

UNIVERSIDAD DEL VALLE DE GUATEMALA  
Facultad de Ingeniería



**Factores que influyen en las evaluaciones de docentes en la  
Universidad del Valle de Guatemala**

Trabajo de graduación presentado por Sara María Paguaga Gonzáles  
para optar al grado académico de Licenciada en Ingeniería en Ciencia de  
la Computación y Tecnologías de la Información

Guatemala,

2025



UNIVERSIDAD DEL VALLE DE GUATEMALA  
Facultad de Ingeniería



**Factores que influyen en las evaluaciones de docentes en la  
Universidad del Valle de Guatemala**

Trabajo de graduación presentado por Sara María Paguaga Gonzáles  
para optar al grado académico de Licenciada en Ingeniería en Ciencia de  
la Computación y Tecnologías de la Información


Guatemala,

2025

Vo.Bo.:

(F)   
Msc. Lynette García Pérez

Tribunal:

(F)   
Msc. Douglas Leonel Barrios

(F)   
Msc. Lynette García Pérez

Fecha de aprobación:

6 de junio de 2025

---

## Agradecimientos

---

Esta etapa es un cierre de ciclo muy importante: un periodo lleno de aprendizajes, retos y desafíos, los cuales me llevaron a tener aprendizajes significativos y experiencias enriquecedoras. Fue un tiempo en el que estuve rodeada y acompañada de personas excepcionales. Es por ello que quiero comenzar agradeciéndole a Dios, porque sus planes siempre son perfectos y su bondad es infinita, así como a la Virgen María por su intercesión y ayuda.

Quiero agradecerle a mis papás porque siempre me acompañaron y guiaron en esta etapa. Creyeron en mí desde un inicio, me alentaron a seguir mis sueños y fueron mi soporte y mi gran ejemplo a seguir.

A mi mamá por enseñarme a nunca rendirme, a seguir adelante y apoyarme siempre sin importar las circunstancias. Gracias por inculcarme ese espíritu de lucha y el deseo de alcanzar mis metas.

A mi papá por inculcarme la pasión por el estudio, por sus sabios consejos y grandes reflexiones.

A mi hermano, Juan Carlos, por siempre inspirarme.

A mi asesora, la ingeniera Lynette García, por todo su apoyo, orientación y mentoría en este proceso.

Al ingeniero, Douglas Barrios, por ser un guía importante a lo largo de la carrera.

A todos aquellos maestros que con sus enseñanzas compartieron no solo conocimientos académicos, sino también valores fundamentales y la preparación necesaria para el mundo profesional.

A mis médicos y fisioterapeutas por su apoyo y darme una mejor calidad de vida.

A todos aquellos familiares y amigos que estuvieron presentes en esta etapa, brindándome siempre apoyo y motivación.

<b>Agradecimientos</b>	<b>III</b>
<b>Lista de figuras</b>	<b>VII</b>
<b>Resumen</b>	<b>VIII</b>
<b>Abstract</b>	<b>IX</b>
<b>1. Introducción</b>	<b>1</b>
<b>2. Justificación</b>	<b>3</b>
<b>3. Objetivos</b>	<b>5</b>
3.1. Objetivo general . . . . .	5
3.2. Objetivos específicos . . . . .	5
<b>4. Antecedentes</b>	<b>6</b>
<b>5. Marco teórico</b>	<b>8</b>
5.1. Evaluación de desempeño docente . . . . .	8
5.1.1. Validez del instrumento de evaluación . . . . .	8
5.1.2. Modelos de evaluación docente . . . . .	9
5.1.3. Criterios de evaluación . . . . .	10
5.1.4. Factores contextuales . . . . .	10
5.1.5. Aspectos didácticos . . . . .	11
5.1.6. Actores en el proceso de evaluación . . . . .	12
5.2. Análisis exploratorio . . . . .	13
5.3. Procesamiento de lenguaje natural . . . . .	13
5.3.1. Análisis de sentimientos . . . . .	13
5.3.2. Estandarización de datos: lematización . . . . .	14
5.4. Aprendizaje de máquinas . . . . .	14
5.4.1. Árboles de decisión . . . . .	14
5.4.2. Reglas de asociación . . . . .	15

5.5.	Análisis de grupos . . . . .	15
5.5.1.	Tipos de clústeres . . . . .	16
5.5.2.	Métodos de agrupamiento . . . . .	16
5.5.3.	Medidas de similitud . . . . .	16
5.5.4.	Evaluación de la calidad del agrupamiento . . . . .	16
5.5.5.	Aplicaciones . . . . .	17
<b>6.</b>	<b>Metodología</b>	<b>18</b>
6.1.	Recolección de datos . . . . .	18
6.2.	Limpieza y preprocesamiento de datos . . . . .	18
6.3.	Privacidad de los datos . . . . .	18
6.4.	Análisis de sentimientos . . . . .	19
6.5.	Análisis exploratorio . . . . .	19
6.6.	Algoritmos de aprendizaje automático no supervisados . . . . .	19
<b>7.</b>	<b>Resultados</b>	<b>20</b>
7.1.	Evaluación docente a lo largo del tiempo . . . . .	22
7.1.1.	Impacto de características del curso en las evaluaciones . . . . .	23
7.2.	Participación de estudiantes en evaluaciones . . . . .	25
7.2.1.	Análisis de comentarios . . . . .	25
7.2.2.	Cantidad de créditos . . . . .	28
7.2.3.	Año académico correspondiente a los cursos . . . . .	29
7.3.	Relación de la nota docente con el desempeño de los estudiantes . . . . .	33
7.4.	Instrumento de evaluación de desempeño docente . . . . .	35
7.4.1.	Árboles de decisión . . . . .	38
7.4.2.	Análisis de cantidad de respuestas por número de estrellas . . . . .	40
7.4.3.	Reglas de asociación . . . . .	43
7.5.	Agrupamiento . . . . .	52
7.6.	Criterios de evaluación . . . . .	55
<b>8.</b>	<b>Conclusiones</b>	<b>65</b>
<b>9.</b>	<b>Recomendaciones</b>	<b>67</b>
<b>10.</b>	<b>Bibliografía</b>	<b>68</b>

---

## Lista de figuras

---

1.	Distribución general de las notas de evaluación docente . . . . .	22
2.	Distribución anual de las notas de evaluación docente . . . . .	23
3.	Cantidad de cursos con mejor evaluación en cada ciclo . . . . .	23
4.	Comparación de cantidad de cursos donde fue mejor teoría o laboratorio por año . . . . .	24
5.	Densidad conjunta entre estudiantes inscritos y evaluadores . . . . .	25
6.	Distribución de polaridad de comentarios . . . . .	26
7.	Relación entre porcentaje de estudiantes que evaluaron y cantidad de comentarios positivos . . . . .	27
8.	Relación entre porcentaje de estudiantes que evaluaron y cantidad de comentarios negativos . . . . .	27
9.	Porcentaje promedio de estudiantes que evaluaron por cantidad de créditos correspondientes a un curso . . . . .	28
10.	Promedio de la nota de evaluación del docente por cantidad de créditos del curso . . . . .	28
11.	Cantidad promedio de profesores evaluados por año académico correspondiente a curso . . . . .	29
12.	Distribución de notas de evaluación por el año académico que le corresponde al curso . . . . .	29
13.	Relación entre el porcentaje de estudiantes que evaluaron y la nota de evaluación del docente . . . . .	30
14.	Relación entre el porcentaje de estudiantes que evaluaron y la nota de evaluación del docente (Notas menores o iguales a 80) . . . . .	31
15.	Relación entre el porcentaje de estudiantes que evaluaron y la nota de evaluación del docente (Notas menores o iguales a 60) . . . . .	31
16.	Distribución de evaluación docente según grupos de porcentaje de estudiantes evaluadores . . . . .	32
17.	Relación entre la nota de evaluación del docente vs nota promedio de la sección	33
18.	Relación entre nota de evaluación del docente y porcentaje de alumnos aprobados . . . . .	34
19.	Relación entre nota de evaluación del docente y porcentaje de alumnos reprobados . . . . .	34
20.	Instrumento de evaluación actualmente (2025, primer ciclo) . . . . .	35

21.	Cantidad de profesores con una nota mayor a 80 vs con menor a 80 . . . . .	36
22.	Correlación de variables con la nota de evaluación . . . . .	37
23.	Comparación de calificaciones promedio por estrellas y estado de la nota del docente, más de 80 o menos de 80 . . . . .	41
24.	Promedio de cantidad de respuestas por nivel de evaluación . . . . .	42
25.	Reglas de asociación (Confianza) . . . . .	45
26.	Reglas de asociación ( <i>Lift</i> ) . . . . .	46
27.	Red de asociación de notas mayores a 80 puntos . . . . .	48
28.	Red de asociación de notas menores a 80 . . . . .	51
29.	Método del codo para determinar el número óptimo de grupos . . . . .	52
30.	Características de los grupos . . . . .	52
31.	Promedio de cantidad de estrellas por pregunta (1 y 2 estrellas) de 2017 a 2019	56
32.	Promedio de cantidad de estrellas por pregunta (3 y 4 estrellas) de 2017 a 2019	57
33.	Promedio de cantidad de estrellas por pregunta (1 y 2 estrellas) de 2020 a 2022	60
34.	Promedio de cantidad de estrellas por pregunta (3 y 4 estrellas) de 2020 a 2022	61
35.	Promedio de cantidad de estrellas por pregunta (1 y 2 estrellas) de 2023 . . .	62
36.	Promedio de cantidad de estrellas por pregunta (3 y 4 estrellas) de 2023 . . .	63

El propósito de este trabajo es identificar factores que inciden en las evaluaciones de desempeño docente, a partir de los resultados obtenidos de las evaluaciones realizadas por estudiantes de licenciatura de la Universidad del Valle del año 2017 al 2023. Asimismo, se busca analizar tanto la relación como la influencia de las variables relevantes disponibles, con el fin de obtener una panorama más completo de los resultados obtenidos en las evaluaciones. Para alcanzar los objetivos propuestos, se realizará un análisis exploratorio de las variables cuantitativas, que incluirá estadísticas descriptivas y medidas de tendencia central y dispersión. Además, para la variable cualitativa del estudio: polaridad de los comentarios, se realizará un etiquetado de los comentarios con el lexicón ML-Senticon y en base a ello un análisis de sentimientos. Se espera que en base a los resultados de los análisis realizados se logren encontrar patrones y métricas relevantes que apunten a los principales factores influyentes en las evaluaciones de desempeño docente.

The purpose of this work is to identify factors that influence teaching performance evaluations, based on the results obtained from the evaluations performed by undergraduate students at Universidad del Valle from 2017 to 2023. Likewise, it seeks to analyze both the relationship and the influence of the relevant variables available, in order to obtain a more complete picture of the results obtained in the evaluations. To achieve the proposed objectives, an exploratory analysis of the quantitative variables will be carried out, including descriptive statistics and measures of central tendency and dispersion. In addition, for the qualitative variable of the study: polarity of the comments, a labeling of the comments with the ML-Senticon lexicon and based on this a sentiment analysis will be performed. It is expected that based on the results of the analyses performed, it will be possible to find relevant patterns and metrics that point to the main influential factors in teacher performance evaluations.

La Universidad del Valle de Guatemala se ha distinguido a lo largo de los años, por su compromiso con valores y competencias que generan un cambio positivo, así como con la integridad en el ámbito académico. Estos atributos se han consolidado gracias al compromiso con una serie de estándares, normativas y evaluaciones continuas, enfocadas en mantener la excelencia académica.

La importancia de evaluar el desempeño docente ha sido reconocida con anterioridad, dado que favorece al proceso de aprendizaje y hay una mayor conciencia sobre las áreas que necesitan ser mejoradas. La herramienta de evaluación docente más utilizada a lo largo del tiempo ha sido el cuestionario. Sin embargo, ha sido sugerido utilizarla en conjunto con otro método evaluativo para poder cubrir un espectro más amplio de los aspectos docentes. Este enfoque ha sido implementado en las evaluaciones de la universidad, donde se complementa el cuestionario con una sección de comentarios para poder obtener retroalimentación directa por parte de los estudiantes [1].

En estudios previos sobre la evaluación del desempeño docente, se han considerado aspectos como el enfoque pedagógico del profesor, la planificación de las clases y las estrategias didácticas empleadas [2]. No obstante, este proyecto plantea otro enfoque al examinar factores cuantificables a través del análisis de datos y textos en los comentarios. Se consideran factores como el rendimiento de los estudiantes (determinado por el promedio de la sección), el grado académico de los estudiantes que efectúan las evaluaciones, el año en que se han realizado dichas evaluaciones, la nota general del docente, la polaridad de los comentarios, entre otros. Estas variables son tomadas en cuenta dado que se pretende determinar su influencia en las evaluaciones de desempeño docente.

Para evaluar el desempeño docente de los catedráticos, se implementan revisiones del rendimiento docente a través de encuestas estudiantiles. Estas abarcan aspectos tanto cuantitativos como cualitativos, por medio de preguntas de escala específicas y la sección de comentarios. Esto posibilita realizar diversos análisis y complementarlos entre sí, para entender de una mejor forma los resultados.

Los datos cualitativos obtenidos permiten realizar un análisis de sentimientos en los comentarios de los estudiantes, identificando la proporción de comentarios neutros, positivos o negativos [3].

Analizar los comentarios de los estudiantes es fundamental porque esta sección del instrumento de evaluación brinda al estudiante, la oportunidad de expresar de manera libre y sin restricciones su opinión personal sobre el desempeño de sus docentes. Además, los estudiantes tienen una perspectiva única para evaluar la enseñanza, ya que pueden reflexionar sobre su propio aprendizaje y comprender de manera integral el proceso de enseñanza-aprendizaje [4].

La evaluación de los comentarios de los estudiantes es valiosa debido a la diversidad de opiniones que puedan proporcionar. Según Zimmaro, Gaede, Heikes, Shim y Lewis [5], los estudiantes no solo expresan críticas o quejas, si no también, recomendaciones para mejorar, lo cual puede ofrecer a los docentes soluciones prácticas para abordar problemas que los métodos cuantitativos no pueden captar. Asimismo, considerar el grado académico como variable permite identificar puntos en común y evaluar el compromiso que cada nivel educativo tiene con la realización de las evaluaciones [6].

En este contexto, los estudiantes, como principales beneficiarios de la enseñanza, desempeñan un papel crucial al ofrecer retroalimentación sobre la calidad educativa. De esta forma, este grupo no solo participa activamente en su formación, sino que también contribuye a la mejora continua de la calidad docente en la universidad.

Este estudio se centra en las evaluaciones de desempeño docente, resaltando su importancia y los factores que influyen en sus resultados. Además, busca mejorar la comprensión del proceso de evaluación y se enfoca en los resultados obtenidos a lo largo de los años. Esto es importante porque, como se menciona en el artículo [7], la profesión pedagógica implica un proceso continuo de adaptación y renovación constante. Asimismo, el docente continúa formándose cada día a través de la experiencia e interacción con sus colegas y estudiantes. Por ello la evaluación docente provee información pertinente para su desarrollo [8].

La evaluación de desempeño docente permite identificar las necesidades de mejora en las habilidades y competencias de los profesores, proporcionando oportunidades de revisión continua [7] y, de este modo, mejorar la calidad del proceso de enseñanza-aprendizaje. Además, mediante la retroalimentación, se proporciona una forma de mantener y mejorar la calidad del trabajo docente, garantizando un avance hacia la excelencia educativa [9]. Esto proporciona razones para explorar, profundizar y analizar los resultados de dicha herramienta.

Se tiene como objetivo fundamental identificar los factores clave que influyen en las evaluaciones del desempeño docente y destacar los aspectos más relevantes de esta herramienta, ofreciendo una retrospectiva sobre su aplicación. Además, se busca analizar la objetividad de los estudiantes en las evaluaciones, debido a ciertas inquietudes planteadas. Por ejemplo, se ha planteado el siguiente cuestionamiento: ¿Tienen los estudiantes criterios relevantes al evaluar el trabajo realizado por los docentes, o influye más la empatía que sienten por el profesor o los resultados de sus calificaciones? [7].

Por otro lado, se argumenta que la visión de los estudiantes es parcial y subjetiva, dado que no son expertos en la disciplina [10]. No obstante, se ha reconocido que los estudiantes son una fuente crucial de información sobre el proceso de enseñanza-aprendizaje y el logro de los objetivos académicos por parte del docente [11]. Estos argumentos subrayan la importancia de evaluar la objetividad de los estudiantes al momento de evaluar a los docentes, lo cual es clave en el proceso de evaluación.

Mediante el análisis de datos, se pretende evidenciar si los cambios hechos al instrumento de evaluación a lo largo del tiempo han influido en los resultados obtenidos, proporcionando así una perspectiva sobre las tendencias educativas de calificación del proceso enseñanza-aprendizaje.

Las herramientas informáticas seleccionadas para este proyecto son esenciales debido a su capacidad para manejar grandes cantidades de datos de forma eficiente. Dichas herramientas no solo facilitan la gestión y automatización del procesamiento de datos, sino que también permiten la identificación de patrones. Por otra parte, su uso es fundamental en la interpretación de los comentarios de los estudiantes, especialmente en el análisis de sentimientos, donde permiten incrementar la precisión y reducir el margen de error en la clasificación de textos según su polaridad.

Con estos procesos, se busca ofrecer datos relevantes sobre la evaluación docente y su relación con el rendimiento estudiantil, aportando evidencias que puedan guiar la mejora continua de las prácticas educativas.

### 3.1. Objetivo general

Determinar por medio de datos, los factores que influyen en la evaluación docente de los catedráticos en la Universidad del Valle de Guatemala.

### 3.2. Objetivos específicos

- Determinar la relación que existe entre evaluación del catedrático y el rendimiento de la sección.
- Descubrir patrones relevantes en los datos mediante el uso de análisis exploratorio y algoritmos de aprendizaje de máquinas no supervisados
- Realizar un análisis de sentimientos en los comentarios de las evaluaciones.

El interés por fortalecer la práctica docente ha sido un tema relevante en las instituciones académicas, con el objetivo de mejorar la calidad educativa. No obstante, en Guatemala, no es común evaluar sistemáticamente los resultados de las evaluaciones docentes. A continuación, se presentan dos estudios con enfoques contrastantes sobre esta temática.

Según Urrutia [12], se realizó un estudio con el objetivo de analizar la evaluación del desempeño de los docentes de la carrera de pedagogía del Centro Universitario de Oriente (CUNORI) en sus dos sedes ubicadas en los departamentos de Zacapa y Chiquimula, pertenecientes a la Universidad de San Carlos de Guatemala, donde se implementó un modelo de evaluación de 360 grados para valorar las competencias profesionales de los docentes de la carrera de Pedagogía. Este modelo incluyó la participación de estudiantes, coordinadores, compañeros docentes (evaluación por pares) y autoevaluación, permitiendo una apreciación integral del desempeño.

El estudio concluyó que, si bien la mayoría de docentes fueron calificados como “competentes” o “destacados”, también evidenció factores críticos que influyen en la percepción del desempeño docente, como la edad del profesorado, la falta de actualización tecnológica y la ausencia de procesos efectivos de retroalimentación. Además, se agregó la necesidad de implementar programas de desarrollo profesional, señalando que los docentes con posgrado o con mayor formación específica tienden a obtener mejores evaluaciones.

Así también, Chen [13] realizó un estudio en la Universidad Agrícola del Sur de China, analizando datos de evaluación docente desde 2016 hasta 2021, con el objetivo de determinar si las evaluaciones realizadas por los estudiantes del desempeño docente conducen a una mejora real en la calidad de la enseñanza.

Sus hallazgos muestran que las mejoras o deterioros en las evaluaciones no necesariamente reflejan cambios en la calidad del docente, sino que pueden explicarse por un efecto de reversión, es decir, un comportamiento en los resultados en el que los docentes, que reciben evaluaciones muy bajas (o muy altas), tienden a obtener resultados contrarios en la siguiente evaluación. Este patrón se destacó dentro del grupo de docentes jóvenes, mujeres o con menor

rango académico (profesores auxiliares o asociados).

Estos estudios resultan pertinentes ya que demuestran cómo variables externas al desempeño pedagógico pueden influir en las evaluaciones de desempeño docente. A su vez, permiten dar a conocer cómo se impactan distintos factores influyentes dependiendo de la cultura.

## 5.1. Evaluación de desempeño docente

La evaluación del desempeño docente es un proceso fundamental en el ámbito educativo, ya que representa una oportunidad para el cambio positivo y fomenta la innovación en las metodologías de enseñanza-aprendizaje a partir de las percepciones de los estudiantes. Este proceso permite evaluar el rendimiento a lo largo del semestre, brindando a los docentes la oportunidad de reflexionar sobre su práctica profesional, identificar fortalezas y debilidades, y favorece a su mejora pedagógica.

Además, abarca dos dimensiones esenciales de la labor docente: el saber y el hacer. El saber se refiere al dominio de los conceptos que el docente enseña, mientras que el hacer engloba su compromiso con la educación, su metodología de enseñanza y su relación con los estudiantes. Comprender estos aspectos no solo fortalece la práctica docente, sino que también contribuye a su crecimiento profesional [14].

### 5.1.1. Validez del instrumento de evaluación

A lo largo del tiempo, el cuestionario ha sido utilizado como una herramienta para recopilar las respuestas de los estudiantes y evaluar la práctica docente. Como se determina en [15], el cuestionario puede contribuir a la autoevaluación del desempeño docente e incentivar la reflexión sobre la práctica profesional. Esta herramienta promueve cambios orientados a fortalecer la ética en la enseñanza, y se convierte en una herramienta válida al mejorar el proceso de aprendizaje y facilitar el logro de los objetivos educativos [16].

## 5.1.2. Modelos de evaluación docente

### 5.1.2.1. Modelo basado en la opinión de los alumnos

Este modelo de evaluación permite recopilar las respuestas de los estudiantes, basadas en su experiencia y percepción. Generalmente, las respuestas se recopilan a través de un cuestionario con preguntas de opción múltiple tipo Likert y puede complementarse con un apartado de comentarios.

En estos comentarios, los estudiantes pueden expresar desde quejas hasta sugerencias de mejora, lo que también permite evaluar su nivel de interés según la cantidad. Los comentarios suelen pertenecer a tres categorías principalmente, tal como se establece a continuación:

1. Opiniones o inconformidades que pueden sugerir modificaciones en la metodología de enseñanza del docente.
2. Observaciones sobre la actitud del docente y su compromiso con su función como docente.
3. Comentarios que no tienen relación con la retroalimentación docente [6].

### 5.1.2.2. Modelo de evaluación a través de pares

Este modelo se basa en la evaluación entre colegas docentes, enfocándose principalmente en las competencias académicas. Las reflexiones y el diálogo resultante permiten un beneficio mutuo y una mejora en sus capacidades como profesores [17]. Este tipo de modelo ofrece una perspectiva distinta desde el ámbito profesional, ya que los colegas docentes pueden comprender mejor la metodología empleada, generando sugerencias más allegadas a lo técnico [18]. Se divide en tres distintos tipos de observación entre pares:

- **Modelo de evaluación:** consiste en que docentes con mayor experiencia identifican áreas de mejora y potencial de desarrollo en la cátedra de otros docentes.
- **Modelo de desarrollo:** consiste en que docentes fomentan la mejora de competencias de profesores principiantes.
- **Modelo colaborativo:** consiste en que colegas catedráticos observan la práctica docente de otros profesores y los evalúan con base en distintas métricas. Además, realizan un discernimiento sobre su trabajo y generan sugerencias [17].

### **5.1.2.3. Modelo de autoevaluación**

Este modelo brinda un espacio de reflexión y autocrítica sobre el desempeño como educador, permitiendo evaluar su eficacia y planificar mejoras en la práctica docente. Para alcanzar los objetivos esperados, la evaluación se realiza a través de materiales diseñados para analizar la enseñanza y, en algunos casos, con el acompañamiento de un asesor [19].

### **5.1.2.4. Modelo de evaluación a través de portafolio**

Este modelo se basa en la recopilación de los trabajos más relevantes o significativos del docente, los cuales evidencian su desempeño y demuestran tanto su conocimiento como las competencias alcanzadas. Esto permite analizar su dominio profesional y, con base en ello, recibir críticas constructivas para mejorar su proceso formativo [20].

### **5.1.3. Criterios de evaluación**

Los criterios de evaluación permiten definir qué constituye un buen desempeño docente, de acuerdo con los estándares o competencias establecidos por la institución académica. De esta manera, se establece un marco para la enseñanza que fomenta la mejora de la profesión docente y profesional, siempre considerando el contexto y las condiciones en las que se lleva a cabo la labor educativa.

Los marcos para la enseñanza definen las competencias y responsabilidades que un docente debe poseer para ejercer su labor tanto en el aula como en la comunidad [21].

Estos marcos sirven como una guía para la evaluación continua y formación de los docentes, constituyen una base para la toma de decisiones y contribuyen a garantizar la excelencia institucional.

### **5.1.4. Factores contextuales**

Si bien el rendimiento de un docente no depende exclusivamente del contexto, los recursos o la infraestructura, estos pueden potenciarlo o dificultarlo. Por tanto, es necesario entender el impacto que estos puedan llegar a tener.

Según [22], la docencia se contextualiza en cuatro niveles: el salón de clases, el contexto institucional y dos contextos sociales: el inmediato y el mediato.

#### **Salón de clases**

Es el entorno más directo para el docente y juega un papel clave en su desempeño. Dos factores principales influyen en su labor: el alumnado y los recursos disponibles. La cantidad de estudiantes y sus características —como capacidades, trayectoria académica, necesidades educativas, expectativas y contexto socioeconómico— impactan la enseñanza. Asimismo, la

infraestructura y los materiales didácticos, tanto tradicionales como tecnológicos, condicionan las estrategias pedagógicas y la dinámica del aula [22].

### **Contexto institucional**

Este nivel se divide en dos dimensiones: el instituto académico y el sistema educativo. Tal como se menciona en Meliá et al. [22, p. 189], "la evaluación de las escuelas o instituciones educativas debe constituir el marco en el que se inserte la evaluación del profesorado". La calidad de la enseñanza no depende solo del profesorado, sino también de las condiciones, normas y estructuras en las que trabajan los docentes.

### **Contexto social inmediato**

El contexto social inmediato comprende los rasgos demográficos del área en la que se encuentra el instituto académico. Esta categoría incluye indicadores como el nivel cultural y socioeconómico de la comunidad, dentro del cual la estructura económica juega un papel clave al influir en las expectativas, las elecciones de la comunidad y el comportamiento del profesorado.

Otro aspecto relevante son las características sociales de la población atendida por el instituto académico, ya que determinan sus aspiraciones y necesidades. Asimismo, la convivencia es un factor esencial, pues influye en la estructura social, los tipos de interacción y los niveles de violencia o cooperación dentro de la comunidad escolar [22].

### **Contexto social mediato**

Para este nivel, se pueden considerar los mismos indicadores considerados en el nivel inmediato, pero a una escala mayor, refiriéndose a las condiciones y políticas de un país [23], ya que los países poseen particularidades culturales que influyen en las aspiraciones y expectativas educativas, así como en el comportamiento docente [22].

#### **5.1.5. Aspectos didácticos**

La didáctica ha evolucionado de modelos tradicionales basados en la transmisión unidireccional del conocimiento hacia enfoques más participativos y centrados en el estudiante. Este cambio ha sido impulsado por el reconocimiento de la diversidad en los estilos de aprendizaje, la incorporación de tecnologías, y la necesidad de responder a contextos sociales cambiantes. Más allá de su dimensión técnica, la didáctica también implica una reflexión crítica sobre qué se enseña, cómo se enseña y con qué finalidad [24].

Los enfoques didácticos son una parte importante en el proceso educativo. Entre los más destacados se encuentran:

- **Enfoque tradicional:** da al docente un rol central como transmisor del conocimiento, mientras que los estudiantes son receptores pasivos. Los contenidos se organizan según la lógica de las disciplinas. La enseñanza se basa en clases magistrales y en la memorización de contenidos. La escuela, en este enfoque, busca conservar y transmitir los valores y saberes establecidos por la sociedad.
- **Enfoque tecnicista:** busca que la enseñanza sea una actividad sistemática y científica, adaptándose a las necesidades del mercado laboral y el ámbito económico. Se enfoca en tener objetivos claros y bien definidos, considerando que los contenidos y las estrategias se eligen en función de esos objetivos. Además, el docente no diseña el plan educativo, sino que solo se encarga de aplicarlo.
- **Enfoque activo:** centra el proceso educativo en el estudiante, considerando sus intereses y necesidades, fomentando su autonomía. El docente actúa como guía y más que memorizar conceptos, se busca que el estudiante adquiera herramientas para aprender por sí mismo.
- **Enfoque problematizador:** tanto el docente como los estudiantes participan en el proceso educativo. Los contenidos se vinculan con los intereses del alumnado, y la planificación se concibe como una guía flexible, abierta a modificaciones durante el desarrollo de la enseñanza. Desde esta perspectiva, la escuela tiene la misión de formar personas críticas, autónomas y comprometidas, con la capacidad de intervenir y transformar su entorno social [25].

#### 5.1.6. Actores en el proceso de evaluación

En el proceso de evaluación docente, existen tres roles fundamentales: el del estudiante, el del docente y el de los directivos. Los estudiantes, como receptores directos de la labor docente, pueden formar un criterio basado en su experiencia académica y la interacción con el docente. En el caso de los docentes, al recibir la retroalimentación de los estudiantes, pueden identificar qué prácticas son efectivas y cuáles no, y, con base en ello, realizar los cambios necesarios para mejorar su enseñanza e influir positivamente en el aprendizaje de los estudiantes. Finalmente, para los directivos, la recopilación de los resultados es importante, ya que les permite obtener una visión general del desempeño docente e identificar aciertos y áreas de mejora a lo largo del semestre [26].

## 5.2. Análisis exploratorio

El análisis exploratorio de datos (AED) es un proceso que permite comprender en mayor profundidad los datos disponibles. En esta fase, es posible examinar supuestos, comprobar hipótesis, identificar patrones y detectar irregularidades o valores atípicos. Este proceso se utiliza precisamente para entender las variables involucradas y cómo se relacionan entre sí.

Existen cuatro tipos principales de análisis exploratorio:

- **Análisis univariado no gráfico:** se evalúa únicamente una variable, con el propósito de describir sus datos y encontrar patrones relevantes.
- **Análisis univariado gráfico:** se utiliza una sola variable, pero apoyándose en gráficos para representar su distribución. Algunos gráficos comunes son:
  - **Histograma:** muestra la frecuencia o proporción de casos dentro de rangos de valores específicos.
  - **Diagrama de caja y bigotes:** representa estadísticas descriptivas clave de la variable, como valor mínimo, máximo, primer cuartil, mediana y tercer cuartil.
- **Análisis multivariado no gráfico:** evalúa dos o más variables usando técnicas como tabulación cruzada o análisis estadísticos.
- **Análisis multivariado gráfico:** representa visualmente dos o más variables para mostrar cómo se relacionan entre sí. Algunos gráficos multivariados comunes son:
  - **Gráfico de dispersión:** muestra la relación entre dos variables numéricas a través de puntos ubicados en los ejes X y Y.
  - **Gráfico de burbujas:** similar al gráfico de dispersión, pero incorpora una tercera variable representada por el tamaño o color de los puntos.
  - **Mapa de calor:** representación gráfica donde los datos son mostrados mediante una escala de colores, permitiendo identificar patrones y concentraciones [27].

## 5.3. Procesamiento de lenguaje natural

El procesamiento del lenguaje natural (PLN) es un campo interdisciplinario que abarca técnicas computacionales para analizar, interpretar y generar lenguaje humano, tanto hablado como escrito. Su desarrollo ha transitado desde enfoques basados en reglas y traducción automática, hacia el uso de modelos estadísticos, y más recientemente hacia arquitecturas de aprendizaje profundo, incluyendo modelos de lenguaje de gran escala (LLMs) que permiten una comprensión más contextual y flexible del lenguaje natural [28].

### 5.3.1. Análisis de sentimientos

El análisis de sentimientos es el proceso de clasificar un texto como positivo, negativo o neutro, basándose en su polaridad total. Esta polaridad se calcula sumando los valores

de polaridad asignados a cada palabra del texto, los cuales provienen de un lexicón, que es una base de datos que asocia palabras con valores específicos de polaridad y con su correspondiente "*part of speech*" (*POS*) [29]. El *POS* permite identificar si una palabra funciona como sustantivo, verbo, adjetivo, adverbio, entre otros, dentro de la oración [30].

Un texto se clasifica como positivo si su polaridad total es mayor que 0, negativo si es menor que 0, y neutro si es igual a 0 [29].

### 5.3.2. Estandarización de datos: lematización

La lematización es un proceso del procesamiento de lenguaje natural (PLN) que transforma palabras en sus diversas formas morfológicas a su forma base, denominada lema [31].

## 5.4. Aprendizaje de máquinas

Se basa en la aplicación de métodos estadísticos y técnicas de optimización para poder analizar conjuntos de datos, identificar patrones y construir modelos predictivos. Dependiendo del tipo de problema y la disponibilidad de datos, el aprendizaje automático puede clasificarse en cuatro categorías [32]:

- **Aprendizaje supervisado:** utiliza un conjunto de datos etiquetados para aprender las relaciones entre las entradas y las salidas, permitiendo que el modelo pueda realizar predicciones sobre nuevos datos. Este tipo de aprendizaje permite resolver tareas de clasificación (identificar la clase a la que pertenece cada dato) y de regresión (predecir valores continuos) [33].
- **Aprendizaje no supervisado:** analiza conjuntos de datos no etiquetados para descubrir patrones, agrupamientos o asociaciones entre variables [34].
- **Aprendizaje semisupervisado:** combina un pequeño conjunto de datos etiquetados con un conjunto más amplio no etiquetado. El modelo utiliza el conocimiento aprendido de los datos etiquetados para generalizar y extraer información del resto del conjunto [35].
- **Aprendizaje por refuerzo:** en este tipo de aprendizaje se aprende mediante prueba y error. Se va aprendiendo de los resultados exitosos y se ajusta progresivamente una política (conjunto de reglas) para tomar las mejores decisiones [35].

### 5.4.1. Árboles de decisión

Los árboles de decisión son un algoritmo supervisado que puede adaptarse a la complejidad de los datos, ya que no dependen de un número fijo de parámetros, lo que los clasifica como modelos no paramétricos. Este tipo de algoritmo se utiliza tanto para tareas de clasificación como de regresión, dividiendo los datos en función de ciertas condiciones. Además, su estructura jerárquica facilita la interpretación y permite comprender de forma transparente el proceso de toma de decisiones del modelo [36].

### 5.4.2. Reglas de asociación

Las reglas de asociación son un algoritmo no supervisado, en el cual se busca identificar patrones relevantes, relaciones o posibles vínculos causales entre conjuntos de elementos dentro de una base de datos o repositorio de información [37].

Este enfoque se ha popularizado especialmente en el análisis de la cesta de mercado, donde el objetivo es descubrir patrones del tipo “Si  $A$ , entonces  $B$ ”, formalizados como reglas  $A \Rightarrow B$ , donde  $A$  es el antecedente y  $B$  el consecuente [38].

Para evaluar la calidad de una regla de asociación, se emplean varias métricas estadísticas, entre las cuales destacan:

- **Soporte (*Support*):** mide la proporción de transacciones que contienen simultáneamente  $A$  y  $B$ . Se define como:

$$\text{support}(A \Rightarrow B) = P(A \cap B)$$

- **Confianza (*Confidence*):** representa la probabilidad de que  $B$  ocurra dado que  $A$  ha ocurrido. Se calcula como:

$$\text{confidence}(A \Rightarrow B) = \frac{P(A \cap B)}{P(A)}$$

- **Elevación (*Lift*):** indica el grado de dependencia entre  $A$  y  $B$ . Un valor mayor a 1 sugiere una correlación positiva:

$$\text{lift}(A \Rightarrow B) = \frac{\text{confidence}(A \Rightarrow B)}{P(B)}$$

- **Convicción (*Conviction*):** mide cuán frecuentemente  $A$  ocurre sin que ocurra  $B$ , en comparación con la expectativa si fueran independientes:

$$\text{conviction}(A \Rightarrow B) = \frac{1 - P(B)}{1 - \text{confidence}(A \Rightarrow B)}$$

- **Apalancamiento (*Leverage*):** evalúa la diferencia entre la frecuencia observada de  $A$  y  $B$  juntos, y la frecuencia esperada si fueran independientes:

$$\text{leverage}(A \Rightarrow B) = P(A \cap B) - P(A) \cdot P(B)$$

Estas métricas permiten filtrar y priorizar las reglas más interesantes, útiles y no triviales en los análisis [38].

## 5.5. Análisis de grupos

Es una técnica de aprendizaje no supervisado que busca agrupar objetos similares en subconjuntos denominados clústeres, de manera que los objetos dentro de un mismo grupo sean más similares entre sí que con los de otros grupos [38].

### 5.5.1. Tipos de clústeres

- **Clústeres bien separados:** cada punto de un clúster está más cercano a cualquier otro punto del mismo clúster que a cualquier punto de otro clúster.
- **Clústeres basados en centroide:** los objetos se agrupan alrededor de un centroide, como en el algoritmo K-means.
- **Clústeres contiguos:** se forman mediante la conectividad de los objetos, utilizando métodos como el de vecinos más cercanos.
- **Clústeres basados en densidad:** se identifican regiones densas separadas por regiones de baja densidad, como en el algoritmo DBSCAN [38].

### 5.5.2. Métodos de agrupamiento

- **Particionales:** dividen el conjunto de datos en un número predefinido de clústeres no superpuestos.
- **Jerárquicos:** generan una estructura en forma de árbol (dendrograma) que representa anidamientos de grupos. Pueden ser aglomerativos o divisivos.
- **Basados en densidad:** detectan grupos de forma arbitraria y manejan bien el ruido [38].

### 5.5.3. Medidas de similitud

La elección de la medida de similitud o disimilitud es crucial en el análisis de clústeres. Algunas de las más comunes incluyen [38]:

- **Distancia euclidiana:** adecuada para variables cuantitativas.
- **Distancia de Manhattan:** suma de las diferencias absolutas entre las coordenadas.
- **Coefficiente de Jaccard:** utilizado para variables binarias asimétricas.

### 5.5.4. Evaluación de la calidad del agrupamiento

La calidad de un agrupamiento puede evaluarse mediante [38]:

- **Medidas internas:** evalúan la cohesión y separación de los clústeres sin referencia externa.
- **Medidas externas:** comparan el resultado del agrupamiento con una estructura conocida previamente.

### 5.5.5. Aplicaciones

El análisis de clústeres tiene aplicaciones en diversas áreas, tales como [38]:

- **Segmentación de mercado:** identificación de grupos de clientes con comportamientos similares.
- **Biología:** clasificación de especies o genes con características similares.
- **Reconocimiento de patrones:** agrupamiento de datos para identificar patrones ocultos.

## **6.1. Recolección de datos**

Para realizar un análisis de sentimientos de los comentarios de evaluación de los estudiantes, así como para estudiar la correlación entre el rendimiento académico de los estudiantes y las calificaciones recibidas por los catedráticos en las evaluaciones de desempeño docente, es necesario recolectar datos específicos. Estos se obtuvieron de las evaluaciones finales de los catedráticos de la facultad de ingeniería en el período de 2017 a 2023, usando el portal de servicios electrónicos de la Universidad del Valle de Guatemala.

## **6.2. Limpieza y preprocesamiento de datos**

Para poder hacer un adecuado análisis de datos se llevó a cabo una limpieza previa en el texto, para poder utilizarlo como entrada del modelo de análisis de sentimientos. Para las tareas de limpieza se eliminaron caracteres especiales, apóstrofes, artículos, preposiciones, conjunciones y signos de puntuación. Por otra parte, en la etapa de preprocesamiento, se convirtió el texto de mayúsculas a minúsculas para estandarizar el formato. La nota de evaluación de cada catedrático se calculó sumando la proporción de respuestas en las categorías 3 y 4 (3 y 4 estrellas) y multiplicando por 100. Esto permitió hacer análisis que involucran el rendimiento del docente en el curso que imparten.

## **6.3. Privacidad de los datos**

El conjunto de datos se obtuvo con el identificador del docente (un número entero de hasta 5 cifras) de manera que se mantiene la privacidad del docente. Además se borraron los nombres de los docentes mencionados en los comentarios, ya que constituyen datos sensibles.

## **6.4. Análisis de sentimientos**

Se realizó un etiquetado de los textos de los comentarios de evaluación de los estudiantes utilizando el recurso léxico ML-Senticon en español [29]. Con este proceso se determinará la proporción de mensajes neutros, positivos y negativos.

## **6.5. Análisis exploratorio**

Se llevó a cabo un análisis exploratorio para encontrar datos relevantes, patrones, variabilidad y posibles sesgos en los datos de las evaluaciones de desempeño docente. Este análisis incluyó procedimientos de estadística descriptiva y se enfocó en analizar los criterios con mayor influencia en la evaluación de los catedráticos. Para poder estudiar la influencia de los posibles factores en la nota de evaluación, se emplearon varias técnicas entre las que se incluyeron el cálculo de coeficientes de correlación entre las variables numéricas, análisis de criterios más influyentes, entre otras. Además, se evaluó la distribución de los datos por medio de medidas de tendencia central y dispersión.

## **6.6. Algoritmos de aprendizaje automático no supervisados**

Se aplicaron 2 técnicas de aprendizaje automático no supervisado para descubrir patrones ocultos en los datos que permitieran determinar los factores que influyen en los resultados de la evaluación docente del catedrático. Se obtuvieron reglas de asociación y algoritmos de agrupamiento.

Para poder obtener los siguientes resultados, se utilizaron distintos conjuntos de datos, los cuales se detallan a continuación:

- **info\_cursos.csv**: este conjunto de datos incluye las siguientes variables: ciclo correspondiente al curso, el año correspondiente al curso, la cantidad de créditos del curso y el código de la materia.
- **resumen\_cursos\_2016\_2024.csv**: este conjunto de datos contiene información académica sobre las calificaciones de los estudiantes y características asociadas, incluyendo el año y ciclo del curso, estadísticas descriptivas de las notas (valor mínimo, promedio, primer cuartil, tercer cuartil, valor máximo), la cantidad de alumnos aprobados, la cantidad de alumnos inscritos y el código de la materia.  
\* Se utilizaron únicamente datos hasta 2023, ya que los demás conjuntos de datos también contenían información hasta ese año.
- **datos\_evaluacion\_prof\_ing.csv**: este conjunto de datos contiene información sobre las respuestas a las evaluaciones docentes, descritas a través de diversas variables y condiciones, incluyendo el año y ciclo en que se realizaron, el código del docente evaluado, el tipo de curso (práctico o teórico), así como datos cuantitativos sobre cada evaluación: cantidad de comentarios, cantidad de respuestas, cantidad de estudiantes que participaron y cantidad de estudiantes inscritos.
- **datos\_evaluacion\_prof\_ing\_items.csv**: este conjunto de datos contiene información sobre las evaluaciones docentes realizadas por los estudiantes. Incluye las siguientes variables: el número correspondiente al criterio evaluado, la cantidad y el porcentaje de alumnos que seleccionaron una cantidad específica de estrellas para el criterio de evaluación, la cantidad de estrellas asignadas, el código del docente correspondiente a la evaluación, el código de la materia, el ciclo y el año de la evaluación.
- **datos\_evaluacion\_prof\_ing\_comments**: este conjunto de datos contiene información sobre los comentarios que los estudiantes han dejado en las evaluaciones docentes.

Incluye las siguientes variables: año y ciclo de la evaluación, el código del docente evaluado y el texto del comentario registrado.

Por otra parte, fue necesario generar la variable de la nota que obtuvo cada docente en cada ciclo, año y curso, a partir del conjunto de datos: `datos_evaluacion_prof_ing_items.csv`. Por tanto, se siguieron los siguientes pasos para calcularla:

1. Agrupar los datos por ciclo, año, código del docente, código de la materia y cantidad de estrellas seleccionadas, para calcular la cantidad total de respuestas correspondientes a cada combinación.
2. Ajustar los valores de estrellas del año 2023 restando 1, ya que en ese año se registraron con un rango de 1 a 5 estrellas, mientras que en los años anteriores se usaba un rango de 0 a 4.
3. Se combinan los conjuntos de datos: `datos_evaluacion_prof_ing_items.csv` y `datos_evaluacion_prof_ing.csv` para obtener las variables necesarias para hacer el cálculo final de la nota docente.
4. Finalmente, para calcular la nota de evaluación, se suman los valores correspondientes a las respuestas con 3 y 4 estrellas, y el resultado se divide por la cantidad total de respuestas, y el resultado se multiplica por 100.

En algunos análisis fue necesario combinar varios conjuntos de datos para relacionar diferentes variables y así generar cálculos o analizar su comportamiento. Por lo tanto, se realizaron las siguientes combinaciones:

- `datos_evaluacion_prof_ing_items.csv` y `datos_evaluacion_prof_ing.csv` para poder calcular la nota de evaluación docente.
- Conjunto de datos que contiene la nota de evaluación e `info_cursos.csv`.
- Conjunto de datos que contiene la nota de evaluación y `resumen_cursos_2016_-2024.csv`.

## 7.1. Evaluación docente a lo largo del tiempo

Esta sección presenta la evolución de las notas de evaluación de desempeño docente del año 2017 al 2023. El estudio contempla tanto la distribución global de las notas como su comportamiento anual, con el propósito de identificar patrones de tendencia y variaciones anuales (Figura 1).

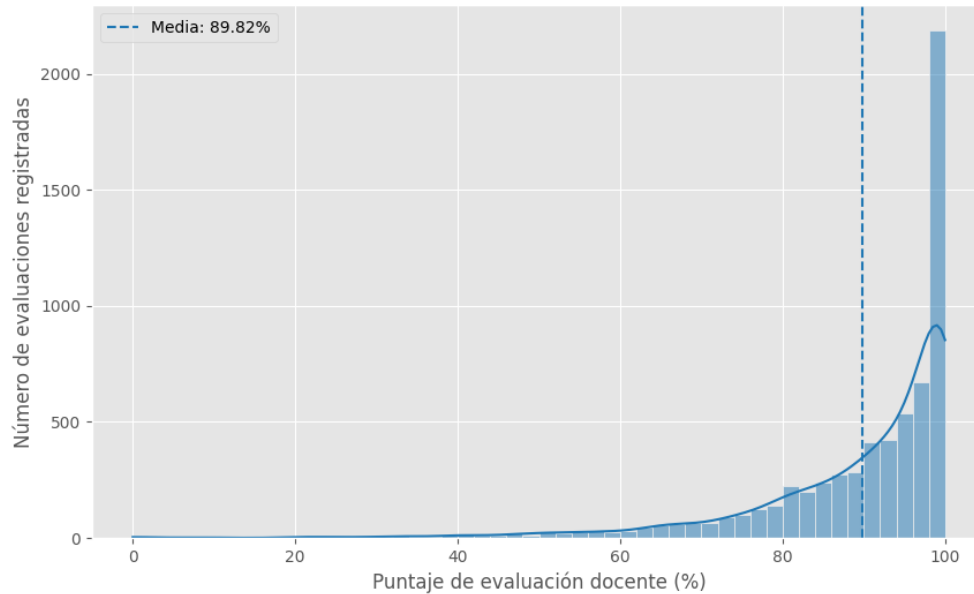


Figura 1: Distribución general de las notas de evaluación docente

La mayoría de las notas están por encima de 90, lo que evidencia una tendencia clara de evaluaciones altas y consistentes. La mediana de las evaluaciones se mantiene en niveles superiores a 90, indicando una percepción predominantemente positiva por parte de los estudiantes. No obstante, existe una minoría de evaluaciones por debajo del 80.

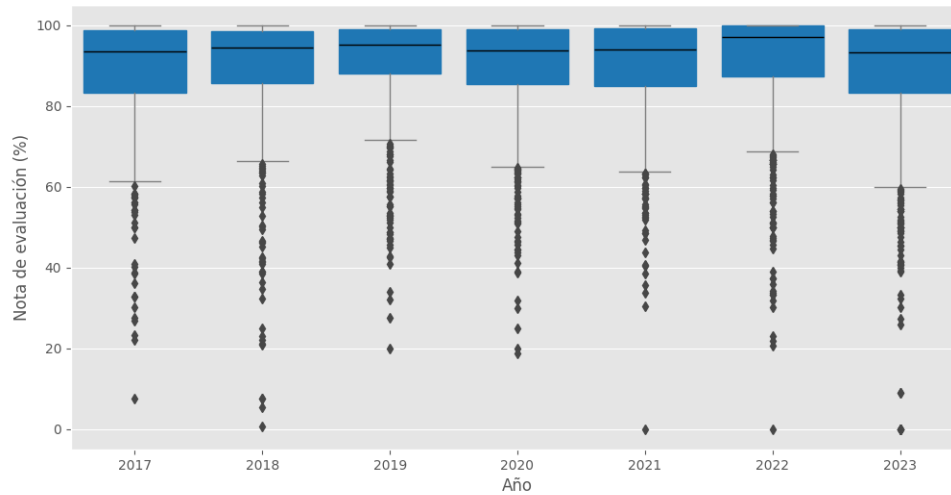


Figura 2: Distribución anual de las notas de evaluación docente

Al observar la evolución anual (Figura 2), se muestra que no existen grandes diferencias entre los años analizados. La forma de las distribuciones es consistente, indicando que el sistema de evaluación se ha mantenido estable tanto en tendencias como en dispersión. La densidad máxima permanece cercana al 100% en todos los períodos, reflejando una alta acumulación de evaluaciones sobresalientes año tras año.

### 7.1.1. Impacto de características del curso en las evaluaciones

En esta sección se muestran los resultados de aquellas características contextuales de los cursos que pueden influir en cómo se percibe el desempeño del docente. Estos elementos, aunque ajenos al control del docente, pueden condicionar la experiencia de los estudiantes y, por ende, incidir en los resultados obtenidos de las evaluaciones docentes.

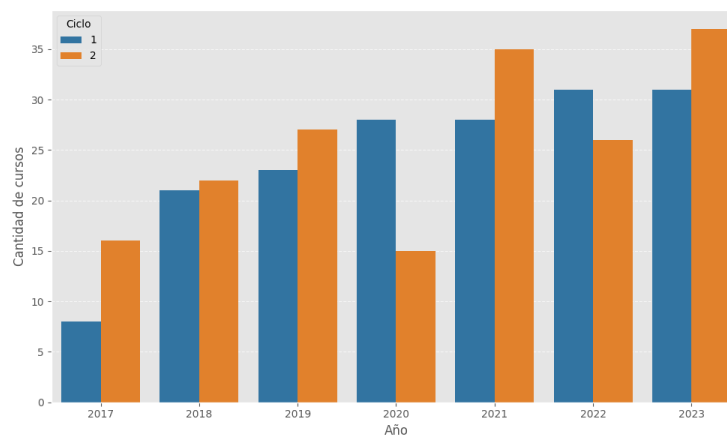


Figura 3: Cantidad de cursos con mejor evaluación en cada ciclo

En la Figura 3 se observa que, en la mayoría de los años analizados, el segundo ciclo presenta una mayor cantidad de cursos con mejores evaluaciones en comparación con el primero, destacando los años 2019, 2021 y 2023. Un caso atípico se presenta en 2020, cuando el ciclo 1 supera de forma notable al ciclo 2. Esta variación podría estar asociada a las condiciones impuestas por la pandemia, que afectaron las condiciones de enseñanza.

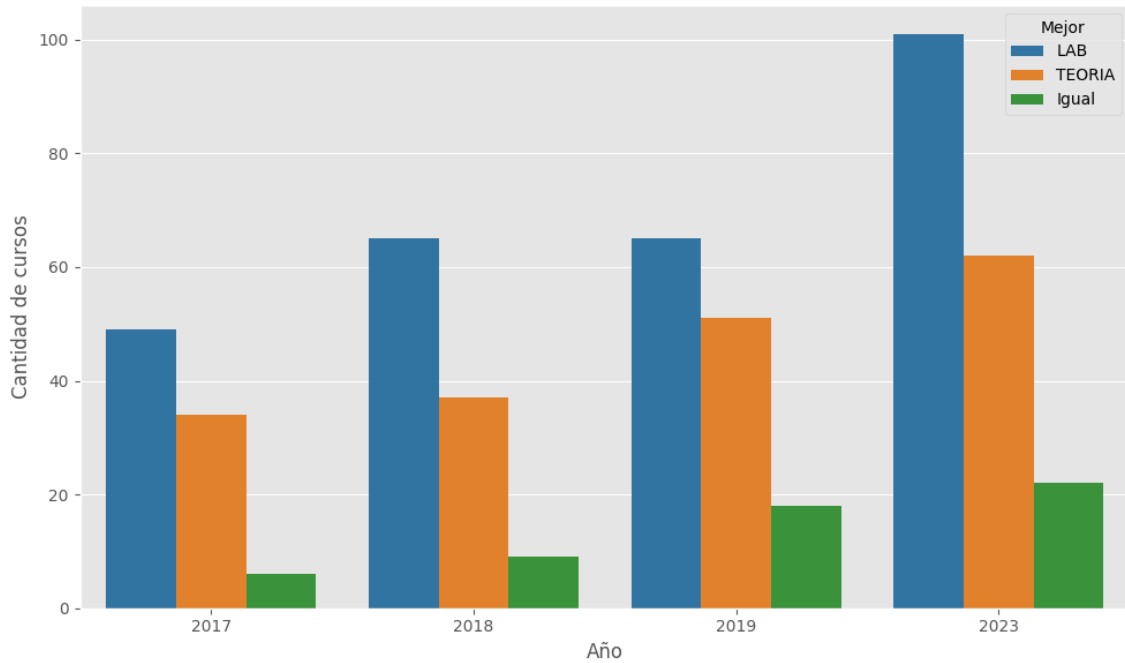


Figura 4: Comparación de cantidad de cursos donde fue mejor teoría o laboratorio por año

La Figura 4 evidencia una preferencia por el componente práctico (laboratorio) en las evaluaciones, lo que podría estar relacionado con una mayor interacción, aplicación directa de conceptos o las metodologías empleadas.

## 7.2. Participación de estudiantes en evaluaciones

En esta sección se analiza cómo la participación estudiantil incide en los resultados de las evaluaciones docentes, a partir del vínculo entre la cantidad de estudiantes evaluadores y diversas variables contextuales. El análisis busca identificar patrones de participación y su impacto en la representatividad de los resultados obtenidos.

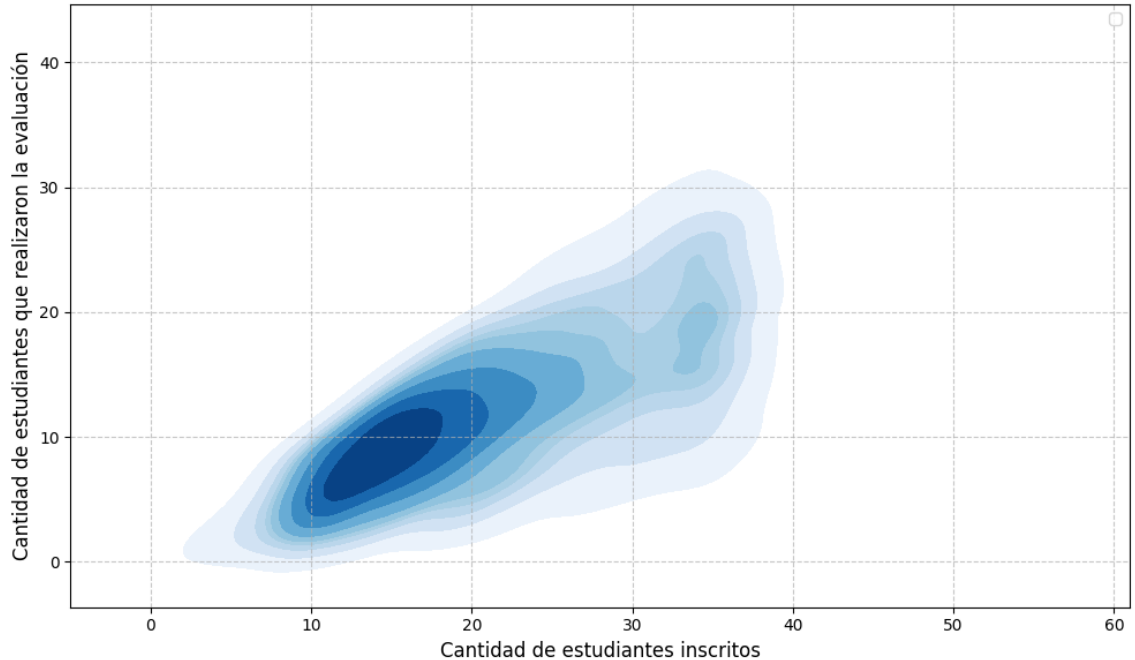


Figura 5: Densidad conjunta entre estudiantes inscritos y evaluadores

La Figura 5 muestra que, a mayor cantidad de estudiantes inscritos, mayor tiende a ser la participación en la evaluación docente, aunque la participación en las secciones es parcial. Además, se observa una concentración de datos entre 10 y 20 estudiantes inscritos, y entre 5 y 15 que participaron en la evaluación. Esto sugiere que la mayoría de las secciones analizadas no corresponden a grupos grandes y tienen una participación moderada en las evaluaciones.

### 7.2.1. Análisis de comentarios

A continuación, se presentan los resultados del análisis sobre la relación entre la cantidad de comentarios positivos y la cantidad de estudiantes que evaluaron al docente, así como el comportamiento observado en los comentarios negativos.

Para obtener estos resultados, se llevó a cabo un análisis de sentimientos en los comentarios, utilizando el lexicón en español ML-Senticon, un recurso multilingüe de polaridades semánticas a nivel de lemas, desarrollado en la Universidad de Sevilla.

Como paso inicial, se realizó un preprocesamiento de los textos, que incluyó: la eliminación de filas con comentarios vacíos, la conversión de los textos a minúsculas, la eliminación de *stopwords* y la anonimización de los nombres de los docentes mencionados en los comentarios.

Como siguiente paso, se llevó a cabo la lematización de los textos. Para cada *token* presente en los comentarios, se almacenó en un objeto el texto original del *token*, su lema y etiqueta gramatical (*part of speech*) correspondiente. La extracción de los lemas y las etiquetas gramaticales se realizó utilizando la biblioteca *spaCy*.

Luego, se extrajeron las polaridades correspondientes a cada palabra, pertenecientes a las capas de los primeros tres niveles del lexicón utilizado, junto con su respectiva etiqueta gramatical (*part of speech*) y se guardaron en un diccionario.

Seguidamente, para cada lema de cada *token* en cada comentario, se buscó su polaridad en el diccionario que contenía las polaridades extraídas del lexicón. En los casos en que una palabra no se encontraba presente, se recurrió al uso de *embeddings* del modelo *word2vec* de *gensim* para obtener las tres palabras más similares y verificar nuevamente si alguna de ellas estaba disponible en el diccionario de polaridades.

Finalmente, con base en el valor final de la polaridad de cada comentario, se clasificó el comentario como positivo si la polaridad era mayor que 0, negativo si era menor que 0, y neutro si la polaridad era igual a 0.

A partir de los comentarios clasificados, se obtuvo la siguiente distribución de comentarios por polaridad (Figura 6):

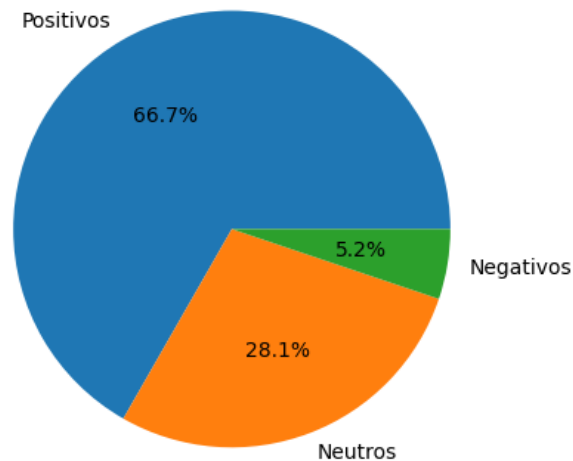


Figura 6: Distribución de polaridad de comentarios

En la Figura 7 se observa una relación positiva, que indica que a mayor cantidad de estudiantes que evalúan, mayor es la cantidad de comentarios positivos recibidos, mostrando una tendencia lineal.

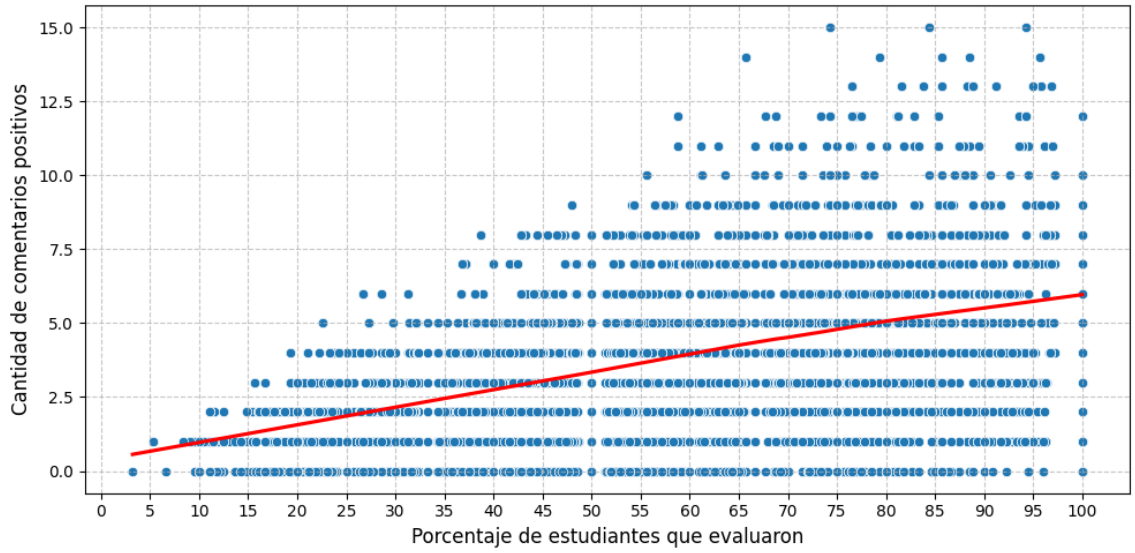


Figura 7: Relación entre porcentaje de estudiantes que evaluaron y cantidad de comentarios positivos

En comparación, la Figura 8 muestra que la cantidad total de comentarios negativos se mantiene baja y su tendencia es casi plana. Incluso en los grupos más numerosos, los comentarios negativos son escasos. Estos resultados sugieren que los estudiantes tienden a utilizar el espacio de comentarios principalmente para brindar retroalimentación positiva, mientras que las críticas negativas son poco frecuentes, independientemente del tamaño del grupo que participa en la evaluación.

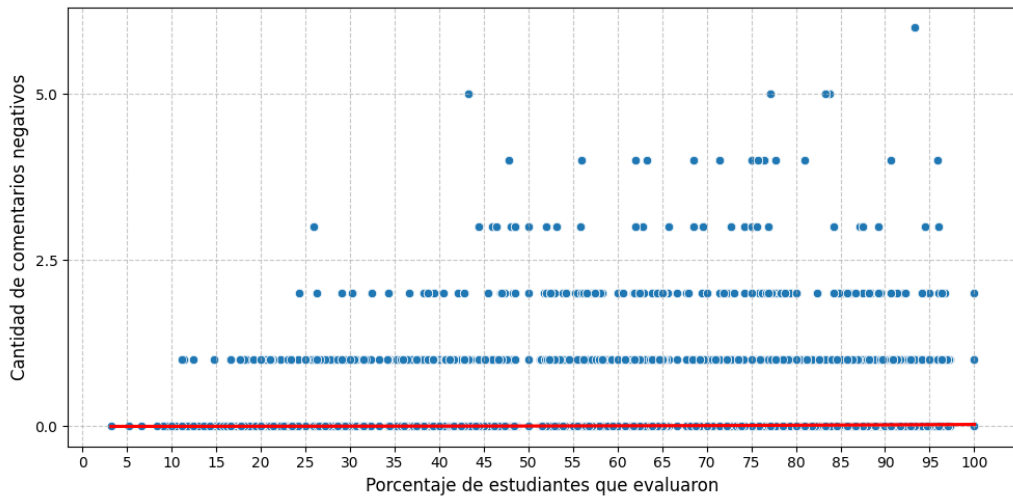


Figura 8: Relación entre porcentaje de estudiantes que evaluaron y cantidad de comentarios negativos

### 7.2.2. Cantidad de créditos

Los cursos en la Universidad del Valle tienen asociados una cantidad de créditos. El estudiante debe obtener un mínimo de créditos para completar su carrera universitaria y los va acumulando a medida que aprueba los cursos. La Figura 9, muestra una tendencia positiva en el porcentaje promedio de estudiantes que evaluaron los cursos con entre 1 y 7 créditos, destacándose especialmente los cursos de 7 créditos. Sin embargo, se observa un descenso significativo en los cursos de 10 y 12 créditos, lo que podría implicar que en cursos de mayor carga académica o dificultad hay una menor participación en las evaluaciones.

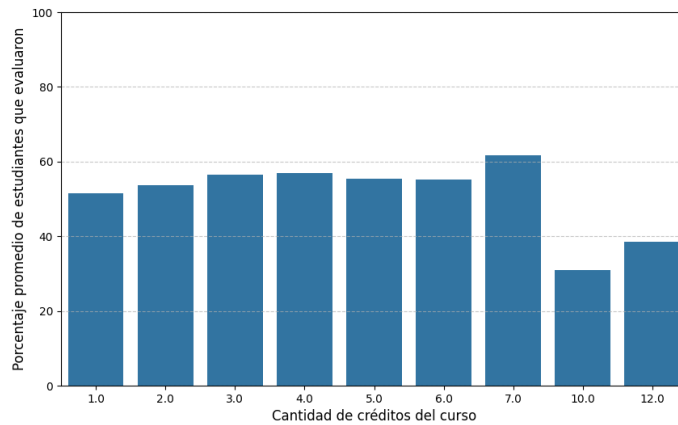


Figura 9: Porcentaje promedio de estudiantes que evaluaron por cantidad de créditos correspondientes a un curso

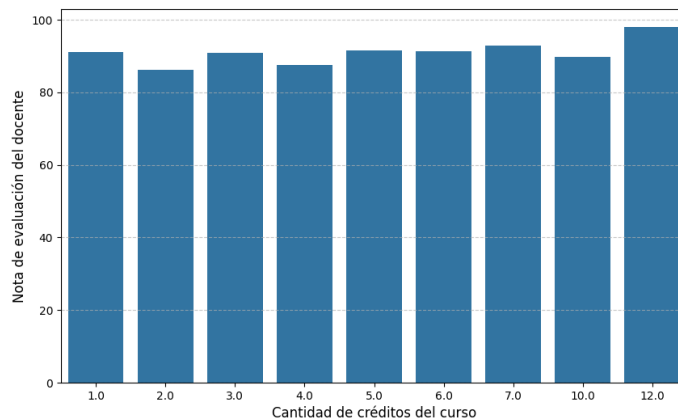


Figura 10: Promedio de la nota de evaluación del docente por cantidad de créditos del curso

Independientemente de la participación, la Figura 10 muestra una tendencia positiva: mientras más créditos se asignan a un curso, tiende a observarse una mejor nota promedio de evaluación del docente. Esto sugiere que la dificultad o carga del curso no necesariamente impacta negativamente en la calificación del docente.

### 7.2.3. Año académico correspondiente a los cursos

Se observa una tendencia ascendente en la cantidad de docentes evaluados que imparten cursos del primer al quinto año, destacando especialmente el cuarto año (Figura 11). En contraste, los cursos correspondientes al sexto año presentan la menor cantidad de docentes evaluados. Este patrón sugiere que los estudiantes que cursan asignaturas de años intermedios (especialmente tercero y cuarto) tienden a evaluar a más docentes. En el caso del sexto año, se espera una participación menor, ya que pocas carreras incluyen cursos en ese nivel.

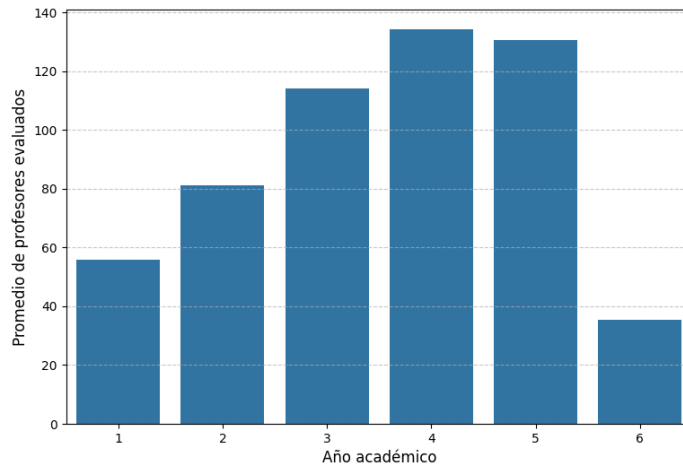


Figura 11: Cantidad promedio de profesores evaluados por año académico correspondiente a curso

La evaluación del docente se mantiene alta (Figura 12) a lo largo de los años académicos, con notas concentradas entre 90 y 100. En todos los años académicos, se muestran atípicos en notas bajas. Para cuarto y quinto año se observa mayor dispersión y tienen más atípicos, lo que puede indicar que hay más variabilidad en la percepción del docente en estos niveles.

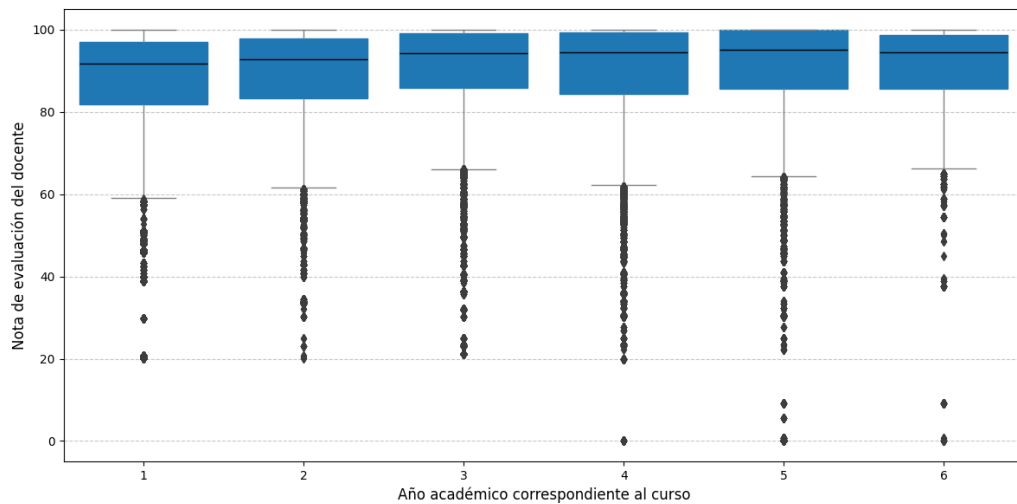


Figura 12: Distribución de notas de evaluación por el año académico que le corresponde al curso

### 7.2.3.1. Relación entre porcentaje de estudiantes que evaluaron al docente y la nota de evaluación docente

A continuación, se presenta específicamente cómo la participación de los evaluadores puede influir en los distintos rangos de notas.

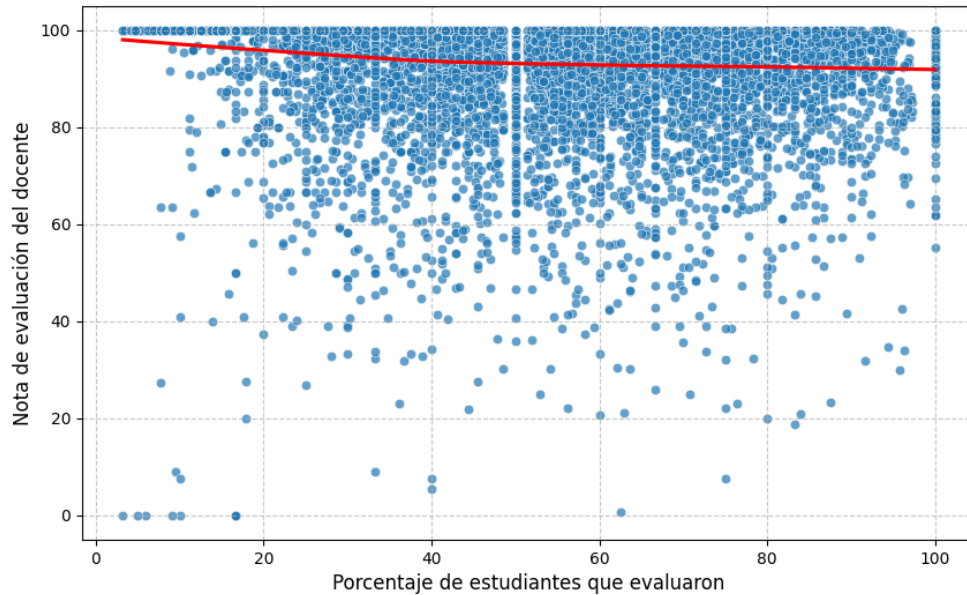


Figura 13: Relación entre el porcentaje de estudiantes que evaluaron y la nota de evaluación del docente

En la Figura 13, no se observa una relación lineal fuerte entre la nota de evaluación del docente y el porcentaje de estudiantes que participaron en la evaluación, tal como establece el coeficiente de correlación de Spearman (-0.14). Aunque se identifica una concentración de notas muy altas (superiores al 95%) en contextos con menor porcentaje de participación, esto sugiere que las mejores calificaciones no siempre reflejan la opinión general del grupo. Asimismo, una alta participación no garantiza necesariamente una evaluación favorable, ni una baja participación implica por defecto una calificación negativa.

Los resultados presentados en la Figura 14 evidencian que, para las notas menores a 80, existe una tendencia positiva; es decir, a mayor porcentaje de participación se observa un incremento moderado en la nota promedio del docente.

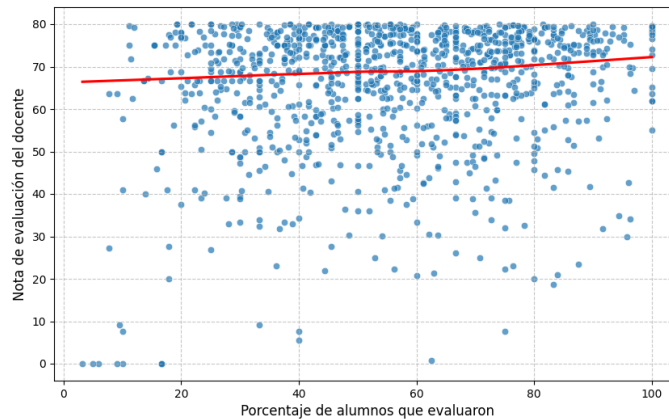


Figura 14: Relación entre el porcentaje de estudiantes que evaluaron y la nota de evaluación del docente (Notas menores o iguales a 80)

Sin embargo, es muy interesante lo que sucede con docentes que obtienen notas muy bajas (menos de 60) en los cursos (Figura 15).

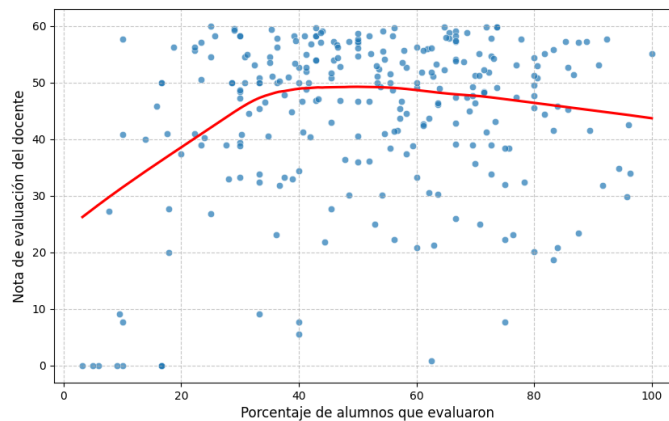


Figura 15: Relación entre el porcentaje de estudiantes que evaluaron y la nota de evaluación del docente (Notas menores o iguales a 60)

La Figura 15 muestra una tendencia ascendente en las calificaciones cuando el porcentaje de estudiantes que evaluaron se encuentra entre aproximadamente 0 % y 40 %, lo que refleja la participación inicial del grupo. La curva alcanza su punto máximo alrededor del 50 % de participación, lo que podría indicar un equilibrio óptimo entre la cantidad de evaluadores y la autenticidad de las respuestas. Sin embargo, a partir de un 60 % de participación, se observa una disminución en la nota promedio, lo cual podría interpretarse como el reflejo de una mayor participación por parte de estudiantes particularmente insatisfechos que responden de forma masiva.

### 7.2.3.2. Distribución de porcentajes de participación

En la Figura 16, se observa que en los grupos con menor participación, hay una mayor cantidad de valores atípicos en el rango de calificaciones bajas. A medida que aumenta el porcentaje de estudiantes evaluadores, la distribución de las calificaciones se vuelve más estable, disminuyendo la presencia de notas extremas. Por lo tanto, una mayor participación permite obtener resultados más representativos y consistentes.

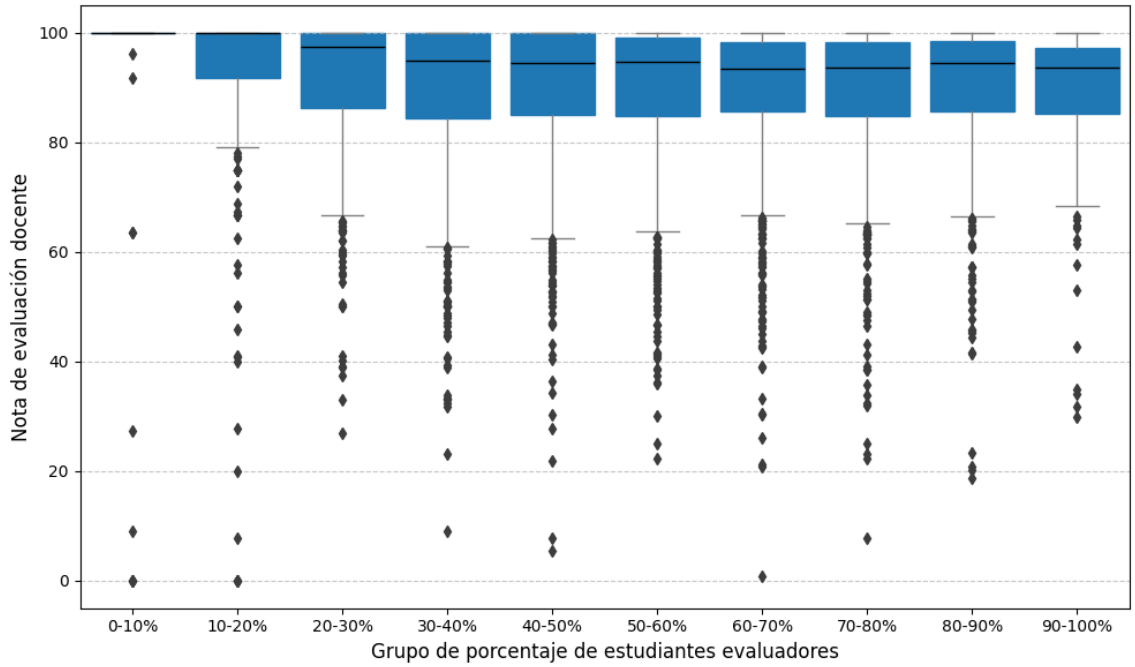


Figura 16: Distribución de evaluación docente según grupos de porcentaje de estudiantes evaluadores

### 7.3. Relación de la nota docente con el desempeño de los estudiantes

En esta sección se analiza la influencia que tienen factores como las calificaciones obtenidas por los estudiantes, así como ser aprobado o reprobado por un docente.

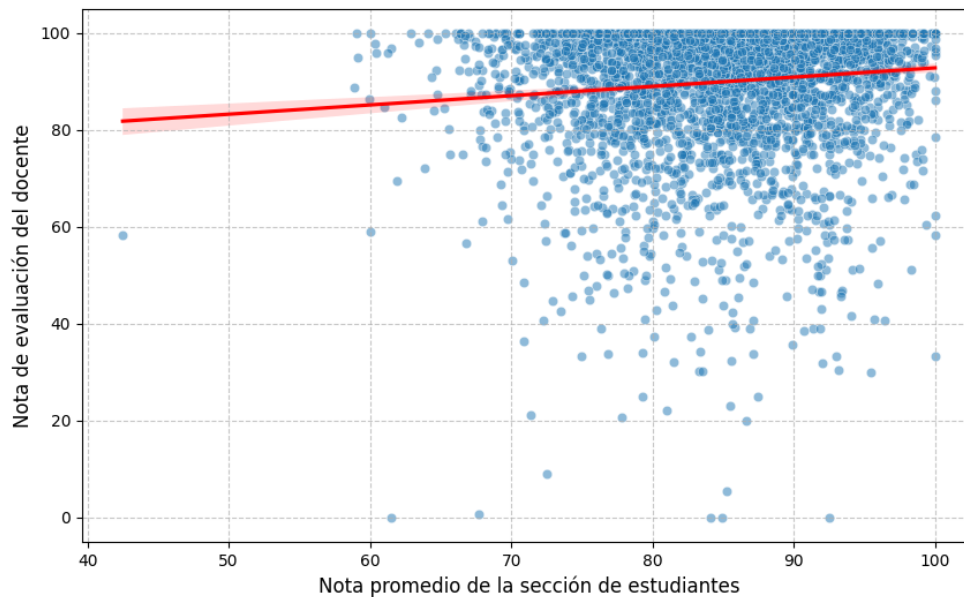


Figura 17: Relación entre la nota de evaluación del docente vs nota promedio de la sección

La línea de regresión de la Figura 17, presenta una pendiente positiva, aunque poco pronunciada, lo que indica una relación débil entre ambas variables, como también lo refleja el coeficiente de correlación de Spearman (0.14). Esto sugiere que, a medida que aumenta la nota promedio de la sección, también tiende a aumentar levemente la nota de evaluación del docente; no obstante, este efecto no influye de manera significativa en el valor final de la nota del docente.

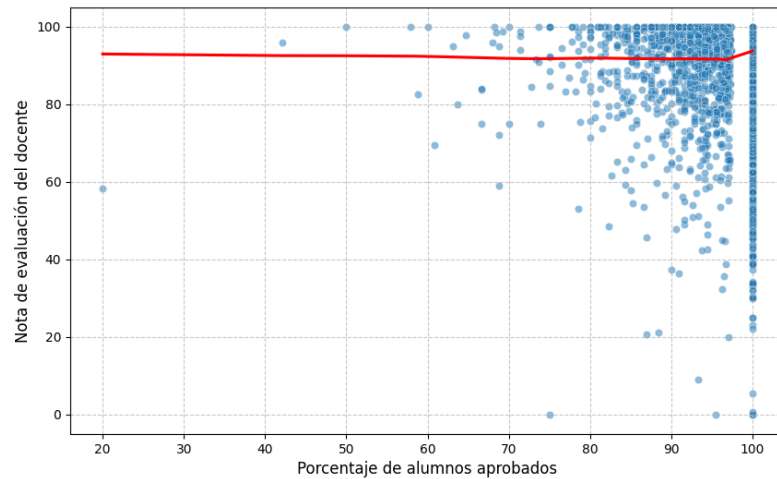


Figura 18: Relación entre nota de evaluación del docente y porcentaje de alumnos aprobados

En la Figura 18, la línea de regresión muestra una tendencia casi plana, lo que indica una relación muy débil entre la nota de evaluación del docente y el porcentaje de alumnos aprobados, confirmada por el coeficiente de correlación de Spearman (0.12). Esto implica que tener un mayor porcentaje de aprobación no garantiza necesariamente obtener una mejor evaluación docente. No obstante, aproximadamente en el rango entre el 98 % y el 100 %, se observa una pendiente positiva, aunque débil, que podría sugerir una ligera influencia favorable sobre la nota del docente.

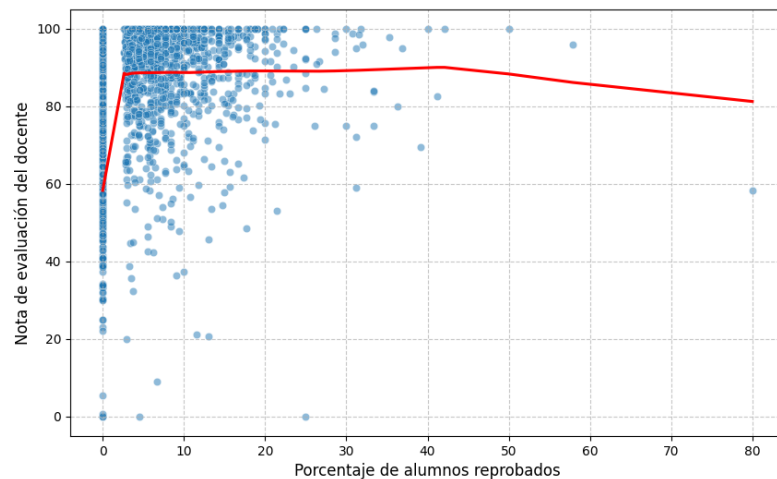


Figura 19: Relación entre nota de evaluación del docente y porcentaje de alumnos reprobados

La línea de tendencia de la Figura 19 muestra que, tras un aumento inicial, las notas de evaluación se mantienen estables alrededor de los 85–90 puntos, con un leve descenso cuando el porcentaje de reprobados supera aproximadamente el 40 %. El valor de correlación de Spearman ( $-0.12$ ) indica que existe una relación muy débil entre el porcentaje de reprobación

y la calificación asignada al docente. Por tanto, el hecho de que un docente repruebe a más o menos estudiantes no parece influir fuertemente en la nota de evaluación que se recibe por parte de los estudiantes.

## 7.4. Instrumento de evaluación de desempeño docente

Evaluación del docente	
MUESTRA DOMINIO DEL CURSO QUE IMPARTE	NUNCA ★☆☆☆☆
ES CREATIVO, FLEXIBLE E INNOVADOR EN SU ESTILO DE ENSEÑANZA	CASI NUNCA ★★☆☆☆
BRINDA AYUDA PARA QUE LOS ESTUDIANTES RESUELVAN DUDAS Y PROGRESEN	OCASIONALMENTE ★★★☆☆
PROMUEVE LA CONFIANZA, LA COMUNICACION Y EL RESPETO EN EL AULA	FRECUENTEMENTE ★★★★☆
ACOMPaña A LOS ESTUDIANTES PARA QUE ALCANCEN LOS APRENDIZAJES PROPUESTOS EN EL PROGRAMA DEL CURSO	MUY FRECUENTE ★★★★★
REALIZA LAS ACTIVIDADES PROPUESTAS EN EL PROGRAMA DEL CURSO	SIEMPRE ★★★★★

Figura 20: Instrumento de evaluación actualmente (2025, primer ciclo)

El instrumento de evaluación que ha sido utilizado para calificar a los docentes (Figura 20) consta de una serie de criterios, cuya cantidad ha variado entre 8 y 13 dependiendo del período de tiempo en que haya sido aplicado. El estudiante puede calificar al profesor en cada uno de estos criterios usando una escala de Likert que va desde 1 a 4 hasta 2022, de 1 a 5 para el 2023 y de 1 a 6 para el 2025. Al estudiante, esta escala se le muestra a modo de estrellas, es decir si el alumno considera que el profesor obtiene la calificación más alta en un criterio, marca la cuarta estrella, y esto hace que se rellenen las otras tres.

El profesor debe obtener un promedio de más 80 puntos en la evaluación docente para que se considere una buena evaluación, por lo que se tomó este umbral para estudiar los factores que influyen en esta calificación.

A lo largo del período comprendido entre 2017 y 2023, del total de 6,494 profesores evaluados, el 84 % obtuvo calificaciones superiores a 80 puntos, mientras que solo el 16 % registró calificaciones inferiores (Figura 21). Este resultado refleja un desempeño predominantemente alto por parte de los docentes evaluados.

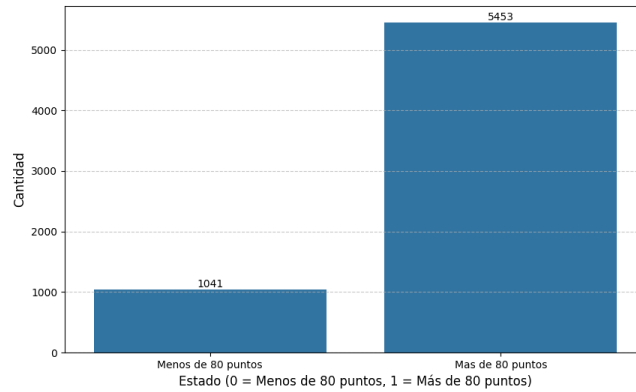


Figura 21: Cantidad de profesores con una nota mayor a 80 vs con menor a 80

Luego se calcularon en promedio la cantidad de respuestas de los estudiantes en cada una de las estrellas en la evaluación para poder analizar si esto influye en la calificación del docente. Dado que la nota del docente se calcula teniendo en cuenta la cantidad de respuestas de 3 y 4 estrellas se elaboró la siguiente gráfica:

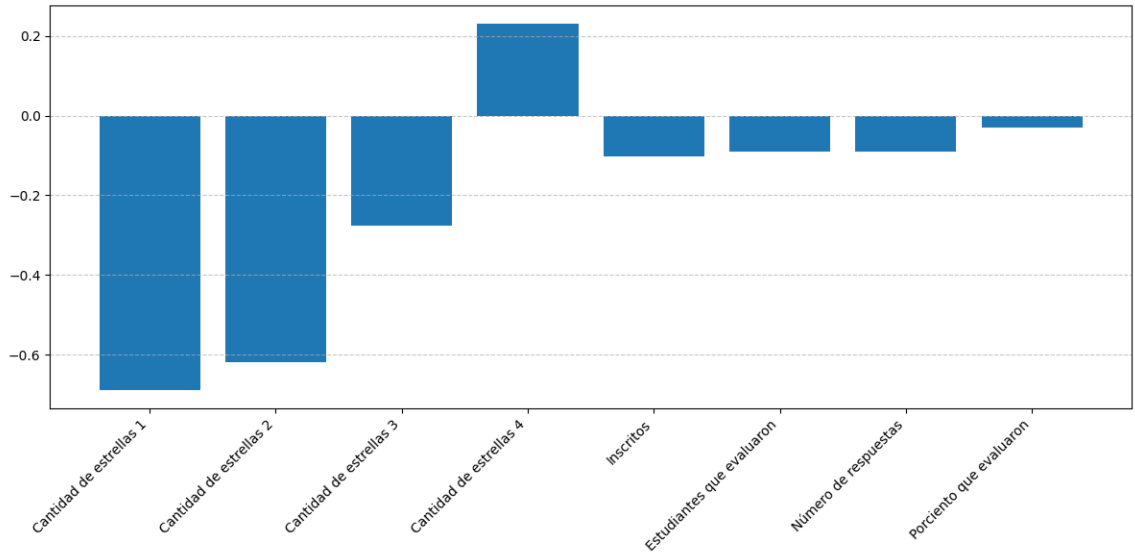


Figura 22: Correlación de variables con la nota de evaluación

En la Figura 22 se pueden observar las variables que tienen alguna influencia sobre la nota del docente. La cantidad de respuestas de 4 estrellas tiene una influencia positiva sobre la nota del docente, pero la correlación no es fuerte (0.23). Sin embargo, la influencia que tienen la cantidad de respuestas de 1 y 2 estrellas con correlaciones de -0.69 y -0.62 respectivamente, sí es fuerte. Esto indica que, mientras mayor sea la cantidad de respuestas de 1 y 2 estrellas en cualquiera de los criterios, la evaluación del profesor será más baja.

### 7.4.1. Árboles de decisión

Se desarrollaron dos árboles de decisión, uno de clasificación y uno de regresión para identificar patrones relacionados con la evaluación del docente. Se tomó como variable respuesta, en el de regresión, la nota de evaluación de desempeño docente, y en el de clasificación la variable que determina si un docente obtiene una nota mayor o menor a 80 puntos.

#### 7.4.1.1. Árbol de regresión

El Cuadro 1 muestra los resultados generales del árbol de decisión generado para la variable de la nota del profesor:

Categoría	Tendencia general en el árbol
Promedio de 1 estrella	<ul style="list-style-type: none"> <li>■ Bajo (<math>\leq 0.28</math>) es muy favorable (predicciones altas para la nota, <math>\sim 90-99\%</math>).</li> <li>■ Alto (<math>&gt; 0.28</math>) reduce fuertemente la predicción de la nota (algunas ramas bajan a <math>&lt; 10\%</math>).</li> </ul>
Promedio de 2 estrellas	<ul style="list-style-type: none"> <li>■ Bajo (<math>\leq 0.32</math>) es muy favorable (<math>\sim 98-99\%</math>)</li> <li>■ Alto (<math>&gt; 0.32</math>) reduce valor, sobre todo combinado con un promedio bajo de 4 estrellas.</li> </ul>
Promedio de 3 estrellas	<ul style="list-style-type: none"> <li>■ Bajo (<math>\leq 0.35, \leq 0.08</math>) es muy favorable (<math>\sim 98-99\%</math>).</li> <li>■ Alto (<math>&gt; 0.35</math>) baja valor, dependiendo de interacción con otras variables.</li> </ul>
Promedio de 4 estrellas	<ul style="list-style-type: none"> <li>■ Alto (<math>&gt; 2.04</math>) mejora valor cuando el promedio de 4 estrellas es bajo.</li> <li>■ Bajo (<math>\leq 0.04</math>) baja la predicción (<math>\sim 0-7\%</math>).</li> </ul>
Promedio de 5 estrellas	<ul style="list-style-type: none"> <li>■ Moderado-alto (<math>&gt; 4.30, &gt; 4.89</math>) mejora valor (<math>\sim 80-90\%</math>)</li> <li>■ Muy bajo (<math>\leq 0.12</math>) o extremadamente alto (<math>&gt; 10.56</math>) reduce valor en algunas ramas.</li> </ul>
Porcentaje de estudiantes evaluadores	<ul style="list-style-type: none"> <li>■ Alto (<math>&gt; 7.5\%</math>) el mejora valor de la nota.</li> <li>■ Bajo (<math>\leq 7.5\%</math>) reduce el valor de la nota, en algunos casos hasta cero.</li> </ul>

Cuadro 1: Resumen por categoría: valores bajos o altos favorables según el árbol de decisión de 5 niveles para predecir la nota de evaluación docente

El análisis del árbol de decisión muestra que mantener bajos los promedios de 1, 2 y 3 estrellas es clave para obtener calificaciones altas en la evaluación docente. En contraste, promedios muy bajos de 3 estrellas afectan negativamente, mientras que valores altos ayudan a mejorar la predicción cuando se combinan con otros factores.

Los valores moderados a altos en el promedio de 4 estrellas suelen favorecer evaluaciones positivas. El promedio de 4 estrellas mejora el valor de predicción cuando los promedios de calificaciones de 1 a 3 estrellas son bajos, actuando como refuerzo positivo. No obstante, cuando las calificaciones más bajas aumentan, el efecto positivo disminuye. Por otra parte, valores extremadamente bajos o altos pueden reducir la calificación esperada, porque el modelo puede interpretar estos altos valores como atípicos, y por ello ajusta la predicción a rangos moderados.

Finalmente, un mayor porcentaje de estudiantes que participan en la evaluación tiene un efecto positivo, mientras que una participación más reducida suele asociarse con peores resultados.

#### 7.4.1.2. Árbol de clasificación

El Cuadro 2 muestra los resultados generales del árbol de decisión generado para la variable categórica obtenida a partir de la calificación del profesor (más de 80 y menos de 80):

Categoría	Tendencia general en el árbol
Promedio de 1 estrella	<ul style="list-style-type: none"> <li>▪ Bajo (<math>\leq 0.22</math>) favorece a tener más de 80 puntos.</li> <li>▪ Alto (<math>&gt; 0.54</math>) generalmente desfavorable.</li> </ul>
Promedio de 2 estrellas	<ul style="list-style-type: none"> <li>▪ Bajo (<math>\leq 0.38</math>) favorece a tener más de 80 puntos.</li> <li>▪ Alto (<math>&gt; 0.52</math>) tiende a tener menos de 80 puntos.</li> </ul>
Promedio de 3 estrellas	<ul style="list-style-type: none"> <li>▪ Bajo-moderado (<math>\leq 1.74</math>) favorece a tener más de 80 puntos.</li> <li>▪ Muy alto (<math>&gt; 2.08</math>) desfavorable.</li> </ul>
Promedio de 4 estrellas	<ul style="list-style-type: none"> <li>▪ Alto (<math>&gt; 2.92</math>) puede elevar la posibilidad de clasificar a más de 80 puntos en ramas específicas.</li> </ul>
Promedio de 5 estrellas	<ul style="list-style-type: none"> <li>▪ Moderado-alto (<math>&gt; 2.74</math>) favorece a tener más de 80 puntos.</li> <li>▪ Muy bajo (<math>\leq 2.74</math>) o muy alto (<math>&gt; 12.86</math>) puede ser desfavorable.</li> </ul>
Porcentaje de estudiantes evaluadores	<ul style="list-style-type: none"> <li>▪ Alto (<math>&gt; 11.62\%</math>) favorece a tener más de 80 puntos.</li> <li>▪ Bajo (<math>\leq 11.62\%</math>) tiende a menos de 80 puntos.</li> </ul>

Cuadro 2: Resumen por categoría: valores bajos o altos favorables según el árbol de decisión de 5 niveles para clasificar si un docente recibirá una nota mayor o menor a 90

Según los resultados del árbol de clasificación, los docentes con promedios bajos de 1, 2 y 3 estrellas tienden a tener calificaciones superiores a 80 puntos. Además, valores moderados o altos en la categoría de 5 estrellas también contribuyen a obtener calificaciones altas,

considerando que no sean valores altos dado que el modelo los llega a tomar como valores atípicos.

Por otra parte, una participación estudiantil superior al 11.6% aparece como un factor diferenciador clave, reforzando la tendencia hacia notas altas. En contraste, combinaciones de promedios altos en estrellas negativas con baja participación tienden a generar calificaciones por debajo de los 80 puntos.

Dado que los patrones encontrados en los árboles de decisión involucran la cantidad de respuestas de 1 y 2 estrellas, se decidió profundizar en el estudio de los valores críticos de estas variables que ponen en riesgo una buena calificación del docente.

#### **7.4.2. Análisis de cantidad de respuestas por número de estrellas**

Se analizaron los comportamientos de la cantidad de estrellas recibidas en general. En la Figura 23 puede observarse que profesores con más de 80 (puntos azules) en la nota de la evaluación obtienen un máximo de 2 respuestas en alguno de los criterios con 1 estrella, en el caso de las respuestas de 2 estrellas puede llegar hasta 4. Para profesores con menos de 80 la dispersión entre la cantidad de respuestas de 1 y 2 estrellas es mayor. Para el caso de la cantidad de respuestas de 3 y 4 estrellas, se puede observar que los profesores que tienen más de 80 tienden a obtener una mayor cantidad de respuestas, que los profesores que obtienen calificaciones menores a 80.

Los gráficos de cajas y bigotes representan que para profesores que obtienen más de 80 puntos la cantidad de respuestas con 1 y 2 estrellas es menor y tienen menos variabilidad que para los que obtienen menos de 80. En el gráfico de la distribución de 3 y 4 estrellas se observa que los profesores que tienen menos de 80 tienen una mayor variabilidad en la cantidad de respuestas de 4 estrellas, pero obtienen una menor cantidad de respuestas que los profesores que obtuvieron una calificación por encima de 80.

Como resumen, se puede observar en la Figura 23 que los profesores que reciben calificaciones por encima de 80 tienden a recibir más cantidad en promedio de respuestas de 3 y 4 estrellas que los profesores que obtienen calificaciones por debajo de 80. Mientras que, estos últimos reciben más respuestas en promedio de 1 y 2 estrellas que los que obtienen los que tienen calificación por encima de 80.

El Cuadro 3 muestra la cantidad promedio de respuestas de 1, 2, 3 y 4 estrellas tanto para los profesores que obtuvieron notas por encima de 80 como para los profesores que quedaron por debajo. El umbral crítico se determina usando la mediana de la cantidad de respuestas en cada una de las estrellas para los profesores que obtuvieron evaluaciones por debajo de 80. Estos umbrales pueden entenderse como límites de riesgo: superarlos puede sugerir la posibilidad de que la evaluación esté por debajo de 80.

Comparación de calificaciones promedio por estrellas y estado

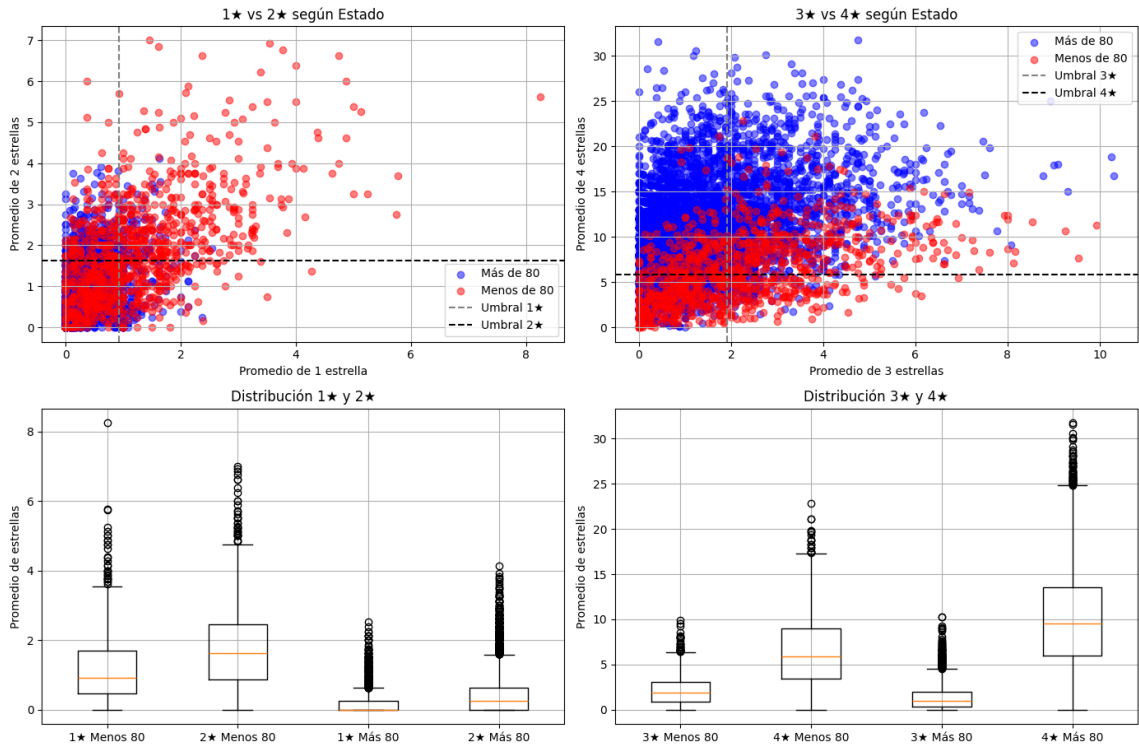


Figura 23: Comparación de calificaciones promedio por estrellas y estado de la nota del docente, más de 80 o menos de 80

Estrella	Promedio de notas menores a 80	Promedio de notas menores a 80	Umbral crítico (Mediana de notas menores a 80)
1	1.20	0.17	0.92
2	1.81	0.46	1.62
3	2.03	1.17	1.91
4	5.17	9.86	5.88

Cuadro 3: Resumen selección de estrellas según notas de evaluaciones de los docentes

La cantidad de respuestas altas de 1 y 2 estrellas sugiere un mayor riesgo de que el profesor obtenga notas por debajo de 80 en las evaluaciones docentes. En el caso de la cantidad de respuestas altas de 4 estrellas, es más común en evaluaciones por encima de 80. El caso de las respuestas de 3 estrellas puede ser ambigua, puesto que es necesario estudiar la combinación con la cantidad de respuestas en otras estrellas para determinar la manera en la que influyen en que un profesor obtenga más o menos de 80.

En la Figura 24 se puede observar que es más alta la cantidad de respuestas promedio de 1, 2 y 3 estrellas para los profesores que obtuvieron menos de 80, no sucede esto en la cantidad promedio de respuestas de 4 estrellas. En este caso, la cantidad promedio de respuestas de 4 estrellas de los profesores que obtuvieron más de 80 puntos es más alta que para los que obtuvieron menos de 80.

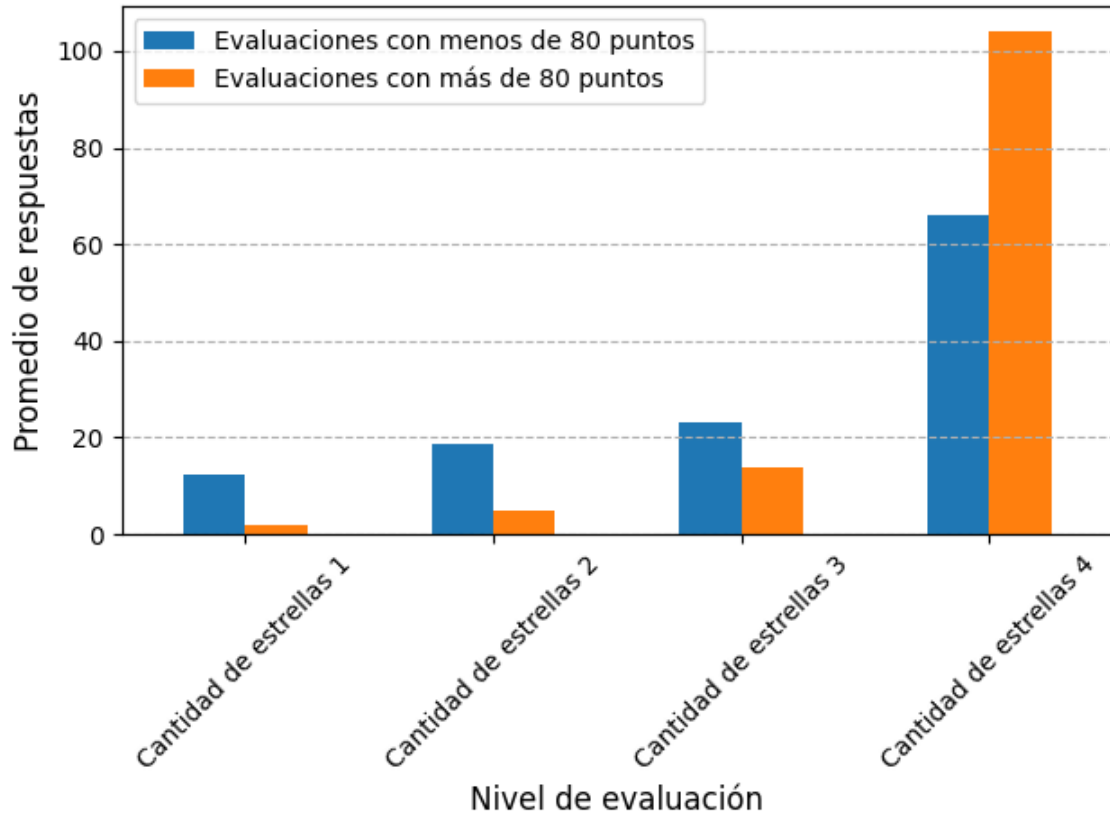


Figura 24: Promedio de cantidad de respuestas por nivel de evaluación

### 7.4.3. Reglas de asociación

Basado en los resultados anteriores, se creó una serie de variables categóricas para poder hacer reglas de asociación (Cuadro 4). Las variables se crearon usando los umbrales críticos descritos anteriormente.

Variable	Descripción
alta_participacion	Porcentaje de estudiantes que evaluaron por encima del 70%
baja_4_estrellas	Cantidad de respuestas de 4 estrellas inferior al umbral crítico
baja_3_estrellas	Cantidad de respuestas de 3 estrellas inferior al umbral crítico
alta_2_estrellas	Cantidad de respuestas de 2 estrellas superior al umbral crítico
alta_1_estrellas	Cantidad de respuestas de 1 estrella superior al umbral crítico
anio_carrera_1	Si el curso está ubicado en el primer año del programa
anio_carrera_2	Si el curso está ubicado en el segundo año del programa
anio_carrera_3	Si el curso está ubicado en el tercer año del programa
anio_carrera_4	Si el curso está ubicado en el cuarto año del programa
anio_carrera_5	Si el curso está ubicado en el quinto año del programa
anio_carrera_6	Si el curso está ubicado en el sexto año del programa

Cuadro 4: Descripción de variables utilizadas en reglas de asociación

Para obtener la variable del año en el que se imparte la asignatura según el programa de estudio, se tomó la moda en el caso de cursos que se imparten en distintos programas en años diferentes.

Se aplicó un algoritmo de aprendizaje de máquinas no supervisado para descubrir patrones ocultos en los datos. Las reglas de asociación se obtuvieron aplicando el método *a priori*, que se utiliza para buscar patrones frecuentes en el conjunto de datos. Para esto se utilizó un soporte mínimo de 0.1 que es para el tamaño del conjunto de datos, un valor que permitirá obtener patrones muy comunes y confiables. De las reglas generadas se obtuvieron aquellas que tienen en el consecuente “mas de 80 puntos en la evaluación docente” o “menos de 80 en la evaluación docente”.

Las 6 reglas más interesantes son las siguientes:

Antecedente	Consecuente	Confianza	Lift	Soporte	Interpretación
alta_2_estrellas, alta_1_estrella	menos_80, alta_ participacion	0.41	8.63	2.87%	A pesar de una alta participación, una proporción elevada de respuestas de 1 y 2 estrellas se asocia con evaluaciones inferiores a 80.

*Continúa en la siguiente página*

<b>Antecedente</b>	<b>Consecuente</b>	<b>Confianza</b>	<b>Lift</b>	<b>Soporte</b>	<b>Interpretación</b>
alta_2_estrellas, baja_4_estrellas	menos_80, alta_1_estrella	0.69	8.48	1.8%	Una baja cantidad de respuestas de 4 estrellas combinada con una alta cantidad de respuestas de 1 y 2 estrellas se asocia fuertemente con evaluaciones inferiores a 80.
anio_carrera_1, alta_1_estrella	alta_2_estrellas, menos_80	0.62	7.61	1.38%	En el primer año de carrera hay una mayor concentración de evaluaciones negativas, con predominio de calificaciones de 1 estrella, lo que tiende a generar evaluaciones inferiores a 80.
alta_ participacion, alta_1_estrella	alta_2_estrellas, menos_80	0.58	7.10	2.87%	La alta participación no garantiza una evaluación positiva cuando predominan las respuestas de 1 y 2 estrellas.
alta_2_estrellas, alta_1_estrella, baja_4_estrellas	menos_80	1.00	6.24	1.8%	La presencia simultánea de respuestas elevadas en 1 y 2 estrellas y bajas en 4 estrellas se asocia consistentemente con evaluaciones inferiores a 80.
alta_1_estrella	menos_80, alta_2_estrellas, anio_carrera_3	0.12	7.05	1.41%	En el tercer año, una alta cantidad de respuestas críticas se asocia con un riesgo moderado de evaluaciones inferiores a 80.

Cuadro 5: Reglas de asociación que tienen en el consecuente variable mas\_80 o menos\_80

### 7.4.3.1. Significación de las reglas de asociación

En las siguientes figuras (Figura 25, Figura 26) se muestran la confianza y el *lift* de las reglas de asociación clave:

La confianza representa la probabilidad condicional de que ocurra el consecuente dado que se dio el antecedente de la regla. La mayoría de estas reglas claves tienen una confianza bastante alta. Es interesante mencionar el caso de la regla 5, en la que, en el 100 % de los casos en los que se da la combinación de alta cantidad de 2 estrellas, alta cantidad de 1 estrella y baja cantidad de 4 estrellas, el docente obtiene una calificación de menos de 80. La mayoría de las reglas más relevantes tienen una confianza de más del 40 %. El caso de la regla 6 no tiene una confianza tan elevada pero tiene un *lift* de más de 7.

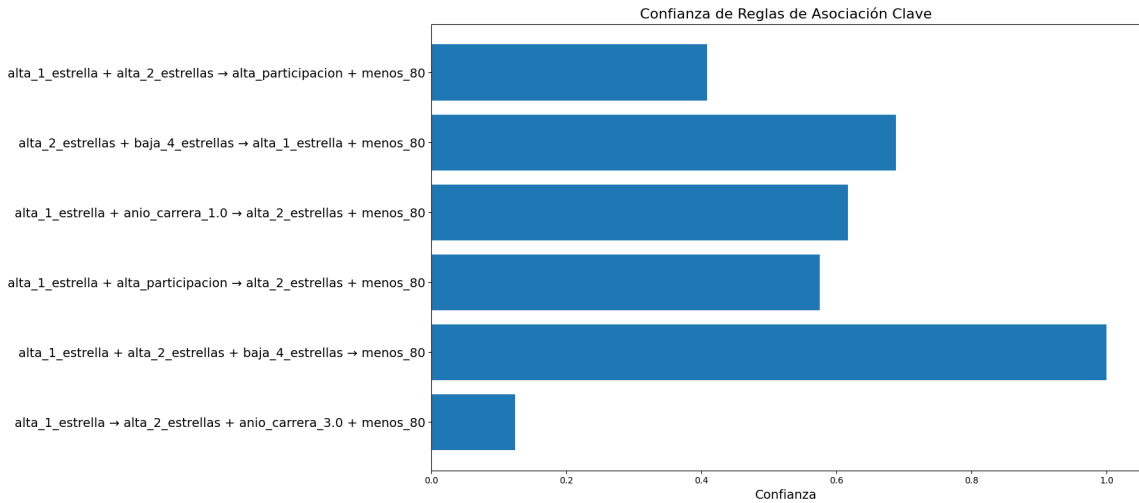


Figura 25: Reglas de asociación (Confianza)

El *lift*, es la métrica que representa la fortaleza de la regla, mientras más grande sea el valor de esta métrica, menos alejada del azar está la ocurrencia de la regla. En el caso de las reglas claves, todas tienen un *lift* mayor a 6, esto significa que es 6 veces más probable que ocurra que si fuera una regla producto del azar. Significa que es un patrón recurrente en los datos. A pesar de que la regla 6 tiene una confianza baja, es 7 veces más probable que ocurra.

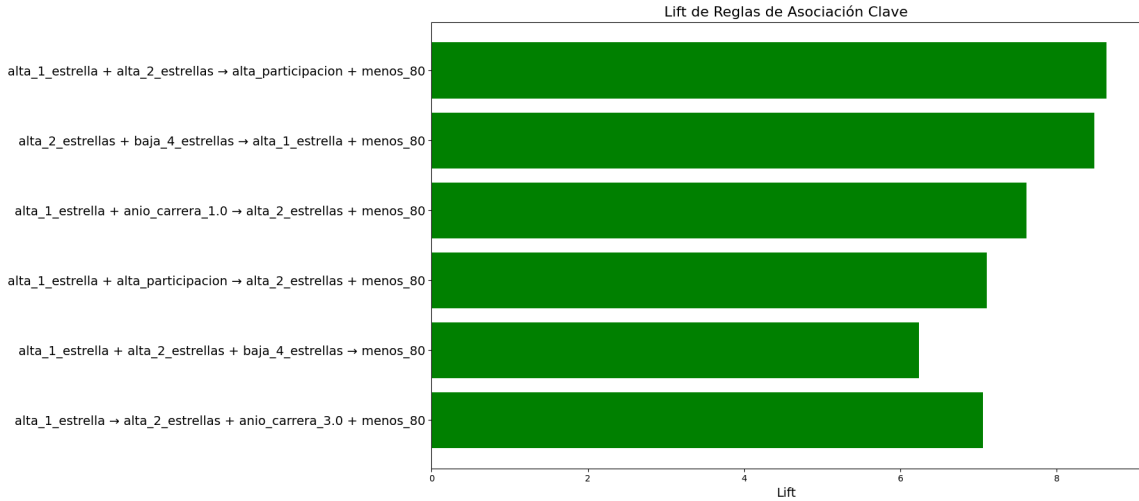


Figura 26: Reglas de asociación (*Lift*)

Analizando estas 6 reglas claves, se puede corroborar que la cantidad de 1 y 2 estrellas son muy buenos predictores de obtener notas en la evaluación docente de menos de 80. Que la participación estudiantil en la evaluación docente no es suficiente para revertir una evaluación de menos de 80. El *lift* de las reglas está por encima de 6.24 por lo que todas estas reglas claves se consideran patrones pues no ocurren al azar.

## Detección de patrones para evaluaciones mayores y menores a 80

Se dividió el conjunto de datos en dos, uno que contenía todas las evaluaciones por encima de 80 y otro con las evaluaciones por debajo de 80. El objetivo fue detectar patrones en ambos grupos por separado. Las reglas de asociación más interesantes se seleccionaron de acuerdo con la confianza, el soporte y el *lift*. Se consideró un *lift* fuerte por encima de 1.5.

### 7.4.3.2. Reglas de asociación para evaluaciones docentes con notas de evaluación docente mayores a 80

Para el caso de las reglas en las que se incluye en el consecuente “nota mayor a 80” no se encontraron con un *lift* mayor a 1.5 por lo que se seleccionaron las que tenían mayor a 1. Las reglas más relevantes para este grupo son las siguientes:

Antecedente	Consecuente	Confianza	<i>Lift</i>	Soporte	Interpretación
baja_4_estrellas	mas_80, baja_3_estrellas	0.93	1.26	22.06 %	Las evaluaciones superiores a 80 pueden presentarse aun cuando la cantidad de respuestas de 3 y 4 estrellas no sea predominante.
anio_carrera_5	mas_80, baja_3_estrellas	0.85	1.15	15.97 %	Los cursos de quinto año presentan una tendencia hacia evaluaciones superiores a 80, aunque la concentración de respuestas de 3 estrellas sea baja.
anio_carrera_4	mas_80, baja_3_estrellas	0.79	1.08	19.22 %	En los cursos de cuarto año, una baja cantidad de respuestas de 3 estrellas no impide obtener evaluaciones superiores a 80.
anio_carrera_3	mas_80, baja_3_estrellas	0.75	1.01	19.31 %	Los cursos de tercer año tienden a recibir evaluaciones superiores a 80, aun cuando la cantidad de respuestas de 3 estrellas sea baja.

Cuadro 6: Reglas de asociación para notas mayores a 80

Las reglas de asociación para este grupo revelan que pueden obtenerse evaluaciones superiores a 80 aun cuando no predominen las respuestas de 3 y 4 estrellas. Este patrón se observa con mayor frecuencia en cursos impartidos en años superiores (tercer, cuarto y quinto año), donde se mantiene una tendencia hacia evaluaciones favorables. Algunos aspectos interesantes son:

- La presencia de pocas respuestas de 3 y 4 estrellas no impide una alta nota, lo que sugiere que otros factores, que pudieran ser la madurez del estudiante o el contenido o relevancia del curso para el programa de estudio, influyen de forma positiva en la obtención de calificaciones por encima de 80.
- A medida que los cursos se imparten en años avanzados la probabilidad de obtener una nota mayor a 80 se incrementa.
- Estas reglas tienen un *lift* moderado (1.01–1.26), pero tienen alta confianza ( $\geq 0.75$ ) y soporte sólido ( $>15$ ), lo que las convierte en patrones útiles para el análisis.

Los aspectos interesantes se refuerzan en la siguiente gráfica:

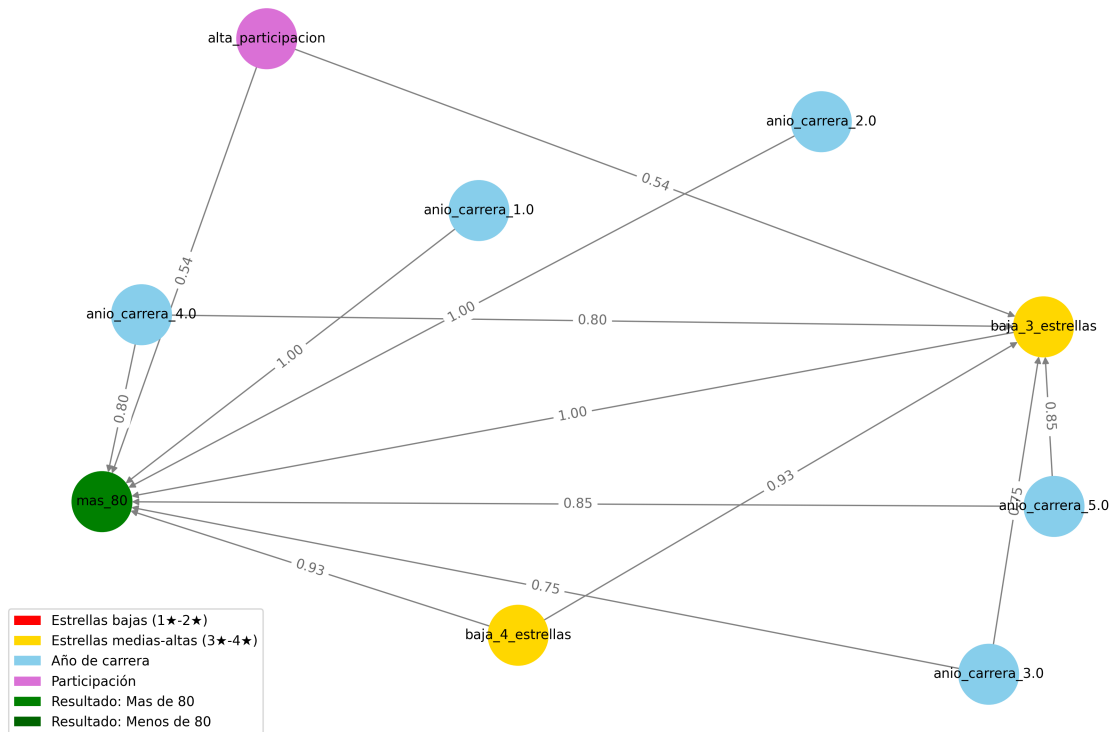


Figura 27: Red de asociación de notas mayores a 80 puntos

Este grafo, representado en la Figura 27, evidencia que las evaluaciones docentes con más de 80 puntos están asociadas a:

- Altos niveles de participación estudiantil.
- Años de carrera iniciales y finales (primero, segundo, cuarto y quinto).

Además, se evidencia que si la cantidad de respuestas de 3 y 4 estrellas son bajas, pueden obtenerse notas por encima de 80, lo que indica que evitar alta cantidad de respuestas de 1 y 2 estrellas tiene más impacto para obtener mejores puntuaciones.

### 7.4.3.3. Reglas de asociación para evaluaciones docentes con notas menores a 80

En el caso de las reglas de asociación para evaluaciones por debajo de 80 se seleccionaron las reglas que tienen un *lift* por encima de 1.5. Las más interesantes son las siguientes:

Antecedente	Consecuente	Confianza	Lift	Soporte	Interpretación
alta_ participación, alta_1_estrella	alta_2_estrellas, menos_80	0.86	1.70	17.96%	A pesar de una alta participación, cuando hay muchas respuestas de 1 estrella, también se dan muchas respuestas de 2 estrellas y se asocia con evaluaciones inferiores a 80.
alta_2_estrellas, alta_ participación	menos_80, alta_1_estrella	0.79	1.55	17.96%	Una alta cantidad de respuestas en 2 estrellas incluso con participación alta, suele estar acompañada de muchas respuestas en 1 estrella y termina en una nota menor a 80.
alta_ participación	alta_2_estrellas, menos_80	0.78	1.53	22.86%	Una alta participación por sí sola no garantiza una evaluación favorable cuando predominan las respuestas de 2 estrellas.

*Continúa en la siguiente página*

Antecedente	Consecuente	Confianza	Lift	Soporte	Interpretación
anio_carrera_5, baja_3_estrellas	menos_80, baja_4_estrellas	0.84	1.68	11.43 %	En cursos de quinto año cuando la cantidad de respuestas de 3 estrellas es baja conlleva a una cantidad de respuestas bajas también de 4 estrellas y el riesgo de obtener calificaciones menores a 80.
alta_2_estrellas, alta_1_estrella	menos_80, alta- participación	0.50	1.70	17.96 %	Una proporción elevada de respuestas de 1 y 2 estrellas se asocia con evaluaciones inferiores a 80 incluso si hubo participación.

Cuadro 7: Reglas de asociación para notas menores a 80

Las reglas de asociación para este grupo evidencian que las evaluaciones de menos de 80 están determinadas principalmente por la cantidad de respuestas en las diferentes estrellas de todos los criterios. Algunos aspectos interesantes son:

- La combinación de muchas respuestas en 1 y 2 estrellas es un predictor consistente de notas menores a 80 incluso cuando la participación de los estudiantes es elevada.
- Si cantidad de respuestas de 3 y 4 estrellas que dan los estudiantes de quinto año es escasa, el docente enfrenta una alta probabilidad de no alcanzar más de 80 en su evaluación.
- La alta participación, en lugar de ser un factor protector, puede reforzar la validez de una evaluación menor a 80.

Los aspectos interesantes se refuerzan en la Figura 28:

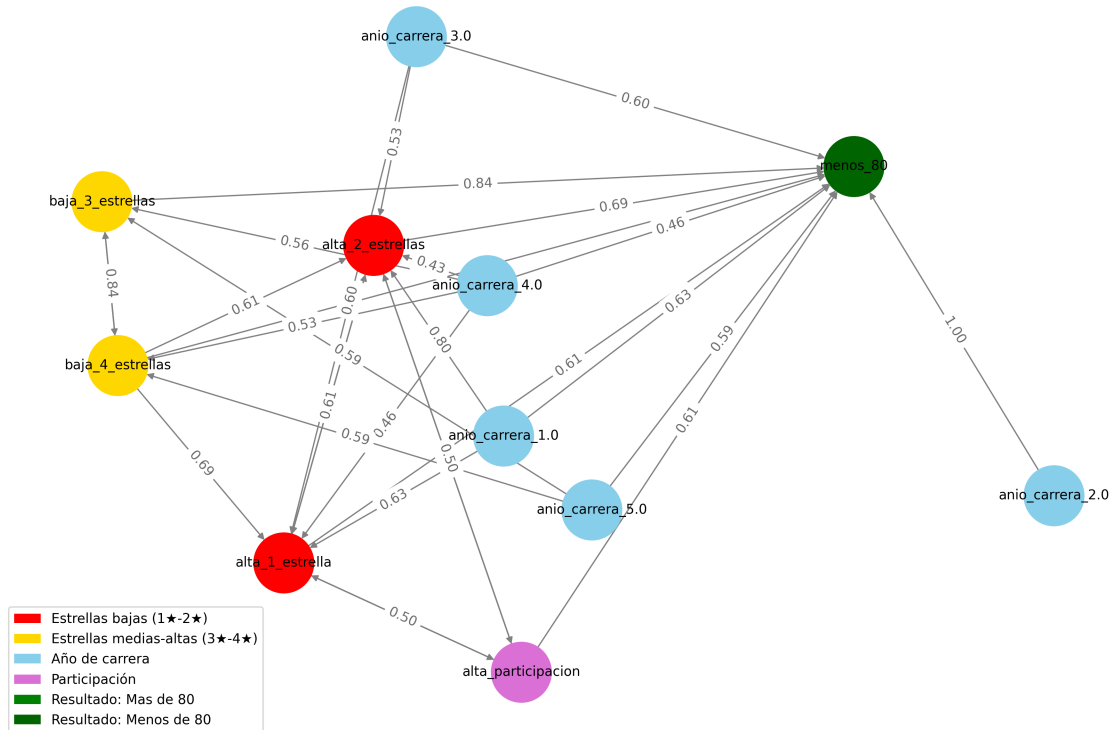


Figura 28: Red de asociación de notas menores a 80

Esta red muestra que las evaluaciones de menos de 80 están fuertemente vinculadas a alta cantidad de respuestas de 1 y 2 estrellas y a veces incluso con alta participación estudiantil. Además:

- La percepción negativa está correlacionada internamente: donde hay muchas respuestas de 2 estrellas también hay muchas de una 1 estrella y escasean la cantidad de respuestas de 3 y 4 estrellas.
- El segundo año de carrera destaca como un punto crítico donde, según los datos, siempre se produce una nota menor a 80.
- La participación no mitiga la obtención de notas inferiores a 80 si la percepción del curso es negativa.

## 7.5. Agrupamiento

Se utilizó un algoritmo de aprendizaje no supervisado (K-medias) para elaborar grupos que permitieran obtener patrones interesantes en los datos. Para esto se elaboró un gráfico de codo (Figura 29) para determinar la cantidad óptima de grupos dadas las características del conjunto de datos.

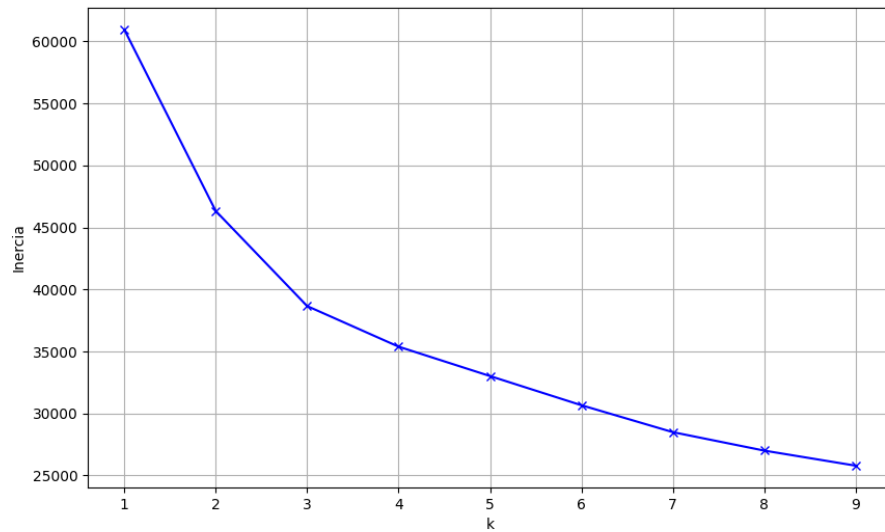


Figura 29: Método del codo para determinar el número óptimo de grupos

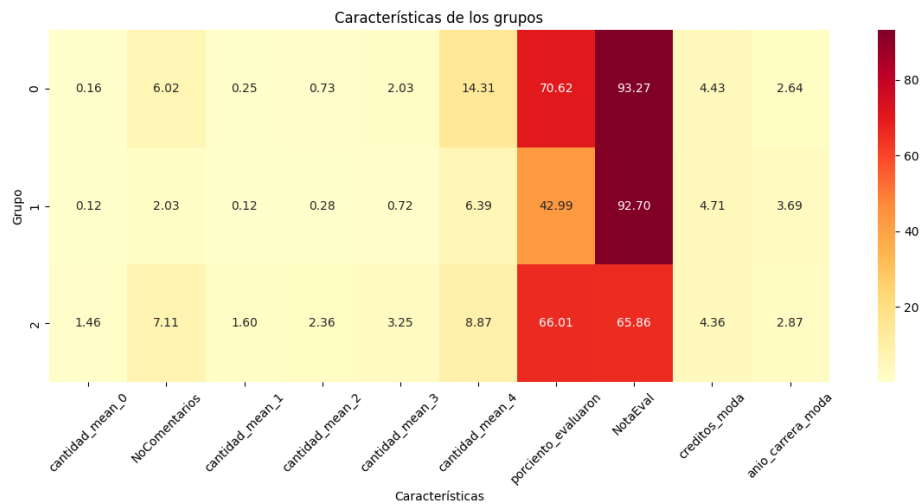


Figura 30: Características de los grupos

Según el diagrama de codo (Figura 29) para este conjunto de datos se pueden calcular 3 grupos. El resultado del agrupamiento se muestra en la Figura 30:

Las características de los grupos son las siguientes:

### **Grupo 1: evaluaciones altamente positivas y participativas**

- Tamaño: 2,253 casos
- Nota promedio: 93.27
- Cantidad de evaluaciones de profesores con notas por encima de 80 puntos: 97.38 %
- **Promedios de estrellas:**
  - Cantidad de respuestas de 4 estrellas: 14.32 (muy alto)
  - Cantidad de respuestas de 3 estrellas: 2.04
  - Cantidad de respuestas de 1 y 2 estrellas: muy bajos (entre 0.25 y 0.73)
- Cantidad de comentarios: 6 en promedio
- Participación: 70.6 %
- Año promedio de carrera: 2.6

Este grupo representa cursos donde los estudiantes participan activamente, emiten comentarios y otorgan principalmente 4 estrellas, lo cual se refleja en notas muy altas en las evaluaciones docentes. En este grupo el 97 % de los docentes tienen notas superiores a 80. Además, los estudiantes suelen encontrarse en años intermedios de carrera (segundo y tercer año).

### **Grupo 2: evaluaciones pasivas pero positivas**

- Tamaño: 3,130 casos
- Nota promedio: 92.69
- Cantidad de evaluaciones de profesores por encima de 80: 89.68 %
- **Promedios de estrellas:**
  - Cantidad de respuestas de 4 estrellas: 6.40 (bajo comparado con el Grupo 1)
  - Cantidad de respuestas de 1, 2 y 3 estrellas: muy bajos, (entre 0.12 y 0.72)
- Comentarios: 2 en promedio
- Participación: 43.0 % (muy baja)
- Año promedio de carrera: 3.7

Este grupo representa evaluaciones con poco involucramiento, es decir, pocos comentarios y baja participación, pero aún así altas notas (en promedio 92). Predomina en años avanzados (segundo y tercero). Esto podría indicar una actitud de cumplimiento formal más que de compromiso reflexivo con la evaluación docente.

### **Grupo 3: evaluaciones críticas y reprobadas**

- Tamaño: 713 casos
- Nota promedio: 65.82
- Cantidad de evaluaciones por encima de 80: 15.57 %
- **Promedios de estrellas:**
  - Cantidad de respuestas de 1 y 2 estrellas : muy altas (entre 1.6 y 2.36)
  - Cantidad de respuestas de 3 estrellas: más alto (3.25)
  - Cantidad de respuestas de 4 estrellas: más bajo (8.86)
- Comentarios: 7.1 (más que cualquier grupo)
- Participación: 66 %
- Año promedio de carrera: 2.86

Este grupo representa evaluaciones altamente críticas, debido a que los estudiantes dejan muchos comentarios, hay una alta participación, y las estrellas bajas dominan. El resultado es una cantidad de evaluaciones por debajo de 80 elevada. Suele ocurrir en cursos de nivel intermedio (segundo y tercer año), lo cual puede coincidir con mayor exigencia y frustración.

## 7.6. Criterios de evaluación

### Período (2017 - 2019)

A lo largo de los años en los que se enmarca este estudio ha habido 3 cambios en los criterios en los que los estudiantes evalúan al profesor. Hubo un cambio de estos criterios en 2020 que se mantuvo hasta 2022. Luego hubo otro cambio en 2023 que es el que se mantiene hasta la actualidad. Dado que los criterios de evaluación cambian de acuerdo al año en el que se desarrolla la evaluación, el análisis de estos se dividió por los períodos en los que estuvo vigente ese instrumento.

En el período de 2017 a 2019 los criterios que se encontraban en el instrumento de evaluación eran los siguientes:

1. Tiene un conocimiento profundo y actualizado de la materia que imparte.
2. Define claramente los aprendizajes a alcanzar al inicio del ciclo en el programa y durante el desarrollo del proceso.
3. Establece claramente la forma de trabajar en la clase.
4. Guía a los estudiantes para que alcancen aprendizajes significativos.
5. Muestra a los estudiantes aplicaciones prácticas de la materia.
6. Emplea efectivamente diversas metodologías y estrategias didácticas innovadoras que promueven el éxito de los estudiantes.
7. Verifica en el proceso que los estudiantes hayan alcanzado los aprendizajes propuestos.
8. Define claramente la forma de evaluar los aprendizajes.
9. Proporciona información oportuna y valiosa al evaluar los aprendizajes utilizando el sistema de registro académico.
10. Tiene disponibilidad para atender y apoyar a los estudiantes.
11. Promueve entre los estudiantes la confianza en sí mismos para aprender la materia.
12. Brinda ayuda para que los estudiantes resuelvan dudas y progresen en su proceso de aprendizaje.
13. Utiliza efectivamente tecnología educativa y las plataformas de Canvas para fortalecer los aprendizajes.

En el instrumento de calificación podrían agruparse los criterios en áreas clave, una posible se muestra en el Cuadro 8:

Área	Preguntas relacionadas	Descripción
Dominio del contenido y planificación	1, 2, 3	Evalúa el conocimiento del docente, claridad en la definición de objetivos y organización de la clase.
Desarrollo didáctico y metodológico	4, 5, 6, 7	Se enfoca en cómo guía el aprendizaje, uso de estrategias didácticas y verificación de aprendizajes.
Evaluación y retroalimentación	8, 9	Examina cómo el docente comunica criterios de evaluación y entrega retroalimentación efectiva.
Relación interpersonal y apoyo	10, 11, 12	Valora la disponibilidad, empatía y acompañamiento que ofrece el docente al estudiante.
Tecnología educativa	13	Evalúa el uso adecuado de plataformas y recursos digitales.

Cuadro 8: Agrupación de preguntas del cuestionario del período 2017-2019 en áreas

Como la cantidad de respuestas en cada una de las estrellas para cada uno de los criterios define la evaluación del profesor, se analizaron que criterios que recibieron mayor cantidad de respuestas en cada una de las 4 estrellas disponibles. La cantidad de respuestas de 1 y 2 estrellas es significativamente menor que la cantidad de respuestas de 3 y 4 estrellas por lo que se realizó un gráfico para cada grupo. La Figura 31 muestra la cantidad promedio de respuestas que hay de 1 y 2 estrellas para cada uno de los criterios que componen el instrumento.

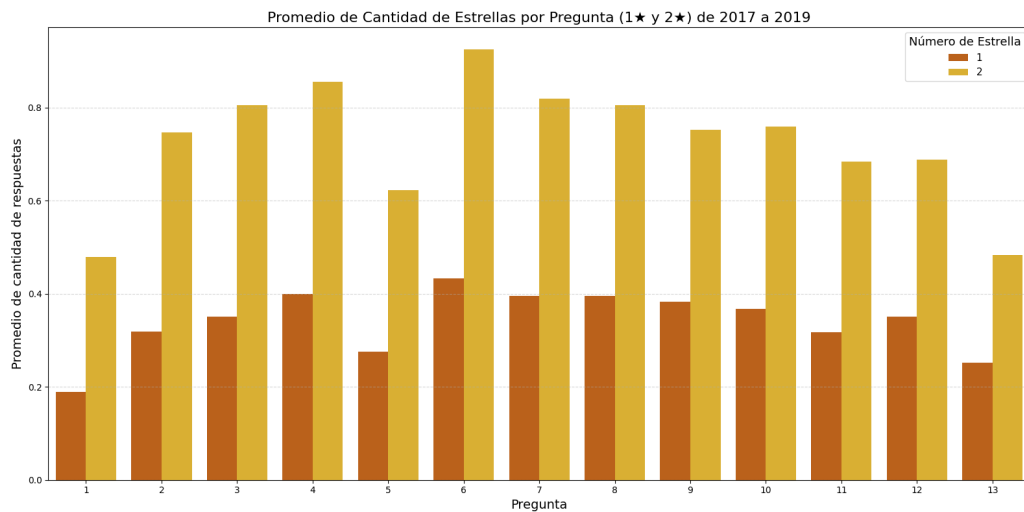


Figura 31: Promedio de cantidad de estrellas por pregunta (1 y 2 estrellas) de 2017 a 2019

En este período, los criterios peor calificados se encuentran en el área metodológica y de verificación de los aprendizajes. Los estudiantes manifiestan mayor insatisfacción cuando perciben deficiencias en estas áreas. Las preguntas asociadas con relación interpersonal (10–12), disponibilidad del docente y conocimiento del contenido (pregunta 1) presentan bajos niveles de calificación de 1 y 2 estrellas. Sin embargo, las preguntas que en general tienen peores calificaciones son:

- Pregunta 6: “Emplea efectivamente diversas metodologías y estrategias didácticas innovadoras que promueven el éxito de los estudiantes”
- Pregunta 4: “Guía a los estudiantes para que alcancen aprendizajes significativos”
- Pregunta 7: “Verifica en el proceso que los estudiantes hayan alcanzado los aprendizajes propuestos”

Todas pertenecen al área de desarrollo didáctico y metodológico del curso.

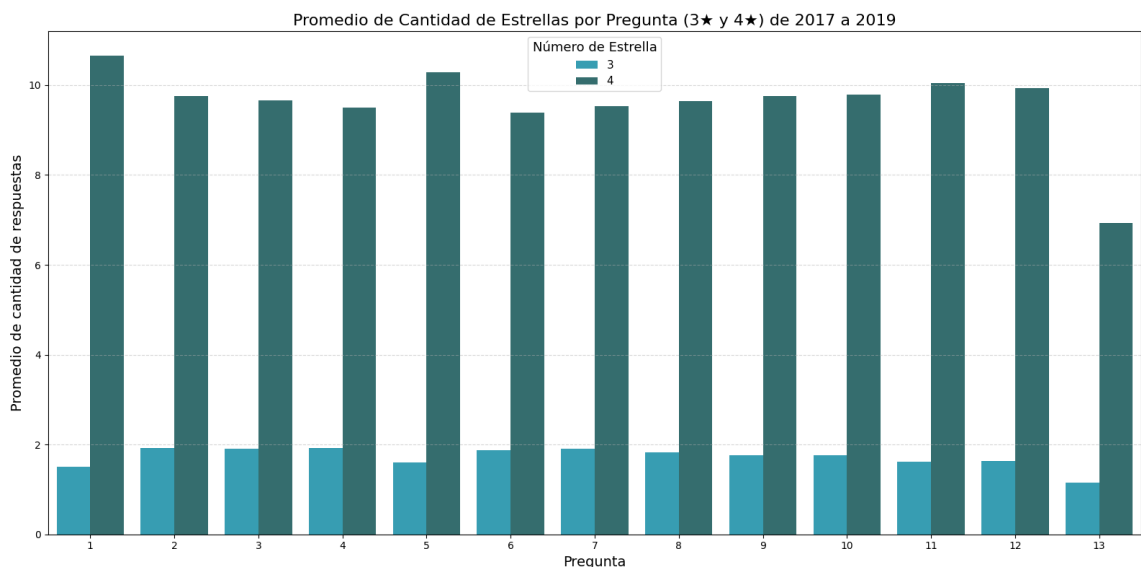


Figura 32: Promedio de cantidad de estrellas por pregunta (3 y 4 estrellas) de 2017 a 2019

Para el caso de 3 y 4 estrellas (Figura 32), el comportamiento es interesante, puesto que hay muchas más respuestas de 4 estrellas que de 3. Los criterios que son mejor evaluados en este período son:

- Pregunta 1: “Tiene un conocimiento profundo y actualizado de la materia que imparte”
- Pregunta 5: “Muestra a los estudiantes aplicaciones prácticas de la materia”

Lo que sugiere que, lo que más valoraban los estudiantes era la seguridad del docente al impartir los contenidos y el hecho de demostrar la aplicación práctica que estos tienen.

El criterio que menos evaluaciones de 4 estrellas recibió fue el 13 (“Utiliza efectivamente tecnología educativa y las plataformas de Canvas para fortalecer los aprendizajes”).

Aunque la pregunta 13 no es la que más calificaciones negativas presenta (1 y 2 estrellas) tiene el promedio más bajo en cantidad promedio de respuestas de 3 y 4 estrellas. Esto sugiere que, previo a la pandemia, los estudiantes no percibían un uso efectivo de plataformas como Canvas.

### **Período (2020 - 2022)**

Con el comienzo de la pandemia en el año 2020, fue necesario adaptar la metodología de los cursos a una modalidad virtual por el confinamiento que se decretó. Fue por eso que el instrumento de evaluación de catedráticos cambió también. En el período de 2020 a 2022, los criterios que se encontraban en el instrumento de evaluación eran los siguientes:

1. Adapta, modifica, o transforma el programa del curso para que sus estudiantes aprendan en un contexto remoto.
2. Logra un equilibrio en la asignación de trabajos individuales o grupales para alcanzar los aprendizajes.
3. Crea y favorece un ambiente apropiado para aprender e intercambiar experiencias dentro y fuera del aula virtual.
4. Mantiene una comunicación positiva y cercana con sus estudiantes.
5. Crea espacios donde responde, escucha y toma en cuenta las necesidades y situaciones particulares de los estudiantes.
6. Evalúa continuamente cómo se avanza en el proceso de los aprendizajes.
7. Realimenta el avance en el aprendizaje
8. Utiliza la plataforma para mantener informados a sus estudiantes de su progreso.

En el instrumento de calificación podrían agruparse los criterios en áreas clave, una posible se describe en el Cuadro 9:

En este período en particular, el país estaba atravesando por la etapa del confinamiento, por lo que la modalidad de clases tuvo que cambiarse de forma abrupta y sin preparación previa a la virtualidad. Se realizó el mismo análisis que para el período anterior.

Las preguntas con mayor cantidad promedio de respuestas de 1 y 2 estrellas en conjunto son:

- Pregunta 7 (“Realimenta el avance en el aprendizaje”) destaca la cantidad más alta de respuestas de 2 estrellas y también un pico en la selección de 1 estrella.
- Las preguntas 2 (“Logra equilibrio en los trabajos”) y 3 (“Crea un ambiente apropiado para aprender”) también presentan valores relativamente altos en cuanto a la cantidad de respuestas promedio de 1 y 2 estrellas.

Área	Preguntas relacionadas	Descripción
Dominio del contenido y planificación	1, 2	Evalúa la capacidad del docente para adaptar el curso al contexto (por ejemplo, remoto), así como planificar de forma equilibrada las actividades para alcanzar los aprendizajes.
Desarrollo didáctico y metodológico	3	Se enfoca en cómo el docente promueve un ambiente adecuado para el aprendizaje y la interacción educativa en entornos virtuales.
Evaluación y retroalimentación	6, 7	Examina si el docente realiza seguimiento continuo al proceso de aprendizaje y ofrece retroalimentación útil para mejorar.
Relación interpersonal y apoyo	4, 5	Valora la cercanía, empatía y disponibilidad del docente para escuchar, comunicarse y atender necesidades individuales del estudiante.
Tecnología educativa	8	Evalúa el uso de plataformas digitales para mantener a los estudiantes informados sobre su progreso académico.

Cuadro 9: Agrupación de preguntas del cuestionario del período 2020–2022 en áreas

La figura 33 sugiere que, aunque el nivel de insatisfacción general no es extremo (los promedios están por debajo de 1 en ambos casos), existen focos recurrentes de evaluación crítica en áreas como la evaluación y retroalimentación del aprendizaje, el dominio del contenido y la planificación y el ambiente educativo virtual.

Para el caso de 3 y 4 estrellas, la figura 32 muestra que en todas las preguntas, el promedio de respuestas de 4 estrellas es notablemente más alto que el de 3, lo que indica una tendencia clara de los estudiantes a calificar con la máxima puntuación cuando perciben un buen desempeño docente. La pregunta con mayor cantidad promedio de respuestas (mayor satisfacción, es decir, estrella 4) es la pregunta 8 (“Utiliza la plataforma para mantener informados a sus estudiantes de su progreso”) con un mayor promedio de 4 estrellas. Seguida de las preguntas:

- Pregunta 5: “Crea espacios donde responde, escucha y toma en cuenta las necesidades y situaciones particulares de los estudiantes”
- Pregunta 4: “Mantiene una comunicación positiva y cercana con sus estudiantes”

Estas dos últimas pertenecen al área de relación interpersonal y de apoyo, lo que sugiere que a pesar del confinamiento y el paso a la virtualidad, el estudiante se sintió acompañado y escuchado por el docente.

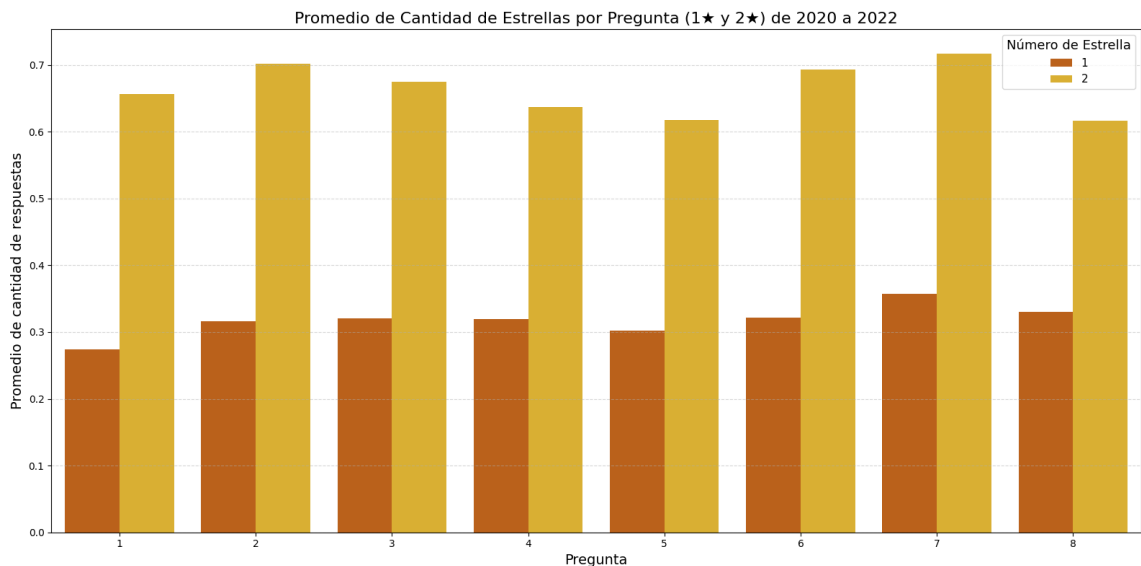


Figura 33: Promedio de cantidad de estrellas por pregunta (1 y 2 estrellas) de 2020 a 2022

### Período (2023)

Para los últimos años de estudio, ya se había cambiado la modalidad de los cursos nuevamente a presencial, por lo que el instrumento de evaluación fue reformulado nuevamente. Desde el año 2023 hasta el momento, el cuestionario de la evaluación docente tiene las siguientes preguntas:

1. Muestra dominio del curso que imparte.
2. Es creativo, flexible e innovador en su estilo de enseñanza.
3. Brinda ayuda para que los estudiantes resuelvan dudas y progresen.
4. Promueve la confianza, la comunicación y el respeto en el aula.
5. Acompaña a los estudiantes para que alcancen los aprendizajes propuestos en el programa del curso.
6. Realiza las actividades propuestas en el programa del curso.
7. Define claramente la forma de evaluar (tareas, verificaciones, foros, exposiciones en clase etc.).
8. Evalúa de diversas formas.
9. Evalúa durante todo el ciclo.
10. La información, recursos (videos, enlaces, artículos..) y calificaciones están accesibles y organizadas en Canvas.
11. Tomaría otro curso con este docente.

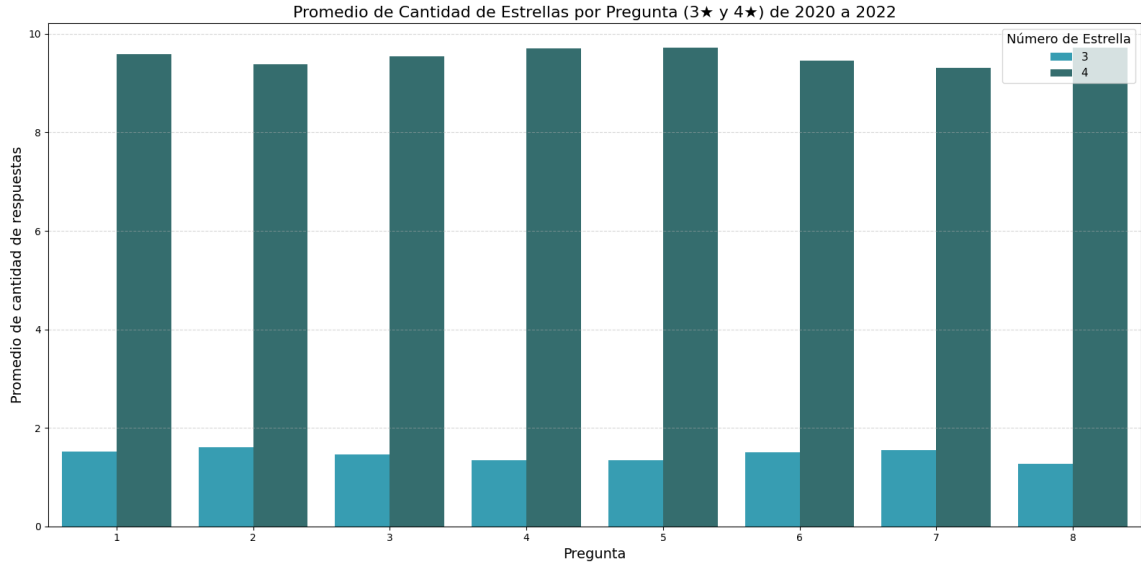


Figura 34: Promedio de cantidad de estrellas por pregunta (3 y 4 estrellas) de 2020 a 2022

Como en años anteriores, es posible agrupar por áreas en las preguntas del cuestionario, la distribución queda de la siguiente forma (Cuadro 10):

Área	Preguntas relacionadas	Descripción
Dominio del contenido y planificación	1, 2, 5, 6	Evalúa conocimiento disciplinar, innovación, y cumplimiento del programa.
Relación interpersonal y apoyo	3, 4	Se enfoca en la empatía, confianza y acompañamiento al estudiante.
Evaluación y retroalimentación	7, 8, 9	Analiza claridad en la evaluación, diversidad de métodos y frecuencia.
Tecnología educativa	10	Evalúa organización y disponibilidad de materiales en plataformas.
Satisfacción global	11	Refleja la intención del estudiante de repetir curso con el docente.

Cuadro 10: Agrupación de preguntas del cuestionario del período 2023 en áreas

Luego de un par de años en confinamiento, la modalidad cambió a clases presenciales. Esto explica la modificación del instrumento de evaluación.

La distribución de bajas calificaciones en cada una de las preguntas del cuestionario (cantidad promedio de 1 y 2 estrellas) se muestra en la Figura 35:

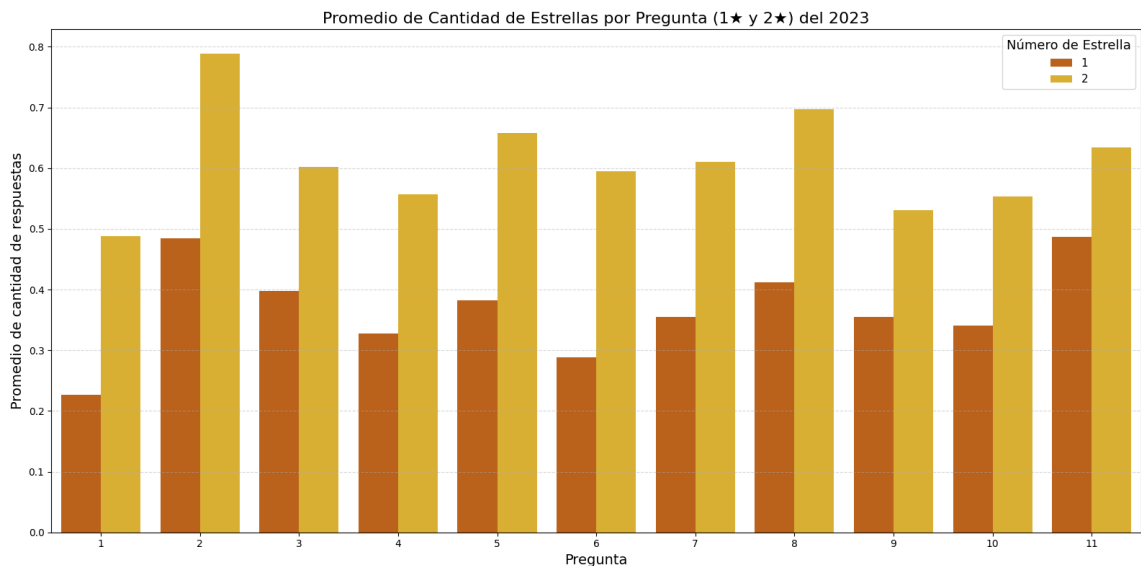


Figura 35: Promedio de cantidad de estrellas por pregunta (1 y 2 estrellas) de 2023

La pregunta 11 (“Tomaría otro curso con este docente”) obtuvo el mayor promedio de valoraciones de 1 estrella, lo que refleja un descontento claro y general hacia el docente. Posteriormente, se presentaron las preguntas:

- Pregunta 2: “Es creativo, flexible e innovador en su estilo de enseñanza”
- Pregunta 8: “Evalúa de diversas formas”
- Pregunta 3: “Brinda ayuda para que los estudiantes resuelvan dudas y progresen”
- Pregunta 5: “Acompaña a los estudiantes para que alcancen los aprendizajes propuestos en el programa del curso”

Estos resultados evidencian deficiencias en el proceso de evaluación, en el apoyo al estudiante y en el acompañamiento pedagógico.

La Figura 36 muestra para 3 y 4 estrellas, la cantidad promedio de respuestas:

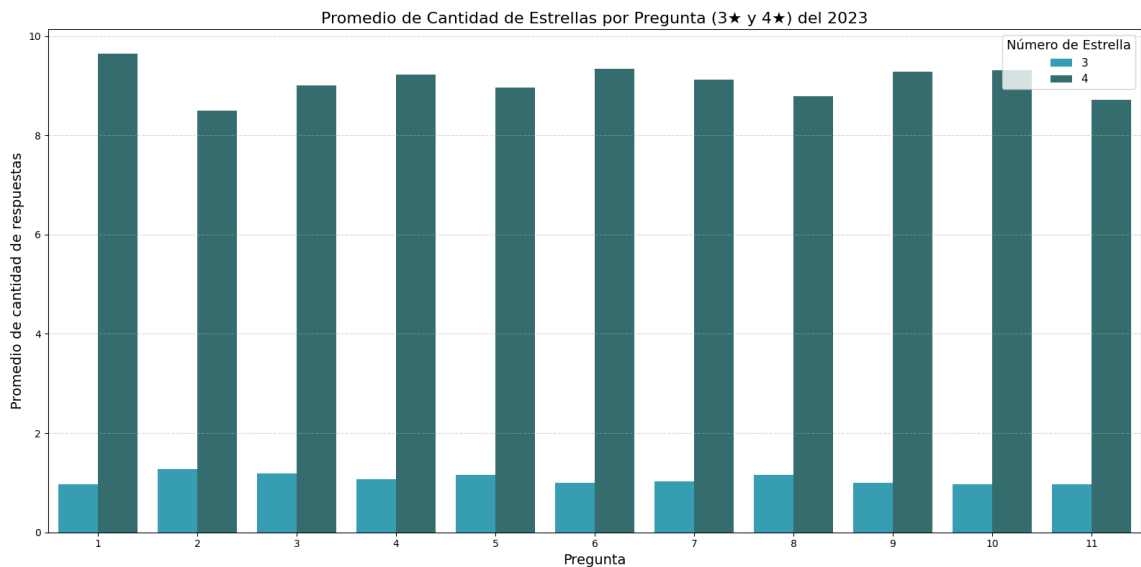


Figura 36: Promedio de cantidad de estrellas por pregunta (3 y 4 estrellas) de 2023

Las preguntas 2 (“Es creativo, flexible e innovador en su estilo de enseñanza”) y 8 (“Evalúa de diversas formas”) presentan mayor cantidad de respuestas de 3 estrellas, lo cual coincide con observaciones críticas previas en estrellas bajas (Figura 35).

Por tanto, esto significa que existe un segmento que siente falta de innovación o variedad en la evaluación; la pregunta 6 (“Realiza las actividades propuestas en el programa del curso”) es una de las mejor valoradas, lo que refleja profesionalismo y cumplimiento del docente. Por otra parte, las preguntas sobre relación interpersonal (3 y 4) mantienen altos niveles, lo que sugiere un ambiente saludable de enseñanza-aprendizaje.

Cuando se analizan los criterios en todo el período de estudio (2017-2023) se pueden identificar patrones clave en todos los formularios. Los criterios con más cantidad de respuestas promedio en 1 y 2 estrellas son:

- **Innovación y metodología didáctica:** en todos los períodos, las preguntas que exploran el uso de estrategias nuevas o creativas presentan mayor cantidad de respuestas de 1 y 2 estrellas.
- **Claridad o variedad en la evaluación:** en los ciclos 2020–2022 y 2023, también se identificaron más respuestas críticas en preguntas sobre formas de evaluación y diversidad de instrumentos.

En cuanto a las áreas con mayor desempeño, en las que había mayor cantidad de respuestas de 3 y 4 estrellas fueron:

- **Dominio del contenido:** el conocimiento profundo del tema por parte del docente se valora altamente en todos los períodos.
- **Relación interpersonal y acompañamiento:** la empatía, comunicación y disponibilidad del docente han recibido altos promedios de respuestas de 3 y 4 estrellas sostenidamente.
- **Cumplimiento del programa:** hay una percepción clara de que la mayoría de docentes cumplen con lo planificado.

El presente estudio tuvo como propósito analizar los factores más influyentes en las evaluaciones de desempeño docente, con el fin de comprender los resultados obtenidos a lo largo de los años y aportar una interpretación más precisa del proceso evaluativo en el contexto universitario. A través del análisis exploratorio de datos históricos, se logró identificar patrones, tendencias y relaciones entre diversas variables académicas y la puntuación asignada a los docentes, por lo que se alcanzaron los objetivos propuestos.

Se identificó la influencia de la participación estudiantil en distintos rangos de calificación docente. Aunque en términos generales no se evidenció una relación significativa entre el porcentaje de participación y la nota docente, una mayor participación puede aportar mayor representatividad y consistencia a los resultados. Esta influencia se vuelve más notable en evaluaciones con calificaciones inferiores a 60 puntos. Para estos casos, una mayor participación estudiantil disminuye aún más la calificación del profesor.

Además, se observó que los estudiantes que participan en las evaluaciones suelen utilizar el espacio de evaluación principalmente para emitir comentarios positivos, más que para expresar críticas negativas. Esto sugiere que tienden a enfocarse en aspectos positivos del curso o del docente, aprovechando el instrumento de evaluación como una oportunidad para reforzar lo que consideran importante en su proceso educativo.

Por otra parte, al analizar el instrumento de evaluación se identificó que los factores más influyentes en la nota final del docente son las cantidades de 1 y 2 estrellas. Seleccionar estas opciones tiene un impacto significativo, ya que una alta proporción de respuestas en este rango puede llevar a que la nota final del docente sea inferior a 80 puntos. Incluso, si un solo estudiante asigna 1 estrella en todos los criterios, la calificación final del docente puede descender por debajo de ese umbral.

En contraste, las evaluaciones con una alta cantidad de 4 estrellas pueden tener un efecto de refuerzo positivo, pero este solo se percibe cuando las puntuaciones más bajas son escasas. Aun así, cuando predominan las calificaciones de 1 y 2 estrellas, la presencia de algunas evaluaciones con 4 estrellas no es suficiente para evitar una nota desfavorable.

Contar con información sobre distintas características del estudiantado permitió analizar su impacto en las evaluaciones docentes y reflexionar sobre el papel que desempeñan.

Tanto las reglas de asociación como el agrupamiento demostraron cómo los estudiantes de distintos años académicos pueden tener una percepción y un enfoque distintos respecto al desempeño docente, evidenciando patrones diferenciados según su nivel de avance en la carrera.

Se identificó una relación positiva débil, entre la calificación del docente y la del grupo que evaluó, lo que indica que obtener calificaciones altas o bajas no tiene un efecto significativo sobre la evaluación del docente.

Finalmente, cabe mencionar que al tener tres períodos de tiempo disponibles en los datos, se logró detectar que en cada una de las etapas hubo características distintivas para cada una.

En 2017-2019, se valoraba la preparación académica y la actualización del docente, mientras que se criticaban aspectos como el seguimiento del aprendizaje y el uso de metodologías innovadoras. En 2020-2022, destacaron la comunicación y el acompañamiento docente, pero hubo críticas hacia la retroalimentación, planificación y ambiente educativo. En 2023, se evidenció el impacto de la pandemia: la creatividad, flexibilidad e innovación fueron tanto valoradas como cuestionadas y, por otra parte, hubieron críticas sobre el seguimiento pedagógico, lo cual se pudo haber visto afectado por los cambios en las dinámicas de enseñanza posteriores a la pandemia.

Todos estos puntos importantes nos permiten obtener un panorama general y una mejor percepción de los resultados obtenidos a través de los años, explicando el porqué de ciertos comportamientos y las relaciones que llegan a tener distintas variables entre sí.

---

### Recomendaciones

---

- Ajustar el método de cálculo de la nota de evaluación para garantizar una mayor objetividad, redistribuyendo de forma más equilibrada el peso asignado a cada categoría de estrellas, con el fin de que evitar que el sistema resulte desproporcionadamente penalizante y agregar más estrellas para que pueda haber una mayor flexibilidad en la asignación de calificaciones.
- Analizar la influencia de los cursos, departamentos académicos y sedes en las evaluaciones docentes.
- Analizar en que cursos las calificaciones son más altas y más bajas.
- Analizar los grupos de docentes con evaluaciones bajas para identificar las características que influyen en sus resultados.
- Investigar la relación puntual entre el nivel de participación estudiantil y las calificaciones otorgadas a los docentes, para esto se podría analizar con mayor precisión los grupos de notas entre 80 y 100 puntos, menores a 80 puntos y menores a 60 puntos.
- Explorar como influyen distintas características de los docentes, tal como, la cantidad de años que ha sido docente, la cantidad de años que ha sido docente en la Universidad del Valle, formación académica o experiencia previa.
- Implementar y comparar modelos de lenguaje a gran escala (LLM) y enfoques basados en léxicos para el análisis de sentimientos de las evaluaciones docentes en español, dado que en el artículo [39] se demostró que los LLM alcanzan una mayor fiabilidad en la clasificación del sentimiento al menos en inglés.

- 
- [1] E. Z. Serrano, “La evaluación del desempeño docente mediante cuestionarios en la universidad: Su legitimidad según la literatura y los requerimientos para que sea efectiva,” *Actualidades Investigativas en Educación*, vol. 21, págs. 1-23, 3 sep. de 2021. DOI: 10.15517/aie.v21i3.46221.
  - [2] B. Acosta, “EVALUACIÓN DE LAS PROPIEDADES PSICOMÉTRICAS DEL CUESTIONARIO DE EVALUACIÓN DOCENTE (DOCENTIA) PARA SU USO EN UNA MUESTRA MEXICANA,” *Revista Iberoamericana de Evaluación Educativa*, 2012, ISSN: 1989-0397. dirección: <http://www.rinace.net/riee/numeros/vol5-num2/art13.pdf>.
  - [3] S. Gupta, *Sentiment Analysis: Concept, Analysis and Applications*, ene. de 2018.
  - [4] M. Scriven, “Student Ratings Offer Useful Input to Teacher Evaluations. - Practical Assessment, Research & Evaluation,” inf. téc., 1995.
  - [5] D. M. Zimmaro, C. S. Gaede, E. J. Heikes, M.-S. P. Shim y K. G. Lewis, “Written Comments Running Head: STUDY OF WRITTEN COURSE EVALUATION COMMENTS A Study of Students’ Written Course Evaluation Comments at a Public University,” inf. téc., 2006.
  - [6] E. P. R. Piñuelas, E. C. G. Machado y P. A. B. Ramírez, “Evaluar la práctica docente a partir de los comentarios de estudiantes universitarios,” *Revista Iberoamericana de Educación / Revista Ibero-americana de Educação*, vol. 76, n.º 2, págs. 117-134, 2018.
  - [7] B. Valentín-Martínez y C. Mayor-Ruiz, “The evaluation of university professors: a look from the students,” vol. 49, 2023, ISSN: 16784634. DOI: 10.1590/S1678-4634202349241907eng.
  - [8] L. M. O. Valdés, “CALIDAD DE LA EDUCACIÓN SUPERIOR: EVALUACIÓN DEL DOCENTE UNIVERSITARIO, UNA APROXIMACIÓN A SUS DIMENSIONES Y MODELOS,” 2020. DOI: 10.22633/rpge.v24iesp2.14339.
  - [9] C. G. de Alvarez, “La Evaluación Docente como Mecanismo de Mejora en la Facultad de Ciencias Económicas de la Universidad Nacional de Asunción,” *Ciencia Latina Revista Científica Multidisciplinar*, vol. 7, págs. 5088-5100, 2 2023, ISSN: 2707-2207. DOI: 10.37811/cl\_rcm.v7i2.5708.

- [10] R. S. Rincón, “Los rasgos a evaluar en el desempeño docente. La voz de los estudiantes,” *Revista Espacio I+D, Innovación más Desarrollo*, págs. 102-118, 28 oct. de 2021, ISSN: 20076703. DOI: 10.31644/IMASD.28.2021.a05. dirección: <https://www.espacioimasd.unach.mx/index.php/Inicio/article/view/277>.
- [11] M. J. G. Garduño, “LOS PROS Y CONTRAS DEL EMPLEO DE LOS CUESTIONARIOS LOS PROS Y CONTRAS DEL EMPLEO DE LOS CUESTIONARIOS PARA EVALUAR AL DOCENTE,” *inf. téc.*, jul. de 2003.
- [12] J. F. U. Chacón, “Evaluación del desempeño de los docentes de la carrera de Pedagogía del Centro Universitario de Oriente -CUNORI-,” *Revista Científica del Sistema de Estudios de Postgrado de la Universidad de San Carlos de Guatemala*, vol. 1, págs. 46-65, 2018, ISSN: 2523-6121. DOI: 10.36958/sep.v1i01.5.
- [13] Y. Chen, “Does students’ evaluation of teaching improve teaching quality? Improvement versus the reversal effect,” *Assessment and Evaluation in Higher Education*, vol. 48, págs. 1195-1207, 8 2023. DOI: 10.1080/02602938.2023.2177252.
- [14] F. J. Tejedor, “EVALUACIÓN DEL DESEMPEÑO DOCENTE,” *Revista Iberoamericana de Evaluación Educativa*, vol. 5, págs. 319-327, 2012, ISSN: 1989-0397. dirección: [http://www.rinace.net/riee/numeros/vol5-num1\\_e/art24.pdf](http://www.rinace.net/riee/numeros/vol5-num1_e/art24.pdf).
- [15] A. D. T. González y J. L. Enríquez, “Validación de un Cuestionario de Evaluación de la Ética Profesional Docente Universitaria,” *Revista Iberoamericana de Evaluación Educativa*, vol. 14, págs. 101-114, 1 2021. DOI: 10.15366/riee2021.14.1.006.
- [16] F. T. Segura, A. M. Díaz y A. S. Moguel, “La evaluación como proceso de legitimidad: la opinión de los alumnos. Reporte de una experiencia,” págs. 7-24, 2007.
- [17] R. Espejo, V. Romo y M. L. Hervias, “La evaluación por pares a través de la observación en aula en educación superior,” *Perspectiva Educacional*, 2021. DOI: 10.4151/07189729-vol.56-iss.2-art.525.
- [18] L. L. Elizalde y R. R. Chávez, “Elementos clave para la evaluación del desempeño de los docentes,” *Revista Electrónica de Investigación Educativa*, *inf. téc.*, 2008. dirección: <http://redie.uabc.mx/NumEsp1/contenido-elizaldereyes.html>.
- [19] M. de Diego Correa y M. R. Beltrán, “La evaluación docente en educación superior: uso de instrumentos de autoevaluación, planeación y evaluación por pares,” *inf. téc.*, 2012.
- [20] F. D. B. Arceo y M. M. P. Rendón, “El portafolio docente a escrutinio: sus posibilidades y restricciones en la formación y evaluación del profesorado,” vol. 4, págs. 6-27, jun. de 2010, ISSN: 1988-5105. dirección: [www.odas.es/site/magazine.php](http://www.odas.es/site/magazine.php).
- [21] M. del Ágel Vázquez Cruz, G. C. Arroyo e Y. E. L. Barajas, “ANÁLISIS COMPARATIVO DE CRITERIOS DE DESEMPEÑO PROFESIONAL PARA LA ENSEÑANZA EN CUATRO PAÍSES DE AMERICA,” *inf. téc.*, sep. de 2014, págs. 1-20.
- [22] J. M. J. Meliá, J. González-Such y P. Sánchez-Delgado, “Factores contextuales que influyen en el desempeño docente,” *Revista Iberoamericana de Evaluación Educativa*, vol. 7, págs. 185-195, 2 2014, ISSN: 1989-0397.
- [23] M. R. Beltrán, A. C. Sánchez, Y. E. L. Barajas y E. L. Serrano, “Condiciones contextuales para el desarrollo de la práctica docente,” *Revista Iberoamericana de Evaluación Educativa*, vol. 7, págs. 171-183, 2 2014, ISSN: 1989-0397.

- [24] W. A. Z. Lascano, J. R. M. Quintanilla, M. E. A. Lagla, E. J. N. León y S. del Pilar Orozco Lata, “LA DIDÁCTICA Y SU RELACIÓN CON EL PROCESO ENSEÑANZA APRENDIZAJE Y EL DESARROLLO DEL CONOCIMIENTO,” *Ciencia Latina Revista Científica Multidisciplinar*, vol. 8, págs. 5435-5452, 6 2024, ISSN: 2707-2215. DOI: 10.37811/cl\_rcm.v8i6.15259.
- [25] G. G. Redcozub, “La enseñanza del Derecho en aulas masivas. El caso de una Facultad de Derecho argentina.” vol. 12, págs. 245-268, 3 2014, ISSN: 1887-4592.
- [26] D. J. Fernandes, C. C. M. Martínez y M. S. Sánchez, “EL PROCESO DE EVALUACIÓN DEL DESEMPEÑO DEL DOCENTE UNIVERSITARIO EN LAS UNIVERSIDADES PÚBLICAS ANGOLANAS: ACTORES INVOLUCRADOS,” 2016.
- [27] IBM, *What is exploratory data analysis (EDA)?* 2021. dirección: <https://www.ibm.com/think/topics/exploratory-data-analysis>.
- [28] P. Wulff, M. Kubsch y C. Krist, “Natural Language Processing and Large Language Models,” en *Applying Machine Learning in Science Education Research*, Springer, 2025, cap. 7, págs. 117-142. DOI: 10.1007/978-3-031-74227-9\_7.
- [29] F. L. Cruz, J. A. Troyano, B. Pontes y F. J. Ortega, “ML-SentiCon: Un lexicon multilingüe de polaridades semánticas a nivel de lemas,” vol. 53, págs. 113-120, 2014, ISSN: 1135-5948.
- [30] B. Dictionary, *Part of speech Definition & Meaning*, 2025. dirección: <https://www.britannica.com/dictionary/part-of-speech>.
- [31] O. Toporkov y R. Agerri, “On the Role of Morphological Information for Contextual Lemmatization,” *Association for Computational Linguistics*, 2024. DOI: 10.1162/coli. dirección: <https://doi.org/10.1162/coli>.
- [32] U. B. S. of Information, *What Is Machine Learning (ML)?* 2020. dirección: <https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/>.
- [33] I. Belcic y C. Stryker, *What is supervised learning?* 2024. dirección: <https://www.ibm.com/think/topics/supervised-learning>.
- [34] IBM, *What is unsupervised learning?* 2021. dirección: <https://www.ibm.com/think/topics/unsupervised-learning>.
- [35] IBM, *What is machine learning?* 2021. dirección: <https://www.ibm.com/think/topics/machine-learning>.
- [36] IBM, *What is a decision tree?* 2021. dirección: <https://www.ibm.com/think/topics/decision-trees>.
- [37] S. Kotsiantis y D. Kanellopoulos, “Association Rules Mining: A Recent Overview,” *GESTS International Transactions on Computer Science y Engineering*, inf. téc., 2006, págs. 71-82.
- [38] P.-N. Tan, M. Steinbach, A. Karpatne y V. Kumar, *Introduction to Data Mining*, 2nd. Pearson, 2018, ISBN: 9780133128901.
- [39] A. S. Morales, J. Yass, T. Jain, J. Nissen, B. V. Dusen y E. Baran, “Understanding Sentiment in User Feedback: Lexicon-Based vs. Generative AI Approaches,” en *Lecture Notes in Computer Science*, vol. 15778 LNAI, Springer Science y Business Media Deutschland GmbH, 2025, págs. 233-244, ISBN: 9783031937231. DOI: 10.1007/978-3-031-93724-8\_18.