Sergio Fuentes
Realizar la limpieza profunda y preprocesamiento avanzado de los datos recopilados.	Sergio Fuentes

Ejecutar análisis exploratorio detallado enfocado en variables de precios, características de inmuebles y correlaciones.	Sergio Fuentes
Seleccionar, implementar, entrenar y evaluar modelos predictivos	Sergio Fuentes
Analizar y documentar la precisión de los modelos	Sergio Fuentes
Proceso detallado de limpieza y transformación de datos.	Sergio Fuentes
Colaborar en la estructuración de datos para optimizar la visualización.	Miguel González
Realizar análisis exploratorio enfocado en tendencias estacionales, temporales y patrones visuales.	Miguel González
Diseñar y desarrollar los dashboards interactivos para presentar los hallazgos clave.	Miguel González
Asegurar la claridad y facilidad de interpretación de las visualizaciones.	Miguel González
Discusión de los insights obtenidos a través de la visualización	Miguel González
Conclusiones	Miguel González y Sergio Fuentes

Fuente: Elaboración propia.

Objetivo del TFE desde el punto de vista de la adquisición de conocimientos

De acuerdo con el planteamiento del trabajo propuesto es posible evidenciar la aplicación de los conocimientos adquiridos en el desarrollo de cada asignatura aplicado al contexto real y laboral planteando los siguientes aspectos así:

Tabla 2. Alineación del conocimiento

Ejecucion	Descripción	Asignatura alineada
Recolección y Gestión de Datos Heterogéneos	manejo de datos de fuentes diversas como portales web y bases de datos públicas asumiendo retos de integración y calidad de datos del mundo real, alineado a la asignatura	ingeniería para el procesado masivo de datos
Preprocesamiento y Limpieza de Datos	Desarrollo de habilidades para tratar datos ruidosos, inconsistentes o incompletos.	bases de datos para el big data

Análisis Exploratorio de Datos (EDA)	Aplicación de técnicas estadísticas y visuales para explorar un dataset, identificando patrones iniciales, distribuciones, correlaciones y datos atípicos.	análisis e interpretación de datos
Modelado Predictivo	implementación y evaluación de técnicas de inteligencia artificial en la selección de algoritmos, alineado a la asignatura	técnicas de inteligencia artificial
Visualización Avanzada de Datos	Diseño y construcción de visualizaciones efectivas para comunicar tendencias de manera clara e intuitiva, utilizando herramientas y principios interiorizados.	visualización interactiva de la información
Interpretación y Comunicación de Resultados	informe de hallazgos técnicos en propuestas para la toma de decisiones.	análisis e interpretación de datos
Consideraciones Éticas y Normativas	Aspectos de privacidad y protección de datos.	gobernanza de datos y toma de decisiones

Mecanismos de coordinación empleados

Con el animo de garantizar un correcto uso de las herramientas y mantener una organización adecuada y sistemática para la elaboración del trabajo, se propone la siguiente tabla que incluye las razones correspondientes para la implementación y uso de cada herramienta.

Tabla 3. Mecanismos de coordinación

Actividad	Descripción	Herramienta
Reunión	semanal obligatoria para sincronizar avances, discutir problemas y planificar la siguiente semana.	Microsoft Teams
Repositorio de Código	Se propone el uso de un repositorio de código compartido para documentar y ampliar información.	Git Hub

Almacenamiento de documentos	Utilizar una carpeta compartida en la nube para la memoria del trabajo, artículos, datasets intermedios y otros documentos.	One drive
Gestor de fuentes	Software especializado y gratuito para la gestión de fuentes bibliográficas para el trabajo	Zotero
Comunicación inmediata	Canal de comunicación rápido para consultas ágiles y correo electrónico para comunicaciones más formales con la directora.	Whastapp Outlook

Pendiente validacion

1. Introducción

El mercado inmobiliario en la ciudad de Bogotá ha experimentado una serie de transformaciones en los últimos años, motivadas por dinámicas socioeconómicas, políticas gubernamentales, y condiciones macroeconómicas como la inflación y las tasas de interés. Estas variaciones han impactado directamente la oferta y demanda de vivienda¹, generando desafíos tanto para los compradores como para los actores del sector constructor e inmobiliario (Portafolio, 2024).

En este contexto, la necesidad de contar con herramientas analíticas basadas en datos se vuelve crucial para comprender las tendencias del mercado y apoyar procesos de toma de decisiones estratégicas (Camacol, 2023)

Este Trabajo Fin de Máster se enfoca en el análisis y visualización de las tendencias del mercado inmobiliario en Bogotá, con el propósito de proporcionar una visión integral del comportamiento de las ventas de vivienda (DANE, 2025), la actividad constructora (La República, 2024), y las políticas públicas que inciden sobre este sector.

A través de la recopilación de datos provenientes de fuentes abiertas y especializadas incluyendo portales inmobiliarios, informes del DANE, Camacol y el Ministerio de Vivienda se construyó un conjunto de datos representativo del entorno inmobiliario de la ciudad.

Con base en este conjunto de datos, se aplican técnicas de análisis exploratorio y modelado predictivo para identificar patrones de comportamiento, correlaciones y niveles de dispersión entre las variables más relevantes. A su vez, se desarrollan visualizaciones interactivas que permiten representar estos hallazgos de forma clara y comprensible, facilitando así su aplicación en contextos empresariales y gubernamentales.

Este piloto experimental busca ser una solución replicable y escalable, orientada a fortalecer la toma de decisiones mediante la integración de tecnologías de análisis de datos masivos en escenarios urbanos reales.

¹ En diciembre de 2024, Camacol reportó que las ventas de vivienda nueva se redujeron un 22,8 % anual, mientras que los lanzamientos cayeron un 51,9 %, evidenciando una contracción sostenida del mercado inmobiliario en Bogotá.

1.1. Motivación

En este apartado se deberá presentar el problema de estudio al que se quiere dar solución y justificar su importancia para la comunidad educativa y científica.

La lectura de este apartado debe dar una idea clara de las razones, motivos e intereses que han llevado a la elección de este tema. Recuerda que para poder justificar este trabajo debe haber referencias a la investigación previa sobre el tema objeto de estudio, independientemente de que luego se profundice en otros apartados.

Las siguientes preguntas puedan ayudar a la redacción de este apartado:

- ▶ ¿Cuál es el problema que quieres tratar?
- ▶ ¿Cuáles crees que son las causas?
- ▶ ¿Por qué es relevante el problema?

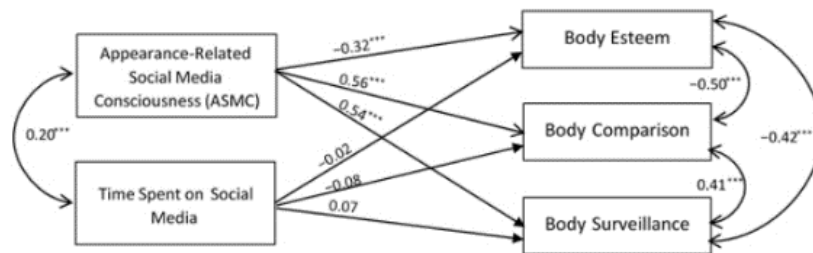
A continuación, se indica con un ejemplo cómo deben introducirse los títulos y las fuentes en Tablas y Figuras.

Tabla 2. *Ejemplo de tabla con sus principales elementos.*

Measure	Urban		Rural		$F(1, 294)$	η^2
	M	SD	M	SD		
Self-esteem	2.91	0.49	3.35	0.35	68.87***	.19
Social support	4.22	1.50	5.56	1.20	62.60***	.17
Cognitive appraisals						
Threat	2.78	0.87	1.99	0.88	56.35***	.20
Challenge	2.48	0.88	2.83	1.20	7.87***	.03
Self-efficacy	2.65	0.79	3.53	0.92	56.35***	.16

Fuente: American Psychological Association, 2020a.

Figura 1. *Ejemplo de figura realizada para nuestro trabajo.*



Fuente: American Psychological Association, 2020b.

1.2.Planteamiento del trabajo

Se debe plantear, de forma breve, el problema / necesidad detectada de la que se parte para proponer la propuesta y la finalidad del TFE. Los objetivos se van a plantear posteriormente, pero en este apartado debe quedar claro qué te planteas con la intervención.

Es necesario que los temas escogidos tengan una vinculación directa con la ingeniería de software, el desarrollo web y/o la ciberseguridad y, por tanto, el tema trabajado debe estar en consonancia con la titulación.

Las siguientes preguntas puedan ayudar a la redacción de este apartado:

- ▶ ¿Cómo se podría solucionar el problema?
- ▶ ¿Qué es lo que se propone? Aquí describes tus objetivos en términos generales.

1.3.Estructura del trabajo

Aquí describes brevemente lo que vas a contar en cada uno de los capítulos siguientes.

2. Contexto y estado del arte

Después de la introducción, se suele describir el contexto de aplicación. Suele ser un apartado (o dos en ciertos casos) en los que se estudia a fondo el dominio de aplicación, citando numerosas referencias. Debe aportar un buen resumen del conocimiento que ya existe en el campo de los problemas habituales identificados. Es el contexto general del trabajo.

Es conveniente que revises los estudios actuales publicados en la línea elegida, y deberás consultar diferentes fuentes. Hay que tener presente los autores de referencia en la temática del trabajo de investigación. Si se ha excluido a alguno de los relevantes hay que justificar adecuadamente su exclusión. Si por la extensión del trabajo no se puede señalar a todos los autores, habrá que justificar por qué se han elegido unos y se ha prescindido de otros.

El capítulo debería concluir con una última sección de resumen de conclusiones, resumiendo las principales averiguaciones del estudio y cómo van a afectar al desarrollo específico del trabajo.

Recuerda que debes referenciar adecuadamente los autores que citas en el texto y que en el aula virtual tienes información sobre cómo referenciar según la normativa APA.

Típicamente este capítulo se puede dividir en tres apartados:

2.1.Contexto del problema

2.2.Estado del arte

Estado del arte (base teórica): antecedentes, estudios actuales, comparativa de herramientas existentes, etc.

2.3.Conclusiones

Conclusiones (nexo de unión de lo investigado con el trabajo a realizar).

3. Objetivos concretos y metodología de trabajo

Este apartado es el puente entre el estudio del dominio y la contribución a realizar. Según el tipo concreto de trabajo, el bloque se puede organizar de distintas formas, pero los siguientes elementos deberían estar presentes con mayor o menor detalle.

3.1. Objetivo general

Los proyectos de análisis de datos aplicados buscan generar un impacto significativo y medible, más allá de la simple creación de herramientas o metodologías. Su propósito es demostrar la eficacia del análisis de datos mediante la propuesta de enfoques innovadores o el desarrollo de nuevas tecnologías de análisis. Así, el objetivo de estos proyectos no debe ser simplemente “crear una herramienta analítica” o “desarrollar una metodología”, sino enfocarse en lograr análisis detallados y bien articulados que ofrezcan resultados observables y aplicables.

Por ejemplo, un objetivo general pertinente podría ser: Desarrollar y aplicar un modelo analítico avanzado para la predicción y comprensión de tendencias en grandes conjuntos de datos, permitiendo su interpretación y manipulación en tiempo real a través de una interfaz interactiva. Este proyecto implicaría no solo la creación de un sistema para el análisis de datos, sino también la integración de técnicas de machine learning y visualización de datos para proporcionar una solución robusta y segura que mejore la toma de decisiones basada en datos. Además, se detallarán los procesos de desarrollo del sistema analítico y las tecnologías utilizadas, asegurando que el análisis sea lo más eficiente y seguro posible, y que esté preparado para abordar y mitigar posibles desafíos y vulnerabilidades.

3.2. Objetivos específicos

Independientemente del tipo de trabajo, el objetivo general típicamente se dividirá en un conjunto de objetivos más específicos analizables por separado. Suelen ser explicaciones de los diferentes pasos a seguir en la consecución del objetivo general.

Con los objetivos específicos, has de concretar qué pretendes conseguir. Se formulan con un verbo en infinitivo más el contenido del objeto de estudio. Se suelen usar viñetas para cada uno de los objetivos. Se pueden utilizar fórmulas verbales, como las siguientes:

Analizar – Calcular – Clasificar – Comparar – Conocer - Cuantificar – Desarrollar - Describir – Descubrir - Determinar – Establecer – Explorar -Identificar –Indagar - Medir – Sintetizar – Verificar.

3.3. Metodología del trabajo

La metodología del trabajo debe describir qué pasos se van a dar, el porqué de cada paso, qué instrumentos se van a utilizar, cómo se van a analizar los resultados, etc.

Para las tipologías de desarrollo práctico y de planificación de un proyecto de desarrollo de software se suele trabajar con técnicas de desarrollo ágil o tradicional como pueden ser Scrum, XP, Proceso Unificado, Métrica v.3, RUP, MSF, Kanban, Scrumban, SAFe o Lean, entre otros. Aunque, no necesariamente hay que seguir una técnica concreta de desarrollo, y también se puede trabajar con un modelo de desarrollo de software como puede ser espiral, iterativo e incremental o CBSE, describiendo las fases o actividades comunes en todo proceso de desarrollo de software. También puede consultar metodologías de desarrollo del tipo CRISP-DM para proyectos de análisis de datos.

Si estás realizando un trabajo de tipo 3. Desarrollo de metodología, es importante que no te confundas: una cosa es la metodología con la que desarrollas el TFE (análisis previo, propuesta de objetivos, planteamiento de la metodología, evaluación) y otra la nueva forma de hacer las cosas en el ámbito de la ingeniería del software, de la ingeniería web y de la seguridad informática que estás proponiendo.

4. Marco normativo

Es necesario que revises la normativa actual sobre privacidad y protección de datos personales. Si has incluido datos personales de terceros identificados o identificables, cuyo tratamiento es indispensable para el desarrollo del TFE, debes hacer especial consideración al RGPD y a la Ley Orgánica 3/2018 (NLOPD).

En términos prácticos, se debe determinar qué datos se van a tratar, con qué finalidades y qué tipo de operaciones de tratamiento se llevarán a cabo, documentando todo el proceso, de manera que pueda demostrarse la diligencia de la actuación ante los interesados (titulares de los datos) y las autoridades de supervisión.

Finalmente, las medidas dirigidas a garantizar el cumplimiento del Reglamento deberán tener en cuenta la naturaleza, el ámbito, el contexto y los fines del tratamiento, así como el riesgo para los derechos y libertades de las personas.

Para aplicar correctamente la normativa debes **revisar en instrucciones** el Anexo N°1 “Guía sobre privacidad y protección de datos personales”.

.

5. Desarrollo específico de la contribución

En este bloque debes desarrollar la descripción de tu contribución. Es muy dependiente del tipo de trabajo concreto (**ver instrucciones**), y puedes contar con la ayuda de tu director para estudiar cómo comunicar los detalles de tu contribución. A continuación, te presentamos la estructura habitual para cada uno de los tipos de trabajo.

Ejemplo de nota al pie².

5.1. “Título 2” del menú de estilos

Texto Normal del menú de estilos.

A continuación, se indica con un ejemplo cómo deben introducirse los títulos y las fuentes en Tablas y Figuras.

5.2. “Título 2” del menú de estilos

Texto Normal del menú de estilos.

5.2.1. “Título 3” del menú de estilos

Texto Normal del menú de estilos.

5.2.1.1. “Título 4” del menú de estilos

Texto Normal del menú de estilos.

² Ejemplo de nota al pie.

6. Código fuente y datos analizados

6.1. Código fuente

Es recomendable que el estudiante incluya en su memoria la URL del repositorio donde tiene alojado el código fuente desarrollado durante el TFE. El estudiante debe ser el único autor del código y único propietario del repositorio. En el repositorio no debe haber commit de ningún otro usuario del repositorio

6.2. Datos Analizados

De igual forma, los datos que hayan utilizado para el análisis, siempre que así se considere oportuno, también deberían estar alojados en el mismo repositorio.

Si el TFE está asociado a una actividad o proyecto de Empresa, se debe justificar en la memoria que, por temas de confidencialidad, no se deja disponible ni el código fuente ni los datos utilizados.

7. Conclusiones

Este último apartado es habitual en todos los tipos de trabajos y presenta el resumen final de tu trabajo y debe servir para informar del alcance y relevancia de tu aportación.

Suele estructurarse empezando con un resumen del problema tratado, de cómo se ha abordado y de por qué la solución sería válida.

Es recomendable que incluya también un resumen de las contribuciones del trabajo, en el que relaciones las contribuciones y los resultados obtenidos con los objetivos que habías planteado para el trabajo, discutiendo hasta qué punto has conseguido resolver los objetivos planteados. Las conclusiones ofrecidas deberán ser consecuencia del trabajo realizado y, por lo tanto, deberán marcar el grado de consecución de los objetivos propuestos (cada objetivo del trabajo se enlazará con una conclusión).

8. Limitaciones y prospectiva

8.1. Limitaciones

Una vez concluido el trabajo, deberás hacer una **valoración crítica sobre el mismo y exponer las limitaciones que has encontrado** y que han marcado la realización de tu trabajo. Aquí se deberán hacer las consideraciones pertinentes sobre qué problemas o carencias se ha encontrado el autor para el desarrollo del trabajo (necesidad de valorar otras variables, ampliar la muestra, utilizar otros instrumentos, etc.); estas serán las limitaciones del trabajo

8.2. Trabajo futuro

Finalmente, se suele dedicar un último apartado a hablar de líneas de trabajo futuro que podrían aportar valor añadido al trabajo realizado. La sección debería señalar las perspectivas de futuro que abre el trabajo desarrollado para el campo de estudio definido. En el fondo, debes justificar de qué modo puede emplearse la aportación que has desarrollado y en qué campos.

Referencias bibliográficas

Una vez que el trabajo está terminado, hay que revisar el apartado “Referencias bibliográficas”. Si has usado un sistema automático (un gestor bibliográfico tipo Endnote, Refworks o Mendeley), inserta la bibliografía en la opción adecuada (APA).

Si lo has ido haciendo manualmente, repasa que todo es correcto: aparecen todas las referencias citadas en el texto, los autores están ordenados alfabéticamente por apellidos, las cursivas son correctas, los artículos tienen números de páginas, no faltan años ni ciudades de edición, se cumple en todas las referencias la normativa APA, etc.

Se recomienda evitar citas que hagan referencia a Wikipedia y que no todas las referencias sean solo enlaces de internet, es decir, que se vea alguna variabilidad entre libros, congresos, artículos y enlaces puntuales de internet.

Ejemplos:

Swanson, E., Barnes, M., Fall, A. M., & Roberts, G. (2017). Predictors of Reading Comprehension Among Struggling Readers Who Exhibit Differing Levels of Inattention and Hyperactivity. *Reading & Writing Quarterly*, 34(2), 132-146.
doi:[10.1080/10573569.2017.1359712](https://doi.org/10.1080/10573569.2017.1359712)

Anexo A. Privacidad y protección de datos

El presente anexo establece las directrices a seguir por el alumno en la elaboración de su memoria, cuando requiera cumplir con la normativa de privacidad y protección de datos personales. (**ver instrucciones**)