



Escuela de Administración  
Escuela de Ingeniería, Ciencia y Tecnología

Generación de un proceso de análisis descriptivo y minería de opinión para el sistema de evaluación docente de pregrado de la Escuela de Administración de la Universidad del Rosario

Trabajo de Grado

Autores:

Viviana Catherinn Bocachica Guzman, Omar Ernesto Manta Ulloa

Bogotá, 13 de junio de 2023



Escuela de Administración  
Escuela de Ingeniería, Ciencia y Tecnología

Generación de un proceso de análisis descriptivo y minería de opinión para el sistema de evaluación docente de pregrado de la Escuela de Administración de la Universidad del Rosario

Trabajo de Grado

Autores:

Viviana Catherinn Bocachica Guzman, Omar Ernesto Manta Ulloa

Tutor:

Diego Guecha Lopez

Título para obtener:  
Magíster en Business Analytics

Maestría en Business Analytics  
Escuela de Administración  
13 de mayo de 2023


Bogotá, Colombia  
2023

## Tabla de contenido

Lista de figuras .....	6
Lista de tablas .....	7
Abreviaturas.....	8
Resumen ejecutivo.....	11
Palabras clave .....	11
Abstract.....	1
1.    Introducción .....	2
2.    Objetivos .....	4
2.1    Objetivos estratégicos.....	4
2.2    Objetivos generales.....	4
2.3    Objetivos específicos .....	4
2.4    Objetivos del proyecto .....	5
3.    Alcance del proyecto aplicado .....	6
4.    Metodología .....	8
4.1    Metodología del proyecto CRISP-DM.....	8
4.2    Comprensión del negocio: .....	8
4.3    Entendimiento de los datos:.....	11
4.4    Preparación de los datos .....	12
4.5    Modelamiento:.....	13
4.6    Evaluación: .....	14
4.7    Despliegue .....	14
5.    Cronograma.....	16
6.    Descripción de las fuentes de información .....	17
6.1    Recopilación de los datos .....	17
6.2    Descripción de los datos .....	18
6.2.1 <i>Compleitud:</i> .....	19
6.2.2 <i>Validez:</i> .....	19


6.2.3	<i>Puntualidad</i> .....	20
6.2.4	<i>Unicidad</i> .....	21
6.2.5	<i>Precisión</i> .....	21
6.2.6	<i>Consistencia</i> .....	21
6.2.7	<i>Exploración de los datos</i> .....	23
7	Resultados modelos y evaluación .....	27
7.1	Proceso y resultados del análisis cuantitativo.....	27
7.2	Análisis estadístico de texto.....	35
7.2.1	<i>Análisis de sentimientos</i> .....	35
7.2.2	<i>Clasificador Naive Bayes</i> .....	37
7.2.3	<i>Regresión logística</i> .....	49
8.	Conclusiones y recomendaciones .....	51
8.1	Conclusiones Dashboard y ETL .....	51
8.2	Conclusiones análisis estadístico de texto .....	52
8.3	Recomendaciones .....	53
	Referencias bibliográficas .....	55
	Anexos Técnicos .....	58
1.	Ejemplo cuestionario evaluación docente: .....	58
2.	Proceso del modelo Cuantitativo en R .....	61
3.	Resumen de la información para obtención de resultados .....	62
4.	Proceso del modelo Cuantitativo en R .....	62
5.	Nube de palabras .....	63
6.	Canvas modelo de negocio.....	65

- a. *Declaración de autonomía:* Incluir una hoja aparte con la siguiente declaración, firmada por el autor de tesis de maestría: “Declaramos bajo gravedad de juramento, que hemos escrito la presente tesis de maestría por nuestra propia cuenta, y que por lo tanto, su
- contenido es original. Declaramos que hemos indicado clara y precisamente todas las fuentes directas e indirectas de información, y que esta tesis de maestría no ha sido entregada a ninguna otra institución con fines de calificación o publicación”.

 / Viviana Catherinn Bocachica Guzman/ 29 junio 2023.

 / Omar Ernesto Manta Ulloa/ 29 junio 2023.

- b. *Declaración de exoneración de responsabilidad:* “Declaramos que la responsabilidad intelectual del presente trabajo es exclusivamente de sus autores. La Universidad del Rosario no se hace responsable de contenidos, opiniones o ideologías expresadas total o parcialmente en él”.

 / Viviana Catherinn Bocachica Guzman/ 29 junio 2023.

 / Omar Ernesto Manta Ulloa/ 29 junio 2023.

## Lista de figuras

<b>Figura 1</b> Metodología Crisp-DM.....	8
<b>Figura 2</b> Entendimiento de la problemática actual de la secretaría académica .....	10
<b>Figura 3</b> Propuesta planteada a la secretaría académica.....	11
<b>Figura 4</b> Estructura To-Be del proyecto .....	15
<b>Figura 5</b> Auditoria de calidad del dato .....	18
<b>Figura 6</b> Validación y resumen de cifras cuantitativas a utilizar .....	24
<b>Figura 7</b> Validación y resumen de cifras cuantitativas a utilizar .....	24
<b>Figura 8</b> Proceso de implementación y almacenamiento .....	28
<b>Figura 9</b> Proceso notebooks .....	29
<b>Figura 10</b> Proceso notebooks - python.....	30
<b>Figura 11</b> Estructura Dashboard.....	33
<b>Figura 12</b> Estructura Dashboard – Resultados generales .....	33
<b>Figura 13</b> Estructura Dashboard – Resultados detallados .....	34
<b>Figura 14</b> Estructura Dashboard – Resultados cualitativos.....	35
<b>Figura 15</b> las características Mas informativas del clasificador Naive Bayes .....	38
<b>Figura 16</b> Naive Bayes Accuracy.....	39
<b>Figura 17</b> Experto en materia .....	41
<b>Figura 18</b> Enseñanza y manejo del curso .....	43
<b>Figura 19</b> Habilidades blandas .....	45
<b>Figura 20</b> Categoría evaluación y retroalimentación .....	47
<b>Figura 21</b> Función de regresión logística .....	49
<b>Figura 22</b> Representación de la ecuación de regresión logística.....	50
<b>Figura 23</b> Calculo de la exactitud del modelo.....	51

## Lista de tablas

<b>Tabla 1</b> Diagrama de Gantt.....	16
<b>Tabla 2</b> Consistencia de las bases.....	21
<b>Tabla 3</b> Diccionario variables base cuantitativa.....	25
<b>Tabla 4</b> Diccionario variables base cualitativa.....	26
<b>Tabla 5</b> Categorías Criterio Experto.....	40

## Abreviaturas

1. **Power BI:** Herramienta de Inteligencia de negocios
2. **ETL:** Extracción Carga y Transformación de los datos (ETL).

## Glosario

- **Dashboard:** Es una interfaz de usuario, que puede presentar algo de semejanza con el panel de control de un coche, donde se organiza y se presenta la información de una manera que es fácil de leer. Este panel de control es más interactivo que el que nos puede presentar un coche, a menos que sea más moderno y esté basado en una pantalla de ordenador. (Armetrics, 2020)
- **Dato:** Un dato es la representación de una variable que puede ser cuantitativa o cualitativa que indica un valor que se le asigna a las cosas y se representa a través de una secuencia de símbolos, números o letras (Editorial Etecé, 2005).
- **Machine learning:** El Machine Learning es una disciplina del campo de la Inteligencia Artificial que, a través de algoritmos, dota a los ordenadores de la capacidad de identificar patrones en datos masivos y elaborar predicciones (análisis predictivo). Este aprendizaje permite a los computadores realizar tareas específicas de forma autónoma, es decir, sin necesidad de ser programados. (CORPORATIVA, 2022)
- **Minería de opinión:** Análisis de sentimiento (también conocido como minería de opinión) se refiere al uso de procesamiento de lenguaje natural, análisis de texto y lingüística computacional para identificar y extraer información subjetiva de los recursos. Desde el punto de vista de la minería de textos, el análisis de sentimientos es una tarea de clasificación masiva de documentos de manera automática, en función de la connotación positiva o negativa del lenguaje empleado en el documento. (Weiss, 2005)

**Regresión:** proceso estadístico para entender como una variable depende de otra variable.

**Backtesting:** es una técnica que se aplica para validar la exactitud del modelo aplicado.

## **Resumen ejecutivo**

La Universidad del Rosario cuenta con un sistema de encuestas que mide el desempeño de los docentes; En estas encuestas, a los estudiantes de pregrado se les permite dar su apreciación tanto cuantitativa como cualitativa, frente a la experiencia del periodo académico en curso. La finalidad de las encuestas es retroalimentar y generar nuevas estrategias de planeación, seguimiento, gestión y toma de decisiones académicas en pro de mejorar el servicio y garantizar los niveles de calidad académica establecidos en la universidad.

Actualmente la secretaría académica de la Escuela de Administración invierte un mayor tiempo al esperado en el análisis y procesamiento de las encuestas, y en algunos casos el análisis de opiniones (respuestas cualitativas) se hace de forma aleatoria según la necesidad del caso, impidiendo tener un contexto más amplio de la información.

Como vía de solución se propone, implementar una herramienta que cubra las necesidades de procesamiento, análisis y visualización de los datos estructurados y no estructurados, de forma más eficiente y automatizada, aplicando analítica descriptiva y prescriptiva.

### **Palabras clave**

Análisis de sentimientos, dashboard, técnicas de procesamiento de texto, evaluación docente.

## **Abstract**

The Rosario university has a system of surveys that measure the performance of teachers; in these surveys, students are allowed to give both quantitative and qualitative feedback on the experience of the current academic period. The purpose of the information received is to provide feedback and generate new strategies for planning, monitoring, management, and academic decision making to improve the service and ensure academic quality levels expected by the university standards.

Currently, the academic secretary of the Management school is taking longer than expected to analyze and process the surveys, and in some cases the analysis of opinions is done randomly according to the needs of the case, not allowing to have a broader context of the information.

As a solution, it is proposed to implement a tool that meets the needs of processing, analysis, and visualization of structured and unstructured data, in a more efficient and automated way, applying descriptive and prescriptive analytics.

### **keywords**

Sentiment analysis, dashboard, word processing techniques, teacher evaluations.

## 1. Introducción

Las Universidades en Colombia buscan mejorar sus índices de excelencia académica, con el objetivo de generar factores diferenciales frente a las demás instituciones y programas de educación superior, en búsqueda de posicionamiento, reconocimiento y mayores niveles de satisfacción. Dentro de las estrategias de mejora continua, se encuentran: la investigación, diferenciación en los planes curriculares, planes directos de acompañamiento y procesos de capacitación - calidad docente.

Colombia, a cifras del 2022 tiene más de 2,6 millones de estudiantes en programas de pregrado, representado en una cobertura del 51% sobre la población total de ciudadanos entre 17 y 22 años. Cifras que hacen fuerte énfasis en niveles de cobertura y de necesidad de calidad educativa para esta gran parte de la población del país. (Ministerio de Educación, 2020)

Un factor fundamental para incrementar la calidad educativa es el desempeño docente, los procesos de calificación y apreciación juegan un rol cada vez más importante, dado que se reconoce a los profesores como elementos claves para garantizar la calidad de los procesos y objetivos institucionales. La Universidad del Rosario busca "Impartir una sólida formación ética, humanística y científica que, unida a la investigación y a una idónea y exigente docencia, permita a esta comunidad educativa formar integralmente personas insignes y actuar en beneficio de la sociedad, con un máximo sentido de responsabilidad." (*La Universidad - Universidad del Rosario*, 2022).

La evaluación docente es una medida de calificación directa que relaciona la percepción y opinión de los estudiantes en aspectos relacionados a la materia y al docente durante el semestre académico, evaluando metodología, contenido de la materia, conocimiento del docente, bienestar estudiantil, opiniones y sugerencias.

Ante esta necesidad de procesamiento y análisis de los datos de la evaluación docente, se busca crear un proceso automatizado donde se almacenen los datos por periodo académico por medio de ETL, incluirá procesamiento, transformación y visualización (dashboard) de datos garantizando así un mayor aprovechamiento de los datos de la evaluación docente de los programas de pregrado proporcionando una fuente de toma de información veraz para la toma de decisiones.

Para desarrollar el proyecto se cuenta con información histórica de los periodos 2022-1 y 2022-2 de las encuestas de los estudiantes de pregrado de los Programa de Administración de Empresas, Programa de Administración de Negocios Internacionales, Programa de Administración en Logística y Producción, Programa de Marketing y Negocios Digitales y Pregrado de Emprendimiento; Con un porcentaje de participación de los estudiantes del 21.21% para primer semestre y del 47.86% para el segundo. El promedio de estudiantes por semestre es de 1963, en donde también se evaluaron 100 docentes y 110 asignaturas de la Escuela de administración. (Universidad del Rosario Dirección de Planeación y Efectividad Institucional, 2022).

La metodología del proyecto será enfocada a través de un método mixto ya que se cuenta con información estructurada y no estructurada, brindando a la universidad un valor agregado a través del desarrollo del proyecto empresarial, donde generará oportunidades de medición, mejoramiento, adaptación en sus prácticas y contenidos educativos alineado al objetivo estratégico de la universidad.

## 2. Objetivos

### 2.1 Objetivos estratégicos

Generar resultados analíticos automatizados y confiables en menor tiempo a través de un tablero de visualización y de un análisis estadístico de texto, el cual permitirán generar resultados que aporten a nuevas estrategias de planeación, seguimiento, gestión y toma de decisiones académicas en pro de mejorar el proceso de enseñanza.

### 2.2 Objetivos generales

Optimizar la gestión de datos para obtener el máximo valor agregado de la información estructurada y no estructurada contenida en las encuestas de evaluación docente para pregrado de la Escuela de Administración de la Universidad del Rosario, a través de la aplicación de técnicas de análisis descriptivo y estadístico de texto.

### 2.3 Objetivos específicos

- Analizar el estado de las bases de información almacenadas a través de las encuestas de calidad docente, identificando mejoras y oportunidades de optimización.
- Definir fuentes y canales de administración adecuados de los datos, para garantizar la integridad, fiabilidad y accesibilidad de estos.
- Evaluar y definir las variables más importantes, como fuentes principales del análisis descriptivo y estadístico de texto con el fin de obtener resultados precisos y significativos.

- Mejorar los entregables actuales para los comités de evaluación y seguimiento académico, mediante la creación de tableros integrados de información y generación de información estadísticos de fácil comprensión.
- Identificar y resaltar las inconsistencias existentes en las bases de datos (cualitativa y cuantitativa), generando recomendaciones para su posterior corrección o mejora.

## **2.4** **Objetivos del proyecto**

- Preparar los datos para su análisis mediante un proceso de limpieza, transformación, estandarización, garantizando calidad de los datos antes de su análisis.
- Determinar las variables claves que se usarán para el análisis del modelo cualitativo y cuantitativo; asegurando que sean validas y representativas para el objetivo del estudio.
- Contar con una ruta de almacenamiento de los datos, que permita a través de ETL, mantener la integridad, trazabilidad y disponibilidad de los datos para su uso instantáneo.
- Generar un tablero de visualización automatizado, donde se visualice la información más relevante de manera clara, coherente y actualizada.
- Generar hipótesis y establecer correlaciones basadas en los resultados de la minería de texto con el objetivo de proveer conclusiones y recomendaciones que añadan valor agregado a la información recolectada, permitiendo una visión enriquecedora para el desarrollo de nuevas estrategias académicas.

### 3. Alcance del proyecto aplicado

El proyecto busca brindar soluciones analíticas para la correcta toma de decisiones a través de herramientas descriptivas y de análisis estadístico de texto. El resultado final estará vinculado a los siguientes alcances:

- **Entregables:** los entregables del proyecto incluirán los datos procesados, transformados e integrados para su uso, los cuales serán presentados en un tablero de visualización con los resultados de análisis descriptivo, así como de las hipótesis generadas con el análisis de las correlaciones obtenidas al evaluar estadísticamente la parte cualitativa de la encuesta. La entrega también comprenderá herramientas, scripts de códigos, además de todos los desarrollos y hallazgos necesarios para la automatización de los procesos y estandarización de las fuentes de información.
- **Análisis:** En este proyecto se emplearán modelos descriptivos y análisis estadístico de texto, con el fin de generar un tablero de visualización a través de herramientas como Power BI. Además, se aplicarán procesos de estructuración de datos los cuales podrán ejecutarse mediante un script en sintaxis python.
- **Temporal:** El proyecto busca ser desarrollado durante los tres semestres académicos del programa de la maestría (2022\_I,2022\_II,2023\_III), realizando avances y entregas durante los mismos, cumpliendo con los requerimientos del programa y de la secretaría académica de la Escuela de Administración de la Universidad del Rosario, siendo el primer semestre del 2023, el tiempo de finalización y entrega de resultados finales a los entes interesados.
- **Geográfico:** El proyecto se limita a la calificación docente de la secretaría académica de la Escuela de Administración de la Universidad del Rosario y a sus cinco programas de

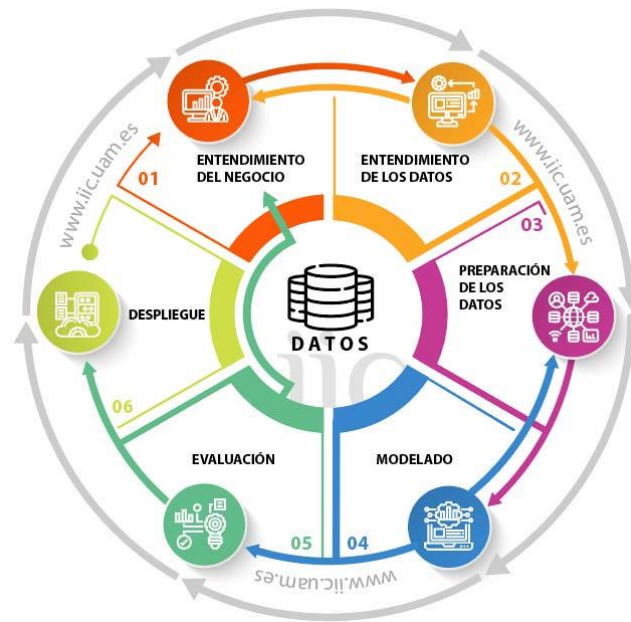
pregrado: Administración de empresas, administración de negocios internacionales, administración en logística y producción, marketing y negocios digitales y pregrado en emprendimiento. Limitando las fuentes de información solo a las encuestas de docencia que se realizan a través de la plataforma de la Universidad del Rosario.

## 4. Metodología

### 4.1 Metodología del proyecto CRISP-DM

El proyecto se desarrolló siguiendo la metodología CRISP-DM, la cual se encuentra estructurada en 6 fases, abordadas de manera secuencial, cabe resaltar que la metodología puede ser aplicada bidireccionalmente, lo que significa que se puede retroceder o volver a pasar por alguna etapa, cuando este sea considerado, ya sea por redefiniciones o necesidades del proyecto, se trabajará el proyecto acorde a estas fases, explicando cómo se realizará la aplicación a las fuentes de información, al planteamiento del problema y a la solución buscada (Manual CRISP-DM de IBM SPSS Modeler, 1994-2012).

**Figura 1** Metodología Crisp-DM



Fuente (Haya, 2021)

### 4.2 Comprensión del negocio:

Fase inicial para la comprensión de los objetivos del proyecto, en esta etapa se cuenta con las bases para la definición del problema y el plan de diseño para cumplir los objetivos.

El proyecto empresarial busca la alineación con los objetivos estratégicos de la Universidad del Rosario, se busca que el proyecto tenga un impacto positivo dentro de la universidad en pro del desarrollo de su visión y misión:


“Seremos una Universidad transformadora de ideas en innovaciones al servicio del bien común, con reconocimiento global y referente en Colombia y América latina por su oferta académica e investigación de excelencia, al igual que, por su compromiso con la construcción de país en conjunto con las regiones. Seremos una universidad sostenible que promueve experiencias generadoras de conocimiento para todos los miembros de su comunidad, que forma líderes con compromiso social y vocación humanista, internacional e intercultural que se extiende más allá de las fronteras”. (*La Universidad - Universidad del Rosario*, 2022)

La calidad docente, los procesos de calificación, seguimiento y mejora continua son fundamentales para asegurar estándares de alta calidad dentro de la universidad. La secretaría académica de la escuela de administración de la Universidad del Rosario cuenta con un proceso estándar de procesamiento de la información de encuestas de calidad docente. En dichas encuestas los estudiantes de pregrado plasman sus opiniones y percepciones en las dos opciones disponibles: bien sean de forma cuantitativa o cualitativa, frente a la experiencia del periodo académico en curso. La finalidad de la información recibida es retroalimentar y generar nuevas estrategias de planeación, seguimiento, gestión y toma de decisiones académicas en pro de mejorar el servicio y garantizar niveles de calidad académica esperados por los estándares de la universidad.

Actualmente la secretaría Académica presenta mayores tiempos a los esperados en el análisis y procesamiento de las encuestas; en algunos casos el análisis de opiniones se hace de forma aleatoria según la necesidad del caso, no permitiendo tener un contexto profundo de la información y el campo de acción es limitado a la información que se requiere. Se plantea abordar


la solución desde un punto de vista de automatización y de generación de diferentes metodologías aplicables y entendibles.

**Figura 2** Entendimiento de la problemática actual de la secretaría académica

PROBLEMATICA EMPRESARIAL 	
<b>Cientes:</b> Secretaria Académica de la Escuela de Administración de la Universidad del Rosario	Quien?
<b>Problema:</b> No se cuenta con una metodología basada en analítica de datos para sacar el mayor aprovechamiento de los datos que son recolectados en cada semestre por parte de los estudiantes.	Que?
<b>Contexto:</b> Se aplica la Evaluación Docente semestralmente, a los estudiantes de pregrado de la Escuela de Administración de la Universidad del Rosario, donde se relaciona la percepción y opinión de los mismos. Secretaria Académica, convoca a un Comité de Evaluación docente donde presenta los resultados a los directivos de la Escuela de Administración y entre todos toman decisiones, crean estrategias de mejora en pro de cumplir los objetivos estratégicos de la Universidad del Rosario.	Cuando?
<b>En que se beneficiará Secretaria Académica de la Escuela de Administración de la UR?:</b>  La universidad recibirá un valor agregado al obtener resultados de calidad por medio del modelo y su análisis automatizados y confiables en menor tiempo, logrando una acertada toma de decisiones.	Por qué? # 1?
<b>En que nos beneficiaremos los Integrantes del Proyecto Académico</b>  Lograr el título Académico de Magister en Business Analytics	Por qué? #2

Fuente: Elaboración propia

**Figura 3** Propuesta planteada a la secretaría académica

Propuesta Planteada 	
<b>Actores:</b> Estudiantes de la Maestría de Business Analytics de la Universidad del Rosario	Quién?
<b>Problema que buscan solucionar:</b> No se cuenta con una metodología basada en analítica de datos para sacar el mayor aprovechamiento de los datos que son recolectados en cada semestre por parte de los estudiantes.	Qué?
<b>Propuesta de Valor</b> Generar resultados analíticos automatizados y confiables en menor tiempo, a través de un modelo analítico cuantitativo y cualitativo, que permita generar nuevas estrategias de planeación, seguimiento, gestión y toma de decisiones académicas en pro de mejorar el servicio y garantizar niveles de calidad académica esperados por los estándares de la universidad.	Cómo?
<b>Ventajas:</b> Se Obtendrán resultados precisos, confiables y fáciles de identificar desde un punto de vista objetivo y profesional, el cual logre plantear estrategias que mejore cada vez mas los resultados de semestre a semestre.	Por qué?
<b>Riesgos:</b> Falta de comunicación por el factor tiempo. No lograr resultados que le den valor agregado a la Secretaria Académica Gestión del Cambio: Falta de claridad y que el personal a cargo, no logre comprender el funcionamiento del modelo. Escases de Recursos, para poder contratar las licencias necesarias si no se cuentan por parte de la Universidad. Los datos recolectados no tengan la consistencia necesaria para el respectivo análisis.	Porque no?
<b>Beneficio/ Impacto:</b> Reducir Tiempo que se toma en analizar los resultados y generar informes obtener precisión y alta confiabilidad de los resultados Grado de Satisfacción de los Estudiantes, evidenciándolo en los resultados obtenidos Mejor desempeño de los docentes dentro de sus metodologías de enseñanza.	Resultados Medibles

Fuente: Elaboración propia

### 4.3 Entendimiento de los datos:

Esta es una de las fases más importantes del proyecto debido a que es esencial conocer los datos recolectados al aplicar la evaluación Docente a los estudiantes; así como también identificar cualquier problema que se pueda presentar y hallar oportunidades de análisis de los datos existen para obtener el máximo valor de los resultados.

Como parte inicial, se deberá aplicar el proceso de **recopilación de los datos**, para así asegurarse que se encuentran en un formato que es de fácil acceso y análisis, identificar fuentes, proveedores y posibles fallas en el proceso de insumos. Como paso siguiente debemos incluir la realización de descripción general y **exploración inicial de los datos**, permitiendo tener una noción

macro de las características claves de tipología, calidad, consistencia y naturaleza de los datos, como, por ejemplo, la descripción de registros, variables, fuentes, tamaño, formato etc.

#### **4.4 Preparación de los datos**

Es la fase para seleccionar, limpiar y generar conjuntos de datos correctos, organizados y preparados para la fase de modelado. Es la fase crítica del proyecto de minería de datos, los errores en los datos que se pasan por alto y que no son resueltos en esta fase, evidentemente se trasladarían hasta la fase de modelado, y por consecuente generaría una reducción en la exactitud de los modelos. (IBM Preparación de datos, 2021)

Una vez culminado el entendimiento de los datos (tanto cuantitativos como cualitativos) procedemos a la preparación de la información para continuar con el modelado y con la evaluación. Se debe definir las tipologías, atributos y niveles de medición de las variables (para detalle de tipología ver anexo tipo variables) de las dos tipologías de bases.

La definición del nivel de medición es importante para realizar cálculos, porcentajes, participaciones etc. La importancia de este paso radica en la lectura de la información para su correcto uso.

La transformación es fundamental para ser inyectadas a cualquier proceso o modelo, en esta etapa también podemos abordar los valores nulos, en el caso de variables de opinión se deben retirar los valores que no sean consistente a opiniones, o que no tengan texto definido. Adicional se debe modificar la información para tener una data de opinión homogénea, bien sea mayúsculas, sin tildes, comas buscando una estandarización para poder entrenar los modelos en base a diccionarios de palabras creados.

Por otro lado, para las variables cuantitativas se busca estandarizar medidas de valor, redondeos a dos cifras etc. También se pueden categorizar variables continuas a través de procesos de

agrupación cuando se desean las variables en forma de grupos para su análisis. Para los valores perdidos se puede incurrir en imputación de valores o en el borrador del proceso.

Una vez analizada y tratada la data de los dos archivos podemos continuar con el proceso de modelaje.

#### **4.5 Modelamiento:**

En esta etapa se busca encontrar los algoritmos de modelado más apropiados al problema de análisis descriptivo de los datos y análisis de texto, Posteriormente se genera un plan de prueba, donde configuramos los valores de los parámetros que se usarán para los algoritmos de aprendizaje automático (Conocido como el Machine Learning), ya que muchos de estos pueden ser configurados para determinar las características del modelo que se generará.

Luego se construyen los modelos, ejecutando los algoritmos seleccionados sobre los datos preparados con el fin de generar uno o más modelos y calcular las métricas. Para finalizar se evalúan los resultados. (IBM corporation, 1994/2012).

Para el proceso de modelaje de análisis de sentimientos sobre la información cuantitativa, se busca el entrenamiento del modelo para el reconocimiento de sentimiento sobre algunas preguntas (principales de las encuestas docente) y poder identificar las acciones y toma decisiones. En este proceso, se crea un diccionario de palabras para poder realizar el entrenamiento de sentimientos, un conjunto de palabras puntuadas positiva y negativamente. El uso de algunos paquetes python de análisis de texto son requeridas para el manejo y uso de las palabras. Al tener estandarizado los diccionarios y las opiniones, podemos empezar a categorizar los sentimientos acordes al modelo requerido que se profundizaran más adelante.

Otra parte del modelado integra procesos de ETL sobre carga de información y bases históricas conjuntas a través de python y Power BI. En esta etapa es importante generar también el modelado

para generar el procesamiento de datos numéricos. Librerías como dplayer serán utilizadas para la utilización y generación de cubos de información para poder ser utilizado en Power BI.

#### **4.6 Evaluación:**

Para esta fase, se evalúa el desempeño de los modelos, teniendo en cuenta el cumplimiento de los criterios de éxito del problema, debe considerarse que la fiabilidad calculada para el modelo se aplica solamente para los datos sobre los que se realizó el análisis. Es preciso revisar el proceso, teniendo en cuenta los resultados obtenidos, para poder repetir algún paso anterior, en el que se pueda haber cometido algún error. (IBM corporation, 1994/2012)

La evaluación de los resultados (Backtesting) tanto de conectividad y proceso, en métricas de ETL como de los resultados descriptivos de texto (pruebas a través de regresiones logísticas) tienen que constatar que el proceso o los resultados obtenidos muestren consistencia, bajo métricas medibles como procesos de flujo o de resultados como el de exactitud.

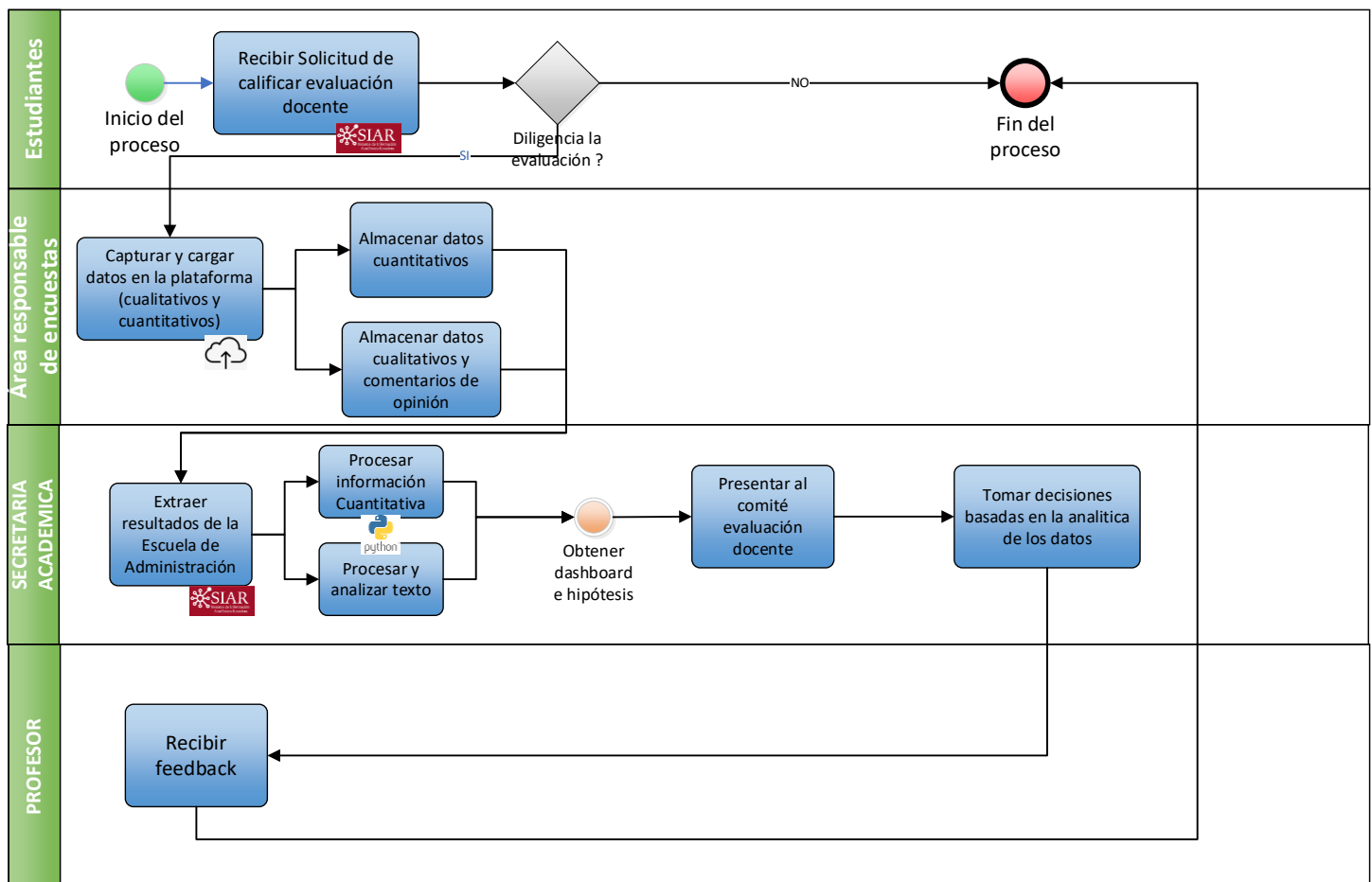
Antes de ser entregado a la Secretaría Académica, se debe validar que en la fase de prueba no se identificaron errores, y si estos ocurren, se debe realizar los debidos ajustes, y actualizaciones de los mismos, ya que el objetivo de esta fase es solucionar los fallos encontrados en la fase anterior; y asegurarse de que todos los requisitos han sido cumplidos adecuadamente, si la secretaria Académica, llegase a modificar el alcance de los resultados, será necesario crear un control de cambios que permita ajustar y lograr los resultados esperados del proyecto.

#### **4.7 Despliegue**

Una vez realizados los modelos y procesos de visualización requeridos tanto con información cualitativa como cuantitativa, y dado el aval del cliente sobre los entregables se continua con la etapa de implementación. Para esta fase, una vez se tienen los resultados deseados en herramientas, visualización, análisis de texto u otras fuentes de salida y de resultado, se procede a implementar su despliegue para que el usuario final pueda acceder al proceso y entregables finales.

Generalmente un proyecto de minería de datos no concluye en la implantación del modelo, ya que se deben documentar y presentar los resultados de manera comprensible para el usuario con el objetivo de lograr un incremento del conocimiento. Por otra parte, en la fase de explotación se debe asegurar el mantenimiento de la aplicación y la posible difusión de los resultados (IBM corporation, 1994/2012).

**Figura 4** Estructura To-Be del proyecto



Fuente: Elaboración propia



5. Validación fuentes de almacenamiento
6. Validación de los datos y almacenamiento en SharePoint
7. Reestructuración varios.

### **Modelo:**

1. Diseñar la versión 1 modelo Cualitativo (minería de Opinión)
2. Diseñar la versión 1 Modelo Cuantitativo
3. Diseño Dashboard versión 1
4. Diseñar la versión 2 Modelo Cualitativo (minería de opinión)
5. Resultados análisis estadístico de texto
6. Diseño Dashboard versión 2.

### **Evaluación**

1. Revisión y retroalimentación Modelo Cuantitativo y Dashboard Cuantitativo versión 1
2. Revisión y retroalimentación Resultados análisis estadístico de texto
3. Definición versión Final Modelo y Dashboard.

### **Despliegue y entregables**

1. Entregable dashboard descriptivo y cualitativo
2. Desarrollo entregables código y ejecutores automatizados
3. Revisión y otros adicionales varios
4. Entrega oficial a secretaria académica de documentación y Modelo Cuantitativo y Dashboard y análisis estadístico de texto.

## **6 Descripción de las fuentes de información**

### **6.1 Recopilación de los datos**

Los datos actualmente son recolectados por medio de una plataforma externa contratada por la Universidad del Rosario, los estudiantes reciben un correo electrónico para seguir un enlace que los llevara hacia la respectiva plataforma, continuando con el diligenciamiento de la evaluación docente unas semanas antes de finalizar el semestre. Los datos son procesados a través de la plataforma SIAR, donde los usuarios de cada área interesada cuentan con permisos acorde a su

necesidad, para así extraer la información en el formato .csv o xlsx permitiendo filtrar por los periodos requeridos.

Las fuentes principales de información son las bases de datos suministradas por secretaría académica en pregrado. De una muestra de datos exploratorios se realizó análisis preliminar de estructura, calidad y cantidad de variables disponibles para el análisis de la información. Las fuentes están discriminadas entre la calificación de resultados cuantitativos y otra base con las opiniones y comentarios de la parte cualitativa de la encuesta.

## 6.2 Descripción de los datos

Este proceso se realizará bajo la metodología de Análisis de dimensiones de la calidad de los datos, aplicando este proceso para las dos bases suministradas del proyecto empresarial:

**Figura 5** Auditoria de calidad del dato



Fuente: (Ruiz Ayuso, 2019)

Se debe considerar que para los cálculos y análisis de calidad se renombraron las variables de las bases de datos por nombres más cortos, el cual permitirá un adecuado y más fácil disposición y manejo de la información, adjuntamos tablas con nombres y nombres codificados (ver adjuntos). Se muestran igualmente resultados y códigos realizados en el proceso a través de python y Excel.

### **6.2.1 Completitud:**

Se realiza la validación de porcentaje de valores nulos tanto en las variables requeridas, como en las no requeridas. Entendiendo valores nulos como registros NA. La completitud a nivel de registros se realiza sobre las primeras 12 variables (variables requeridas) y a nivel de variable sobre la totalidad de las dos bases.

#### **✓ Base cuantitativa :**

Acorde a los resultados tenemos que para la base se identifican 14 variables con total completitud al 100% y 15 con 84% de completitud. A nivel de registros tenemos un 100 por ciento de completitud.

#### **✓ Base Cualitativa:**

Para la base cualitativa presentamos valores altos de nulos. Principalmente en las variables canalizadas de opinión y texto de los estudiantes. Sin embargo, a nivel de registro tenemos toda la información requerida y con completitud del 100%.

### **6.2.2 Validez:**

Para la verificación de validez se implementaron una serie de reglas a nivel de variable, principalmente en los campos que son autocompletados por la plataforma externa que contrata la universidad. Estos campos están limitados a la información de la escuela de administración, materia, profesor, código materia, etc. Variables con campos y resultados limitados.

Para la información de valores las reglas son porcentajes de calificación de 0 a 5 por lo cual serán permitidos valores que entren dentro de este rango como calificaciones promedio resultantes del procesamiento de la plataforma de la universidad. a continuación, se muestran algunas reglas de validación para esta información.

La **validación** del registro se ejecuta sobre las primeras 12 variables en la base de datos cuantitativa y 9 variables en la base de datos cualitativa.

✓ **Base cuantitativa:**

Se obtiene un resultado similar al de completitud, los valores están dentro de los esperados para el campo de la variable correspondiente.

✓ **Base Cualitativa:**

para la validación cualitativa se verifican reglas adicionales en las variables de texto, que al menos contengan dos letras por lo menos para hacer valido el registro de la variable. En este caso si tenemos diferencias y porcentajes menores de valides en las variables no requeridas, por los campos de texto.

### ***6.2.3 Puntualidad***

Las encuestas de evaluación docente, que se aplican a los estudiantes de la Escuela de administración de la Universidad del Rosario, se realizan por medio de un software externo a la universidad, es por esto por lo que, para nosotros no es posible medir el tiempo máximo esperado para acceder a la información solicitada. Sin embargo, secretaría Académica definió, que las evaluaciones se deben presentar cuando la clase ya tiene un 60% de avance; También secretaría Académica debe asegurarse que la mayor cantidad de estudiantes por curso, puedan responder la evaluación docente, a través de mensajes recordatorios por correo electrónico, visitas por parte de secretaría académica a los salones o solicitar al profesor asignado a la materia el recordatorio de esta a sus estudiantes. y con esto poder tener la información a tiempo generar reportes y realizar los respectivos análisis.

Es importante resaltar que el software procesa automáticamente la información y la sube para permitir el acceso de los interesados.

#### 6.2.4 Unicidad

Se realiza validación en las dos bases acorde a los valores requeridos para garantizar que el registro contenga un único valor. Variables como facultad, docente, cédula y materia son las variables con las cuales

Dentro de los resultados tenemos que en ambas bases se respeta la unicidad de registros, es decir de la base cuantitativa tenemos 341 registros únicos y de la cualitativa 5481.

#### 6.2.5 Precisión

La base de datos correspondiente a los resultados de la evaluación docente no cuenta con información relevante para definir o identificar una variable de negocio condicional, en la cual permita la toma de decisiones de manera acertada, en este caso las calificaciones son más los resultados agrupados por criterios de evaluación.

#### 6.2.6 Consistencia

**Tabla 2** Consistencia de las bases

ID_ASIGNATURA	CUANTITATIVA1 ASIGNATURA	CUALITATIVA1 ASIGNATURA	ID_PROFESOR	CUANTITATIVA1 Nombre_Profesor	CUALITATIVA1 Nombre_Profesor
10810002	CONTABILIDAD PARA LA TOMA DE DECISIONES	CONTABILIDAD PARA LA TOMA DE DECISIONES	7965485	RODRIGUEZ CAÑAS GONZALO ANDRES	RODRIGUEZ CAÑAS GONZALO ANDRES
10810003	MARKETING	MARKETING	79643160	RODRIGO PINZON OSCAR JAVIER	RODRIGO PINZON OSCAR JAVIER
10810009	GERENCIA FINANCIERA	GERENCIA FINANCIERA	79532648	ORTEGON ROJAS JORGE ARMANDO	ORTEGON ROJAS JORGE ARMANDO
15210001	ESTRATEGIA DE EMPRESA I	ESTRATEGIA DE EMPRESA I	75034308	RIVERA RODRIGUEZ HUGO ALBERTO	RIVERA RODRIGUEZ HUGO ALBERTO
15210003	CONTABILIDAD FINANCIERA	CONTABILIDAD FINANCIERA	194415	CASTILIBLANCO OMAR EDUARDO	CASTILIBLANCO OMAR EDUARDO
15210004	SIMULACION GERENCIAL	SIMULACION GERENCIAL	7988887	GIRALDO VELEZ ALAN DAVID	GIRALDO VELEZ ALAN DAVID
15210005	SISTEMAS DE INFORMACIÓN ADMINISTRATIVOS	SISTEMAS DE INFORMACIÓN ADMINISTRATIVOS	80051309	ALDANA MIRANDA DANIEL ORLANDO	ALDANA MIRANDA DANIEL ORLANDO
15210006	MÉTODOS DE INVESTIGACIÓN CIENTÍFICA	MÉTODOS DE INVESTIGACIÓN CIENTÍFICA	6758056	GONZALEZ RODRIGUEZ JAVIER LEONARDO	GONZALEZ RODRIGUEZ JAVIER LEONARDO
15210007	PRODUCTIVIDAD	PRODUCTIVIDAD	5220740	PARDÓ MARTINEZ CLARA INÉS	PARDÓ MARTINEZ CLARA INÉS
15210008	PENSAMIENTO ESTRATÉGICO	PENSAMIENTO ESTRATÉGICO	75034308	RIVERA RODRIGUEZ HUGO ALBERTO	RIVERA RODRIGUEZ HUGO ALBERTO
15210010	MEJORAMIENTO EN AMBIENTE T.O.C.	MEJORAMIENTO EN AMBIENTE T.O.C.	19158292	RODRIGUEZ ACOSTA LUIS OBDULIO	RODRIGUEZ ACOSTA LUIS OBDULIO
15210015	MATEMÁTICAS FINANCIERAS	MATEMÁTICAS FINANCIERAS	51662363	HERNANDEZ NINO FANNY CONSUELO	HERNANDEZ NINO FANNY CONSUELO
15210025	PLANEACIÓN ESTRATÉGICA POR ESCENARIOS	PLANEACIÓN ESTRATÉGICA POR ESCENARIOS	75034308	RIVERA RODRIGUEZ HUGO ALBERTO	RIVERA RODRIGUEZ HUGO ALBERTO
15210030	LOGÍSTICA EMPRESARIAL	LOGÍSTICA EMPRESARIAL	79983850	MEJIA DELGADO OSCAR ARMANDO	MEJIA DELGADO OSCAR ARMANDO
15210033	COMPORTAMIENTO DEL CONSUMIDOR	COMPORTAMIENTO DEL CONSUMIDOR	52030353	LOMBANA GONZALEZ MARIA EMMA	LOMBANA GONZALEZ MARIA EMMA
15210049	DESARROLLO EMPRESARIAL COLOMBIANO	DESARROLLO EMPRESARIAL COLOMBIANO	79131343	PEREZ GOMEZ CARLOS HERNAN	PEREZ GOMEZ CARLOS HERNAN
15210055	SEMINARIO PRACTICA PROFESIONAL	SEMINARIO PRACTICA PROFESIONAL	5296748	TULA LAVERDE CAROLINA	TULA LAVERDE CAROLINA
15210056	SEMINARIO DE TRABAJO DE GRADO	SEMINARIO DE TRABAJO DE GRADO	51808946	CONTRERAS TORRES FRANCOISE VENEZIA	CONTRERAS TORRES FRANCOISE VENEZIA
15210060	TEORÍA DE JUEGOS	TEORÍA DE JUEGOS	332154	RIVES ORTIZ GIOVANNI EFRAIN	RIVES ORTIZ GIOVANNI EFRAIN
15210079	MERCADO DE CAPITALES Y ADMINISTRACIÓN DE PORTAFOLIO	MERCADO DE CAPITALES Y ADMINISTRACIÓN DE PORTAFOLIO	79732484	LEON VILLANUEVA ALEXANDER	LEON VILLANUEVA ALEXANDER
15210098	FUNDAMENTOS DE MERCADERO	FUNDAMENTOS DE MERCADERO	79933683	URIBE AREVALO ABEL EDUARDO	URIBE AREVALO ABEL EDUARDO
15210138	GOOGLE ADS & SOCIAL MEDIA	GOOGLE ADS & SOCIAL MEDIA	1053792008	BETANCUR RAMIREZ EDWIN	BETANCUR RAMIREZ EDWIN
15210142	EXCEL BÁSICO	EXCEL BÁSICO	19434383	SANCHEZ BARRERO JUAN MANUEL	SANCHEZ BARRERO JUAN MANUEL
15210150	INVESTMENT BANKING	INVESTMENT BANKING	1020501817	CASTRO TORRES WILSON JAVIER	CASTRO TORRES WILSON JAVIER
15210171	DIGITAL TRANSFORMATION	DIGITAL TRANSFORMATION	80203811	ROA GONZALEZ EFRAIN ANDRES	ROA GONZALEZ EFRAIN ANDRES
15210177	MERCADO DE OPCIONES Y DERIVADOS	MERCADO DE OPCIONES Y DERIVADOS	80871644	MACIAS VARGAS DIEGO FERNANDO	MACIAS VARGAS DIEGO FERNANDO
15210183	SEGMENTACIÓN E INTELIGENCIA DE MERCADOS B2C Y B2B	SEGMENTACIÓN E INTELIGENCIA DE MERCADOS B2C Y B2B	32256520	AMARDO MATEUS MARLEIF	AMARDO MATEUS MARLEIF
15210185	MARKETING EMOCIONAL	MARKETING EMOCIONAL	51769094	ARIAS LEVIN GLORIA ISABEL	ARIAS LEVIN GLORIA ISABEL
15220001	ESTUDIOS ORGANIZACIONALES	ESTUDIOS ORGANIZACIONALES	79460536	ESGUERRA PEREZ GUSTAVO ADOLFO	ESGUERRA PEREZ GUSTAVO ADOLFO
15220002	COSTOS Y PRESUPUESTOS	COSTOS Y PRESUPUESTOS	79747473	TINOCO BERNAL CAMILO ERNESTO	TINOCO BERNAL CAMILO ERNESTO
16910021	TRABAJO DE GRADO (TNG)	TRABAJO DE GRADO (TNG)	79511975	CRUZ BERNAL HERNAN ALBERTO	CRUZ BERNAL HERNAN ALBERTO
84210029	INTRODUCCIÓN A SUPPLY CHAIN MANAGEMENT	INTRODUCCIÓN A SUPPLY CHAIN MANAGEMENT	79983850	MEJIA DELGADO OSCAR ARMANDO	MEJIA DELGADO OSCAR ARMANDO
84210081	GERENCIA DE CALIDAD	GERENCIA DE CALIDAD	52585533	MARTINEZ SANCHEZ PALOMA MARIA TERESA	MARTINEZ SANCHEZ PALOMA MARIA TERESA
84210086	SISTEMAS DE INFORMACIÓN EMPRESARIAL	SISTEMAS DE INFORMACIÓN EMPRESARIAL	79467783	SANTOS CASTELLANOS WEIMAR	SANTOS CASTELLANOS WEIMAR
10810008	LEADERSHIP AND SUSTAINABILITY	LEADERSHIP AND SUSTAINABILITY	79466560	ESCORBAR SALAS MAURICIO EDUARDO	ESCORBAR SALAS MAURICIO EDUARDO
10810018	MARKETING PLAN	MARKETING PLAN	19222348	MANGILLA MEJIA CARLOS ALFREDO	MANGILLA MEJIA CARLOS ALFREDO
15210002	ESTRATEGIA DE EMPRESA II	ESTRATEGIA DE EMPRESA II	19058448	VILLACHES CARDENAS GUILLERMO	VILLACHES CARDENAS GUILLERMO
15210053	GERENCIA DE LOS NEGOCIOS INTERNACIONALES	GERENCIA DE LOS NEGOCIOS INTERNACIONALES	79498976	BUTRAGO RUBIANO RICARDO ERNESTO	BUTRAGO RUBIANO RICARDO ERNESTO
15210058	COMERCIO INTERNACIONAL	COMERCIO INTERNACIONAL	38142655	PULIDO LOPEZ ALEJANDRA	PULIDO LOPEZ ALEJANDRA
15210100	DECISIONES DE MERCADERO	DECISIONES DE MERCADERO	19222348	MANGILLA MEJIA CARLOS ALFREDO	MANGILLA MEJIA CARLOS ALFREDO

Fuente: elaboración propia

Se definió que las dos llaves primarias dentro de la bases cualitativa y cuantitativa son:

- ID\_Asignatura
- ID\_Profesor

Para determinar la consistencia de las llaves primarias se compararon tanto los IDs de la base cuantitativa como la cualitativa y se determinó que las dos bases son consistentes ya que la calidad de los datos coherente, tanto en la digitación y orden de estos.

Tras el análisis inicial de calidad de datos realizado en las dos bases tanto cuantitativa como cualitativa, suministradas para la implementación del proyecto, se concluye lo siguiente:

Base de datos cuantitativa cumple con los con los requerimientos mínimos de calidad, es decir, la información puede ser trabajada y tratada al obtener el visto bueno de dimensiones en cuanto a completitud, validez, unicidad y consistencia; la base de datos presenta las condiciones requeridas en estos aspectos, salvaguardando algunos campos no requeridos como principales en completitud y veracidad. Las dimensiones de Precisión y puntualidad no son medibles en este caso, puesto que la información es suministrada a través de una plataforma de la universidad y esta es la que efectúa estas validaciones. Al ser usuario final de la información ya tenemos ciertas restricciones en cuanto a permisos de la secretaría académica de la Escuela de Administración de la Universidad del Rosario y permisos de extracción y validaciones de cargues.

Por otro lado, se debe tener claro que las bases de datos cualitativas cuentan con campos de opinión que presentan gran porcentaje de nulos en campos de texto, dado los factores de no diligenciamiento obligatorio de parte de los estudiantes; también se debe considerar que desde un principio esta encuesta fue implementada sin una asesoría en el correcto uso de la analítica de datos y otras indicaciones claves para el análisis de sentimientos. Es por esto que se debe analizar muy

bien los tipos de preguntas y sus respectivos resultados, validar si cumplen con los requisitos mínimos para poder aplicar un análisis de texto. Es importante tener en cuenta que la incorrecta recolección de los datos generara inconvenientes que podrían ser un impedimento para el análisis requerido de la información y de los análisis de sentimientos esperados.

### **6.2.7 Exploración de los datos**

Se obtienen una muestra de datos exploratorios, con los cuales se hace un análisis preliminar de estructura, calidad y cantidad de variables disponibles para el análisis de la información. Se tiene dos bases suministradas por la secretaría Académica de la Escuela de Administración, de los periodos 2022-1 y 2022-02, discriminada entre la calificación de resultados cuantitativos y otra base con las opiniones y comentarios de la parte cualitativa de la encuesta.

#### **✓ Tipos de datos y tamaño de la base**

Informe estadístico base cuantitativa

Base agrupada para el semestre 2022-01 y 2022-02 con resultados y promedios, cuenta con 696 registros y 27 variables, de las cuales 11 variables son de carácter nominal y sirven para identificar programa académico, semestre, idioma, e identificación del profesor, 2 variables de agrupación de cantidad de estudiantes que realizaron las pruebas, finalmente los promedios resultantes de las encuestas para agrupaciones vinculadas (grupos de preguntas) y promedio de puntaje de las preguntas individuales. Para ver la totalidad de las variables y tipología remitirse a la tabla 7 y 8.

**Figura 6** Validación y resumen de cifras cuantitativas a utilizar

var	count	mean	std	min	25%	50%	75%	max
total_est_curso	354	15,7	9,6	1	9,00	14,00	25,00	33
total_est_rta	354	7,5	5,8	0	3,00	6,00	11,00	28
participacion_est	354	0,5	0,3	0	0,30	0,45	0,64	1
Promedio_general	354	4,2	1,3	0	4,22	4,69	4,88	5
promedio_dim1	354	4,1	1,3	0	4,05	4,61	4,87	5
promedio_dim2	354	4,1	1,3	0	4,08	4,62	4,83	5
promedio_dim3	354	4,2	1,3	0	4,29	4,75	4,95	5
promedio_dim4	354	4,4	1,3	0	4,50	4,83	5,00	5
promedio_p1	354	4,1	1,3	0	4,00	4,50	4,90	5
promedio_p2	354	4,2	1,3	0	4,12	4,63	4,92	5
promedio_p3	354	4,2	1,3	0	4,09	4,67	4,93	5
promedio_p4	354	4,2	1,3	0	4,25	4,71	5,00	5
promedio_p5	354	4,0	1,3	0	3,89	4,50	4,83	5
promedio_p6	354	4,2	1,3	0	4,19	4,75	5,00	5
promedio_p7	354	4,2	1,3	0	4,10	4,67	5,00	5
promedio_p8	354	4,3	1,3	0	4,44	4,82	5,00	5
promedio_p9	354	4,4	1,3	0	4,57	4,87	5,00	5
promedio_p10	354	4,3	1,3	0	4,50	4,82	5,00	5

Fuente: elaboración propia

**Figura 7** Validación y resumen de cifras cuantitativas a utilizar

var	describe
total_est_curso	nro estudiantes en el curso
total_est_rta	nro estudiantes que respondieron
participacion_est	participacion
Promedio_general	promedio general
promedio_dim1	promedio dimensión 1 orientación del aprendizaje preguntas 1 2 y 3
promedio_dim2	promedio dimensión 2 evaluación y retroalimentación preguntas 4 y 5
promedio_dim3	promedio dimensión 3 ambiente de aprendizaje preguntas 6 7 y 8
promedio_dim4	promedio dimensión 4 responsabilidad y cumplimiento preguntas 9 y 10
promedio_p1	promedio p1 aplico metodologías y estrategias que me motivaron a aprender y que me retan a seguir aprendiendo y investigando
promedio_p2	promedio p2 promovió mi pensamiento crítico analítico o creativo
promedio_p3	promedio p3 brindo espacios para aclarar dudas o ampliar temas
promedio_p4	promedio p4 aplico evaluaciones y desarrollo actividades que contribuyeron a mi proceso de aprendizaje
promedio_p5	promedio p5 brindo una retroalimentación útil y oportuna que me permitió identificar mis aciertos y mis errores
promedio_p6	promedio p6 promovió un ambiente de aprendizaje basado en la confianza y la tolerancia hacia los demás
promedio_p7	promedio p7 demostró un interés por atender el bienestar de los estudiantes durante las actividades académicas
promedio_p8	promedio p8 el ambiente del curso se caracterizo por el respeto y la cordialidad
promedio_p9	promedio p9 asistió puntualmente a las clases y aprovecho el tiempo para desarrollar las actividades programadas
promedio_p10	promedio p10 desarrollo los temas o actividades para el curso de acuerdo con los objetivos previstos

Fuente: elaboración propia

✓ **Informe estadístico Base cuantitativo**

**Tabla 3** *Diccionario variables base cuantitativa*

<b>Variab</b> les	<b>Descripción</b>	<b>Tipología</b>	<b>Razón</b>
Periodo	fecha de calificación docente	cuantitativa	intervalo
Unidad Académica	facultad	cuantitativa	nominal
Área	plan de estudios pregrado	cuantitativa	nominal
Idioma	idioma habilitado de pregrado	cuantitativa	nominal
Id asignatura	Código de la asignatura	cuantitativa	nominal
Asignatura	nombre de la asignatura	cuantitativa	nominal
Estrategia	estrategia perteneciente al plan académico	cuantitativa	nominal
Grupo	semestre académico del programa de pregrado	cuantitativa	nominal
Actividad	Actividad asignada al programa de pregrado	cuantitativa	nominal
Id_profesor	Numero Id profesor	cuantitativa	nominal
Profesor	Nombre profesor	cuantitativa	nominal
Nro. estudiantes en el curso	cantidad de estudiantes por el curso	cuantitativa	nominal
Nro. estudiantes que respondieron	cantidad de estudiantes que respondieron el cuestionario	cuantitativa	razón
Participación	% participación por grupo	cuantitativa	razón
Promedio general	promedio general de las 10 preguntas	cuantitativa	razón
Promedio dimensión 1 - Orientación del aprendizaje (preguntas 1,2 y3)	promedio puntuación preguntas 1,2,3 agrupadas	cuantitativa	razón
Promedio dimensión 2 - Evaluación y retroalimentación (preguntas 4 y 5)	promedio puntuación preguntas 4,5 agrupadas	cuantitativa	razón
Promedio dimensión 3 - Ambiente de aprendizaje (preguntas 6,7 y 8)	promedio puntuación preguntas 6,7,8 agrupadas	cuantitativa	razón
Promedio dimensión 4 - Responsabilidad y cumplimiento (preguntas 9 y 10)	promedio puntuación preguntas 9,10 agrupadas	cuantitativa	razón
Promedio P1 - Aplicó metodologías y estrategias que me motivaron a aprender y que me retan a seguir aprendiendo / investigando.	promedio pregunta 1	cuantitativa	razón
Promedio P2 - Promovió mi pensamiento crítico, analítico o creativo.	promedio pregunta 2	cuantitativa	razón
Promedio P3 - Brindó espacios para aclarar dudas o ampliar temas.	promedio pregunta 3	cuantitativa	razón
Promedio P4 - Aplicó evaluaciones y desarrolló actividades que contribuyeron a mi proceso de aprendizaje.	promedio pregunta 4	cuantitativa	razón
Promedio P5 - Brindó una retroalimentación útil y oportuna que me permitió identificar mis aciertos y mis errores.	promedio pregunta 5	cuantitativa	razón
Promedio P6 - Promovió un ambiente de aprendizaje basado en la confianza y la tolerancia hacia los demás.	promedio pregunta 6	cuantitativa	razón
Promedio P7 - Demostró un interés por atender el bienestar de los estudiantes durante las actividades académicas.	promedio pregunta 7	cuantitativa	razón
Promedio P8 - El ambiente del curso se caracterizó por el respeto y la cordialidad.	promedio pregunta 8	cuantitativa	razón
Promedio P9 - Asistió puntualmente a las clases y aprovechó el tiempo para desarrollar las actividades programadas.	promedio pregunta 9	cuantitativa	razón

Fuente: elaboración propia

### ✓ Informe estadístico base cualitativa

Base individualizada con 11054 registros y 34 variables, de las cuales 20 variables son de tipo texto, con los resultados de opiniones y comentarios de los estudiantes referentes a los programas y profesores. Las otras variables al igual que en el caso anterior, presentan la información de programa, facultad, materia e identificación del profesor. Para ver la totalidad de las variables y tipología remitirse a la tabla N 8.

**Tabla 4** *Diccionario variables base cualitativa*

Variables	Descripción	Tipología	Razón
Periodo	fecha de calificación docente	cuantitativa	intervalo
Unidad	facultad	cualitativa	nominal
Área	plan de estudios pregrado	cualitativa	nominal
Id_asignatura	Código de la asignatura	cualitativa	nominal
ESTRATEGIA	nombre de la asignatura	cualitativa	nominal
Asignatura	estrategia perteneciente al plan académico	cualitativa	nominal
Grupo	semestre académico del programa de pregrado	cualitativa	nominal
Actividad	Actividad asignada al programa de pregrado	cualitativa	nominal
Id_profesor	Id profesor	cualitativa	nominal
Profesor	nombre profesor	cualitativa	nominal
Justifica la contribución del diálogo formativo	pregunta cualitativa de opinión	cualitativa	texto
¿Qué aspectos destacas de tu profesor(a)?	pregunta cualitativa de opinión	cualitativa	texto
¿Qué aspectos consideras que podría mejorar tu profesor(a)?	pregunta cualitativa de opinión	cualitativa	texto
Reflexiones imagen 1	comentarios asociados a la imagen Numero 1	cualitativa	texto
Reflexiones imagen 2	comentarios asociados a la imagen Numero 2	cualitativa	texto
Reflexiones imagen 3	comentarios asociados a la imagen Numero 3	cualitativa	texto
Reflexiones imagen 4	comentarios asociados a la imagen Numero 4	cualitativa	texto
Reflexiones imagen 5	comentarios asociados a la imagen Numero 5	cualitativa	texto
Reflexiones imagen 6	comentarios asociados a la imagen Numero 6	cualitativa	texto
Reflexiones imagen 7	comentarios asociados a la imagen Numero 7	cualitativa	texto
Reflexiones imagen 8	comentarios asociados a la imagen Numero 8	cualitativa	texto
Reflexiones imagen 9	comentarios asociados a la imagen Numero 9	cualitativa	texto
Reflexiones imagen 10	comentarios asociados a la imagen Numero 10	cualitativa	texto
Reflexiones imagen 11	comentarios asociados a la imagen Numero 11	cualitativa	texto
Otra ¿Cuál?	pregunta cualitativa de opinión	cualitativa	texto
Reflexiones imagen otra	pregunta cualitativa de opinión	cualitativa	texto
¿Cómo la(s) cualidad(es) / valor(es) elegido(s) fueron aplicada(s) por el (la) profesor(a) en su proceso de enseñanza?	pregunta cualitativa de opinión	cualitativa	texto
justifique_cuanto_aprendio	pregunta cualitativa de opinión	cualitativa	texto
Reflexiones estrategia hyflex	pregunta cualitativa de opinión	cualitativa	texto
Otras reflexiones	pregunta cualitativa de opinión	cualitativa	texto

Fuente: elaboración propia

## 7 Resultados modelos y evaluación

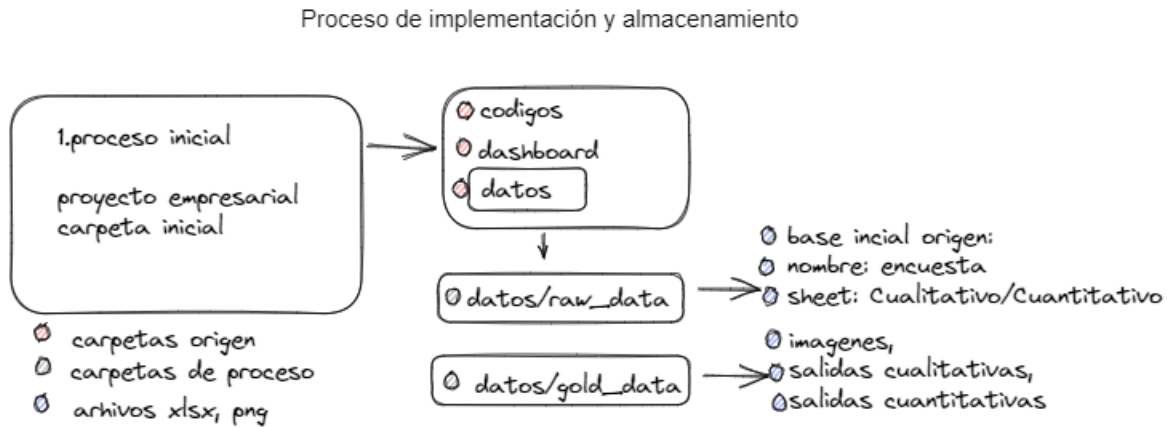
### 7.1 Proceso y resultados del análisis cuantitativo

“ETL busca simplificar la gestión, procesamiento y análisis de la información; permitiendo aplicar criterios de calidad y consistencia en un solo almacenamiento; mejora la productividad al facilitar la lectura de los datos, optimizando procesos, garantizando accesibilidad a los datos con estándares requeridos de la entidad. ETL contiene tres procesos que son:

- ✓ Extracción: determina las fuentes, orígenes de datos a procesar, durante este proceso se realiza un análisis de limpieza y diferenciación de la información. Los datos pueden provenir de distintas fuentes XML, JSON, CSV, TXT, SQL, SAS, API, sitios web entre otros.
- ✓ Transformación: se transforma, corrigen y modifican los datos para su mejor uso. Suelen usarse procesos de eliminación, filtrado, cambio de formato requeridos para obtener el dato como se requiere y como mejor se pueda usar dentro de cualquier proceso.
- ✓ Carga: después de la extracción y transformación se requiere la carga de los datos en bases destino, se usa la nube o físicamente en una instalación o disco local” (Santos, 2023).

Al aplicar un proceso de ETL en el proyecto, se busca tener facilidad de acceso y de fuentes, la implementación se hará en unidad local, con la posibilidad de transferir a un drive de ser requerido por la secretaría académica. Se cubren los procesos de extracción, transformación y cargue dentro de carpetas respetando la naturaleza y orden de los procesos, adicional se incluirán dentro de la misma, notebook con código fuente, dashboard de visualización y carpetas de origen y de distribución de la data.

**Figura 8** *Proceso de implementación y almacenamiento*

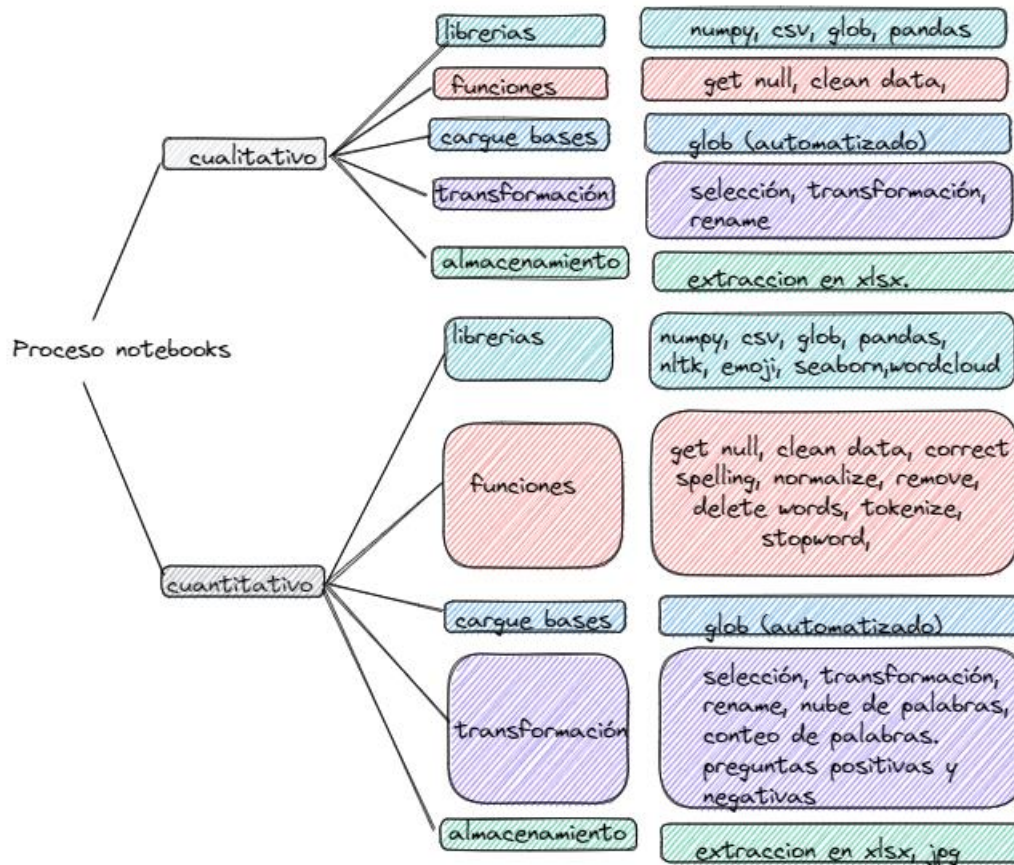


Fuente: elaboración propia

El proceso de almacenamiento inicia desde la extracción de insumos (base cuantitativa y cualitativa encuestas docentes de pregrado) para el semestre. El requerimiento de estructura inicial del archivo debe incluir: documento en xlsx que inicie con el nombre “encuesta” y por las hojas de información tanto “Cualitativas” como “Cuantitativas” respectivamente para que el proceso de cargue de información detecte las fuentes automáticamente y se realice adecuadamente el cargue de las fuentes.

Dentro de la fase inicial de carpetas encontramos los notebooks de procesos, estos son códigos que incluyen el cargue, transformación y resultado de los datos. Teniendo una buena administración de información dentro de un proyecto de ingeniería o ETL de data, se incluyen etapas de almacenamiento y procesamiento de la data, como inicial o data bruta tenemos raw\_data y como carpeta de almacenamiento final de resultados como es la gold\_data. esto podemos verlo a más detalle en las siguientes imágenes:

**Figura 9** Proceso notebooks



Fuente: elaboración propia

Dentro de la estructura del proyecto encontramos los códigos fuente que realizan la transformación de toda la data, tanto cualitativo como cuantitativo; estos códigos son los encargados de orquestar el proceso desde el archivo en bruto hasta los archivos e insumos procesados requeridos para análisis y visualización. Ambos contienen las mismas estructuras definidas que son:

**Figura 10** *Proceso notebooks - python*

Select items to perform actions on them.			Upload	New	Refresh
Desktop / proyecto_emp			Name	Last Modified	File size
<input type="checkbox"/>	..			hace unos segundos	
<input type="checkbox"/>	datos			hace 12 horas	
<input type="checkbox"/>	1_proceso_cuantitativo.ipynb	Running		hace una hora	40.9 kB
<input type="checkbox"/>	2_proceso_cualitativo.ipynb			hace una hora	137 kB
<input type="checkbox"/>	dashboard_esc_doc_fac_adm.pbix			hace 36 minutos	351 kB

Fuente: elaboración propia

Librerías, instaladores: generalmente hablando en python el uso de librerías más comunes recaen en pandas y numpy, son fuentes de código simples para la depuración, transformación y formulación de fuentes. Otras librerías importantes en el proceso descarguen son glob, que es una librería de expansión del patrón de nombres de ruta tipo unix, permitiendo realizar cargues y concatenados de archivos con fuentes en común. Finalmente se usan librerías como emoji y nltk para análisis y extracción de texto, así como cloud para la creación de las nubes de palabras.

Funciones: bajo una buena gestión de ingeniería de datos se requiere tener en orden las funciones que se van a usar y crear en todos los procesos, para luego ser llamadas dentro del código y poder tener una buena ejecución y versionamiento de los procesos. Las funciones que se crean acá incluyen, lineamiento de nombres de variables, limpieza, duplicados, y validación de textos.

Cargues de base: la fuente de automatización del proceso radica en la librería glob, para el caso del proceso definido tomamos la búsqueda inicial de archivos con nombre encuesta y luego el cargue de la pestaña correspondiente a cada ruta requerida, tanto cualitativa como cuantitativa.

A continuación, mostramos parte del código PY usado para el cargue a través de glob, especificando fuente de archivo y ruta de acceso:

```
paths=glob.glob("./datos/01_raw/encuesta*.xlsx")
base_cuanti_ls=[]
for path in paths:
    df=pd.read_excel(path,sheet_name="Cuantitativo")
    base_cuanti_ls.append(df)

base_cuanti_df=pd.concat(base_cuanti_ls)
```

Transformación: procesos donde las bases, sufren el mayor tratamiento y modificación requerida para generar las fuentes de información final, depuraciones, cálculos, frecuencias, cambios de nombre, creación de nubes de palabras, son algunos de los procesos que encontramos en esta etapa. Dentro de este proceso cabe mencionar las métricas de transformación de texto, estas incluyen:

- **Normalización:** transformacional o separación de las palabras raíz de cada frase.
- **Corrección:** funciones de transformación de palabras que se identifican están mal escritas y permiten ganar más data. Se identificaron regionalismos más comunes, usos incorrectos de “s” o “c” y se usaron fuentes de identificación de errores más comunes a la hora de comentar textos.
- **Borrado:** funciones de eliminación de palabras que se identifican no aportan nada al análisis y tampoco son stop words.
- **Stop words:** palabras identificadas como: que, más, aun, la etc. que generalmente son muy usadas y generan distorsiona el análisis.

Almacenamiento: etapa que fija la carpeta gold o de procesamiento final para almacenar las fuentes del dashboard o bases depuradas.

## ✓ **Dashboard**

“Un dashboard es una herramienta que gestiona la información con la finalidad de monitorizar, analizar y tomar decisiones de manera visual indicadores o cifras relevantes para el negocio y la toma de decisiones, es usado actualmente por todas las áreas de las empresas, departamentos, campañas o procesos específicos. Cumple principalmente con las siguientes características:

personalizado: cifras y datos relevantes del negocio, resolviendo preguntas que desea responder el destinatario.

Visual: los datos se presentan en forma de gráfica y con indicadores claros a través del color y la figura idónea.

Practico: orientar a acciones y toma de decisiones con la información necesaria en tiempo real” (Ortiz, 2022)

Uno de los puntos a cubrir es la necesidad visual y detallada de información requerida por la secretaría académica; aspectos de visualización, detalles importantes para la toma de acciones y retroalimentaciones requeridas en tiempos competitivos al mundo actual. El dashboard de las encuestas cualitativas y cuantitativas, representa una mejora en tiempos y en posibilidades de explotación de la información.

**Figura 11 Estructura Dashboard**



Fuente: elaboración propia

Cuantitativo - Métricas generales: en este apartado del dashboard encontramos cifras generales de calificación por dimensión (preguntas asociadas a un énfasis particular: orientación, evaluación, aprendizaje y responsabilidad-cumplimiento) tanto como algunos valores y cifras generales, máximos de calificación, calificación media, estudiantes calificados y estudiantes con respuestas, para poder ver también el detalle agrupado de las preguntas que alimentan la encuesta.

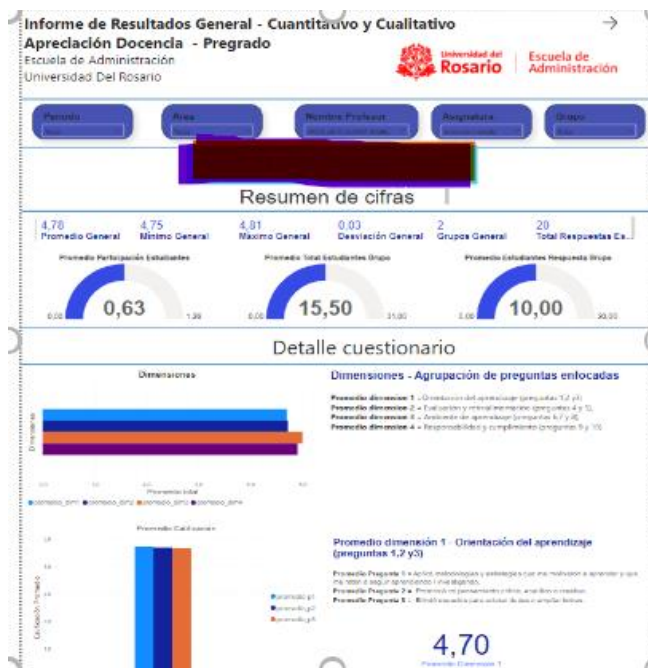
**Figura 12 Estructura Dashboard – Resultados generales**



Fuente: elaboración propia

Cuantitativo - Métricas detalladas: se encuentran los mismos indicadores mencionados anteriormente, pero a nivel docente, permitiendo ver las cifras de cada docente a nivel individual, junto a promedios y respuestas de cada encuesta a nivel cuantitativo.

**Figura 13** Estructura Dashboard – Resultados detallados

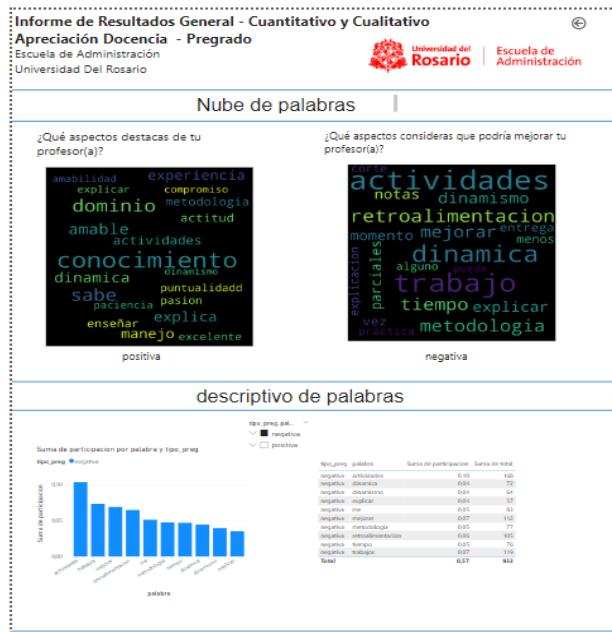


Fuente: elaboración propia

Cualitativo – proceso global base: para este apartado se hace uso de minería de texto y frecuencia de palabras para abordar las dos preguntas principales del proceso cualitativo: aspectos a destacar y aspectos de mejora de los docentes. Generando un detalle a nivel de cada pregunta por medio de nubes de palabras con el proceso de depuración y transformación antes mencionado, que

permite identificar las palabras más usadas para la búsqueda de un profesor de alto nivel y en contra parte las palabras que más se encuentran en oportunidades de mejora.

**Figura 14** Estructura Dashboard – Resultados cualitativos



Fuente: elaboración propia

## 7.2 Análisis estadístico de texto

### 7.2.1 Análisis de sentimientos

El análisis de sentimientos o minería de opinión es un proceso evaluativo de los comentarios, donde se extrae información subjetiva por medio de un algoritmo en lenguaje natural, con la capacidad de generar resultados identificando y clasificando palabras positivas, negativas o neutras (Leonardo Diego, 2019).

Sin embargo, el análisis puede resultar complejo debido a ciertas características gramaticales como la morfología o el modo de expresar las opiniones. También, las opiniones pueden ser

explícitas o implícitas, donde se expresa directa o indirectamente la opinión respectivamente. (Cortes Diaz, 2021).

Como se mencionó en capítulos anteriores, los datos no estructurados se encuentran en la base de datos cualitativa; Esta base contiene variables de información general del semestre, las variables de opinión constan de 20 preguntas, de las cuales 10 son de interpretación de test gráficos psicológicos y las otras preguntas de evaluación, interpretación y reflexión del estudiante hacia el docente.

Como recomendación de la secretaría Académica, se enfatizó sobre dos variables que enmarcan en mayor nivel de importancia para su análisis y toma de decisiones. estas son: ¿Qué aspectos destacas de tu profesor (a)? y ¿Qué aspectos consideras que podría mejorar tu profesor (a)?, de esta manera, todo el análisis se aplicó en base a estas dos variables.

En la fase de exploración y entendimiento de los datos, se identificó que los datos no cumplen con los parámetros necesarios para un análisis de sentimientos, debido a que las preguntas de la evaluación docente solo destacan o aspectos positivos o aspectos por mejorar de los docentes, dando un espacio a preguntas sesgadas y no de opinión requeridas para este tipo de modelos. Razón por la que se propone realizar un análisis estadístico de texto con la información de las preguntas seleccionadas realizando el planteamiento de hipótesis sustentadas en los comentarios de estas dos preguntas.

Como primera medida se busca justificar y comprobar que las hipótesis presentadas cuentan con un respaldo analítico y estadístico por medio de modelos de machine learning, y así llevarlas a un contexto de negocio-académico, las cuales aportaran conclusiones y recomendaciones que generaran valor agregado a secretaría académica.

### 7.2.2 *Clasificador Naive Bayes*

“Tanto en probabilidad como en minería de datos, un clasificador ingenuo Bayesiano (clasificador naive bayes) es un método probabilístico que tiene sus bases en el teorema de Bayes y recibe el apelativo de ingenuo dadas algunas simplificaciones adicionales que determinan la hipótesis de independencia de las variables predictoras. El argumento de Bayes no es que el mundo sea intrínsecamente probabilístico o incierto, sino que aprendemos sobre el mundo a través de la aproximación, acercándonos cada vez más a la verdad, a medida que recogemos más evidencias.

En términos sencillos, el clasificador ingenuo de Bayes asume que la presencia o ausencia de una característica particular no está relacionada con la presencia o ausencia de cualquier otra característica” (Camacho, 2021)

Acorde a los requerimientos del análisis, se aplicó la **clasificación Naive Bayes** especialmente para tener consistencia estadística de las ponderaciones y distribuciones de las frecuencias de las palabras. Con estos resultados, se realizó una agrupación de las palabras más importantes en base a nuestro criterio experto, en diferentes categorías (Experto en Materia, Enseñanza y manejo del curso, Habilidades blandas y Evaluación y retroalimentación), acorde a su clasificación, lo que permitió usar métodos de visualización en el comportamiento de cada una de estas opiniones.

En la siguiente imagen podemos observar “el proceso de clasificación que se aplica como el corpus de subjetividad con el algoritmo de Naïve Bayes para determinar el score de los términos que influyen en la clasificación del comentario. Por lo tanto el algoritmo de Naïve Bayes, además de considerar la cantidad de términos negativos y positivos en el corpus de subjetividad, también considera el score de cada término calculado” (Gutiérrez Esparza et al 2017).

**Figura 15** las características Mas informativas del clasificador Naive Bayes

```

Most Informative Features
contains(conocimiento) = True      1.0 : 0.0 = 67.1 : 1.0
contains(dominio) = True          1.0 : 0.0 = 63.2 : 1.0
contains(amable) = True           1.0 : 0.0 = 43.4 : 1.0
contains(gran) = True             1.0 : 0.0 = 42.5 : 1.0
contains(pasion) = True           1.0 : 0.0 = 40.3 : 1.0
contains(menos) = True            0.0 : 1.0 = 32.9 : 1.0
contains(compromiso) = True       1.0 : 0.0 = 29.9 : 1.0
contains(amabilidad) = True       1.0 : 0.0 = 29.5 : 1.0
contains(conocimientos) = True    1.0 : 0.0 = 27.3 : 1.0
contains(diapositivas) = True     0.0 : 1.0 = 26.9 : 1.0
contains(conoce) = True           1.0 : 0.0 = 24.0 : 1.0
contains(calificar) = True        0.0 : 1.0 = 23.3 : 1.0
contains(amplio) = True           1.0 : 0.0 = 22.9 : 1.0
contains(calificacion) = True     0.0 : 1.0 = 22.1 : 1.0
contains(falta) = True            0.0 : 1.0 = 18.5 : 1.0
contains(tal) = True              0.0 : 1.0 = 18.2 : 1.0
contains(aveces) = True           0.0 : 1.0 = 17.3 : 1.0
contains(mejorar) = True          0.0 : 1.0 = 17.3 : 1.0
contains(pueda) = True            0.0 : 1.0 = 16.1 : 1.0
contains(ninguno) = True          0.0 : 1.0 = 16.0 : 1.0
contains(creo) = True             0.0 : 1.0 = 15.1 : 1.0
contains(respetuoso) = True       1.0 : 0.0 = 15.0 : 1.0
contains(dejar) = True            0.0 : 1.0 = 14.7 : 1.0
contains(ninguna) = True          0.0 : 1.0 = 14.7 : 1.0

```

Fuente: elaboración propia

Gracias a que la base contaba con una cantidad similar de comentarios negativos y positivos, no fue necesario aplicar técnicas de sobre muestreo o submuestreo, que consiste en crear o eliminar datos en la parte minoritaria o mayoritaria de los datos respectivamente. Para el entrenamiento del modelo se usó una separación de 80-20, 80% de los datos para el entrenamiento, y el 20 % restante para las pruebas de este. Como resultado el modelo tuvo una precisión del 0.713800. la exactitud del modelo respalda la toma de decisiones en negocios, ya que su finalidad, es validar si el modelo es adecuado para ser aplicado a nuestra base de datos.

**Figura 16** *Naive Bayes Accuracy*

```
# now we can test our Naive Bayes' classifier and find it Accuracy
• print('NB model accuracy: {}'.format(
    nltk.classify.accuracy(NBclassifier, test_set)))

NB model accuracy: 0.7138001359619307
```

Fuente: Elaboración propia

En base a nuestro criterio experto, se categorizaron las palabras con mayor porcentaje de frecuencia positiva o negativa, para poder identificar patrones que nos ayudaran a entender las fortalezas y debilidades de los profesores; así crear recomendaciones las cuales permitan a la secretaría académica implementar acciones de mejora de acuerdo con cada necesidad o categoría.

Dentro de estas categorías encontramos:

- **Experto en materia:** Se refiere al nivel de conocimiento avanzado de un profesor. Esta categoría se caracteriza por el tiempo y esfuerzo que los profesores han dedicado a su preparación convirtiéndose en expertos reconocidos en el campo de la educación.
- **Enseñanza y manejo del curso:** Esta categoría hace referencia, a las estrategias y métodos utilizados por el profesor para el desarrollo de las clases, la cual conlleva a una correcta administración, gestión y planificación del contenido y recursos necesarios, con el fin de facilitar el aprendizaje y lograr una participación activa un aprendizaje efectivo por parte de los estudiantes.
- **Habilidades blandas:** Se busca resaltar la importancia de las habilidades blandas en el ámbito académico, dado que estas se refieren a las habilidades no técnicas ni profesionales de un profesor, pero que son indispensables para la interacción efectiva, empática y

coherente entre profesores y estudiantes, logrando enfrentar desafíos y obstáculos en el proceso de enseñanza y aprendizaje.

- **Evaluación y retroalimentación:** Principalmente busca medir el avance y el logro de los objetivos académicos, brindando acompañamiento a los estudiantes para que identifiquen tanto sus fortalezas como sus áreas de mejora

### Categorización en base al criterio experto

**Tabla 5** *Categorías Criterio Experto*

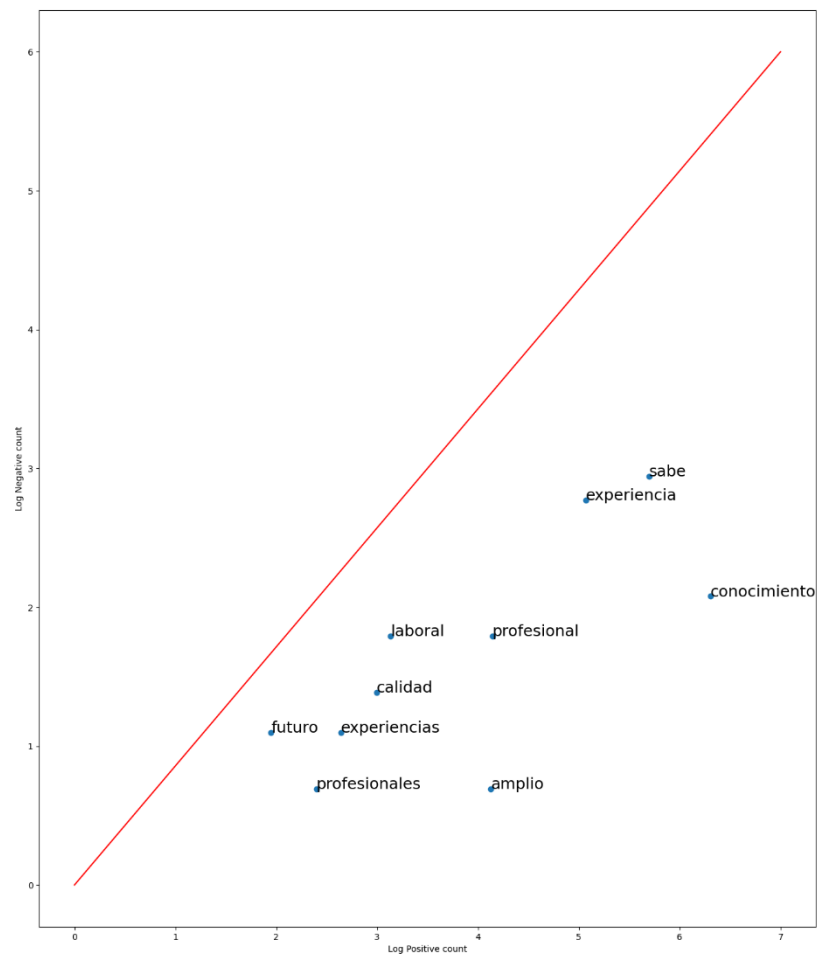
Categoría	Palabras positivas	Palabras negativas
<b>Experto en materia</b>	Conocimiento, conocimientos, amplio, sabe, profesional, perfecto, laboral, experiencia, profesionalismo, experiencias, calidad, profesionales, futuro, mundo, perfecto.	
<b>Enseñanza y manejo del curso</b>	Dominio, pensamiento, demuestra, ayuda, interesante, area, ambiente, creatividad, profesora, resolver, buena, calidad, interesantes, enseña, prepara, actitud, paciencia, critico, esfuerzo, ejemplos.	enseñar, rapido, monotono, entendamos, aburrido, monotona, aburrida, implementar, leer, ritmo, tediosa, laboratorios, monotonas, actividad, abordar, profundizacion, programa, proponer, presentaciones,
<b>Habilidades blandas</b>	Amable, Pasión, buena, amabilidad, compromiso, respetuoso, amena, paciente, atento, comprensiva, responsabilidad, cordial, puntual, respeto, final, enseñar, demuestra, humano, dedicado, comprensivo, respetuosa, creatividad, trasmite, resolver, ideas, interesantes, enseña, confianza, idioma, paciencia, critico, empatico, esfuerzo, chevere, cordialidad, creativa.	pasion, aburrido, estricto, chistes, exigir, incentivar, miedo, flexible, tiempo.
<b>Evaluación y retroalimentación</b>	Conceptos, calidad, critico, aprende, feedback.	Calificación, ninguno, calificar, notas, parciales, plataforma, teoricas, tareas, academica, retroalimentar, evaluar, e-aulas, errores, estricto, retroalimentacion, evaluaciones, cantidad, comentarios, talleres, laboratorios, actividad, entregar, asistencia, exigir, miedo, profundizacion, proponer, comunicacion, exponer, flexible, tiempo, lecturas, instrucciones.

Fuente: Elaboración propia

A continuación, se presentan gráficas, con escala logarítmica, donde permiten visualizar y analizar la relación que existe entre los comentarios tanto positivos como negativos. La escala logarítmica permite una representación apropiada y precisa. Los resultados bajo la figura lineal son comentarios que resaltaron los estudiantes como positivos. Por el contrario, los que se encuentran encima de la figura lineal representa los comentarios negativos.

### Categoría Experto en materia

**Figura 17** *Experto en materia*



Fuente: Elaboración propia.

Este criterio muestra claramente el alto nivel de satisfacción de los estudiantes frente a la experiencia tanto profesional como la experticia de sus docentes, se resalta palabras como

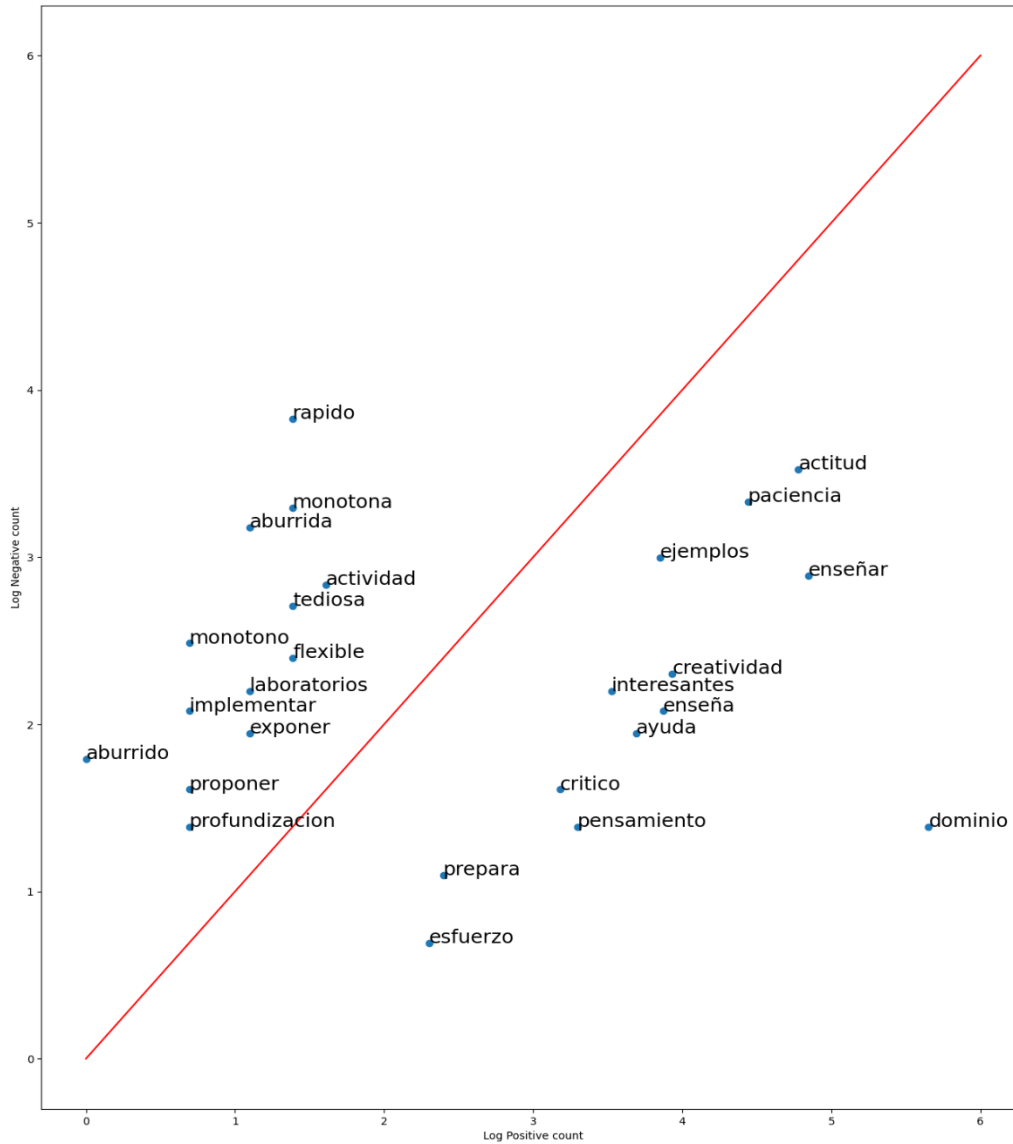
‘**conocimiento**’ una de las más elegidas con una frecuencia de 493 veces, ‘**experiencia**’ con una frecuencia de 120 veces, ‘**sabe**’ (que expresa experticia), con 230 veces y ‘**profesionalismo**’ 66 veces. La percepción negativa de los estudiantes para este criterio no presenta opiniones de gran peso, lo que demuestra la buena gestión y preparación de los procesos de contratación de cada programa de la universidad de Rosario.

Para un docente universitario es de su responsabilidad, asegurar la calidad de la educación, motivar e inspirar al estudiante, comunicarse asertivamente con ellos, con el fin de preparar profesionales íntegros comprometidos, con capacidades diversas para participar en el mercado laboral tanto en su país como en el mundo.

En conclusión, este resultado se relaciona positivamente con la estrategia de la ruta UR 2025, y su plan integral de desarrollo, que define las 6 variables priorizadas dentro de las megas de desarrollo académica, principalmente aportando en la construcción de una de ellas: “Excelencia Académica” y en su objetivo de: “Atraer, retener y consolidar un cuerpo profesoral de excelencia que movilice fronteras de conocimiento, dinamice procesos de enseñanza-aprendizaje y sirva a la comunidad académica y a su entorno”(Plan integral de desarrollo UROSARIO, 2021).

## Categoría Enseñanza y Manejo del curso

**Figura 18** Enseñanza y manejo del curso



Fuente: elaboración propia

En esta categoría encontramos que los estudiantes resaltan positivamente características como: el **“dominio”** con una numero de 219 palabras, **“pensamiento crítico”** alrededor de 20 veces, la **“actitud”** 100 veces, entre otras.

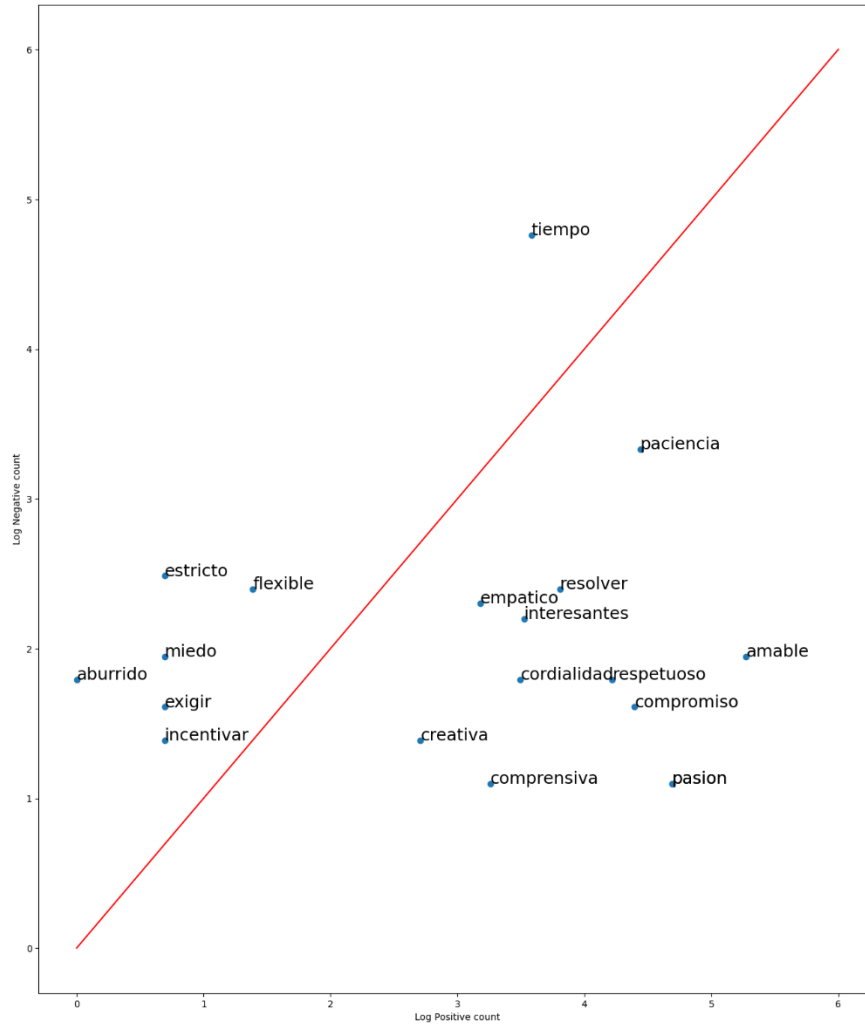
Por otro lado, se observa que varios estudiantes coinciden con una percepción negativa con que algunas clases fueron “**monótonas**” con un número de 40 veces, “**aburrido**” 23 veces, con falta de “**comunicación**” 23 veces y “**tediosas**” 13 veces.

Podemos evidenciar que, aunque los profesores cuenten con buen dominio de los temas, buena actitud frente al grupo, paciencia con los estudiantes para procesar la información, las clases tienden a ser muy monótonas, aburridas, no muy claras, quizás con poco tiempo para captar lo importante, poca adaptabilidad del profesor a las necesidades individuales de cada estudiante.

Se recomienda que la universidad del Rosario brinde capacitaciones y/o charlas a sus profesores en técnicas y manejo de grupos, donde puedan aplicar más de tecnología educativa, nuevos métodos de enseñanza como: juegos de rol, retos en equipo, proyectos dinámicos, donde mantenga a los estudiantes activos, receptivos y motivados, por otro lado también es importante que se sensibilice a crear espacios de confianza en el aula de clases, que sean retadores y participativos, para que así mismo el profesor pueda brindar una correcta retroalimentación, que le aporte valor al aprendizaje de cada estudiante.

## Categoría de Habilidades Blandas

Figura 19 *Habilidades blandas*



Fuente: elaboración propia

En la actualidad, las habilidades blandas en un individuo son tan importantes como el contar con un excelente perfil profesional (educativo + experiencia), ya que estas, permiten la comunicación asertiva, el trabajo en equipo, la comprensión de diferentes puntos de vista entre estudiantes y

profesores, como también liderazgo y comportamiento en el aula, lo cual conlleva al éxito del objetivo propuesto.

Acorde a los resultados del análisis para la categoría de habilidades blandas, se observa que la Universidad del Rosario cuenta con profesores con estas características. Los estudiantes calificaron positivamente a sus profesores con una mayor participación dentro de la data de estudio, como: “**amables**” con una frecuencia de 150 veces, “**apasionados**” 92 veces, “**comprometidos**” 64 veces, “**respetuosos**” 130 veces, “**creativos**” 52 veces, “**empáticos**” 16 veces, “**pacientes**” 120 veces, “**compresivos**” 56 veces, “**cordiales**” 50 veces entre otros.

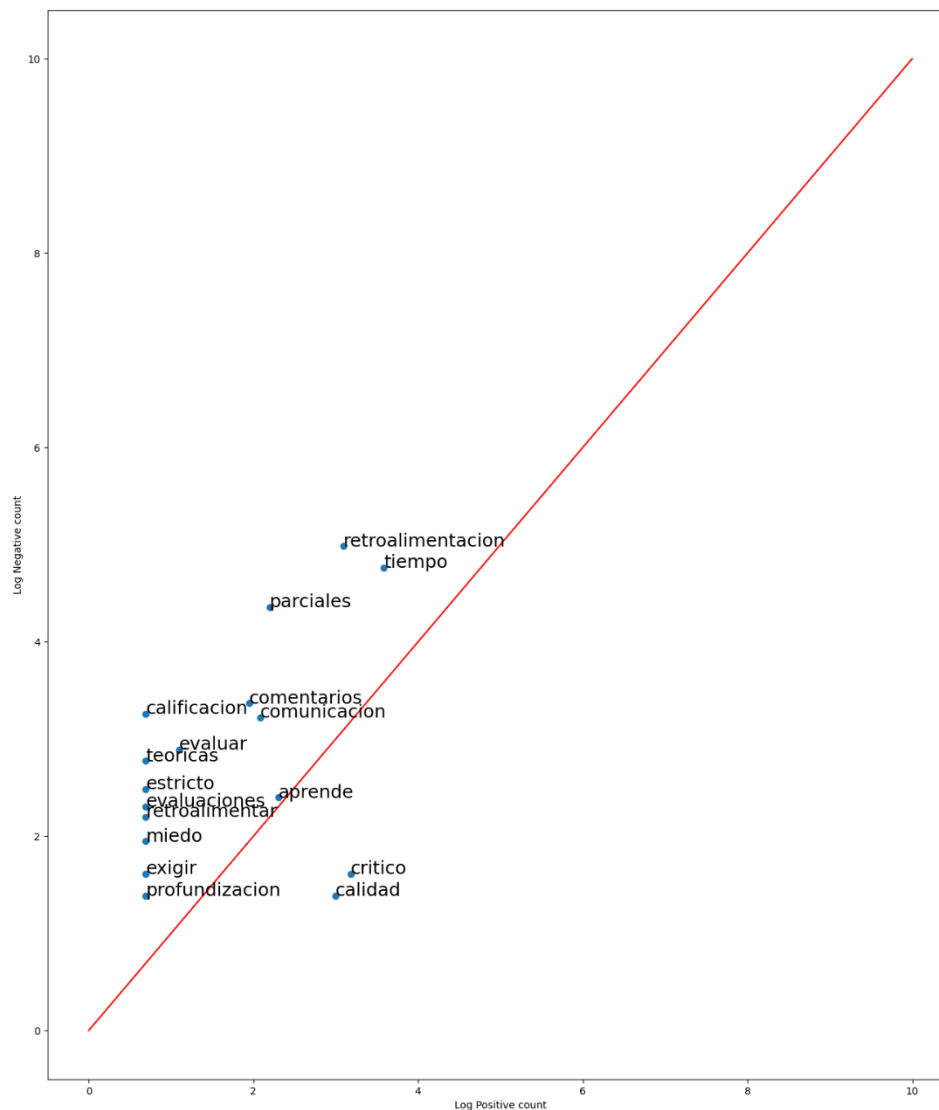
Acorde a nuestro criterio experto, evidenciamos que la mayoría de los profesores de la universidad del rosario, cuentan con habilidades blandas, requeridas para aportar a la calidad de la educación. Sin embargo, resaltan negativamente las siguientes palabras, “**aburrido**” con una frecuencia de 5 veces, “**estricto**” 8 veces, “**chistes**” 3 veces, “**exigir**” 3 veces y “**miedo**” 6 veces. Los estudiantes resaltan que sus profesores son amables y cordiales, los cuales permiten un espacio de confianza para crear un ambiente de aprendizaje agradable, empáticos y comprensivos, lo cual entienden cuando hay dificultades en el aprendizaje y buscan la mejor forma de que el estudiante entienda nuevamente. Adicionalmente, destacan la creatividad como nuevo método de enseñanza para mantener a sus estudiantes interesados en el aprendizaje continuo. Por otro lado, también se reflejaron algunas percepciones negativas, las cuales resaltaban el comportamiento de algunos profesores, como irrespetuosos, empleando chistes inadecuados en sus clases, algunos otros generando ambientes incómodos como de miedo, poca flexibilidad, poca empatía o aburridos.

Las habilidades blandas pueden desarrollarse y mejorarse, es por esto que se recomienda que la universidad pueda crear talleres, actividades prácticas o cursos en línea, que aborden las habilidades que más necesiten fortalecer. Otra oportunidad de mejora podría consistir en charlas, talleres o contratación de profesionales en mentoría y coaching, ya que se enfocan en ayudar profesionales

con gran experiencia, en esas falencias que deben mejorar. Por último, sería de gran valor fomentar la retroalimentación entre los mismos colegas, es decir ver la posibilidad de asistir a las clases de otros profesores, observar y proveer retroalimentación constructiva basada en experiencia, permitiendo ver diferentes enfoques que aporten a la mejora de su metodología profesional.

### Categoría de evaluación y retroalimentación

**Figura 20** Categoría evaluación y retroalimentación



Fuente: elaboración propia

En esta categoría se evidencia que es una de las que requiere más atención tanto de la universidad como de los profesores, ya que los estudiantes han expresado su descontento en la parte evaluativa y su retroalimentación. Se encuentran las siguientes palabras con mayor participación en la parte negativa: **“retroalimentación”** 140 veces, **“notas”** 65 veces, **“parciales”** 58 veces, **“calificación”** 47 veces, **“comentarios”** 25 veces, **“evaluar”** 16 veces, **“plataforma”** 11 veces.

Para los estudiantes, su experiencia en cuanto a la parte evaluativa y de retroalimentación, está siendo muy frustrante, y se demuestra en la evaluación docente, ya que, si un estudiante se dedica a entregar un trabajo o prepararse para un examen, no está recibiendo la respectiva retroalimentación y esto puede afectar directamente la calidad del aprendizaje, ya que no reconoce sus áreas de mejora y fortalezas. Por otro lado, se evidencia que no están contentos con la forma de evaluar de sus docentes, puede que la forma que manejen la parte evaluativa les genere presión, miedo, alto niveles de estrés, lo que resultara afectando negativamente los resultados de estas evaluaciones.

Es por esto que se propone que los profesores creen lazos de confianza con los estudiantes, apliquen diferentes técnicas evaluativas, entreguen retroalimentaciones, y en el caso que no tengan el tiempo, tener a su disposición tutores (estudiantes) destacados de últimos semestres, los cuales puedan colaborar con la retroalimentación y si es necesario aportar a una nueva explicación, lo cual genere autoconciencia y autorreflexión para el estudiante y le permita aplicar mejores técnicas de estudio.

La validación de las hipótesis y la asignación de palabras, se validaron bajo una regresión logística, donde se pudo evidenciar que la asignación de palabras estaba clasificada como positivas y negativas según la base de datos. Se visualizaron las percepciones en las gráficas de regresión y la distribución de frecuencias de cada categoría analizada, Y donde también se respaldaron estas

hipótesis a través la calidad y precisión de los modelos que se entrenaron y testearon con los datos obtenidos de la evaluación docente de la escuela de administración de la Universidad del rosario.

### 7.2.3 *Regresión logística*

La regresión logística, al igual que otras técnicas estadísticas multivariadas, da la posibilidad de evaluar la influencia de cada una de las variables independientes sobre la variable dependiente o de respuesta y controlar el efecto del resto. Tendremos, por tanto, una variable dependiente, llámense Y, que puede ser dicotómica. y una o más variables independientes, llámense X, que pueden ser de cualquier naturaleza, cualitativas o cuantitativas. Si la variable Y es dicotómica, podrá tomar el valor "0" si el hecho no ocurre y "1" si el hecho ocurre. Este proceso es denominado binomial ya que sólo tiene dos posibles resultados, siendo la probabilidad de cada uno de ellos constante en una serie de repeticiones. (Alderete, 2006)

#### ✓ **Función de regresión logística**

La regresión logística es un modelo estadístico que utiliza la función logística, o función logit, en matemáticas como la ecuación entre  $x$  e  $y$ . La función logit mapea  $y$  como una función sigmoidea de  $x$ .

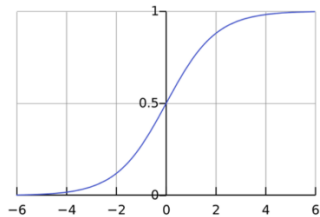
**Figura 21** *Función de regresión logística*

$$f(x) = \frac{1}{1 + e^{-x}}$$

Fuente: elaboración propia

Si representa esta ecuación de regresión logística, obtendrá una curva en S como la que se muestra a continuación.

**Figura 22** Representación de la ecuación de regresión logística



Fuente: elaboración propia

Como puede ver, la función logit devuelve solo valores entre 0 y 1 para la variable dependiente, al margen de los valores de la variable independiente. Así es como la regresión logística estima el valor de la variable dependiente. Los métodos de regresión logística también modelan ecuaciones entre múltiples variables independientes y una variable dependiente (AWS, 2023).

Se realizó una asignación de valores entre 0 y 1 acorde a la preguntas. Para la pregunta positiva ¿Qué aspectos destacas de tu profesor (a)? se le asignó el valor 1 y para la pregunta negativa ¿Qué aspectos consideras que podría mejorar tu profesor (a)? se le asignó el valor 0, para la prueba de hipótesis se realizó una regresión logística, donde la variable aleatoria dependiente de salida Y se estableció como un valor binario (0 ó 1) siendo 0 negativo y 1 positivo, y donde la variable aleatoria X son las palabras y sus frecuencias.

Para el modelo de regresión logística se usó el 80% de los datos para el entrenamiento del modelo, y el 20 % restante para las pruebas de este. De igual forma esta exactitud permite validar que este modelo es correcto para mis datos y soporta las hipótesis que se generaron del análisis, en base a las predicciones que el modelo permite obtener como resultados de una precisión del modelo

del 0.823249. Al final se compararon los resultados obtenidos con las percepciones reales de los estudiantes.

**Figura 23** *Calculo de la exactitud del modelo*

```
# Calculate the accuracy of the model
accuracy = logreg_model.score(X_test, test_y)

print("Accuracy:", accuracy)
```

Accuracy: 0.82324949014276

Fuente: elaboración propia

Se concluye que basado en dos modelos que se entrenaron con los datos cualitativos de la Evaluación docente de la escuela de administración, estos cuentan con la capacidad y la complejidad de predicción. Con el fin de respaldar análisis más profundos en base al criterio experto, el cual nos lleven a mejorar los procesos de enseñanza para la escuela de administración de la universidad del Rosario.

## 8. Conclusiones y recomendaciones

### 8.1 Conclusiones Dashboard y ETL

Para el análisis de información de las encuestas de satisfacción docente cualitativa y cuantitativa de la escuela de administración de la universidad del rosario se concluye lo siguiente:

- La Realización de procesos esperados y resultados para el análisis de las encuestas docentes de la escuela logran reflejar efectividad en tiempos, procesamiento y tratamiento de la data

con entregables de alto valor (dashboard) a través de metodologías correctas de extracción transformación y carga (ETL), permitiendo a la secretaría enfocarse en la gestión y toma de decisiones.

- La extracción y código organizado permite un mayor dinamismo en estructura, procesamiento y control de cambios. Permitiendo a los usuarios el entendimiento de los procesos dada las buenas prácticas de código y estructuración.
- Los resultados presentados en el dashboard presentan las cifras de impacto tanto generales como específicas, adicional con la visualización de texto en el apartado cualitativo permite visualizar las indoles positivas y negativas de mayor impacto, permitiendo tomar decisiones sobre los principales puntos de atención con las palabras que más se presentan en los análisis y nubes de palabras.

## 8.2 Conclusiones análisis estadístico de texto

- Los modelos a nivel de análisis de sentimientos no son factibles con el tipo de cuestionarios que presenta actualmente la universidad del Rosario, dado que cada pregunta se enfoca individualmente en una única dirección o foco directamente positivo o negativo. Adicional que no permiten la interacción entre una pregunta abierta (opinión de calidad) y un rating o variable de salida para validar el sentimiento, posibilidad que se podría dar conectando y relacionando las preguntas de la parte cuantitativa y cualitativa.
- Después del análisis estadístico de texto sobre las evaluaciones docentes se generan conclusiones significativas respecto a algunas características que influyen en la calidad académica. Dentro de Estas características se muestran:
- **La experticia docente:** acorde a los resultados de las evaluaciones y el criterio experto que se aplicó, se evidenció la importancia de los docentes con excelente preparación y

conocimiento, ya que los estudiantes resaltaron positivamente su dominio del tema, además de la forma clara y precisa de transmitirlos.

- **Percepción estudiantil:** La percepción de los estudiantes fue clara y permitió identificar opiniones muy similares en cuanto a fortalezas y aspectos por mejorar de sus docentes. Estas consistencias permitirán a secretaría académica crear planes de acción a corto y mediano plazo, impactando positivamente la estrategia institucional.
- **Desarrollo de habilidades blandas:** las evaluaciones destacaron la importancia de las habilidades blandas, como, por ejemplo, el ser empático, apasionado, y contar con comunicación asertiva, esto motiva a los estudiantes y genera un mayor compromiso de su parte.

### 8.3 Recomendaciones

Dentro de las recomendaciones u oportunidades de mejora, tenemos:

- Mejorar la interacción de los procesos cuantitativos y cualitativos, esto significa que los dos se comuniquen conjuntamente y permita tomar decisiones para realizar procesos de análisis más profundos y estructurados. Las preguntas tipo abiertas generan oportunidades de aplicar modelos más estructurados de análisis de sentimientos y predecir el sentimiento adecuado. las preguntas con una sola dirección positiva o negativa limitan en gran medida estos resultados deseados. Permitir una conexión de pregunta tipo abierta con un score (calificación de 1 a 5 donde 1 es bueno y 5 es malo) de la misma forma facilita el proceso de análisis de texto y sentimientos, para no asumir calificaciones y tener una variable de salida real para aplicar metodologías como BERT, BETO u otros.
- La completitud de estudiantes totales por semestre versus estudiantes que culminan las encuestas es baja, lo que afecta directamente las fuentes de información y de datos

disponibles para los modelos y descriptivos. Un mayor número de datos permite tener un mayor número de sentimientos y ratings de calificación con mayores aportes, palabras y sentimientos que enriquezcan los modelos y permitan realizar calibraciones más adecuadas, adicional de poder llegar a un modelo final con respuesta a nivel docente e incluso a nivel docente-materia.

- Poder tener acceso a la totalidad de las fuentes o encuestas de los estudiantes de la universidad enriquece el proceso analítico (otras facultades, posgrado, etc) y permite tener mayor población de estudio y de comentarios, lo cual enriquecería los modelos y permitiría tener mayor nivel de predictibilidad en los resultados.
- La recolección constante de información, las mejoras de versiones en los procesos entregados (códigos, etl, dashboard, análisis descriptivos) permiten tener impactos positivos en tiempos y en calidad de información para la toma de decisiones. Por lo cual consideramos las experimentaciones iniciales, logros del proceso, hallazgos de hipótesis y otras entregables fuentes de trabajo y mejora dentro de lo que se realizaba previo a la culminación del proyecto empresarial.
- En base a las conclusiones del análisis de texto se presentan las siguientes recomendaciones:
- **Métodos de enseñanza innovadores:** es importante implementar ambientes innovadores con el fin promover la participación y aprendizaje efectivo por parte de los alumnos.
- **Monitoreo de la satisfacción de los estudiantes:** es importante continuar con la realización de las evaluaciones docentes, ya que permite controlar y medir las mejoras implementadas, al punto de lograr el 99.9% de satisfacción en los estudiantes.
- **Retroalimentación y capacitación de docentes:** implementar espacios de retroalimentación a los docentes con calificaciones bajas y percepciones negativas, puede

mejorar el desempeño profesional, generando mejores ambientes en las aulas y permitiendo aprendizaje efectivo por parte de los estudiantes.

### Referencias bibliográficas

- Alderete, A. M. (2006). *Fundamentos del Análisis de Regresión Logística en la Investigación Psicológica* / *Revista Evaluar*.  
<https://revistas.unc.edu.ar/index.php/revaluar/article/view/534>
- Armetrics. (2020, enero 29). *Qué es Dashboard—Definición, significado y ejemplos*.  
<https://www.armetrics.com/glosario-digital/dashboard>
- AWS. (2023). *¿Qué es la regresión logística? - Explicación del modelo de regresión logística - AWS*. Amazon Web Services, Inc. <https://aws.amazon.com/es/what-is/logistic-regression/>
- Camacho, por J. A. (2021, agosto 19). *Clasificador Naive Bayes* / *JacobSoft*.  
[https://www.jacobsoft.com.mx/es\\_mx/clasificador-naive-bayes/](https://www.jacobsoft.com.mx/es_mx/clasificador-naive-bayes/)
- CORPORATIVA, I. (2022, 01). *Discover the main benefits of Machine Learning*. Iberdrola.  
<https://www.iberdrola.com/innovation/machine-learning-automatic-learning>
- Cortes Diaz, N. (2021). *USO DEL ANÁLISIS DE SENTIMIENTOS PARA MEDIR EL POSICIONAMIENTO DE UNA MARCA A TRAVÉS DE LA INFORMACIÓN EXTRAÍDA DE LA PLATAFORMA TWITTER*.
- Editorial Etecé. (s/f). *Dato—Qué es, concepto, ejemplos y tipos de datos*. *Concepto*. Recuperado el 19 de junio de 2022, de <https://concepto.de/dato/>
- Gutiérrez Esparza, G. O., Margain Fuentes, M. de L., Ramírez del Real, T. A., & Canul Reich, J. (2017). Un modelo basado en el Clasificador Naïve Bayes para la evaluación del desempeño docente. *RIED. Revista iberoamericana de educación a distancia*.  
<https://doi.org/10.5944/ried.20.2.17717>

- Haya, P. (2021, noviembre 29). La metodología CRISP-DM en ciencia de datos—IIC. *Instituto de Ingeniería del Conocimiento*. <https://www.iic.uam.es/innovacion/metodologia-crisp-dm-ciencia-de-datos/>
- La Universidad—Universidad del Rosario*. (2022, diciembre 1). <https://urosario.edu.co/static/La-Universidad/Inicio/>
- Leonardo Diego, P. (2019). *ANÁLISIS DE SENTIMIENTOS: APLICACIÓN SOBRE TEXTOS EN REDES SOCIALES*.
- Ministerio de Educación. (2020). *Informes e indicadores de educación superior*.  
<https://snies.mineducacion.gov.co/portal/ESTADISTICAS/Informes-e-indicadores/>
- Ortiz, D. (2022). *¿Qué es un dashboard y para qué se usa? (2023)*.  
<https://www.cyberclick.es/numerical-blog/que-es-un-dashboard>
- Plan integral de desarrollo UROSARIO.pdf*. (s/f). Recuperado el 15 de mayo de 2023, de  
[https://ruta2025.urosario.edu.co/ruta2025/media/Ruta2025/PID\\_Urosario\\_2021\\_v2019-min.pdf](https://ruta2025.urosario.edu.co/ruta2025/media/Ruta2025/PID_Urosario_2021_v2019-min.pdf)
- Ruiz Ayuso, P. (2019, noviembre 6). Cómo viajar en la transformación digital (IV).  
*Ibermática/digital an Ayesa company*. <https://ibermaticadigital.com/como-viajar-en-la-transformacion-digital-iv/>
- Santos, D. (2023). *Qué es ETL: Definición, proceso y herramientas*.  
<https://blog.hubspot.es/marketing/que-es-etl>
- Universidad del Rosario Dirección de Planeación y Efectividad Institucional. (2022). Boletín Estadístico 2021. *instname:Universidad del Rosario*.  
[https://doi.org/10.48713/10336\\_34244](https://doi.org/10.48713/10336_34244)

Weiss, S. (2005). Análisis de sentimiento. En *Wikipedia, la enciclopedia libre*.

[https://es.wikipedia.org/w/index.php?title=An%C3%A1lisis\\_de\\_sentimiento&oldid=1436](https://es.wikipedia.org/w/index.php?title=An%C3%A1lisis_de_sentimiento&oldid=1436)

66384

## Anexos Técnicos

### 1. Ejemplo cuestionario evaluación docente:



Dirección  
Académica

CEAP  
Centro de Enseñanza, Aprendizaje  
y Trayectoria Profesional

#### FORMULARIO DE PREGRADO

Para la Universidad del Rosario es muy importante conocer tu percepción sobre el avance en tu aprendizaje y las reflexiones acerca del rol que tú y tus profesores (as) han asumido en este semestre tanto en el desarrollo de las clases presenciales como en modalidad de clases en acceso remoto o en estrategia Hyflex. Esta encuesta es anónima y confidencial y tus respuestas estarán destinadas a nuestro aprendizaje colectivo. Te pedimos que al momento de contestar las preguntas lo hagas con claridad, sinceridad y asertividad.

#### SOBRE LA EXPERIENCIA GLOBAL DEL CURSO

Teniendo en cuenta lo desarrollado a lo largo de todo el semestre o módulo, y con una mirada integral que incluya las clases presenciales, en acceso remoto o en estrategia Hyflex, contesta las siguientes preguntas:

✚ Si el profesor(a) realizó el diálogo formativo de la 4ª a la 6ª semana ¿Justifica tu respuesta si consideras que este diálogo contribuyó o no al cumplimiento de los resultados de aprendizaje?\*

1. En una escala de 1 a 5, en la que 1 es la menor y 5 la mayor opción, valora los siguientes aspectos relacionados con el rol de tu profesor(a) y sus aportes a tu proceso de aprendizaje (puedes registrar decimales):

APRECIACIÓN DE LA DOCENCIA	
El profesor/ La profesora:	
Orientación del aprendizaje	1. Aplicó metodologías y estrategias que me motivaron a aprender y que me retan a seguir aprendiendo / investigando.*
	2. Promovió mi pensamiento crítico, analítico o creativo.*
	3. Brindó espacios para aclarar dudas o ampliar temas.*
Evaluación y retroalimentación	4. Aplicó evaluaciones y desarrolló actividades que contribuyeron a mi proceso de aprendizaje.*
	5. Brindó una retroalimentación útil y oportuna que me permitió identificar mis aciertos y mis errores.*
Ambiente de aprendizaje	6. Promovió un ambiente de aprendizaje basado en la confianza y la tolerancia hacia los demás.*
	7. Demostró un interés por atender el bienestar de los estudiantes durante las actividades académicas.*
	8. El ambiente del curso se caracterizó por el respeto y la cordialidad.*
Responsabilidad y cumplimiento	9. Asistió puntualmente a las clases y aprovechó el tiempo para desarrollar las actividades programadas.*
	10. Desarrolló los temas o actividades para el curso de acuerdo con los objetivos previstos.*

¿Qué aspectos destacas de tu profesor(a)?

¿Qué aspectos consideras que podría mejorar tu profesor(a)?

2. ¿Cuál o cuáles de las siguientes imágenes representan tu experiencia de aprendizaje en este curso? Puedes elegir una o más imágenes. Cuentas con un espacio para ampliar tus reflexiones.



Amplía tus reflexiones para las imágenes 1 a la 11

3. ¿Cuáles de las siguientes cualidades / valores del profesor(a) hicieron que en general consideres este curso como una buena experiencia de aprendizaje?

- Motivación
- Dominio del tema
- Compromiso social - ambiental
- Pasión
- Integridad



Universidad del  
**Rosario**

**Dirección  
Académica**

**CEAP**  
Centro de Enseñanza, Aprendizaje  
y Trayectoria Profesional

- Dinamismo
- Innovación
- Empatía
- Humano
- Reflexivo
- Ninguna de las anteriores
- OTRA: SI O NO\*

SI MARCÓ SI \* ¿Cuál? \_\_\_\_\_ \*

¿Cómo la(s) cualidad(es) / valor(es) elegido(s) fueron aplicada(s) por el (la) profesor(a) en su proceso de enseñanza?

4. En general ¿Cuánto consideras que aprendiste en este curso o actividad académica?

- Más de lo esperado
- Menos de lo esperado

Justifica tu respuesta

5. Teniendo en cuenta que la Universidad ha venido avanzando en la aplicación de la estrategia Hyflex que dispone recursos y escenarios presenciales, remotos y mixtos para que el estudiante pueda acceder a la formación y a los recursos según sean sus propios intereses y necesidades, te pedimos nos retroalimentes acerca de las ventajas y oportunidades de mejora que le has visto a esta estrategia al interior de nuestra Universidad. (En el caso que no se haya aplicado la estrategia Hyflex en este curso te pedimos registrar la respuesta No Aplica)

6. Espacio para que nos compartas tus reflexiones acerca de tu experiencia en este curso, los contenidos, sobre el profesor o sobre algún otro aspecto no contemplado en esta encuesta

## 2. Proceso del modelo Cuantitativo en R

Se adjuntan anexos de unas partes del código desarrollado

```

library(readxl)
library(reporttools)
library(dplyr)
##install.packages("xlsx")
##install.packages("readr")
##install.packages("writexl")
library(readr)
library(xlsx)
##install.packages('reporttools')
##analisis ini

###cargue inicial de las bases a procesar
#####base cuantitativas####
setwd("C:\\Users\\omarm\\Documents\\Maestria Business Analytics\\proyecto_empresarial\\proceso_hist")

base_cuanti_2019_1 <- read_excel("2019-1/Reporte final apreciación de la docencia 2019-1 - Administración.xlsx",sheet="ESC. ADMIN")
colnames(base_cuanti_2019_1)

base_cuanti_2019_2 <- read_excel("2019-2/base_2019_2.xlsx")
colnames(base_cuanti_2019_2)

base_cuanti_2020_1 <- read_excel("2020-1/ESCUELA DE ADMINISTRACIÓN.xlsx")
colnames(base_cuanti_2020_1)

base_cuanti_2020_2 <- read_excel("2020-2/Reporte final Cuali_Cuanti - Escuela de Administración 2020-2S.xlsx",sheet = "Cuanti")
colnames(base_cuanti_2020_2)

base_cuanti_2021_1 <- read_excel("2021-1/base.xlsx",sheet = "2021-1")
colnames(base_cuanti_2021_1)

base_cuanti_2021_2 <- read_excel("2021-2/Resultados_cuali y cuanti.xlsx",sheet = "Cuantitativo")
colnames(base_cuanti_2021_2)

#file_choose()
base_cuanti_2022_1 <- read_excel("2022-1/Resultados_cuantitativos_cualitativo2022.xlsx",sheet = "Cuantitativo")
colnames(base_cuanti_2022_1)

base_final<-rbind(base_cuanti_2022_1)
base_final<-base_final %>% distinct()####colocar las bases requeridas en el analisis
summary(base_final)
scolnames(base_final)

###nombres homologación:
base_final2<-base_final %>% rename(Periodo='Periodo',
'Area'='Área',
'Unidad_academica'='Unidad Académica',
'Idioma'='Idioma',
'Id_asignatura'='Id asignatura',
'Asignatura'='Asignatura',
'Estrategia'='ESTRATEGIA',
'Grupo'='grupo',
'Actividad'='Actividad',
'Id_profesor'='Id_profesor',
'Nom_profesor'='Profesor',
'total_est_curso'='Nro. estudiantes en el curso',
'total_est_rta'='Nro. estudiantes que respondieron',
'participacion_est'='Participación',
'Promedio_general'='Promedio general',
'promedio_dim1'='Promedio dimensión 1 - Orientación del aprendizaje (preguntas 1,2 y3)',
'promedio_dim2'='Promedio dimensión 2 - Evaluación y retroalimentación (preguntas 4 y 5)',
'promedio_dim3'='Promedio dimensión 3 - Ambiente de aprendizaje (preguntas 6,7 y 8)',
'promedio_dim4'='Promedio dimensión 4 - Responsabilidad y cumplimiento (preguntas 9 y 10)',
'promedio_p1'='Promedio P1 - Aplicó metodologías y estrategias que me motivaron a aprender y que me retan a seguir aprendiendo / investigando.',
'promedio_p2'='Promedio P2 - Promovió mi pensamiento crítico, analítico o creativo.',
'promedio_p3'='Promedio P3 - Brindó espacios para aclarar dudas o ampliar temas.',
'promedio_p4'='Promedio P4 - Aplicó evaluaciones y desarrolló actividades que contribuyeron a mi proceso de aprendizaje.',
'promedio_p5'='Promedio P5 - Brindó una retroalimentación útil y oportuna que me permitió identificar mis aciertos y mis errores.',
'promedio_p6'='Promedio P6 - Promovió un ambiente de aprendizaje basado en la confianza y la tolerancia hacia los demás.',
'promedio_p7'='Promedio P7 - Demostró un interés por atender el bienestar de los estudiantes durante las actividades académicas.',
'promedio_p8'='Promedio P8 - El ambiente del curso se caracterizó por el respeto y la cordialidad.',
'promedio_p9'='Promedio P9 - Asistió puntualmente a las clases y aprovechó el tiempo para desarrollar las actividades programadas.',
'promedio_p10'='Promedio P10 - Desarrolló los temas o actividades para el curso, de acuerdo con los objetivos previstos.')

```

```

> summary(base_final)
  Período          Unidad Académica          Área          Idioma          Id asignatura          Asignatura          ESTRATEGIA
Length:341      Length:341      Length:341      Length:341      Min. :10810001      Length:341      Length:341
Class :character  Class :character  Class :character  Class :character  1st Qu.:10810011      Class :character  Class :character
Mode :character   Mode :character   Mode :character   Mode :character   Median :13210015      Mode :character   Mode :character
                                                Mean :19450262
                                                3rd Qu.:16910076
                                                Max. :84220004

  Grupo          Actividad          Id_profesor          Profesor          Nro. estudiantes en el curso          Nro. estudiantes que respondieron
Min. :1.000      Length:341      Min. :1.944e+05      Length:341      Min. :1.00          Min. :0.000
1st Qu.:1.000    Class :character  1st Qu.:5.177e+07      Class :character  1st Qu. :8.00        1st Qu. :1.000
Median :2.000    Mode :character   Median :7.953e+07      Mode :character   Median :15.00       Median :2.000
Mean :3.097      Mean :2.232e+08
3rd Qu.:5.000    3rd Qu.:8.087e+07
Max. :10.000     Max. :1.137e+09
                                                Mean :16.07
                                                3rd Qu.:25.00
                                                Max. :35.00
                                                Mean :3.507
                                                3rd Qu.:5.000
                                                Max. :20.000

Participación          Promedio general          Promedio dimensión 1 - orientación del aprendizaje (preguntas 1,2 y3)
Min. :0.0000          Min. :1.000          Min. :1.000
1st Qu.:0.0741        1st Qu.:4.450          1st Qu.:4.330
Median :0.1667        Median :4.760          Median :4.670
Mean :0.2189          Mean :4.607            Mean :4.509
3rd Qu.:0.2857        3rd Qu.:4.930          3rd Qu.:4.960
Max. :1.0000          Max. :5.000            Max. :5.000
                                                NA's :56
                                                NA's :56

Promedio dimensión 2 - Evaluación y retroalimentación (preguntas 4 y 5)          Promedio dimensión 3 - Ambiente de aprendizaje (preguntas 6,7 y 8)
Min. :1.000          Min. :1.000
1st Qu.:4.190        1st Qu.:4.560
Median :4.690        Median :4.830
Mean :4.495          Mean :4.676
3rd Qu.:5.000        3rd Qu.:5.000
Max. :5.000          Max. :5.000
NA's :56              NA's :56

Promedio dimensión 4 - Responsabilidad y cumplimiento (preguntas 9 y 10)
Min. :1.000
1st Qu.:4.670
Median :5.000
Mean :4.761
3rd Qu.:5.000
Max. :5.000
NA's :56

Promedio P1 - Aplicó metodologías y estrategias que me motivaron a aprender y que me retan a seguir aprendiendo / investigando.
Min. :1.00
1st Qu.:4.17
Median :4.60
Mean :4.44
3rd Qu.:5.00
Max. :5.00
NA's :56

```

### 3. Resumen de la información para obtención de resultados

```

colnames(base_final2)
base_final3<-base_final2 %>% select(Período,Área, Unidad_academica, Idioma,Id_asignatura, Asignatura, Estrategia, Grupo, Actividad, Id_profesor, Nom_profesor,
total_est_curso,total_est_rta,participacion_est,Promedio_general,promedio_dim1,promedio_dim2,promedio_dim3,promedio_dim4,promedio_p1,promedio_p2,
promedio_p3,promedio_p4,promedio_p5,promedio_p6,promedio_p7,promedio_p8,promedio_p9,promedio_p10)

colSums(is.na(base_final3))
base_final3<-base_final3 %>% distinct()

base_final3b<-base_final3 %>% filter(total_est_rta==0)

write.csv(base_final3a,"base_final_cuantitativa.txt",row.names = FALSE)
write.csv(base_final3b,"base_final_cuantitativa_sd.txt",row.names = FALSE)
library (xlsx)

```

### 4. Proceso del modelo Cuantitativo en R

```

####Grupo
base_cuali_2022_1 <- read_excel("2022-1/Resultados_cuantitativos_cualitativo2022.xlsx",sheet = "cualitativa")
colnames(base_cuali_2022_1)

base_cuali_2022_1b<-base_cuali_2022_1 %>% rename('Período'=Período,
'Área'=Área,
'Unidad_academica'=Unidad,
'Id_asignatura'=Id_asignatura,
'Asignatura'=Asignatura,
'Estrategia'=ESTRATEGIA,
'Grupo'=Grupo,
'Actividad'=Actividad,
'Id_profesor'=Id_profesor,
'Nom_profesor'=Profesor,
'preg_1'='Justifica la contribución del diálogo formativo',
'preg_2'='¿Qué aspectos destacas de tu profesor(a)?',
'preg_3'='¿Qué aspectos consideras que podría mejorar tu profesor(a)?',
'preg_4'='¿Cómo la(s) cualidad(es) / valor(es) elegido(s) fueron aplicada(s) por el (la) profesor(a) en su proceso de enseñanza?',
'preg_5'='Justifique cuánto aprendió',
'preg_6'='Reflexiones estrategia hyflex') %>% select (Período,Área,Unidad_academica,Id_asignatura,Asignatura,Actividad,Id_profesor)

base_cuali_2022_1b<-base_cuali_2022_1b %>% distinct()
####Grupo
base_final_cuali<-rbind(base_cuali_2021_2b,base_cuali_2022_1b)
base_final_cuali<-base_final_cuali %>% distinct()

```

```
summary(base_final_cuali)
colSums(is.na(base_final_cuali))
base_final_cuali2<-base_final_cuali %>% mutate(val_p1=ifelse((is.na(preg_1 )|nchar(preg_1 )<=1),0,1),
val_p2=ifelse((is.na(preg_2 )|nchar(preg_2 )<=1),0,1),
val_p3=ifelse((is.na(preg_3 )|nchar(preg_3 )<=1),0,1),
val_p4=ifelse((is.na(preg_4 )|nchar(preg_4 )<=1),0,1),
val_p5=ifelse((is.na(preg_5 )|nchar(preg_5 )<=1),0,1),
val_p6=ifelse((is.na(preg_6 )|nchar(preg_6 )<=1),0,1))
```

```
##### Limpieza de texto #####
#####

## Eliminación de saltos de línea y tabulaciones
Discurso_text <- gsub("[[:cntrl:]]", " ", texto_palabras)

## Conversión todo a minúsculas
Discurso_text <- tolower(Discurso_text)

## Se eliminan preposiciones y muletillas
Discurso_text <- removeWords(Discurso_text, words = stopwords("spanish"))

## Se eliminan puntuaciones
Discurso_text <- removePunctuation(Discurso_text)

## Remoción de números
Discurso_text <- removeNumbers(Discurso_text)

## Eliminación de espacios vacíos
Discurso_text <- stripWhitespace(Discurso_text)

## Análisis del texto
## Definimos el texto que se va a analizar
Discurso_corpus<- Corpus(VectorSource(Discurso_text))
Discurso_corpus

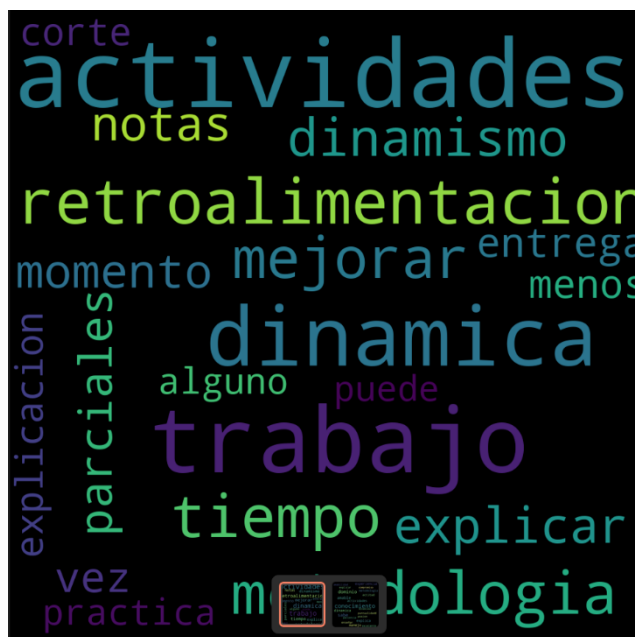
## Se mapea la información del documento
Discurso_ptd<-tm_map(Discurso_corpus, PlainTextDocument)

## Nube de palabras
win.graph()
wordcloud(Discurso_corpus, min.freq = 1, random.order = FALSE)
dev.off()

## Mas depuración
Discurso_text <- removeWords(Discurso_text, words = c("ves","mas","asi","ser","deben","asi","ser","tan","toda","aún","año","peor","tasa","atender","dicen","peor","tener","dos","Fa

#Discurso_corpus <- Discurso_text %>% VectorSource() %>% Corpus()
#Discurso_ptd <- Discurso_corpus %>% tm_map(PlainTextDocument)
```

## 5. Nube de palabras



```
##### modelo #####
#####

sentimientos_df <- get_nrc_sentiment(Discurso_text, lang="spanish")
head(sentimientos_df)

##### resultado #####
#####

sentimientos_df2<-sentimientos_df %>% mutate(enfadado=anger,
                                             preparado=anticipation,
                                             disgusto=disgust,
                                             miedo=fear,
                                             alegria=joy,
                                             tristeza=sadness,
                                             sorpresa=surprise,
                                             confianza=trust,
                                             negativo=negative,
                                             positivo=positive)

sentimientos_df3<-sentimientos_df2 %>% group_by() %>% summarise(enfadado=mean(enfadado),preparado=mean(preparado),disgusto=mean(disgusto)
                                                                ,miedo=mean(miedo),alegria=mean(alegria)
                                                                ,tristeza=mean(tristeza),sorpresa=mean(sorpresa),
                                                                confianza=mean(confianza),negativo=mean(negativo),
                                                                positivo=mean(positivo))

###mod preg 3
texto_palabras <- get_tokens(preg3_esp)
head(texto_palabras)
length(texto_palabras)
sentimientos_df <- get_nrc_sentiment(texto_palabras, lang="spanish")
head(sentimientos_df)


sentimientos_df2<-sentimientos_df %>% mutate(enfadado=anger,
                                             preparado=anticipation,
                                             disgusto=disgust,
                                             miedo=fear,
                                             alegria=joy,
                                             tristeza=sadness,
                                             sorpresa=surprise,
                                             confianza=trust,
                                             negativo=negative,
                                             positivo=positive)

sentimientos_df3<-sentimientos_df2 %>% group_by() %>% summarise(enfadado=mean(enfadado),preparado=mean(preparado),disgusto=mean(disgusto)
                                                                ,miedo=mean(miedo),alegria=mean(alegria)
                                                                ,tristeza=mean(tristeza),sorpresa=mean(sorpresa),
                                                                confianza=mean(confianza),negativo=mean(negativo),
                                                                positivo=mean(positivo))

sentimientos_df4<-sentimientos_df3 %>% mutate(id_profesor=preg3_docentes_1) %>% select (id_profesor,1:10)
```

## 6. Canvas modelo de negocio

Se aplicó el modelo de negocio Canvas, para poder organizar nuestras ideas en cuanto a la solución que queríamos aportar, metodología vista en la materia Analytics producto management.

Canvas Modelo del Negocio 		
Nombre del Proyecto	Supuestos claves en el proyecto	Problema Clave
<p>Generación de un proceso de análisis descriptivo y minería de opinión para el sistema de evaluación docente de pregrado de la Escuela de Administración de la Universidad del Rosario</p>	<p><b>Creemos que a través del Desarrollo de este modelo Analítico Automatizado tanto Cualitativo como cuantitativo la Universidad garantizará:</b></p> <ol style="list-style-type: none"> <li>1. Información de alta calidad, muy precisa ya que estas metodologías son basadas en el criterio experto matemático lo cual genera un alto grado de confianza</li> <li>2. Menor tiempo en la consolidación y análisis de la información generada para el comité de evaluación.</li> <li>3. Adecuada toma de decisiones con resultados exitosos en corto, mediano y largo plazo.</li> <li>4. Implementación y uso de herramientas y metodologías, de vanguardia cumpliendo con los objetivos institucionales</li> </ol>	<p>Como podemos aprovechar los datos obtenidos durante la aplicación de la Evaluación Docente que aplica Secretaría Académica de la Escuela de Administración de la Universidad del Rosario, para la elaboración del informe que es presentado al comité directivo al finalizar cada semestre, con el fin de tomar las mejores decisiones en implementación de nuevas estrategias en pro de mejorar el Servicio Educativo y la Calidad de los programas.</p>
<p><b>Resultados Esperados</b></p> <ul style="list-style-type: none"> <li>• Lograr que Secretaría Académica, obtenga los resultados de alta calidad, en menor tiempo para el desarrollo de informe que será socializado en el Comité de Evaluación Docente al finalizar cada semestre.</li> <li>• Desarrollar del modelo Cuantitativo y Cualitativo, determinando variables claves que será relacionadas para generar resultados reflejados en el dashboard.</li> <li>• Capacitar a Secretaría Académica, en las herramientas que serán implementadas en el Proyecto, el cual los convertirá en el Punto Focal del Proyecto.</li> <li>• Entender el correcto funcionamiento del modelo, y como se estructuran y analizan los obtenidos, para lograr resultados tan confiables y precisos.</li> </ul>		<p><b>Métricas (Cumplimiento &amp; Progreso)</b></p> <p><b>Mayor eficiencia Operativa:</b> Productividad relacionada al tiempo que se toma en realizar el reporte con los resultados obtenidos.</p> <p><b>Mejorar la Satisfacción de los estudiantes</b> frente a los profesores: estrategias implementadas / plan de acción y seguimientos</p>
<p><b>Contexto</b></p> <p>Actualmente Secretaría Académica de la Escuela de Administración de la Universidad del Rosario, realiza un proceso Manual de análisis de los datos recolectados por medio de la evaluación docente, para preparar el Reporte que será presentado ante el Comité.</p>	<p><b>Restricciones</b></p> <p>Presupuesto disponible para instalación de software y herramientas (si no existen) necesarias para el correcto desarrollo del proyecto Equipos con la capacidad requeridas por las herramientas. Uso de nube o servidores de la Universidad</p>	<p><b>Recursos</b></p> <p>Se cuenta con bases de datos Cualitativos y cuantitativos, de diferentes periodos desde el 2019-1 hasta el 2022-2. Contamos con conocimiento de profesores con experiencia tanto en la parte teórica como profesional, lo cual nos apoyan durante el Desarrollo del Proyecto. Contamos con equipos y los programas como R, Python y Power BI. Además acceso a la biblioteca de la Universidad para consultar todo el material necesario en pro de que el Proyecto tenga resultados exitosos</p>
<p><b>Estudiantes Proyecto de Grado</b></p> <p>1) Omar Manta Ulloa - Estudiante Maestría Business Analytics</p> <p>2) Viviana Bocachica Guzmán - Estudiante Maestría Business Analytics</p>	<p><b>Miembros de Secretaría Académica</b></p> <p>1) Sandra Liliana Amaya- Secretaria Académica</p> <p>2) Andres Felipe Rodríguez- Auxiliar de Servicios Académicos</p>	<p><b>Asesor Proyecto de Grado</b></p> <p>Diego Guecha Lopez</p> <p><b>Directores Maestría Business Analytics</b></p> <p>Carlos Alberto Franco Diego Alejandro Ortiz</p>

Fecha: 10-12-22