

# GCG

The Journal of Globalization,  
Competitiveness, and Governability

20

## Ciencia de Datos Aplicada a Prácticas de Gestión de Riesgo en Sociedades Chilenas que Reportan a la Comisión de Mercados Financieros

### AUTORES

#### Joseline J.

**Sepúlveda-Araya<sup>1</sup>**  
Facultad de Ciencias Sociales y Económicas, Universidad Católica del Maule, Chile  
jsepulveda@ucm.cl

#### Valeria A.

**Orostiga-Lazo**  
Facultad de Ciencias Sociales y Económicas, Universidad Católica del Maule, Chile  
vorostiga@ucm.cl

#### Armin Herrera

Facultad de Ciencias de la Ingeniería, Universidad Católica del Maule, Chile

#### Cristian A.

#### Martinez

Facultad de Ciencias de la Ingeniería, Universidad Católica del Maule, Chile  
cmartinez@di.unsa.edu.ar

#### Ricardo J.

#### Barrientos

Facultad de Ciencias de la Ingeniería, Universidad Católica del Maule, Chile  
cmartinez@di.unsa.edu.ar

1. Autora de contacto:  
Facultad de Ciencias Sociales y Económicas, Universidad Católica del Maule, Av. San Miguel 3605, Talca, Chile.  
Código Postal 3460000

ÁREA: 2  
TIPO: Aplicación

*Data science applied to risk management practices in Chilean companies reporting to the Financial Markets Commission*

*Ciéncia de dados aplicada a práticas de gerenciamento de riscos em empresas chilenas que prestam contas à Comissão de Mercados Financeiros*

*El presente estudio analiza el nivel de cumplimiento de prácticas de gobierno corporativo relacionadas a la gestión de riesgos y su desempeño financiero, empleando técnicas de Ciencia de Datos (Machine Learning) en empresas chilenas. La metodología utilizada abarca las fases del proceso de descubrimiento de conocimiento en base de datos. Como resultado, se obtuvieron modelos basados en Aprendizaje No Supervisado y Supervisado, que permitieron caracterizar y luego predecir con un alto nivel de precisión, el grado de cumplimiento de prácticas. Asimismo, se determina una agrupación de empresas con diferentes características en cuanto al nivel de cumplimiento y rendimiento financiero.*

*This study analyzes the level of compliance with corporate governance practices related to risk management and financial performance, using Machine Learning techniques in Chilean companies. The methodology used covers the phases of the knowledge discovery process in a database. As a result, models based on Unsupervised and Supervised Learning were obtained, which allowed characterizing and then predicting with a high level of accuracy, the degree of compliance with practices. Also, a grouping of companies with different characteristics in terms of the level of compliance and financial performance is determined.*

*Este estudo analisa o nível de conformidade com as práticas de governança corporativa relacionadas à gestão de riscos e ao desempenho financeiro, usando técnicas de Machine Learning em empresas chilenas. A metodologia utilizada abrange as fases do processo de descoberta de conhecimento em um banco de dados. Como resultado, foram obtidos modelos baseados em Aprendizagem Não Supervisionada e Supervisionada, que permitiram caracterizar e, em seguida, prever, com um alto nível de precisão, o grau de conformidade com as práticas. Além disso, foi determinado um agrupamento de empresas com características diferentes em termos de nível de conformidade e desempenho financeiro.*

DOI

10.58416/GCG.2026.V20.N1.01

RECIBIDO

2024.12.20

ACEPTADO

2025.04.05

## 1. Introducción

En Chile, la adopción de prácticas de gobierno corporativo ha sido limitada (Moraga y Ropero, 2018). Moraga-Flores y Rossi-Undurraga (2019) analizaron la NCG 385 y concluyeron que no existe relación entre el área productiva de la empresa y la adopción de prácticas como tampoco entre la adopción de prácticas y la solvencia de las empresas. De manera similar, Arenas-Torres et al. (2021) determinaron que su implementación no afecta a la rentabilidad de las mismas.

Desde finales de los años 90, los avances en TIC han impulsado una generación exponencial de datos. Su gestión y análisis han sido posibles gracias a disciplinas como estadística, matemática, programación e inteligencia artificial, dando origen a la Ciencia de Datos (Barbaglia et al., 2021). Su principal objetivo es extraer información y conocimiento a partir del análisis de datos. Si bien ha tenido aplicaciones exitosas en Salud, Educación, Agricultura, Cadena de Suministro y Deportes (Subrahmanyam et al., 2022; Liu y Huang, 2017; Cravero y Sepúlveda, 2021; Jahani et al., 2023; D'Urso et al., 2023), sus aportes en Economía y Finanzas son aún limitados.

En esta investigación se aplican técnicas de Ciencia de Datos para analizar a las sociedades chilenas que informan a la Comisión de Mercado Financiero, con el fin de mejorar su comprensión considerando el cumplimiento de prácticas de gobierno corporativo, gestión de riesgos y datos financieros. El estado del arte da cuenta que no existen estudios previos que aborden esta problemática desde la perspectiva propuesta en esta investigación.

Además de esta introducción, el presente trabajo se estructura de la siguiente manera: en Marco Teórico se abordan conceptos de gestión de riesgos y Ciencia de Datos; en Metodología se describe la base de datos y las técnicas aplicadas; en Resultados se presentan los valores obtenidos, los cuales se analizan en la Discusión; y finalmente, en Conclusión, se destacan los aportes del estudio.

**PALABRAS CLAVE**  
**Ciencia de Datos,  
Empresas, Gestión  
de Riesgos,  
Machine Learning**

**KEYWORDS**  
**Data Science,  
Companies, Risk  
Management,  
Machine Learning**

**PALAVRAS-CHAVE**  
**Ciência de  
Dados, Empresas,  
Gerenciamento de  
Riscos, Machine  
Learning**

## 2. Marco Teórico

### 2.1. Gestión de riesgos y desempeño financiero

La gestión de riesgos proporciona herramientas para afrontar los diversos riesgos a los que están expuestos las organizaciones (Robles et al., 2019) contribuyendo a maximizar el valor cuando la dirección encuentra un equilibrio entre rentabilidad, crecimiento y riesgos, empleando los recursos de forma eficaz y eficiente para el logro de los objetivos.

Gestionar el riesgo implica mucho más que identificarlos, clasificarlos, valorarlos y monitorearlos antes de que surjan; también significa mitigar el efecto de los que se presentan y reflexionar sobre cómo mejorar la organización en el futuro (Mora Navarro, 2022).

**Códigos JEL**  
**C38-C45-G34-  
L25-M14**

Estudios como los de Ramlee y Ahmad (2020) destacan que la implementación de una gestión de riesgos en empresas no financieras es una buena práctica de gobernanza, como también ciertas características del comité de gestión de riesgos (CGR) en el desempeño financiero. Ayuningtyas y Harymawan (2022) señalan que su existencia aumentó la divulgación sobre riesgos. Ahmad et al. (2021) concluyen que dar a conocer información sobre gestión de riesgo empresarial, refleja que la gerencia ha administrado la entidad bajo un buen juicio.

La adopción de buenas prácticas corporativas podría tener efectos positivos en el rendimiento financiero de la empresa a futuro (Arenas-Torres et al., 2021). Los estudios sobre la gestión de riesgos y desempeño financiero han mostrado resultados heterogéneos (Horvey y Ankamah, 2020). Malik et al. (2020) al estudiar el impacto de la gestión de riesgos empresariales en empresas del Reino Unido, observan que su eficacia influye positivamente en el rendimiento y que un comité de riesgos estructuralmente sólido a nivel de consejo amplifica su efecto. En contraste, Agustina y Baroroh (2016) señalan que no existe una influencia significativa entre la gestión de riesgos empresarial y el rendimiento financiero.

Desde una mirada teórica, diversos estudios han abordado la gestión de riesgos, desde la teoría de agencia, donde los accionistas delegan la responsabilidad de administración de la empresa en altos ejecutivos. En este sentido, Bui y Krajcsák (2024) indican que "los accionistas esperan que los gerentes tomen decisiones que los beneficien. Sin embargo, las prioridades de los gerentes a veces no son las mismas que la de los accionistas; sus propios objetivos pueden diferir en cuanto a aumentar el valor de la empresa" (p.19). A su vez, Ramlee y Ahmad (2020) indican que el comité de gestión de riesgo, actúa como un agente del consejo de administración, resguardando que las decisiones referentes a gestión de riesgos que toman se implementen y controlen.

En el estudio de las prácticas de gobierno corporativo mediante técnicas estadísticas y de Aprendizaje Automático (ML), se revisaron algunos trabajos clave, tales como el de Moraga Flores y Ropero Moriones (2017) que analizaron la adopción de estas prácticas en empresas chilenas entre 2015 y 2016. En dicho trabajo, propusieron un indicador de adopción (IAGCE) basado en una relación lineal del promedio de las prácticas de gobierno y lo correlacionaron con la rentabilidad neta, el retorno sobre el patrimonio (ROE) y el índice Q de Tobin. Los resultados indicaron que la adopción de estas prácticas no influye directamente en dichos indicadores; Gouiaa (2018) investigó cómo ciertos atributos del Consejo de Administración afectan la gestión de riesgos en empresas canadienses entre 2012 y 2013, desarrollando un modelo de regresión para predecir el manejo de riesgos. El autor destacó que los Consejos más grandes e independientes, con un comité de riesgos independiente, gestionan mejor los riesgos financieros; Horvey y Ankamah (2020) estudiaron la relación entre gestión de riesgos y desempeño empresarial en el mercado de valores de Ghana entre 2010 y 2016, utilizando tres modelos lineales de regresión para predecir ROA, ROE y el índice de Tobin, con el índice de manejo de riesgo (ERMI) como variable dependiente. La investigación reveló que las empresas grandes y de capital extranjero presentan mejores valores de ROE y Q de Tobin; recientemente, Navarrete Oyarce et al. (2023) analizaron la relación entre la adopción de prácticas de gobierno corporativo y el desempeño financiero en empresas chilenas entre 2013 y 2019, utilizando el índice IAGCE para medir el cumplimiento de estas prácticas. Los autores concluyeron que las empresas con buenas prácticas de gobierno tienen un mejor desempeño financiero, especialmente en términos de ROE.

Aunque estos trabajos aportan valiosos hallazgos, no profundizan en los conjuntos de datos, ni aplican técnicas que permitan agrupar las empresas de manera homogénea o clasificarlas según su cumplimiento de las prácticas de gobierno corporativo y su rendimiento financiero. En base en lo anterior, surgen las siguientes preguntas de investigación:

¿Qué variables empresariales describen mejor el cumplimiento de las prácticas de gobierno corporativo y su situación financiera?

¿Es posible agrupar empresas con características comunes, considerando variables de gobierno corporativo y situación financiera, bajo criterios numéricos y de manera homogénea?

¿Se puede aplicar una técnica predictiva para estimar con alta precisión el nivel de cumplimiento de las prácticas de gobierno y la situación financiera de una empresa en particular?

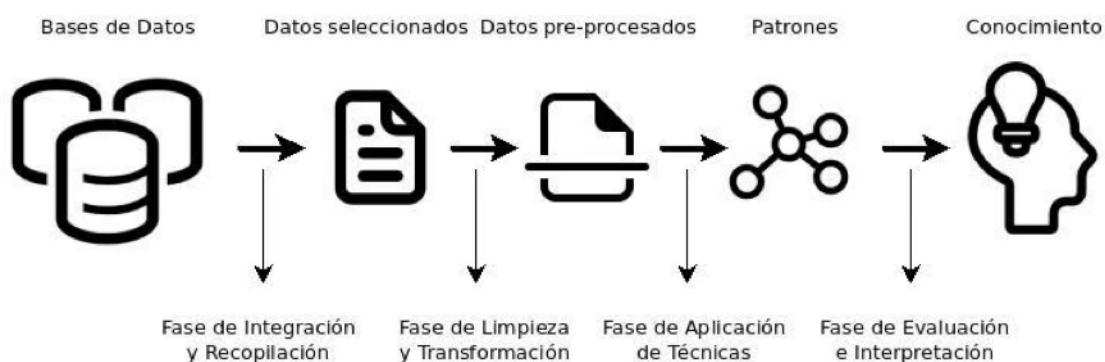
## 2.2. Ciencia de Datos aplicado a los negocios

El descubrimiento de conocimiento en base de datos (KDD) es un proceso no trivial de identificar patrones válidos, novedosos, potencialmente útiles y entendibles a partir de los datos (Fayyad y Stolorz, 1997). Este proceso consta de (Ibarra, 2020):

- Integración y recopilación de datos
- Selección, limpieza y transformación de datos
- Aplicación de técnicas
- Evaluación e interpretación de resultados
- 

En la **Figura 1** se muestran las etapas del proceso KDD, el cual es un proceso iterativo y de interacción continua con el interesado del proyecto.

Figura 1. - Etapas en el proceso KDD (Ibarra, 2020)



Dentro de las técnicas usadas, las provenientes principalmente de la Estadística son las más conocidas: correlación, estadística descriptiva y análisis multivariado. Últimamente, la Inteligencia Artificial (IA) ha tenido una participación más que destacada. Dentro de ella, las áreas más conocidas son los Sistemas Expertos, el Aprendizaje Automático, la Robótica, los Sistemas de soporte para la toma de decisión y el Reconocimiento de Patrones. El Aprendizaje Automático (Machine Learning o ML) se define como un conjunto de algoritmos y modelos que pueden aprender a identificar patrones y tomar decisiones para resolver tareas específicas usando datos relacionados con dichas tareas (Plathottam et al., 2023). De acuerdo a los métodos de aprendizaje, se puede dividir en supervisado, no supervisado, semi-supervisado y por refuerzo. El aprendizaje supervisado está referido al uso de datos etiquetados para entrenar modelos (basados en métodos específicos) para predecir el tipo o valor de un nuevo dato.

Entre los métodos, podemos citar a Random Forests, Support Vector Machines y Extreme Learning Machine (Breiman, 2001; Flach, 2001; Wang et al. 2022). En el aprendizaje no supervisado, los modelos son entrenados con datos no etiquetados para descubrir patrones ocultos de manera automática. Las tareas más conocidas con el aprendizaje no supervisado se relacionan con la reducción de dimensión, el agrupamiento de datos, la asociación y la detección de datos anómalos. Entre los métodos se destacan Componentes Principales, Feature Importance, K-Means y BIRCH (Ghahramani, 2004; Shobha y Rangaswamy, 2018; Zhang et al., 1997).

### 3. Metodología

La presente investigación es de carácter cuantitativo transversal, con enfoque descriptivo.

Los datos fueron obtenidos a través de la base de datos pública de la Comisión del Mercado Financiero (CMF), sobre prácticas de gobierno corporativo de sociedades anónimas abiertas fiscalizadas por esta entidad y que continuaban reportando bajo la Norma de Carácter General N°385, con información disponible hasta diciembre del 2024. Dicha información, se encuentra disponible a través de un formulario anualizado compuesto por 99 prácticas. Sin embargo, para el estudio se usaron 22 del cuestionario que corresponden a la categoría de control y gestión de los riesgos correspondiente a los períodos 2022 y 2023, los cuales incluyen a la mayor cantidad de respuestas de las sociedades. Las respuestas son de carácter dicotómica, indicando "Sí" en el caso de que la sociedad esté adoptando la práctica y "No" en caso contrario. A efectos del análisis de los datos, Si = 1 y No = 0. En cuanto a la información financiera de las sociedades, ésta se obtuvo de las estadísticas públicas de la CMF, referente a los indicadores financieros bajo estándar IFRS de emisores de valores, correspondiente al cierre de los estados financieros al 31 de diciembre de 2023. A la fecha, la información del período 2024 no se encuentra disponible.

La base de datos consta de 22 preguntas y respuestas sobre diferentes prácticas de gobierno corporativo de gestión de riesgos, información financiera y rubro por cada empresa. Para un análisis más general, se agruparon en 4 subcategorías que miden el cumplimiento de las prácticas. Por otra parte, considerando que los datos recolectados corresponden a cumplimiento de prácticas empresariales y datos financieros, la cantidad de empresas con datos completos asciende a 105 y corresponden al período 2022. En **Tabla 1** se listan los atributos.

**Tabla 1. - Descripción de la base de datos**

Atributo	Tipo de dato	Operación de cálculo
Procedimiento de gestión y control de riesgos (VAR1)	Numérico (%)	Sumatoria de prácticas adoptadas 9 prácticas
Canal de denuncias (VAR2)	Numérico (%)	Sumatoria de prácticas adoptadas 4 prácticas
Diversidad y sucesión (VAR3)	Numérico (%)	Sumatoria de prácticas adoptadas 5 prácticas

Estructura salarial y compensación de la alta gerencia (VAR4)	Numérico (%)	Sumatoria de prácticas adoptadas 4 prácticas
Rentabilidad sobre el patrimonio (ROE)	Numérico (%)	Ganancia (pérdida) *100 Patrimonio total - Ganancia (pérdida)
Rentabilidad sobre los activos (ROA)	Numérico (%)	Ganancia (pérdida) *100 Total de activos
Capital (CT)	Numérico (%)	(Activos corrientes totales - Pasivos corrientes totales) / 1000
Razón de endeudamiento (RE)	Numérico (%)	Total pasivos Patrimonio total
Industria	Carácter	Rubro de la empresa

Se realizó un Análisis Exploratorio de datos (EDA) de manera de tener un primer contacto con los datos (Meloun y Militký, 2011). A través de estadística descriptiva, correlación, gráfica univariada, bivariada y multivariada y del análisis de los mismos, se avanzó en técnicas que describen mejor la relación entre los datos.

Considerando que la base de datos no tiene datos etiquetados, se propusieron 2 modelos basados en K-Means y BIRCH para agrupar datos de manera de obtener conocimiento (oculto hasta el momento) de las empresas. La calidad de los agrupamientos fue medida con los índices de Silhouette y Calinski-Harabasz (Wang y Xu, 2019), para luego proceder al análisis e interpretación de los resultados alcanzados.

A partir del agrupamiento obtenido previamente, se desarrollaron 3 modelos basados en Random Forests, Support Vector Machines y Extreme Learning Machine para predecir condiciones de las empresas. Respecto a los atributos de entrada, se aplicó Feature Importance para reducir y elegir los más significativos. En cuanto a los modelos, se usó Cross-validation para el entrenamiento intensivo y Grid Search para la selección de los mejores hiper-parámetros (Belete y Huchaiah, 2021). Se usaron las métricas Accuracy, Precisión, Recall y F1-Score (Shobha y Rangaswamy, 2018) para medir la calidad de las predicciones.

## 4. Resultados

Como parte del análisis exploratorio de datos, se presenta la **Tabla 2** con estadística descriptiva.

**Tabla 2. - Estadística descriptiva**

Atributo	Promedio	D. Estándar	Mínimo	Máximo
VAR1	0,432804	0,409098	0,000000	1,000000
VAR2	0,561905	0,455708	0,000000	1,000000
VAR3	0,276190	0,367299	0,000000	1,000000

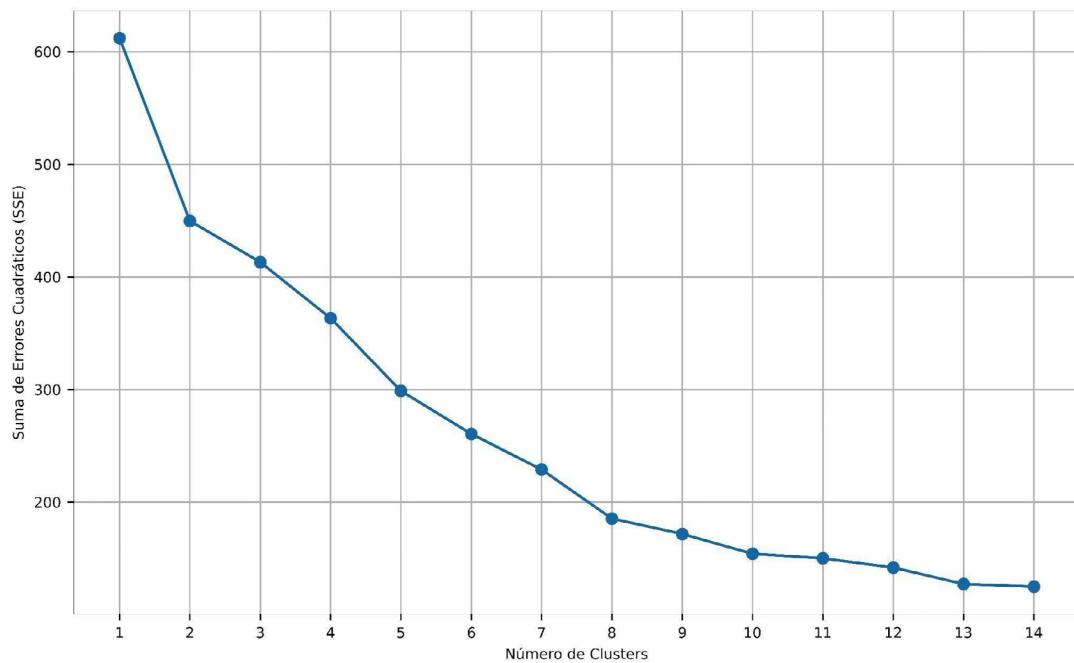
VAR4	0,185714	0,245215	0,000000	1,000000
ROE	12,187552	24,291913	-32,254500	172,401900
ROA	5,971389	14,459569	-28,178800	113,152000
CT	1,652334e+07	6,413463e+07	-4,254594e+08	2,101914e+08
RE	1,066659	1,130163	0,000200	4,837800

A nivel general, las empresas en promedio no superan el 56,2% de adopción de prácticas de gobierno corporativo, siendo el canal de denuncias la práctica con mayor nivel de adopción (56,2%), en contraste con la estructura salarial y compensación de la alta gerencia que es la menos adoptada (18,6%). En relación al rendimiento financiero, se observa que en promedio son resultados positivos. Sin embargo, los indicadores financieros presentan una alta dispersión de datos con respecto de la media.

En cuanto a rubros, las empresas relevadas pertenecen mayoritariamente a inversión (21,9% del total), inmobiliaria (9,5%), construcción (7,6%) y agropecuaria y forestal (7,6%). En el extremo opuesto se encuentran salud, supermercado, productos químicos, energía, servicios financieros, factoring, aguas sanitarias y servicios sanitarios que representan el 1% (del total) en cada caso.

Luego de una introducción a los datos, se procedió a desarrollar 2 modelos no supervisados basados en K-Means y BIRCH. Previo a esto, se utilizó el método de Elbow (Syakur et al., 2018) (**Figura 2**) de manera de tener un número apropiado de grupos a considerar.

**Figura 2. - Método de Elbow**



El número recomendado de clusters se sitúa entre 2 y 6. Considerando el tamaño de la base de datos, se optó por 4. En **Tabla 3** se presentan las métricas respecto a la calidad de los agrupamientos alcanzados por K-Means y BIRCH. Debido a la correlación detectada en los atributos durante el análisis exploratorio de datos, se usaron VAR2, VAR3, VAR4, ROA, CT y RE, los cuales fueron estandarizados.

**Tabla 3. - Calidad de agrupamientos para K=4**

Métrica	Métodos de agrupamiento	
	K-Means	BIRCH
Silhouette	0,18	0,27
Calinski-Harabasz	22,36	20,31

Se observa que según el índice de Silhouette, BIRCH alcanza un mejor resultado en tanto para Calinski-Harabasz, K-Means obtiene mejores agrupaciones. Ante esta paridad, se analizó la cantidad de casos por grupo, siendo el modelo basado en K-Means el elegido para el agrupamiento de datos.

En las **Tablas 4, 5, 6 y 7** se presentan un resumen estadístico descriptivo de los grupos de empresas obtenidos por K-Means.

**Tabla 4. - Estadística descriptiva del grupo 0 (35 empresas)**

Atributo	Promedio	D. Estándar	Mínimo	Máximo
VAR1	0,793651	0,286646	0,000000	1,000000
VAR2	0,900000	0,244047	0,000000	1,000000
VAR3	0,657143	0,364841	0,000000	1,000000
VAR4	0,442857	0,227565	0,000000	1,000000
ROE	8,843366	18,794909	-32,254500	50,991500
ROA	2,031783	6,961358	-23,347200	12,609400
CT	3,102286e+06	8,152703e+07	-4,254594e+08	1,375303e+08
RE	1,593509	1,229585	0,013800	4,837800

**Tabla 5. - Estadística descriptiva del grupo 1 (24 empresas)**

Atributo	Promedio	D. Estándar	Mínimo	Máximo
VAR1	0,319444	0,365314	0,000000	0,888889
VAR2	0,375000	0,423495	0,000000	1,000000
VAR3	0,133333	0,152277	0,000000	0,400000
VAR4	0,104167	0,145898	0,000000	0,500000
ROE	31,100583	36,291482	6,328400	172,401900
ROA	12,863142	5,509574	5,297000	26,618400
CT	4,981694e+06	8,159113e+06	1,864700e+04	2,573950e+07
RE	0,538062	0,499292	0,016900	1,600900

**Tabla 6. - Estadística descriptiva del grupo 2 (7 empresas)**

Atributo	Promedio	D. Estándar	Mínimo	Máximo
VAR1	0,444444	0,435087	0,000000	1,000000
VAR2	0,500000	0,478714	0,000000	1,000000
VAR3	0,057143	0,097590	0,000000	0,200000
VAR4	0,178571	0,188982	0,000000	0,500000
ROE	15,930671	7,769984	4,338000	23,102800
ROA	10,530371	7,421531	1,757100	18,645800
CT	1,723815e+08	3,646851e+07	1,322333e+08	2,101914e+08
RE	0,736057	0,801811	0,006500	2,032200

**Tabla 7. - Estadística descriptiva del grupo 3 (36 empresas)**

Atributo	Promedio	D. Estándar	Mínimo	Máximo
VAR1	0,163580	0,252656	0,000000	0,888889
VAR2	0,388889	0,456870	0,000000	1,000000
VAR3	0,038889	0,093435	0,000000	0,400000
VAR4	0,006944	0,041667	0,000000	0,250000
ROE	2,102328	10,952458	-28,787200	27,094400
ROA	0,072647	6,502638	-28,178800	7,110500
CT	6,770795e+06	1,433232e+07	-5,617038e+06	6,139936e+07
RE	0,992381	1,229045	0,000200	4,294700

Del análisis numérico de clusters, se observa que en promedio, el grupo 0 agrupa a empresas con cumplimiento alto de las políticas (~70%), rentabilidad baja sobre el patrimonio (~8,8%), capital de trabajo bajo (~3,1M CLP) y rentabilidad baja sobre los activos (~2%); el grupo 1 a empresas con cumplimiento bajo de las políticas (~23,3%), rentabilidad alta sobre el patrimonio (~31,1%), capital de trabajo bajo (~5M CLP) y rentabilidad alta sobre los activos (~12,9%); el grupo 2 a empresas con bajo cumplimiento (~29,5%), rentabilidad elevada sobre el patrimonio (~15,9%), capital de trabajo alto (~172M CLP) y rentabilidad elevada sobre los activos (~10,5%). Finalmente, el grupo 3 a empresas con muy bajo cumplimiento (~15%), rentabilidad baja sobre el patrimonio (~2,1%), capital de trabajo bajo (~6,7M CLP) y rentabilidad baja sobre los activos (~0,1%).

Se desarrollaron 3 modelos predictivos basados en Random Forests (RF), Support Vector Machines (SVM) y Extreme Learning Machine (ELM) para predecir características de las empresas, a partir del agrupamiento anterior.

Para la selección de los mejores atributos de entrada, se usó la técnica Feature Importance. En la **Tabla 8** se presentan los valores de importancia.

**Tabla 8. - Feature Importance: importancia de los atributos**

Atributo	Puntuación
VAR4	0,25
ROA	0,24
CT	0,14
VAR3	0,10
VAR1	0,08
ROE	0,08
RE	0,04
Industria	0,03
VAR2	0,03

Los modelos predictivos reciben los mismos atributos de entrada. En este caso, se eligieron la menor cantidad de atributos y a la vez, los más significativos para los modelos: VAR4, ROA, CT, VAR3 y ROE. Fueron entrenados con el 75% de los datos disponibles y el 25% restante se utilizó para la prueba de los mismos. Para los modelos basados en Random Forests y Support Vector Machines, se intensificó el entrenamiento mediante Cross-Validation con 5 Folds y 3 repeticiones. Para Extreme Learning Machine, esta opción fue descartada. Para obtener modelos más precisos, se propusieron diferentes valores para los hiper-parámetros más relevantes de cada una de las técnicas consideradas. Los modelos fueron entrenados y evaluados en el entorno Google Colab.

En **Tabla 9** se muestra el reporte de clasificación durante la fase de prueba.

**Tabla 9. - Reporte de clasificación para los modelos predictivos**

Modelo	Métrica	Modelo	Grupo			
			0	1	2	3
RF	Accuracy (%)	85				
	Precision (%)		90	88	100	71
	Recall (%)		75	100	100	83
	F1-Score (%)		82	93	100	77
SVM	Accuracy (%)	85				
	Precision (%)		100	70	50	100
	Recall (%)		67	100	100	100
	F1-Score (%)		80	82	67	100
ELM	Accuracy (%)	96				
	Precision (%)		100	88	100	100
	Recall (%)		92	100	100	100
	F1-Score (%)		96	93	100	100

Los valores de los hiper-parámetros asociados al modelo RF son:

- Número de árboles: 700
- Máxima profundidad del árbol: 3
- Criterio: entropy
- Número máximo de predictores: sqrt

El tiempo total de CPU para entrenamiento y evaluación del modelo fue 639 segundos, en tanto que el valor de accuracy en fase de entrenamiento fue 97%.

Respecto al modelo SVM, los mejores valores de los hiper-parámetros son:

- Peso de clases: balanceado
- Gamma: 0,1
- kernel: sigmoid

El tiempo total de CPU para entrenamiento y evaluación del modelo fue 14 segundos y el accuracy en fase de entrenamiento fue 96%.

En cuanto a ELM, los valores de los hiper-parámetros son:

- Número de neuronas: 1000
- Función de activación: relu
- C (sesgo): 1

El tiempo de CPU fue de 786 segundos y el valor de accuracy durante el entrenamiento fue 95%.

El modelo RF alcanza un valor de accuracy del 85%, en tanto que, respecto a las clases/grupos, logra valores de f1-score entre 77% (para el grupo 3) y 100% (grupo 2). El modelo SVM también logra un valor de accuracy de 85%. En cuanto a f1-score, consigue valores entre 67% (grupo 2) y 100% (grupo 3). Finalmente, el modelo ELM obtiene un valor de accuracy de 96% con f1-score entre 93% (clase 1) y 100% (clase 2 y 3).

Como se mencionó anteriormente, para obtener modelos de agrupamiento de datos y predictivos más precisos, se usó la base de datos de sociedades anónimas chilenas del año 2022 por ser la más reciente y con mayor cantidad de casos<sup>2</sup>. En **tabla 10** se muestra el resultado de algunas predicciones del modelo basado en Extreme Learning Machine para sociedades relevadas en el año 2023 con y sin relevamiento en el periodo anterior.

**Tabla 10. - Predicciones del modelo ELM sobre sociedades anónimas chilenas en 2023**

Nombre	VAR3	VAR4	ROE	ROA	CT	Resultado (clase predicha)	Obs.
<i>Entidad 1</i>	0,4	0,3	39,5679	10,6601	1918821	1	Sin datos disponibles en 2022
<i>Entidad 2</i>	0,0	0,0	0,6331	0,5273	47960356	3	
<i>Entidad 3</i>	0,2	0,0	-6,1211	-3,9558	0	3	Mantiene clasificación de 2022
<i>Entidad 4</i>	0,0	0,0	2,9711	2,0122	28645	3	Cambia de clasificación respecto a 2022

Respecto a sociedades que no fueron relevadas en 2022, para la Entidad 1 el modelo ELM predice el grupo/clase 1 que agrupa a sociedades con rentabilidades altas; para la Entidad 2, predice el grupo 3 que corresponde a sociedades con rentabilidades bajas. En cuanto a sociedades relevadas en 2022, para la Entidad 3, el modelo devuelve el grupo 3 el cual coincide respecto al año anterior debido a un rendimiento similar; en cambio para la Entidad 4, predice el cluster 3. Esto difiere de 2022 donde la sociedad tuvo un mejor rendimiento financiero por lo que el modelo ELM le había asignado el grupo 1.

## 5. Discusión

A nivel global, la dimensión de prácticas de gestión y control de riesgos, relacionada al canal de denuncias, se presenta como la dimensión con mayor adopción por parte de las empresas. En cambio, el menor nivel de adopción está representado por la dimensión de estructura salarial y compensación de alta gerencia, así lo reafirma Arenas Torres et al. (2019).

Considerando el cumplimiento en gestión de riesgos, el cluster 0 (**Tabla 4**), incluye empresas con alto nivel de cumplimiento de políticas (>70%), rentabilidad baja sobre el patrimonio (8,8%) capital de trabajo bajo (3,1M CLP) y rentabilidad baja sobre los activos, esto en contraste con lo que señala Musallam (2020) indicando que la gestión de riesgos mejora el desempeño corporativo, así lo reafirma, Nahar et al. (2020); Horvey y Ankamah (2020) mencionan que se presenta una relación no lineal entre ambos al no ser estable a largo plazo, debido a factores externos que pueden afectar al ERM, mientras que por otra parte Agustina y Baroroh (2016) manifiestan que la implementación de la Gestión de Riesgos Empresariales (ERM) no tiene una influencia significativa en el valor de la empresa, asimismo lo reafirma, Moraga y Ropero (2018). No obstante, Ayuningtyas y Harymawan (2022) destacan la relación positiva entre la gobernanza de la empresa con la divulgación sobre gestión de riesgo, dado que una buena gobernanza permite un mejor control en la entidad. Por último, sólo el cluster 3 (Tabla 7) considera a las empresas con cumplimiento bajo de políticas, rentabilidad baja sobre el patrimonio y sobre los activos en comparación a los demás clusters, así lo reafirma Orostiga Lazo y Lillo (2024).

Los modelos basados en Random Forests, Support Vector Machines y Extreme Learning Machine para predecir características de las empresas, alcanzaron altos niveles de precisión (a nivel de modelo, se alcanzaron niveles de accuracy del 85% y 96%) del nivel de cumplimiento de prácticas de gestión y control de los riesgos, utilizando un reducido número de atributos (12 de un total de 27). Esto en contraste, con lo presentado por Navarrete Oyarce et al. (2023) donde formulan un índice para medir el grado de adherencia a las prácticas de gobierno corporativo de la NCG 385, considerando sus 99 prácticas en conjunto para estimar su efecto sobre el desempeño financiero de las empresas. Los excelentes resultados de los modelos supervisados, se basan en una apropiada transformación y selección de atributos, entrenamiento intensivo, configuración de hiper-parámetros y técnicas predictivas eficaces. Asimismo, los tiempos de CPU requeridos son bajos en los 3 casos.

Finalmente, los resultados evidencian que las empresas chilenas que reportan a la Comisión del Mercado Financiero, han tenido un avance poco significativo, por lo que, aún se encuentran en una etapa temprana de desarrollo respecto de prácticas de gobierno corporativo (**Tabla 2**), tal como lo

indica Arenas-Torres et al. (2021). Asimismo, Severino González y Acuña Moraga (2019) mencionan que el bajo cumplimiento se debe a que las empresas no tienen un mejor conocimiento de las nuevas acciones que deben realizarse internamente, lo que no permite generar las condiciones necesarias para el cumplimiento de las acciones declaradas en la norma.

## 6. Conclusiones

La investigación realizada tuvo por objetivo analizar el nivel de cumplimiento de prácticas de gobierno corporativo relacionadas a la gestión de riesgos y desempeño financiero de las empresas chilenas, utilizando técnicas de Ciencia de Datos. Los resultados obtenidos muestran que la adopción de prácticas de gestión y control de riesgos en las sociedades chilenas que reportan a la Comisión del Mercado Financiero es limitada y se encuentra en una etapa temprana de desarrollo. Si bien la dimensión sobre el canal de denuncias presenta el mayor nivel de cumplimiento, la estructura salarial y compensación de la alta gerencia es la menos adoptada. Esta situación indica una disparidad en la implementación de políticas de gobierno corporativo, lo que podría estar vinculado con una falta de internalización cultural de la gestión de riesgos en las empresas.

El análisis de agrupamiento reveló la existencia de 4 categorías de empresas diferenciadas según el grado de cumplimiento de prácticas de gestión de riesgos y su desempeño financiero. Las empresas con mayor cumplimiento tienden a presentar rentabilidades más bajas, en contraste con aquellas con menor adopción, las cuales paradójicamente, registran mayores indicadores de rentabilidad. Este hallazgo sugiere que el impacto de la gestión de riesgos sobre el rendimiento financiero no es lineal ni homogéneo, respaldando estudios previos que plantean resultados contradictorios sobre esta relación.

En cuanto a los modelos supervisados, éstos mostraron un alto nivel de precisión en la clasificación de las empresas (siendo los atributos más significativos VAR4, ROA, CT, VAR3 y ROE), destacándose el modelo de Extreme Learning Machine con una precisión del 96%. Este resultado confirma la efectividad del uso de la Ciencia de Datos para mejorar el conocimiento organizacional y la capacidad de anticipación en la gestión de riesgos.

Finalmente, el bajo nivel general de cumplimiento observado reafirma la necesidad de fortalecer las estrategias de gobernanza y fomentar una mayor adopción de prácticas de gestión de riesgos. La investigación futura podría centrarse en ampliar la muestra de empresas e incorporar variables adicionales como métricas ESG, satisfacción del cliente, compromiso de los empleados, para pronosticar mejor el cumplimiento y la eficacia de las prácticas de gobierno corporativo y su desempeño empresarial.

## Referencias

- Agustina, L.; Baroroh, N. (2016). "The relationship between Enterprise Risk Management (ERM) and firm value mediated through the financial performance", *Review of Integrative Business and Economics Research*, Vol. 5, Num. 1, pp. 128-138. Obtenido de [https://sibresearch.org/uploads/2/7/9/9/2799227/ribcr\\_h15-081\\_128-138.pdf](https://sibresearch.org/uploads/2/7/9/9/2799227/ribcr_h15-081_128-138.pdf)
- Abmad, A; Muhammad, M.; Narullia, D. (2021). "Corporate Risk Disclosure: The Effect of Corporate Governance", *Journal of Applied Managerial Accounting*, Vol. 5, Num. 1, pp. 101-113. <https://doi.org/10.30871/jama.v5i1.2794>
- Arenas Torres, F.; Campos Troncoso, R.; Santander Ramírez, V. (2019). "Adopción de prácticas de gobierno corporativo del mercado chileno en el trienio 2015-2017", *Revista Científica Visión De Futuro*, Vol. 23, Num. 2, pp. 37-55. Obtenido de <https://visiondefuturo.fee.unam.edu.ar/index.php/visiondefuturo/article/view/371>
- Arenas-Torres, F.; Bustamante-Ubilla, M.; Campos-Troncoso, R. (2021). "La incidencia de la adopción de prácticas de gobierno corporativo en las rentabilidades de las sociedades anónimas chilenas", *Journal of Globalization, Competitiveness and Governability*, Vol. 15, Num. 3, pp. 79-93. <http://doi.org/10.3232/GCG.2021.V15.N3.04>
- Ayuningtyas, E.; Harymawan, I. (2022). "Risk Management Committee and Textual Risk Disclosure", *Risks*, Vol. 10, Num. 2, 30. <https://doi.org/10.3390/risks10020030>
- Barbaglia, L.; Consoli, S.; Manzan, S.; Reforgiato Recupero, D.; Saisana, M.; Tiozzo Pezzoli, L.; (2021). "Data Science technologies in Economics and Finance: A gentle walk-in", in: Consoli, S.; Reforgiato Recupero, D.; Saisana, M.; (Eds.), *Data Science for Economics and Finance: Methodologies and Applications*, Springer Nature, Switzerland. [https://doi.org/10.1007/978-3-030-66891-4\\_1](https://doi.org/10.1007/978-3-030-66891-4_1)
- Belete, D.; Huchaiah, M. (2021). "Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results", *International Journal of Computers and Applications*, Vol. 44, Num. 9, pp. 875-886. <https://doi.org/10.1080/1206212X.2021.1974663>
- Breiman, L. (2001). "Random Forests", *Machine Learning*, Vol. 45, pp. 5-32. <https://doi.org/10.1023/A:1010933404324>
- Bui, H.; Krajcsák, Z. (2024). "The impacts of corporate governance on firms performance: from theories and approaches to empirical findings", *Journal of Financial Regulation and Compliance*, Vol. 32, Num. 1, pp. 18-46. <https://doi.org/10.1108/JFRC-01-2023-0012>
- Cravero, A.; Sepúlveda, S. (2021). "Use and adaptations of Machine Learning in Big Data applications in real cases in Agriculture", *Electronics*, Vol. 10, Num. 5, 552. <https://doi.org/10.3390/electronics10050552>
- D'Urso, P.; De Giovanni, L.; Swartz, T. (2023). "Editorial: Big data and data science in sport", *Annals of Operational Research*, Vol. 325, pp. 1-7. <https://doi.org/10.1007/s10479-023-05344-z>
- Fayyad, U.; Stolorz, P. (1997). "Data mining and KDD: promise and challenges", *Future Generation Computer Systems*, Vol. 13, Num. 2-3, pp. 99-115. [https://doi.org/10.1016/S0167-739X\(97\)00015-0](https://doi.org/10.1016/S0167-739X(97)00015-0)
- Flach, P. (2001). "On the state of the art in Machine Learning: a personal review", *Artificial intelligence*, Vol. 131, Num. 1-2, pp. 199-222. [https://doi.org/10.1016/S0004-3702\(01\)00125-4](https://doi.org/10.1016/S0004-3702(01)00125-4)
- Ghabramani, Z. (2004). "Unsupervised Learning" In: Bousquet, O., von Luxburg, U., Rätsch, G. (eds) *Advanced Lectures on Machine Learning. ML 2003. Lecture Notes in Computer Science*, Vol. 3176. Springer, Berlin. [https://doi.org/10.1007/978-3-540-28650-9\\_5](https://doi.org/10.1007/978-3-540-28650-9_5)
- Gouiaa, R. (2018). "Analysis of the effect of corporate governance attributes on risk management practices", *Risk Governance and Control: Financial Markets & Institutions*, Vol. 8, Num. 1, pp. 14-23. <https://doi.org/10.22495/rgev8i1art2>
- Horvey, S.; Ankamah, J. (2020). "Enterprise risk management and firm performance: empirical evidence from Ghana equity market", *Cogent Economics & Finance*, Vol. 8, Num. 1, 1840102. <https://doi.org/10.1080/23322039.2020.1840102>

- Liu, M.; Huang, Y. (2017). "The use of data science for education: the case of social-emotional learning", *Smart Learning Environments*, Vol. 4, 1. <https://doi.org/10.1186/s40561-016-0040-4>
- Ibarra, C. (2020). "Técnicas de Data Mining aplicadas a la deserción de los estudiantes de la Facultad de Ciencias Exactas", tesis de maestría, Universidad Santo Tomás de Aquino, Argentina. <https://doi.org/10.13140/RG.2.2.29986.66244>
- Jahani, H.; Jain, R.; Ivanov, D. (2023). "Data science and big data analytics: a systematic review of methodologies used in the supply chain and logistics research", *Annals of Operations Research*, pp. 1-58. <https://doi.org/10.1007/s10479-023-05390-7>
- Malik, M.; Zaman, M.; Buckby, S. (2020). "Enterprise risk management and firm performance: Role of the risk committee," *Journal of Contemporary Accounting and Economics*, Vol. 16, Num. 1, 100178. <https://doi.org/10.1016/j.jcae.2019.100178>
- Meloun, M.; Militký, J. (2011). "The exploratory and confirmatory analysis of univariate data", In: Meloun, S.; Militký, J. (Eds.), *Statistical Data Analysis*, pp. 25-71. Woodhead, India. <https://doi.org/10.1533/9780857097200.25>
- Mora Navarro, O. (2022). "Gestión de riesgos: un desafío para las organizaciones", *Administración & Desarrollo*, Vol. 52, Num. 1, pp. 4-19. <https://doi.org/10.22431/25005227.vol52n1.1>
- Moraga Flores, H.; Ropero Moriones, E. (2017). "Gobierno Corporativo y desempeño financiero de las empresas chilenas", *CAPIC REVIEW*, Vol. 15, pp. 31-43. <https://doi.org/10.35928/cr.vol15.2017.7>
- Moraga, H.; Ropero, E. (2018). "Gobierno Corporativo y desempeño financiero de las empresas más importantes del mercado bursátil chileno", *Revista Venezolana de Gerencia*, Vol. 23, Num. 81, pp. 145-162. <https://doi.org/10.31876/revista.v23i81.23473>
- Moraga-Flores, H.; Rossi-Undurraga, M. (2019). "Gobierno corporativo y riesgo de quiebra en las empresas chilenas", *Journal of Globalization, Competitiveness and Governability*, Vol. 13, Num. 1, pp. 112-125. <https://doi.org/10.3232/GCG.2019.V13.N1.06>
- Musallam, S. (2020). "Effects of board characteristics, audit committee and risk management on corporate performance: evidence from Palestinian listed companies", *International Journal of Islamic and Middle Eastern Finance and Management*, Vol. 13, Num. 4, pp. 691-706. <https://doi.org/10.1108/IJIFM-12-2017-0347>
- Nahar, S.; Azim, M.; Hossain, M. (2020). "Risk disclosure and risk governance characteristics: evidence from a developing economy", *International Journal of Accounting & Information Management*, Vol. 28, Num. 4, pp. 577-605. <https://doi.org/10.1108/IJAIM-07-2019-0083>
- Navarrete Oyarce, J.; Moraga Flores, H.; Gallegos Mardones, J. (2023). "Index to degree of adhesion to good practices of corporate governance and their effect on financial performance: evidence for Chilean companies", *Economic Research-Ekonomska Istraživanja*, Vol. 36, Num. 1, pp. 2527-2544. <https://doi.org/10.1080/1331677X.2022.2101016>
- Orostiga Lazo, V.; Lillo, F. (2024). "Clustering model to classify firms according to their Management and Risk Management Practices from a Financial Perspective Societies' Risk Control: a financial Clustering model for classification", *Review of Economics and Finance*, Vol. 22, pp. 180-188. <https://doi.org/10.55365/1923.x2024.22.21>
- Plathottam, S.; Rzonca, A.; Lakhnori, R.; Iloeje, C. (2023). "A review of Artificial Intelligence applications in manufacturing operations", *Journal of Advanced Manufacturing and Processing*, Vol. 5, Num. 3, e10159. <https://doi.org/10.1002/amp.2.10159>
- Ramlee, R.; Ahmad, N. (2020). "Malaysian Risk Management Committees and Firms' Financial Performance", *Asia-Pacific Management Accounting Journal*. Vol. 15, Num. 2, pp. 147-167. <https://doi.org/10.24191/APMAJ.v15i2-07>
- Robles, A.; Castañeda, A.; Carrizo, J. (2019). "Gestión de riesgos corporativos y la necesidad de su regulación en las empresas argentinas", *Revista de Ciencias Empresariales-Universidad Blas Pascal*, Vol. 4, Num. 4, pp. 53-67. [https://doi.org/10.37767/2468-9785\(2019\)005](https://doi.org/10.37767/2468-9785(2019)005)

Severino González, P.; Acuña Moraga, O. (2019). "Corporate governance: Ranking of chilean companies with greater transaction in the stock exchange", *Revista Lasallista de Investigación*, Vol. 16, Num. 1, pp. 103–114. <https://doi.org/10.22507/RJI.V16N1A3>

Shobha, G.; Rangaswamy, S. (2018). Machine learning. In: Gudivada, V; Rao, C. (Eds), *Handbook of Statistics*, Vol. 38, pp. 197-228. Elsevier. <https://doi.org/10.1016/bs.host.2018.07.004>

Subrahmanyam, S.; Shetty, D.; Patil, V.; Hameed, B.; Paul, R.; Smriti, K.; Naik, N.; Somani, B. (2022). "The role of data science in healthcare advancements: applications, benefits, and future prospects", *Irish Journal of Medical Science*, Vol. 191, pp. 1473-1483. <https://doi.org/10.1007/s11845-021-02730-z>

Syakur, M.; Khotimah, B.; Rochman, E.; Satoto, B. (2018), "Integration K-Means clustering method and Elbow method for identification of the best customer profile cluster", *IOP Conference Series: Materials Science and Engineering*, Vol. 336, 012017. <https://doi.org/10.1088/1757-899X/336/1/012017>

Wang, J.; Lu, S.; Wang, S.; Zhang, Y. (2022). "A review on Extreme Learning Machine", *Multimedia Tools and Applications*, Vol. 81, pp. 41611-41660. <https://doi.org/10.1007/s11042-021-11007-7>

Wang, X.; Xu, Y. (2019). "An improved index for clustering validation based on Silhouette index and Calinski-Harabasz index", *IOP Conference Series: Materials Science and Engineering*, Vol. 569, 052024. <https://doi.org/10.1088/1757-899X/569/5/052024>

Zhang, T.; Ramakrishnan, R.; Livny, M. (1997). "BIRCH: a new data clustering algorithm and its applications", *Data mining and knowledge discovery*, Vol. 1, pp. 141-182. <https://doi.org/10.1023/A:1009783824328>

.....

### Notas

2. La cantidad de sociedades anónimas chilenas con datos completos respecto al periodo 2023 asciende a 52