Escuela Superior Politécnica del Litoral

Facultad de Ingeniería en Electricidad y Computación

Implementación de un sistema de anonimización, usando técnicas de protección de datos sensibles del departamento de ventas de una fábrica de bebidas

Proyecto de Titulación

Previo la obtención del Título de:

Magíster en Seguridad Informática

Presentado por:

Ing. Ángel Adrián Ayala López

Ing. Andrea Lisette Cáceres Molina

Guayaquil – Ecuador

Año: 2025

Agradecimiento

Quiero expresar mi más sincero agradecimiento a todas las personas que han sido parte fundamental en la realización de esta tesis.

En primer lugar, a mi padre, mis hermanas, familiares y amistades, quienes siempre han estado a mi lado brindándome su amor, apoyo y comprensión. Gracias por ser mi refugio en los momentos difíciles y por celebrar conmigo cada pequeño logro.

Ing. Angel Adrian Ayala López.

Agradezco a Dios y a mi familia por su apoyo incondicional.

También un agradecimiento a la fábrica de bebidas por ofrecernos los recursos necesarios para el cumplimiento de este proyecto.

Ing. Andrea Cáceres Molina.

Dedicatoria

A mis queridas hijas, Ainhoa y Mishell, a quienes deseo inspirar y son mi motivo de esfuerzo diario. Su amor y alegría me han dado la fuerza para superar cada obstáculo en este camino académico.

Ing. Angel Adrian Ayala Lopez

A mis padres, quienes, gracias a sus consejos y sus esfuerzos, han sido un pilar fundamental en mi formación personal y profesional.

Ing. Andrea Cáceres M.

Declaración Expresa

Nosotros Ángel Adrián Ayala López y Andrea Lisette Cáceres Molina acordamos y reconocemos que:

La titularidad de los derechos patrimoniales de autor (derechos de autor) del proyecto de graduación corresponderá al autor o autores, sin perjuicio de lo cual la ESPOL recibe en este acto una licencia gratuita de plazo indefinido para el uso no comercial y comercial de la obra con facultad de sublicenciar, incluyendo la autorización para su divulgación, así como para la creación y uso de obras derivadas. En el caso de usos comerciales se respetará el porcentaje de participación en beneficios que corresponda a favor del autor o autores.

La titularidad total y exclusiva sobre los derechos patrimoniales de patente de invención, modelo de utilidad, diseño industrial, secreto industrial, software o información no divulgada que corresponda o pueda corresponder respecto de cualquier investigación, desarrollo tecnológico o invención realizada por nosotros durante el desarrollo del proyecto de graduación, pertenecerán de forma total, exclusiva e indivisible a la ESPOL, sin perjuicio del porcentaje que nos corresponda de los beneficios económicos que la ESPOL reciba por la explotación de nuestra innovación, de ser el caso.

En los casos donde la Oficina de Transferencia de Resultados de Investigación (OTRI) de la ESPOL comunique los autores que existe una innovación potencialmente patentable sobre los resultados del proyecto de graduación, no se realizará publicación o divulgación alguna, sin la autorización expresa y previa de la ESPOL.

Guayaquil, 24 de agosto del 2025.

Firmado electrónicamente por:
ANGEL ADRIAN AYALA LOPEZ
Validar únicamente con FirmaEC

Ing. Ángel Adrián Ayala López

Firmado electrónicamente por:
ANDREA LISETTE CACERES MOLINA
Validar únicamente con FirmaEC

Ing. Andrea Lisette Cáceres Molina

Evaluadores

Firmado electrónicamente por:
LENIN EDUARDO FREIRE COBO
Validar únicamente con FirmaEC

M.Sc. Lenin Eduardo Freire Cobo

Tutor

Firmado electrónicamente por:
JUAN CARLOS GARCIA PLUA
Validar únicamente con FirmaEC

M.Sc. Juan Carlos García Plúa

Revisor

Resumen

Las empresas en Ecuador deben implementar acciones con el objetivo de proteger y resguardar la información sensible, garantizando la total privacidad de los datos. Esto es esencial para evitar cualquier actividad ilícita que pueda ser sancionada por las entidades de control. Para lograr este propósito, se contó con el apoyo y colaboración del Departamento de Ventas, así como el departamento de IT. Se establecieron medidas y protocolos de seguridad de la información para evitar el mal uso de los datos recopilados en los sistemas de ventas. El enfoque principal consistió en disminuir la vulnerabilidad de la empresa y prevenir cualquier posible violación de los datos. Un aporte valioso en este sentido fue la segmentación adecuada de la base de datos de los clientes, lo cual genera un clima de confianza en cuanto al uso de sus datos, en pleno cumplimiento de la ley.

El presente proyecto centró su enfoque principal en la protección y seguridad de la información mediante la implementación de un eficiente y efectivo sistema de anonimización de datos. Este sistema tiene como objetivo primordial disminuir la vulnerabilidad y prevenir cualquier tipo de violación de la información del departamento de ventas de la fábrica de bebidas ubicada en la ciudad de Guayaquil, Ecuador. La implementación de este sistema de anonimización de datos resulta esencial para evitar posibles sanciones derivadas del mal uso, tratamiento inadecuado o pérdida de la información sensible perteneciente a los clientes de la compañía. El incumplimiento de los estándares de confidencialidad y seguridad adecuados podría dejar a la empresa en una situación de desventaja frente a su competencia, generando así un clima de desconfianza y, en consecuencia, provocando la pérdida de clientes fundamentales para el crecimiento y desarrollo del negocio. En este sentido, fue indispensable llevar a cabo la implementación de una propuesta sólida y completa que permitió identificar y evaluar los posibles riesgos a los que se encuentra expuesta la información de la empresa, así como también

desarrollar el marco de políticas efectivas para disminuir significativamente cualquier vulnerabilidad existente. Esto, a su vez, brindará la confianza necesaria tanto a la administración del departamento de ventas como a los clientes, asegurando así un uso óptimo y apropiado de la información y, en definitiva, fomentando la fidelización de los clientes como una ventaja competitiva bajo el cumplimiento de las leyes y regulaciones vigentes.

Palabras Clave: Datos sensibles, segmentación, vulnerabilidad, riesgos, leyes.

Índice General

Abreviatura

DCP	Datos de Carácter Personal
GDPR	General Data Protection Regulation
LODPD	Ley Orgánica de Protección de Datos

Índice de figuras

Índice de tablas

Introducción

El manejo inadecuado de datos sensibles representa una amenaza significativa para la estabilidad y la reputación de las empresas en la era digital. Las organizaciones enfrentan desafíos significativos derivados de la creciente exposición a ciberataques, fuga de información y el mal uso de datos. Aquellos desafíos pueden comprometer tanto la confianza de sus clientes como la estabilidad operativa.

Esta propuesta aborda la problemática de la vulnerabilidad en el manejo de datos sensibles, planteando la implementación de un sistema de anonimización que salvaguarde los datos sensibles. El siguiente trabajo tiene como objetivo reducir los riesgos de reidentificación de los datos mediante técnicas de anonimización. La ejecución de técnicas de anonimización ayudará a fortalecer la seguridad de la fábrica y al seguir una apropiada seguridad en los datos. Al aplicar medidas adecuadas de seguridad en datos garantizará el cumplimiento a normas internacionales tales como la ISO 27001 e ISO 27002 y además de la Ley Orgánica de Protección de Datos de Ecuador.

Este trabajo está estructurado de la siguiente manera:

En el capítulo 1, se presentan los antecedentes del problema, descripción del problema, objetivos generales, objetivos específicos y la solución propuesta al problema.

En el capítulo 2, se establece el marco teórico de la propuesta de solución del trabajo y descripción de criterios para la elección de un método para la técnica de anonimización.

En el capítulo 3, se selecciona el método que se usará para la implementación de la técnica de anonimización, basados en criterios descritos en el capítulo 2.

Finalmente, en el capítulo 4, se desarrolla el marco de políticas y su correcta difusión dentro de la fábrica de bebidas de consumo masivo.

Capítulo 1

Generalidades

1.1 Antecedentes

La empresa de bebidas de consumo masivo ubicada en la ciudad de Guayaquil desde hace 137 años cuenta con un aproximado de 2.000 empleados.

En los últimos años, con el desarrollo de plataformas en la nube, las empresas que manejan gran cantidad de datos o información se han visto en la necesidad de hacer uso de estas, lo que conlleva que sean susceptibles a ataques o robo de información.

Los diferentes tipos de amenazas que enfrentan las empresas, tales como, robo de información, ciberataques, plagios, phishing, entre otros, ha puesto al descubierto la vulnerabilidad de ciertos departamentos de la fábrica de bebidas de consumo masivo. En años anteriores la falta de herramientas, técnicas o políticas en la ciberseguridad ha causado que exista filtración de información fuera de la fábrica causando pérdidas de clientes esenciales que aportan significativamente en el crecimiento de la fábrica. A raíz de este robo, se ha registrado que existe un decremento del -10,2% en sus ventas en el mes de septiembre del año 2024, en comparación con los dos años anteriores situados en el mismo mes. Este decremento ha afectado significativamente en el departamento de ventas a cumplir sus objetivos como equipo de trabajo.

Adicionalmente, a pesar de que la fábrica cuenta con campañas de concientización y capacitaciones a todo nivel de su estructura organizacional, es necesario contar con técnicas de anonimización. La ausencia de estas

técnicas podría dar como resultado en la pérdida de cliente, extorsión, deficiencia en el departamento de ventas y por consiguiente pérdidas económicas significativas en la fábrica de bebidas de consumo masivo.

1.2 Descripción del problema

En respuesta a esta problemática, la presente propuesta está enfocada en la implementación de una de las técnicas del sistema de anonimización para los datos sensibles administrados por el departamento de ventas de la fábrica de bebidas de consumo masivo. Esta fábrica está expuesta a varias vulnerabilidades, una de las cuales y la más común, es la fuga de datos. Esta fuga se da ya sea por una persona externa o interna sin autorización y toma información sensible, como datos de clientes. En el estudio de Morgan et al. [1], se señala que una fuga de datos podría costar más de cientos de millones de dólares en una empresa, por lo que las empresas deben hacer todo lo posible por tener las debidas protecciones.

La implementación adecuada de una de las técnicas de anonimización, asegurará la confianza en el equipo de trabajo, clientes y aliados estratégicos. La Ley Orgánica de Protección de Datos en Ecuador [2] indica qué se debe asegurar y ofrecer mecanismos suficientes para garantizar el derecho a la protección de datos personales. Es por ello que un sistema de anonimización ayudará a reducir el riesgo de identificación a una persona, tal como lo describe Demir et al. [3], mitigando así los riesgos que afectan a la privacidad de las personas.

Esta propuesta es viable ya que el equipo cuenta con apertura en la fábrica de bebidas de consumo masivo que permitirá implementar la técnica de

anonimización en un periodo no superior a 4 meses. Esto facilitará su difusión e interés entre futuros desarrolladores dentro del departamento de sistemas.

1.3 Solución Propuesta

La solución que se propone es la implementación de una de las técnicas de anonimización, previo a su debido análisis. Esta técnica ayudará a la protección de datos personales administrados por el departamento de ventas de una fábrica de bebidas.

En este proyecto, se analizará 4 técnicas de anonimización basados en el estudio de H. Tahir et al [4]:

1. Reemplazo de datos originales por datos no comunes
2. Conservación de solo datos necesarios
3. Sustitución
4. Ahogamiento de datos mediante la generalización

El análisis que se realizará a las cuatro técnicas será de manera comparativa. Se centrará en evaluar los siguientes criterios: la naturaleza de los datos, nivel de protección requerido, compatibilidad con la legislación vigente, costos y recursos requeridos para la anonimización. El complimiento y evaluación de estos criterios permitirá identificar la técnica adecuada para una protección a los datos personales de los clientes.

Una vez seleccionada e implementada de una de las cuatro técnicas de anonimización ayudará a que, de existir fuga de datos en el departamento de ventas, la información no será descifrable.

En la figura1.3, se describe el flujo de trabajo con la técnica de anonimización seleccionada. El proceso inicia con un archivo excel que contiene información sensible, luego pasa por una técnica de anonimización para así ocultar o modificar la información sensible. Finalmente, se obtiene la información anonimizada asegurando así la privacidad de los clientes.

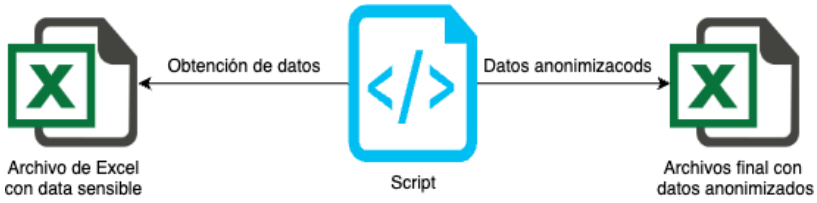


Figura 1.3 Esquema básico de la solución

Fuente: Autor

Esta solución da cumplimiento al uso correcto sobre el derecho de la protección de datos que exige la Ley Orgánica de Protección de Datos [2]. Así mismo, la implementación de una técnica de anonimización, ayudará a que se cumplan estándares internacionales, tales como la ISO 27001 y la ISO 27002.

La implementación de una técnica de anonimización, da mejor utilidad a los datos, según lo demuestra el estudio de Pawar et al [5], debido a estas técnicas permitirá realizar análisis estadístico sin comprometer la privacidad de las personas.

1.4 Objetivo General

Implementar un sistema de anonimización, basado en las normativas ecuatorianas vigentes sobre la protección de datos personales, luego del análisis de las técnicas existentes que garantizará la privacidad y seguridad de la información de los datos personales de los clientes y proveedores del departamento de ventas.

1.5 Objetivos Específicos

1) Analizar las diferentes técnicas de anonimización disponibles para el departamento de ventas de la fábrica objetivo.
2) Determinar las políticas necesarias para desarrollar la técnica seleccionada.
3) Evaluar el marco de políticas establecidas para determinar la efectividad del sistema de anonimización mediante pruebas, haciendo uso de la técnica seleccionada en un entorno controlado, asegurando que la información no pueda ser re-identificada.
4) Difundir el marco político de la técnica implementada de tal manera que sirva de base en su desarrollo para el departamento de ventas.

1.6 Metodología

El presente trabajo de titulación se realizará una revisión exhaustiva sobre técnicas de anonimización, sus aplicaciones y casos de estudios relevantes haciendo uso de políticas y normas vigentes. El diseño de este proyecto es un estudio de tipo transversal, debido a que se le pedirá al departamento de sistemas, un conjunto de datos de clientes de la fábrica de bebidas en un determinado rango de tiempo.

Este trabajo se enfocará en analizar datos de clientes, dentro de los últimos 5 años, administrados por el departamento de ventas. Se realizará una muestra de tipo no probabilística por conveniencia. Esto es posible debido a qué se tiene acceso a estos datos y nuestro interés sobre esta muestra tendrá como grupo objetivo clientes estratégicos, tales como centros de distribución, supermercados y bodegas. Estos datos tendrán las siguientes características: aportación económica, historial de compra, datos financieros y datos personales.

Se usará dos tipos de estudio para la obtención de los datos, primero se aplicará una encuesta a un grupo objetivo de ciento setenta y cuatro personas para evaluar si los campos que se determinaron como sensibles están correctos. Adicional, se usará el levantamiento de información documental para establecer un conjunto de datos sensibles de los clientes estratégicos.

Luego del levantamiento de datos, se va a analizar la sensibilidad de la información, segmentarla por categorías, tipos de datos e historial de compras (a nivel monetario), de tal manera que permita identificar la criticidad de cada uno de ellos. En consecuencia, luego de este análisis, se evaluará la aplicación de una de las técnicas del sistema de anonimización.

Capítulo 2

Marco Teórico

En este capítulo se presenta el marco teórico que sustenta la investigación, abordando temas cruciales para la comprensión y aplicación de 4 técnicas de anonimización y respectivo concepto basado en el contexto de la protección de datos. Primero, se examina la Ley de Protección de Datos del Ecuador, proporcionando un panorama claro de las regulaciones y obligaciones legales que rigen el manejo de información personal y sensible en Ecuador. A continuación, se revisan las normativas aplicables ISO 27001 y 27002, que proporcionan estándares internacionales para la gestión de la seguridad de la información, asegurando un enfoque integral y conforme a las mejores prácticas globales. Asimismo, se exploran las técnicas de anonimización, detallando los métodos utilizados para proteger la identidad de los individuos en los conjuntos de datos. Se analizan las ventajas y desventajas de estas técnicas, destacando sus fortalezas y limitaciones en diferentes escenarios para su posterior selección y aplicación. Finalmente, se establecen los criterios de selección de técnicas de anonimización, ofreciendo una guía para elegir la metodología más adecuada según la naturaleza de los datos, nivel de protección requerido, compatibilidad con la legislación vigente y recursos requeridos para la anonimización.

2.1 Ley Orgánica de Protección de Datos en Ecuador y su relación con los datos sensibles

La Constitución de Ecuador protege los datos personales e introduce responsabilidades complejas y severas para las organizaciones. Los ciudadanos tienen derechos sobre sus datos personales que necesitan

protección especial. Las entidades privadas y públicas deben asumir y demostrar responsabilidades de protección de esta información personal. Un planteamiento sólido hacia la seguridad y protección de datos es la creación y mantenimiento de programas de seguridad que sean integrales, dinámicos y en evolución. La ausencia de un marco legal nacional de protección de datos personales representa un riesgo tanto para los ciudadanos como para el país.

El actual contexto ha llevado a plantear el verdadero valor de los datos personales, considerando las enormes aportaciones que se realizan en la búsqueda de soluciones a problemas cotidianos. En el ámbito legal, Ecuador emitió su Ley Orgánica de Protección de Datos Personales [2], cuya finalidad consiste en establecer las garantías al tratamiento de datos de carácter personal. Esto ha sido derivado de una creciente problemática respecto a la administración y procesamiento ilícito de datos por la constante digitalización de los sectores productivos.

Es relevante saber que existen datos sensibles y no sensibles, ya que las regulaciones respecto a su tratamiento son sustancialmente diferentes. Los datos sensibles son aquellos datos que, por su naturaleza íntima o reserva, solamente sean obtenibles en virtud de una autorización expresa y que faciliten información atinente al ámbito personal y a las garantías que deben reconocérsele a los titulares de estos. Mientras que los datos no sensibles son aquellos que no revelan información especialmente protegida sobre una persona.

Se debe considerar que, en ningún caso, basándose en el tipo de dato se podrá llevar a cabo decisiones para el usuario basadas únicamente en el tratamiento automatizado, que produzcan efectos jurídicos sobre su esfera personal o que le afecten negativamente, salvo que la autoridad independiente de control así se manifieste y lo indique la ley.[4]

2.2 Definición y técnicas de anonimización: Tipos, ventajas y desventajas

2.2.1 Definición

La anonimización es el proceso mediante el cual permite desvincular la relación que existe de una persona con sus datos personales, tal como lo definen varios autores [4], [5], [6]. La anonimización tiene como principio que la reidentificación sea imposible y su objetivo principal es que se garantice la protección de datos personales. En este contexto, los datos personales pueden ser compartida con investigaciones científicas o en la medicina, la anonimización debe garantizar la protección de los datos para cumplir marcos legales.

2.2.2 Técnicas de anonimización: Tipos, ventajas y desventajas

En la actualidad, las bases de datos de las empresas contienen información de datos personales. Estos datos sensibles pueden estar en entornos de producción o de pruebas aumentando así la probabilidad de robo, pérdida o divulgación. Las técnicas de anonimización ayudan a que los datos sensibles estén protegidos bajo amenazas. De acuerdo con el estudio de A. Pawar et al [4], la implementación de estas técnicas de anonimización ayudan a evitar

ataques como: Ataque de homogeneidad, ataque de similitud, ataque de conocimiento de fondo y ataque de inferencia probabilística.

A continuación, en la tabla 1 se presentan algunos métodos de anonimización y una descripción.

Tabla 1. Métodos de anonimización

Método de anonimización	Descripción
Supresión	Este método elimina atributos o registros que contienen datos sensibles.
Reemplazado de caracteres	Este método realiza el reemplazo de caracteres con símbolos establecidos para ocultar parcialmente información.
Aleatorización	Este método de manera aleatoria reorganiza los valores para dificultar la asociación de registros.
Adición de ruido	Este método modifica los registros originales con datos adicionales.
Generalización	Este método cambia registros específicos por información más general.
K-Anonymity	Este método agrupa registros que contienen al menos “k” combinaciones simulares.

L-Diversity	Este método es la extensión de k-anonymity en el cual cada grupo contiene al menos “l” valores distintos para un registro sensible.
Sustitución	Este método reemplaza los registros originales por valores alternativos.
Eliminación de datos útiles	Este método elimina información sensible que no requiera un análisis o intercambio específico.
Uso de datos sintéticos	Este método genera datos ficticios reemplazando a los datos originales.

Por consiguiente, este trabajo de titulación se enfoca en estos cuatro tipos de anonimización:

1. Reemplazo de datos originales por datos no comunes

El reemplazo de datos originales por datos no comunes consiste en reemplazar los datos con los que se trabaja por otros que no guarden relación alguna con los verdaderos, para evitar, de este modo, la identificación de las personas [5]. La ventaja más evidente de esta técnica es que los datos no conservan ninguna información sensible o directa, por lo que son completamente seguros desde el punto de vista de la confidencialidad. Sin embargo, la principal desventaja de utilizar esta técnica es que, a la hora de trabajar con información sintética, ya no se puede asegurar que los resultados son de

aplicación válida a la realidad original. Tampoco se pueden responder a preguntas individuales sobre el comportamiento de personas o entidades concretas en el tiempo.

2. Conservación de solo datos necesarios

Esta técnica consiste en el análisis exhaustivo y detallado de cuáles son los datos realmente necesarios para conservar en la organización, con el objetivo primordial de evitar tener que compartir información sensible y confidencial [8]. Mediante la implementación de esta técnica, logramos alcanzar altos niveles de eficiencia y economía, ya que se simplifica enormemente el tratamiento y gestión de los datos, reduciendo así los costes asociados al almacenamiento, copia de seguridad, protección, gestión y explotación de estos.
No obstante, es importante tener en cuenta algunas desventajas que pueden surgir. En primer lugar, a medida que establecemos umbrales más restrictivos para la retención de datos, nos encontramos con un menor índice de aprovechamiento de estos. Esto implica que la utilidad de los datos se ve cuantificada en función de un umbral determinado. Es decir, cuantos más atributos descartemos, mayor será el riesgo de conservar datos que podrían resultar perjudiciales o incluso perder la utilidad de aquellos datos que son realmente importantes para la organización.

3. Sustitución

Esta técnica de anonimización consiste en reemplazar, con un valor aleatorio o con un diccionario de palabras los valores originales.

Varios autores muestran que una de las ventajas de esta técnica es su sencillez y fácil implementación [5], [8], [9]. Debido a esta ventaja, es posible de implementar directamente desde una consulta SQL. Sin embargo, la conservación del formato puede indicar posibles semejanzas con sus valores originales, siendo esta característica una desventaja.

4. Ahogamiento de datos mediante la generalización

Esta técnica de anonimización consiste en eliminar cierta información y de generalizar información de la persona. En el estudio de J. F. Marques y J. Bernardino [7], describe dos técnicas consideradas de tipo generalización: K-Anonymity y la L-Diversity. La técnica K-Anonymity, agrupa registros de K individuos en categorías con una misma combinación de atributos.

Por otro lado, la técnica L-Diversity es una evolución de la K-Anoymity y exige al menos L valores distintos para cada grupo equivalente en los atributos sensibles. Ambas técnicas tienen la ventaja de anonimizar información sensible con distintos enfoques. Sin embargo, estas técnicas son susceptibles a ataques de tipo de inferencia probabilística.

2.3 Normativas de aplicación de políticas basado en la norma ISO 27001 e ISO 27002

La ISO 27001 [10], es una norma internacional centrada en la gestión de seguridad de la información. Estas normas, establecen prácticas y requisitos para proteger la confidencialidad, integridad y disponibilidad de los datos.

Ahora bien, la ISO 27001 y la ISO 27002 se complementan de tal forma que ambas ayudan a establecer políticas para la protección de los datos personales. La implementación de ambas normas ayudará a las organizaciones para crear, ajustar o modificar procesos con la finalidad de reducir el riesgo de exposición de datos personales.

Dentro de este marco, los controles de la ISO 27001 como la clasificación, la manipulación y controles criptográficos aportan una protección para la anonimización de datos. Una clasificación adecuada permite priorizar datos sensibles que requieren mayor resguardo. Un conjunto adecuado de procedimientos ayudará que la manipulación de datos este dentro de los esquemas adoptados por la organización. Desarrollar e implementar controles criptográficos ayudará a proteger los datos de la empresa. Por su parte, la ISO 27002 [11], nos ofrece recomendaciones y guías de cómo podemos realizar una anonimización a la información de datos de una organización.

Como ya se expuso, la ISO 27002 ofrece recomendaciones y guías para realizar la anonimización de los datos. Sin embargo, estas implementaciones pueden conllevar a retos dentro de la organización, tales como elegir la técnica adecuada, costos de implementación, mantenimiento y adaptación a regulaciones cambiantes. Para alcanzar a superar estos retos, es importante el compromiso de la organización, una formación continua, recursos adecuados y adaptación a las normativas cambiantes.

2.4 Criterios de selección de técnicas de anonimización

A continuación, se detallan los criterios para la selección de la técnica de anonimización:

2.4.1 Naturaleza de los datos

El concepto de la naturaleza de los datos se refiere a las características intrínsecas de los propios datos, sin considerar la información que estos proporcionan. En este sentido, la aplicación de técnicas de anonimización o supresión de elementos específicos dependerá del tipo de datos con los que se esté trabajando. Es importante destacar que los datos de interés son confidenciales, por lo que el responsable del análisis debe determinar con precisión en qué momento un dato deja de ser anónimo. Este concepto de anonimato está estrechamente relacionado con la naturaleza de los datos y su comprensión es fundamental para garantizar la privacidad y seguridad en su utilización. Es además, es crucial considerar los diferentes contextos en los que se manejan los datos y la necesidad de adaptar las técnicas de anonimización necesario tener en cuenta las implicaciones éticas y legales que rodean el tratamiento de datos confidenciales y la importancia de implementar medidas adecuadas de protecciónonimización según cada caso específico. En conclusión, comprender la naturaleza de los datos y aplicar técnicas adecuadas de anonimización es esencial para salvaguardar la privacidad y los derechos de las personas en el manejo de la información. [12], [13], [14]

2.4.2 Nivel de protección requerido

El principal desafío al identificar niveles de protección digital de datos radica en los diversos perfiles de usuarios que deben ser protegidos, los cuales varían desde lo general hasta lo particular. En este contexto, podemos dividir a los usuarios en dos grupos. El primer grupo incluye a aquellos que generalmente requieren un mayor esfuerzo para el tratamiento de sus datos y el cumplimiento de todas las leyes vigentes, garantizando así la seguridad de la información a lo largo de toda la cadena de suministro. Por otro lado, el segundo grupo está compuesto por particulares, quienes manejan información mucho más sensible y son los principales destinatarios de las leyes de protección de datos. Debido a las diferencias entre estos dos grupos y la variabilidad en sus perfiles y escalas, es posible establecer un nivel de protección específico para las empresas. [15], [16], [17], [18]

2.4.3 Compatibilidad con la legislación vigente

Varios países ya cuentan con normas o leyes que ayudan a regular cómo se da tratamiento a los datos personales. El Reglamento General de Protección de Datos (GDPR) de la Unión Europea, establece un marco legal para la recopilación, procesamiento y almacenamiento de datos personales a los ciudadanos de la Unión Europea. En Ecuador, desde el 2021, se tiene una Ley Orgánica de Protección de Datos Personales (LOPDP), la cual se basa en

normativas europeas. En ambos marcos legales, se puede encontrar que la anonimización es un método efectivo para ayudar a cumplir con obligaciones legales.

De acuerdo con el estudio de Britton et al. [19], al aplicar anonimización, la información ya no se considerará como dato personal, por lo cual el GDPR o alguna otra ley ya no se aplicarían. De tal manera, la ley ecuatoriana establece que datos anonimizados no se consideran datos personales, siempre que se garantice que no puedan ser reidentificados.

Dentro de este contexto, es importante resaltar que, si los datos anonimizados se les aplica técnicas para recuperar la identificación del titular, esta información está sujeta a obligaciones establecidas tanto por la GDPR o de la Ley Orgánica de Protección de Datos en Ecuador.

2.4.4 Costos y recursos requeridos para la anonimización

En un proceso de anonimización se requiere una inversión de tiempo, recursos y de conocimientos especializados, los cuales garantizarán la protección de los datos personales. La implementación de las técnicas de anonimización, puede requerir, además, inversiones en infraestructura tecnológicas y de software especializado. De acuerdo con el estudio de Ni et al. [20], cada técnica de anonimización pueden variar en su complejidad y eficiencia, lo que conlleva afectar costos operativos.

Cada técnica de anonimización tendrá su consumo de recursos computacionales y es importante evaluarlos para que así, no existan interrupciones operacionales o gastos adicionales para el departamento de IT, a continuación, se agrupan las técnicas y criterios en la gráfica 2.1.

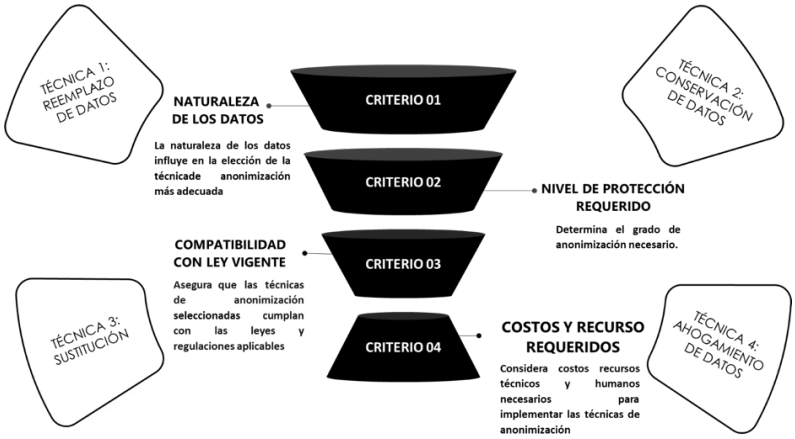


Figura 2.1 Gráfico de criterios de selección y las técnicas

Fuente: Autor

En conclusión, el marco teórico presentado proporciona una base sólida para comprender y aplicar técnicas de anonimización en el contexto de la protección de datos. La Ley de Protección de Datos del Ecuador establece el marco legal necesario para garantizar la privacidad y seguridad de la información personal, subrayando la importancia de cumplir con las regulaciones vigentes. Las técnicas de anonimización ofrecen diversas metodologías para proteger la identidad de los individuos, cada una con sus propias ventajas y desventajas, lo que permite seleccionar la técnica más adecuada según las necesidades específicas.

Los criterios de selección de técnicas de anonimización proporcionan una guía práctica para evaluar y elegir las técnicas más efectivas de acuerdo con cada caso de estudio. Por tanto, este marco teórico facilita la comprensión de los conceptos claves, proporciona una guía práctica para la implementación efectiva de técnicas de anonimización y asegura el cumplimiento legal y la protección adecuada de los datos personales.

Capítulo 3

Elección y Desarrollo de Técnica de Anonimización dentro del Marco De Políticas Establecida

En esta etapa, se procederá a realizar una propuesta de diversas maneras de obtener una base de datos que sea totalmente anónima y segura. Se prestará especial atención al resultado de las técnicas empleadas en la base de datos, al mismo tiempo que se identificarán las funciones reconocidas en las diferentes normativas legales que se han mencionado en los capítulos anteriores. En la selección de la técnica de anonimización más adecuada para los datos que se manejan, se han establecido varios criterios que resultan altamente relevantes y prioritarios. En primer lugar, resalta la naturaleza de los datos, que sea altamente robusta y la facilidad de implementación, aspectos que son esenciales para garantizar que la técnica seleccionada se aplique de forma efectiva. También es esencial asegurarse de que la técnica de anonimización no esté sujeta a restricciones por parte de los facilitadores de los datos, lo que permite su implementación y uso efectivo en diversas aplicaciones.

En un análisis técnico mediante la implementación, usando código de programación con las diferentes técnicas de forma dinámica y obteniendo datos estadísticos, se logró determinar que con la técnica de sustitución se dieron tiempos reducidos de anonimización, cambio completo de caracteres de los datos anonimizados en varios archivos aleatorios y conservación de datos originales seleccionados para otras aplicaciones.

3.1 Elección de la técnica

Explicación de diagrama de flujo

1. Inicio:
    - Comienzo del proceso de selección de la técnica de anonimización de las cuatro previamente seleccionadas en el capítulo anterior.
2. Naturaleza de los datos:
    - Identificación de los datos: Clasifica los datos según su tipo (personales, sensibles, financieros, etc.).
    - Evaluación de riesgos: Si los datos son personales, evalúa los riesgos asociados con su exposición y su sensibilidad.
3. Nivel de protección:
    - Requisitos de seguridad: Establece los niveles de protección necesarios según la clasificación de los datos.
    - Selección de técnicas: Identifica si requiere alta protección, y selecciona técnicas avanzadas de anonimización. Si no, selecciona técnicas básicas y se descarta.
4. Compatibilidad con la ley de protección de datos vigentes y las normas 27001 y 27002:
    - Revisión legal: Verifica las leyes y regulaciones aplicables (GDPR, CCPA, etc.).
    - Cumplimiento: Asegura de que la técnica de anonimización seleccionada cumpla con las normativas legales estudiadas en el capítulo anterior.
5. Costos y recursos para la implementación:

- Análisis de costos: Evalúa los costos asociados con la implementación de la técnica de anonimización.
- Recursos necesarios: Identifica los recursos técnicos y humanos necesarios para llevar a cabo la implementación.

6. Selección de técnica de anonimización:
   - Si los costos son aceptables y los recursos están disponibles, implementa la técnica de anonimización.
   - Si no, revisa los recursos y ajusta el plan según sea necesario.

Considerando cada uno de los criterios, se determina que la técnica de anonimización a usar es la de sustitución debido de que es la más adecuada en la implementación de este proyecto, ya que establece que los datos son altamente sensibles, requiere un alto nivel de protección y además deben cumplir con estrictas leyes de protección de datos, lo cual puede justificar una inversión a gran escala en esta técnica avanzada tal como se demuestra en la figura 3.1:

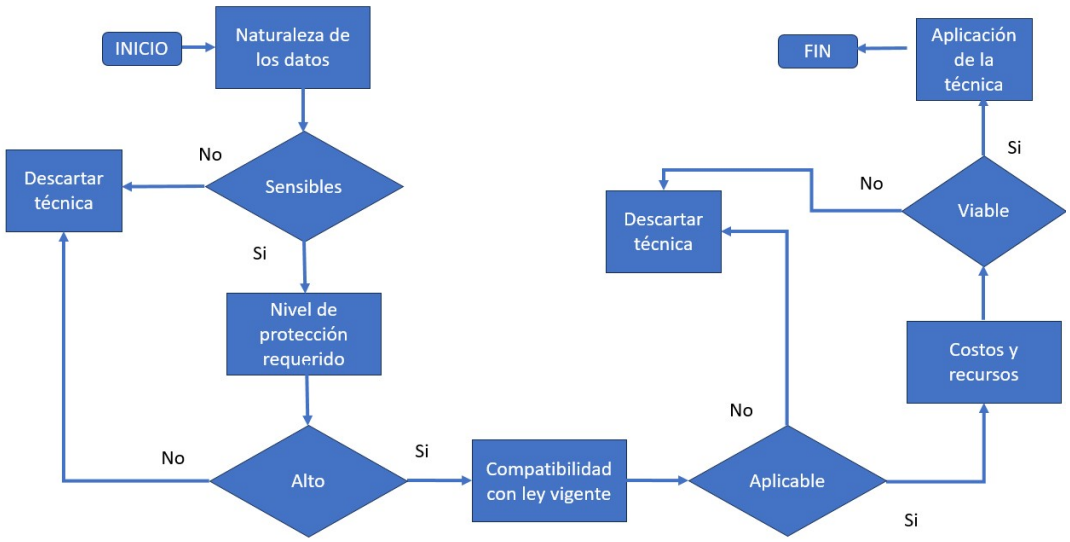


Figura 3.1 Flujograma de criterios para selección de técnica

Fuente: Autor

3.2 Plan de implementación de técnica seleccionada

La técnica sustitución, ayudará a la conservación de la información y garantiza la privacidad de la información de los clientes y proveedores de la fábrica de bebidas de consumo masivo.

Dentro de este contexto, la implementación de la técnica sustitución sigue los siguientes pasos:

3.2.1 Campos elegidos para anonimización

Para la selección de los datos en los campos a anonimizar se realiza una encuesta a ciento setenta y cuatro personas dentro del departamento de sistemas y ventas de la fábrica de consumo masivo de bebidas.

La encuesta se preparó con veinte preguntas, en las cuales se formula la consulta de si el campo contiene información sensible. Todas las preguntas tenían solo dos opciones de respuesta, “Sí” o “No”.

Aquellas preguntas que tuvieron una puntuación de la opción “Sí” mayor o igual al setenta por ciento, se considera que aquellos campos contienen datos sensibles.

A continuación, se muestran en la tabla 2 los campos con su respectiva puntuación:

Tabla 2. Resumen de la encuesta

Campos	Sí	No
CLIENTE_SAP	13,8%	86,2%
COD_REGIONAL	14,9%	85,1%
NOMBRE_DIRECCION	67,8%	32,2%
NOMBRE_GERENCIA_GENERAL	59,2%	40,8%
COD_GERENCIA_VENTAS	44,8%	55,2%
ZONA_DE_VENTAS	66,7%	33,3%
NOMBRE_GERENCIA_VENTAS	34,5%	65,5%
NOMBRE_1	74,1%	25,9%

NOMBRE_3	73%	27%
IDENTIFICACION	75,5%	24,7%
PROVINCIA	66,1%	33,9%
POBLACION	20,7%	79,3%
CALLE_3	70,7%	29,3%
CALLE_Y_NO	71,3%	28,7%
CALLE_4	70,7%	29,3%
TELEFULL	73,6%	26,4%
LATITUD	67,2%	32,8%
LONGITUD	68,4%	31,6%
NOM_DISTRIBUIDOR	69%	31%
CORREO	70,7%	29,3%

Por lo tanto, los siguientes campos: NOMBRE_1, NOMBRE_3, IDENTIFICACION, CALLE_3, CALLE_Y_NO, CALLE_4, TELEFULL y CORREO son considerados con información sensible y adecuados para aplicar la técnica de anonimización sustitución.

En el Anexo A, se puede encontrar la encuesta realizada a las ciento setenta y cuatro personas.

3.2.2 Elección de información y uso de lenguaje de programación para la técnica seleccionada

La fábrica de bebidas de consumo masivo facilita un archivo en formato Excel con información de sus clientes. Este archivo de Excel cuenta con un total de 10,000 filas y 20 columnas.

El código para la implementación de la técnica de sustitución fue NodeJS. NodeJS es un lenguaje de programación multiplataforma, de código abierto y está basado en el lenguaje de programación JavaScript.

Una de las ventajas que tiene NodeJS es la capacidad de poder manejar grandes volúmenes de datos. Así cómo también es de fácil adopción ya que proviene de un lenguaje conocido que es JavaScript.

3.3 Aplicación de la técnica

Para la aplicación de la técnica de anonimización se utiliza un script con el lenguaje de programación NodeJS usando la versión 20. El script usa una librería llamada Faker. Faker genera datos falsos con el objetivo de realizar de manera aleatoria la sustitución de datos reales con datos ficticios.

El script se puede ejecutar enviándole un parámetro de cuantos registros de manera aleatoria se necesita anonimizar. El script abre el archivo Excel usando la librería xlsx y con el valor pasado como parámetro, obtiene la cantidad de registros para su respectiva anonimización. Finalmente se genera un archivo en Excel con la información anonimizada usando la librería xlsx, como se muestra en la figura 3.3.

En el Anexo B se puede encontrar el código implementado que se utilizó para anonimizar la información sensible.

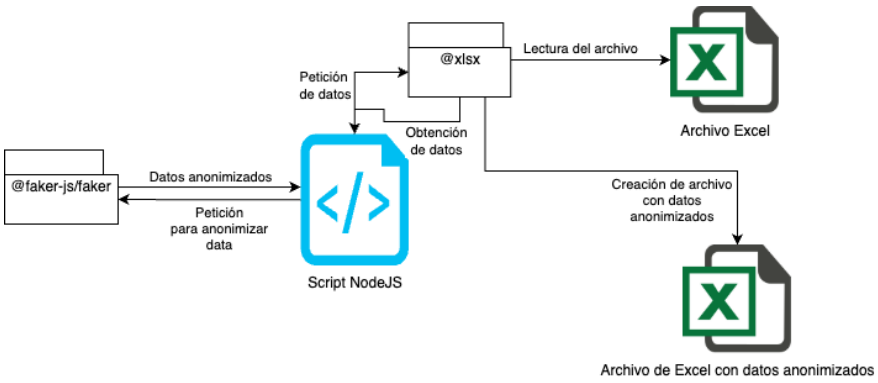


Figura 3.3 Esquema del plan de implementación.

Fuente: Autor

3.4 Resultados de la técnica de anonimización

Luego de aplicarse la técnica de sustitución, se observa en la siguiente figura 3.4, que se ha reducido su relación o vínculo hacia un cliente o proveedor.

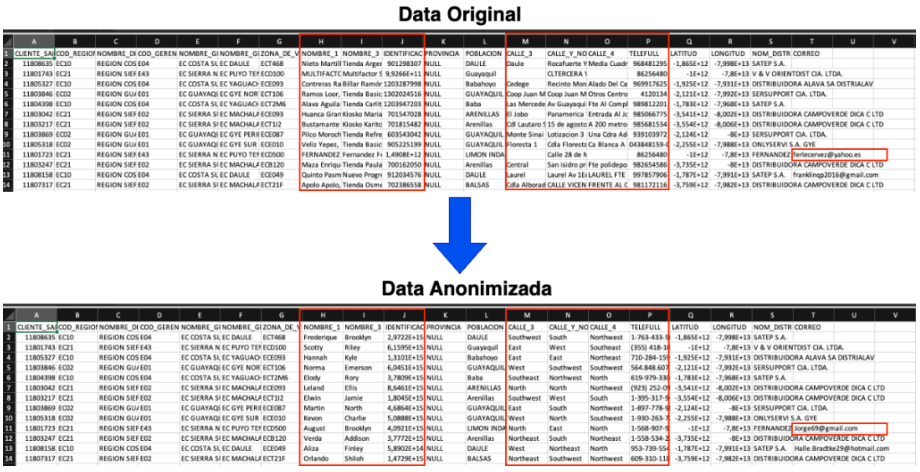


Figura 3.4 Resultados de la anonimización

Fuente: Autor

Se realizaron dos tipos de pruebas, la primera prueba fue tomando un archivo de 10,000 registros y la segunda prueba fue tomando múltiples archivos. Dentro de la primera prueba, se sacó un promedio de tres tomas de tiempo de ejecución del código tomando muestras aleatorias de 3,000, 6,000 y 9,000 registros. A continuación, se presenta la figura 3.5, dónde se ve un gráfico de dispersión.

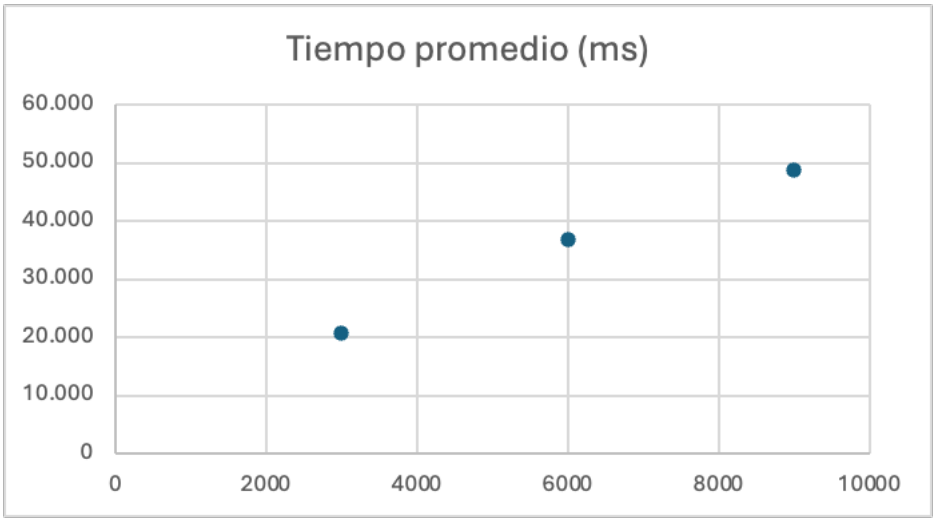


Figura 3.5 Gráfico de dispersión de tiempos de ejecución

Fuente: Autor

La segunda prueba se sacó un promedio de tres tomas de tiempo usado cinco archivos. Cada archivo contenía 2,000 registros. La ejecución del código tomo por cada archivo, 1,500 registros de manera aleatoria. A continuación, se presenta la figura 3.6, dónde se ve un gráfico de los resultados de los tiempos en milisegundos.

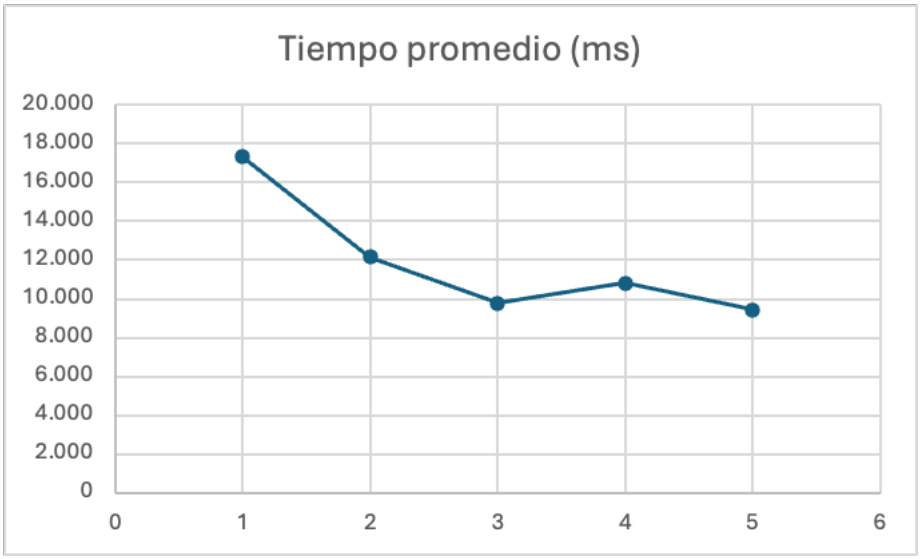


Figura 3.6 Gráfico lineal de tiempos de ejecución.

Fuente: Autor

En la figura 3.6, se observa que el tiempo es directamente proporcional a la cantidad de archivos para ser procesados.

3.5 Ejecución de código usando diferentes técnicas

El código creado, esta implementado de tal manera que puede aceptar múltiples archivos en formato Excel y ser dinámico en el uso de la técnica establecidas en este proyecto.

La manera de ejecutar el código se presenta a continuación:

node index --tecnique=substitution --totalRows=1500, donde el parámetro --tecnique puede tener un único valor al momento de ejecutarse, como se presenta en la figura 3.7. El parámetro --totalRows indicará la cantidad de registros que se necesita anonimizar de manera aleatoria.

```
switch( techniqueChoosen ) {
  case 'replace':
    anonymizedData = await replaceDataPromise( techniqueChoosen, suffixFileName, randomRows );
    break;
  case 'keep':
    anonymizedData = await keepUsefullDataPromise( techniqueChoosen, suffixFileName, randomRows );
    break;
  case 'drowing':
    anonymizedData = await drowingDataPromise( techniqueChoosen, suffixFileName, randomRows );
    break;
  case 'substitution':
    anonymizedData = await sustitutionPromise( techniqueChoosen, suffixFileName, randomRows );
    break;
  default:
    console.log('Please choose a valid technique: replace, keep, drowing or substitution');
    break;
}
```

Figura 3.7 Switch de técnicas implementadas

3.5.1 Registro de pruebas usando diferentes técnicas de anonimización

Se registró una prueba utilizando las cuatro técnicas de anonimización seleccionadas, haciendo uso del mismo formato de la primera prueba redactadas en la sección 3.4.

A continuación, se presentan los resultados en la siguiente tabla.

Tabla 3 Resultado de tiempos usando diferentes técnicas

Técnica	Carga archivo (ms)	Anonimización (ms)	Exportación (ms)	Total, de registros
Reemplazo de datos	502.940	11.071	84.176	3000
Conservación de datos	507.930	0	47.259	
Sustitución	502.070	20.676	79.208	

Ahogamiento de datos	502.993	19.564	94.385	
Reemplazo de datos	517.492	21.050	175.655	6000
Conservación de datos	526.557	1.433	93.637	
Sustitución	511.834	36.772	152.070	
Ahogamiento de datos	499.720	31.964	155.414	
Reemplazo de datos	519.493	31.265	268.502	9000
Conservación de datos	508.059	1.773	149.014	
Sustitución	500.638	48.874	239.250	
Ahogamiento de datos	510.917	49.401	247.063	

Capítulo 4

Difusión de Marco Político de la Técnica Seleccionada

4.1 Desarrollo de marco político

Mediante este marco de políticas a establecer, la fábrica de bebidas de consumo masivo obtendrá una capa adicional sobre la información sensible que maneja de sus clientes y proveedores.

Alcance

Este marco de políticas de protección de datos personales se aplicará a todas las bases de datos y/o archivos que contengan datos personales que sean objeto de tratamiento por parte de los departamentos y grupos de ventas y sistemas de la fábrica de bebidas, cada una individualmente considerada como responsable del tratamiento de datos personales.

Como se determina y se describe en el capítulo 3, la técnica de anonimización usando el método de sustitución, cumple con la ley de protección de datos vigente y con las normas ISO 27001 e ISO 27002.

Definiciones para implementación de marco de políticas

- Autorización: consentimiento previo, expreso e informado del titular para llevar a cabo el tratamiento de datos personales.
- Aviso de privacidad: comunicación verbal o escrita generada por el responsable, dirigida al titular para el tratamiento de sus datos personales, mediante la cual se le informa acerca de la existencia de las políticas de tratamiento de información que le serán aplicables, la

forma de acceder a las mismas y las finalidades del tratamiento que se pretende dar a los datos personales.

- Base de datos: conjunto organizado de datos personales que sea objeto de tratamiento.
- Clientes: persona natural o jurídica, pública o privada con los cuales la empresa tiene una relación comercial. Comprende las tiendas, supermercados, minimercados, entre otros.
- Consumidores: persona natural que consume los bienes producidos por la empresa.
- Dato personal: cualquier información vinculada o que pueda asociarse a una o varias personas naturales determinadas o determinables. Algunos ejemplos de datos personales los siguientes: nombre, cédula de ciudadanía, dirección, correo electrónico, número telefónico, estado civil, datos de salud, huella dactilar, salario, bienes, estados financieros, etc.
- Dato sensible: información que afecta la intimidad del titular o cuyo uso indebido puede generar su discriminación, tales como aquellos que revelen el origen racial o étnico, la orientación política, las convicciones religiosas o filosóficas, la pertenencia a sindicatos, organizaciones sociales, de derechos humanos o que promueva intereses de cualquier partido político o que garanticen los derechos y garantías de partidos políticos de oposición así como los datos relativos a la salud, a la vida sexual y los datos biométricos, entre

otros, la captura de imagen fija o en movimiento, huellas digitales, fotografías, iris, reconocimiento de voz, facial o de palma de mano, etc. sin embargo los datos personales también pueden considerarse como sensibles.

- Encargado del tratamiento: persona natural o jurídica, pública o privada, que por sí misma o en asocio con otros, realice el tratamiento de datos personales por cuenta del responsable del tratamiento. en los eventos en que el responsable no ejerza como encargado de la base de datos, se identificará expresamente quién será el encargado.
- Reclamo: solicitud del titular del dato o de las personas autorizadas por éste o por la ley para corregir, actualizar o suprimir sus datos personales o para revocar la autorización en los casos establecidos en la ley.
- Términos y condiciones: marco general en el cual se establecen las condiciones para los participantes de actividades promocionales o afines.
- Titular: persona natural cuyos datos personales sean objeto de tratamiento.
- Transferencia: la transferencia de datos tiene lugar cuando el responsable y/o encargado del tratamiento de datos personales, ubicado en Ecuador, envía la información o los datos personales a un receptor, que a su vez es responsable del tratamiento y se encuentra dentro o fuera del país.

- Transmisión: tratamiento de datos personales que implica la comunicación de estos dentro o fuera del territorio de la república de Ecuador cuando tenga por objeto la realización de un tratamiento por el encargado por cuenta del responsable.

- Tratamiento: cualquier operación o conjunto de operaciones sobre datos personales, tales como la recolección, almacenamiento, uso, circulación o supresión.

4.2 Principios aplicables al tratamiento de datos personales

Para el tratamiento de los datos personales, la empresa aplicará los principios que se mencionan a continuación, los cuales constituyen las reglas a seguir en la recolección, manejo, uso, tratamiento, almacenamiento e intercambio, de datos personales:

- Legalidad: el tratamiento de datos personales se realizará conforme a las disposiciones legales aplicables (ley de protección de datos del ecuador y las normas ISO 27001 e ISO 27002).

- Finalidad: los datos personales recolectados serán utilizados para un propósito específico y explícito el cual debe ser informado al titular o permitido por la ley. El titular será informado de manera clara, suficiente y previa acerca de la finalidad de la información suministrada.

- Libertad: la recolección de los datos personales solo podrá ejercerse con la autorización, previa, expresa e informada del titular.

- Veracidad o calidad: la información sujeta al tratamiento de datos personales debe ser veraz, completa, exacta, actualizada, comprobable y comprensible.
- Transparencia: en el tratamiento de datos personales se garantiza el derecho del titular a obtener en cualquier momento y sin restricciones, información acerca de la existencia de datos que le conciernan.
- Acceso y circulación restringida: el tratamiento de datos personales solo podrá realizarse por las personas autorizadas por el titular y/o por las personas previstas en la ley.
- Seguridad: los datos personales sujetos a tratamiento se manejarán adoptando todas las medidas de seguridad que sean necesarias para evitar su pérdida, adulteración, consulta, uso o acceso no autorizado o fraudulento.
- Confidencialidad: todos los funcionarios que trabajen en la empresa están obligados a guardar reserva sobre la información personal a la que tengan acceso con ocasión de su puesto de trabajo en empresa.

Tratamiento y finalidades de los datos personales por la empresa

La empresa, como responsable del tratamiento de datos personales, recolecta, almacena, utiliza, circula y elimina datos personales de personas naturales con las que tiene o ha tenido relación, incluyendo empleados, sus familiares, accionistas, consumidores, clientes, distribuidores, proveedores,

acreedores y deudores. Esto se realiza para desarrollar adecuadamente sus actividades comerciales y fortalecer sus relaciones con terceros.

Marco de políticas generales para el tratamiento de datos personales

- Facilitar la participación de los titulares de los datos en actividades de clasificación y categorización.
- Analizar la calidad del servicio, realizar estudios de mercado sobre hábitos de consumo y llevar a cabo análisis estadísticos para uso interno.
- Controlar el acceso a las instalaciones de la empresa y establecer medidas de seguridad, incluyendo áreas con videovigilancia.
- Responder a consultas, peticiones, quejas y reclamos de los titulares y organismos de control, y transmitir los datos personales a las autoridades competentes según la ley aplicable.
- Contactar, por correo electrónico u otros medios, a personas naturales con quienes se tiene o ha tenido relación, como empleados, familiares, accionistas, consumidores, clientes, distribuidores, proveedores, acreedores y deudores, para los fines mencionados.
- Transferir la información recolectada a diferentes áreas de la empresa y a sus compañías vinculadas en el extranjero cuando sea necesario para sus operaciones (recaudo de cartera, cobros administrativos, tesorería, contabilidad, entre otros).

- Atender requerimientos judiciales o administrativos y cumplir con mandatos legales o judiciales.
- Registrar los datos personales en los sistemas de información de la empresa y en sus bases de datos comerciales y operativas.
- Realizar cualquier otra actividad similar a las mencionadas que sea necesaria para el desarrollo del objeto social y económico de la empresa.

Marco de políticas de derechos de los titulares de los datos personales

Las personas naturales cuyos datos personales sean objeto de tratamiento por parte de la empresa, tienen los siguientes derechos, los cuales pueden ejercer en cualquier momento:

- Conocer los datos personales que la empresa está tratando. El titular puede solicitar en cualquier momento la actualización o rectificación de sus datos si son parciales, inexactos, incompletos, fragmentados, inducen a error, o si su tratamiento está prohibido o no ha sido autorizado.
- Todo dato o información sensible debe pasar por la técnica de anonimización usando el método de sustitución antes de su análisis o intercambio interno o externo.
- Todo dato o información sensible además de pasar por un proceso de anonimización, debe tener acceso limitado.

- Solicitar prueba de la autorización otorgada a la empresa para tratar sus datos personales.
- Ser informado por la empresa, previa solicitud, sobre el uso que se ha dado a sus datos personales.
- Presentar quejas ante la entidad correspondiente por infracciones a la Ley de Protección de Datos Personales.
- Solicitar a la empresa la eliminación de sus datos personales y/o revocar la autorización para su tratamiento, mediante la presentación de un reclamo, de acuerdo con los procedimientos establecidos en esta política. Sin embargo, la solicitud de eliminación y la revocación de la autorización no procederán si el titular tiene un deber legal o contractual de permanecer en la base de datos, o mientras la relación entre el titular y la empresa esté vigente.
- Acceder de forma gratuita a sus datos personales que hayan sido objeto de tratamiento. Los derechos de los titulares pueden ser ejercidos por un representante o apoderado, previa acreditación de la representación o apoderamiento.

Marco de políticas de la empresa como responsable del tratamiento de datos personales

La empresa debe recordar que los datos personales pertenecen a las personas a las que se refieren y solo ellas pueden decidir sobre su uso. Por lo tanto, la empresa utilizará los datos personales conforme a la Política de

Protección de Datos Personales, asegurando que se empleen únicamente para los fines establecidos dentro del marco legal.

Área responsable de la implementación y observancia de este Marco de Políticas

El área de Compliance es responsable de desarrollar, implementar, capacitar y supervisar este marco político. Todos los empleados que manejan datos personales en las distintas áreas de la empresa deben reportar estas bases de datos al área de Compliance y remitir de inmediato cualquier petición, queja o reclamo recibido de los titulares de los datos personales. Además, el área de Compliance ha sido designada por la empresa para atender peticiones, consultas, quejas y reclamos, permitiendo a los titulares de la información ejercer sus derechos de conocer, actualizar, rectificar y suprimir sus datos, así como revocar la autorización.

Políticas para Atención y Respuesta a Peticiones, Consultas, de los Titulares de Datos Personales

Los titulares de datos personales gestionados por la empresa tienen el derecho de acceder a sus datos y conocer los detalles de su tratamiento. También pueden corregir y actualizar sus datos si son inexactos, solicitar su eliminación si consideran que son excesivos o innecesarios para los fines que justificaron su obtención, u oponerse a su tratamiento para fines específicos.

Para garantizar estos derechos, se han implementado los siguientes canales para la presentación de solicitudes:

- Comunicación dirigida a la empresa. Solicitud enviada al correo electrónico: DepartamentoSGI_VENTA@co.ab-inbev.com.
- Solicitud presentada telefónicamente al área de compliance.

Estos canales pueden ser utilizados por los titulares de datos personales o por terceros autorizados por ley para actuar en su nombre, con el fin de ejercer los siguientes derechos:

Procedimiento para la realización de peticiones y consultas

- El titular puede consultar sus datos personales en cualquier momento. Para ello, debe presentar una solicitud indicando la información que desea conocer, a través de cualquiera de los mecanismos mencionados.
- El titular o sus causahabientes deben acreditar su identidad, la de su representante, o la representación a favor de otro. Si la solicitud es formulada por una persona distinta del titular y no se acredita que actúa en representación de este, se considerará no presentada.
- La consulta o petición debe incluir al menos el nombre y dirección de contacto del titular, o cualquier otro medio para recibir la respuesta, así como una descripción clara y precisa de los datos personales sobre los cuales se desea ejercer el derecho de consulta o petición.
- Si la consulta o petición está incompleta, la empresa requerirá al interesado dentro de los cinco días siguientes a la recepción de la solicitud para que subsane las fallas. Si el solicitante no presenta la

información requerida dentro de dos meses desde la fecha del requerimiento, se entenderá que ha desistido de su consulta.

- Las peticiones y consultas serán atendidas por la empresa en un plazo máximo de diez días hábiles a partir de la fecha de recepción.
- Si no es posible atender la solicitud dentro de este plazo, se informará al solicitante, explicando los motivos de la demora y señalando la fecha en que se atenderá la petición, la cual no podrá superar los cinco días hábiles siguientes al vencimiento del primer plazo.

4.3 Difusión de marco político

La difusión del marco de políticas para garantizar una capa adicional a los datos sensibles, se lo debe realizar mediante el uso de medios digitales, grupos de conversación y correos masivos de difusión.

4.4 Casos de Éxitos

En este capítulo se presentan dos casos de éxito utilizados como referencia para validar técnicamente la efectividad del nuevo algoritmo de anonimización, aplicado en sistemas de recolección y virtualización de información del departamento de ventas. Cada caso se estructura bajo su propio formato, incluyendo la descripción del sistema, su funcionamiento, las áreas responsables de su administración y de la gestión de la información de clientes, el personal operativo involucrado y el volumen de datos sensibles procesados.

Caso 1: Sistema SAP-PR3

El primer caso corresponde a la implementación del proceso de pruebas en el sistema SAP-PR3, diseñado y evaluado detalladamente para la gestión de clientes y proveedores. Este sistema era utilizado por aproximadamente 3.000 funcionarios distribuidos en 18 provincias, abarcando más de 2.000 puntos de custodia de tiendas. SAP-PR3 soportaba dos tipos principales de procedimientos operativos y transacciones, orientados a satisfacer de forma secuencial y eficiente las necesidades de una amplia base de clientes.

Durante un mes de operación productiva, el sistema gestionaba un volumen significativo de registros de clientes y transacciones, todos ellos de carácter altamente sensible. Esta condición exigía la implementación de medidas de protección robustas para prevenir cualquier fuga de información, garantizando así el cumplimiento estricto de las políticas de cumplimiento normativo (compliance) de la empresa y asegurando la confidencialidad y privacidad de los datos tratados.

Caso 2: Sistema SAP PR-IME

El segundo caso se refiere a la implementación del proceso de pruebas en el sistema SAP PR-IME, diseñado y validado específicamente para las áreas de ventas, gestión de servicios y manejo de materias primas. Este sistema innovador era utilizado por cerca de 3.000 operarios distribuidos en tres países: Perú, Colombia y Ecuador, con presencia en 23 provincias ecuatorianas, lo que evidencia su amplia cobertura y relevancia operativa.

El sistema gestionaba un alto volumen de incidencias y transacciones críticas, especialmente aquellas relacionadas con movimientos financieros

como órdenes de compra, pagos y otros procedimientos esenciales para el funcionamiento de la organización. Estas operaciones eran ejecutadas por personal especializado en planta, lo que requería un enfoque particular en la protección de los datos sensibles involucrados, su cargue a SAP de los datos anonimizados se realizó en entornos de prueba con SAP Test Data Migration Server (TDMs) y para trabajar con reportes, se cargaron los datos anonimizados a SAP BW/4HANA sin comprometer la privacidad.

Conclusiones y Recomendaciones

Conclusiones

A continuación, se enumeran las conclusiones del proyecto:

1. Se establece un análisis de las cuatro técnicas de anonimización, viendo sus ventajas y desventajas, así como también un método de selección a aplicar basado en cuatro criterios. Todo esto con el objetivo de elegir una técnica de anonimización que ayude a cumplir con las normas y leyes de protección de datos vigentes en el Ecuador.
2. Este proyecto realiza una identificación adecuada y precisa de campos con información sensible, gracias a la colaboración de los departamentos de sistemas y ventas. Esto permite definir los datos sensibles que requieren un nivel de protección adicional.
3. Las pruebas realizadas en un entorno controlado dan como resultado que la técnica seleccionada dará un apoyo significativo para proteger y dar una capa adicional a información sensible de datos.

Recomendaciones

A continuación, se enumeran las recomendaciones del proyecto:

1. La aplicación de técnicas de anonimización en cualquier base de datos con información sensible, conlleva a que la información tenga una mayor seguridad. Esta aplicación ayudará a dar cumplimiento con las leyes vigentes en el Ecuador como también a normas internacionales como la ISO 27001 e ISO 27002.

2. La aplicación de técnicas de anonimización, permite que se fortalezca cada uno de los tres pilares de la seguridad de la información que son, la confidencialidad, la integridad y la disponibilidad.

3. Dentro de este proyecto de graduación, la implementación del método de sustitución tiene un alcance de lectura de un archivo de Excel y de anonimización de ciertos campos. Una mejora sería que la implementación pueda leer más de un archivo y de que la anonimización de la información de los campos sea de manera dinámica.

Bibliografía

[1] M. D. Morgan, M. M. Chowdhury, y S. Latif, «Protecting Business from Data Breach», en *2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, 2021, pp. 1-5. doi: 10.1109/ICECCME52200.2021.9590975.

[2] «1162059_-_LEY_ORGÁNICA_DE_PROTECCIÓN_DE_DATOS_PERS_202107011248165227».

[3] L. Demir, A. Kumar, M. Cunche, y C. Lauradoux, «The Pitfalls of Hashing for Privacy», *IEEE Commun. Surv. Tutor.*, vol. 20, n.º 1, pp. 551-565, 2018, doi: 10.1109/COMST.2017.2747598.

[4] A. Pawar, S. Ahirrao, y P. P. Churi, «Anonymization Techniques for Protecting Privacy: A Survey», en *2018 IEEE Punecon*, 2018, pp. 1-6. doi: 10.1109/PUNECON.2018.8745425.

[5] H. Tahir y P. Brézillon, «A Context Approach to Improve the Data Anonymization Process», en *2022 International Conference on Engineering and Emerging Technologies (ICEET)*, 2022, pp. 1-6. doi: 10.1109/ICEET56468.2022.10007410.

[6] A. Stam y B. Kleiner, «Data anonymization: legal, ethical, and strategic considerations», *Guide No 11 Version 10 Lausanne Swiss Cent. Expert. Soc. Sci. FORS*, 2020.

[7] J. F. Marques y J. Bernardino, «Analysis of Data Anonymization Techniques.», en *KEOD*, 2020, pp. 235-241.

[8] Z. Aslanyan y M. S. Boesgaard, «Privacy Analysis of Format-Preserving Data-Masking Techniques», en *2019 12th CMI Conference on Cybersecurity and Privacy (CMI)*, 2019, pp. 1-6. doi: 10.1109/CMI48017.2019.8962143.

[9] A. Anant y R. Prasad, «Public Private Data Partnerships enabling Privacy Technologies», en *2022 25th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, 2022, pp. 86-91. doi: 10.1109/WPMC55625.2022.10014859.

[10] «Norma-ISO-27001-2017».

[11] «Norma-ISO-27002-2022».

[12] J. L. Córdova-Real y G. M. López-Sevilla, «Técnicas de anonimización y pseudonimización en la protección de datos personales», *MQRInvestigar*, vol. 8, n.º 1, pp. 204-235, 2024.

[13] M. R. Martínez y J. I. Pincay-Ponce, «Buenas Prácticas de Seguridad para la Protección de la Privacidad con Datos Abiertos», *Rev. Científica Informática ENCRIPTAR-ISSN 2737-6389*, vol. 7, n.º 14, pp. 187-205, 2024.

[14] E. Bodean y G. P. Henning, «Tiempos de traslado y de servicio en la distribución de mercadería de última milla. Estimación mediante técnicas de minería de datos», *Mem. Las JAIIO*, vol. 10, n.º 14, pp. 418-421, 2024.

[15] M. I. Espinoza Quintana y C. E. Piedra Ramírez, «Propuesta de un plan de seguridad industrial y salud ocupacional para la estación de servicio´´ Centenario´´ comercializadora Terpel ubicada en la ciudad de Guayaquil», 2023.

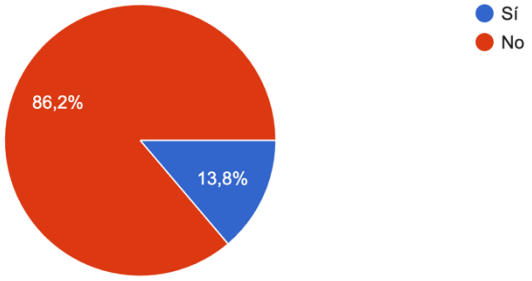
[16] G. D. L. C. Rodríguez, R. A. M. Fernández, y A. C. M. Fernández, «Seguridad de la información en el comercio electrónico basado en ISO 27001: Una revisión sistemática», *Innov. Softw.*, vol. 4, n.º 1, pp. 219-236, 2023.

[17] D. J. Palma Chávez, «Modelo de gestión documental para la empresa pública de la Universidad Laica Eloy Alfaro de Manabí», Jipijapa-Unesum, 2023.
[18] J. Rodríguez Hernández, «Impacto de la comunicación en Twitter en el movimiento ambientalista durante la COP15», *Rev. Comun.*, vol. 23, n.º 1, pp. 485-505, 2024.
[19] F. C. Britton, S. Dowling, y M. Frain, «A contribution towards the regulation of anonymised datasets within the framework of GDPR», en *2022 Cyber Research Conference - Ireland (Cyber-RCI)*, 2022, pp. 1-6. doi: 10.1109/Cyber-RCI55324.2022.10032680.
[20] C. Ni, L. S. Cang, P. Gope, y G. Min, «Data anonymization evaluation for big data and IoT environment», *Inf. Sci.*, vol. 605, pp. 381-392, 2022, doi: https://doi.org/10.1016/j.ins.2022.05.040.

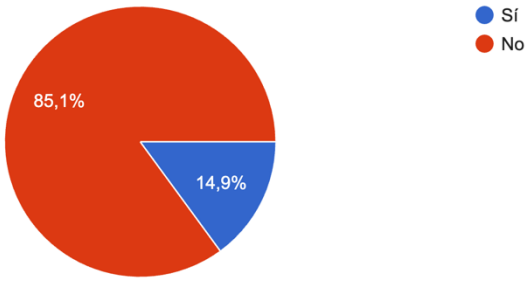
Anexos

Anexo A: Encuesta para evaluar que campos son considerados con información sensible

¿El campo CLIENTE_SAP, considera usted que contiene información sensible?
174 respuestas

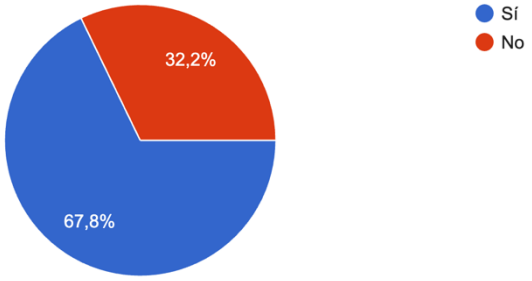


¿El campo COD_REGIONAL, considera usted que contiene información sensible?
174 respuestas



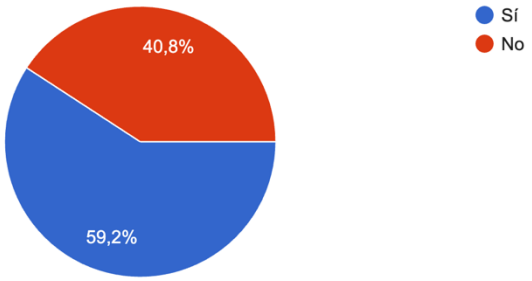
¿El campo NOMBRE_DIRECCION, considera usted que contiene información sensible?

174 respuestas



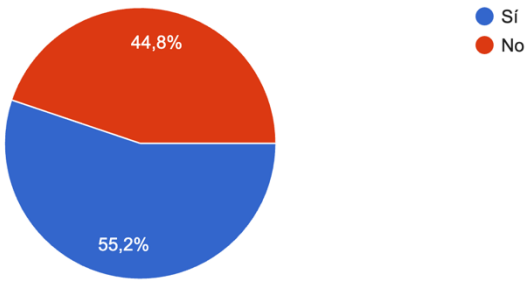
¿El campo NOMBRE_GERENCIA_GENERAL, considera usted que contiene información sensible?

174 respuestas



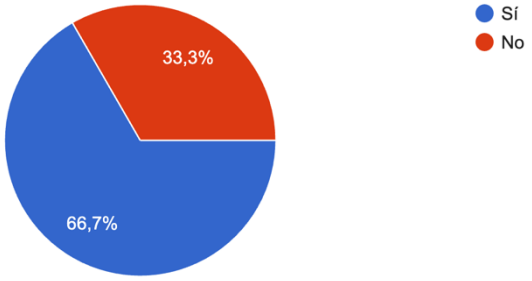
¿El campo COD_GERENCIA_VENTAS, considera usted que contiene información sensible?

174 respuestas



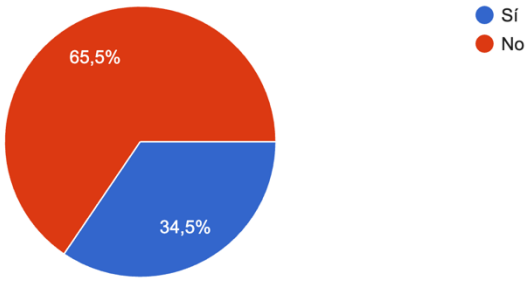
¿El campo ZONA_DE_VENTAS, considera usted que contiene información sensible?

174 respuestas



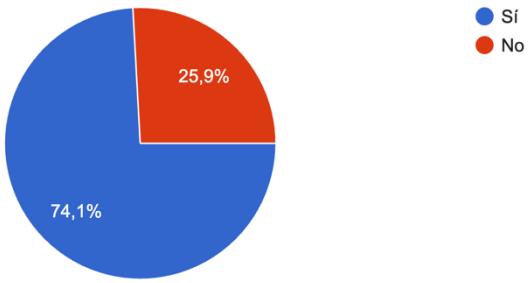
¿El campo NOMBRE_GERENCIA_VENTAS, considera usted que contiene información sensible?

174 respuestas



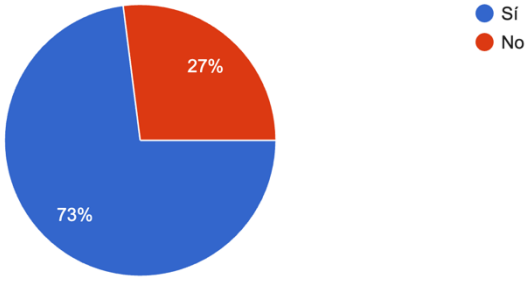
¿El campo NOMBRE_1, considera usted que contiene información sensible?

174 respuestas



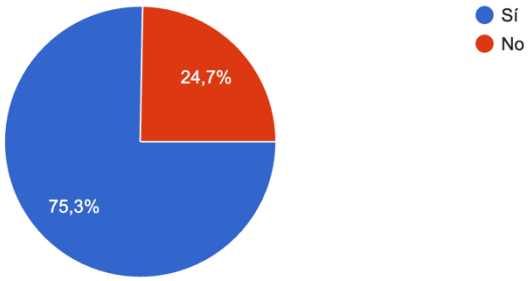
¿El campo NOMBRE_3, considera usted que contiene información sensible?

174 respuestas



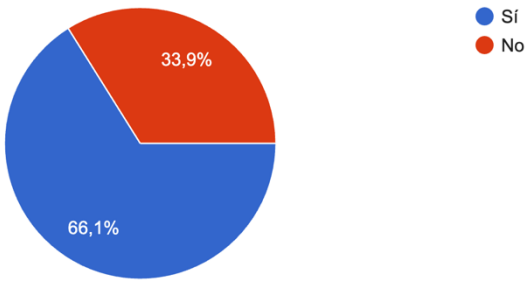
¿El campo IDENTIFICACION, considera usted que contiene información sensible?

174 respuestas



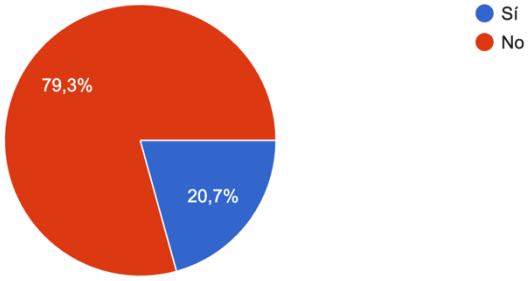
¿El campo PROVINCIA, considera usted que contiene información sensible?

174 respuestas



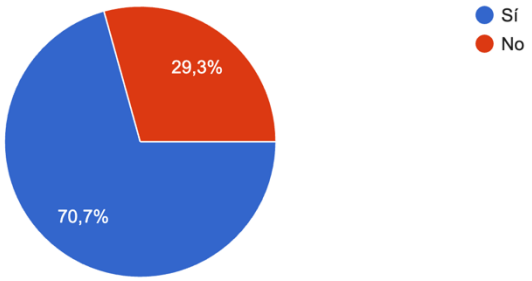
¿El campo POBLACION, considera usted que contiene información sensible?

174 respuestas



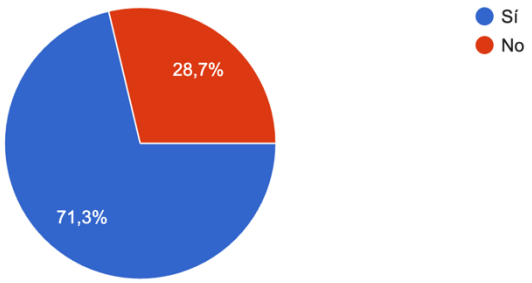
¿El campo CALLE_3, considera usted que contiene información sensible?

174 respuestas



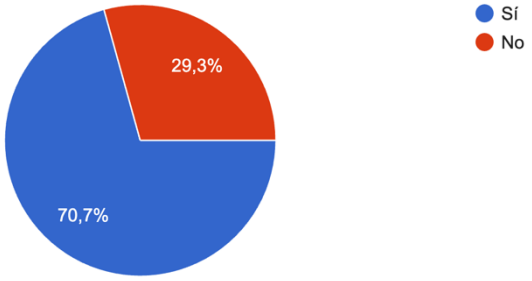
¿El campo CALLE_Y_NO, considera usted que contiene información sensible?

174 respuestas



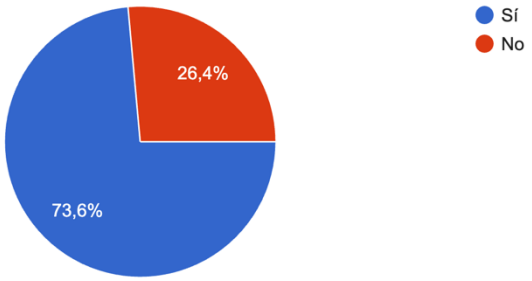
¿El campo CALLE_4, considera usted que contiene información sensible?

174 respuestas



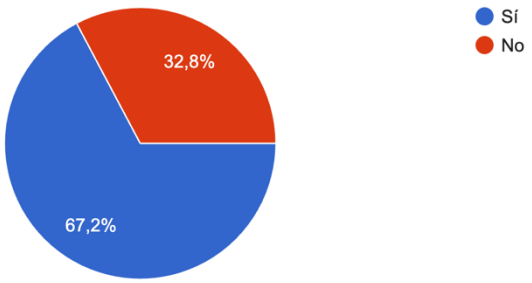
¿El campo TELEFULL, considera usted que contiene información sensible?

174 respuestas



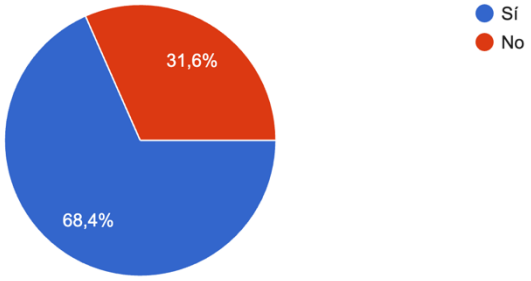
¿El campo LATITUD, considera usted que contiene información sensible?

174 respuestas



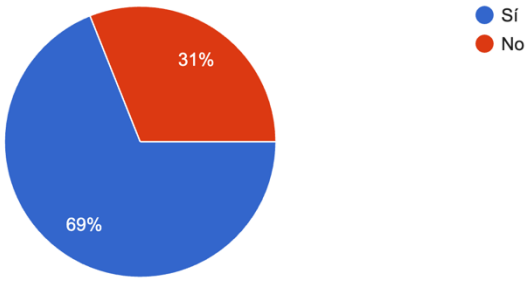
¿El campo LONGITUD, considera usted que contiene información sensible?

174 respuestas



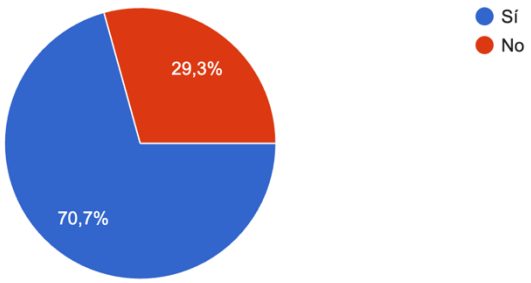
¿El campo NOM_DISTRIBUIDOR, considera usted que contiene información sensible?

174 respuestas



¿El campo CORREO, considera usted que contiene información sensible?

174 respuestas



Anexo B: Código de la técnica de anonimización usando el método de sustitución

Código fuente del archivo index.js

```
const { sustitutionPromise } =
require('./techiques/substitutionTecnique');
const { replaceDataPromise } =
require('./techiques/replaceDataTecnique');
const { keepUsefullDataPromise } =
require('./techiques/keepUsefullDataTecnique');
const { drowingDataPromise } =
require('./techiques/drowingDataTecnique');
const { getRandomRowsFromXlsx } = require('./utils/openFileHelper');
const { exportDataToXlsx } = require('./utils/createFileHelper');

const fs = require('fs');
const fileInputPath = './resource/input/';
const fileOutputPath = './resource/output/';
const files = fs.readdirSync( fileInputPath ).filter( file =>
file.endsWith( '.xlsx' ));
const techique = process.argv[2] ? process.argv[2] : '';
const totalRows = process.argv[3] ? parseInt(process.argv[3]) :
2000;

files.forEach( file => {
  console.log( `Processing file: ${file}` );
  const suffix = Date.now();
  const suffixFileName = file.split('.xlsx')[0]+'-'+suffix+'.xlsx';
  let anonymizedData = '';
  const techniqueChoosen = techique.toString().toLowerCase();
  (async () => {
    try {
      const randomRows = await getRandomRowsFromXlsx(
fileInputPath+file, totalRows );
      const exportDataOriginal = await exportDataToXlsx( 'original'
,fileOutputPath+'Original-'+suffixFileName, randomRows );
      switch( techniqueChoosen ) {
        case 'replace':
```

```
          anonymizedData = await replaceDataPromise(
techniqueChoosen, suffixFileName, randomRows );
          break;
        case 'keep':
          anonymizedData = await keepUsefullDataPromise(
techniqueChoosen, suffixFileName, randomRows );
          break;
        case 'drowing':
          anonymizedData = await drowingDataPromise(
techniqueChoosen, suffixFileName, randomRows );
          break;
        case 'substitution':
          anonymizedData = await sustitutionPromise(
techniqueChoosen, suffixFileName, randomRows );
          break;
        default:
          console.log('Please choose a valid technique: replace,
keep, drowing or substitution');
          break;
      }
      if ( techique.toString().toLowerCase() !== '' ) {
        const exportData = await exportDataToXlsx( techniqueChoosen,
fileOutputPath+techniqueChoosen+'-Anonymize-'+suffixFileName,
anonymizedData );
        console.log( exportData );
      }
    } catch ( error ) {
      console.error( 'Error:', error );
    }
  })();
});
```

Código fuente de la técnica de sustitución

```
const { faker } = require('@faker-js/faker');

async function sustitutionPromise ( techniqueChoosen, filePath,
dataSet ) {
  return new Promise(( resolve, reject ) => {
```

```
    try {
      console.time( `Time to anonymize data using
${techniqueChoosen}: ${filePath}` );
      const randomData = structuredClone(dataSet);
      randomData.forEach(data => {
        data.NOMBRE_1 = (data.NOMBRE_1) ? faker.person.firstName() :
null;
        data.NOMBRE_3 = (data.NOMBRE_3) ? faker.person.middleName()
: null;
        data.IDENTIFICACION = (data.IDENTIFICACION) ?
faker.number.int() : null;
        data.CALLE_3 = (data.CALLE_3) ? faker.location.direction() :
null;
        data.CALLE_Y_NO = (data.CALLE_Y_NO) ?
faker.location.direction() : null;
        data.CALLE_4 = (data.CALLE_4) ? faker.location.direction() :
null;
        data.TELEFULL = (data.TELEFULL) ? faker.phone.number() :
null;
        data.CORREO = (data.CORREO) ? faker.internet.email() : null;
      });
      console.timeEnd( `Time to anonymize data using
${techniqueChoosen}: ${filePath}` );
      resolve( randomData );
    } catch ( error ) {
      reject ( error );
    }
  });
};

module.exports = { sustitutionPromise };
```

Código fuente de la técnica de ahogamiento de datos

```
const generalizeName = ( name ) => {
  const firstName = name.split(" ")[0];
  return firstName.charAt(0).toUpperCase() + ".";
```

```
}

const generalizeID = ( id ) => {
  try {
    const idStr = id.toString().trim();
    if (idStr.length <= 4) {
      return '*'.repeat(idStr.length - 1) + idStr.slice(-1);
    }
    const visible = idStr.slice(-4);
    return '*'.repeat(idStr.length - 4) + visible;
  } catch (error) {
    return '';
  }
}

const generalizeAddress = ( address ) => {
  try {
    let generalized = address.replace(/\d+/g, '');
    const specificWords = [

/\b(casa|dept?o?\.?|departamento|edificio|block|bloque|torre|piso)\s
*\d*\b/gi,
      /\b(mz|manzana|lote|solar)\s*\d*\b/gi,
      /\b(km|kilómetro)\s*\d+\.?\d*\b/gi
    ];

    specificWords.forEach(pattern => {
      generalized = generalized.replace(pattern, '');
    });
    generalized = generalized.replace(/\s+/g, ' ').trim();
    return generalized;
  } catch (error) {
    return '';
  }
}

const generalizePhone = ( phone ) => {
  const phoneStr = phone.toString();
  const maskChar = '*';
  if (phoneStr.includes('-')) {
    const phoneNumbers = phoneStr.split('-');
```

```
    const maskedNumbers = phoneNumbers.map(num => {
      return num.length > 4 ? maskChar.repeat(num.length - 4) +
num.slice(-4) : maskChar.repeat(num.length);
    });
    return maskedNumbers.join('-');
  } else {
    return phoneStr.length > 4 ? maskChar.repeat(phoneStr.length -
4) + phoneStr.slice(-4) : maskChar.repeat(phoneStr.length);
  }
}

const generalizeEmail = ( email ) => {
  const [user, domain] = email.split("@");
  if (!user || !domain) return '';
  return user.slice(0, Math.min(2, user.length)) + '****' +
'@ejemplo.com';
}

async function drowingDataPromise ( techniqueChoosen, filePath,
dataSet ) {
  return new Promise(( resolve, reject ) => {
    try {
      console.time( `Time to anonymize data using
${techniqueChoosen}: ${filePath}` );
      const randomData = structuredClone(dataSet);
      randomData.forEach(data => {
        data.NOMBRE_1 = ( data.NOMBRE_1 ) ? generalizeName(
data.NOMBRE_1 ) : null;
        data.NOMBRE_3 = ( data.NOMBRE_3 ) ? generalizeName(
data.NOMBRE_1 ) : null;
        data.IDENTIFICACION = ( data.IDENTIFICACION ) ?
generalizeID( data.IDENTIFICACION ) : null;
        data.CALLE_3 = ( data.CALLE_3 ) ? generalizeAddress(
data.CALLE_3 ) : null;
        data.CALLE_Y_NO = ( data.CALLE_Y_NO ) ? generalizeAddress(
data.CALLE_Y_NO ) : null;
        data.CALLE_4 = ( data.CALLE_4 ) ? generalizeAddress(
data.CALLE_4 ) : null;
        data.TELEFULL = ( data.TELEFULL ) ? generalizePhone(
data.TELEFULL ) : null;
```

```
        data.CORREO = ( data.CORREO ) ? generalizeEmail( data.CORREO
) : null;
      });
      console.timeEnd( `Time to anonymize data using
${techniqueChoosen}: ${filePath}` );
      resolve( randomData );
    } catch ( error ) {
      reject( error.message );
    }
  });
}

module.exports = { drowingDataPromise };
```

Código fuente de la técnica de conservación de datos necesarios

```
async function keepUsefullDataPromise ( techniqueChoosen, filePath,
dataSet ) {
  return new Promise(( resolve, reject ) => {
    try {
      console.time( `Time to anonymize data using
${techniqueChoosen}: ${filePath}` );
      const randomData = dataSet;
      randomData.forEach( data => {
        delete data.NOMBRE_1;
        delete data.NOMBRE_3;
        delete data.IDENTIFICACION;
        delete data.CALLE_3;
        delete data.CALLE_Y_NO;
        delete data.CALLE_4;
        delete data.TELEFULL;
        delete data.CORREO;
      });
      console.timeEnd( `Time to anonymize data using
${techniqueChoosen}: ${filePath}` );
      resolve( randomData );
    } catch ( error ) {
```

```
      reject( error );
    }
  });
};

module.exports = { keepUsefullDataPromise };
```

Código fuente de la técnica reemplazo de datos

```
const maskChar = '*';
const patterns = {
  name: /\b[A-Z][a-z]{2,15}(?:\s+[A-Z][a-z]{2,15})?\b/g,
  id: /\b\d{6,}\b/g,
  address: /\b\d+\s+[A-Za-z\s,.-
]+(?:Street|St|Avenue|Ave|Road|Rd|Boulevard|Blvd|Drive|Dr|Lane|Ln|Co
urt|Ct|Place|Pl|Way|Calle)\b/gi,
  phone: /\b(?:\d{8,10}(?:-\d{8,10})*)\b/g,
  email: /\b[A-Za-z0-9._%+-]+@[A-Za-z0-9.-]+\.[A-Z|a-z]{2,}\b/g
};
const maskName = ( name ) => {
  return name.replace(/\b[A-Za-z]+\b/g, (word) => {
    return word[0] + maskChar.repeat(Math.max(word.length - 1, 2));
  });
};
const maskId = ( id ) => {
  const idStr = id.toString();
  return idStr.length > 4 ? maskChar.repeat(idStr.length - 4) +
idStr.slice(-4) : maskChar.repeat(idStr.length);
};
const maskAddress = ( address ) => {
  if (typeof address !== 'string') {
    return address;
  }
  return address.replace(/\b\d+\b/g, '***').replace(/\b[A-Za-
z]{3,}\b/g, (word) => {
    if (['Av', 'COOP', 'Ava', 'Via', 'Km', 'MZ',
'Cdla'].includes(word)) {
```

```
      return word;
    }
    return maskChar.repeat(Math.min(word.length, 5));
  });
};
const maskPhone = ( phone ) => {
  const phoneStr = phone.toString();
  if (phoneStr.includes('-')) {
    const phoneNumbers = phoneStr.split('-');
    const maskedNumbers = phoneNumbers.map(num => {
      return num.length > 4 ? maskChar.repeat(num.length - 4) +
num.slice(-4) : maskChar.repeat(num.length);
    });
    return maskedNumbers.join('-');
  } else {
    return phoneStr.length > 4 ? maskChar.repeat(phoneStr.length -
4) + phoneStr.slice(-4) : maskChar.repeat(phoneStr.length);
  }
};
const maskEmail = ( email ) => {
  return email.replace( patterns.email, ( match ) => {
    const [username, domain] = match.split('@');
      return username[0] + maskChar.repeat(Math.max(username.length
- 1, 3)) + '@' + domain;
  });
};
async function replaceDataPromise( techniqueChoosen, filePath,
dataSet) {
  return new Promise(( resolve, reject ) => {
    try {
      console.time( `Time to anonymize data using
${techniqueChoosen}: ${filePath}` );
      const randomData = dataSet;
      randomData.forEach( data => {
        data.NOMBRE_1 = ( data.NOMBRE_1 ) ? maskName( data.NOMBRE_1
) : null;
        data.NOMBRE_3 = ( data.NOMBRE_3 ) ? maskName( data.NOMBRE_3
) : null;
        data.IDENTIFICACION = ( data.IDENTIFICACION ) ? maskId(
data.IDENTIFICACION ) : null;
```

```
        data.CALLE_3 = ( data.CALLE_3 ) ? maskAddress( data.CALLE_3
) : null;
        data.CALLE_Y_NO = ( data.CALLE_Y_NO ) ? maskAddress(
data.CALLE_Y_NO ) : null;
        data.CALLE_4 = ( data.CALLE_4 ) ? maskAddress( data.CALLE_4
) : null;
        data.TELEFULL = ( data.TELEFULL ) ? maskPhone( data.TELEFULL
) : null;
        data.CORREO = ( data.CORREO ) ? maskEmail( data.CORREO ) :
null;
      });
      console.timeEnd( `Time to anonymize data using
${techniqueChoosen}: ${filePath}` );
      resolve( randomData );
    } catch ( error ) {
      reject( error );
    }
  });
}

module.exports = { replaceDataPromise };
```

Código fuente de la función para obtención de filas de manera aleatoria

```
const xlsx = require('xlsx');

async function getRandomRowsFromXlsx( file, totalRows ) {
  return new Promise(( resolve, reject ) => {
    try {
      console.time( `Time to open file ${file}` );
      const workbook = xlsx.readFile( file );
      const firstSheetName = workbook.SheetNames[0];
      const sheetData = xlsx.utils.sheet_to_json(
workbook.Sheets[firstSheetName] );
      const randomRows = sheetData.sort(() => Math.random() -
Math.random()).slice(0, totalRows);
      console.timeEnd( `Time to open file ${file}` );
      resolve( randomRows );
```

```
    } catch ( error ) {
      reject( error );
    }
  });
}

module.exports = { getRandomRowsFromXlsx };
```

Código fuente de la función para exportar datos a un archivo en Excel.

```
const xlsx = require('xlsx');

async function exportDataToXlsx( techniqueChoosen, outputFilePath,
data ) {
  return new Promise(( resolve, reject ) => {
    try {
      console.time( `Time to export data ${techniqueChoosen}:
${outputFilePath}` );
      if (!Array.isArray(data)) {
        throw new TypeError( 'La data debe ser un arreglo de
objetos' );
      }
      const workbook = xlsx.utils.book_new();
      const worksheet = xlsx.utils.json_to_sheet( data );
      xlsx.utils.book_append_sheet( workbook, worksheet, 'Libro1' );
      xlsx.writeFile( workbook, outputFilePath );
      console.timeEnd( `Time to export data ${techniqueChoosen}:
${outputFilePath}` );
      resolve( `Archivo creado exitosamente en el directorio:
${outputFilePath}` );
    } catch ( error ) {
      reject( error );
    }
  });
}

module.exports = { exportDataToXlsx };
```