

**Evaluación de la asimilación de competencias no académicas dentro de un programa de postgrado en una universidad de Colombia con técnicas de minería de datos**

**Diana Carolina Barrera Alfonso**

**Escuela Colombiana de Ingeniería Julio Garavito  
Decanatura de Ingeniería de Sistemas  
Maestría Gestión de Información  
Bogotá D.C., 2020**

**Evaluación de la asimilación de competencias no académicas dentro de un programa de postgrado en una universidad de Colombia con técnicas de minería de datos**

**Diana Carolina Barrera Alfonso**

**Trabajo de investigación para optar al título de  
Magíster en Gestión de Información**

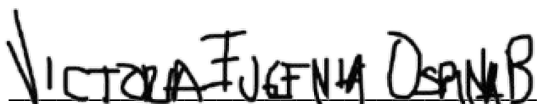
***Directores***

**PhD. Victoria Eugenia Ospina Becerra  
PhD. Dante Conti**

**Escuela Colombiana de Ingeniería Julio Garavito  
Decanatura de Ingeniería de Sistemas  
Maestría en Gestión de Información  
Bogotá D.C., 2020**


## **PÁGINA DE ACEPTACIÓN**

El trabajo de grado de maestría titulado “Evaluación de la asimilación de competencias no académicas dentro de un programa de postgrado en una universidad de Colombia con técnicas de minería de datos”, presentado por Diana Carolina Barrera Alfonso, cumple con los requisitos establecidos y recibe nota aprobatoria para optar al título de Magíster en Gestión de información.



**Victoria Eugenia Ospina**

**Director del Trabajo de Grado**



**Dante Conti**

**Director del Trabajo de Grado**



**Oswaldo Castillo Navetty**

**Jurado**



**Daniela de la Rosa**

**Jurado**

Bogotá, D.C., 01 de septiembre de 2020

## Resumen

Los programas de posgrados, especialmente en Ingeniería y Gerencia preparan profesionales para asumir labores de colaboración, resolución, planificación y liderazgo en organizaciones, convirtiéndolos en entes facilitadores del cambio e integradores de la ingeniería, la administración y las tecnologías, a través de la aplicación de conocimientos, metodologías y herramientas para el modelado, diseño y administración de sistemas complejos que involucran personas, procesos, tecnologías e información y que se orientan tanto al logro de objetivos organizacionales como a la innovación y emprendimiento con sentido ético y responsabilidad social. Bajo este contexto, los programas de posgrado profundizan en la necesidad de implementar o reforzar dentro de sus cursos, competencias “no académicas” para de esta manera equilibrar las enseñanzas técnicas y académicas de cada programa con competencias más blandas, buscando obtener una combinación de competencias globales para perfiles más completos que puedan interactuar con una eficacia mayor dentro de las organizaciones.

En este trabajo se abordó la evaluación de competencias blandas con foco en su asimilación temporal entre 80 profesionales que participaron en el programa de Maestría de Gerencia de Ingeniería de la Universidad de la Sábana (Colombia). Para ello se propuso un marco metodológico/experimental en el contexto del *KDD (Knowledge Discovery in Databases)* con técnicas de minería de datos a los fines de encontrar perfiles de asimilación de las competencias blandas mediante clusterización durante y luego de un año de culminación de estudios, así como, su posible transición temporal en dichos instantes temporales. Se realizaron 4 experimentos al respecto y mediante técnicas de post-procesamiento (centroides de los grupos y visualización) se pudo extraer hallazgos que ofrecen el estatus evaluativo a nivel agrupado de la asimilación de las competencias blandas en el caso de estudio. Las competencias blandas estudiadas son: Planear, Alcanzar, Actuar, Adaptabilidad, Motivación y Liderazgo, las cuales aparecieron en líneas generales reforzadas al cabo de un año de la culminación de estudios.

Los resultados obtenidos además de medir el grado de asimilación de dichas competencias y su cambio o transición al cabo de un año, abren el camino para soportar decisiones o posibles cambios en el reforzamiento o mantenimiento de las prácticas internas actuales dentro del programa de maestría evaluado. Finalmente, esta investigación ofrece líneas futuras de investigación para iniciar análisis más exhaustivos, de ingeniería reversa y de carácter predictivo cuando se pueda contar con mayor volumetría de datos.

**Palabras claves:** *Gestión del Conocimiento en Bases de Datos (KDD), clustering, post-procesamiento, competencias blandas o “no académicas”, programas de posgrado.*

## Tabla de contenido

Capítulo 1 – MARCO INTRODUCTORIO .....	6
1.1 Objetivos .....	8
1.2. Diseño metodológico .....	8
1.3 Solución del problema.....	9
Capítulo 2 – MARCO CONTEXTUAL.....	10
2.1 Competencias.....	10
2.1.1 Tipos de competencias.....	11
2.1.2 Gestión por competencias .....	11
2.2 Estado del arte – Estudios basados en Modelado de Datos y Minería de datos .....	12
2.3 Marco Teórico – Aspectos básicos de Minería de Datos .....	14
2.3.1 Minería de datos .....	14
2.3.2 Clustering .....	16
2.3.3. Algoritmos de Clustering.....	16
2.3.4 Técnicas de visualización y Análisis de Perfiles .....	18
Capítulo 3 – MARCO EXPERIMENTAL .....	20
3.1 Descripción de los datos.....	21
3.2 Etapa selección .....	23
3.3 Etapa datos objetivo .....	23
3.4 Etapa de preprocesamiento/limpieza, procesamiento y transformación .....	24
3.5 Etapa de minería de datos – Extracción de Patrones y su Interpretación .....	25
3.6 Etapa de conocimiento (Visualización) .....	34
Capítulo 4 – CONCLUSIONES Y RECOMENDACIONES .....	38
Referencias Bibliográficas .....	41
Anexos .....	44

## Tabla de ilustraciones

Figura 1. KDD (Knowledge Discovery in Databases) .....	8
Figura 2. Dendrograma del análisis de conglomerado jerárquico aplicando el método de Ward.....	18
Figura 1. KDD (Knowledge Discovery in Databases) .....	20
Figura 1 KDD (Knowledge Discovery in Databases) .....	21
Figura 1 KDD (Knowledge Discovery in Databases) .....	23
Figura 1 KDD (Knowledge Discovery in Databases) .....	23
Figura 1 KDD (Knowledge Discovery in Databases) .....	24
Figura 1 KDD (Knowledge Discovery in Databases) .....	25
Figura 3. Resultados Experimento 1. ....	26
Figura 4. Resultados Experimento 2. ....	28
Figura 5. Centroides del experimento 2 .....	28
Figura 6. Comportamiento por competencias del Experimento 2.....	29
Figura 7. Dendrograma y Boxplots para el Experimento 3 .....	30
Figura 8. Centroides Experimento 3.....	31
Figura 9. Comportamiento por competencias del Experimento 3.....	31
Figura 1 KDD (Knowledge Discovery in Databases) .....	34
Figura 10. visualización de competencias de los estudiantes durante y después de la maestría.....	35
Figura 11. Competencias estudiante grado v1 .....	36
Figura 12. Competencias estudiante grado v2 .....	36
Figura 13. Visualización estudiante graduado por competencias.....	37
Figura 14. Visualización resultados de perfilamiento .....	37
Figura 10. Visualización de competencias de los estudiantes durante y después de la maestría.....	44
Figura 11. Competencias estudiante grado v1 .....	45
Figura 12. Competencias estudiante grado v2 .....	45
Figura 15. visualización perfilamiento de estudiantes en experimento 1 .....	46
Figura 13 visualización estudiante graduado por competencias .....	46
Figura 16. Visualización un año después.....	47
Figura 14 visualización resultados de perfilamiento .....	47

## Capítulo 1 – MARCO INTRODUCTORIO

La innovación organizacional y de gestión educativa se considera una estrategia para el continuo crecimiento en la educación superior. El uso, cada día más presente de herramientas digitales y de Tecnologías de la Información en el entorno universitario, genera datos a una escala cada día más presente que están dando soporte a la gestión educativa y, por tanto, en su mejora con enfoques emergentes de análisis y modelado, dentro de los que se incluyen la Minería de datos y la Inteligencia de Negocios.

A medida que la tecnología avanza, las universidades comienzan a abordar dichos enfoques de análisis de datos generando investigaciones orientadas en estudiar el comportamiento de sus estudiantes y su evolución dinámica y altamente variable, en los diferentes procesos de captación y asimilación de conceptos, así como en las habilidades, competencias y recursos académicos extracurriculares. Los resultados permiten crear herramientas de soporte en la toma de decisiones y así, ayudar en el mantenimiento, monitoreo y evaluación de la gestión académica, el crecimiento de la educación y la excelencia de los programas.

La globalización, la sociedad del conocimiento, los procesos sociales y económicos, impulsaron el surgimiento del modelo por competencias, este modelo surge como una respuesta a las necesidades de sistematizar los saberes para generar un puente entre las necesidades sociales y el desarrollo de saberes en las instituciones educativas. La concepción del proceso de enseñanza aprendizaje al igual que los saberes evoluciona, se requiere de la participación de profesores y estudiantes, un modelo centrado en el saber, saber hacer, en el saber ser y el saber estar. A pesar de la ausencia de un consenso de su concepción o de las controversias y cuestionamientos en torno a sus fundamentos teóricos, el modelo basado en competencias ha sido un escenario propicio para relacionar las necesidades sociales con lo individual así como con la estructura educativa, una triangulación que involucra las competencias afectivas, cognoscitivas, psicológicas, creativas, etc., que además involucra cambios en los sistemas de evaluación, de planes curriculares, todo acompañado del constante cambio, adaptación e innovación. (López, S, 2019).

En Colombia se apuesta a la innovación y la tecnología para elevar la calidad en la educación y Conectividad en el Sistema Educativo, el Ministro (E) de Educación, se realizan esfuerzos para acompañar a estudiantes y docentes en la implementación de las tecnologías de la información y las comunicaciones en el aula de clase a lo largo y ancho del país. "Estos últimos dos años el Ministerio de Educación Nacional ha logrado las cifras más altas de incorporación de TIC`s en la educación de la historia del país, con más de 200 mil docentes formados en el uso de estas herramientas, el 76% de la matrícula escolar pública con acceso a internet, y el 100% de la matrícula de educación conectada a la red", estos datos evidencian el avance que se ha tenido aunque aún no es suficiente. (Ministerio de Educación Nacional, 2020).

En el país hay muchas instituciones que ofrecen el servicio de formación y capacitación, cuyas principales falencias son su alto nivel de dispersión, su heterogeneidad y la falta de interrelación, a lo que se agrega la ausencia de un Sistema de equivalencias que permita la movilidad entre el proceso de formación y el mundo laboral. Se plantea el fortalecimiento y la articulación de la “Educación para el trabajo y el desarrollo humano” con el resto del sistema, con miras a garantizar que la fuerza de trabajo se capacite y actualice sus conocimientos de forma permanente. (Ministerio de Educación Nacional, Colombia, 2010).

Estos análisis no solo cubren aspectos propios de las competencias académicas propias de cada programa de estudio, sino también en el conjunto de habilidades o competencias profesionales “no académicas” incluidas en los programas y que se focalizan en una formación integral de sus participantes.

Bajo este contexto, este proyecto de grado pretende abordar la evaluación y el estatus en el grado de asimilación de las competencias “no académicas” dentro de un programa de postgrado en una universidad colombiana en un horizonte temporal que corta dos etapas claramente definidas: (1) periodo previo a la finalización del programa y (2) periodo posgraduación al cabo de un año. La recopilación y centralización de los datos proviene de encuestas y mecanismos internos de evaluación de la universidad aplicados a una muestra de estudiantes. Esa data alimentará modelos de minería de datos capaces de detectar patrones, perfiles y relaciones, los cuales persiguen segmentar y detectar el grado de asimilación de las habilidades no académicas, antes y después de la finalización del programa, y por ende, evaluar de forma exploratoria, que tan efectivo ha sido el desempeño en la captación de esas habilidades a lo largo del tiempo. Es importante recalcar que el alcance del proyecto no radica en evaluar la efectividad intrínseca de cómo la universidad mide la captación de esas habilidades, sino que se limita a un modelo evaluativo como consecuencia del desempeño centrado en las calificaciones obtenidas por los estudiantes y la ponderación fijada que hace la universidad en esos mecanismos de evaluación para las competencias definidas como “no académicas”.

El proyecto entonces pretende contextualizar la situación actual y evaluar mediante perfiles de estudiantes, el grado de asimilación de dichas competencias y detectar posibles relaciones mediante la interpretación multidimensional de los resultados. El proyecto abordará dicho análisis con técnicas de clusterización y el post-procesamiento de los perfiles obtenidos.



## 1.1 Objetivos

### Objetivo General

Proponer un modelo exploratorio y de minería de datos para evaluar la asimilación de las competencias propuestas en la Maestría en Gerencia de Ingeniería de la Universidad de La Sabana

### Objetivos Específicos

- Recepción de la información basada en las competencias académicas propuestas por la maestría en Gerencia de Ingeniería de la Universidad de La Sabana.
- Limpiar, preparar y transformar la data en una vista de datos “minable” como input para el modelado.
- Diseñar el modelo exploratorio y de minería de datos para el caso de uso.
- Validar el prototipo identificando la satisfacción del modelo junto a los usuarios y expertos que gestionan la maestría.
- Crear un dashboard exploratorio para resumir resultados y consumo de la información obtenida durante la investigación.

## 1.2. Diseño metodológico

El problema antes citado se aborda con técnicas de Minería de Datos y Gestión del conocimiento. Para ello se propone al KDD por sus siglas en inglés como marco metodológico. El KDD o Descubrimiento de Conocimiento en Bases de Datos (Knowledge Discovery in Databases – KDD) es un proceso del tipo “paso” a “paso” capaz de encontrar un “modelo” válido, útil y entendible que describa patrones de acuerdo con la información, y su posterior consumo como soporte en procesos de toma de decisiones. La figura 1 ilustra los pasos del KDD seguidos en este trabajo. (Landa, J, 2016)



Figura 1. KDD (Knowledge Discovery in Databases)

### Pasos del KDD:

- Como en cualquier tipo de investigación, es fundamental tener muy claros los límites y objetivos de la investigación, de tal manera de no perder el rumbo en el océano infinito de datos a nuestra disposición o en alcances del problema bien delimitados (Minerva, J, 2018). Aquí, el foco es la obtención

de perfiles que permitan extraer información relacionada a la asimilación de competencias para el caso de estudio de la investigación.

- Creación de un set de datos objetivo: Los datos relevantes de nuestro estudio deben seleccionarse, integrarse y transformarse para que puedan convertirse en el input del modelado basado en datos (Minerva, J, 2018).
- Limpieza y procesamiento de datos: El preprocesamiento y la limpieza tienen el objetivo de mejorar la calidad de los datos y los resultados de la minería. (Minerva, J, 2018)
- Minería de datos: con la vista minable se procede a la ejecución de motores de análisis con técnicas y algoritmos que pueden incluir Estadística, Aprendizaje de Máquinas (Machine Learning), entre otros. Para esta investigación se usan algoritmos de agrupamiento (Clustering), Interpretación de Perfiles (Post-procesamiento de perfiles) y visualización de resultados mediante Tableros o Dashboards.
- Interpretación de los patrones minados y su validación ante el usuario final.
- Conocimiento y su presentación como soporte al proceso de toma de decisiones.

### 1.3 Solución del problema

*¿Cómo se puede cualificar si las competencias propuestas por la MGI (Maestría en Gerencia de Ingeniería de la Universidad de La Sabana son realmente logradas por el estudiante al final de su periodo académico?*

Esta investigación propone la aplicación de un motor robusto analítico que usa técnicas de clusterización con el fin de discriminar perfiles, antes y después de la finalización del programa (al cabo de un año), para a partir de allí, (a) hacer análisis comparativo entre los perfiles obtenidos (sus centroides característicos) y los umbrales máximos ideales previamente establecidos por la universidad dentro de la nota máxima y los pesos dados a cada competencia, y (b) medir estadísticamente (vía frecuencia relativa) el grado de propensión promedio a que las competencias se mantengan en el tiempo según los umbrales ideales establecidos por la universidad. Los resultados se exportan a un tablero sencillo que pueda dar soporte en la identificación de sesgos y por ende, ayudar y reforzar los mecanismos de control asociados al seguimiento y monitoreo de la calidad del programa académico y sus variantes centradas en las competencias “no académicas” impartidas en la Maestría.

El proyecto entonces determinará los perfiles actuales (basado en la muestra histórica de los estudiantes sometidos en el proceso de monitoreo) para medir el grado de asimilación de 6 grandes competencias inmersas dentro del evolutivo académico de cada estudiante (calificaciones obtenidas mientras cursan o cursaron el programa) y los resultados de un mecanismo de evaluación posgraduación realizado al cabo de un año. Las competencias definidas para el análisis son: **Planear, Alcanzar, Actuar, Adaptabilidad, Motivación y Liderazgo.**

## Capítulo 2 – MARCO CONTEXTUAL

### 2.1 Competencias

Las competencias son aquellas habilidades, capacidades y conocimientos que una persona tiene para cumplir eficientemente determinada tarea. Son características que capacitan en un determinado campo; no solo incluyen aptitudes teóricas, sino también definen el pensamiento, el carácter, los valores y el buen manejo de las situaciones problemáticas. Dichas competencias son adquiridas mediante el aprendizaje y la formación de las personas; son una herramienta fundamental para el ejercicio en el campo donde sean necesarias y permiten que la persona combine las destrezas y la capacidad en desempeñar una función de forma efectiva y transversal en el tiempo, generalmente definidas en el ámbito profesional (Según link de Significados, 2017).

Las competencias:

- **Articulan conocimiento conceptual, procedimental y actitudinal:** El hecho de acumular conocimientos no implica ser competente necesariamente. El sumatorio de saberes y capacidades no nos lleva a la competencia. El ser competente implica un paso más: supone, de todo el acervo de conocimiento que uno posee (o al que puede acceder), seleccionar el que resulta pertinente en aquel momento y situación (desestimando otros conocimientos que se tienen pero que no nos ayudan en aquel contexto) para poder resolver el problema o reto que enfrentamos (Cano E, 2008).
- **Se vinculan a rasgos de personalidad, pero en realidad se aprende:** El hecho de poseer de forma innata ciertas inteligencias es un buen punto de partida, pero no me garantiza ser competente. Las competencias deben desarrollarse con formación inicial, con formación permanente y con experiencia a lo largo de la vida. Se puede ser competente hoy y dejarlo de ser mañana o serlo en un contexto y dejarlo de ser en otro contexto que no me resulta conocido. Las competencias tienen, pues, un carácter recurrente y de crecimiento continuo. Nunca se “es” competente para siempre (Cano, E., 2008).
- **Toman sentido en la acción pero con reflexión:** El hecho de tener una dimensión aplicativa (en tanto que suponen transferir conocimientos a situaciones prácticas para resolverlas

eficientemente) no implica que supongan la repetición mecánica e irreflexiva de ciertas pautas de actuación. Al contrario, para ser competente es imprescindible la reflexión, que nos aleja de la estandarización del comportamiento (Cano, E., 2008).

### **2.1.1 Tipos de competencias**

Las competencias son definidas según el área en la cual son ejecutadas. Existen varios tipos de competencias que pueden ser divididas en (Según link de Significados, 2017)

- Las competencias básicas, o también llamadas competencias para la vida, ayudan al individuo a insertarse adecuadamente en un determinado contexto social el cual suele asociarse a valores universales
- Las competencias genéricas aquellas que son útiles en todo tipo de profesión o trabajo como el trabajo en equipo, la proactividad, la empatía o la creatividad.
- Las competencias específicas se refieren a aquellas que son necesarias en un ámbito profesional como competencias laborales, competencias docentes y competencias comunicativas.

### **2.1.2 Gestión por competencias**

Los sistemas educativos convencionales usualmente se enfocan en el estudio teórico, que sustenta a la enseñanza más en la repetición, en la recordación por memorización de la información, que en la interiorización por reflexión y comprensión. Un gran esfuerzo hacia "las notas" para evidenciar el impacto en las evaluaciones de conocimientos mediante pruebas escritas y orales para alcanzar finalmente, un certificado impreso (Yturralde, E, 2019).

La metodología del Aprendizaje Experiencial basada en el Constructivismo es una nueva tendencia para enfocar la educación, formación y capacitación de una manera muy práctica basada en la experiencia y la reflexión, en la cual fundamentamos nuestra propuesta de capacitación, generando resultados de alto impacto (Yturralde, E., 2019).

En el mundo real, las exigencias van más allá de la recordación de información, de los conocimientos que adquirimos, de los pasos a seguir para un resultado, de los saberes, puesto que para alcanzar el éxito se requiere tener las habilidades blandas, competencias sociales, habilidades transversales muy desarrolladas, como se ha comprobado por investigaciones de la Universidad de Harvard en conjunto con *The Carnegie Foundation* y *Stanford Research Institute* que concluyeron en la importancia de las actitudes frente a las aptitudes basados en un estudio

publicado por Charles Riborg Mann. El 15% del éxito en las actividades profesionales y laborales las genera el conocimiento técnico especializado en cada individuo, en sus competencias técnicas, mientras que el 85% radica en las habilidades blandas, en su actitud, en su autodeterminación, en su capacidad de comunicación, de trabajo en equipo, de pensamiento crítico, resolución de conflictos y muchos más aspectos importantes (Yturalde, E., 2019).

- **Habilidades blandas** o también llamadas Soft Skills son aquellas habilidades interpersonales que hacen de un trabajador, un excelente empleado para adaptarse mejor a la cultura de la empresa y al resto del equipo. se demuestran en la ejecución del trabajo, y no están relacionadas, únicamente, con los conocimientos, sino con la puesta en práctica de una combinación de habilidades sociales, habilidades de comunicación, aptitudes y capacidad de acercamiento con los demás. Estas habilidades facilitan la interacción con otras personas, generando un ambiente de trabajo más grato y un clima organizacional de entendimiento y cooperación (Según portal Orientación Universia, 2018).
- **Habilidades duras** o también llamadas hard skills son aquellas destrezas que se aprenden a través de cursos de formación, junto con la experiencia profesional de desempeñar un puesto de trabajo. Al igual que las habilidades blandas, las hard skills son imprescindibles para aportar valor a las tareas que vas a desarrollar para la empresa (Según portal Orientación Universia, 2018).

## **2.2 Estado del arte – Estudios basados en Modelado de Datos y Minería de datos**

Desde un punto de vista más general y pragmático, la educación superior en Colombia en los próximos 15 años se enfrenta en entender los desafíos y retos que exige el siglo XXI, desarrollar acciones y estructurar estrategias a corto, mediano y largo plazo, para así movilizar los recursos y fuentes de financiamiento necesarios para el cumplimiento de las metas y mejora continua de la Educación no solo en términos de competencias académicas, sino de aquellas complementarias o llamadas “no académicas”. De ahí la importancia de una agenda donde se relacionen las acciones prioritarias a seguir para garantizar un financiamiento sostenible que permita la expansión y diversificación del sistema y el fortalecimiento de la calidad en la educación superior con el soporte y ayuda de técnicas emergentes y multidisciplinarias (Ayala M, 2010). Un ámbito decisivo para cualquier proyecto de cambio es el de su puesta en práctica, el de su desarrollo en centros y

aulas, y el de su transformación en procesos y resultados de aprendizaje para todos los sujetos implicados: profesores, alumnos, expertos, etc. De este modo, hoy constituye un imperativo pensar y decidir sobre el cambio tanto en las fases de su construcción y diseño como en las correspondientes a su implementación, evaluación, institucionalización (Escudero, J). El desarrollo de las competencias de los estudiantes durante el proceso de formación se ha convertido en el centro de la actividad educativa en el diseño de los planes de estudios modernos, los cuales se deben comparar con los demás programas de las universidades (López S, 2019). Es de esta manera como han surgido últimamente diversos estudios basados en modelos de datos, analítica avanzada y minería de datos que persiguen dar solución al contexto pragmático citado previamente y contribuir en el proceso de toma de decisiones.

Bajo una óptica más detallada y puntual pueden citarse diversos estudios que van en línea con la justificación y alcance de esta investigación. Los procesos de creación de conocimiento se implantan de manera visual, teniendo en cuenta que esto facilita la interpretación de resultados al implementar modelos y sistemas de clasificación en procesos de minería de datos, los cuales se componen de diferentes técnicas estadísticas y de inteligencia artificial. Las reglas de clústeres, por ejemplo, aplicadas con un enfoque de minería de datos muestran cómo pueden clasificarse y compararse diferentes programas de Ingeniería Industrial en Colombia considerados de alta calidad (De la Hoz, E., 2018).

La minería de datos en la educación no es un tópico nuevo. El uso de técnicas de Minería de Datos permite predecir cualquier fenómeno como la detección de patrones de deserción estudiantil a partir de los datos socioeconómicos, académicos, disciplinares e institucionales de los estudiantes de los programas de pregrado (Tiraman, R., 2013), en dicha investigación se descubrieron perfiles socioeconómicos y académicos de los estudiantes que desertan utilizando la técnica de clasificación basada en árboles de decisión. El conocimiento generado permitió soportar la toma de decisiones eficaces de las directivas universitarias enfocadas a formular políticas y estrategias relacionadas con los programas de retención estudiantil.

Por otra parte, la utilización de diferentes técnicas de clasificación genera modelos predictivos para cada competencia que permiten descubrir conocimiento para predecir tendencias sobre el desempeño académico de los actuales y futuros estudiantes de los programas profesionales, esto con el fin de identificar riesgos y oportunidades que ayuden a las instituciones gubernamentales y de educación superior a tomar decisiones para el mejoramiento de la educación superior en Colombia (Timaran, S, et al, 2016) y (Timaran R, 2016). Otro ejemplo puede verse mediante la aplicación de técnicas de clustering a un curso en específico para evaluar cada una de las competencias y crecimiento académico que el estudiante interioriza de acuerdo con el programa académico de dicha asignatura (Montoya, A., et al, 2015).

Es así como, el proceso de minería de datos educativos busca la adecuación para que la analítica de datos permita interactuar con las variables con el fin de obtener visualizaciones, resolver preguntas o relacionar variables de acuerdo con los datos obtenidos. El modelo metodológico para Minería de Datos Educativos en un aula de educación media presencial cuya implementación y resultados contribuyeron a la reflexión sobre el acto pedagógico en relación con la gestión de datos en el contexto escolar y evaluación del aprendizaje, es otro ejemplo en esta línea del estado del arte (Aristizábal J, 2017).

El uso de técnicas de clasificación en minería de datos de los factores asociados a la percepción que el recién egresado tiene de la utilidad de los conocimientos y destrezas adquiridos a lo largo de sus estudios, que forman parte vital en su rol laboral fue objeto de estudio en un trabajo liderado en la Universidad de Caldas (Bedoya et al., 2016). Los enfoques investigativos como el exploratorio y el descriptivo, juntamente con cuatro técnicas de minería de datos de clasificación y un repositorio de datos permitió cuantificar dicho grado de percepción de utilidad de las habilidades y destrezas adquiridas durante sus estudios.

El rendimiento académico es otro tema estudiado desde hace mucho tiempo. En 2018, en una universidad del Perú (Yamao E, 2018), se realizó la predicción del rendimiento académico de los alumnos que ingresaron en el primer ciclo utilizando minería de datos. Se extrajeron datos que fueron agrupados en tres factores: sociales, económicos y académicos y se realizaron predicciones a través de tres técnicas: regresión lineal, árbol de decisiones y máquinas de vectores de soporte. De los diferentes factores, los que más incluyeron en el rendimiento académico fueron los siguientes: nota de examen de admisión, género, edad, modalidad de ingreso y desplazamiento. Un par de estudios finales también enfocan temas de deserción o abandono en la educación universitaria y patrones de aprendizaje/ estudio. Al respecto, se cita a (Daza A, 2016) y (Vergara, C., et al, 2018).

## **2.3 Marco Teórico – Aspectos básicos de Minería de Datos**

### **2.3.1 Minería de datos**

La minería de datos (*Data Mining*) es un proceso para detectar información sobre conjuntos de datos, de la manera más automáticamente posible. Su objetivo es encontrar patrones, tendencias o reglas que expliquen el comportamiento de los datos en un contexto específico (Garassino A, 2015). Tiene cuatro componentes principales: (Según link Sinergia e Inteligencia de negocios, 2007).

- **Determinación de los objetivos:** Trata de la delimitación de los objetivos que el cliente desea bajo la orientación del especialista en data Mining (Científico de Datos).
- **Preprocesamiento de los datos:** Se refiere a la selección, la limpieza, el enriquecimiento, la reducción y la transformación de las bases de datos. Esta etapa consume generalmente alrededor del setenta por ciento del tiempo total de un proyecto de data mining.
- **Determinación del modelo:** Se comienza realizando unos análisis estadísticos de los datos, y después se lleva a cabo una visualización gráfica de los mismos para tener una primera aproximación. Según los objetivos planteados y la tarea que debe llevarse a cabo, pueden utilizarse algoritmos desarrollados en diferentes áreas de la Inteligencia Artificial, Matemática Aplicada, Estadística, Machine Learning, entre otros.
- **Análisis de los resultados:** Verifica si los resultados obtenidos son coherentes y los coteja con los obtenidos por los análisis estadísticos y de visualización gráfica. El cliente determina si son novedosos y si le aportan un nuevo conocimiento que le permita considerar su apoyo en el proceso de toma de decisiones.

## Técnicas de minería de datos

En el ámbito de la investigación, las técnicas de minería de datos pueden ayudar a los científicos a clasificar y segmentar datos y a formar hipótesis. Permiten encontrar información escondida en los datos que no siempre resulta aparente (Según Clinic-Cloud, 2017).

Las técnicas pueden ser de dos tipos:

**Métodos descriptivos:** Buscan patrones interpretables para describir datos. Son los siguientes: clustering, descubrimiento de reglas de asociación y descubrimiento de patrones secuenciales. Los métodos descriptivos se han utilizado, por ejemplo, para ver qué productos suelen adquirirse conjuntamente en el supermercado (el famoso problema de la cesta de la compra o carrito del supermercado).

**Métodos predictivos:** Usan algunas variables para predecir valores futuros o desconocidos de otras variables. Son los siguientes: clasificación, regresión y detección de outliers. Los métodos predictivos pueden emplearse en tareas como clasificar tumores en benignos o malignos, por citar un ejemplo.



### 2.3.2 Clustering

Las técnicas de Clustering son las que utilizando algoritmos matemáticos se encargan de agrupar objetos con la información que brindan las variables que pertenecen a cada objeto. Se mide la similitud entre los mismos, y una vez hecho esto se colocan en clases que son muy similares internamente (entre los miembros de la misma clase) y a la vez diferente entre los miembros de las diferentes clases. (Enciclopedia colaborativa en la red cubana, 2012).

El análisis de clústers o clustering, también llamado segmentación de data, tiene una variedad de objetivos, todos ellos relacionados con agrupar o segmentar una colección de objetos en subconjuntos o “clústers”, tal que aquellos objetos dentro de cada clúster están más cercanamente relacionados que los asignados a clústers diferentes. Un objeto puede ser descrito por un conjunto de medidas, o por su relación con otros objetos. Adicionalmente, el objetivo puede ser ordenar los clústers en una jerarquía natural. Esto involucra agrupar los clústers sucesivamente, de modo que, en cada nivel de la jerarquía, los clústers en un mismo grupo son más similares entre ellos, que aquellos en diferentes grupos. Centro de todos los objetivos del clustering, es la noción de grado de similitud (o diferencia) entre los objetos individuales a ser clusterizados, y por ello es fundamental para todas las técnicas de clustering, la elección de la medida de distancia o similitud entre dos objetos. Un método de clustering intenta agrupar los objetos basados en la definición de similitud que se le provee. La situación es algo parecida a la especificación de una función de pérdida o costo, en problemas de predicción (aprendizaje supervisado). El costo asociado con una predicción inexacta depende de consideraciones externas a la data. (Fernández, J, & Velásquez, J, 2007)

### 2.3.3. Algoritmos de Clustering

**Clustering Jerárquico:** El algoritmo de clúster jerárquico agrupa los datos basándose en la distancia entre cada uno y buscando que los datos que están dentro de un clúster sean los más similares entre sí. En una representación gráfica los elementos quedan anidados en jerarquías con forma de árbol llamado Dendograma.

Los métodos o agrupamientos jerárquicos van generando grupos en cada una de las fases del proceso buscando el número de clúster hasta converger en una agrupación óptima.

El agrupamiento jerárquico es capaz de fijar por sí solos el número de clústeres (interpretando el Dendograma), por ello se pueden utilizar de forma exploratoria y posteriormente aplicar un análisis no jerárquico con el número de clúster ya fijado.

Las estrategias para conseguir este objetivo se dividen en: estrategias aglomerativas y divisivas. Los Métodos son (Según portal Duk2 – estrategias de trading, 2019):

- Método del vecino más próximo
- Método del vecino más lejano
- Método de agrupación de vinculación promedio
- Método de Ward o Varianza mínima
- Método del Centroides
- Método de la Mediana

**Distancia euclídea:** La distancia euclidiana se calcula desde el centro de la celda de origen hasta el centro de cada una de las celdas circundantes. El algoritmo euclidiano funciona del siguiente modo: para cada celda, la distancia a cada celda de origen se determina al calcular la hipotenusa con  $x_{max}$  y  $y_{max}$  con los otros dos lados del triángulo. Este cálculo deriva la verdadera distancia euclidiana, en vez de la distancia de la celda. Se determina la distancia más corta a un origen, y si es menor que la distancia máxima especificada, el valor se asigna a la ubicación de la celda en el ráster de salida. (Link ArcMap, 2016)

**Método Ward:** Ward propuso que la pérdida de información que se produce al integrar los distintos individuos en clústers puede medirse a través de la suma total de los cuadrados de las desviaciones entre cada punto (individuo) y la media del clúster en el que se integra, de tal manera que el proceso de clusterización resulte óptimo en el sentido de que los grupos formados no distorsionen los datos originales. El método de Ward es uno de los más utilizados en la práctica; posee casi todas las ventajas del método de la media y suele ser más discriminativo en la determinación de los niveles de agrupación el cual es capaz de acertar mejor con la clasificación óptima que otros métodos (mínimo, máximo, media y centroides). (Según Uv.es, 2011)<sup>1</sup>

Este es un ejemplo de identificar perfiles dentro del alumnado según sus niveles de bienestar, se utiliza un análisis de conglomerado jerárquico aglomerativo, empleando el método de Ward, utilizando específicamente la distancia euclídea al cuadrado como medida entre los datos. Este se usó para identificar el número de

---

<sup>1</sup> K-means: construye una partición de las observaciones en  $k$  conjuntos ( $k \leq n$ ) a fin de minimizar la suma de los cuadrados dentro de cada grupo y x-means: agrupar los datos / observaciones en unos pocos segmentos.

conglomerados, que, según el dendrograma, Figura 2, indicó la presencia de tres grupos de estudiantes según sus resultados (H, Paula et al., 2015).

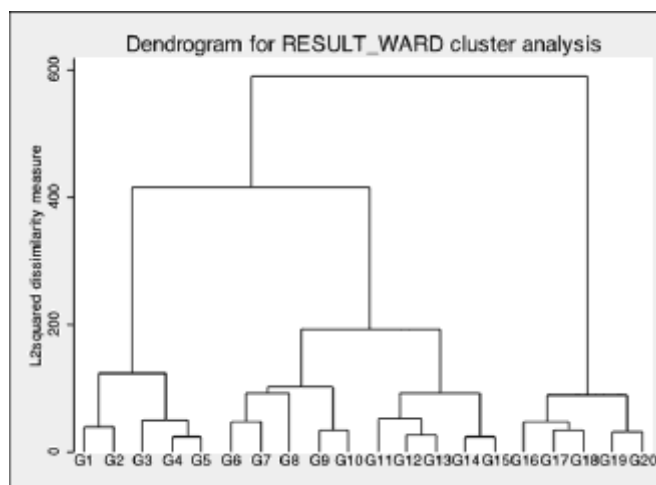


Figura 2. Dendrograma del análisis de conglomerado jerárquico aplicando el método de Ward.

### 2.3.4 Técnicas de visualización y Análisis de Perfiles

La visualización de datos es una representación gráfica de la información y los datos. Mediante el uso de elementos visuales, como gráficos y mapas, la visualización de datos ofrece una manera accesible para detectar y comprender las tendencias, los valores atípicos y los patrones en los datos. (Según link Tableau, 2003) y es pieza fundamental para interpretar y postprocesar perfiles.

El Análisis de Perfiles de Comportamiento Organizacional ofrece una metodología efectiva que guía en el estudio de aspectos esenciales, cuando de definir el comportamiento de una persona en su relacionamiento con otras se trata, así como en el ajuste de sus competencias personales con las que exige el desempeño del puesto de trabajo u ocupación actual o del que pasaría a cumplir probablemente en un futuro. (Sotelo J, 2006)

Las "dimensiones" o variables del Análisis de Perfiles tienen una estrecha relación con aspectos definidos del ciclo de vida de una persona. Nace como una técnica operacional de la Administración de la Educación y tiene excepcionales condiciones de integración con otras técnicas de administración, de estudio y mejoramiento de conductas personales. (Sotelo J, 2006)

El diseño inicial del cual nació el Análisis de Perfiles fue una respuesta a las necesidades de disponer de un instrumento efectivo de evaluación de las

competencias laborales de profesores, médicos, personal de servicios, administrativos, pianistas y otros tipos de personal con diversas responsabilidades, formación y experiencias. El Análisis de Perfiles se puede ejecutar en base a ciclos de aplicación, Cada uno de los ciclos avanza más profundamente sobre cada dimensión. (Sotelo J, 2006)

La aplicación del "primer ciclo" permite realizar un diagnóstico para conocer la situación de la persona en estudio en función de cada una de las variables de comportamiento utilizadas por la técnica. Con las conclusiones de cada estudio se pueden diseñar, entre otras posibilidades, lo siguiente: (Sotelo J, 2006)

- Programas de selección
- Programas de recapitación
- Programas de reconversión de competencias
- Programas de perfeccionamiento.

En una fase más avanzada de aplicación del Análisis de Perfiles, él mismo permite la elaboración de series de seguimiento del progreso realizado por la persona que está siendo evaluada, así como una evaluación continuada del programa de capacitación (Sotelo J, 2006).

## Capítulo 3 – MARCO EXPERIMENTAL

### Diseño metodológico

Para la solución del problema documentado en el capítulo 1, se propone un enfoque de diseño metodológico con técnicas de Minería de Datos y Gestión del conocimiento. Se utiliza el KDD (por sus siglas en inglés, Knowledge Discovery in Databases – KDD) ya que es un proceso del tipo “paso” a “paso” capaz de encontrar un “modelo” válido, útil y entendible que describa patrones de acuerdo con la información, y su posterior consumo como soporte en procesos de toma de decisiones. La figura 1 ilustra los pasos del KDD seguidos en este trabajo. (Landa, J, 2016)



Figura 1. KDD (Knowledge Discovery in Databases)

Es importante recalcar que el diseño metodológico es aplicable y reproducible a problemas similares que quieran gestionar y descubrir patrones. Para ello, el analista debe centrar sus esfuerzos en el objetivo, tipo de datos y tipología o asociaciones que se quieran hallar dependiendo del problema.

Por tanto, en este trabajo se reproduce las respuestas mínimas que un analista debe considerar al abordar el KDD. A continuación, se hace referencia de dichas respuestas en relación con cada uno de los pasos del KDD.

#### *Pasos del KDD:*

- Como en cualquier tipo de investigación, es fundamental tener muy claros los límites y objetivos de la investigación, de tal manera de no perder el rumbo en el océano infinito de datos a nuestra disposición o en alcances del problema bien delimitados (Minerva, J, 2018). Aquí, el foco es la obtención de perfiles que permitan extraer información relacionada a la asimilación de competencias para el caso de estudio de la investigación, pero en otros trabajos podría ser la deserción escolar, evaluación hacia docentes, entre otros.
- Creación de un set de datos objetivo: Los datos relevantes de nuestro estudio deben seleccionarse, integrarse y transformarse para que puedan convertirse en el input del modelado basado en datos (Minerva, J, 2018). Aquí se centra la investigación con las calificaciones y promedios de las 6 competencias sujetas al análisis en 80 estudiantes. Pero en otros trabajos

podría ser calificaciones de profesores, indicadores de asignaturas entre otros.

- Limpieza y procesamiento de datos: El preprocesamiento y la limpieza tienen el objetivo de mejorar la calidad de los datos y los resultados de la minería. (Minerva, J, 2018)
- Minería de datos: con la vista minable se procede a la ejecución de motores de análisis con técnicas y algoritmos que pueden incluir Estadística, Aprendizaje de Máquinas (Machine Learning), entre otros. Para esta investigación se usan algoritmos de agrupamiento (Clustering), Interpretación de Perfiles (Post-procesamiento de perfiles) y visualización de resultados mediante Tableros o Dashboards. En otros trabajos puede sugerirse árboles de decisión, regresión logística, entre otros.
- Interpretación de los patrones minados y su validación ante el usuario final.
- Conocimiento y su presentación como soporte al proceso de toma de decisiones.

Teniendo como base el esquema metodológico KDD (descrito en la Figura 1), en este capítulo se detalla el paso a paso experimental aplicado al caso de estudio de la investigación, desde la captura de los datos hasta los patrones obtenidos en el proceso de Minería propiamente dicha con sus resultados, hallazgos y visualizaciones.

Bajo las premisas anteriores, a continuación se detalla el paso a paso del KDD de esta investigación que pretende obtener conocimiento para evaluar la asimilación de competencias “no académicas” dentro del programa de maestría en Gerencia de Ingeniería en la Universidad de La Sabana, que se ha utilizado como caso de estudio y aplicación.

### 3.1 Descripción de los datos

Etapa datos



Figura 1 KDD (Knowledge Discovery in Databases)

Los datos fueron suministrados en un archivo plano bajo formato Microsoft Excel. Este archivo contiene 4 hojas con información que contiene: 17 asignaturas obligatorias, 12 asignaturas complementarias; cada una con sus respectivos estudiantes y profesores, así como sus respectivas fechas de inicio y fin. También contamos con la evaluación que cada estudiante realizó de la asignatura hacia el

profesor responsable. Esta última información no es utilizada en el marco/alcance de esta investigación. Del archivo fuente, se consideraron únicamente los datos relacionados a las calificaciones por competencia obtenidas por los estudiantes al momento de aún estar cursando la maestría y posteriormente, sus evaluaciones un año después de su titulación (La tabla adjunta lista los descriptores de las variables en los datos).

<b>DATOS</b>	<b>CARACTERISTICAS</b>	<b>DESCRIPCION</b>
Asignatura y docente	<b>ASIGNATURA</b>	listado de asignaturas
	<b>TIPO</b>	obligatorias o complementarias
	<b>DOCENTE</b>	dato anonimizado
	<b>Créditos</b>	peso dentro de pensum
	<b>%</b>	peso dentro de pensum
	<b>fecha reporte de nota 1</b>	fecha de reporte
	<b>fecha reporte de nota 2</b>	fecha de reporte
	<b>fecha reporte de nota 3</b>	fecha de reporte
	<b>fecha reporte de nota 4</b>	fecha de reporte
	<b>fecha reporte de nota 5</b>	fecha de reporte
	<b>fecha reporte de nota 6</b>	fecha de reporte
Estudiantes	<b>codigo_estudiante</b>	identificación del estudiante en el sistema anonimizado
	<b>Nombre</b>	nombres del estudiante en el sistema anonimizado
	<b>Fecha grado</b>	fecha de grado
	<b>promedio nota</b>	nota de grado
	<b>fecha evaluacion_1añodespues</b>	fecha de evaluación año después
	<b>promedio evaluación 1años despues</b>	nota de evaluación año después
	<b>Inicio</b>	fecha de matricula
	<b>Fin</b>	fecha de grado
Notas estudiantes	<b>codigo_estudiante</b>	identificación del estudiante en el sistema anonimizado
	<b>Nombre</b>	nombres del estudiante en el sistema anonimizado
	<b>Nota</b>	nota asignatura
	<b>Asignatura</b>	asignatura
	<b>fecha registro</b>	fecha registro asignatura
	<b>Planear</b>	peso competencia planear
	<b>Alcanzar</b>	peso competencia alcanzar
	<b>Actuar</b>	peso competencia actuar
	<b>Adaptabilidad</b>	peso competencia adaptabilidad
	<b>Motivación</b>	peso competencia motivación
	<b>Liderazgo</b>	peso competencia liderazgo

### 3.2 Etapa selección



Figura 1 KDD (Knowledge Discovery in Databases)

Se realiza una revisión a la base de datos para determinar de estos datos, cuáles son los que vamos a manejar para la toma de decisiones, así como también cuáles de ellos nos van a permitir el cumplimiento de los objetivos planteados.

### 3.3 Etapa datos objetivo



Figura 1 KDD (Knowledge Discovery in Databases)

Los datos objetivo son las notas que cada estudiante obtuvo en cada una de las asignaturas, las cuales van atadas a un porcentaje (%) de cumplimiento de cada competencia aplicado a las asignaturas vistas por cada estudiante según los créditos de dichas asignaturas. Es un proceso estándar aplicado por la Maestría bajo estudio y no está sujeto a validación dentro de la investigación. Las competencias para evaluar en este caso son:

**Planear** los procesos relacionados con la investigación, diseño, desarrollo y producción de productos y servicios para garantizar altos niveles de productividad y agregar valor en cada eslabón de la cadena haciendo uso de herramientas de ingeniería en cualquier tipo y tamaño de organización.

**Alcanzar** el cumplimiento de objetivos organizacionales que garantice un sano desarrollo y crecimiento a través del trabajo interdisciplinario; interpretando y aplicando correctamente indicadores financieros, de operaciones, sociales, del medio ambiente, de mercadeo, de sistemas complejos y de gerencia del conocimiento.

**Actuar** ética y responsablemente tanto en la organización como en la comunidad o país donde se desempeña con conciencia ambiental y sostenible siguiendo regulaciones y leyes, y respetando la cultura organizacional y política para garantizar que las consecuencias de las propuestas generan beneficios a todas las partes involucradas sin riesgo alguno.

**Adaptarse** con flexibilidad a diferentes culturas consciente de las problemáticas globales y la diversidad social.



**Motivar** a otros en organizaciones altamente productivas con habilidades de comunicación efectiva, negociación, solución de conflictos y trabajo en equipo.

**Liderar** y apoyar sus equipos de trabajo con visión, pensamiento estratégico, veracidad, integridad y entusiasmo.

### 3.4 Etapa de preprocesamiento/limpieza, procesamiento y transformación



Figura 1 KDD (Knowledge Discovery in Databases)

En la base de datos se encuentran algunos datos de ruido y campos vacíos, por ejemplo, estudiantes sin calificaciones o estudiantes con calificaciones únicamente en un periodo de análisis. Estos estudiantes no son considerados para el modelado. Se limpia la data y se obtiene información completa para 80 estudiantes plenamente identificados en los 2 periodos de interés del estudio, es decir: calificaciones cursando la maestría y calificaciones al transcurrir un año de su titulación. Asimismo, en esta etapa se enriquece la data con variables indicadoras de umbrales (llamadas valores ideales) para así medir la desviación faltante de cada uno de los estudiantes hacia el ideal o puntaje máximo de cada competencia.

DATOS	OBSERVACIÓN
Estudiantes	se le calcula la diferencia entre promedio nota y promedio evaluación año después
eval estudian grado	se genera la nota ideal y diferencia por cada una de las competencias
eval estu año desp	se genera la nota ideal y diferencia por cada una de las competencias
notas_estudiantes	se genera la nota ideal y diferencia por cada una de las competencias
Estudiante	Se calcula la diferencia entre la nota de graduación vs un año después de la obtención de conocimientos.
Evaluación estudiante grado	Se calcula el % de cada una de las competencias por cada estudiante, tanto el ideal, el que obtuvo el estudiante y la diferencia entre estos datos.
Evaluación estudiante año después	Se calcula el % de cada una de las competencias por cada estudiante, tanto el ideal, el que obtuvo el estudiante y la diferencia entre estos datos.
Notas estudiantes	Se calcula el % de cada una de las competencias por cada una de las asignaturas según sus pesos académico o número de créditos a los que equivale la asignatura, tanto el ideal, el que obtuvo el estudiante y la diferencia entre estos datos.

### 3.5 Etapa de minería de datos – Extracción de Patrones y su Interpretación



Figura 1 KDD (Knowledge Discovery in Databases)

Aquí se procede a ejecutar el clustering para determinar el perfilamiento de los estudiantes en base a sus calificaciones en las competencias en diversos experimentos. El objetivo es agrupar a los 80 estudiantes en base a su similitud asociada al rendimiento de sus competencias. Los grupos obtenidos permitirán etiquetar a los estudiantes para luego postprocesar los centroides de cada grupo y extraer el perfilamiento interpretativo y de allí, establecer comparaciones en: (a) *Perfiles en la etapa “cursando la maestría”*, (b) *Perfiles un año después de la graduación o culminación de la Maestría* y (c) *Evolución y/o contabilización de la transición de pasar de un grupo a otro grupo habiendo transcurrido un año de haber finalizado estudios, esto con el objeto de explorar la probabilidad de transición entre grupos y poder verificar intuitivamente si un estudiante mejoró, empeoró o se mantuvo constante en la asimilación de las competencias impartidas en la maestría.*

Estos experimentos se hicieron en el lenguaje R con sus librerías de clustering y de visualización de resultados. El clustering se soporta en el *método Ward, con distancia Euclidiana* y el análisis se hace con el cálculo de los centroides de cada grupo y cajas del tipo “boxplot” grupo a grupo.

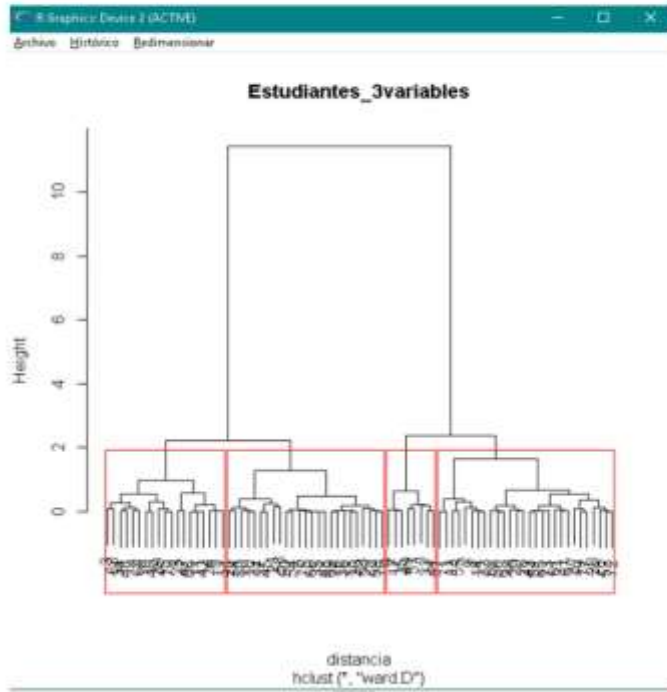
#### Experimento 1: Segmentación exploratoria básica

Este experimento resume la agrupación de los 80 estudiantes considerando el promedio de sus calificaciones durante la etapa de sus estudios, el promedio un año después y la diferencia entre ambos promedios. Son 3 las variables discriminantes y este agrupamiento no refleja en detalles diferencias competencia a competencia, es más bien, un experimento preliminar para inducir hallazgos básicos e iniciales. Este experimento se etiquetó como “resultados estudiantes 3 variables”.

La figura adjunta ilustra el dendograma de agrupamiento y el corte de este donde se determina que se forman 4 grupos bien diferenciados. Al lado se muestran los diagramas de caja calculados para cada grupo a los fines de determinar diferencias y extraer perfilamiento de cada grupo.

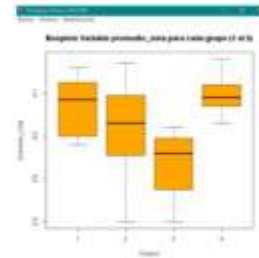
# CLÚSTER

## resultados\_estudiantesvariables

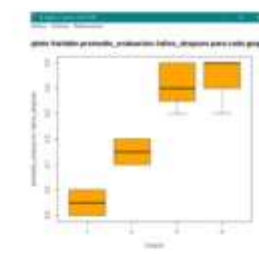


# BOX

Promedio



Promedio grado



Diferencia

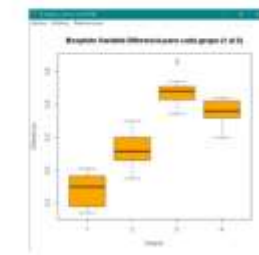


Figura 3. Resultados Experimento 1.

### Perfilamiento y Resultados Experimento 1:

- Los estudiantes se clasifican en 4 clústeres, compuestos por: Grupo1 = 8, Grupo2 = 28, Grupo3 = 19 y Grupo4 = 25 esto para un total de 80 estudiantes, los cuales nos arrojan los siguientes perfiles en base a sus centroides (promedio en cada una de las variables dentro de cada grupo, es decir, entre los estudiantes que caen en cada grupo):
- Perfil 1 → Estudiantes que caen en el grupo o clúster 1 con un centroide para el promedio de 4,07, un centroide para el promedio de grado (un año después) de 3,09 teniendo una diferencia de -0,12 en su centroide promedio en la diferencia de promedios. De manera similar se tienen los perfiles subsiguientes,
- Perfil 2 → Estudiantes con un promedio del 4.017, un promedio de grado del 4,15 teniendo una diferencia del 0,13.

- Perfil 3 → promedio del 3,93 con un promedio de grado del 4,41 teniendo una diferencia del 0,48.
- Perfil 4 → promedio del 4,09 con un promedio de grado del 4,45 teniendo una diferencia del 0,35
- En base a los centroides (ver boxplot) de la variable promedio (durante los estudios) hay diferencias entre los 4 grupos discriminados (véase que cada una de las medianas en las cajas no se intercepta gráficamente entre ellas). Los mejores estudiantes en base a este análisis son los del perfil 4 seguidos por 1, 2 y finalmente los del grupo 3. Similarmente al cabo de un año los mejores estudiantes son los ubicados en los perfiles 4, 3, 2 y 1 respectivamente.
- Sin embargo, al validar la variable diferencia, se puede apreciar que las competencias (medidas en promedio) un año después, se fortalecieron únicamente en los grupos 4, 3 y 2 (la diferencia es positiva). Nótese que en los estudiantes del grupo 1, sus competencias al cabo de un año disminuyen a pesar de haber sido los segundos mejores (durante los estudios) y muestran una diferencia negativa, la cual indica que según el instrumento de medición un año después, disiparon las competencias aprendidas en la maestría. No obstante, este sesgo o desviación negativa en solo 8 de los 80 estudiantes evaluados.
- Este primer experimento permitió obtener desde una primera vista, hallazgos sobre el cambio en el mantenimiento o no en términos de promedio sobre las competencias no académicas impartidas en la maestría. No ofrece detalles a nivel de competencias por separado. Puede concluirse, que a efectos globales el 90% de los estudiantes en promedio (72 estudiantes de 80) mejoran el promedio al cabo de un año, y por ende, en promedio puede indicarse que la asimilación de las competencias se fortalecen con el tiempo.

### Experimento 2: Segmentación Detallada de Estudiantes por competencias durante sus estudios de Maestría (cursando la Maestría):

Este experimento se enriquece con las calificaciones detalladas de cada una de las 6 competencias, en los 80 estudiantes de base muestral durante su periodo como estudiantes de maestría. Las variables discriminantes son multidimensionales en este experimento. Se tiene, el promedio de las competencias, calificaciones de las 6 competencias por separado y 6 variables extras determinadas como el diferencial entre la calificación de las competencias y el máximo ideal determinado por la maestría para un total de 13 variables introducidas al motor de clustering.

De manera similar al experimento 1, se obtuvo el dendograma y los centroides con los siguientes hallazgos:

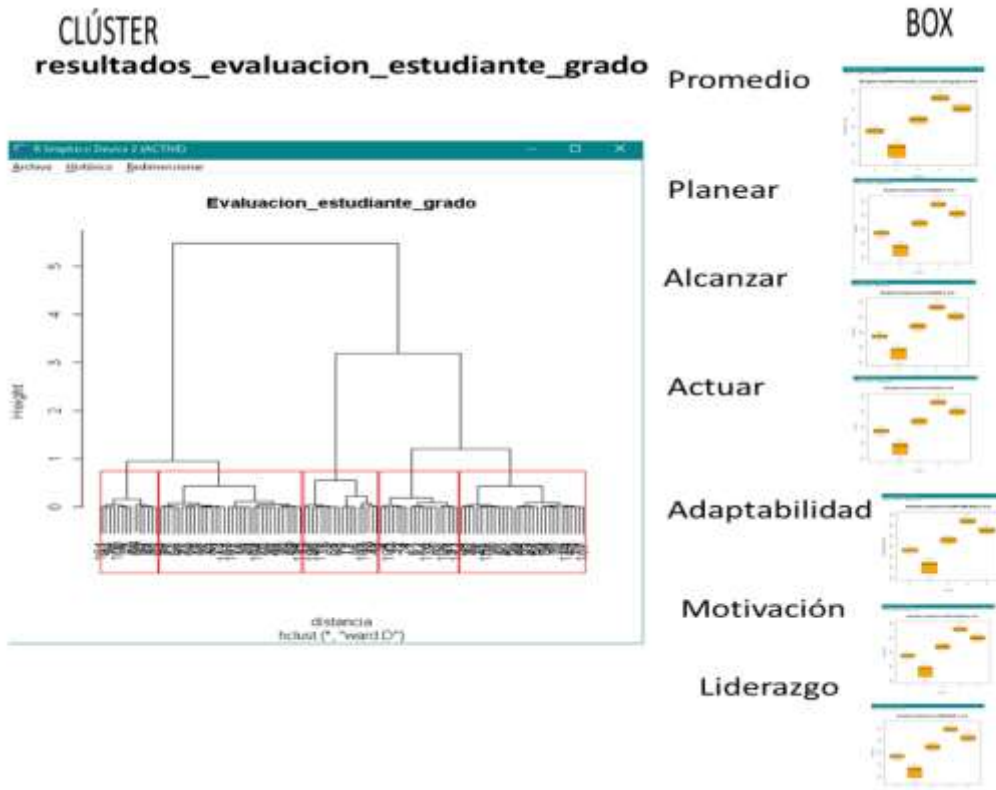


Figura 4. Resultados Experimento 2.

- Los estudiantes se clasifican en 5 clústeres: Grupo 1 con 14 estudiantes, Grupo 2 con 13 estudiantes, Grupo 3 con 22 estudiantes, Grupo 4 con 8 estudiantes y Grupo 5 con 23 estudiantes, respectivamente.

Los centroides de cada grupo o clúster se muestran en la siguiente figura (Figura 5). En relación con dichos centroides y el boxplot se aprecia que, en base a promedio, los estudiantes de los grupos 4 y 5 resultan los mejores, mientras que los del grupo 2 resultan los peores. Sin embargo, los valores en promedio de los 5 grupos pueden aproximarse a un valor de 4 para todos, por tanto, podría asumirse que en líneas generales, los 80 estudiantes durante su fase de estudios de Maestría absorben bastante bien las competencias no académicas ofrecidas por el programa.

etiquetas	difer_liderazgo	difer_motivacion	difer_adaptabilidad	difer_actuar	difer_alcanzar	difer_planear	Liderazgo	Motivacion	Adaptabilidad	Actuar	Alcanzar	promedio_nota	Planear
1	0,12297815	0,10248179	0,15372269	0,20496359	0,21521176	0,22545994	0,47702185	0,39751821	0,59627731	0,79503641	0,83478824	3,97518207	0,87454006
2	0,13470769	0,11225641	0,16838462	0,22451282	0,23573846	0,2469641	0,46529231	0,38774359	0,58161538	0,77548718	0,81426154	3,8774359	0,8530359
3	0,11522449	0,09602041	0,14403061	0,19204082	0,20164286	0,2112449	0,48477551	0,40397959	0,60596939	0,80795918	0,84835714	4,03979592	0,8887551
4	0,10038242	0,08365201	0,12547802	0,16730403	0,17566923	0,18403443	0,49961758	0,41634799	0,62452198	0,83269597	0,87433077	4,16347985	0,91596557
5	0,10761829	0,0896819	0,13452286	0,17936381	0,188332	0,19730019	0,49238171	0,4103181	0,61547714	0,82063619	0,861668	4,10318095	0,90269981

Figura 5. Centroides del experimento 2

- Si se calculan los promedios de la variable diferencia para las 6 competencias y grupo por grupo, puede notarse que los perfiles 4 y 5 agrupan a los estudiantes que más se acercan a la idealidad o absorción completa de las competencias, mientras que los peores se ubican en el grupo 2. Los estudiantes del grupo 1 y 3 necesitan mejorar un poco más para acercarse a los líderes en calificaciones (grupos 4 y 5)

El perfilamiento a nivel de cada competencia puede verse con el gráfico adjunto.

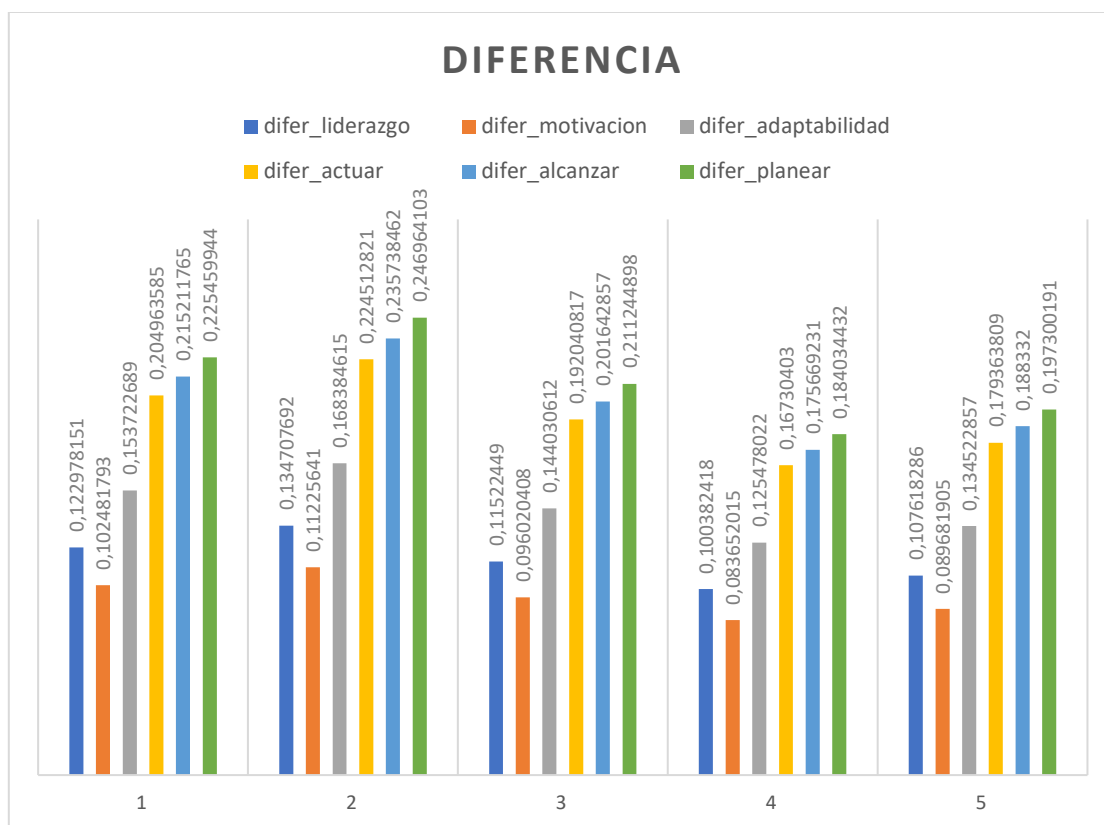


Figura 6. Comportamiento por competencias del Experimento 2

- En este grafico podemos evidenciar como las competencias fueron interiorizadas de forma diferente, las competencias que tuvieron menos fuerza al finalizar la parte de formación fueron: Planear, Alcanzar y Actuar (nótese que su diferencia con el ideal tiene los valores más altos en los 5 grupos). Las demás competencias como los son Liderazgo, Motivación y Adaptabilidad fueron las mejores al evidenciar que están más cerca del ideal (dado que la diferencia es menor), por tanto, se observa que los estudiantes las absorbieron y asimilaron de mejor manera durante la finalización de sus estudios de Maestría. Este hallazgo es común en los 5 grupos y podría sugerirse al programa de Maestría hacer un mejor monitoreo en la impartición de dichas competencias que pierden fuerza en relación con las restantes.

### Experimento 3: Segmentación Detallada de Estudiantes por competencias transcurrido un año luego de haber finalizado sus estudios de Maestría:

Este experimento es análogo al experimento 2, salvo que el análisis se hace en base a los resultados del instrumento de medición de competencias en los mismos 80 estudiantes, una vez transcurrido un año de su titulación de la maestría. El dendograma, boxplots y los centroides se muestran a continuación:

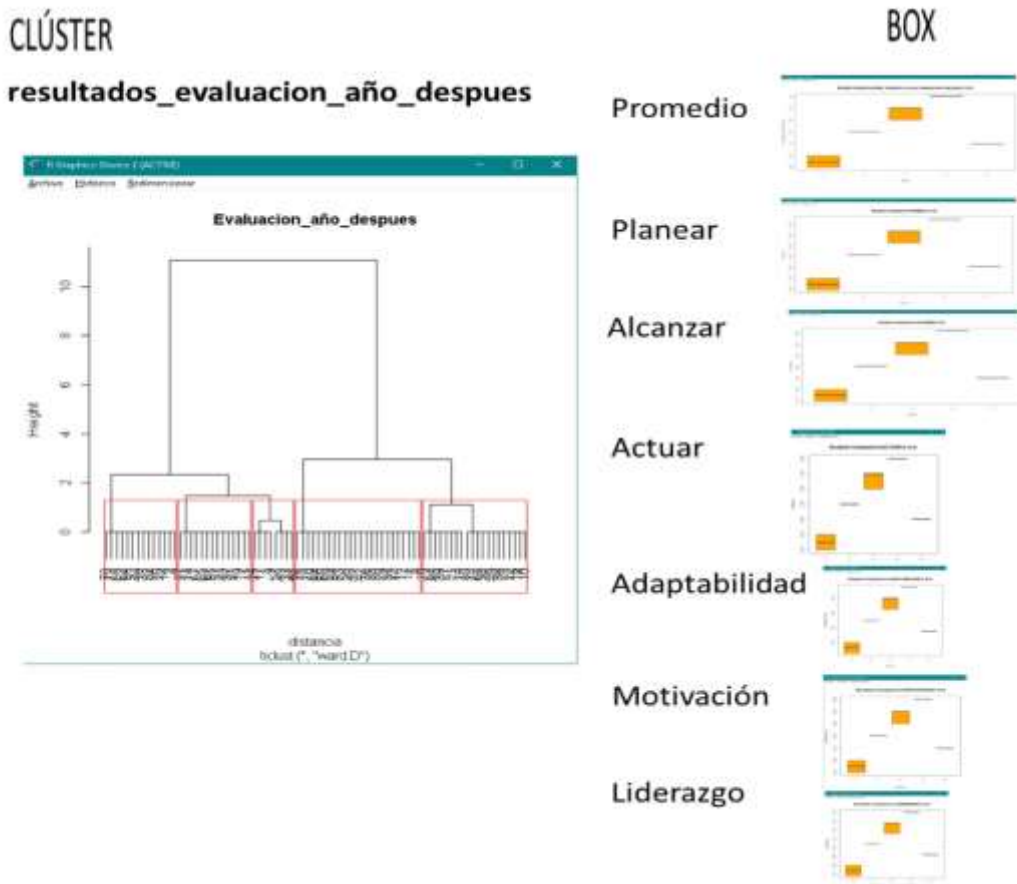


Figura 7. Dendograma y Boxplots para el Experimento 3

#### **Perfilamiento y hallazgos**

- Los estudiantes se clasifican nuevamente en 5 clústeres compuestos por: Grupo1 = 8, Grupo2 = 14, Grupo3 = 20 y Grupo4 = 24 y Grupo5 = 14, esto para un total de 80 estudiantes.

Los centroides de cada grupo o clúster se muestran en la siguiente figura (figura 8). En relación con dichos centroides y el boxplot se aprecia que, en base a promedio,

los estudiantes de los grupos 4 y 3 resultan los mejores, mientras que los del grupo 1 resultan los peores. Sin embargo, los valores en promedio de los 5 grupos pueden aproximarse a un valor de 4 para todos, por tanto, podría asumirse que en líneas generales, los 80 estudiantes durante su fase de estudios de Maestría absorben bastante bien las competencias no académicas ofrecidas por el programa (sugerencia o hallazgo ya advertido desde el experimento 1).

etiquetas	promedio_e	difer_lideraz	difer_motiv	difer_alcanz	difer_planea	difer_actuar	difer_adapt	Liderazgo	Motivacion	Actuar	Adaptabilid	Alcanzar	Planear
1	3,95	0,126	0,105	0,2205	0,231	0,21	0,1575	0,474	0,395	0,79	0,5925	0,8295	0,869
2	4,2	0,096	0,08	0,168	0,176	0,16	0,12	0,504	0,42	0,84	0,63	0,882	0,924
3	4,35	0,078	0,065	0,1365	0,143	0,13	0,0975	0,522	0,435	0,87	0,6525	0,9135	0,957
4	4,5	0,06	0,05	0,105	0,11	0,1	0,075	0,54	0,45	0,9	0,675	0,945	0,99
5	4,1	0,108	0,09	0,189	0,198	0,18	0,135	0,492	0,41	0,82	0,615	0,861	0,902

Figura 8. Centroides Experimento 3

- Si se calculan los promedios de la variable diferencia para las 6 competencias y grupo por grupo, puede notarse que los perfiles 3 y 4 agrupan a los estudiantes que más se acercan a la idealidad o absorción completa de las competencias, mientras que los peores se ubican en el grupo 1. Los estudiantes del grupo 2 y 5 necesitan mejorar un poco más para acercarse a los líderes en calificaciones (grupos 3 y 4).

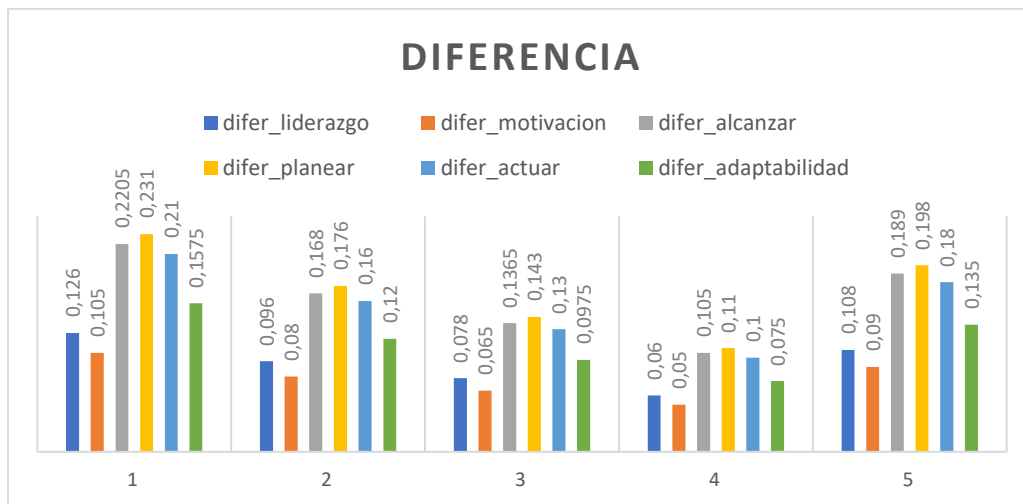


Figura 9. Comportamiento por competencias del Experimento 3.

En este grafico podemos evidenciar como las competencias fueron interiorizadas de forma diferente (similar al experimento 2), las competencias que tuvieron menos fuerza al año después de la obtención del título en la Maestría fueron: Planear, Alcanzar y Actuar. Las demás competencias como los son Liderazgo, Motivación y Adaptabilidad, fueron las mejores al evidenciar que están más cerca del ideal, de esta manera se observa que los estudiantes las mantienen y refuerzan al cabo de un año. Resulta entonces evidente que, al cabo de un año, el problema con las competencias Actuar, Planear y Alcanzar persiste, y esto debería en algún sentido, advertir a las autoridades del programa de Maestría en actuar o activar mecanismos



auditores para revisar cómo dichas competencias se están impartiendo en las disciplinas que conforman el pensum de estudio.

#### Experimento 4: Cuantificación de la transición de estudiantes de un perfil a otro después de un año de graduado.

En los experimentos anteriores se puede hacer un análisis de perfiles identificando hallazgos y descriptores básicos para 2 situaciones en particular, pero de forma separada y excluyente entre los 2 momentos asociados a la recolección de la data, es decir, la fotografía en la asimilación de las competencias durante los estudios y la consecuente fotografía de las competencias y su "status" de asimilación después de haber culminado la maestría.

Considerando los mismos 80 estudiantes y sus etiquetados previos obtenidos en los experimentos 2 y 3, se procedió a inducir a través de un análisis simple de frecuencias relativas (conteo), la transición de que un estudiante habiendo pertenecido a un perfil "X" durante su fase de estudios se traslade a un perfil "Y" al cabo de un año de haber culminado los estudios en la maestría. Este indicador puede ayudar a evaluar de forma más simple y directa la propensión a que las competencias mejoren, empeoren o permanezcan estables al cabo de un año de posevaluación como consecuencia de haber perfilado claramente en 5 perfiles la asimilación de las competencias durante y después de los estudios.

Para realizar este análisis se procedió a clasificar los 5 grupos del experimento 2 y los 5 grupos del experimento 3 en 5 niveles discriminantes u ordinales en base a los hallazgos descritos en los experimentos precedentes. El etiquetado de esos grupos se hizo sustancialmente con la variable promedio para facilitar la experimentación. De esta manera se reclasificaron los grupos en base a un desempeño por niveles quedando así: Alto (A), Medio Alto (MA), Medio (M), Medio Bajo (MB) y Bajo (B), respectivamente. Es importante notar que dichos niveles son netamente cualitativos y se usan como una primera aproximación que pudiese mejorarse con ulteriores experimentos, mayor cantidad de estudiantes (datos) o reglas de negocio establecidas por los criterios de evaluación internos de los responsables de la maestría. Es por ello, que un grupo etiquetado o que tenga nivel "BAJO" no significa literalmente que es pésimo, sino que es el que menor desempeño tuvo en la muestra (dicho valor no necesariamente expresa una situación extrema o muy negativa), por ende, debe medirse en términos de una cierta escala o medida cuantitativa que debe establecer el dueño o usuario del problema.

Los perfiles según este desempeño y para los 2 experimentos quedaron así:

<b>Perfiles/grupos “Cursando Estudios” Experimento 1</b>	<b>“Etiquetado por Desempeño en la asimilación de Competencias”</b>
Grupo 1	Medio Bajo (MB)
Grupo 2	Bajo (B)
Grupo 3	Medio (M)
Grupo 4	Alto (A)
Grupo 5	Medio Alto (MA)

<b>Perfiles/grupos “Un año después” Experimento 2</b>	<b>“Etiquetado por Desempeño en la asimilación de Competencias”</b>
Grupo 1	Bajo (B)
Grupo 2	Medio (M)
Grupo 3	Medio Alto (MA)
Grupo 4	Alto (A)
Grupo 5	Medio Bajo (MB)

Con los grupos etiquetados por niveles, se procedió a construir una tabla cruzada y a contabilizar las transiciones para cada una de las celdas. Esto se hizo con un barrido de la transición reportada por cada uno de los 80 estudiantes desde su evaluación mientras cursaba la maestría versus su posición una vez transcurrido un año de culminación de estudios.

Los resultados se muestran en la tabla adjunta:

	1 año después					
Grado	Grupo1 (Bajo)	Grupo2 (Medio)	Grupo3 (MedioAlto)	Grupo4 (Alto)	Grupo5 (MedioBajo)	Total
Cluster1 (MedioBajo)	2	4	3	3	2	14
Cluster2 (Bajo)	0	1	6	2	4	13
Cluster3 (Medio)	2	4	7	4	5	22
Cluster4 (Alto)	1	1	1	4	1	8
Cluster5 (MedioAlto)	3	4	3	11	2	23
<b>Total</b>	<b>8</b>	<b>14</b>	<b>20</b>	<b>24</b>	<b>14</b>	<b>80</b>

Puede apreciarse de la tabla anterior que de un total 14 estudiantes pertenecientes a un nivel de asimilación medio bajo (MB) durante sus estudios empeoraron (pasaron a bajo) o se mantuvieron estables (medio bajo) en una proporción o frecuencia de 2 + 2 casos es decir 4 de un total de 14, por lo que 10 estudiantes en dicha situación mejoraron la asimilación de las competencias al cabo de un año.

Este tipo de análisis le permite al dueño del problema tener indicios sobre la trayectoria evolutiva en la asimilación de las competencias aprendidas e impartidas dentro de la maestría.

Por citar un par de ejemplos más, puede verse que los 13 estudiantes con peor desempeño lograron mejorar su asimilación al cabo de un año en un 100% de los casos. No obstante, solo el 50% (4 de 8) de estudiantes que mejor habían asimilado las competencias durante sus estudios (Alto) pudieron mantener ese nivel al cabo de un año, mientras que el restante de ellos (4) empeoraron su situación.

### 3.6 Etapa de conocimiento (Visualización)



Figura 1 KDD (Knowledge Discovery in Databases)

El visual analytics es una representación gráfica de la información y los datos. Mediante el uso de elementos visuales, la visualización de datos ofrece una manera accesible para detectar y comprender las tendencias, los valores atípicos y los patrones en los datos. (Según link Tableau, 2003). La visualización de datos se aplica en prácticamente todos los campos del conocimiento. En diversas disciplinas se utilizan técnicas y herramientas informáticas para modelar eventos complejos y

visualizar fenómenos que no se pueden observar directamente, como los patrones climáticos, las condiciones médicas, relaciones entre variables o las relaciones matemáticas, los movimientos de un activo financiero, entre muchos otros.

Los resultados y perfiles obtenidos en esta investigación se exportaron a un tablero (visualizador) con el objeto de que el usuario final (Maestría en Gerencia de Ingeniería de la Universidad de La Sabana) pudiese ver de manera más amigable los hallazgos obtenidos con el proceso de minería de datos.

La visualización se muestra en Microsoft Power Bi. A continuación, se listan pantallas de dicho visualizador. Mayores detalles se incluyen en los anexos del documento. Asimismo, esta investigación incluirá el enlace de dicho dashboard como fuente de acceso a consultar por los usuarios que lo consideren pertinentes y sean autorizados por el cliente final (Universidad de La Sabana, Colombia)



Figura 10. visualización de competencias de los estudiantes durante y después de la maestría

El análisis nos muestra una nota promedio de todos los estudiantes de 4.03 al estar en la etapa de formación, un año después se realiza una nueva evaluación, la cual tiene una variación, pero en promedio está en 4.29 dicha nota lo que indica que los estudiantes al cabo de un año mejoran sus competencias en términos generales, mas no específicos como se detectó en los experimentos 2 y 3.

Demostrando ser un líder altamente efectivo, analítico e innovador que interpreta problemas y necesidades, establece estrategias y diseña soluciones organizacionales aplicando principios y técnicas de la ingeniería, para lo cual, involucra tecnologías e integra recursos de sus áreas de trabajo en forma armónica, con sentido ético y comprometido con el medio ambiente.



**ANÁLISIS DE LAS COMPETENCIAS DE LOS ESTUDIANTE GRADO EN RELACIÓN CON LOS VALORES DE LAS COMPETENCIAS IDEALES**

codigo\_estudiante

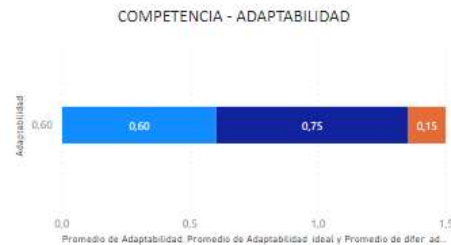


Figura 11. Competencias estudiante grado v1



**ANÁLISIS DE LAS COMPETENCIAS DE LOS ESTUDIANTE GRADO EN RELACIÓN CON LOS VALORES DE LAS COMPETENCIAS IDEALES**

codigo\_estudiante



Figura 12. Competencias estudiante grado v2

En el análisis de competencias de los estudiantes graduados se evidencia que las competencias que continúan en su punto más alto son: Motivación, Liderazgo y Adaptabilidad.

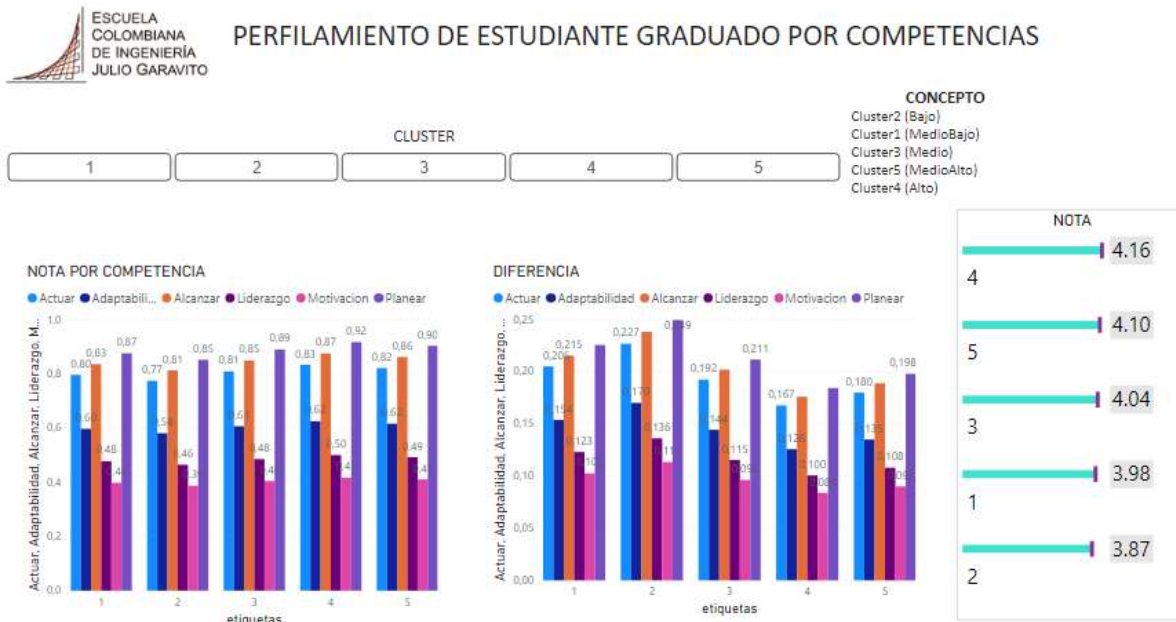


Figura 13. Visualización estudiante graduado por competencias



Figura 14. Visualización resultados de perfilamiento

## Capítulo 4 – CONCLUSIONES Y RECOMENDACIONES

Una vez finalizada la investigación, se listan a continuación las conclusiones, aportes, recomendaciones y líneas futuras de investigación.

- La contextualización del problema junto a la revisión del estado del arte y su solución mediante la puesta en marcha de un diseño de experimentos sencillo pero robusto, permitió evaluar y detallar el estatus sobre el grado de asimilación de las seis competencias “no académicas” impartidas dentro de las disciplinas impartidas en el programa de Maestría en Gerencia de Ingeniería de la Universidad de la Sabana (Bogotá, Colombia).
- La minería de datos aplicada a la base muestral de 80 estudiantes con técnicas de clustering, posprocesamiento de perfiles y visualización de resultados obtuvo hallazgos e indicadores que pueden soportar procesos de toma de decisiones a las autoridades encargadas de la Maestría. Al respecto debe citarse lo siguiente:
- Desde un punto de vista global (basado en el promedio de las 6 competencias), los estudiantes suelen absorber bastante bien las competencias en su último año de Maestría y posterior a un año de la culminación de estudios. Solo 8 de 80 estudiantes suelen empeorar (Experimento 1) la absorción o mantenimiento de las competencias, por tanto, puede indicarse que existe una tasa alta de efectividad de mantenimiento o refuerzo de las 6 competencias en su conjunto, es decir 90%.
- Los experimentos 2 y 3, refuerzan aún más lo indicado previamente, pero con matices (mayor conocimiento detallado). En el experimento 2 (Evaluación durante la maestría), solo 13 de 80 estudiantes se etiquetan en un clúster con el peor desempeño (13/80 → aprox. 16%). Sin embargo, es un clúster relativo ya que el promedio a pesar de ser el peor, es muy cercano a 4 (el máximo o mejor grupo lo reporta en 4.17), por tanto, esos 13 estudiantes tampoco muestran un sesgo excesivamente negativo. Algo similar ocurre un año después con la evaluación posterior a la titulación (Experimento 3). Allí solo 8 de los 80 estudiantes (10%) caen en el peor grupo, pero aquí nuevamente su desviación en comparación con el resto de los grupos es menor, a menos que el analista, le de un peso excesivo al efecto de los decimales en la puntuación. Esto hace concluir, que en líneas globales se tienen las siguientes dos afirmaciones (a): existe cierta homogeneidad que hace pensar que los estudiantes asimilan en promedio las competencias y (b) Al cabo de un año y explorando los 2 experimentos por separado, se aprecia que los estudiantes refuerzan las competencias manteniéndolas o mejorándolas. Nótese que de 13 estudiantes que se etiquetan en el peor grupo, se pasa a únicamente 8 estudiantes en el peor grupo al cabo de un año. La mejora es evidente.
- Retomando los experimentos 2 y 3, se extrae otro hallazgo más particular y discriminante. Tanto en el instante “cursando la maestría” como “después de

un año de la titulación”, las competencias Planear, Alcanzar y Actuar poseen un efecto sutilmente diferenciado y contrario en relación con las tres competencias restantes (Adaptarse, motivar y liderar). Las tres primeras tienen menor desempeño. Se recomienda entonces a los decisores o actores influyentes en la maestría a activar mecanismos de control o de monitoreo, en particular en cómo se están impartiendo o evaluando dichas competencias, con el fin de validar la hipótesis “de menor” desempeño de las competencias Actuar, Planear y Alcanzar. Su efecto suele diluirse en el tiempo y por tanto su asimilación podría ser menor a las competencias de Adaptarse, motivar y liderar.

- Retomando las definiciones de las competencias “con menor desempeño” podemos sugerir que algo dentro de los contenidos programáticos podría estar fallando o quizás no está siendo bien impartido en la docencia. Así se tiene: (a) “**Planear** los procesos relacionados con la investigación, diseño, desarrollo y producción de productos y servicios para garantizar altos niveles de productividad y agregar valor en cada eslabón de la cadena haciendo uso de herramientas de ingeniería en cualquier tipo y tamaño de organización”. Se sugiere reevaluar o modificar los criterios con lo que está competencia se inserta en las asignaturas de la maestría. Es importante recalcar que el efecto de esa competencia va ligado a una enseñanza más técnica y por ello, es quizás más fácil de mejorar. (b) De manera similar y retomando las definiciones para “alcanzar” y “actuar”, se aprecia que la definición intrínseca de las mismas hace que corregirlas sea un proceso más fácil de explorar y hallar que en las otras tres competencias que se han absorbido muy bien. “**Alcanzar** el cumplimiento de objetivos organizacionales que garantice un sano desarrollo y crecimiento a través del trabajo interdisciplinario; interpretando y aplicando correctamente indicadores financieros, de operaciones, sociales, del medio ambiente, de mercadeo, de sistemas complejos y de gerencia del conocimiento”. “**Actuar** ética y responsablemente tanto en la organización como en la comunidad o país donde se desempeña con conciencia ambiental y sostenible siguiendo regulaciones y leyes, y respetando la cultura organizacional y política para garantizar que las consecuencias de las propuestas generan beneficios a todas las partes involucradas sin riesgo alguno”.
- El punto anterior conlleva a puntualizar que las 3 mejores competencias absorbidas o asimiladas (Adaptarse, Motivar y Liderar) están siendo mejor impartidas en el programa, a pesar de que pudiesen etiquetarse como más subjetivas o blandas, si les compara con las otras 3, que en términos prácticos son menos blandas o técnicas.
- El experimento 4 permite evidenciar la transición de un estado a otro en el grado de asimilación de las consecuencias. Si bien es un experimento exploratorio y no inferencial, permite denotar que los estudiantes al cabo del tiempo mejoran o asimilan en mayor grado las competencias.



En esta investigación también se puede concluir que:

- Se ha cumplido con los objetivos propuestos. Los hallazgos anteriormente mencionados pueden usarse para validar con el usuario final políticas de seguimiento, monitoreo y control dentro de la maestría usada como caso de estudio o reporte de caso.
- El desarrollo de un marco experimental y sencillo arroja resultados basados en grupos o perfiles que facilitan el trabajo y el análisis transversal que pudiese complicarse si se hace por cada estudiante por separado.
- Se entrega al usuario final, un dashboard preliminar que puede usarse para describir resultados de manera amigable. Dashboard que puede mejorarse en el tiempo y enriquecerse con mayor volumetría de datos (más estudiantes, por ejemplo).

Las conclusiones y aportes anteriores conllevan a sugerir posibles líneas de investigación a futuro:

- A través del enfoque con minería de datos propuesto y cuando la volumetría de los datos lo permita se podrían realizar modelos de enfoque predictivo en varios niveles discriminantes. A saber, modelos para: (a) Clasificar automáticamente nuevos estudiantes en base a las etiquetas de clustering halladas en esta investigación. (b) Usar el punto (a) para predecir la etiqueta que tendrá un estudiante nuevo al cabo de un año de culminar los estudios y (c) proponer modelos de Markov o de transición de probabilidades de estado basados en el experimento 4 realizado en esta investigación, agregando un instante adicional al problema, (estudiantes a inicio-mitad de maestría)
- Similarmente, pueden hacerse estudios de ingeniería bajo la óptica de sistemas organizacionales que incluyan análisis exhaustivos de ingeniería reversa. Esto podría determinar las causas del menor desempeño en las competencias definidas como Planear, Actuar y Alcanzar. Para ello se involucrarían el histórico de las calificaciones que han dado los estudiantes hacia los profesores, así como la revisión de los contenidos programáticos de las asignaturas y la ponderación que actualmente maneja la maestría para cuantificar la “nota” y el “peso” de cada competencia.
- Finalmente, los aportes y hallazgos de este trabajo se han resumido y adaptado para someterse a revista científica dentro de la lista disponible en Colombia para revistas indexadas y arbitradas del repositorio Latindex o Scielo.

## Referencias Bibliográficas

1. ArcMap, (2016). Comprender el análisis de distancia euclidiana, Recuperado de
2. <https://desktop.arcgis.com/es/arcmap/10.3/tools/spatial-analyst-toolbox/understanding-euclidean-distance-analysis.htm>
3. Aristizábal J, (2017). Diseño y aportes de un modelo para minería de datos educativos en aulas de educación media de carácter presencial. Universidad Santo Tomás, vicerrectoría de universidad abierta y a distancia, facultad de educación, doctorado en educación. Tesis doctoral.
4. Ayala M, (2010). Financiamiento de la educación superior en Colombia reflexiones para un próximo futuro. Recuperado de
5. [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S0185-27602010000400006](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-27602010000400006)
6. Bedoya, O. M., López Trujillo, M. & Marulanda Echeverry, C. E. (2016). Minería de datos en egresados de la Universidad de Caldas. Revista Virtual Universidad Católica del Norte, 49, 110-124.
7. Cano E, (2008). La evaluación por competencias en la educación superior. Revista de Currículum y Formación de Profesorado, vol. 12, núm. 3, pp. 1-16 Universidad de Granada, España
8. Clinic-cloud, Minería de datos, (2017). Recuperado de <https://clinic-cloud.com/blog/data-mining-que-es-definicion-mineria-de-datos/>
9. Daza A, (2016). Un modelo basado en árboles de decisión para predecir la deserción estudiantil en la Educación Superior Privada. UCV - SCIENTIA, Vol. 8, N.º 1, págs. 59-73.
10. De La Hoz, E, Mendoza, A., Herrera, T. (2018). Aplicación de Minería de Datos para la Clasificación de Programas Universitarios de Ingeniería Industrial Acreditados en Alta Calidad en Colombia. Información Tecnológica. 29. DOI: 10.4067/S0718-07642018000300089.
11. Duck2 - Estrategias de trading, (2019). Algoritmos de Data Mining para agrupar datos – Clustering Jerárquico, Recuperado de
12. <https://estrategiastrading.com/clustering-jerarquico/>
13. Enciclopedia colaborativa en la red cubana, (2012). Clustering. Recuperado de <https://www.ecured.cu/Clustering>
14. Escudero, J, La innovación y la organización escolar. Recuperado de [http://www.terras.edu.ar/biblioteca/17/17GSTN\\_Escudero\\_Unidad\\_3.pdf](http://www.terras.edu.ar/biblioteca/17/17GSTN_Escudero_Unidad_3.pdf)
15. Garassino A, Syloper C. (2015). ¿Para qué sirve la minería de datos o data mining? Recuperado de <https://www.syloper.com/blog/recursos/para-que-sirve-la-mineria-de-datos/>
16. H, Paula Pérez, Cristhian Parra, Paula Ortiz M, Liliana Betancourt, Olga C, Peter A, Graciela & K, Andrea. (2015). Relación entre el bienestar y el rendimiento académico en alumnos de primer año de medicina. Revista médica de Chile. 143. 930-937.
17. Landa, J, (2016). Tratamiento de los datos, Recuperado de <http://fcojlanda.me/es/ciencia-de-los-datos/kdd-y-mineria-de-datos-espanol/>

18. López, S. (2019). El modelo de competencias en educación superior, análisis, funcionalidad e importancia, Recuperado de
19. <https://revista.universidadabierta.edu.mx/2019/04/22/el-modelo-de-competencias-en-educacion-superior-analisis-funcionalidad-e-importancia/>
20. Minerva Data Mining, (2018). Recuperado de <https://mnrva.io/kdd-platform.html>
21. Ministerio de Educación Colombia, (2010). Diseño de currículos por competencias, anexo, contexto normativo, El Plan Nacional de Desarrollo 2006 – 2010, Recuperado de [https://www.mineducacion.gov.co/1759/articles-299212\\_archivo\\_pdf.pdf](https://www.mineducacion.gov.co/1759/articles-299212_archivo_pdf.pdf)
22. Ministerio de Educación Colombia, (2020). Colombia le apuesta a la innovación y la tecnología para elevar la calidad en la educación, Recuperado de
23. [https://www.mineducacion.gov.co/1759/w3-article-358498.html?\\_noredirect=1](https://www.mineducacion.gov.co/1759/w3-article-358498.html?_noredirect=1)
24. Montoya Lince Adrián, Vargas Bonilla Jesús Francisco, Contreras Olivares Lyda Yaneth, (2015). Educational Data Mining (EDM) para la determinación de comportamientos en estudiantes de ingeniería del modelo ude@, Universidad de Antioquia, Medellín, Colombia, Recuperado de
25. <https://acofipapers.org/index.php/eiei2015/2015/paper/viewFile/1104/391>
26. Orientación Universia, (2018). Portal de Universia, habilidades blandas, Recuperado de
27. <https://orientacion.universia.net.co/infodetail/consejos/orientacion/que-son-las-habilidades-blandas-4545.html>
28. Significados, (2017). Significado de Competencias, Recuperado de <https://www.significados.com/poesia/>
29. Sinergia e Inteligencia de negocios, (2007). Recuperado de [https://www.sinnexus.com/business\\_intelligence/datamining.aspx](https://www.sinnexus.com/business_intelligence/datamining.aspx)
30. Sotelo J, (2006). Análisis de Perfiles: Diagnóstico, capacitación y evaluación de personal, Boletín electrónico sobre desarrollo de RRHH. Recuperado de <https://groups.google.com/forum/#!topic/cartadesotelo/cYmaWeSYN5c>
31. Tableau, (2003). Data visualización, Guía de visualización de datos para principiantes: definición, ejemplos y recursos de aprendizaje, Recuperado de <https://www.tableau.com/es-mx/learn/articles/data-visualization#:~:text=La%20visualizaci%C3%B3n%20de%20datos%20es,y%20patrones%20en%20los%20datos>
32. Timaran R, Calderón A, Jiménez J, (2013). La minería de datos como un método innovador para la detección de patrones de deserción estudiantil en programas de pregrado en instituciones de educación superior, Recuperado de <https://acofipapers.org/index.php/acofipapers/2013/paper/viewFile/211/112>
33. Timaran, S, Hernández, I, Caicedo, S, Hidalgo A, Alvarado J, (2016). Descubrimiento de patrones de desempeño académico con árboles de decisión en las competencias genéricas de la formación profesional. Fondo editorial ediciones universidad cooperativa de Colombia, Recuperado de <http://ediciones.ucc.edu.co/index.php/ucc/catalog/download/36/40/220-1>

34. Tiraman, R, Calderón, A, Jiménez, J, (2013). Aplicación de la minería de datos en la extracción de perfiles de deserción estudiantil, Ventana informática, universidad de Manizales.
35. Universidad de Valencia, Uv.es, (2011). Clustering: métodos de análisis, Recuperado de <https://www.uv.es/ceaces/multivari/cluster/metodos.htm>
36. Vergara, C., Oviedo, A., Carmona, C., Vélez, G., & Amón, I (2018). Estilos de aprendizaje y minería de datos: un estudio preliminar en el contexto universitario. Ingeniería e Innovación, 6(1), 13-18
37. Yamao, E, Celi, L, Campos, R. (2018). Predicción del rendimiento académico mediante minería de datos en estudiantes del primer ciclo en una universidad peruana. Campus. 23. 151-160.
38. Yturalde E, (2019). Gestión del talento, habilidades blandas – competencias sociales, Recuperado de <http://www.habilidadesblandas.com/>

## Anexos

Anexo Único: Captura de pantalla del dashboard exploratorio para visualizar resultados al cliente/usuario final.



Figura 10. Visualización de competencias de los estudiantes durante y después de la maestría

Se describen los promedios generales por competencias de los estudiantes hallando los promedios durante y después de finalizado el programa educativo.

**ANÁLISIS DE LAS COMPETENCIAS DE LOS ESTUDIANTE GRADO EN RELACIÓN CON LOS VALORES DE LAS COMPETENCIAS IDEALES**

codigo\_estudiante  
est0000014mgj



Figura 11. Competencias estudiante grado v1

**ANÁLISIS DE LAS COMPETENCIAS DE LOS ESTUDIANTE GRADO EN RELACIÓN CON LOS VALORES DE LAS COMPETENCIAS IDEALES**

codigo\_estudiante  
est0000014mgj



Figura 12. Competencias estudiante grado v2

Se obtienen los valores de las competencias y se comparan con respecto al valor ideal de cada competencia, para de esta manera determinar cuales son las competencias mejor absorbidas.

PERFILAMIENTO DE ESTUDIANTES

codigo\_estudiante  
Todas

CLUSTER  
1 2  
3 4

- CONCEPTO  
Cluster2 (Bajo)  
Cluster1 (MedioBajo)  
Cluster3 (Medio)  
Cluster5 (MedioAlto)  
Cluster4 (Alto)

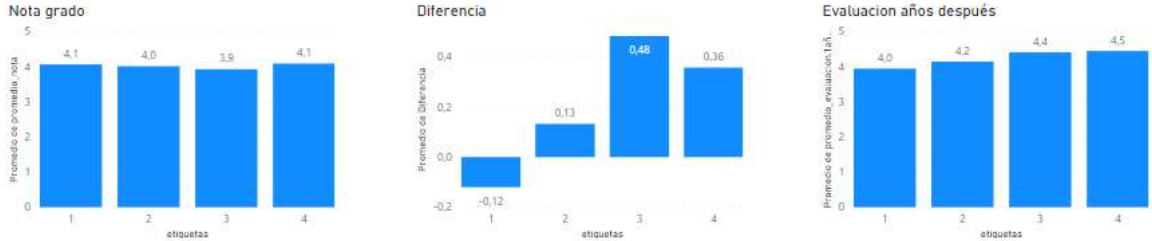


Figura 15. visualización perfilamiento de estudiantes en experimento 1

En el perfilamiento evidenciamos que los estudiantes se ubican en 4 clúster considerando el perfilamiento básico del Experimento1.

PERFILAMIENTO DE ESTUDIANTE GRADUADO POR COMPETENCIAS

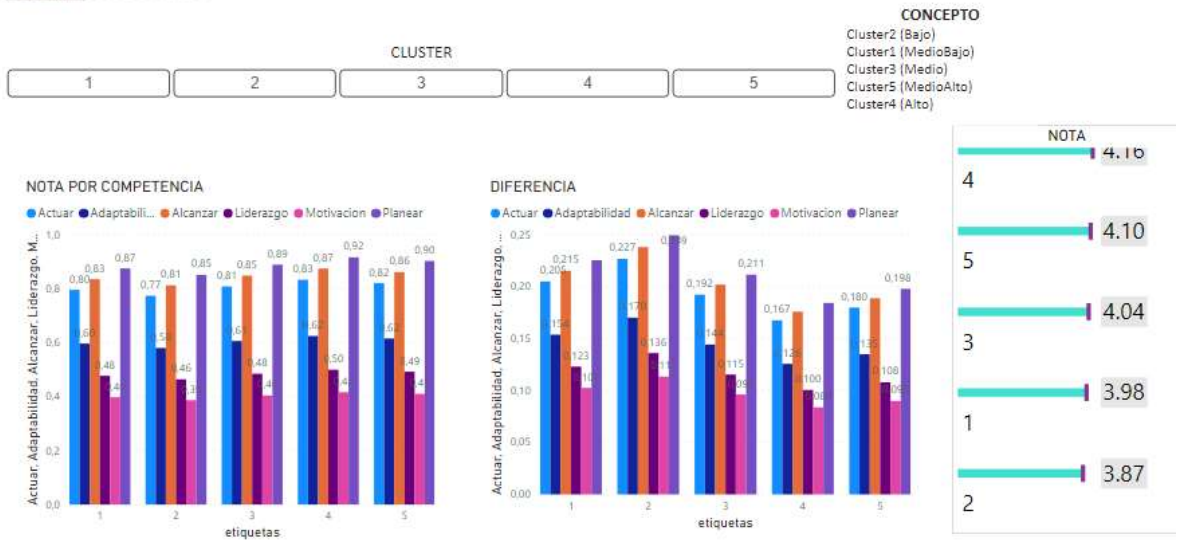


Figura 13 visualización estudiante graduado por competencias

Al realizar el perfilamiento por clúster se puede ver el comportamiento de las competencias en cada uno de los estudiantes y la diferencia que manejan sobre la nota ideal.

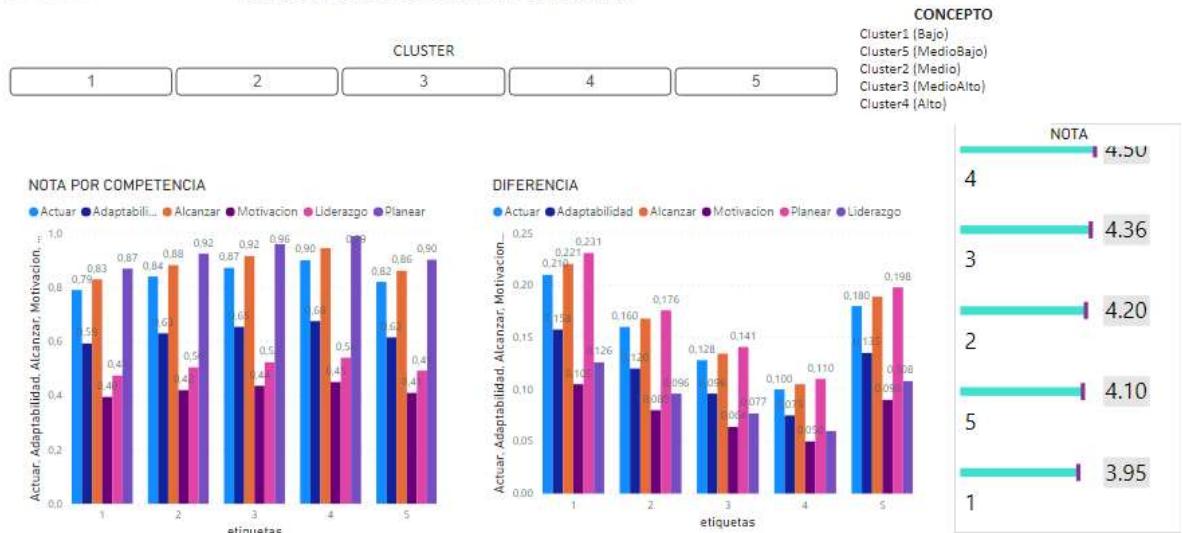


Figura 16. Visualización un año después

Al realizar el perfilamiento por clúster se puede ver el comportamiento de las competencias en cada uno de los estudiantes después de un año de la finalización del programa académico y la diferencia que manejan sobre la nota ideal.

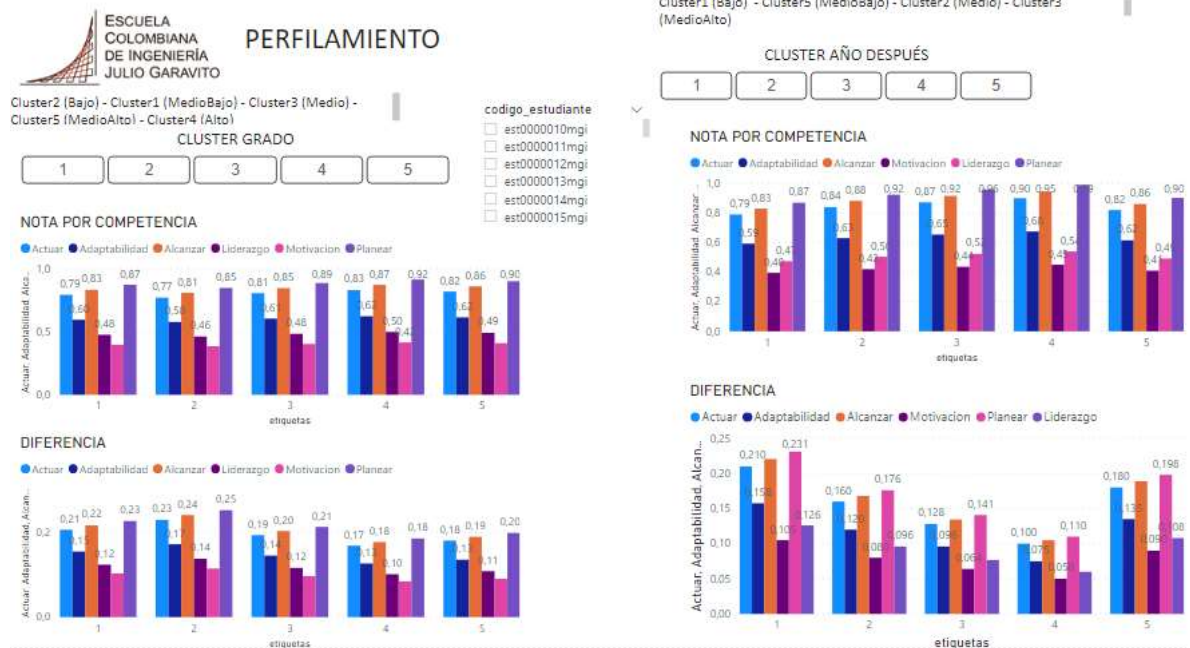


Figura 14 visualización resultados de perfilamiento



Al realizar el perfilamiento comparativo de los clústeres obtenidos en los instantes de evaluación realizados, es decir, cursando o finalizando el programa académico y la evaluación un año después, posterior a la culminación de estudios. Aquí se puede ver el comportamiento de las competencias en cada uno de los estudiantes, el cambio o detención en el clúster y la diferencia que manejan sobre la nota ideal.