



Universidad Autónoma de Querétaro  
Facultad de Informática

Modelo Inteligente de Tutorías (MIT)  
en el Instituto Tecnológico de Morelia

TESIS

Que como parte de los requisitos para obtener el  
grado de Maestro en  
Ingeniería de Software Distribuido

Presenta

Graciela Isolina Martínez Avila

Santiago de Querétaro, Junio del 2015



Universidad Autónoma de Querétaro  
Facultad de Informática  
Maestría en Ingeniería de Software Distribuido

**Modelo Inteligente de Tutorías (MIT)  
en el Instituto Tecnológico de Morelia**

TESIS

Que como parte de los requisitos para obtener el  
grado de Maestro en Ingeniería de Software Distribuido

Presenta:

**Graciela Isolina Martínez Avila**

Dirigido por:

**M.C. Cristhian Torres Millarez**

SINODALES

**M.C. Cristhian Torres Millarez**  
Presidente

Firma

**M.I.S.D. Jesús Armando Rincones**  
Secretario

Firma

**M.S.I. Elisa Morales Portillo**  
Vocal

Firma

**M.S.I. Ernesto Rubalcava Durán**  
Suplente

Firma

**Dr. Ubaldo Chávez Morales**  
Suplente

Firma

**MISD. Juan Salvador Hernández Valerio**  
Director de la Facultad

**Dra. Ma. Guadalupe Flavia Loarca Piña**  
Directora de Investigación y Posgrado

Centro Universitario  
Querétaro, Qro.  
Junio, 2015  
MÉXICO

## RESUMEN

Este documento de tesis, propone un Modelo Inteligente de Tutorías, utilizando técnicas de minería de datos, que contribuya a disminuir la deserción y reprobación de estudiantes a nivel superior, aumentando la eficiencia terminal, a partir de la creación de cuatro perfiles de estudiantes, los cuales se congregarán en grupos con características y necesidades semejantes, para que sean guiados por tres perfiles de tutores, según corresponda, dependiendo del semestre que estén cursando. Todos los perfiles se crearán a partir de las bases de datos generadas; para los alumnos, con el examen de ingreso, estudios psicopedagógicos, bitácoras de tutorías, asesorías y de las evaluaciones que se vayan registrando en su trayectoria académica, de tal manera que el mismo sistema vaya consolidando dichos perfiles a través del tiempo. En el caso de los profesores se toman como bases de datos, no sólo su currículum, sino su desempeño académico dentro y fuera de la Institución, su experiencia laboral, y exámenes psicopedagógicos, de tal manera que el sistema genere conocimiento significativo de los dos actores del proceso enseñanza – aprendizaje, para la toma de decisiones y coadyuve a que las Instituciones de Educación Superior, generen capital humano de calidad. La metodología utilizada se enfocó en el Perfil del Aspirante, aplicando la técnica predictiva Estadística Descriptiva y las técnicas descriptivas como Análisis de Componentes Principales, k – means y Análisis Discriminante Lineal. El caso de estudio práctico fue analizar los datos de la generación de estudiantes que ingresó en junio 2011, a Ingeniería en Tecnologías de la Información del Instituto Tecnológico de Morelia, a partir de la aplicación del Examen de Ingreso (EXANI – II), que aplica el Centro Nacional de Evaluación para la Educación Superior (CENEVAL) y las evaluaciones del primer semestre. Los resultados fueron comparados con los Perfiles de Egreso de las carreras en Tecnologías, que la Asociación Nacional de Instituciones de Educación en Tecnologías de Información (ANIEI) propone, concluyendo que existen coincidencias entre perfiles. Este modelo es aplicable a cualquier Institución de Educación Superior, lo que varía son los perfiles generados, dado que dependerá no sólo de las bases de datos para cada caso, sino de su circunstancia y necesidad académica.

**(Palabras clave:** tutoría académica, minería de datos)



## SUMMARY

This study proposes an Intelligent Tutorial Model based on data mining techniques, which contributes to reduce grade failure and school drop-out rates in students in higher education. Thus increasing graduation rate based on the creation of four profiles of students, which are grouped on the basis of their similar characteristics and needs in order to be led by three different profiles of tutors as the case may be, depending on what semester they are studying. All profiles are created on the basis of the data base generated; for students, the factors considered are the entrance examination, psychological studies, tutorial, advisory and assessment records during their educational path so that the system consolidates the profile through time. In the case of professors not only the data base, but also their curriculum, their academic performance within and outside the institution, their work experience and psychological examinations are taken into consideration so that the system generates a significant knowledge about both actors in the teaching-learning process for decision making and contributing with higher education institutions to generate higher quality human capital. Methodology was based on the Candidates' Profile using the Descriptive Statistics predictive technique and descriptive techniques such as Principal Components Analysis, k – means, and the Lineal Discriminate Analysis. The practical study case consisted in analyzing the data from the generation of students which joined the school in June 2011 in the engineering degree in information technology at the Technological Institute in Morelia, beginning with the application of the examination (EXANI – II) by the National Center of Assessment for Higher Education) (CENEVAL in Spanish) and the first semester evaluations. Results were compared with the profile of graduation of the Degrees in Information Technology, proposed by the National Association of Information Technology Educational Institutions (ANIEI in Spanish), which lead us to conclude that both profiles match. This model can be applied to any Higher Education Institution, what varies is the profiles as they will depend not only on the data base for each case, but also on their circumstances and academic needs.

(Key Words: academic tutors, data mining)



---

**Dedico este esfuerzo a mis hermosos hijos Gabriel y Gisela  
y por supuesto a mi amado esposo  
Gerardo Maldonado Vidal**

---

## AGRADECIMIENTOS

La motivación inicial de este trabajo fue a través de la propuesta del M.C. Martin Larios Osorio, profesor investigador de la Universidad Tecnológica de Querétaro, quien fungió como maestro y asesor del protocolo de tesis en las materias de Proyecto de Tesis I y II del Posgrado de Ingeniería en Sistemas Distribuidos de la Facultad de Informática de la UAQ, quien me inspiró para trabajar en el tema de Minería de Datos y al que le agradezco sus valiosas aportaciones.

Posteriormente y de manera desinteresada el M.C. Cristhian Torres Millarez, Jefe del Departamento de Servicios Escolares del Instituto Tecnológico de Morelia, asesor de mi tesis de posgrado, retoma mi proyecto, y con total disposición, tolerancia y amplios conocimientos del tema, fue mi guía, acompañando mi proceso hasta ahora culminado. Gracias de corazón por haber creído en mí.

No menos importante quiero agradecer a la L.I. Ana María Díaz Álvarez, Asistente de la Jefatura de División de Estudios de Posgrado de la Facultad de Informática de la UAQ, por ser tan amable al facilitarme todos los trámites administrativos de mi proceso de titulación, así como a mis sinodales que depositaron en mí la confianza al darme su voto aprobatorio.

También agradezco al M.C. José Luis Campos Villafuerte, amigo y compañero de la Universidad Tecnológica de Morelia, que de manera desinteresada revisó la forma y redacción de este documento, en un sentido crítico, metódico y responsable.

En general agradezco y bendigo a todas y cada una de las personas que de alguna manera estuvieron presentes desde el inicio hasta el desarrollo de esta tesis y que con sus comentarios, consejos, recomendaciones y porque no, algunas llamadas de atención, me motivaron a concluir este reto y lograr mi grado de Maestría.

# INDICE

	<b>Página</b>
Resumen	i
Summary	ii
Dedicatorias	iii
Agradecimientos	iv
Índice	v
Índice de tablas	vi
Índice de figuras	vii
I. INTRODUCCIÓN	1
II. REVISIÓN DE LITERATURA	3
2.1 La Educación Superior y el Gobierno Federal.	3
2.2 La Educación Superior y los CIEES.	6
2.3 La Educación Superior y COPAES.	9
2.4 La Educación Superior y CENEVAL.	11
2.5 La Educación Superior y las Tutorías.	12
2.6 El Instituto Tecnológico de Morelia (ITM).	21
2.7 El Instituto Tecnológico de Morelia y las Tutorías.	23
2.8 La Minería de Datos.	29
2.8.1 Tipos de Datos para Minería de Datos.	30
2.8.2 Técnicas de Minería de Datos.	34
2.9 La minería de datos y el proceso de descubrimiento del conocimiento.	36
2.9.1 Relación con otras disciplinas.	37
2.9.2 Aplicaciones.	39
2.10 El ITM, las Tutorías y la Minería de Datos.	42
2.11 Herramienta de cómputo XLSTAT 2014.	44
III. METODOLOGÍA	46
3.1 El proceso de extracción del conocimiento.	47
3.2 Fase 1. Integración y recopilación.	49
3.3 Fase 2. Limpieza y transformación.	56
3.4 Fase 3. Minería de Datos.	69
3.5 Fase 4. Evaluación e Interpretación.	95
3.6 Fase 5. Difusión, uso y monitorización.	100
IV. RESULTADOS Y DISCUSIÓN	102
4.1 Fase 1.- De Aspirantes a Alumnos de Nuevo Ingreso.	104
4.2 Fase 2.- Alumnos en Formación Inicial con Tutor para Formación.	105
4.3 Fase 3.- Alumnos en Seguimiento con Tutor para Seguimiento.	107
4.4 Fase 4.- Alumnos en Cierre con Tutor para Cierre.	108
LITERATURA CITADA	
APENDICE	

## ÍNDICE DE TABLAS

<b>Figura</b>		<b>Página</b>
2.1	Diferentes áreas para la extracción del conocimiento.	39
3.1	Variables y su descripción de información del EXANI – II.	50
3.2	Secciones en las que se agrupan las variables de la base de datos del EXANI – II.	54
3.3	Clave y nombre de materias del primer semestre de ITIC.	55
3.4	Campos no significativos de primera instancia.	59
3.5	Campos no significativos en segunda instancia.	60
3.6	Atributos relevantes del primer componente principal	79
3.7	Atributos relevantes del segundo componente principal	79
3.8	Atributos relevantes del tercer componente principal	79
3.9	Atributos derivados de las funciones clasificadoras en general.	88
3.10	Atributos derivados de las funciones clasificadoras de la materia de Cálculo Diferencial e Integral.	90
3.11	Atributos derivados de las funciones clasificadoras de la materia de Fundamentos de Programación.	91
3.12	Atributos derivados de las funciones clasificadoras de la materia de Matemáticas Discretas.	92
3.13	Atributos derivados de las funciones clasificadoras de la materia de Ética.	93
3.14	Atributos derivados de las funciones clasificadoras de la materia de Fundamentos de la Investigación.	94
3.15	Tabla comparativa de atributos del Análisis de Componentes Principales vs Análisis Discriminante Lineal.	96



## ÍNDICE DE FIGURAS

Figura		Página
2.1	Esquema de Acreditación de COPAES.	10
2.2	Esquema de Técnicas de Minería de Datos.	35
2.3	Proceso de Knowledge Discovery Databases (KDD).	37
2.4	Disciplinas que contribuyen a la minería de datos.	37
3.1	Fases del proceso KDD.	48
3.2	Registros generados en EXANI – II Junio 2011.	57
3.3	Filtrado de Registros de la carrera de ITIC.	57
3.4	Acta de Calificaciones de Cálculo Diferencial.	58
3.5	Registros de alumnos activos que presentaron EXANI – II en ITIC.	59
3.6 (a)	Campos depurados para análisis.	63
3.6 (b)	Inserción de nuevos campos.	63
3.7	Opción: Descripción de datos – Estadísticos descriptivos.	64
3.8	Cuadro de Diálogo de Estadísticos descriptivos.	65
3.9	Cuadro de Diálogo de Confirmación de Análisis de Datos Cualitativos	65
3.10	Clasificación de los datos por estadística descriptiva.	66
3.11	Gráficas de frecuencia de cada dato analizado.	67
3.12	Análisis y gráfica de la variable MATRICULA.	67
3.13	Análisis y gráfica de la variable VIVE_MAD.	68
3.14	Análisis y gráfica de la variable 1T4 (Introducción a las TIC's).	69
3.15	Opción: Análisis de Datos – Análisis de Componentes Principales (ACP).	71
3.16	Cuadro de diálogo de Análisis de Componentes Principales.	72
3.17	Cuadro de diálogo de Confirmación de Tabla de Observaciones.	72
3.18	Elección de Ejes. (a) F1 vs F2, (b) F1 vs F3 y (c) F2 vs F3	73
3.19 (a)	Matriz de correlación primer rango.	74
3.19 (b)	Matriz de correlación segundo rango.	74
3.19 (c)	Matriz de correlación tercer rango.	75
3.19 (d)	Matriz de correlación cuarto rango.	75
3.20	Valores Propios de los Componentes Principales.	77
3.21 (a)	Cargas Factoriales primer rango.	78
3.21 (b)	Cargas Factoriales segundo rango.	78
3.22	Opción: Análisis de Datos – Clusters de k – means.	80
3.23	Cuadro de diálogo de Clusters de k – means.	81
3.24	Cuadro de diálogo de Confirmación de Selecciones.	82
3.25 (a)	Adaptación de la hoja para demostrar la relación.	83
3.25 (b)	Preparación de los datos para comparar con k – means.	83
3.26 (a)	Observaciones de las dos clases numéricas.	84
3.26 (b)	Gráfica de dispersión de promedios del primer semestre.	84
3.27	Opción: Análisis Discriminante Lineal (LAD).	85
3.28	Cuadro de diálogo de Análisis Discriminante Lineal.	86

<b>Figura</b>		<b>Página</b>
3.29	Cuadro de diálogo para la elección de gráficos.	86
3.30	Cuadro de diálogo de Confirmación de Selecciones.	87
3.31	Identificación de funciones clasificadoras general.	88
3.32	Identificación de funciones clasificadoras de la materia de Calculo Diferencial e Integral.	90
3.33	Identificación de funciones clasificadoras de la materia de Fundamentos de Programación.	91
3.34	Identificación de funciones clasificadoras de la materia de Matemáticas Discretas.	92
3.35	Identificación de funciones clasificadoras de la materia de Ética.	93
3.36	Identificación de funciones clasificadoras de la materia de Fundamentos de la Investigación.	94
4.1	Nomenclatura utilizada en el MIT.	103
4.2	Fase de Formación de Grupo de Alumnos en Formación Inicial del MIT.	104
4.3	Fase de Perfil del Tutor para Formación del MIT.	105
4.4	Procedimiento para las Tutorías Académicas (PTA).	105
4.5	Fase de Formación de Grupo de Alumnos en Formación (GALF) del MIT.	106
4.6	Procedimiento para las Tutorías Académicas para Grupo de Alumnos en Formación	106
4.7	Fase de Formación de Grupo de Alumnos en Seguimiento (GALS) del MIT.	107
4.8	Fase de Perfil del Tutor para Seguimiento del MIT.	108
4.9	Procedimiento para las Tutorías Académicas para Grupo de Alumnos en Seguimiento.	108
4.10	Fase de Formación de Grupo de Alumnos en Cierre (GALC) del MIT.	109
4.11	Fase de Perfil del Tutor para Cierre del MIT.	109
4.12	Propósito del MIT para que el estudiante obtenga el grado de Licenciatura.	103
4.13	Modelo Inteligente de Tutorías Académicas en Instituciones de Educación Superior.	103

---

## I. INTRODUCCIÓN

Las Instituciones de Educación Superior en México, tienen el compromiso de egresar individuos que respondan a las necesidades, no sólo del sector productivo, sino de la misma sociedad, la cual demanda profesionistas calificados.

El Gobierno en sus diferentes niveles, se preocupa por generar condiciones óptimas para la formación universitaria, estableciendo ejes estratégicos, que son ejecutados, supervisados y evaluados por diversos organismos que otorgan certificaciones y acreditaciones, garantizando la calidad de los Programas Educativos ofertados.

Las universidades y la educación en general, son los puntos más relevantes que se deben considerar para tener un país más productivo, que permita la formación de gente innovadora (alumnos), que cuente con el apoyo de investigadores (docentes) y que genere desarrollos tecnológicos (vinculación gobierno – empresa – escuela). El que una Institución de Educación Superior (IES) a nivel nacional, realice una evolución de este tipo tiene un gran significado para todos ya que se tiene que replantear lo ya constituido y trabajar para cumplir las nuevas expectativas, logrando sus objetivos.

La implementación del sistema tutorial en diversas IES mexicanas tiene variantes surgidas de la necesidad de adecuar la tutoría a las condiciones institucionales, como su historia, planta docente, modelo curricular, recursos humanos y tradiciones, ente otros.

El uso de las tecnologías de la información no podría ser ajeno a este sistema, desde la implementación de sitios Web, hasta el uso de aplicaciones móviles que puedan contribuir al quehacer tutorial.

---

El presente tema de tesis propone un Modelo Inteligente de Tutoría (MIT), que aplicando minería de datos, proporcione perfiles de aspirantes, alumnos y tutores idóneos para crear grupos con características particulares de acuerdo a su contexto, disminuyendo la diferencia entre el ingreso y la eficiencia terminal, aplicado a cualquier IES que implemente el MIT y que su método de selección para el nuevo ingreso de alumnos tome como base el análisis del Examen de Ingreso, que proporciona el Centro Nacional de Evaluación (CENEVAL).

El caso práctico que demuestra un primer perfil considerado en el MIT, es sobre la base de datos proporcionadas por el Instituto Tecnológico de Morelia, generada en Junio del 2011, tomando como referencia los datos obtenidos de los aspirantes a la Ingeniería en Tecnologías de la Información y que haciendo un comparativo con las materias cursadas para el primer semestre de la carrera, se pudo determinar cuáles son los atributos que se deben tomar en consideración, para saber cuáles son los aspirantes idóneos que serán aceptados e inscritos en la carrera, siendo los discriminados los que podrán ser atendidos a través de la aplicación de un instrumento de evaluación tutorial o de asesoría, con el objeto de que se preparen y puedan ser candidatos para el siguiente ciclo.

Las técnicas utilizadas para el análisis de los datos fueron Estadística Descriptiva, Análisis de Componentes Principales (ACP) complementado con k – means y por último Análisis Discriminante Lineal (ADL). Las observaciones realizadas a partir de estos análisis son sólo respecto al primer perfil del MIT, identificado como Perfil del Aspirante (PAS) y sólo teniendo como insumo el EXANI – II de CENEVAL.

Las pretensiones de este modelo son ambiciosas, ya que su demostración dependerá no sólo de la aplicación de estadística y algoritmos de minería de datos, sino también del contexto educativo y del tiempo, para generar cada uno de los perfiles, para que estos interactúen de manera armoniosa y logren con éxito su objetivo en común.

---

## II. REVISIÓN DE LITERATURA

*“Educar a un joven no es hacerle aprender algo que no sabía, sino hacer de él alguien que no existía”.*

*Ruskin John*

### **2.1 La Educación Superior y el Gobierno Federal.**

El Plan Nacional de Desarrollo 2013 – 2018, contienen cinco metas principales; la tercera hace referencia a tener un *“México con educación de calidad”* que permita a los niños y jóvenes la superación y el éxito en sus vidas, a partir de la premisa *“para mover a México hay que fomentar los valores cívicos, elevar la calidad de la enseñanza y promover la ciencia, la tecnología y la innovación.”* (Poder Ejecutivo, 2013).

Esta meta plantea como *Objetivo 3.1 “Desarrollar el potencial humano de los mexicanos con educación”*, considerando seis estrategias para lograrlo, de las cuales es importante resaltar la que se describe en la *Estrategia 3.1.1 “Establecer un sistema de profesionalización docente que promueva la formación, selección, actualización y evaluación del personal docente y de apoyo técnico – pedagógico”*.

Sobre esta estrategia se puede resaltar algunas líneas de acción que impactan en las Instituciones de Nivel Superior (IES) como las referentes a la atención del personal docente para *“Estimular el desarrollo profesional de los maestros, centrado en la escuela y en el aprendizaje de los alumnos, en el marco del Servicio Profesional Docente”*; *“Impulsar la capacitación permanente de los docentes para mejorar la comprensión del modelo educativo, las prácticas pedagógicas y el manejo de las tecnologías de la información con fines educativos”*.

---

Además de *“Estimular los programas institucionales de mejoramiento del profesorado, del desempeño docente y de investigación, incluyendo una perspectiva de las implicaciones del cambio demográfico”*; *“Mejorar la supervisión escolar, reforzando su capacidad para apoyar, retroalimentar y evaluar el trabajo pedagógico de los docentes”*.

Respecto a la *Estrategia 3.1.5 “Disminuir el abandono escolar, mejorar la eficiencia terminal en cada nivel educativo y aumentar las tasas de transición entre un nivel y otro”*, se tienen como líneas de acción *“Ampliar la operación de los sistemas de apoyo tutorial, con el fin de reducir los niveles de deserción de los estudiantes y favorecer la conclusión oportuna de sus estudios”*; *“Implementar un programa de alerta temprana para identificar a los niños y jóvenes en riesgo de desertar”*; *“Establecer programas remediales de apoyo a estudiantes de nuevo ingreso que presenten carencias académicas y que fortalezcan el desarrollo de hábitos de estudio entre los estudiantes”*; *“Definir mecanismos que faciliten a los estudiantes transitar entre opciones, modalidades y servicios educativos”*.

La *Estrategia 3.1.6 “Impulsar un Sistema Nacional de Evaluación que ordene, articule y racionalice los elementos y ejercicios de medición y evaluación de la educación”*, permite tener como única línea de acción *“Garantizar el establecimiento de vínculos formales de interacción entre las instancias que generan las evaluaciones y las áreas responsables del diseño e implementación de la política educativa”*.

Con estas tres estrategias, se refuerza la opción que abre la certificación de competencias profesionales y laborales; además de operar los sistemas de apoyo tutoriales que den seguimiento desde que una persona aspira a ingresar a una IES, hasta que se titula en una carrera universitaria, con el fin de reducir los niveles de deserción de los estudiantes y favorecer la eficiencia terminal.

---

Lo anterior se logra a través de la implementación de un Sistema Nacional de Evaluación que no sólo determine si un profesor es competente o no en su labor, sino que también mida el desempeño de sus estudiantes en función de sus aspiraciones profesionales.

Las evaluaciones a los programas de estudio se perfeccionan, propiciando el apoyo de aquellos que obtuvieron los mejores resultados, reduciendo la brecha entre éstos y los que necesitan ser fortalecidos.

En paralelo se tiene, el Programa Sectorial de Educación 2013 – 2018, que establece seis objetivos, de los cuales sólo dos están relacionados con la Educación Superior, y son el Objetivo 2.- *“Fortalecer la calidad y pertinencia de la educación media superior, superior y formación para el trabajo, a fin de que contribuyan al desarrollo de México”* y el Objetivo 6.- *“Impulsar la educación científica y tecnológica como elemento indispensable para la transformación de México en una sociedad del conocimiento”*. (PSE, 2013)

Desde hace ya varios años, el Gobierno también ha trabajado para articular un Sistema Nacional de Evaluación, Acreditación y Certificación, con el fin de fortalecer las prácticas de autoevaluación, evaluación externa de pares, acreditación formal y exámenes nacionales estandarizados a los egresados, las cuales son efectuadas por instancias como los Comités Interinstitucionales para la Evaluación de la Educación Superior (CIEES), el Consejo para la Acreditación de la Educación Superior (COPAES) y el Centro Nacional de Evaluación para la Educación Superior (CENEVAL), entre otros.

La acreditación de la calidad en la oferta educativa es esencial para mantener el valor social y económico de la educación superior.

---

Con el fortalecimiento de los mecanismos normativos, regulatorios y de aseguramiento de la calidad, se pretende frenar la proliferación de programas que no siguen los estándares, limitando la formación de los ciudadanos productivos y competitivos que requiere el país.

Adicionalmente, el fortalecimiento de los sistemas de evaluación constituye una base fundamental para asegurar y generalizar la rendición de cuentas a los estudiantes, padres de familia, empleadores y sociedad en general.

## **2.2 La Educación Superior y los CIEES.**

La Comisión Nacional de Evaluación de la Educación Superior puso en marcha, en 1990, un sistema de evaluación que asumió la tarea de coordinar la participación colegiada de las autoridades gubernamentales y de las propias IES. (CIEES, 2013).

Este sistema incluye tres actividades fundamentales: la primera es la evaluación global del sistema y los subsistemas de este nivel, encomendada a la entonces Subsecretaría de Educación Superior e Investigación Científica, actualmente Subsecretaría de Educación Superior; a la Subsecretaría de Educación e Investigación Tecnológicas; a la Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES), y al Consejo del Sistema Nacional de Educación Tecnológica. La segunda es la autoevaluación, encomendada a las propias IES. Y la tercera es la evaluación interinstitucional, que se puso en manos de los Comités Interinstitucionales para la Evaluación de la Educación Superior (CIEES), creados en 1991.

Los CIEES realizan sus tareas a través de grupos interinstitucionales colegiados, cuyos integrantes son pares representativos de la comunidad académica nacional.



---

Se cuenta con un comité de evaluación por cada una de las áreas del conocimiento en las que está organizada la educación superior; a saber: Arquitectura, Diseño y Urbanismo, Artes, Educación y Humanidades, Ciencias Agropecuarias, Ciencias de la Salud, Ciencias Naturales y Exactas, Ciencias Sociales y Administrativas, así como Ingeniería y Tecnología. Adicionalmente, se formaron otros dos comités, uno encargado de evaluar la Difusión y Extensión de la Cultura, y otro con la encomienda de valorar la Administración y la Gestión de las IES.

El principal objetivo de la evaluación que realizan los comités consiste en contribuir a mejorar la calidad y eficiencia de la educación superior en el país. Es importante precisar que no se trata de una evaluación fiscalizadora ni tiene propósitos de administración externa. Su tarea se inscribe en la búsqueda de modelos de educación superior que respondan no solo a la evaluación del conocimiento y la cultura, sino a las exigencias y necesidades sociales del país.

Dadas las características complejas de las instituciones y de los procesos educativos, los CIEES han optado por un enfoque particular que combina el análisis histórico y las aportaciones del modelo sistémico. Además, parte de la premisa de que cada programa educativo representa una unidad académico-administrativa donde la instancia académica es la que hace posible la actividad docente y se rige por un plan de estudios formalmente establecido, el cual aspira a satisfacer un perfil del estudiante y del egresado, que responda a las necesidades académicas, laborales, sociales y culturales de una profesión específica en nuestra sociedad actual y en el contexto nacional.

Por otro lado, se asume que la estructura administrativa, es la que permite resolver y organizar los diversos recursos del programa, facilitando con ello que se alcancen los objetivos académicos determinados por el propio plan de estudios.

---

Con base en las aportaciones de dicho modelo sistémico, en su perspectiva de sistema abierto, los CIEES definen las características de los profesores, alumnos e instalaciones, así como los productos del proceso educativo; o sea, las proporciones de egreso, titulación y las labores de investigación, considerándolos parte de sistemas más amplios.

El análisis de los insumos y productos se supedita al de los procesos históricos, que subyacen en la construcción institucional del programa. En otras palabras, se identifican las principales decisiones e inercias de los actores fundamentales del programa educativo desde su inició, todo lo cual puede explicar su situación actual.

La interacción de los diferentes elementos de los programas debe corresponder a las decisiones académicas y administrativas plasmadas en los reglamentos, y también a los usos establecidos tanto por los sujetos directamente relacionados con la enseñanza, como por las autoridades centrales.

Los CIEES retoman el concepto de estructura curricular, como referente teórico y metodológico, permite advertir la relación que se da entre los objetivos del programa, la participación de los profesores, los tiempos que lo rigen, las instalaciones que tienen, los recursos didácticos y las formas de evaluación. Igualmente, hace posible delimitar los espacios curriculares reales que se establecen para el aprendizaje.

---

### **2.3 La Educación Superior y COPAES.**

El Consejo para la Acreditación de la Educación Superior A.C. (COPAES) es la única instancia autorizada por el Gobierno Federal a través de la Secretaría de Educación Pública (SEP), para conferir reconocimiento formal y supervisar a organizaciones cuyo fin sea acreditar programas educativos del tipo superior que se impartan en México, en cualquiera de sus modalidades (escolarizada, no escolarizada y mixta). (COPAES, 2014).

La autorización le fue conferida el 8 de noviembre del año 2000 mediante el convenio de concertación firmado con la Secretaría de Educación Pública (SEP).

Durante la primera década, el COPAES operó al amparo de la estructura de los CIEES, sin embargo, atendiendo a las acciones prioritarias del Programa Sectorial de Educación 2007–2012.

Sin embargo, el 26 de febrero de 2010, la Asamblea General del COPAES de la cual forma parte la SEP, por conducto de la Subsecretaría de Educación Superior, tomó la decisión de separar orgánica y estructuralmente a los dos organismos, a fin de articular el quehacer de las diferentes instancias de evaluación y acreditación existentes, y concretar la creación del Sistema Nacional de Evaluación, Acreditación y Certificación de la Educación Superior. Sus principales funciones son:

- a)** Garantizar que las organizaciones que acreditan programas de educación superior, cumplan con mecanismos que aseguren rigor académico e imparcialidad, el profesionalismo de los evaluadores, la pertinencia de los procedimientos, la imparcialidad de la evaluación y, en general, todo lo que propicie su actuación de acuerdo con los principios éticos correspondientes.

- b) Fungir como órgano de consulta de la SEP, en materia de acreditación de la educación superior.
- c) Realizar investigación relacionada con su objeto: La calidad de la educación superior.
- d) Informar a la sociedad acerca de los organismos reconocidos formalmente por el COPAES y sus actividades, así como de los programas de educación superior acreditados en el país.
- e) Interactuar con organismos análogos internacionales.

Su proceso de acreditación se plasma en la Figura 2.1:



Figura 2.1 Esquema de Acreditación de COPAES.

---

## **2.4 La Educación Superior y CENEVAL.**

El Centro Nacional de Evaluación para la Educación Superior (CENEVAL) es una asociación civil sin fines de lucro cuya actividad principal es el diseño y la aplicación de instrumentos de evaluación de conocimientos, habilidades y competencias, así como el análisis y la difusión de los resultados que arrojan las pruebas (CENEVAL, 2014).

Desde el 28 de abril de 1994, a partir de una recomendación de la ANUIES, proporciona información confiable y válida sobre los conocimientos y habilidades que adquieren las personas como beneficiarios de los programas educativos de diferentes niveles de educación formal e informal.

Su máxima autoridad es la Asamblea General, constituida por instituciones educativas, asociaciones y colegios de profesionales, organizaciones sociales y productivas y autoridades educativas gubernamentales. Cuenta con un Consejo Directivo que garantiza la adecuada marcha cotidiana del Centro. Su director general es la autoridad ejecutiva del mandato emanado de la Asamblea General.

Sus instalaciones centrales se encuentran en la Ciudad de México y sus actividades se sustentan en los últimos avances e investigaciones de la psicometría y otras disciplinas, así como en la experiencia y compromiso de su equipo, integrado por casi 600 personas.

Los instrumentos de medición que elabora el Centro proceden de procesos estandarizados de diseño y construcción y se apegan a las normas internacionales; en su elaboración participan numerosos cuerpos colegiados integrados por especialistas provenientes de las instituciones educativas más representativas del país y organizaciones de profesionales con reconocimiento nacional.

---

Actualmente cuenta con 300 instrumentos de evaluación, uno de ellos es el Examen Nacional de Ingreso a la Educación Superior (EXANI – II), instrumento que se utiliza en procesos de selección de aspirantes a cursar estudios superiores para más de 150 instituciones en la República Mexicana.

## **2.5 La Educación Superior y las Tutorías.**

El sistema tutorial en el nivel de licenciatura se inició en la Universidad Nacional Autónoma de México (UNAM) dentro del sistema de Universidad Abierta (SUA), en dos modalidades distintas: individual y grupal. En la primera se atienden las dudas surgidas en el proceso de estudio particular del alumno y en la grupal se favorece la interacción de los estudiantes con el tutor para la solución de problemas de aprendizaje o para la construcción de conocimientos. (UADY, 2006)

El SUA de la UNAM, tiene sus antecedentes en la Open University de Inglaterra y la UNED de España.; fue creada en 1972 y funciona como una modalidad alternativa al sistema escolarizado en 17 licenciaturas.

Este sistema tiene como objetivo fundamental apoyar al estudiante en el desarrollo de metodologías de estudio; crear un clima de confianza que permita conocer otros aspectos de la vida personal del alumno, así como señalar y sugerir actividades extracurriculares para un desarrollo profesional integral y brindar información académico – administrativas según las necesidades del estudiante.

En agosto de 2000 se inició en la Universidad Autónoma de Nuevo León (UANL) la implantación de un Programa Tutorial en la Facultad de Ingeniería Mecánica (FIME), con la finalidad de apoyar la formación integral de los estudiantes y orientarlos en la solución de sus problemas académicos o personales.

---

Además este programa, pretende como consecuencia de un mejor desempeño de los estudiantes, que se disminuya la deserción y se incremente la eficiencia académica.

En esta Universidad las tutorías tienen como meta desarrollar un programa que permita la orientación de los estudiantes que ingresan a la institución, mediante la conformación de un grupo institucionalizado de profesores – tutores, que en conjunto con asesores apoyen el desarrollo personal y académico de los estudiantes.

El programa tutorial reconoce la existencia de cuatro grandes clases de actividades relacionadas con la tutoría (ANUIES, 1999), que demandan características y habilidades distintas en cada profesor – tutor:

- ▶ **Tutor:** Recibe a los estudiantes de primer ingreso y les da apoyo y seguimiento hasta que son capaces de resolver sus problemas por sí mismos o con una ayuda mínima.
- ▶ **Tutor asignado (tutor de carrera):** Encargado de orientar, de auxiliar a los alumnos en la elección de sus materias, además de ofrecerles orientación curricular.
- ▶ **Tutor principal (tutor de graduación):** Orienta en tópicos específicos a estudiantes que estén en la posibilidad de hacer prácticas profesionales o de titulación.
- ▶ **Asesor (profesor – asesor):** Es un profesor especialista en una disciplina y apoya al alumno en un campo del conocimiento específico.

---

Dada la trascendencia de la función de la tutoría, así como de su impacto con los estudiantes, la administración de la FIME considera que los tutores deben ser personas propositivas, sensibles, honestas, responsables, con un espíritu de servicio, comprometida consigo mismo, con la institución y con los estudiantes.

Lo anterior les exige poseer conocimientos y habilidades, y desarrollarse como especialistas para este trabajo académico fundamental en la formación del estudiante, así como ejercer el respeto de la confidencialidad de la información personal.

Desde la implementación del Programa Tutorial FIME se reconocen los siguientes resultados:

- ▶ Conformación de un cuerpo de tutores, con la formación y perfil deseables para la FIME.
- ▶ Establecimiento de un programa de inducción, paralelo al Curso Propedéutico ofrecido a los estudiantes de primer ingreso a la FIME.
- ▶ Asignación y adecuación de un área física que permita la atención a los estudiantes, en las mejores condiciones posibles.
- ▶ Establecimiento de un programa de entrevistas calendarizado para cada estudiante.
- ▶ Identificación de problemas académicos o personales de los estudiantes.

Por su parte, la Universidad Autónoma del Estado de Hidalgo ofrece un modelo de tutoría integral donde se le proporciona al estudiante asistencia a lo largo de su trayectoria escolar, con apoyo de un asesor psicológico, un trabajador social y un maestro orientador. La tutoría opera como sustento en aspectos académicos, económicos, sociales y personales (UAEH, 2013).



---

Para su operación, este modelo definió la figura del tutor como un académico poseedor de un alto valor moral entre los alumnos y estableció una batería de pruebas para detectar problemas que pudiesen afectar el desempeño académico de los alumnos. De dicha prueba se establece una escala que abarca de alumnos con alta probabilidad de concluir con sus estudios hasta los que probablemente no concluirán.

El tutor comienza a trabajar con los alumnos clasificados en el último grupo. Este tutor no es un experto, pero canaliza los casos con especialistas. Uno de los resultados obtenidos desde la implantación del programa se observa en la disminución del índice de reprobación de 65% al 16% y de deserción del 23% al 9.3%.

En la Universidad de Guadalajara la función de tutoría se inició en 1992, estableciéndose estatualmente como una obligación de todo miembro del personal académico el desempeñarse como tutor académico de los alumnos para procurar su formación integral (UDG, 2004).

Asimismo, se acordó que las funciones tutoriales comprendan las siguientes modalidades:

- ▶ Tutorías curriculares ligadas a cursos regulares.
- ▶ Curso o talleres de nivelación.
- ▶ Asesoría o consultoría académica con el aval del departamento.
- ▶ Orientación para estudios de licenciatura y posgrado con el aval del departamento.
- ▶ Apoyo a estudiantes en el trabajo de laboratorio.
- ▶ Responsabilidad sobre alumnos que cumplen servicio social.

---

A partir de la preocupación que ha existido a nivel internacional, por mejorar la calidad y eficiencia de las instituciones de educación superior, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, en su Declaración Mundial sobre la Educación Superior en el Siglo XXI: Visión y Acción, señala que es necesario modificar el proceso de aprendizaje, ya que considera que debido a la velocidad de los cambios sociales es urgente desarrollar una nueva visión y un nuevo modelo de enseñanza superior que se centre en las necesidades y capacidades del estudiante (UNESCO, 1998).

Además sugiere, que los establecimientos de educación superior deberán ser capaces de proporcionar a sus alumnos, asesoría, orientación y consejo, cursos de recuperación, formación para el estudio y otras formas de apoyo a los estudiantes, que haga posible el mejoramiento de sus condiciones de vida.

Por su parte, la Asociación Nacional de Universidades e Instituciones a Nivel Superior (ANUIES), publicó un libro denominado Programas Institucionales de Tutorías. En esta propuesta se describe la conformación, organización y funcionamiento de dicho Programa, en las instituciones de educación superior, en donde hace referencia, respecto a que la educación superior mexicana requiere transformarse teniendo como eje una nueva visión y un nuevo paradigma para la formación de los estudiantes, donde la atención personalizada de éstos puede ayudar a abatir los índices de reprobación y rezago escolar, a disminuir las tasas de abandono de los estudios y a mejorar la eficiencia terminal (ANUIES, 2000).

En este documento se define que la *“Tutoría es un acompañamiento personal y académico a lo largo del proceso formativo para mejorar el rendimiento académico, facilitar que el estudiante solucione sus problemas escolares, desarrolle hábitos de estudio, trabajo, reflexión y convivencia social.”*

---

Con base a estos antecedentes, en Agosto del 2006, un grupo de docentes del Sistema Nacional de Educación Superior Tecnológica (SNEST) y personal académico de la Dirección General de Educación Superior Tecnológica (DGEST), pertenecientes a la Secretaría de Educación Pública (SEP), se dieron a la tarea de formular el Programa Nacional de Tutorías (PNT), cuyo propósito era *“contribuir, a través de la acción tutorial, al mejoramiento del rendimiento académico de los estudiantes, coadyuvar en el logro de su formación integral, e incidir en las metas institucionales relacionadas con la calidad educativa, como son, la disminución de los índices de reprobación, deserción, y rezago, además de favorecer con ello la eficiencia terminal”*; dicho Programa debía ser implementado por todos los Institutos Tecnológicos del país (DGEST, Programa Nacional de Tutoría, 2006).

Posteriormente en la Reunión Nacional de Jefes de Departamento de Desarrollo Académico realizada en Tehuacán, Puebla, en 2007, el Instituto Tecnológico de Hermosillo aplicó una encuesta, en una muestra de 122 participantes, para obtener un Diagnóstico Nacional de Tutoría, el cual proporcionó información sobre la situación del PNT y de las siguientes necesidades (SNEST, 2012):

1. Un Manual para la práctica de la tutoría.
2. Cursos de formación y actualización para tutores.
3. Necesidad de psicólogos para apoyo del programa.
4. Apoyos institucionales que favorezcan la tutoría.
5. Un reglamento de tutoría.
6. Investigación sobre los procesos, alcances e impacto de la tutoría.
7. Talleres y/o cursos continuos dirigidos a los estudiantes sobre temas de la situación actual del estudiantado.

En diciembre de 2009, el mismo Instituto, realizó el “Taller para la revisión del PNT y el Manual del Tutor para su implementación bajo el enfoque por competencias”.

---

El 21 de noviembre del 2012, la Subsecretaría de Educación Superior, a través de la todavía Dirección General de Educación Superior Tecnológica (DGEST), ahora denominado Sistema Nacional de Institutos Tecnológicos (SNIT), emitió una circular a todos los Institutos Tecnológicos del país, en donde les hacía de su conocimiento que estaba disponible el “Manual del Tutor del SNIT”, cuya última actualización es del mes de febrero del 2013, teniendo como propósito *“Orientar a los docentes – tutores, en la planeación, organización, seguimiento y evaluación de la actividad tutorial, facilitando el desarrollo académico, personal y profesional de los estudiantes, con la finalidad de fortalecer su formación integral”* (SNIT, 2013).

Dicho manual, considera que la tutoría *“es un proceso de acompañamiento grupal o individual que un tutor le brinda al estudiante durante su estancia en el Instituto Tecnológico con el propósito de contribuir a su formación integral e incidir en las metas institucionales relacionadas con la calidad educativa; elevar los índices de eficiencia terminal, bajar los índices de reprobación y deserción.”*

La tutoría contempla tres ejes fundamentales: desarrollo académico, desarrollo personal y desarrollo profesional que se ofrece en cada Instituto Tecnológico.

A través de la acción tutorial, los estudiantes aprenden a: identificar retos en el campo profesional; localizar el conocimiento disponible; recuperar información relevante; desarrollar modelos conceptuales para visualizar posibles abordajes a los problemas; buscar conexiones con otros campos disciplinarios y tener un enfoque interdisciplinario; verificar con una metodología sistemática los resultados de las intervenciones profesionales; reflexionar en y sobre la acción.

---

La tutoría, dentro de sus objetivos, busca reducir los índices de reprobación y deserción, potenciar las competencias del estudiante mediante acciones preventivas y correctivas, así como apoyar al estudiante en el proceso de toma de decisiones relativas a la construcción de su trayectoria formativa, de acuerdo con su vocación, intereses y competencias, mediante la atención personalizada y/o grupal en donde se apoye la formación del tutorado centrada en prácticas metacognitivas, es decir, orientada a que los estudiantes mejoren en forma continua a partir de la propia reflexión sobre su desempeño de tal suerte que no sólo es importante poseer saber, sino comprender cuál es la naturaleza de éste.

De acuerdo a la atención que un tutor puede brindar a un estudiante, se consideran las siguientes modalidades de tutoría:

**Tutoría individual:** La tutoría individual es una situación en la que se establece una relación directa cara a cara entre tutor y estudiante sobre cuestiones académicas individuales, de su situación personal, social o profesional. Se da una interacción, verbal y no verbal.

**Tutoría grupal:** La tutoría grupal es el proceso de acompañamiento de un grupo de estudiantes con la finalidad de abrir un espacio de comunicación, conversación y orientación grupal, donde ellos tengan la posibilidad de revisar y discutir, junto con su tutor, temas que sean de su interés, inquietud, preocupación, así como también para mejorar el rendimiento académico, solucionar problemas escolares, desarrollar hábitos de estudio, reflexión y convivencia social.

Por otro lado, se busca que en la tutoría grupal se oriente a los tutorados hacia lo que el estudiante va a ser capaz de conocer (área cognitiva), lo que el estudiante va a ser capaz de hacer (habilidades y destrezas) y lo que el estudiante va a ser capaz de ser (actitudes y responsabilidades).

---

**Tutoría con padres de familia:** Se busca que los padres identifiquen al tutor del grupo al cual pertenece su hijo, con la finalidad de fortalecer y vincular los esfuerzos en la formación integral del estudiante durante su estancia en el instituto, así mismo se busca mantenerle informado del conjunto de acciones coordinadas que se ofrecen en el tecnológico entre las autoridades y los docentes, así como los servicios de apoyo que inciden en la formación de los estudiantes.

Además de abordar un proceso de capacitación del personal académico, perfil y funciones del tutor, las IES deberán instrumentar acciones complementarias sin las cuales, el PIT no lograría el efecto deseado, tales son:

1. Fortalecer y orientar los servicios institucionales a estudiantes a fin de que respondan eficientemente a las necesidades del programa tutorial.
  
2. De gran relevancia será lograr que cada institución cuente con un conjunto de apoyos a los estudiantes, a fin de que las actividades tutoriales se vean fortalecidas. Se considera que es el tutor quien estará en contacto con las necesidades más sentidas y urgentes del estudiantado, por lo que dichos servicios requerirán del establecimiento de una coordinación efectiva con los grupos de tutores, y éstos deberán conocer las maneras de hacer uso de ellos, de forma que sean aprovechados al máximo, beneficiando la formación integral del estudiantado. Entre otros servicios se encuentran los siguientes:
  - ▶ Orientación educativa.
  - ▶ Servicios médicos.
  - ▶ Trabajo social.
  - ▶ Asistencia psicológica.
  - ▶ Servicio social y residencia profesional.
  - ▶ Becas.
  - ▶ Bolsa de trabajo.

- 
- ▶ Educación continua (cursos y talleres de apoyo al programa tutorial).

Es necesario que estos servicios sean proporcionados por personal profesionalizado en las áreas que atiende, que cuenten con procedimientos sencillos, para su uso y que sean accesibles, a fin de resultar un apoyo real a las necesidades del estudiante.

En síntesis, sería recomendable, al respecto, promover la integración entre el PIT y los distintos servicios estudiantiles, que permita la accesibilidad a estos y su uso oportuno y adecuado al logro de los objetivos educativos del estudiante.

## **2.6 El Instituto Tecnológico de Morelia (ITM).**

El ITM es una institución educativa pública de educación superior que forma parte del Sistema Nacional de Institutos Tecnológicos de México (SNIT) y que se regula a través del Tecnológico Nacional de México (TNM), perteneciente a la Subsecretaría de Educación Superior (TNM, 2014).

Inició sus actividades el 4 de abril de 1965, con carreras de educación en el trabajo, con una captación de 220 alumnos (ITM, 2009).

Actualmente cuenta con una matrícula de 5,016 alumnos y tiene una oferta educativa de quince carreras a nivel licenciatura, trece con sistema escolarizado y dos en sistema abierto, además de seis maestrías y dos doctorados.

---

Su misión es “Formar personas competentes y competitivas a nivel internacional en el campo científico – tecnológico y humano, que contribuyan al desarrollo sustentable de México y al conocimiento global, fomentando valores humanos y culturales.” (ITM, 2014).

Su visión es “Consolidarnos como una institución pública de educación superior de alto desempeño, líder en el campo científico – tecnológico, y referente del sistema educativo nacional por la calidad y pertinencia de los servicios que ofrecemos a la sociedad”.

A fin de guiar y orientar las acciones cotidianas de todo su personal, el Instituto Tecnológico de Morelia define los siguientes valores institucionales:

- ▶ **Honestidad:** La actuación auténtica y justa con apego a la verdad.
- ▶ **Lealtad:** Ser fieles y nobles, con nosotros mismos y con nuestro trabajo.
- ▶ **Respeto:** La tolerancia a las ideas, creencias, salud y dignidad, reflejados en la libertad, la armonía y la paz para asegurar la convivencia entre los individuos.
- ▶ **Responsabilidad:** El cumplimiento efectivo y eficiente de los compromisos asumidos, que implican el estar conscientes de nuestras obligaciones y del efecto de nuestros actos.
- ▶ **Solidaridad:** La integración de los miembros de la comunidad tecnológica mediante necesidades y expectativas de nuestros clientes internos y externos, proporcionándoles servicios de excelencia con base en nuestra filosofía y estrategia de mejora continua.
- ▶ **Libertad:** La facultad de expresión y actuación de individuo dentro de un marco de responsabilidad social.
- ▶ **Cultura Ecológica:** Respeto, conservación y creación de medios ambientes favorables.
- ▶ **Justicia:** Igualdad para todos.



---

Su compromiso es lograr la formación integral de sus estudiantes, entendiendo ésta como un proceso continuo de desarrollo de todas las potencialidades del ser humano, buscando la plenitud en el aprender a ser, aprender a hacer, aprender a aprender, aprender a emprender y aprender a convivir, convirtiéndolo en un ciudadano consiente, responsable y solidario. (DGEST, 2011).

En el año 2010, en el Departamento de Sistemas y Computación, se reacreditó el plan de Estudios de Ingeniería en Sistemas Computacionales por el Consejo de Acreditación de la Enseñanza de la Ingeniería (CACEI).

En Noviembre del 2012, el ITM recibió por parte de la Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES), el premio SEP-ANUIES 2012, por ser un plantel que tiene el 100 por ciento de sus programas reconocidos por su calidad (García, 2012).

Mantener al 100% los programas acreditados y reconocidos por su buena calidad, sólo ha sido posible gracias a que se atienden las observaciones de los organismos acreditadores y se mantienen los procesos de reacreditación de los programas que podrían vencer año con año.

## **2.7 El Instituto Tecnológico de Morelia y las Tutorías.**

Una de las estrategias que se han adoptado para abatir los índices de deserción en las carreras que oferta el Departamento de Sistemas y Computación, del ITM (Ingeniería en Sistemas Computacionales, Licenciatura en Informática, Ingeniería en Informática e Ingeniería en Tecnologías de la Información y Comunicaciones), es la redefinición total en torno al Programa Institucional de Tutorías Académicas.

---

En el PNT incorpora desde 1997 de manera oficial, el concepto de Tutorías Académicas que tenía como objetivo “Establecer las normas y lineamientos para proporcionar tutorías académicas, con la finalidad de orientar al alumno de licenciatura técnica y licenciatura cuando así lo requiera y, de ésta manera crear las condiciones para elevar el nivel académico del alumno” (DGEST, 2006).

En este procedimiento se define la tutoría como las “actividades de asesoría que realizan los profesores para apoyar a todos los alumnos en su desarrollo académico, profesional y personal durante su permanencia en la institución”.

La tutoría como una estrategia de acompañamiento podrá incidir en el incremento de la calidad del proceso formativo, y aumentar el rendimiento de los estudiantes, haciendo más viable alcanzar niveles de eficiencia terminal más satisfactorios.

La aplicación responsable de la tutoría permitirá conocer las características individuales, los intereses personales, las expectativas profesionales y laborales y los proyectos de vida de los jóvenes, y con ello establecer relaciones interpersonales entre el profesor – tutor y el estudiante, que redunde no solo en el beneficio de ambos sino también del ambiente en el que se desarrollan.

La tutoría en los términos, estructura y operación que se plantean en el PNT, va más allá de solamente proporcionar información a los estudiantes; promueve alternativas de atención de carácter preventivo no sólo correctivo, que fortalecen la formación integral, cuenta con un sistema de evaluación, que permite obtener a nivel nacional un diagnóstico de la efectividad del Programa.

---

Los tutores deben participar en este Programa, para tener la posibilidad de reconocer su rol docente, con un sentido mucho más amplio, en un aula y en un horario determinado en un período de tiempo, en respuesta a las actividades administrativas de la acción docente, para ser un facilitador del aprendizaje, cuya acción se expande a diversos escenarios educativos, con la finalidad de atender las diferentes necesidades de formación de sus estudiantes y de sus tutorados.

Algunos tecnológicos han desarrollado propuestas que resultaron diversas en su extensión, contenido, cobertura y pertinencia, dificultando con ello una evaluación de las tutorías académicas a nivel nacional.

Sin embargo en el ámbito local, algunos tecnológicos han alcanzado experiencias significativas, que permiten reconsiderar la implementación de la tutoría bajo un esquema más flexible y humanista, que propicie el trabajo colaborativo y la corresponsabilidad entre los actores educativos.

En el caso en particular del ITM, se han hecho esfuerzos aislados en diversos Departamentos, para consolidar el PNT en su contexto. Uno de ellos fue el que elaboró el Departamento de Sistemas y Computación, el cual puso en marcha a partir de junio del 2005 y lo denominó Proyecto Departamental de Tutorías Académicas, atendiendo al *Objetivo 2.- Fortalecimiento del proceso enseñanza – aprendizaje* del Plan Estratégico de Desarrollo 2005 – 2010, abarcando los siguientes puntos (Computación, 2005):

- ▶ Cubrir en su totalidad planes y programas de estudio.
- ▶ Elevar los índices de aprobación así como la eficiencia terminal.

Sin embargo, a nueve años de operación, se detectaron las siguientes deficiencias:

- 
- ▶ El alumno después del primer semestre de su carrera, pierde comunicación con el tutor, a partir de que elige las materias que más le convenga por semestre, por lo que se forman grupos de alumnos por materias, más no por semestre de acuerdo a la retícula. A través del tiempo el alumno no sabe quién es su tutor.
  - ▶ No existe compromiso de los alumnos para llevar a cabo las actividades originalmente establecidas para los Programas de Tutorías.
  - ▶ No existen las condiciones apropiadas, como espacio, tiempo y recursos, para el desarrollo y aplicación de estrategias que permitan al Programa de Tutorías cumplir con sus objetivos.
  - ▶ A la fecha existen pocos indicadores que permitan determinar qué o cuáles estrategias están funcionando.
  - ▶ Existen pocas o nulas actividades establecidas actualmente en el Programa de Tutorías y éstas se enfocan en la fase de integración y no en que el alumno concrete metas progresivamente.
  - ▶ No hay un organismo que regule, supervise y evalúe el trabajo tutorial de los profesores de tiempo completo, a los que se les asignó, muchas veces por sorteo, un grupo de tutoría.
  - ▶ Existe un ambiente de apatía generalizado sobre la efectividad del Programa de Tutorías.

---

Actualmente esta situación no ha cambiado mucho, a pesar de que ya está establecido un Procedimiento de Tutorías Académicas, sin embargo, presenta deficiencias en su operación, adolece de un sistema de información con procesamiento especializado, que proporcione al tutor conocimiento significativo y trascendente, a través de una herramienta de cómputo para la toma de decisiones, utilizando tecnologías de información de vanguardia.

Esta situación provoca, que se tengan algunos problemas como la mala interpretación de funciones asignadas; esto se observa cuando el personal docente está más preocupado por la entrega a tiempo de formatos relacionados, que por el contenido mismo, el cual requiere de interpretación y análisis para la gestión y toma de decisiones, en beneficio del alumno.

Al momento de gestionar una asesoría, se tiene desconocimiento de los planes y programas de estudio de la carrera, es decir, desde el momento que un alumno reprueba una evaluación parcial, el tutor no conoce a fondo los temas de materias que no tienen que ver con su área de conocimiento, por lo tanto se le dificulta canalizar al alumno a diferentes opciones de asesoría sobre dicho tema.

El tutor se enfrenta al desconocimiento de los perfiles profesionales y experiencia laboral de la planta docente; lo que le impide también, ofrecerle al alumno de manera oportuna, una solución de asesoría, de temas particulares que quizá puedan dominar más de un profesor adscrito a la institución y que no necesariamente sea el profesor de la materia o de la carrera en cuestión.

Aunado a lo anterior, el tutor tiene desconocimiento de la disponibilidad de horarios extraclase, tanto de profesores de tiempo completo como de los profesores que cubren interinatos, por lo tanto es difícil contactarlos fuera de su horario asignado y más aún solicitarles sin previo aviso el apoyo.

---

Es más que evidente el desperdicio de recursos, por ejemplo en el caso de necesitar un documento que demuestre que se llevó a cabo la tutoría, se imprimen una serie de formatos que se almacenan y que solo sirven para sacarlos de evidencia en algún proceso de acreditación o para obtener el Perfil PROMEP (Programa para el Mejoramiento del Profesorado), sin embargo, su propósito real deberá ser la generación de conocimiento de cada alumno a través del registro de bitácoras, que le permitan al tutor dar un verdadero seguimiento, haciendo uso de las nuevas tecnología, durante toda la trayectoria académica de su tutorado.

Cabe mencionar que la labor tutorial no sólo se avoca al desempeño académico del alumno, sino que también funge como un escucha y una guía para resolver o canalizar situaciones personales expuestas por su tutorado.

Sin embargo, otro de los problemas que se presentan es que el tutor no cuenta con información sobre las funciones que realizan los diferentes departamentos internos y/o instancias de apoyo externas, que lo puedan auxiliar al respecto.

En la experiencia de los funcionarios de departamentos académicos universitarios relacionadas con actividades de control de tutorías y asesorías, es evidente la necesidad de disponer de un sistema de gestión y seguimiento, basado en un modelo inteligente de tutoría.

Dicho modelo puede servir como herramienta en la toma de decisiones más adecuada y a corto plazo para establecer estrategias a partir del conocimiento oportuno de los perfiles tanto de alumnos como de profesores, que resulten de la aplicación de minería de datos.

---

## 2.8 La Minería de Datos.

Se define la minería de datos como *“el proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos”* (Clark & Boswell, 2000), es decir, la tarea fundamental de la minería de datos es encontrar modelos que puedan ser entendidos a partir de los datos.

Una segunda definición describe que la minería de datos *“es el proceso analítico, diseñado para explorar grandes cantidades de datos, con el objetivo de encontrar relaciones entre las diferentes variables para aplicarlas a nuevos conjuntos de datos”* (Pérez, 2006).

Es un área multidisciplinaria donde convergen diferentes paradigmas de computación como son la construcción de árboles de decisión, la inducción de reglas, las redes neuronales artificiales, el aprendizaje basado en instancias, el aprendizaje bayesiano, la programación lógica y los algoritmos estadísticos (Romero Morales, Ventura Soto, & Hervas Martínez, 2012).

La tarea fundamental de la minería de datos es encontrar modelos inteligibles a partir de los datos. Para que este proceso sea efectivo debería ser automático o semi – automático (asistido) y el uso de los patrones descubiertos debería ayudar a tomar decisiones más seguras que reporten por tanto un beneficio a la organización.

---

Al obtener dicho conocimiento significativo es necesario administrarlo, para esto surge la Gestión del conocimiento y se define como *“la aplicación de un proceso sistemático que se basa en la capacidad de seleccionar, organizar, presentar y usar la información, por parte de los miembros de la organización, con el objeto de utilizar en forma cooperativa los recursos de conocimiento basados en el capital intelectual propio, con la finalidad de desarrollar las aptitudes organizacionales y la generación de valor”* (Passoni, 2012).

Para que este proceso pueda ser efectivo debería ser automático o asistido y el uso de los patrones descubiertos debería ayudar a tomar decisiones más seguras que reporten por tanto, algún beneficio a la organización.

Dos son los retos de la minería de datos: por un lado trabajar con grandes volúmenes de datos, procedentes mayoritariamente de sistemas de información con los problemas que ello conlleva y por el otro usar técnicas adecuadas para analizar los mismos y extraer conocimiento novedoso y útil (Hernández Orallo, Ramírez Quintana, & Ferri Ramírez, 2004).

De una manera simplista pero ambiciosa se puede decir que el objetivo de la minería de datos es la extracción del conocimiento.

### **2.8.1 Tipos de Datos para Minería de Datos.**

**Base de datos relacionales:** Es una colección de relaciones (tablas). Cada tabla consiste en un conjunto de atributos (columnas o campos) y puede contener un gran número de tuplas (registros o filas). Cada tupla representa un objeto el cual se describe a través del valor de sus atributos y se caracteriza por poseer una clave única o primaria que lo identifica. Una relación puede además tener claves ajenas, es decir, atributos que hagan referencia a otra relación.



---

Una de las principales características de las bases de datos relacionales es la existencia de un esquema asociado, es decir, los datos deben seguir una estructura.

La integridad de los datos se expresa a través de restricciones de integridad. Estas pueden ser de dominio (restringen el valor que puede tomar un atributo respecto a su dominio y puede tomar valores nulos o no), de identidad (por ejemplo la clave primaria tiene que ser única) y referencial (los valores de las claves ajenas se deben corresponder con uno y sólo un valor de la tabla referenciada).

Las bases de datos relacionadas son la fuente de datos para la mayoría de aplicaciones de minería de datos. Lógicamente, mediante una consulta podemos combinar en una sola tabla o vista minable, aquella información de varias tablas que se requiera para cada tarea concreta de minería de datos.

Por lo tanto, la representación tabular, también llamada atributo – valor, es la más utilizada por las técnicas de minería de datos. De las técnicas más habituales interesa distinguir sólo entre dos tipos:

- ▶ **Atributos numéricos:** Contienen valores enteros o reales. Por ejemplo, salario o edad.
- ▶ **Atributos categóricos o nominales:** Toman valores en un conjunto finito y preestablecido de categorías. Por ejemplo, sexo o nombre del departamento.

**Base de datos espaciales:** Contienen información relacionada con el espacio físico en un sentido amplio, como datos geográficos, imágenes médicas, redes de transporte o información de tráfico.

---

La minería de datos sobre estas bases de datos permite encontrar patrones sobre los datos, como por ejemplo, las características de las casas en una zona montañosa, la planificación de nuevas líneas del metro en función de las distancias de las distintas áreas a las líneas existentes, etc.

**Bases de datos temporales:** Almacenan datos que incluyen muchos atributos relacionados con el tiempo o en el que éste es muy relevante. Estos atributos pueden referirse a distintos instantes o intervalos temporales. En este tipo de base de datos las técnicas de minería de datos pueden utilizarse para encontrar las características de la evolución o las tendencias del cambio de distintas medidas o valores de las bases de datos.

**Base de datos documentales:** Contienen descripciones para los objetos (documentos de texto) que pueden ir desde las simples palabras claves a los resúmenes. Esta base de datos puede contener documentos no estructurados (biblioteca digital), semi – estructurados (extraer información por partes) o estructurados (fichas bibliográficas).

Las técnicas de minería de datos pueden utilizarse para obtener asociaciones entre los contenidos, agrupar o clasificar objetos textuales. Para ello los métodos de minería se integran con otras técnicas de recuperación de información y con la construcción y uso de jerarquías específicas para datos textuales, como los diccionarios.

**Base de datos multimedia:** Almacenan imágenes, audio y vídeo. Soportan objetos de gran tamaño. Para la minería de estas bases de datos también es necesario integrar los métodos de minería con técnicas de búsqueda y almacenamiento.

---

**World Wide Web:** Es el almacén de información más grande y diverso de los existentes en la actualidad, por lo tanto hay gran cantidad de datos en la Web de los que se puede extraer conocimiento relevante y útil.

La minería de datos o descubrimiento del conocimiento se ha utilizado en el desarrollo actual de los sistemas de enseñanza basada en Web o sistemas denominados e – learning, como herramientas para poder guiar a los estudiantes durante su aprendizaje y maximizarlo.

Minar la Web no es un problema sencillo, debido a que muchos de los datos son no estructurados o semi – estructurados, a que muchas páginas contienen datos multimedia, y a que estos datos pueden residir en diversos servidores o en archivos.

Otros aspectos que dificultan la minería Web son cómo determinar a qué páginas debemos acceder y cómo seleccionar la información que va a ser útil para extraer conocimiento. Toda esta diversidad hace que la minería Web se organice en tres categorías:

- ▶ **Minería del contenido:** Encontrar patrones de los datos de las páginas Web.
- ▶ **Minería de la estructura:** Entendiendo como estructura los hipervínculos y URL's.
- ▶ **Minería del uso:** La navegación de las páginas Web por parte del usuario.

La minería de datos tiene como objetivo analizar los datos para extraer conocimiento. Este conocimiento se puede representar en forma de relaciones, patrones o reglas inferidas de los datos previamente desconocidos, o bien en forma de descripción más concisa como un resumen.

---

## 2.8.2 Técnicas de Minería de Datos.

La clasificación inicial de las técnicas de minería de datos distingue entre técnicas predictivas, en la que las variables pueden clasificarse inicialmente en dependientes e independientes, técnicas descriptivas, en las que todas las variables tienen inicialmente el mismo estatus y técnicas auxiliares (Pérez, 2006).

Las **técnicas predictivas** especifican el modelo para los datos con base en un conocimiento teórico previo. El modelo supuesto para los datos debe contratarse después del proceso de minería de datos antes de aceptarlo como válido. Formalmente, la aplicación de todo modelo debe superar las fases de *identificación objetiva* (a partir de los datos se aplican reglas que permitan identificar el mejor modelo posible que ajuste los datos), *estimación* (proceso de cálculo de los parámetros del modelo elegido de los datos en la fase de identificación), *diagnos*is (proceso de contraste de la validez del modelo estimado) y *predicción* (proceso de utilización del modelo identificado, estimado y validado para predecir valores futuros de las variables dependientes). En algunos casos, el modelo se obtiene como mezcla del conocimiento obtenido antes y después de la Minería de Datos y también debe contrastarse antes de aceptarse como válido.

En las **técnicas descriptivas** no se asigna ningún papel predeterminado a las variables. No se supone la existencia de variables dependientes ni independientes y tampoco se supone la existencia de un modelo previo para los datos. Los modelos se crean automáticamente partiendo del reconocimiento de patrones.

Tanto las técnicas predictivas como las técnicas descriptivas están enfocadas al descubrimiento del conocimiento embebido<sup>1</sup> en los datos.

---

<sup>1</sup> Embebido: Dicho de una cosa inmaterial: Incorporar, incluir dentro de sí a otra. *Real Academia Española*.

Las **técnicas auxiliares** son herramientas de apoyo más superficiales y limitadas. Se trata de nuevos métodos basados en técnicas estadísticas descriptivas, consultas e informes y enfocados en general hacia la *verificación*.

En la Figura 2.2 se muestra una clasificación de técnicas conocidas:

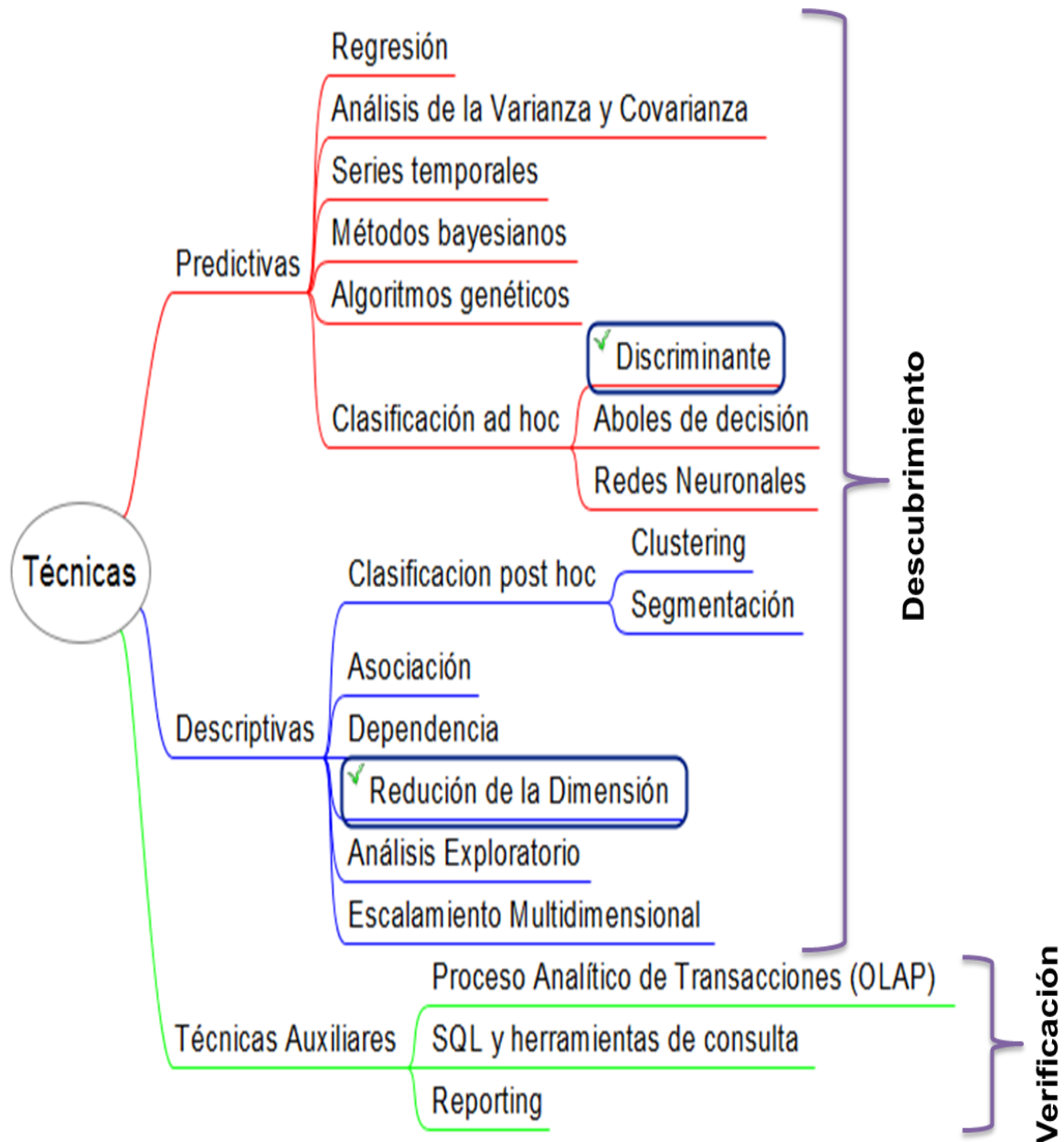


Figura 2.2 Esquema de Técnicas de Minería de Datos.

---

Se observa que las técnicas de clasificación pueden pertenecer tanto al grupo de técnicas predictivas como a las descriptivas. Las técnicas de clasificación predictivas suelen denominarse técnicas de clasificación *ad hoc*<sup>2</sup> ya que clasifican individuos u observaciones dentro de grupos previamente definidos.

Las técnicas de clasificación descriptivas se denominan técnicas de clasificación *post hoc*<sup>3</sup> porque realizan clasificación sin especificación previa de los grupos.

## **2.9 La minería de datos y el proceso de descubrimiento del conocimiento.**

Existen términos que se utilizan como sinónimos de la minería de datos. Uno de ellos es el “análisis (inteligente) de datos” que suele hacer un mayor hincapié en las técnicas de análisis estadístico.

Otro término muy utilizado, y más relacionado con la minería de datos, es la extracción o “descubrimiento de conocimiento en base de datos”, “Knowledge Discovery Databases (KDD)”, que se refiere a un proceso que consta de una serie de fases, siendo la minería de datos sólo una de estas fases.

En Fayyad et al. (1996), se define el KDD como “el proceso no trivial de identificar patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles a partir de los datos”. En esta definición se resumen cuáles deben ser las propiedades deseables del conocimiento extraído.

Como se deduce de la anterior definición, el KDD es un proceso complejo que incluye no sólo la obtención de los modelos o patrones (objetivo de la minería de datos), sino también la evaluación y posible interpretación de los mismos, tal y como se refleja en la Figura 2.3:

---

<sup>2</sup> ad hoc: Que es apropiado o está dispuesto especialmente para un fin. Real Academia Española (RAE, 2005).

<sup>3</sup> post hoc: Significa “después de esto, por lo tanto, a consecuencia de esto”. Wikipedia.

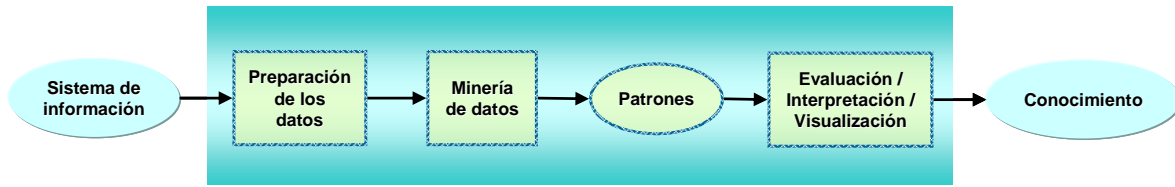


Figura 2.3 Proceso de Knowledge Discovery Databases (KDD).

Así los sistemas de KDD permiten la selección, limpieza, transformación y proyección de los datos; analizar los datos para extraer patrones y modelos adecuados; evaluar e interpretar los patrones para convertirlos en conocimiento; consolidar el conocimiento resolviendo posibles conflictos con conocimiento previamente extraído; y hacer el conocimiento disponible para su uso.

### 2.9.1 Relación con otras disciplinas.

La minería de datos es un campo multidisciplinar que se ha desarrollado en paralelo o como prolongación de otras tecnologías. Por ello, la investigación y los avances en la minería de datos se nutren de los que se producen en estas áreas relacionadas, como se muestra en la Figura 2.4:



Figura 2.4 Disciplinas que contribuyen a la minería de datos.

---

**Las bases de datos:** Conceptos como almacenes de datos, el procesamiento analítico en línea, las técnicas de indexación y de acceso eficiente a los datos son muy relevantes para el diseño de algoritmos eficientes de minería de datos.

**La recuperación de información:** Consiste en obtener información a partir de datos textuales, por lo que su desarrollo histórico se ha basado en el uso efectivo de bibliotecas (recientemente digitales) y en la búsqueda por Internet.

**La estadística:** Esta disciplina ha proporcionado muchos de los conceptos, algoritmos y técnicas que se utilizan en minería de datos, como por ejemplo, la media, la varianza, las distribuciones, el análisis multivariante, la regresión lineal y no lineal, la teoría del muestreo, la validación cruzada, entre otras.

**Los sistemas para la toma de decisión:** Son herramientas y sistemas de información que asisten a los directivos en la solución de problemas y en la toma de decisiones. El objetivo es proporcionar la información necesaria para realizar decisiones efectivas en el ámbito empresarial o en tareas de diagnóstico.

**La visualización de datos:** El uso de técnicas de visualización permite al usuario descubrir, intuir o entender patrones que serían difíciles de ver a partir de descripciones matemáticas o textuales de los resultados. Existen técnicas de visualización como, por ejemplo, las gráficas (diagramas de barras, gráficas de dispersión, histogramas, etc.), las icónicas (basadas en figuras, colores, etc.), las basadas en píxeles (cada dato se representa como un único píxel), las jerárquicas (dividiendo el área de representación en regiones dependiendo de los datos), entre otras muchas.



**La computación paralela y distribuida:** El costo computacional de las tareas más complejas de minería de datos se reparte entre diferentes procesadores o computadoras. Una de las ventajas del procesamiento en paralelo y distribuido es la escalabilidad de los algoritmos, lo que lo hace idóneo para estas aplicaciones.

**Otras disciplinas:** Dependiendo del tipo de datos a ser minados o del tipo de aplicación, la minería de datos usa también técnicas de otras disciplinas como el lenguaje natural, el análisis de imágenes, el procesamiento de señales, los gráficos por computadora, etc.

### 2.9.2 Aplicaciones.

La integración de las técnicas de minería de datos en las actividades del día a día se está convirtiendo en algo habitual. En la Tabla 2.1, se muestran las áreas en las que se puede usar la minería de datos:

---

**Tabla 2.1 Diferentes áreas para la extracción del conocimiento**

---

- ▶ Aplicaciones financieras y banca.
  - ▶ Análisis de mercado, distribución, y en general, comercio.
  - ▶ Seguros y salud privada.
  - ▶ Educación.
  - ▶ Procesos industriales.
  - ▶ Medicina.
  - ▶ Biología, bioingeniería y otras ciencias.
  - ▶ Telecomunicaciones.
  - ▶ Correo electrónico y agendas electrónicas.
  - ▶ Recursos humanos.
  - ▶ Web.
  - ▶ Turismo.
  - ▶ Tráfico.
  - ▶ Hacienda.
  - ▶ Policiales.
  - ▶ Deportes.
  - ▶ Política.
  - ▶ Etcétera.
-

---

Es importante destacar algunas de las aplicaciones en mención, a partir del impacto que estas tienen y que cada una se vuelve una actividad natural en los sistemas de información.

**Aplicaciones financieras y la banca:**

- ▶ Obtención de patrones de uso fraudulento de tarjetas de crédito.
- ▶ Determinación del gasto en tarjeta de crédito por grupos.
- ▶ Cálculo de correlaciones entre indicadores financieros.
- ▶ Identificación de reglas de mercado de valores a partir de históricos.
- ▶ Análisis de riesgos en créditos.

**Análisis de mercado, distribución y, en general, comercio:**

- ▶ Análisis del carrito de compras.
- ▶ Evaluación de campañas publicitarias.
- ▶ Análisis de la fidelidad de los clientes.
- ▶ Segmentación de clientes.
- ▶ Estimación de stock<sup>4</sup>, de costos, de ventas, etcétera.

**Educación:**

- ▶ Selección o captación de estudiantes.
- ▶ Detección de deserción y reprobación.
- ▶ Estimación del tiempo de estancia en la institución.

**Biología, bioingeniería y otras ciencias:**

- ▶ Análisis de secuencias de genes.
- ▶ Análisis de secuencias de proteínas.
- ▶ Predecir si un compuesto químico causa cáncer.
- ▶ Clasificación de cuerpos celestes.
- ▶ Predicción de recorrido y distribución de inundaciones.
- ▶ Modelo de calidad de aguas, indicadores ecológicos.

---

<sup>4</sup> Cantidad de mercancías que se tienen en depósito. *Real Academia Española*

### Telecomunicaciones:

- ▶ Establecimiento de patrones de llamadas.
- ▶ Modelo de cargas en redes.
- ▶ Detección de fraude.

### Otras áreas:

- ▶ Correo electrónico y agendas personales: Clasificación y distribución automática de correo, detección de spam<sup>5</sup>
- ▶ Recursos Humanos: Selección de personal.
- ▶ Web: Análisis del comportamiento de los usuarios, detección de fraude en el comercio electrónico, análisis de los logs<sup>6</sup> de un servidor web.
- ▶ Turismo: Determinar las características socioeconómicas de los turistas en un determinado destino o paquete turístico, así como determinar patrones de reservas.
- ▶ Tráfico: Modelos de tráfico a partir de fuentes diversas como cámaras y GPS<sup>7</sup>.
- ▶ Hacienda: Detección de evasión fiscal.
- ▶ Policiales: Identificación de posibles terroristas en un aeropuerto.
- ▶ Deportes: Estudio de la influencia de jugadores y de cambios. Planificación de eventos.

Todos estos ejemplos muestran la gran variedad de aplicaciones donde el uso de la minería de datos puede ayudar a entender mejor el entorno donde se desenvuelve la organización y, en definitiva, la toma de decisiones en dicho entorno.

---

<sup>5</sup> Son mensajes no solicitados, no deseados o de remitente desconocido y que son sumamente molestos. [www.seguridadpc.net](http://www.seguridadpc.net)

<sup>6</sup> Uno o más ficheros de texto automáticamente creados y administrados por un servidor, en donde se almacena toda la actividad que se hace sobre éste. [www.alwgsa.com](http://www.alwgsa.com)

<sup>7</sup> Sistema que permite conocer la posición de un objeto móvil gracias a la recepción de señales emitidas por una red de satélites (Global Positioning System). *Real Academia Española*

---

## **2.10 El ITM, las Tutorías y la Minería de Datos.**

A partir de la instrucción que se dio por parte de las autoridades del ITM, de instrumentar un Programa Institucional de Tutorías, con base en el PNT 2006, no ha sido posible operar de manera adecuada.

En cada uno de los Departamentos Académicos, se han articulado esfuerzos aislados, sin lograr resultados medibles.

Por otra parte no se ha sabido aprovechar la información que se genera a partir de la aplicación del Examen de Ingreso (EXANI-II) de CENEVAL como mecanismo de selección a nivel superior en el ITM.

Toda la información que se obtiene desde el momento que el aspirante paga su ficha de inscripción, ha sido desaprovechada, ya que a partir de los datos captados y almacenados se puede aplicar minería de datos para extraer información significativa que le permita al ITM obtener perfiles de alumnos de nuevo ingreso que los categoricen y que se les pueda brindar a través de la tutoría un mejor servicio de enseñanza – aprendizaje, enfocado a las necesidades que requiera dicho perfil y obtener resultados satisfactorios en los indicadores de desempeño escolar.

El ITM aplica dicho examen en dos períodos de ingreso al año y si se considera el ejercicio de Junio del 2011, se registraron 1,452 aspirantes, de los cuales aplicaron 1,405 sustentantes y aproximadamente el 85% fue inscrito para ser alumno del Instituto.

---

La forma de elegir a los aspirantes es sólo los que obtienen mayor puntuación, sin embargo, a partir de un estudio que se hizo respecto a la ponderación de las disciplinas que evalúa el EXANI – II, resulta que tiene la misma ponderación el Idioma Inglés que las Matemáticas, lo cual provoca que tenga más ventaja un aspirante que domine ese idioma, respecto al que no, cuando se considera que las matemáticas deberán tener una mayor ponderación a partir de que es la base de cualquier Ingeniería.

Otro aspecto a considerar, es que el ITM tiene un capital intelectual docente de aproximadamente 500 personas, de las cuales 350 son Profesores de Tiempo Completo y 150 Profesores por Asignatura, lo cual permite tener acceso a 500 currículos, en donde se describen aspectos no sólo técnicos, profesionales o de experiencia laboral, sino también aspectos de personalidad, que influyen en el trabajo tutorial, el cual se ofrece como servicio a los alumnos.

Por lo anterior, se propone en esta tesis, un Modelo Inteligente de Tutoría (MIT), en donde se aplica la minería de datos en los almacenes de datos descritos, generando perfiles de tutores y perfiles de alumnos idóneos para crear grupos con características particulares de acuerdo a su circunstancia y contexto, guiados por un profesor adecuado, para lograr disminuir la deserción y la diferencia entre el número de alumnos que ingresan con el número de alumnos que se titulan en una carrera profesional.

Con este proyecto se rompen paradigmas, al aplicar tecnologías de información, procesamiento de datos e inteligencia de negocios, para que junto con instrumentos psicométricos como 16 – FP y test proyectivos; se genere una base de conocimiento de los últimos dos años con los estudiantes de ingreso del 2011 a la fecha.

---

A todos los datos se les ha aplicado un pre–procesamiento, en el cual se ha subsanado cualquier omisión en el proceso de ingreso de datos por todas las partes que se encuentran involucradas.

Una vez que se ha llevado a cabo el pre – procesamiento, toda la información ha sido concentrada en una sola base de datos a donde se aplica toda la algoritmia de minería de datos que ha sido pre–seleccionada para conformar clústeres de estudiantes, así como de perfiles de docentes tutores.

A partir de que han quedado conformados los grupos de tutoría, simplemente es momento de asignar uno de los tutores sugeridos por el sistema y entonces, operar el Modelo Inteligente de Tutorías, retroalimentando en cada una de sus iteraciones, convirtiéndole en un sistema experto.

Solo el tiempo será quien determine si esta estrategia avanzada será finalmente la solución para el serio problema de Tutorías, que ha perdurado por varios años. La propuesta está terminada, es momento de iterar, y esperar los mejores resultados, para los fines académicos en cuanto Acreditación y Reacreditación de Planes de Estudio de Licenciatura presenciales, todo esto en pro de que el ITM no pierda su estandarte de un Instituto Tecnológico de alto desempeño, que por tanto tiempo ha mantenido, como un sello institucional.

## **2.11 Herramienta de cómputo XLSTAT 2014.**

XLSTAT es un complemento de estadística para Microsoft Excel que se ha venido desarrollando desde 1993 para mejorar las funciones de análisis de Excel. XLSTAT es un software estadístico modular creado con la base de XLSTAT-Pro, el producto principal de Addinsoft.

---

XLSTAT se apoya en Excel para la entrada de datos y la visualización de los resultados, pero los cálculos se realizan empleando componentes autónomos de software.

El uso de Excel como interfaz hace que XLSTAT sea un paquete de análisis de datos multivariados orientado al usuario y con funciones estadísticas altamente eficaces. La calidad de los cálculos es idéntica a la que ofrece el software clásico de análisis estadístico científico.

Es posible añadir un amplio abanico de módulos a XLSTAT – Pro para necesidades específicas de análisis. Además, se pueden personalizar módulos específicos a sus necesidades comerciales sobre estadística, desde el desarrollo de una función hasta la automatización de análisis rutinarios. Todo esto garantiza la constante evolución de XLSTAT para satisfacer sus requerimientos estadísticos. (XLSTAT, 1995 - 2014).

Fue precisamente esta herramienta que se utilizó para aplicar tres métodos estadísticos y uno algorítmico, en la base de datos generada por el EXANI – II de CENEVAL, aplicado a los aspirantes de una carrera en el año 2011 por el ITM.

---

### III. METODOLOGÍA

*“Estamos ahogándonos de información, pero famélicos de conocimiento”*

*John Naisbitt*

En esta tesis se propone un Modelo Inteligente de Tutorías (MIT), que se compone de cuatro fases: Captación, Formación, Seguimiento y Cierre, las cuales se llevan a cabo a través de los nueve semestres, que en condiciones normales, un estudiante cursa en su carrera profesional en el ITM.

Dichas fases tienen como resultado, a partir de un proceso de obtención del conocimiento, crear cuatro perfiles de estudiantes: Aspirante, Alumno en Formación, Alumno en Seguimiento y Alumno en Cierre, además de tres perfiles de profesores: Tutor de Formación, Tutor de Seguimiento y Tutor de Cierre.

El alcance técnico que tiene esta propuesta, en cuanto a la aplicación de minería de datos, es abarcar parte de la Fase 1 del MIT, denominada Captación, para poder crear el perfil del Aspirante; para lograrlo se utilizó el proceso de extracción del conocimiento, ya mencionado con anterioridad y conocido por sus siglas en inglés KDD, además de la herramienta de cómputo XLSTAT, para los análisis correspondientes.

Para la creación tanto del MIT como del Perfil de Aspirante, se realizó un cronograma de actividades (Anexo A.1), el cual tuvo un avance del 85% en tiempo y forma, el resto de las actividades se concretaron fuera del tiempo estimado, sin embargo, se logró al 100%.



### 3.1 El proceso de extracción del conocimiento.

El KDD, es un proceso iterativo e interactivo. Es iterativo ya que la salida de alguna de las fases puede hacer volver a pasos anteriores y porque a menudo son necesarias varias iteraciones para extraer conocimiento de alta calidad. Es interactivo porque el usuario, o un experto en el dominio del problema, debe ayudar en la preparación de los datos, validación del conocimiento extraído, etcétera.

El proceso del KDD se organiza en torno a cinco fases como se ilustra en la Figura 3.1. En la **fase de integración y recopilación de datos** se determinan las fuentes de información que pueden ser útiles y dónde conseguirlas. A continuación se transforman todos los datos a un formato común, frecuentemente mediante un almacén de datos que consiga unificar toda la información recogida, detectando y resolviendo las inconsistencias. Este almacén de datos facilita enormemente la “navegación” y visualización previa de sus datos para discernir qué aspectos pueden interesar que sean estudiados.

Dado que los datos provienen de diferentes fuentes, pueden contener valores erróneos o faltantes. Estas situaciones se tratan en la **fase de selección, limpieza y transformación**, en la que se eliminan o corrigen los datos incorrectos y se decide la estrategia a seguir con los datos incompletos.

Además, se proyectan los datos para considerar únicamente aquellas variables o atributos que van a ser relevantes, con el objetivo de hacer más fácil la tarea propia de minería y para que los resultados de la misma sean útiles. La selección incluye tanto una criba<sup>8</sup> o fusión horizontal (filas / registros) como vertical (columnas / atributos). Las dos primeras fases se suelen englobar bajo el nombre de “preparación de datos”.

---

<sup>8</sup> Criba: Selección rigurosa. Someter a criba. Superar pasar la criba. *Real Academia Española*

En la **fase de minería de datos**, se decide cuál es la tarea a realizar (clasificar, agrupar, etcétera) y se elige el método que se va a utilizar.

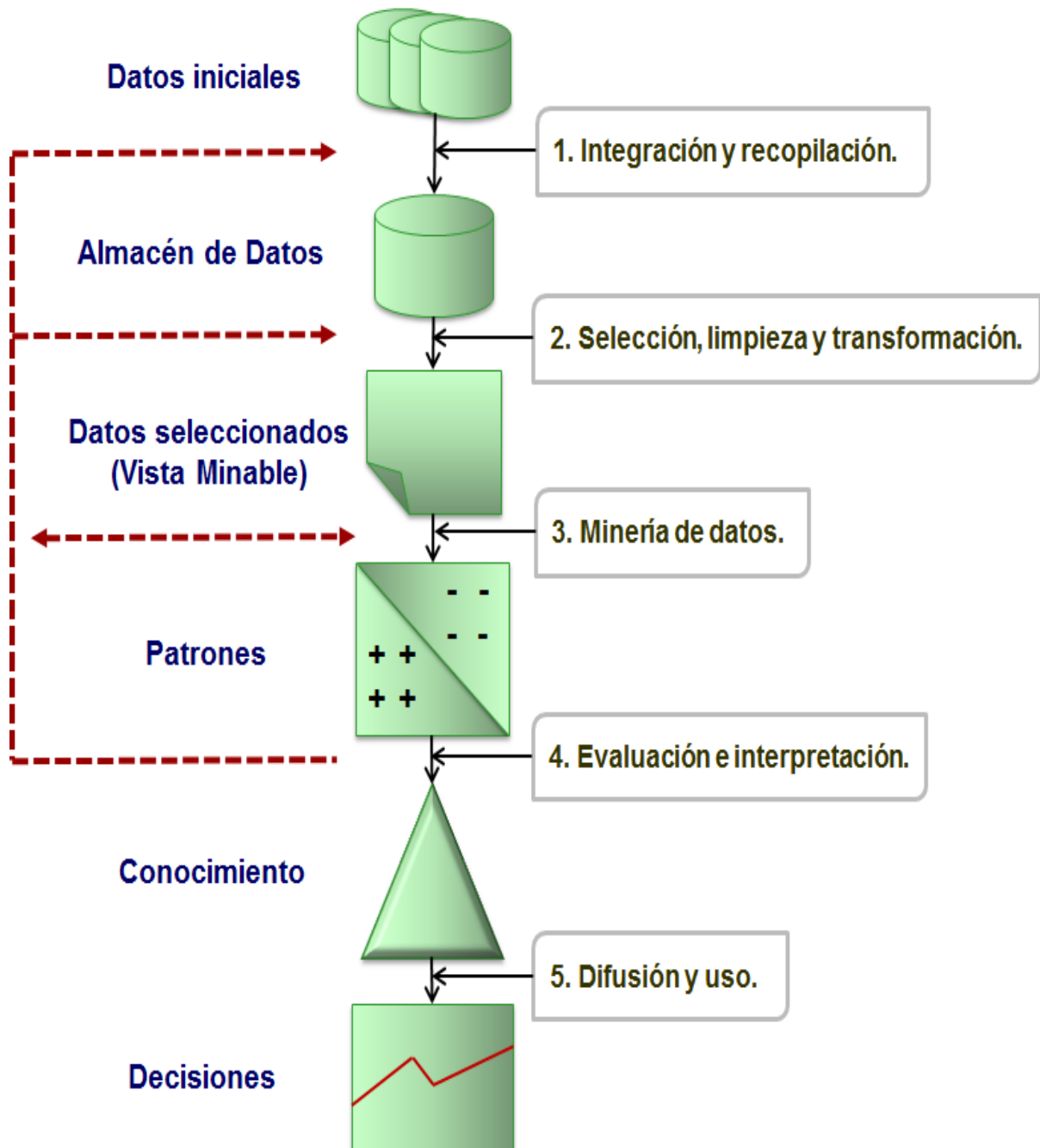


Figura 3.1 Fases del proceso KDD.

---

En la **fase de evaluación e interpretación** se evalúan los patrones y se analizan por los expertos, y si es necesario se vuelve a las fases anteriores para una nueva iteración. Esto incluye resolver posibles conflictos con el conocimiento que se disponía anteriormente.

Finalmente, en la **fase de difusión y uso** se aplica el nuevo conocimiento y se hace partícipe de él a todos los posibles usuarios.

Para cada una de estas fases se emplean distintas técnicas de las diferentes disciplinas relacionadas y mencionadas en la Figura 2.4. Además de las fases descritas, frecuentemente se incluye una fase previa de análisis de las necesidades de la organización y definición del problema, en la que se establecen los objetivos de minería de datos.

### **3.2 Fase 1. Integración y recopilación.**

Para analizar y extraer algo útil de los datos, es preciso disponer de ellos. En general, el problema de reunir un conjunto de datos que posibilite la extracción del conocimiento requiere decidir, entre otros aspectos, de qué fuentes internas y externas se van a obtener los datos, cómo se van a organizar, cómo se van a mantener con el tiempo y finalmente, de qué forma se van a poder extraer parcialmente o totalmente, en detalle o agregados, con distintas “vistas minables” a las que se pueda aplicar las herramientas concretas de minería de datos (Hernández Orallo, Ramírez Quintana, & Ferri Ramírez, 2004).

Como ya se mencionó con anterioridad, la fuente de datos externa fue el EXANI – II, diseñado y aplicado por CENEVAL, a un total 1,086 sustentantes que sacaron ficha en Junio del 2011, para ingresar a las diversas carreras que ofrece el ITM.

Los resultados de dicho examen se registran en un archivo en Excel, el cual contienen en su Hoja 1, una tabla de 134 columnas, cada una de ellas representa un atributo o variable de la evaluación, además de 1,086 filas que representan los registros de cada uno de los aspirantes a los que se les aplicó.

En la Tabla 3.1 se enlistan los campos o atributos, así como su descripción por orden de aparición de izquierda a derecha.

**Tabla 3.1 Variables y su descripción de información del EXANI – II.**

No. C.	Atributo	Descripción
1	TIPO_EXA	TIPO DE EXAMEN
2	OPC_APLI	OPCIÓN DE APLICACIÓN
3	ANO_VER	AÑO DE LA VERSIÓN
4	TIPO_REG	TIPO DE REGISTRO
5	TIPO_RESP	TIPO DE RESPUESTAS
6	CVE_BPM	CLAVE DE IDENTIFICACIÓN DE LA APLICACIÓN
7	APLI	NÚMERO DE APLICACIÓN
8	FECHA_APLI	FECHA DE APLICACIÓN
9	CVE_INST	CLAVE DE LA INSTITUCIÓN EN DONDE SE LLEVÓ A CABO LA APLICACIÓN
10	IDENTIFICA	IDENTIFICACIÓN DE LA APLICACIÓN
11	DESC_IDENT	DESCRIPCIÓN DE LA CLAVE IDENTIFICA
12	LPOS_IMG	LOCALIZACIÓN DE LAS IMÁGENES PARA HOJAS DE RESPUESTA Y DE REGISTRO EN LA BASE DE DATOS
13	FOLIO	FOLIO
14	MATRICULA	ANOTE LA CLAVE O MATRÍCULA DE LA INSTITUCIÓN
15	APE_PAT	APELLIDO PATERNO
16	APE_MAT	APELLIDO MATERNO
17	NOMBRE	NOMBRE
18	DIA_NAC	DÍA DE NACIMIENTO
19	MES_NAC	MES DE NACIMIENTO
20	ANO_NAC	AÑO DE NACIMIENTO
21	SEXO	SEXO
22	CON_DCE	IMPEDIMENTO O DIFICULTAD PARA CAMINAR O SUBIR ESCALERAS
23	CON_IMPE	IMPEDIMENTO O DIFICULTAD PARA ESCRIBIR

No. C.	Atributo	Descripción
24	CON_ESC	PROBLEMAS PARA ESCUCHAR, AÚN A CORTA DISTANCIA
25	CON_VER	PROBLEMAS PARA VER (QUE NO SE CORRIJAN CON LENTES)
26	CON_LEN	PROBLEMAS DE LENGUAJE
27	LI_MAD	¿SU MADRE HABLA UNA LENGUA INDÍGENA?
28	LI_PAD	¿SU PADRE HABLA UNA LENGUA INDÍGENA?
29	EDO_PROC	¿EN QUÉ ENTIDAD CONCLUYÓ EL BACHILLERATO?
30	NOM_PROC	NOMBRE DE LA INSTITUCIÓN EN LA QUE CONCLUYÓ EL BACHILLERATO
31	CIU_PROC	NOMBRE DE LA CIUDAD DONDE SE UBICA LA INSTITUCIÓN EN LA QUE CONCLUYÓ EL BACHILLERATO; SI FUE EN EL EXTRANJERO INDIQUE EL PAÍS.
32	CVE_PROC	CLAVE DE LA INSTITUCIÓN EN LA QUE CONCLUYÓ SUS ESTUDIOS DE BACHILLERATO
33	REG_PROC	¿EN QUÉ RÉGIMEN DE ESCUELA ESTUDIO EL ÚLTIMO AÑO DE BACHILLERATO?
34	ANO_BAC	¿EN QUÉ AÑO CONCLUYÓ EL BACHILLERATO?
35	MOD_BAC	¿EN QUÉ MODALIDAD OBTUVO SU CERTIFICADO DEL BACHILLERATO?
36	PROM_BAC	¿CUÁL FUE SU PROMEDIO GENERAL EN EL BACHILLERATO?
37	BEC_BDAC	¿RECIBIÓ BECA EN EL BACHILLERATO POR DESEMPEÑO ACADÉMICO?
38	BEC_BNE	¿RECIBIÓ BECA EN EL BACHILLERATO POR NECESIDAD ECONÓMICA?
39	BEC_BHD	¿RECIBIÓ BECA EN EL BACHILLERATO POR HABILIDAD DEPORTIVA O ARTÍSTICA?
40	POR_CCE	CRITERIOS DE EVALUACIÓN (Aspecto que dio a conocer los profesores al inicio de los cursos, en el último año de bachillerato)
41	POR_OBAP	OBJETIVOS DE APRENDIZAJE (Aspecto que dio a conocer los profesores al inicio de los cursos, en el último año de bachillerato)
42	POR_BIB	BIBLIOGRAFÍA (Aspecto que dio a conocer los profesores al inicio de los cursos, en el último año de bachillerato)
43	PPR_ACM	¿QUÉ PORCENTAJE DE LOS PROFESORES TENÍA AMPLIOS CONOCIMIENTOS DE LA MATERIA?
44	PPR_EJE	¿QUÉ PORCENTAJE DE LOS PROFESORES DABA CON FRECUENCIA EJEMPLOS QUE PERMITÍAN ENTENDER LA APLICACIÓN DE LOS CONCEPTOS REVISADOS?
45	PPR_TAR	¿QUÉ PORCENTAJE DE LOS PROFESORES SOLICITABA TAREAS O TRABAJOS QUE CONSOLIDABAN LOS TEMAS REVISADOS?
46	PPR_PUN	¿QUÉ PORCENTAJE DE LOS PROFESORES ASISTÍA REGULARMENTE A CLASE?
47	PPR_DUD	¿QUÉ PORCENTAJE DE LOS PROFESORES PERMITÍA QUE LOS ESTUDIANTES PRESENTARAN SUS DUDAS?
48	PPR_ASI	¿QUÉ PORCENTAJE DE LOS PROFESORES ASISTÍA PUNTUALMENTE A CLASE?
49	FEV_ENO	LOS ESTUDIANTES NO PRESTABAN ATENCIÓN A LO QUE LOS MAESTROS DECÍAN (¿En qué porcentaje ocurría?)
50	FEV_RUDE	HABÍA RUIDO Y DESORDEN (¿En qué porcentaje ocurría?)
51	FEV_TIES	LOS PROFESORES TENÍAN QUE ESPERAR MUCHO TIEMPO PARA QUE LOS ESTUDIANTES PUSIERAN ATENCIÓN (¿En qué porcentaje ocurría?)
52	FEV_TEFE	EL TIEMPO EFECTIVO DE TRABAJO EN CLASE ERA MUY POCO (¿En qué porcentaje ocurría?)
53	NIV_IDAT	IDENTIFICAR DATOS Y HECHOS (Nivel de habilidad para realizar la actividad al leer)
54	NIV_EREL	ESTABLECER RELACIONES ENTRE DATOS, HECHOS Y JUICIOS (Nivel de habilidad para realizar la actividad al leer)
55	NIV_RIDE	RECONOCER IDEAS CENTRALES (PRINCIPALES Y SECUNDARIAS) (Nivel de habilidad para realizar la actividad al leer)
56	NIV_RPUN	RECONOCER LOS PUNTOS DE VISTA IMPLÍCITOS Y EXPLÍCITOS DEL AUTOR (Nivel de habilidad para realizar la actividad al leer)
57	CUA_LIB	EN LOS ÚLTIMOS SEIS MESES, ¿CUÁNTOS LIBROS HA LEÍDO (NO CUENTE LOS QUE LE DEJARON LEER EN LA ESCUELA)?

No. C.	Atributo	Descripción
58	NIV_COH	DE FORMA COHERENTE (Nivel de habilidad para redactar un texto)
59	NIV_ERR	SIN ERRORES ORTOGRÁFICOS (Nivel de habilidad para redactar un texto)
60	NIV_RESL	EL RESUMEN DE UN LIBRO (Nivel de habilidad para escribir)
61	NIV_ENSA	UN ENSAYO (Nivel de habilidad para escribir)
62	NIV_CART	UNA CARTA (Nivel de habilidad para escribir)
63	NIV_REPO	UN REPORTE DE INVESTIGACIÓN (Nivel de habilidad para escribir)
64	NIV_DOCO	UN DOCUMENTO OFICIAL (POR EJEMPLO UNA QUEJA) (Nivel de habilidad para escribir)
65	NIV_EXP	EXPONER UN TEMA ACADÉMICO FRENTE A UN GRUPO (Nivel de habilidad para ...)
66	NIV_DUDA	EXPONER UNA DUDA A SU MAESTRO (Nivel de habilidad para ...)
67	NIV_DEBA	DEFENDER SU PUNTO DE VISTA EN UN DEBATE (Nivel de habilidad para ...)
68	HRS_TRAB	MIENTRAS ESTUDIABA EL BACHILLERATO, ¿CUÁNTAS HORAS A LA SEMANA DEDICABA USTED A TRABAJAR?
69	EST_ALCA	¿CUÁL ES EL NIVEL MÁXIMO DE ESTUDIOS QUE LE GUSTARÍA ALCANZAR?
70	FAM_EXA	¿QUÉ TAN FAMILIARIZADO ESTÁ CON LOS EXÁMENES DE OPCIÓN MÚLTIPLE?
71	PRE_EXA1	¿USTED PRESENTÓ EL EXAMEN NACIONAL DE INGRESO A LA EDUCACIÓN MEDIA SUPERIOR (EXANI - I) PARA INGRESAR AL BACHILLERATO?
72	PRE_EXA2	¿CUÁNTAS VECES HA PRESENTADO EL EXAMEN NACIONAL DE INGRESO A LA EDUCACIÓN SUPERIOR (EXANI – II)?
73	VIVE_MAD	¿VIVE CON SU MADRE?
74	VIVE_PAD	¿VIVE CON SU PADRE?
75	TRAB_MAD	¿A QUÉ SE DEDICA SU MADRE O TUTORA ACTUALMENTE?
76	TRAB_PAD	¿A QUÉ SE DEDICA SU PADRE O TUTOR ACTUALMENTE?
77	ESCO_MAD	¿CUÁL ES EL MÁXIMO NIVEL DE ESTUDIOS ALCANZADO POR SU MADRE (AUNQUE HAYA FALLECIDO)?
78	ESCO_PAD	¿CUÁL ES EL MÁXIMO NIVEL DE ESTUDIOS ALCANZADO POR SU PADRE (AUNQUE HAYA FALLECIDO)?
79	CUAN_LIB	¿CUÁNTOS LIBROS HAY EN SU CASA? (No consideres revistas, periódicos o sus libros de texto)
80	CUAN_PELI	¿CUÁNTAS PELÍCULAS HAY EN SU CASA?
81	EXP_PAD	¿QUÉ NIVEL EDUCATIVO QUISIERAN SUS PADRES QUE ALCANZARA?
82	SER_TELE	LÍNEA TELEFÓNICA (Bienes y servicios con los que cuenta en casa)
83	SER_LAV	LAVADORA DE ROPA (Bienes y servicios con los que cuenta en casa)
84	SER_REF	REFRIGERADOR (Bienes y servicios con los que cuenta en casa)
85	SER_HOR	HORNO DE MICROONDAS (Bienes y servicios con los que cuenta en casa)
86	SER_INTE	INTERNET (Bienes y servicios con los que cuenta en casa)
87	SER_CABL	TELEVISIÓN DE PAGA (CABLEVISIÓN, SKY, ETC.) (Bienes y servicios con los que cuenta en casa)
88	SER_DVD	¿CUÁNTOS REPRODUCTOR DE DVD HAY EN SU CASA?
89	BIEN_PC	¿CUÁNTAS COMPUTADORAS HAY EN SU CASA?
90	SER_TV	¿CUÁNTOS TELEVISORES HAY EN SU CASA?
91	SER_AUTO	¿CUÁNTOS AUTOMÓVILES HAY EN SU CASA?
92	SER_BANO	¿CUÁNTOS BAÑOS COMPLETOS HAY EN SU CASA?
93	CINE	ASISTIR AL CINE (Actividad realizada en los últimos seis meses)
94	MUSEO	ASISTIR A MUSEOS (Actividad realizada en los últimos seis meses)

No. C.	Atributo	Descripción
95	ESPEC	ASISTIR A ESPECTÁCULOS ARTÍSTICOS (TEATRO, DANZA, CONCIERTO, ETC.) (Actividad realizada en los últimos seis meses)
96	VAC_RM	EN LOS ÚLTIMOS 2 AÑOS, ¿CUÁNTAS VECES HA SALIDO DE VACACIONES DENTRO DE LA REPÚBLICA MEXICANA? (Actividad realizada en los últimos seis meses)
97	EDO_REP	¿CUÁNTOS ESTADOS DE LA REPÚBLICA MEXICANA HA VISITADO COMO TURISTA? (Actividad realizada en los últimos seis meses)
98	LI_INTER	HABILIDAD PARA LEER INFORMACIÓN EN INTERNET EN INGLÉS
99	LI_TACA	HABILIDAD PARA LEER TEXTOS ACADÉMICOS (LIBROS, REVISTAS) EN INGLÉS
100	LI_NOTI	HABILIDAD PARA LEER NOTICIAS EN INGLÉS
101	HAB_PTEX	HABILIDAD EN LA COMPUTADORA PARA CREAR Y EDITAR UN DOCUMENTO UTILIZANDO UN PROCESADOR DE TEXTO
102	HAB_PRES	HABILIDAD PARA UTILIZAR PROGRAMAS PARA HACER PRESENTACIONES
103	HAB_FBAS	HABILIDAD PARA EMPLEAR FUNCIONES BÁSICAS EN HOJA DE CÁLCULO (CAPTURA DE DATOS, FORMATO, ORDENAMIENTOS Y USO DE FÓRMULAS)
104	HAB_BAJ	HABILIDAD PARA BAJAR PROGRAMAS DE INTERNET
105	POS_SEL	POSICIÓN ALCANZADA POR EL SUSTENTANTE EN EL EXAMEN DE SELECCIÓN
106	ICNE	CALIFICACIÓN EN ÍNDICE CENEVAL DEL EXAMEN DE SELECCIÓN
107	PERCEN	PERCENTIL DEL EXAMEN DE SELECCIÓN
108	PORCECNE	% > CNE DEL EXAMEN DE SELECCIÓN
109	PCNE	CALIFICACIÓN EN PORCENTAJE DE ACIERTOS DEL EXAMEN DE SELECCIÓN
110	PRLM	CALIFICACIÓN DE RAZONAMIENTO LÓGICO MATEMÁTICO EN PORCENTAJE DE ACIERTOS
111	PMAT	CALIFICACIÓN DE MATEMÁTICAS (SELECCIÓN) EN PORCENTAJE DE ACIERTOS
112	PRV	CALIFICACIÓN DE RAZONAMIENTO VERBAL EN PORCENTAJE DE ACIERTOS
113	PESP	CALIFICACIÓN DE ESPAÑOL EN PORCENTAJE DE ACIERTOS
114	PTIC	CALIFICACIÓN DE TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN EN PORCENTAJE DE ACIERTOS
115	IRLM	CALIFICACIÓN DE RAZONAMIENTO LÓGICO MATEMÁTICO EN ÍNDICE CENEVAL
116	IMAT	CALIFICACIÓN DE MATEMÁTICAS (SELECCIÓN) EN ÍNDICE CENEVAL
117	IRV	CALIFICACIÓN DE RAZONAMIENTO VERBAL EN ÍNDICE CENEVAL
118	IESP	CALIFICACIÓN DE ESPAÑOL EN ÍNDICE CENEVAL
119	ITIC	CALIFICACIÓN DE TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN EN ÍNDICE CENEVAL
120	PDD_MG_CAL	CALIFICACIÓN DE CÁLCULO EN PORCENTAJE DE ACIERTOS
121	PDD_MG_FIS	CALIFICACIÓN DE FÍSICA EN PORCENTAJE DE ACIERTOS
122	PDD_MG_MAT	CALIFICACIÓN DE MATEMÁTICAS "DIAGNÓSTICO" EN PORCENTAJE DE ACIERTOS
123	PDD_MG_QUI	CALIFICACIÓN DE QUÍMICA EN PORCENTAJE DE ACIERTOS
124	PDD_MG_ING	CALIFICACIÓN DE INGLÉS EN PORCENTAJE DE ACIERTOS
125	IDD_MG_CAL	CALIFICACIÓN DE CÁLCULO EN ÍNDICE DIAGNÓSTICO
126	IDD_MG_FIS	CALIFICACIÓN DE FÍSICA EN ÍNDICE DIAGNÓSTICO
127	IDD_MG_MAT	CALIFICACIÓN DE MATEMÁTICAS "DIAGNÓSTICO" EN ÍNDICE DIAGNÓSTICO
128	IDD_MG_QUI	CALIFICACIÓN DE QUÍMICA EN ÍNDICE DIAGNÓSTICO
129	IDD_MG_ING	CALIFICACIÓN DE INGLÉS EN ÍNDICE DIAGNÓSTICO

No. C.	Atributo	Descripción
130	DDD_MG_CAL	DICTAMEN DE CÁLCULO
131	DDD_MG_FIS	DICTAMEN DE FÍSICA
132	DDD_MG_MAT	DICTAMEN DE MATEMÁTICAS "DIAGNÓSTICO"
133	DDD_MG_QUI	DICTAMEN DE QUÍMICA
134	DDD_MG_ING	DICTAMEN DE INGLÉS

A su vez cada variable pertenece a un Identificador, Sección o Calificación, según sea el caso; estos son apartados que se enumeran de la I a la VI, omitiendo la V, por lo que se pueden contabilizar cinco, los cuales se describen en la Tabla 3.2.

**Tabla 3.2 Secciones en las que se agrupan las variables de la base de datos del EXANI – II.**

Identificadores (Apartado del I al III)	Cuestionario de contexto (Apartado IV)	Calificación del Examen de Selección (Apartado VI)
<ul style="list-style-type: none"> <li>▶ Identificación de la aplicación. (1 – 8)</li> <li>▶ Identificación de la Institución SEDE. (9 – 12)</li> <li>▶ Identificación del sustentante. (13 y 14)</li> </ul>	<ul style="list-style-type: none"> <li>▶ <b>Sección 1.</b> Datos Generales. (15 – 28)</li> <li>▶ <b>Sección 2.</b> Datos Escolares. (29 – 39)</li> <li>▶ <b>Sección 3.</b> Características de la Escuela de Procedencia. (40 – 52)</li> <li>▶ <b>Sección 4.</b> Comprensión Lectora. (53 y 57)</li> <li>▶ <b>Sección 5.</b> Expresión Escrita. (58 y 64)</li> <li>▶ <b>Sección 6.</b> Expresión Oral. (65 – 67)</li> <li>▶ <b>Sección 7.</b> Situación Laboral. (68)</li> <li>▶ <b>Sección 8.</b> Características Personales. (69)</li> <li>▶ <b>Sección 9.</b> Exámenes de Logro. (70 – 72)</li> <li>▶ <b>Sección 10.</b> Capital Cultural y Económico. (73 – 97)</li> <li>▶ <b>Sección 11.</b> Inglés. (98 – 100)</li> <li>▶ <b>Sección 12.</b> Uso de la Computadora. (101 – 104)</li> </ul>	<ul style="list-style-type: none"> <li>▶ Módulo del Examen Diagnóstico. (105)</li> <li>▶ Calificación del Examen de Selección. (106 – 120)</li> <li>▶ Calificación del Examen Diagnóstico (Módulo Ingeniería y Tecnología). (121 – 134)</li> </ul>



En la Hoja 2 del archivo del EXANI – II, también se presenta el Diccionario de Datos, el cual contiene los siguientes campos:

- ▶ Número consecutivo de variable.
- ▶ Número del segundo nivel de la variable.
- ▶ Variable.
- ▶ Tipo de dato (Carácter, Fecha y Numérico).
- ▶ Longitud con base al tipo de dato.
- ▶ Descripción de la variable.
- ▶ Valores de las variables (Texto, Número o Rango).
- ▶ Etiquetas de las variables.
- ▶ Fuente (1. Cuestionario de contexto, 2. Hoja de Respuestas, 3. Área de procesos ópticos, 4. Área de Calificación, 5. Dir. Gral. Adjunta, 6. Técnica y de Investigación y 7. Sistemas de Internet).

Los valores registrados en la base de datos, en su mayoría son los valores numéricos categorizados, de acuerdo a las opciones que cada sustentante eligió, en cada reactivo.

De igual forma, se tomaron los promedios finales registrados en las Actas de Evaluación de seis materias que cursaron los alumnos en el primer semestre de la carrera de Ingeniería en Tecnologías de la Información y Comunicaciones (ITIC), como se aprecia en la Tabla 3.3, en el semestre Agosto – Diciembre 2011.

**Tabla 3.3 Clave y nombre de materias del primer semestre de ITIC.**

<b>Clave</b>	<b>Nombre de la Materia</b>
1T1	Cálculo Diferencial e Integral
1T2	Fundamentos de Programación
1T3	Matemáticas Discretas I
1T4	Introducción a las TIC's
1T5	Taller de Ética
1T6	Fundamentos de Investigación

---

### 3.3 Fase 2. Limpieza y transformación.

La recopilación de los datos debe ir acompañada de una limpieza e integración de los mismos, para que estos estén en condiciones para su análisis.

Los beneficios del análisis y de la extracción del conocimiento a partir de los datos dependen, en gran medida, de la calidad de los datos recopilados.

Además, generalmente, debido a las características propias de las técnicas de minería de datos, es necesario realizar una transformación de los datos para obtener una “materia prima” que sea adecuada para el propósito concreto y las técnicas que se quieren emplear. En definitiva, el éxito de un proceso de minería de datos depende, no sólo de tener los datos necesarios (una buena recopilación), sino de que éstos estén íntegros, completos y consistentes (una buena limpieza e integración) (Hernández Orallo, Ramírez Quintana, & Ferri Ramírez, 2004).

En la Figura 3.2, se puede apreciar una parte de la base de datos generada por el EXANI – II, en Junio del 2011, en donde resalta el número total de filas registradas, menos uno, debido a que se debe restar el encabezado de cada columna, que significa el campo o variable.

Posteriormente, se realizó un filtrado automático en Microsoft Excel para sólo mostrar la carrera de ITIC, cuya clave es 16000013, la cual se registra en el campo **IDENTIFICA** de la base de datos, obteniendo un total de 70 registros, de los 1,086 generados en dicho examen, como se muestra en la Figura 3.3.

Base de Datos Final - Microsoft Excel

TIPO_EXA	OPC_APLI	ANO_VER	TIPO_REG	TIPO_RESP	CVE_BPM	APLI	FECHA_APLI	CVE_INST	IDENTIFICA	DESC_IDENT	LPOS_IMG	FOLIO	MATRICULA	API
1066	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000063	000000EX206530430	322909712	LIR	
1067	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000093	000000EX206530430	322909810	EZE	
1068	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000103	000000EX206530430	326919012	PIM	
1069	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000103	000000EX206530430	326909519	ZAI	
1070	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000093	000000EX206530430	322909517	MA	
1071	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000063	000000EX206530430	322909016	RO	
1072	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000053	000000EX206530430	325919814	GO	
1073	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000043	000000EX206530430	326919818	ALL	
1074	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000103	000000EX206530430	325929010	MC	
1075	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000063	000000EX206530430	326919716	MA	
1076	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000073	000000EX206530430	324839511	CA	
1077	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000103	000000EX206530430	324859510	GA	
1078	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430	324859612	GU	
1079	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000083	000000EX206530430	324859011	GO	
1080	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000103	000000EX206530430	325779913	DIA	
1081	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000103	000000EX206530430	324779614	MIA	
1082	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430	325779917	ARI	
1083	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000083	000000EX206530430	323819018	DIA	
1084	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000083	000000EX206530430	323819115	VIL	
1085	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000053	000000EX206530430	325809010	TO	
1086	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000103	000000EX206530430	324829714	RO	
1087	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000023	000000EX206530430	325829114	VL	

Figura 3.2 Registros generados en EXANI – II Junio 2011.

Base de Datos Final - Microsoft Excel

TIPO_EX	OPC_APLI	ANO_VER	TIPO_REG	TIPO_RESP	CVE_BPM	APLI	FECHA_APLI	CVE_INST	IDENTIFICA	DESC_IDENT	LPOS_IMG
779	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX2065304302201106326096810000757000000000
802	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX2065304302201106323056216000771000000000
808	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX2065304302201106326336915000704000000000
817	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX2065304302201106326296218000769000000000
849	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX2065304302201106324886114000749000000000
861	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX2065304302201106324866319000715000000000
901	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX2065304302201106325038718000742000000000
907	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX2065304302201106325458518000738000000000
920	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632327841000076800000000
921	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632630811600072400000000
925	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632530891600072900000000
932	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX20653043022011063276581100070500000000
940	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632454801100071600000000
950	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632556891500071300000000
962	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632290891500072500000000
1004	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632408921000077600000000
1042	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632453901800073600000000
1049	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632456941000072200000000
1057	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632495921700074500000000
1065	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632488921100070600000000
1078	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632485961200073900000000
1082	EX2	08	2011	1	2	NO DISPONIBLE	065304302	10/06/2011	161200601	16000013	000000EX206530430220110632579931700073500000000

Figura 3.3 Filtrado de Registros de la carrera de ITIC.

Se hizo un comparativo entre los registros filtrados y las Actas de Evaluación del primer semestre de ITIC, a través del número de control que adquiere cada alumno al ser inscrito al ITM, para obtener una nueva tabla con los registros de alumnos que presentaron el EXANI – II y que además se mantuvieron activos en el ITM, obteniendo como resultado 31 registros.

En la Figura 3.4 se observa el Acta de Calificaciones de la materia de Cálculo Diferencial, en donde se resaltan algunos de los números de control y nombres completos de los alumnos, y en la Figura 3.5 se aprecian los mismos nombres en la base de datos del EXANI – II, los cuales también se resaltan para su identificación.

El nombre completo de cada alumno se encuentra por separado en los campos de APE\_PAT, APE\_MAT y NOMBRE, precedido del campo NO\_CONTROL, el cual fue insertado en la tabla para fines de búsqueda, posteriormente se omitió.

DEPARTAMENTO: DEPARTAMENTO DE METAL MECANICA  
 MATERIA: CALCULO DIFERENCIAL  
 PROFESOR: GUTIERREZ RODRIGUEZ JESUS ORLANDO  
 PERIODO: AGO-DIC/2011

FOLIO: 06111288-F  
 CLAVE: 1T1  
 GRUPO: A  
 ALUMNOS: 34

LUNES		MARTES		MIÉRCOLES		JUEVES		VIERNES		SABADO		DOMINGO	
HORA	AULA	HORA	AULA	HORA	AULA	HORA	AULA	HORA	AULA	HORA	AULA	HORA	AULA
17:00 - 18:00	K4	17:00 - 18:00	K4	17:00 - 18:00	K4	17:00 - 18:00	K4	17:00 - 18:00	K4				
No.	No. CONTROL	NOMBRE DEL ALUMNO						CARRERA	REP.	ESP.	ORD.	REG.	EXT.
1	11121381	ALVARADO SANCHEZ ROMARIO						ITIC			NA	75	--
2	11121382	ANGUIANO GARCIA DIEGO ORLANDO						ITIC			NA	NA	--
3	11121514	ARRIAGA GONZALEZ MARIA YIZEL						ITIC			80	--	--
4	11121383	ARROYO NAVARRETE GRACIELA DIORANY						ITIC			95	--	--
5	11121384	BEJARANO HERNANDEZ SERGIO ALI						ITIC			NA	75	--
6	11121385	BELTRAN ROBLEDLO BERNARDO						ITIC			NA	77	--
7	11121386	BENITEZ CORREA GUILLERMO						ITIC			NA	--	--
8	11121388	CAMBRON BARRALES ANNA PAULETTE						ITIC			96	--	--
9	11121389	CORTES GASCA JORGE EDGAR						ITIC			NA	77	--
10	11121562	DIAZ CORTES ELIAS						ITIC			NA	75	--
11	11121390	FIERROS MENDEZ NOHELY SARAHÍ						ITIC			76	--	--
12	11121391	GONZALEZ MALDONADO HECTOR DANIEL						ITIC			77	--	--
13	11121392	HEREDIA CANDIDO ELIAS MISAEL						ITIC			NA	80	--
14	11121393	HERNANDEZ FUENTES ANA LAURA						ITIC			95	--	--

Figura 3.4 Acta de Calificaciones de Cálculo Diferencial.

FOLIO	NO_CONTROL	APE_PAT	APE_MAT	NOMBRE	DIA_NAC	MES_NAC	ANO_NAC	SEXO	CON_DCE	CON_IMPE	CON_ESC	CON_VER
322951817	11121381	ALVARADO	SANCHEZ	ROMARIO	12	08	1993	1	2	2	2	2
327104016	11121382	ANGUIANO	GARCIA	DIEGO ORLANDO	18	04	1992	1	2	2	2	2
325308916	11121514	ARRIAGA	GONZALEZ	MARIA YIZEL	19	05	1993	2	2	2	2	2
322908915	11121383	ARROYO	NAVARRETE	GRACIELA DIORANY	19	11	1993	2	2	2	2	2
322961314	11121384	BEJARANO	HERNANDEZ	SERGIO ALI	15	04	1993	1	2	2	2	2
323191017	11121385	BELTRAN	ROBLEDO	BERNARDO	17	10	1993	1	2	2	2	2
325930017	11121386	BENITEZ	CORREA	GUILLERMO	18	09	1993	1	2	2	2	2
324089210	11121388	CAMBRO	BARRALES	ANNA PAULETTE	05	07	1993	2	2	2	2	2
326943215	11121389	CORTES	GASCA	JORGE EDGAR	16	08	1993	1	2	2	2	2
322927519	11121390	FIERROS	MENDEZ	NOHELY SARAHÍ	07	05	1993	2	2	2	2	2
324202311	11121391	GONZALEZ	MALDONADO	HECTOR DANIEL	12	03	1993	1	2	2	2	2
326308116	11121392	HEREDIA	CANDIDO	ELIAS MISAE	16	06	1993	1	2	2	2	2
324091715	11121393	HERNANDEZ	FUENTES	ANA LAURA	04	02	1993	2	2	2	2	2
322995415	11121394	HERNANDEZ	TORRES	JOSE RICARDO	19	06	1992	1	2	2	2	2
325294913	11121397	LOPEZ	PAZ	FRANCISCO JAVIER	05	02	1993	1	2	2	2	2
323056216	11121398	MAGAÑA	GUEVARA	MANUEL	24	02	1992	1	2	2	2	2
323802619	11121399	MALINDO	TAPIA	ARIADNA STEPHANIE	14	01	1993	2	2	2	2	2
325568915	11121400	MEDINA	GARABITO	MILED YAMIL	27	08	1993	1	2	2	2	2
322937011	11121401	MEJIA	CORONA	RICARDO	22	02	1990	1	2	2	2	2
322954117	11121402	MENDOZA	HEREDIA	JOSE ALFREDO	26	03	1993	1	2	2	2	2

Figura 3.5 Registros de alumnos activos que presentaron EXANI – II en ITIC.

Se identificaron y consideraron de primera instancia, algunos campos no significativos para el análisis de los datos, en virtud de que son datos propios de las características del examen como tal, y no propiedades que puedan aportar al Perfil del Aspirante, como se puede apreciar en la Tabla 3.4.

Tabla 3.4. Campos no significativos de primera instancia.

No. C.	Atributo	Descripción
1	TIPO_EXA	TIPO DE EXAMEN
2	OPC_APLI	OPCIÓN DE APLICACIÓN
3	ANO_VER	AÑO DE LA VERSIÓN
4	TIPO_REG	TIPO DE REGISTRO
5	TIPO_RESP	TIPO DE RESPUESTAS
6	CVE_BPM	CLAVE DE IDENTIFICACIÓN DE LA APLICACIÓN
7	APLI	NÚMERO DE APLICACIÓN
8	FECHA_APLI	FECHA DE APLICACIÓN
9	CVE_INST	CLAVE DE LA INSTITUCIÓN EN DONDE SE LLEVÓ A CABO LA APLICACIÓN
10	IDENTIFICA	IDENTIFICACIÓN DE LA APLICACIÓN
11	DESC_IDENT	DESCRIPCIÓN DE LA CLAVE IDENTIFICA
12	LPOS_IMG	LOCALIZACIÓN DE LAS IMÁGENES PARA HOJAS DE RESPUESTA Y DE REGISTRO EN LA BASE DE DATOS

Haciendo una depuración más exhaustiva de ellos y considerando el interés de crear el Perfil del Aspirante, como se aprecia en la Tabla 3.5, se omitieron otro conjunto de campos que no son significativos.

**Tabla 3.5 Campos no significativos en segunda instancia.**

No. C.	No. A.	Atributo	Descripción
1	13	FOLIO	FOLIO
2	15	APE_PAT	APELLIDO PATERNO
3	16	APE_MAT	APELLIDO MATERNO
4	17	NOMBRE	NOMBRE
5	18	DIA_NAC	DÍA DE NACIMIENTO
6	19	MES_NAC	MES DE NACIMIENTO
7	22	CON_DCE	IMPEDIMENTO O DIFICULTAD PARA CAMINAR O SUBIR ESCALERAS
8	23	CON_IMPE	IMPEDIMENTO O DIFICULTAD PARA ESCRIBIR
9	24	CON_ESC	PROBLEMAS PARA ESCUCHAR, AÚN A CORTA DISTANCIA
10	25	CON_VER	PROBLEMAS PARA VER (QUE NO SE CORRIJAN CON LENTES)
11	26	CON_LEN	PROBLEMAS DE LENGUAJE
12	27	LI_MAD	¿SU MADRE HABLA UNA LENGUA INDÍGENA?
13	28	LI_PAD	¿SU PADRE HABLA UNA LENGUA INDÍGENA?
14	30	NOM_PROC	NOMBRE DE LA INSTITUCIÓN EN LA QUE CONCLUYÓ EL BACHILLERATO
15	31	CIU_PROC	NOMBRE DE LA CIUDAD DONDE SE UBICA LA INSTITUCIÓN EN LA QUE CONCLUYÓ EL BACHILLERATO; SI FUE EN EL EXTRANJERO INDIQUE EL PAÍS.
16	32	CVE_PROC	CLAVE DE LA INSTITUCIÓN EN LA QUE CONCLUYÓ SUS ESTUDIOS DE BACHILLERATO
17	34	ANO_BAC	¿EN QUÉ AÑO CONCLUYÓ EL BACHILLERATO?
18	40	POR_CCE	CRITERIOS DE EVALUACIÓN (Aspecto que dio a conocer los profesores al inicio de los cursos, en el último año de bachillerato)
19	41	POR_OBAP	OBJETIVOS DE APRENDIZAJE (Aspecto que dio a conocer los profesores al inicio de los cursos, en el último año de bachillerato)
20	42	POR_BIB	BIBLIOGRAFÍA (Aspecto que dio a conocer los profesores al inicio de los cursos, en el último año de bachillerato)
21	43	PPR_ACM	¿QUÉ PORCENTAJE DE LOS PROFESORES TENÍA AMPLIOS CONOCIMIENTOS DE LA MATERIA?
22	44	PPR_EJE	¿QUÉ PORCENTAJE DE LOS PROFESORES DABA CON FRECUENCIA EJEMPLOS QUE PERMITÍAN ENTENDER LA APLICACIÓN DE LOS CONCEPTOS REVISADOS?
23	45	PPR_TAR	¿QUÉ PORCENTAJE DE LOS PROFESORES SOLICITABA TAREAS O TRABAJOS QUE CONSOLIDABAN LOS TEMAS REVISADOS?
24	46	PPR_PUN	¿QUÉ PORCENTAJE DE LOS PROFESORES ASISTÍA REGULARMENTE A CLASE?
25	47	PPR_DUD	¿QUÉ PORCENTAJE DE LOS PROFESORES PERMITÍA QUE LOS ESTUDIANTES PRESENTARAN SUS DUDAS?
26	48	PPR_ASI	¿QUÉ PORCENTAJE DE LOS PROFESORES ASISTÍA PUNTUALMENTE A CLASE?
27	49	FEV_ENO	LOS ESTUDIANTES NO PRESTABAN ATENCIÓN A LO QUE LOS MAESTROS DECÍAN (¿En qué porcentaje ocurría?)
28	50	FEV_RUDE	HABÍA RUIDO Y DESORDEN (¿En qué porcentaje ocurría?)

No. C.	No. A.	Atributo	Descripción
29	51	FEV_TIES	LOS PROFESORES TENÍAN QUE ESPERAR MUCHO TIEMPO PARA QUE LOS ESTUDIANTES PUSIERAN ATENCIÓN (¿En qué porcentaje ocurría?)
30	52	FEV_TEFE	EL TIEMPO EFECTIVO DE TRABAJO EN CLASE ERA MUY POCO (¿En qué porcentaje ocurría?)
31	53	NIV_IDAT	IDENTIFICAR DATOS Y HECHOS (Nivel de habilidad para realizar la actividad al leer)
32	54	NIV_EREL	ESTABLECER RELACIONES ENTRE DATOS, HECHOS Y JUICIOS (Nivel de habilidad para realizar la actividad al leer)
33	55	NIV_RIDE	RECONOCER IDEAS CENTRALES (PRINCIPALES Y SECUNDARIAS) (Nivel de habilidad para realizar la actividad al leer)
34	56	NIV_RPUN	RECONOCER LOS PUNTOS DE VISTA IMPLÍCITOS Y EXPLÍCITOS DEL AUTOR (Nivel de habilidad para realizar la actividad al leer)
35	58	NIV_COH	DE FORMA COHERENTE (Nivel de habilidad para redactar un texto)
36	59	NIV_ERR	SIN ERRORES ORTOGRÁFICOS (Nivel de habilidad para redactar un texto)
37	60	NIV_RESL	EL RESUMEN DE UN LIBRO (Nivel de habilidad para escribir)
38	61	NIV_ENSA	UN ENSAYO (Nivel de habilidad para escribir)
39	62	NIV_CART	UNA CARTA (Nivel de habilidad para escribir)
40	63	NIV_REPO	UN REPORTE DE INVESTIGACIÓN (Nivel de habilidad para escribir)
41	64	NIV_DOCO	UN DOCUMENTO OFICIAL (POR EJEMPLO UNA QUEJA) (Nivel de habilidad para escribir)
42	65	NIV_EXP	EXPONER UN TEMA ACADÉMICO FRENTE A UN GRUPO (Nivel de habilidad para ...)
43	66	NIV_DUDA	EXPONER UNA DUDA A SU MAESTRO (Nivel de habilidad para ...)
44	67	NIV_DEBA	DEFENDER SU PUNTO DE VISTA EN UN DEBATE (Nivel de habilidad para ...)
45	70	FAM_EXA	¿QUÉ TAN FAMILIARIZADO ESTÁ CON LOS EXÁMENES DE OPCIÓN MÚLTIPLE?
46	71	PRE_EXA1	¿USTED PRESENTÓ EL EXAMEN NACIONAL DE INGRESO A LA EDUCACIÓN MEDIA SUPERIOR (EXANI - I) PARA INGRESAR AL BACHILLERATO?
47	72	PRE_EXA2	¿CUÁNTAS VECES HA PRESENTADO EL EXAMEN NACIONAL DE INGRESO A LA EDUCACIÓN SUPERIOR (EXANI - II)?
48	75	TRAB_MAD	¿A QUÉ SE DEDICA SU MADRE O TUTORA ACTUALMENTE?
49	76	TRAB_PAD	¿A QUÉ SE DEDICA SU PADRE O TUTOR ACTUALMENTE?
50	79	CUAN_LIB	¿CUÁNTOS LIBROS HAY EN SU CASA? (No consideres revistas, periódicos o sus libros de texto)
51	80	CUAN_PELI	¿CUÁNTAS PELÍCULAS HAY EN SU CASA?
52	81	EXP_PAD	¿QUÉ NIVEL EDUCATIVO QUISIERAN SUS PADRES QUE ALCANZARA?
53	82	SER_TELE	LÍNEA TELEFÓNICA (Bienes y servicios con los que cuenta en casa)
54	83	SER_LAV	LAVADORA DE ROPA (Bienes y servicios con los que cuenta en casa)
55	84	SER_REF	REFRIGERADOR (Bienes y servicios con los que cuenta en casa)
56	85	SER_HOR	HORNO DE MICROONDAS (Bienes y servicios con los que cuenta en casa)
57	87	SER_CABL	TELEVISIÓN DE PAGA (CABLEVISIÓN, SKY, ETC.) (Bienes y servicios con los que cuenta en casa)
58	88	SER_DVD	¿CUÁNTOS REPRODUCTOR DE DVD HAY EN SU CASA?
59	90	SER_TV	¿CUÁNTOS TELEVISORES HAY EN SU CASA?
60	91	SER_AUTO	¿CUÁNTOS AUTOMÓVILES HAY EN SU CASA?
61	92	SER_BANO	¿CUÁNTOS BAÑOS COMPLETOS HAY EN SU CASA?
62	93	CINE	ASISTIR AL CINE (Actividad realizada en los últimos seis meses)
63	94	MUSEO	ASISTIR A MUSEOS (Actividad realizada en los últimos seis meses)

No. C.	No. A.	Atributo	Descripción
64	95	ESPEC	ASISTIR A ESPECTÁCULOS ARTÍSTICOS (TEATRO, DANZA, CONCIERTO, ETC.) (Actividad realizada en los últimos seis meses)
65	96	VAC_RM	EN LOS ÚLTIMOS 2 AÑOS, ¿CUÁNTAS VECES HA SALIDO DE VACACIONES DENTRO DE LA REPÚBLICA MEXICANA? (Actividad realizada en los últimos seis meses)
66	97	EDO_REP	¿CUÁNTOS ESTADOS DE LA REPÚBLICA MEXICANA HA VISITADO COMO TURISTA? (Actividad realizada en los últimos seis meses)
67	105	POS_SEL	POSICIÓN ALCANZADA POR EL SUSTENTANTE EN EL EXAMEN DE SELECCIÓN
68	106	ICNE	CALIFICACIÓN EN ÍNDICE CENEVAL DEL EXAMEN DE SELECCIÓN
69	107	PERCEN	PERCENTIL DEL EXAMEN DE SELECCIÓN
70	108	PORCECNE	% > CNE DEL EXAMEN DE SELECCIÓN
71	109	PCNE	CALIFICACIÓN EN PORCENTAJE DE ACIERTOS DEL EXAMEN DE SELECCIÓN
72	115	IRLM	CALIFICACIÓN DE RAZONAMIENTO LÓGICO MATEMÁTICO EN ÍNDICE CENEVAL
73	116	IMAT	CALIFICACIÓN DE MATEMÁTICAS (SELECCIÓN) EN ÍNDICE CENEVAL
74	117	IRV	CALIFICACIÓN DE RAZONAMIENTO VERBAL EN ÍNDICE CENEVAL
75	118	IESP	CALIFICACIÓN DE ESPAÑOL EN ÍNDICE CENEVAL
76	119	ITIC	CALIFICACIÓN DE TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN EN ÍNDICE CENEVAL
77	125	IDD_MG_CAL	CALIFICACIÓN DE CÁLCULO EN ÍNDICE DIAGNÓSTICO
78	126	IDD_MG_FIS	CALIFICACIÓN DE FÍSICA EN ÍNDICE DIAGNÓSTICO
79	127	IDD_MG_MAT	CALIFICACIÓN DE MATEMÁTICAS "DIAGNÓSTICO" EN ÍNDICE DIAGNÓSTICO
80	128	IDD_MG_QUI	CALIFICACIÓN DE QUÍMICA EN ÍNDICE DIAGNÓSTICO
81	129	IDD_MG_ING	CALIFICACIÓN DE INGLÉS EN ÍNDICE DIAGNÓSTICO

En conclusión, de los 134 campos creados por el EXANI – II, se quitaron 12 variables en primera instancia y 81 variables en segunda instancia, dando un total de 93 campos omitidos, por lo que sólo quedan 41 variables para su análisis, como se muestra en la Figura 3.6 (a), en donde resalta el número columnas.

Después de haber depurado la base de datos, se procedió a la inserción de seis columnas, al final de los campos originales, cada una correspondiente a cada materia cursada en el primer semestre, en donde se registraron cada una de las calificaciones finales (Ordinarias y Extraordinarias) de los 31 alumnos activos.

Se insertó un campo denominado **Clase numérica**, la cual toma el valor de 0 si el alumno tienen más de dos materias con promedio final menor a 70, y el valor de 1 si se cumple lo contrario, esto es determinante para saber si el alumno permaneció o no en el ITM, como se muestra en la Figura 3.6 (b).



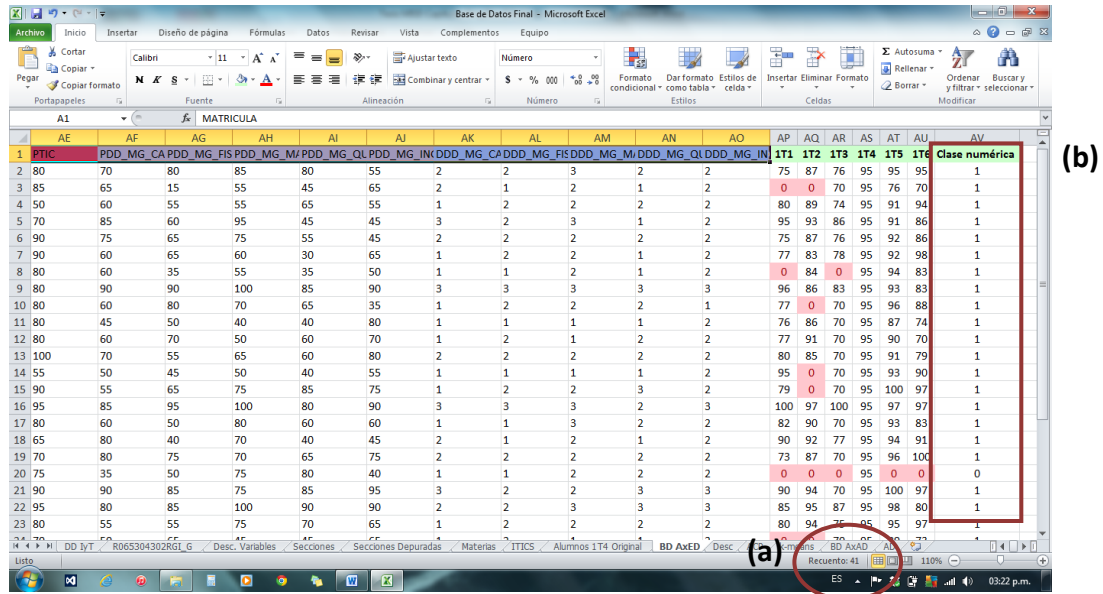


Figura 3.6 (a) Campos depurados para análisis.  
(b) Inserción de nuevos campos.

A partir de este momento se tiene formada una matriz de datos de 31 registros x 47 campos, siendo el primero, el número de registros (alumnos) y el segundo, el número de campos (variables), para lo cual se procede al análisis de los datos a través de **Estadística Descriptiva**.

Para lograrlo se utilizó la herramienta XLSTAT, la cual se sincroniza con Microsoft Excel al ser instalada; después se realizó la siguiente configuración para lograr activar las fuentes de confianza, requisito previo para utilizar XLSTAT:

1. Hacer clic en el botón "Office" en la esquina superior izquierda de la ventana de Excel para abrir el menú Office.
2. Hacer clic en el botón "Opciones de Excel" al final de la ventana del menú de Office.
3. Aparecerá la ventana "Opciones de Excel". Seleccionar "Centro de confianza" en la mitad de la barra de menú de la izquierda.
4. Hacer clic en el botón "Configuración del centro de confianza..." a la derecha de la ventana.

5. Aparecerá la ventana "Centro de confianza". Seleccionar "Configuración de macros" en la mitad de la barra del menú de la izquierda. Seleccionar también la segunda opción en "Configuración de macros", la cual dice "Deshabilitar todas las macros con notificación".
6. Activar la opción "Confiar en el acceso al modelo de objeto de proyectos de VBA".
7. Hacer clic en OK para cerrar la ventana "Centro de confianza", y luego clic en OK para cerrar la ventana "Opciones de Excel".

Este tipo de análisis se denomina también análisis exploratorio de datos, el cual ofrece modos de presentar y evaluar las características principales de los datos a través de tablas y gráficos e identificar qué datos se comportan con una desviación estándar significativa o nula.

Un requisito de preparación fue convertir los datos en texto para que tuvieran la propiedad de ser cualitativos. La base de datos para este análisis se denominó **BD AxED**, etiqueta que se muestra en la Hoja de Excel; después se eligieron las opciones que se observan en la Figura 3.7.

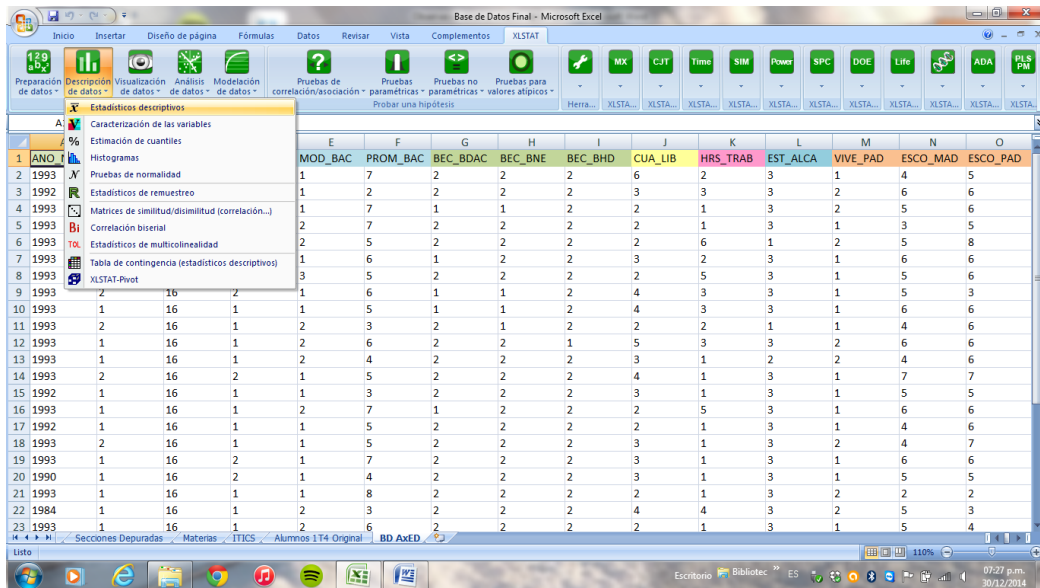


Figura 3.7 Opción: Descripción de datos – Estadísticos descriptivos.

En la Figura 3.8, se muestra su cuadro de diálogo, en donde se determinan los Datos cualitativos, utilizando el rango de la base de datos que incluye los campos depurados del EXANI – II, las seis materias elegidas del primer semestre y la Clase numérica.

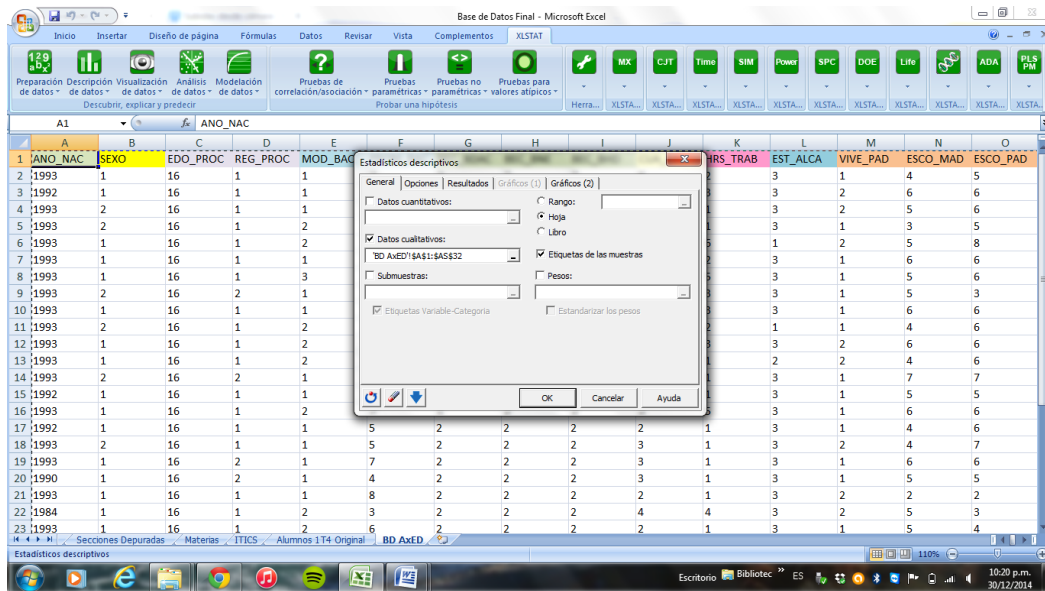


Figura 3.8 Cuadro de Diálogo de Estadísticos descriptivos.

En la Pestaña de **Gráfico (2)**, se debe seleccionar todos los tipos de gráficos propuestos.

En la Figura 3.9, se muestra el número de filas y columnas para su análisis, en cuanto se elige el botón Continuar, se generan las estadísticas descriptivas, cuyos resultados parciales se muestran en la Figura 3.10.

Se puede observar sólo una parte de la tabla generada, en el archivo original se encuentra la clasificación de todos los campos, así como sus gráficas correspondientes, como se puede apreciar en la Figura 3.11.

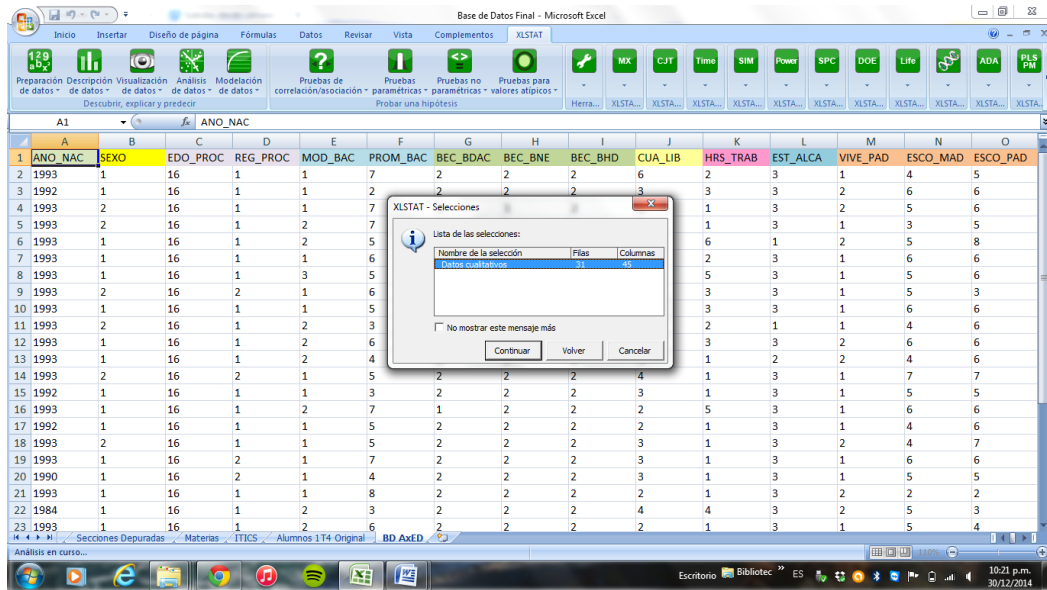


Figura 3.9 Cuadro de Diálogo de Confirmación de Análisis de Datos Cualitativos.

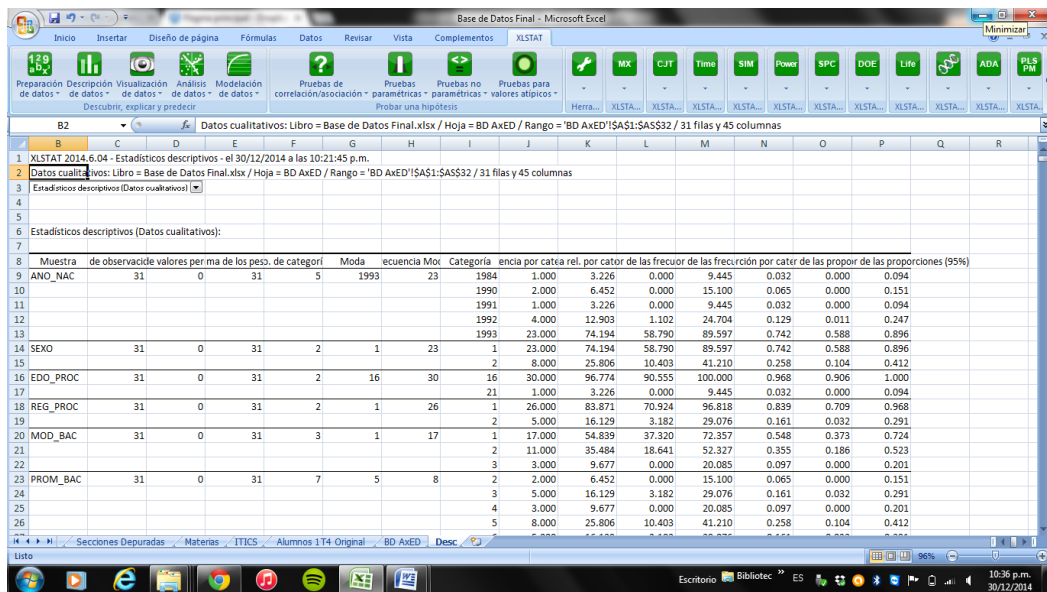


Figura 3.10 Clasificación de los datos por estadística descriptiva.

Al observar cada uno de los datos generados, se detectó que existen tres campos cuyos valores son constantes por lo que no aporta significado para el objetivo.

En la Figura 3.12, se muestra la tabla y el gráfico en barras, de la variable **MATRICULA**, la cual tiene una frecuencia de **1.000**, lo que significa que el valor es único para cada registro, por lo que no puede haber rangos que se puedan categorizar, esto la convierte en una variable no significativa.

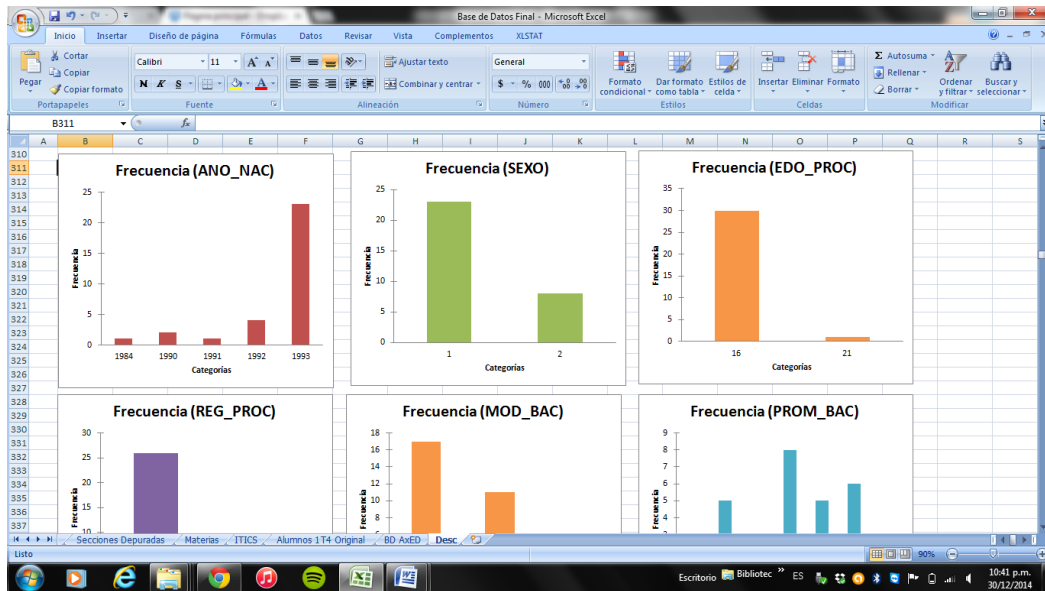


Figura 3.11 Gráficas de frecuencia de cada dato analizado.

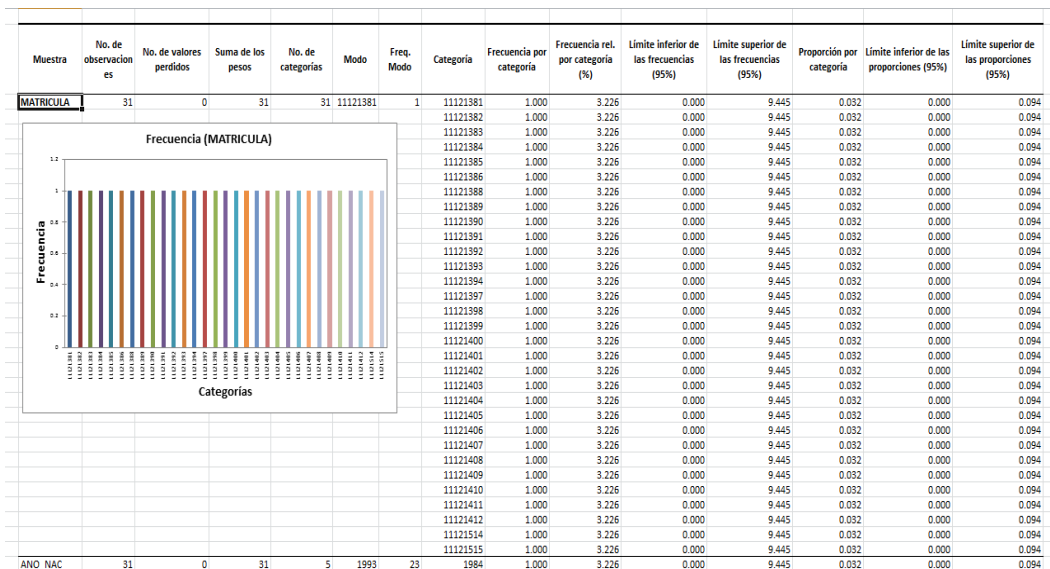


Figura 3.12 Análisis y gráfica de la variable **MATRICULA**.

Posteriormente aparece la variable denominada **VIVE\_MAD**, cuyo reactivo es *¿VIVE CON LAS SIGUIENTES PERSONAS? (Madre o Padre)* y la opción 1 es Madre y la opción 2 es Padre la cual también se omite dado que a diferencia de la anterior, el valor de la frecuencia es **31**, y la Categoría es **1 (Si)** lo que significa que todos los encuestados viven con su madre.

Lo anterior también conduce a omitir dicha variable, ya que resulta no significativa para crear el Perfil del Aspirante, como se muestra en la Figura 3.13.

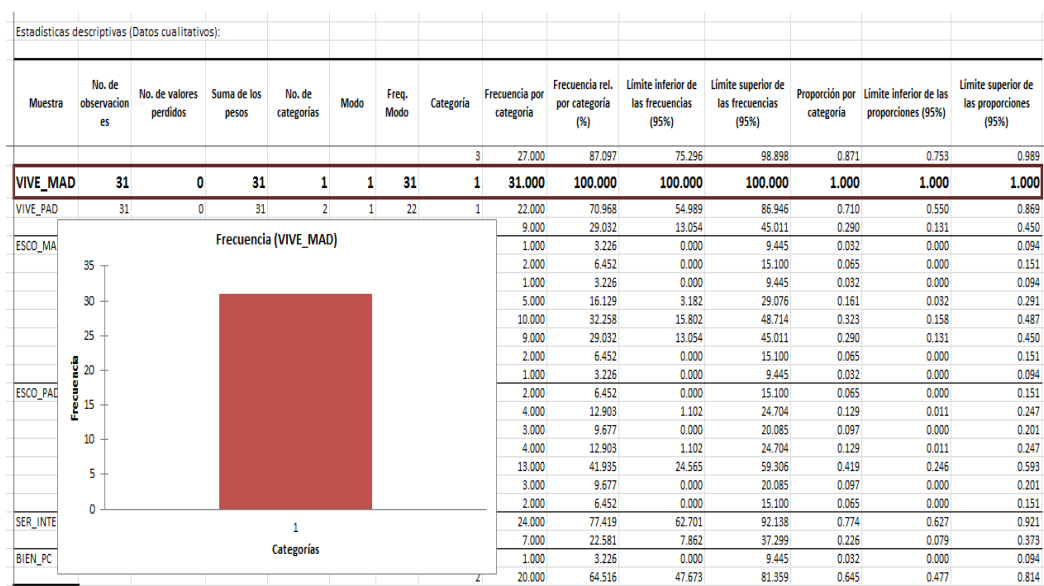


Figura 3.13 Análisis y gráfica de la variable **VIVE\_MAD**.

Dentro de este preprocesamiento de los datos también se encontró la variable denominada **1T4**, la cual contiene las evaluaciones finales de los aspirantes que cursaron la materia de “Introducción a las TIC’s”, en el primer semestre y curiosamente todos los alumnos tienen la misma evaluación de 95, por lo que al igual que la variable anterior, no resulta significativa, esto se puede apreciar en la Figura 3.14.

Al aplicar estadística descriptiva en el almacén de datos, se detectan las variables no significativas y reducir la dimensionalidad, aún cuando no se aplique algoritmos de minería de datos, por lo que para los siguientes análisis se omitirán dichas variables, quedando así una matriz de 44 campos x 31 registros.

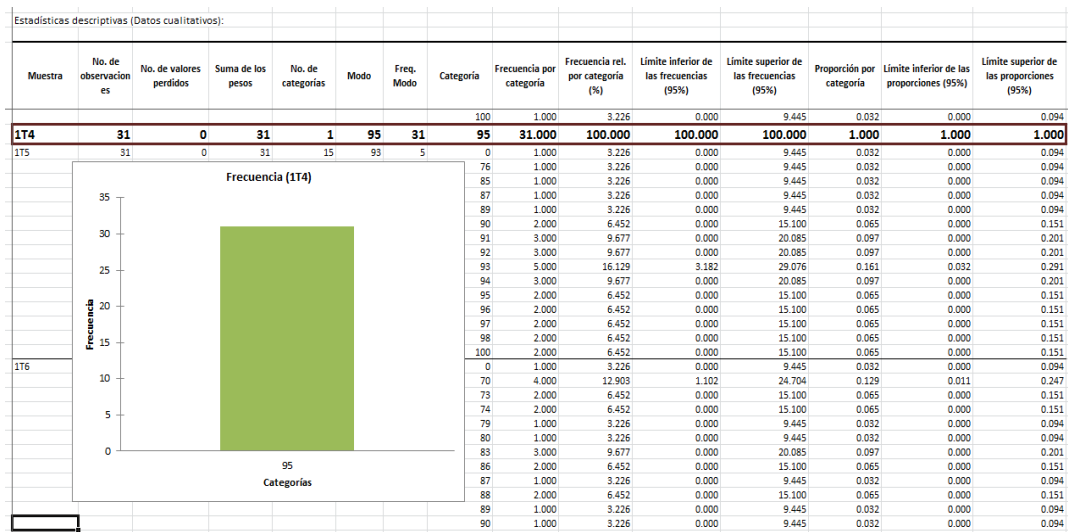


Figura 3.14 Análisis y gráfica de la variable 1T4 (Introducción a las TIC's).

### 3.4 Fase 3. Minería de Datos.

Esta fase es la más característica del KDD y, por esta razón, muchas veces se utiliza esta fase para nombrar todo el proceso. El objetivo de esta fase es producir nuevo conocimiento que pueda utilizar el usuario. Esto se realiza construyendo un modelo basado en los datos recopilados para este efecto. El modelo es una descripción de los patrones y relaciones entre los datos que pueden usarse para hacer predicciones, para entender mejor los datos o para explicar situaciones pasadas. Para ello es necesario tomar una serie de decisiones antes de empezar el proceso, como determinar el tipo de tarea de minería de datos más apropiado, elegir el tipo de modelo y elegir el algoritmo de minería (Hernández Orallo, Ramírez Quintana, & Ferri Ramírez, 2004).

---

El segundo análisis fue a través de **Componentes Principales**, identificada en inglés como PCA (Principal Component Analysis); técnica estadística inferencial de síntesis de la información, o reducción de la dimensión de un conjunto de datos (número de variables).

Se utiliza para hallar las causas de la variabilidad de un conjunto de datos y ordenarlas por importancia. Técnicamente, el PCA busca la proyección según la cual los datos queden mejor representados en términos de mínimos cuadrados. Se emplea sobre todo en análisis exploratorio de datos para construir modelos predictivos. El PCA realiza el cálculo de la descomposición en autovalores de la matriz de covarianza, normalmente tras centrar los datos en la media de cada atributo.

Este análisis construye una transformación lineal que escoge un nuevo sistema de coordenadas para el conjunto original de datos en el cual la varianza de mayor tamaño del conjunto de datos es capturada en el primer eje (llamado el Primer Componente Principal), la segunda varianza más grande es el segundo eje (llamado el Segundo Componente Principal), y así sucesivamente. Para construir esta transformación lineal debe construirse primero la matriz de covarianza o matriz de coeficientes de correlación. Debido a la simetría de esta matriz existe una base completa de vectores propios de la misma. La transformación que lleva de las antiguas coordenadas a las coordenadas de la nueva base, es precisamente la transformación lineal necesaria para reducir la dimensionalidad de datos. Además las coordenadas en la nueva base dan la composición en factores subyacentes de los datos iniciales.

Una de las ventajas del PCA para reducir la dimensionalidad de un grupo de datos, es que retiene aquellas características del conjunto de datos que contribuyen más a su varianza, manteniendo un orden de bajo nivel de los componentes principales e ignorando los de alto nivel.



El objetivo es que esos componentes de bajo orden a veces contienen el aspecto "más importante" de esa información. Nuevamente se utilizó XLSTAT para aplicar dicho análisis, como se muestra en la Figura 3.15.

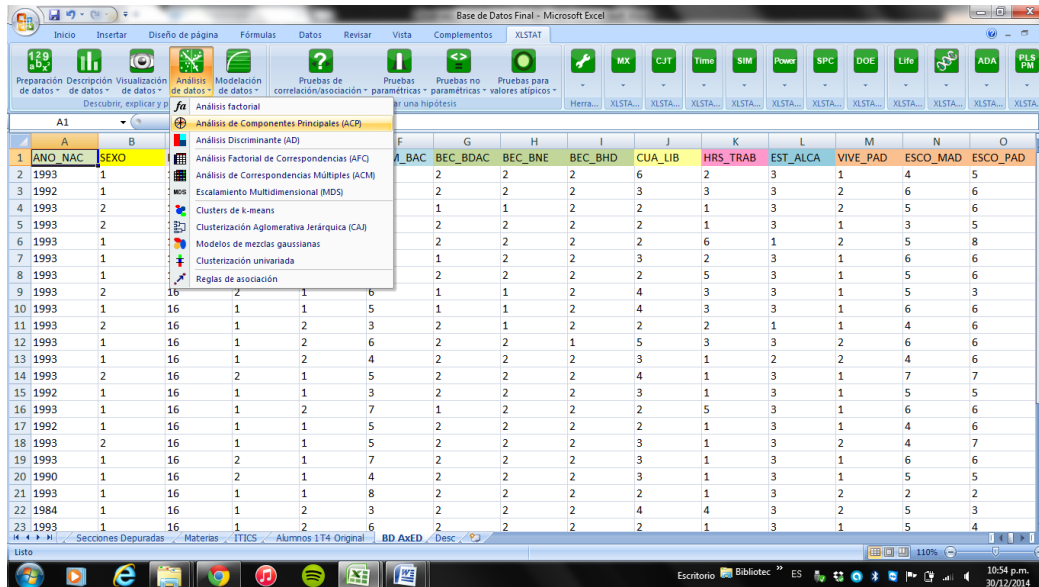


Figura 3.15 Opción: Análisis de Datos – Análisis de Componentes Principales (ACP).

Después de elegir la opción correspondiente, aparece el cuadro de diálogo de la Figura 3.16, en donde se seleccionan nuevamente los datos depurados del EXANI – II y las evaluaciones de las materias, junto con la Clase Numérica. Dicho rango se especifica en el cuadro de texto etiquetado como **Tabla de Observaciones/Variables**, además de activar todas las opciones de la pestaña de **Gráficos**. En la Figura 3.17 sólo se confirma la tabla de observaciones y se elige el botón de Continuar.

Posteriormente y de manera consecutiva se realiza la elección de los ejes, que significa comparar el primer componente principal o vector propio con el segundo (F1 vs F2), después el primero con el tercero (F1 vs F3) y por último el segundo con el tercero (F2 vs F3), como se muestra en la Figura 3.18 (a) que es la comparación de F1 vs F2, Figura 3.18 (b) que es la comparación de F1 vs F3 y Figura 3.28 (c) que es la comparación de F2 vs F3.

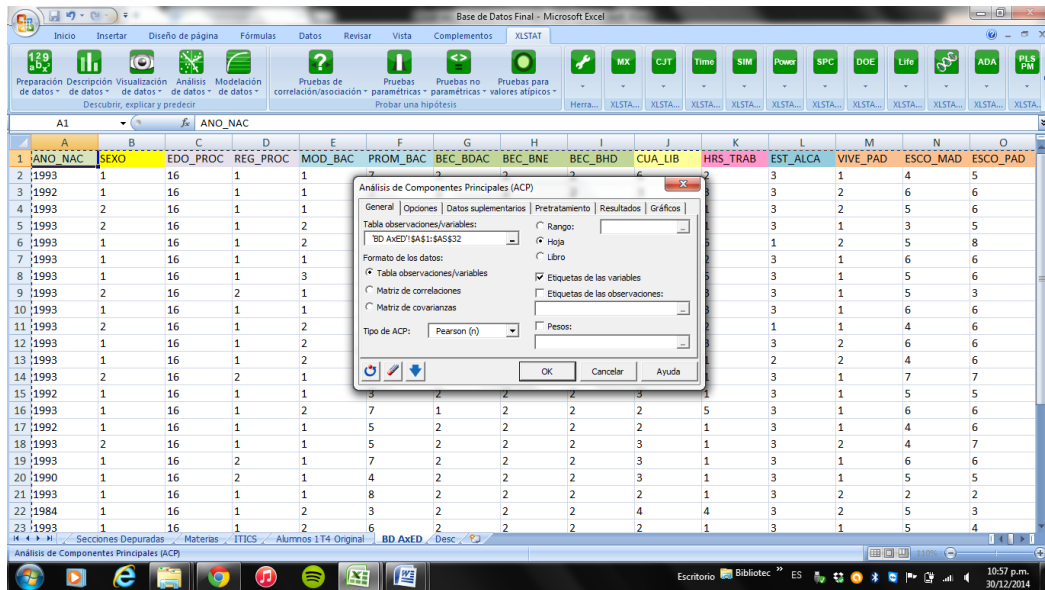


Figura 3.16 Cuadro de diálogo de Análisis de Componentes Principales.

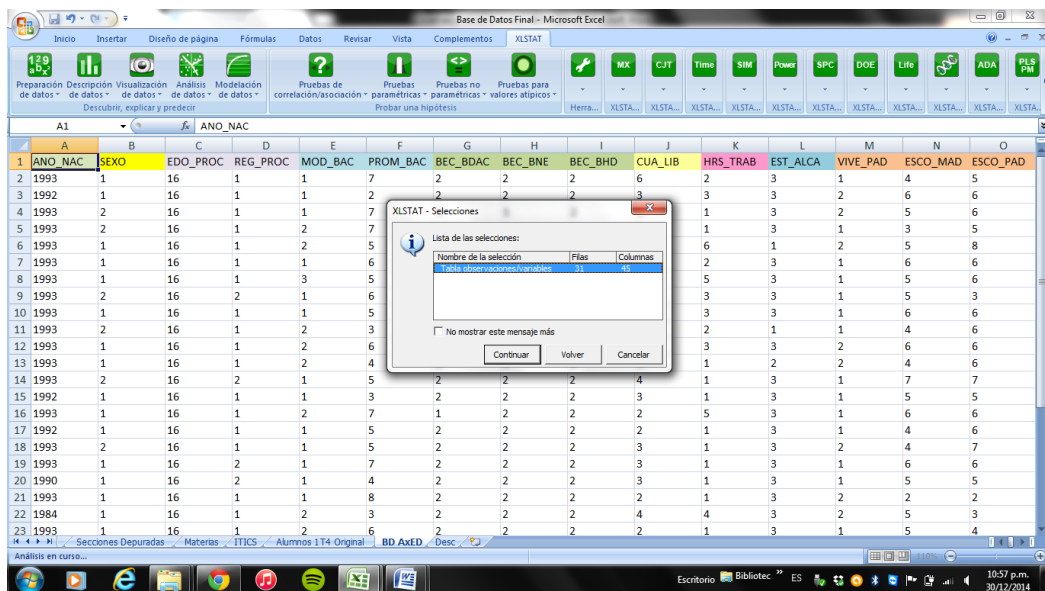


Figura 3.17 Cuadro de diálogo de Confirmación de Tabla de Observaciones.

La elección de los tres primeros ejes es debido a que son las variables más representativas respecto a las demás, por lo que se realiza la combinación para saber qué tan relevantes son, dentro de esas variables debe haber pistas que de alguna manera haga significativa la muestra.

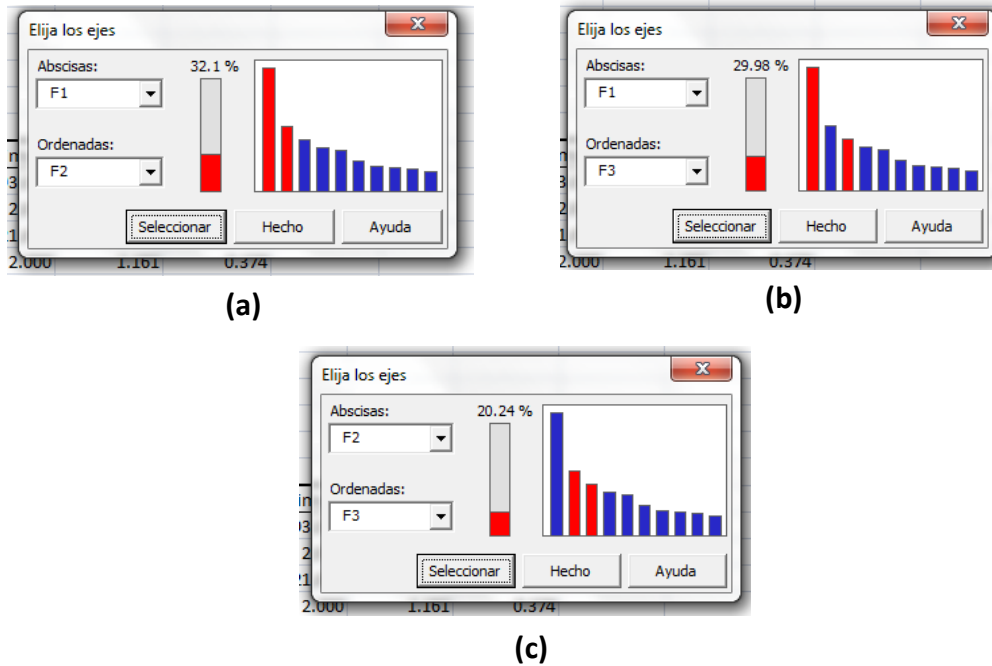


Figura 3.18 Elección de Ejes. (a) F1 vs F2, (b) F1 vs F3 y (c) F2 vs F3

Al elegir en cada caso el botón **Seleccionar**, la herramienta realiza el Análisis de Componentes Principales, como mecanismo para reducir la dimensionalidad, obteniendo la Matriz de Correlaciones Pearson, la cual se aprecia como una matriz conformada por  $n$  filas y por  $n$  columnas. Además es una matriz simétrica; es decir, que los valores de los elementos  $a_{ij}$  de la matriz, son el mismo valor que los elementos  $a_{ji}$  de la misma.

La matriz de correlación explica cómo se encuentran relacionadas cada una de las variables con otra variable. Su diagonal siempre contendrá el valor de 1. Si tiene un valor 0, indica que no tiene ninguna relación con esa variable, por lo menos no lineal; es decir, puede que tenga una relación cuadrática o de otro grado.

Cuando la correlación es positiva, indica que su proyección de la regresión lineal va a tender a crecer conjuntamente con la contra variable.

Cuando la correlación es negativa, indica que su proyección de la regresión lineal va a tender a decrecer conjuntamente con la contra variable.

Al observar la Figura 3.19 en sus incisos (a), (b), (c) y (d), se toma como valores significativos los que son mayores de 0.3, resaltados en celdas rellenas de color verde y mayores de -0.3, resaltados en celdas rellenas de color anaranjado.

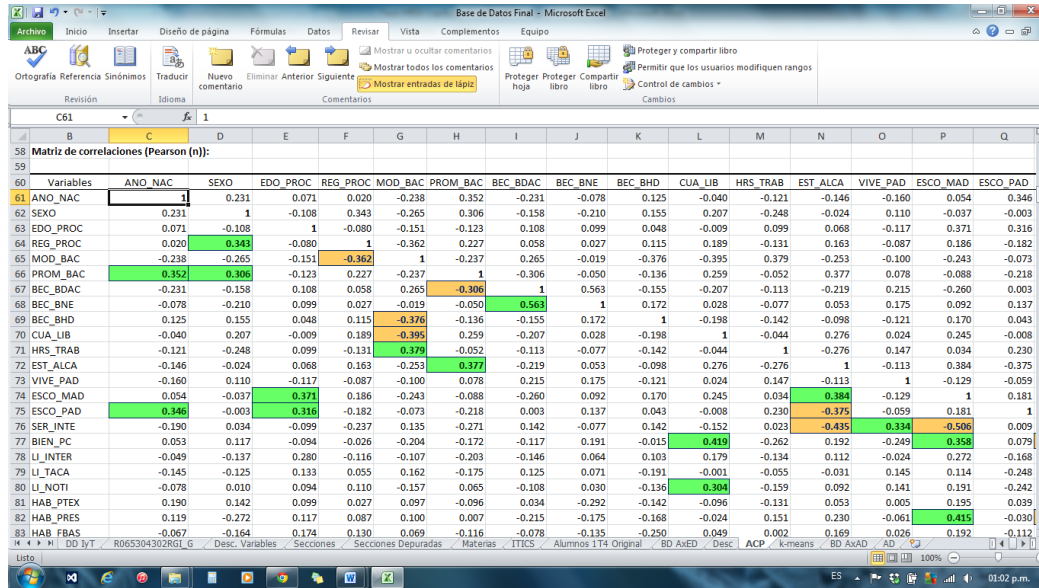


Figura 3.19 (a) Matriz de correlación primer rango.

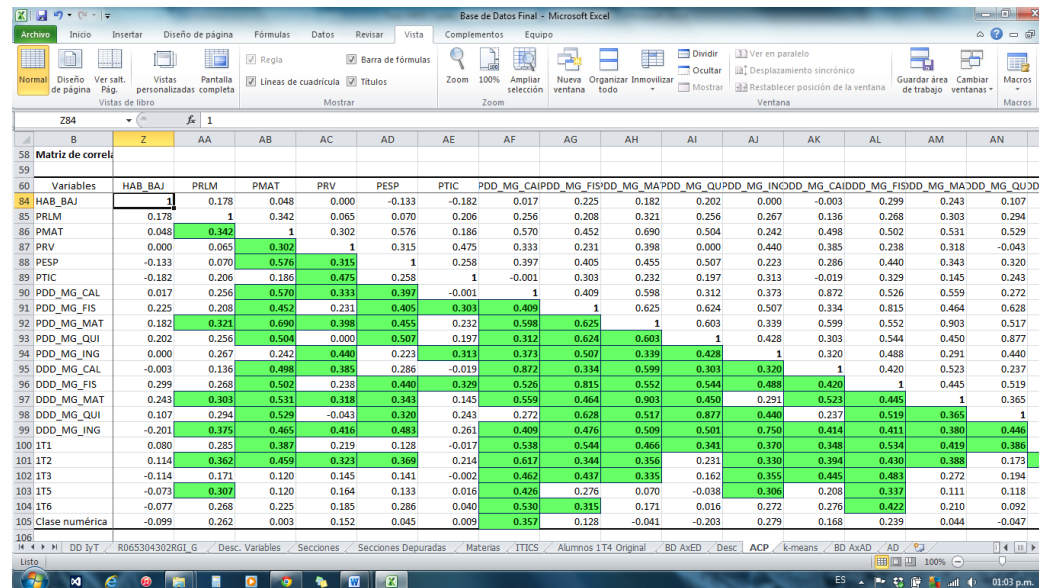


Figura 3.19 (b) Matriz de correlación segundo rango.

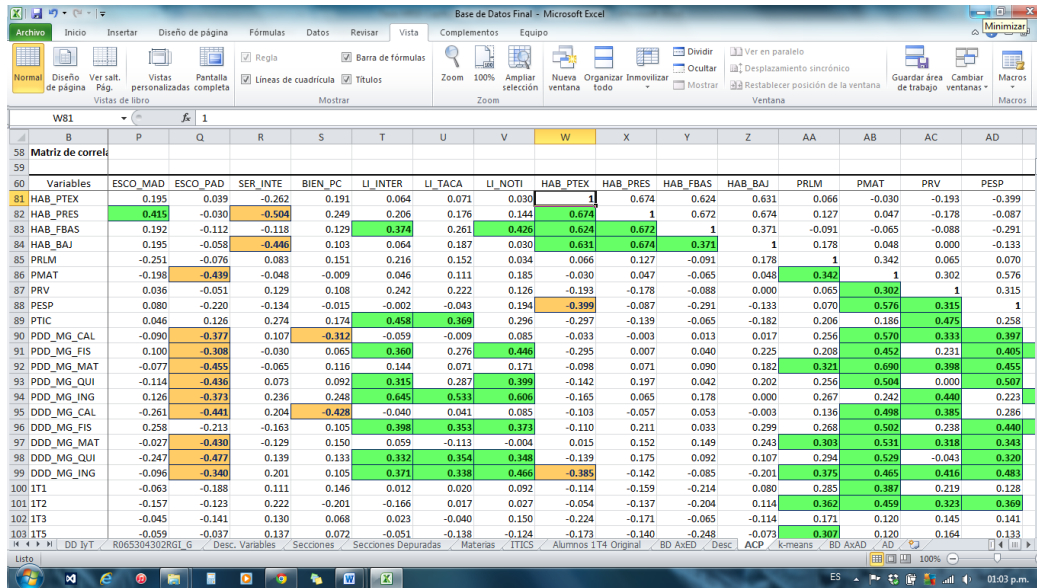


Figura 3.19 (c) Matriz de correlación tercer rango.

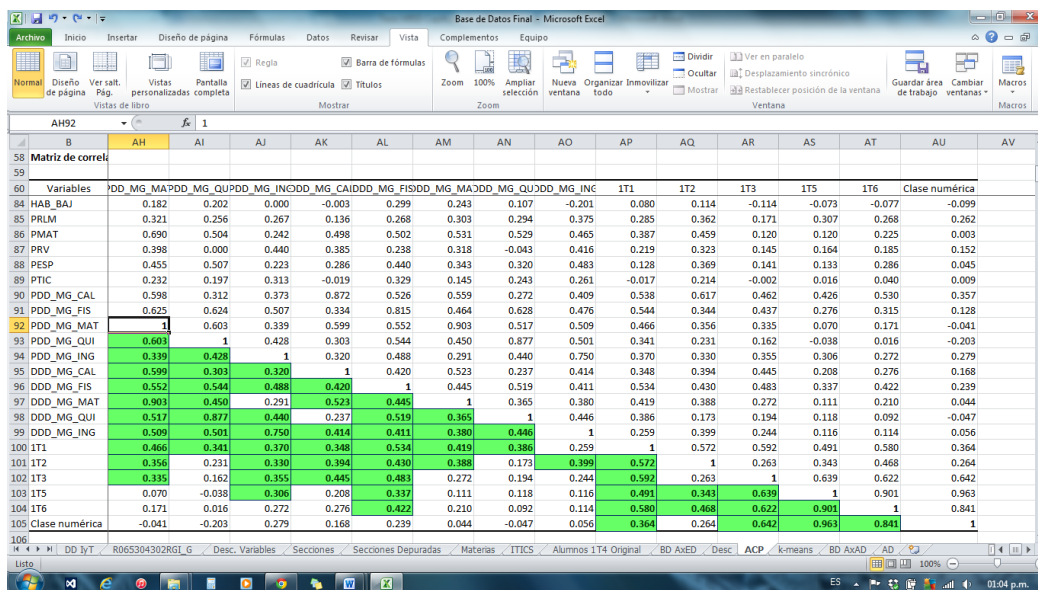


Figura 3.19 (d) Matriz de correlación cuarto rango.

A partir de esta matriz se puede establecer la siguiente relación de variables, categorizándolas en grupos de interés para la formación del Perfil del Aspirante, las cuales se muestran en su totalidad en el Anexo A.2, a partir de las siguientes conjeturas:

---

Se puede decir que la variable Año de Nacimiento (ANO\_NAC) del estudiante está correlacionada positivamente (crecen juntas) con el Promedio de Bachillerato (PROM\_BAC) y el grado máximo de Escolaridad del Padre (ESCO\_PAD). Esta relación puede establecer la hipótesis que a mayor edad del estudiante, mayor será su Promedio de Bachillerato y más alto será el nivel de Estudios del Padre. No obstante también se puede observar que el Año de Nacimiento del estudiante, está correlacionado negativamente con el Dictamen de Química (DDD\_MG\_QUI), lo que hace suponer que entre menos edad, existe la posibilidad de que tenga un mayor promedio en el Dictamen de Química.

La variable Sexo (SEXO) del estudiante está correlacionada positivamente con el Régimen de escuela que estudió el último año de bachillerato (REG\_PROC) y el Promedio de Bachillerato (PROM\_BAC). Esta relación puede establecer la hipótesis de que si el estudiante viene de una escuela privada, es probable que tenga un mejor promedio en relación a la Escuela pública.

La variable Estado de Procedencia (EDO\_PROC) del estudiante está correlacionada positivamente con el nivel máximo de Escolaridad de la Madre (ESCO\_MAD) y con el nivel máximo de Escolaridad del Padre (ESCO\_PAD). Esta relación puede establecer la hipótesis de que depende del Estado de la República Mexicana en el que el estudiante terminó el último año de bachillerato, del nivel máximo de estudios que tenga el padre y la madre, incluso en un análisis más profundo podemos establecer que si la zona económica a la que pertenece dicho Estado es de un poder adquisitivo alto, el nivel máximo de estudios de los padres del estudiante, también es alto (Licenciatura o Posgrado). Sin embargo también se puede observar que el Estado de Procedencia (EDO\_PROC) del estudiante está correlacionado negativamente con el Dictamen de Química (DDD\_MG\_MAT), lo que hace suponer que si el Estado de Procedencia tiene un nivel económico bajo, el promedio en el Dictamen de Matemáticas será alto, o viceversa.

A partir de este análisis se puede adquirir conocimiento relevante para la toma de decisiones que apunten a crear el Perfil del Aspirante, considerando todas las correlaciones positivas (directa) y negativas (inversa) que existen entre las variables principales, sin embargo, este método estadístico, no resuelve en su totalidad lo que se busca, que es obtener una clasificación que se derive en dos clases dentro de la población, la primera que haga referencia a los alumnos que en el primer semestre reprobaron más de dos materias y que de acuerdo al reglamento del ITM, son alumnos que entran en un proceso de baja de la Institución y la segunda que haga referencia a los alumnos que aprobaron todas las materias o que sólo tienen hasta dos materias reprobadas, las cuales pueden regularizar a través de un proceso denominado “nivelación académica”, guiado y gestionado por una labor tutorial, sin embargo, este método estadístico descriptivo es la entrada para generar Componentes Principales que es un modelo inferencial.

Al aplicar este modelo, se debe considerar un margen de confiabilidad del análisis del 95%, con un 5% de error, por lo que se puede apreciar en la Figura 3.20, que se reduce la dimensionalidad, considerando los valores hasta F20, en un sentido optimista, sin embargo, siguiendo el grado de confiabilidad se tiene que las variables significativas llegan hasta F19 = 95.091%.

	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26
Valor propio	0.887	0.864	0.749	0.686	0.551	0.520	0.454	0.402	0.250	0.247	0.216	0.181	0.136
Variabilidad (%)	1.971	1.919	1.664	1.525	1.225	1.156	1.009	0.894	0.556	0.549	0.480	0.403	0.303
% acumulado	87.602	89.521	91.185	92.710	93.934	95.091	96.100	96.994	97.550	98.100	98.580	98.983	99.286

Figura 3.20 Valores Propios de los Componentes Principales.

Al observar la Figura 3.21 en sus incisos (a) y (b), donde se muestran las Cargas Factoriales, se puede apreciar que los datos relevantes son mayores a +0.4 o mayores a -0.4, por lo que no se presentan a partir de la columna F15, por lo que el porcentaje de confiabilidad será del 89.521%, discriminando de alguna manera el resto (de la F16 a la F30).

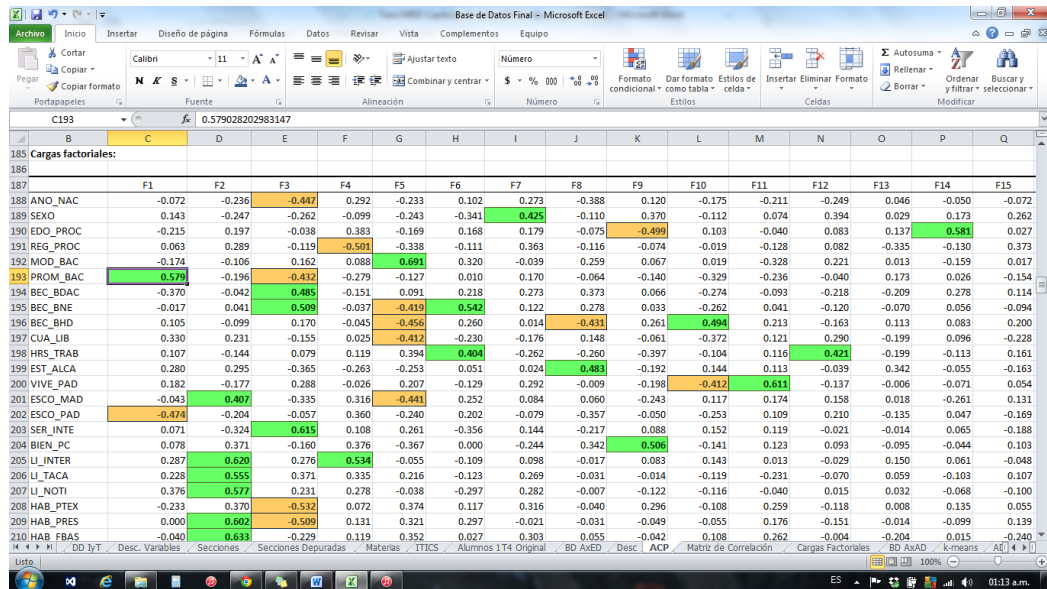


Figura 3.21 (a) Cargas Factoriales primer rango.

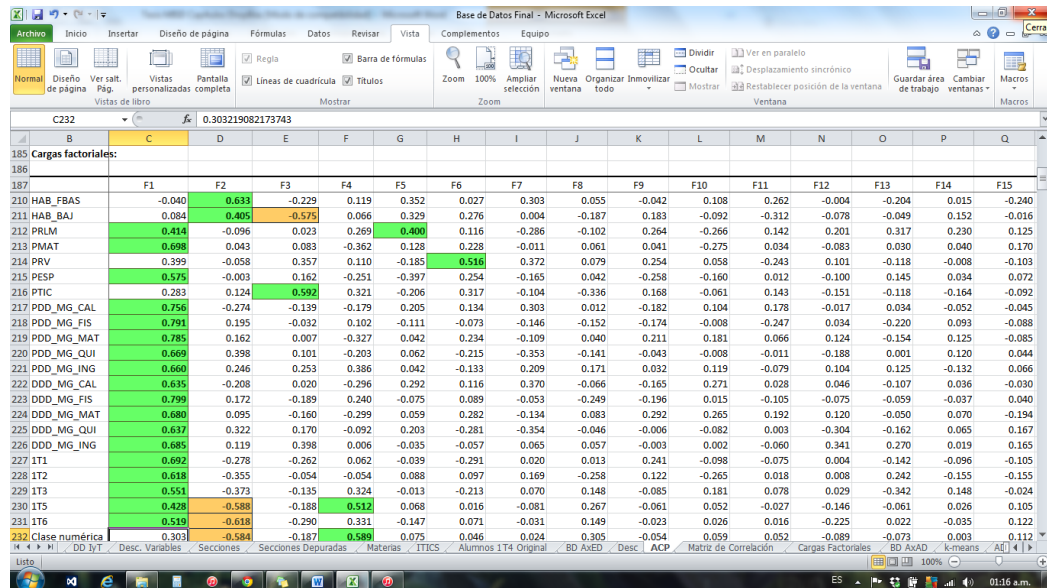


Figura 3.21 (b) Cargas Factoriales segundo rango.



Al tomar los tres componentes principales (F1, F2 y F3), se puede determinar cuáles son las variables representativas de la muestra, las cuales se tomarán en cuenta para establecer el Perfil del Aspirante. Estas variables se muestran en las Tablas 3.6, 3.7 y 3.8, clasificadas por Función o Componente Principal.

**Tabla 3.6 Atributos relevantes del primer componente principal.**

No.	Atributo	F1	Descripción
1	PROM_BAC	0.579	¿CUÁL FUE SU PROMEDIO GENERAL EN EL BACHILLERATO?
2	ESCO_PAD	-0.474	¿CUÁL ES EL MÁXIMO NIVEL DE ESTUDIOS ALCANZADO POR SU PADRE (AUNQUE HAYA FALLECIDO)?
3	PRLM	0.414	CALIFICACIÓN DE RAZONAMIENTO LÓGICO MATEMÁTICO EN PORCENTAJE DE ACIERTOS
4	PMAT	0.698	CALIFICACIÓN DE MATEMÁTICAS (SELECCIÓN) EN PORCENTAJE DE ACIERTOS
5	PESP	0.575	CALIFICACIÓN DE ESPAÑOL EN PORCENTAJE DE ACIERTOS
6	PDD_MG_CAL	0.756	CALIFICACIÓN DE CÁLCULO EN PORCENTAJE DE ACIERTOS
7	PDD_MG_FIS	0.791	CALIFICACIÓN DE FÍSICA EN PORCENTAJE DE ACIERTOS
8	PDD_MG_MAT	0.785	CALIFICACIÓN DE MATEMÁTICAS "DIAGNÓSTICO" EN PORCENTAJE DE ACIERTOS
9	PDD_MG_QUI	0.669	CALIFICACIÓN DE QUÍMICA EN PORCENTAJE DE ACIERTOS
10	PDD_MG_ING	0.660	CALIFICACIÓN DE INGLÉS EN PORCENTAJE DE ACIERTOS
11	DDD_MG_CAL	0.635	DICTAMEN DE CÁLCULO
12	DDD_MG_FIS	0.799	DICTAMEN DE FÍSICA
13	DDD_MG_MAT	0.680	DICTAMEN DE MATEMÁTICAS "DIAGNÓSTICO"
14	DDD_MG_QUI	0.637	DICTAMEN DE QUÍMICA
15	DDD_MG_ING	0.685	DICTAMEN DE INGLÉS

**Tabla 3.7 Atributos relevantes del segundo componente principal.**

No.	Atributo	F2	Descripción
1	ESCO_MAD	0.407	¿CUÁL ES EL MÁXIMO NIVEL DE ESTUDIOS ALCANZADO POR SU MADRE (AUNQUE HAYA FALLECIDO)?
2	LI_INTER	0.620	HABILIDAD PARA LEER INFORMACIÓN EN INTERNET EN INGLÉS
3	LI_TACA	0.555	HABILIDAD PARA LEER TEXTOS ACADÉMICOS (LIBROS, REVISTAS) EN INGLÉS
4	LI_NOTI	0.577	HABILIDAD PARA LEER NOTICIAS EN INGLÉS
5	HAB_PRES	0.602	HABILIDAD PARA UTILIZAR PROGRAMAS PARA HACER PRESENTACIONES
6	HAB_FBAS	0.633	HABILIDAD PARA EMPLEAR FUNCIONES BÁSICAS EN HOJA DE CÁLCULO (CAPTURA DE DATOS, FORMATO, ORDENAMIENTOS Y USO DE FÓRMULAS)
7	HAB_BAJ	0.405	HABILIDAD PARA BAJAR PROGRAMAS DE INTERNET

**Tabla 3.8 Atributos relevantes del tercer componente principal.**

No.	Atributo	F3	Descripción
1	ANO_NAC	-0.447	AÑO DE NACIMIENTO
2	PROM_BAC	-0.432	¿CUÁL FUE SU PROMEDIO GENERAL EN EL BACHILLERATO?
3	BEC_BDAC	0.485	¿RECIBIÓ BECA EN EL BACHILLERATO POR DESEMPEÑO ACADÉMICO?
4	BEC_BNE	0.509	¿RECIBIÓ BECA EN EL BACHILLERATO POR NECESIDAD ECONÓMICA?
5	SER_INTE	0.615	INTERNET (Bienes y servicios con los que cuenta en casa)
6	HAB_PTEX	-0.532	HABILIDAD EN LA COMPUTADORA PARA CREAR Y EDITAR UN DOCUMENTO UTILIZANDO UN PROCESADOR DE TEXTO
7	HAB_PRES	-0.509	HABILIDAD PARA UTILIZAR PROGRAMAS PARA HACER PRESENTACIONES
8	HAB_BAJ	-0.575	HABILIDAD PARA BAJAR PROGRAMAS DE INTERNET
9	PTIC	0.592	CALIFICACIÓN DE TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN EN PORCENTAJE DE ACIERTOS

Como complemento del Análisis de Componentes Principales, se realizó otro análisis denominado **k – means**, el cual funciona como un método o algoritmo de agrupamiento, que tiene como objetivo la partición de un conjunto de  $n$  observaciones en  $k$  grupos en el que cada observación pertenece al grupo más cercano a la media.

Lo que hace el algoritmo es encontrar un centroide, que es un punto donde se encuentran atraídos todos los datos. Cada clase tiene  $n$  observaciones a partir de un umbral (threshold) que los divide.

Cada clase maneja su propio centroide (centros de gravedad) y funciona como una Gaussiana, la cual trata de crecer lo suficiente para evitar que se traslapen, para tener un componente aleatorio.

Se procedió a utilizar la herramienta para ejecutar el análisis como se muestra en la Figura 3.22, en donde aparece el cuadro de diálogo que permite elegir el rango para analizar, que abarca todos los campos generados y depurados del EXANI – II y las cinco materias, más no el promedio ni la Clase numérica derivada del mismo, como se puede apreciar en la Figura 3.23.

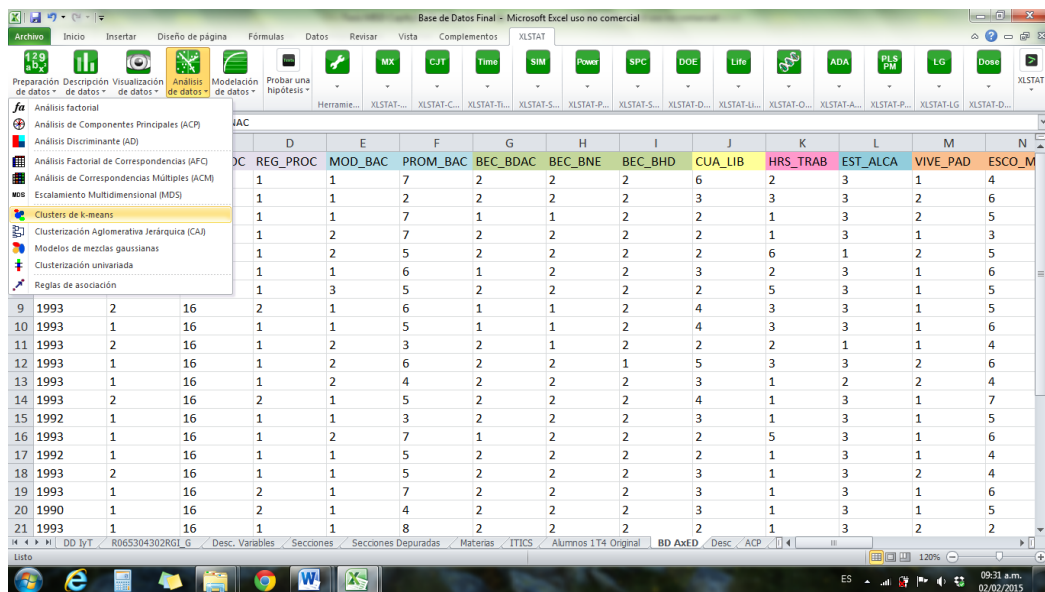


Figura 3.22 Opción: Análisis de Datos – Clusters de k – means.

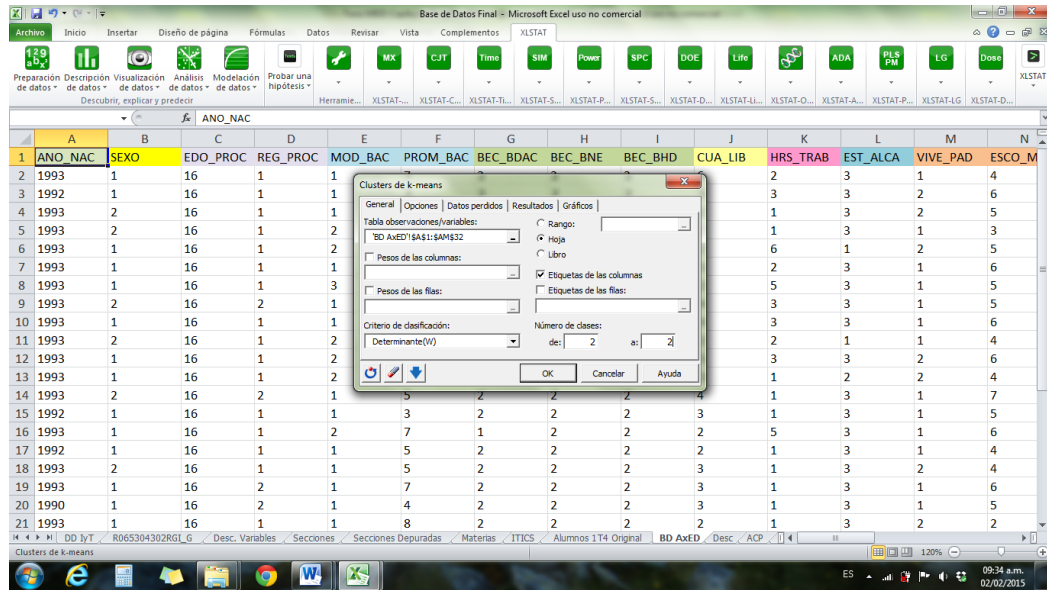


Figura 3.23 Cuadro de diálogo de Clusters de k – means.

Mientras que el Análisis de Componentes Principales identifica los atributos importantes dentro de la muestra poblacional, k – means ratifica que hay dos clases, a partir de todos los datos de la muestra poblacional, que se caracterizan por los que están por debajo o por encima del umbral 70.

Como este método no es paramétrico, se tienen nubes de datos y se espera que sólo se distribuyan las nubes y que muestren cuáles se encuentran en un lado y cuáles del otro (Clase 1 y Clase 2), para lograrlo se ratifica la selección como se muestra en la Figura 3.24, para que posteriormente se visualicen las observaciones correspondientes a cada una de las clases.

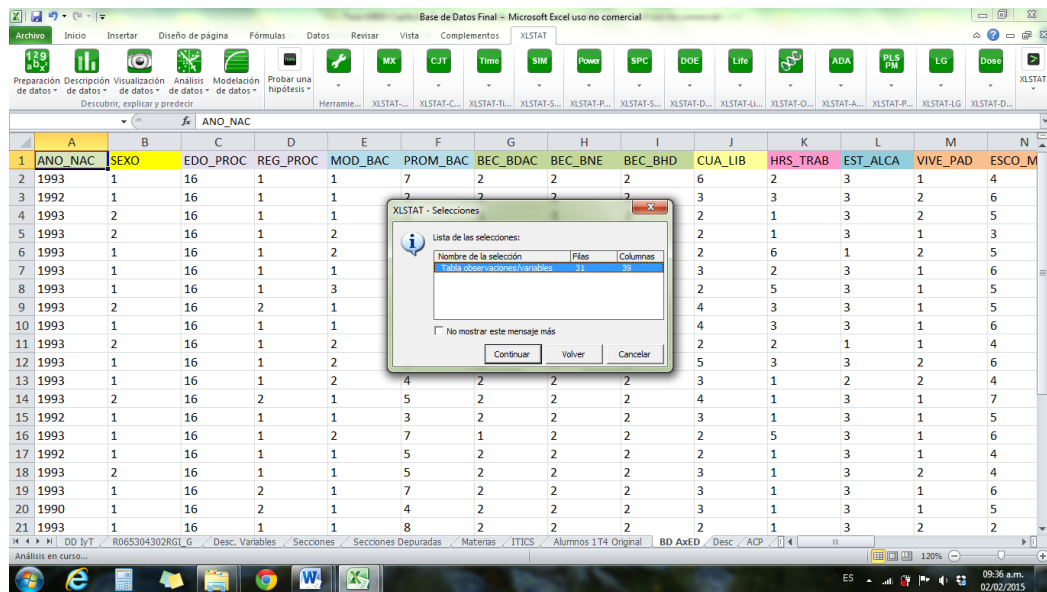


Figura 3.24 Cuadro de diálogo de Confirmación de Selecciones.

Previo a la visualización de los resultados del análisis, se procedió a la adecuación de la base de datos, por lo que se realizó una copia de la tabla en Excel denominada **BD AxED**, cambiándole el nombre a la Hoja, la cual se etiquetó como **BD AxAD**, como se aprecia en la Figura 3.25 (a). Posteriormente se incluyó un campo denominado **Promedio**, el cual se identificó como atributo discriminante, en donde por cada registro se calculó el promedio de las cinco materias del semestre.

Esto permitió que se pudieran determinar las dos clases numéricas, considerando la Clase 1 como los alumnos que están por debajo del valor 70, que significa que es la evaluación mínima aprobatoria y la Clase 2, los que se encuentran con promedio mayor o igual a 70, que son los que permanecerán en el ITM, como se aprecia en la Figura 3.25 (b).

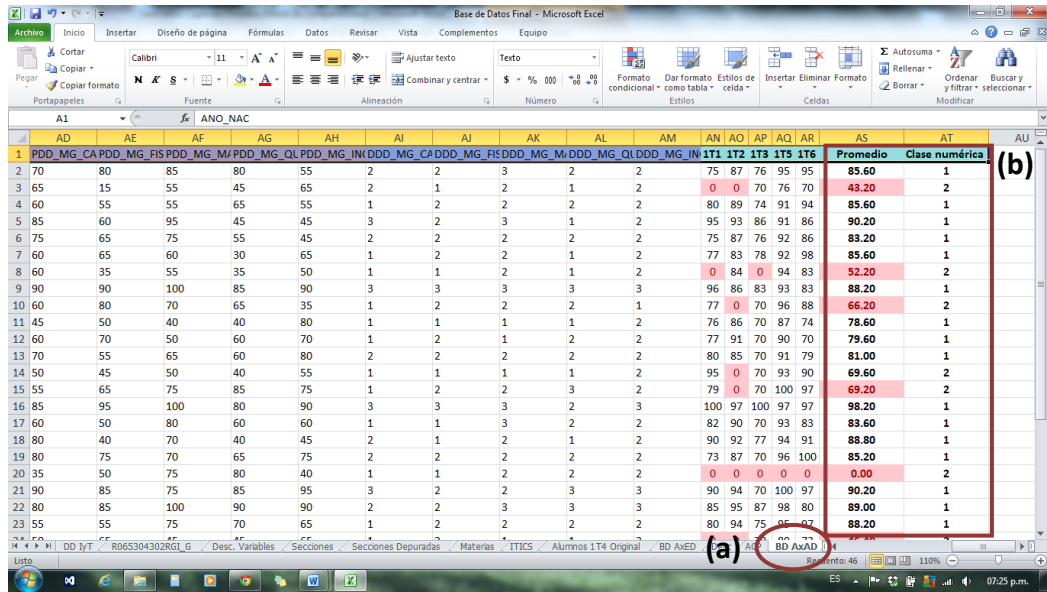


Figura 3.25 (a) Adaptación de la hoja para demostrar la relación.  
 (b) Preparación de los datos para comparar con k – means.

En la Figura 3.26 (a), se aprecia cada una de las observaciones, resultado del análisis de k-means, en donde la Clase 1 representada con el color azul, suma un total de 20, que significan los estudiantes que están por encima del umbral de 70 y que permanecen y los 11 son los que están por debajo y que son candidatos de tutoría. Cabe mencionar que estas deducciones son a través del algoritmo que tomo como referencia todos los datos, más nunca el promedio.

En la Figura 3.26 (b) se aprecia de una mejor manera esta situación al utilizar una gráfica de dispersión, que sí toma como origen de datos el promedio de calificación final y a pesar que en k-means no los toma, al final los datos coinciden en el número de clases y en el número de observaciones de cada una.

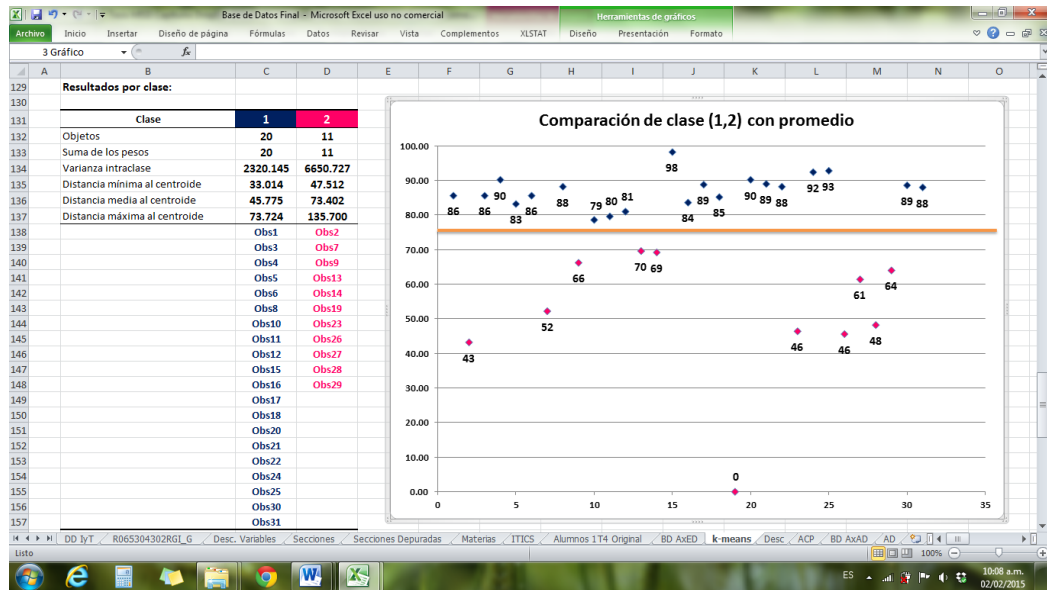


Figura 3.26 (a) Observaciones de las dos clases numéricas.  
 (b) Gráfica de dispersión de promedios del primer semestre.

El valor del umbral que fue 70, sirvió para tener variabilidad en los datos y así poder generar dos nubes y de ahí las dos clases.

La última técnica inferencial de minería de datos fue la denominada **Análisis Discriminante Lineal**, la cual es capaz de determinar qué variables permiten diferenciar a los grupos y cuántas de estas variables son necesarias para alcanzar la mejor clasificación posible. La pertenencia a los grupos, conocida de antemano, se utiliza como variable dependiente (una variable categórica con tantos valores discretos como grupos), en este caso es el promedio de las cinco materias del primer semestre.

Las variables en las que se supone que se diferencian los grupos, se utilizan como variables independientes o variables de clasificación (también llamadas variables discriminantes), en este caso serán todos los atributos depurados de la base de datos del EXANI – II, incluyendo las calificaciones de las cinco materias del primer semestre.

El objetivo de esta técnica es encontrar la combinación lineal de las variables independientes que mejor permite diferenciar (discriminar) a los grupos. Una vez encontrada esa combinación (la función discriminante) podrá ser utilizada para clasificar nuevos casos. Se trata de una técnica de análisis multivariante que es capaz de aprovechar las relaciones existentes entre una gran cantidad de variables independientes para maximizar la capacidad de discriminación.

Se utiliza nuevamente la herramienta XLSTAT y como se muestra en la Figura 3.27, se elige la opción correspondiente para ejecutar este análisis. Como se observa en el cuadro de diálogo de la Figura 3.28, la variable dependiente es determinada por la categorización de 1 o 2, a partir del umbral 70, basado en el promedio de las cinco materias.

En el caso de las variables independientes se toman el rango de los datos depurados del EXANI – II, junto con las calificaciones de cada una de las materias de los 31 registros; se selecciona el tipo de Gráficos a mostrar (Figura 3.29) y se corroboran los rangos de análisis (Figura 3.30).

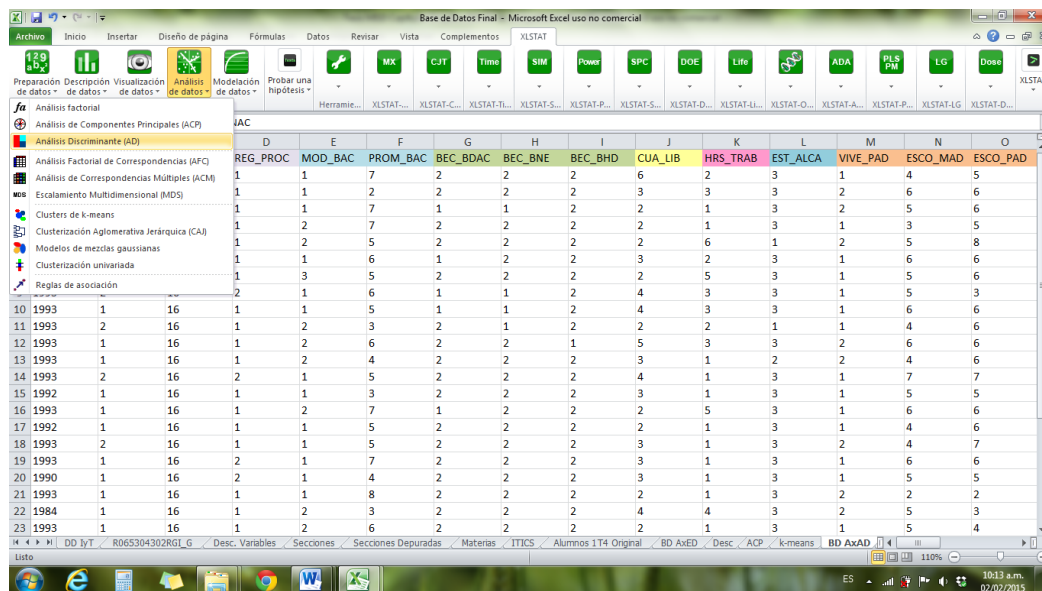


Figura 3.27 Opción: Análisis Discriminante Lineal (ADL)

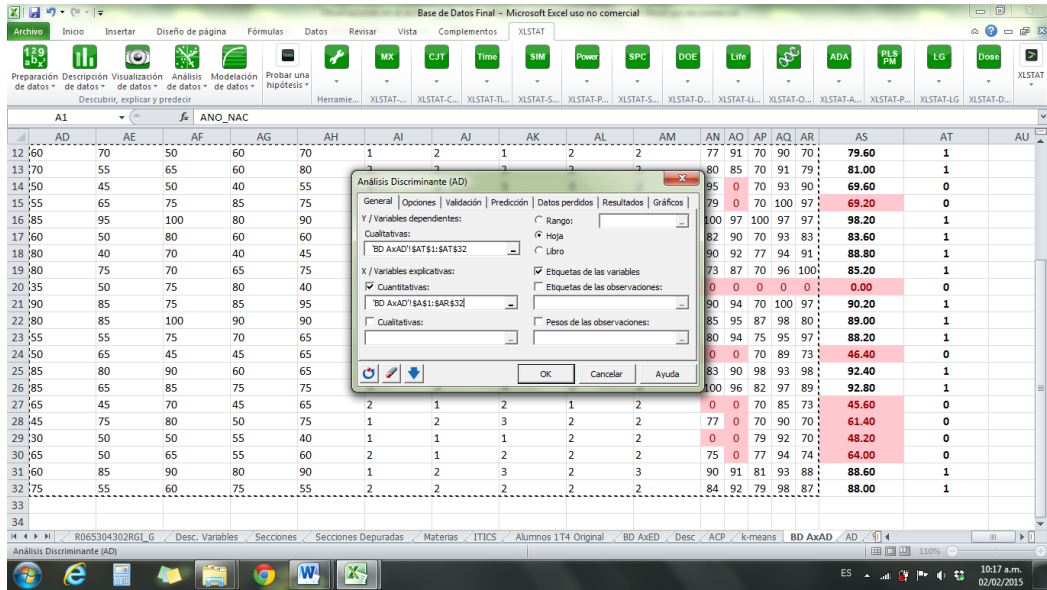


Figura 3.28 Cuadro de diálogo de Análisis Discriminante.

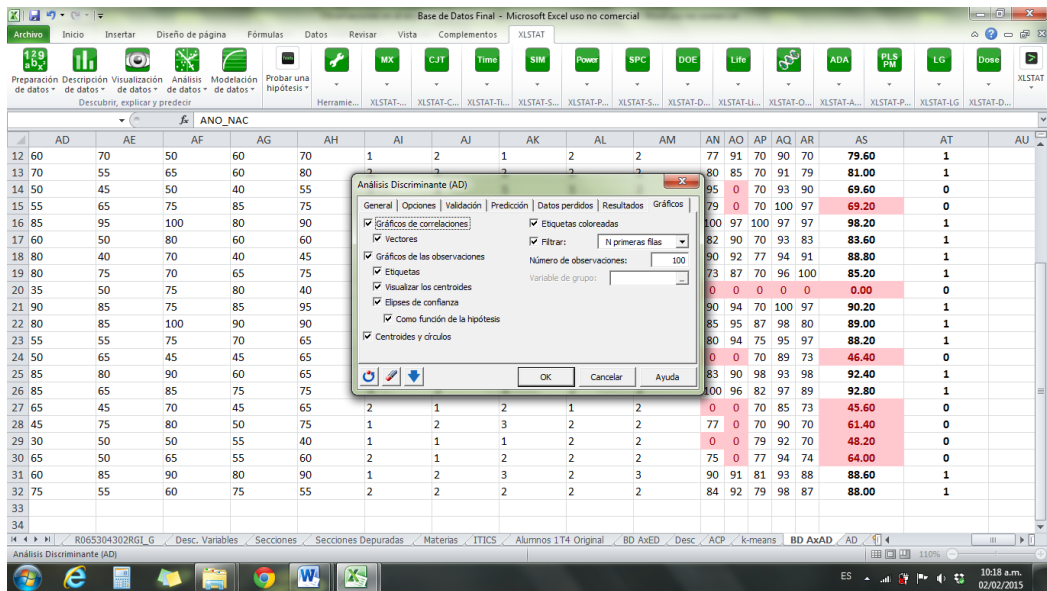


Figura 3.29 Cuadro de diálogo para la elección de gráficos.



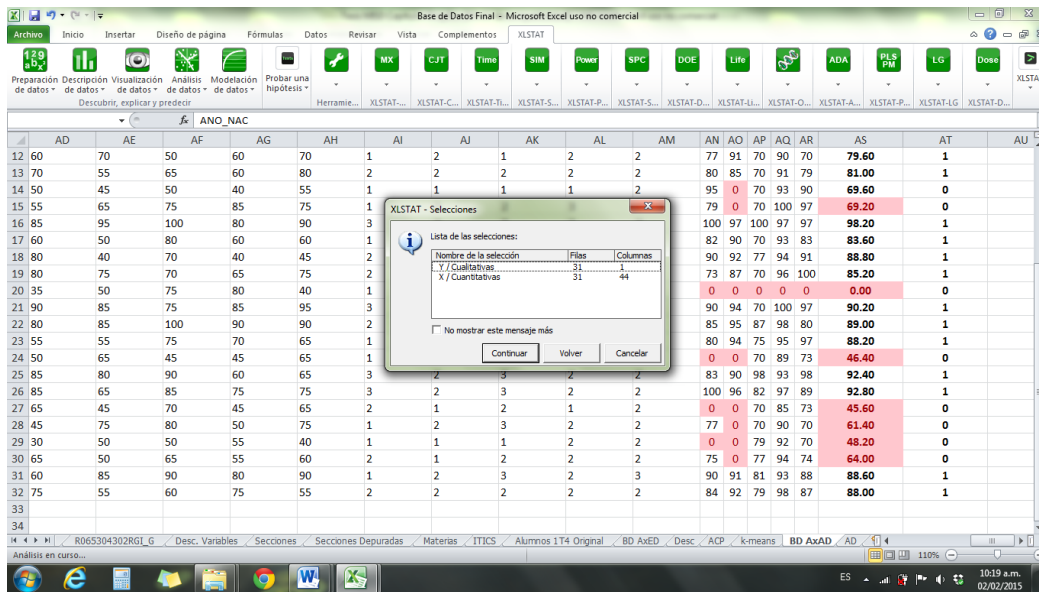


Figura 3.30 Cuadro de diálogo de Confirmación de Selecciones.

Los resultados de este análisis en primera instancia, resaltan que el Análisis Discriminante Lineal no es ajeno al k-means, dado que en las dos se demuestra que existen dos clases a partir del umbral 70, incluso se visualiza que existen 11 observaciones de la Clase 0 (menor a 70) y 20 observaciones de la Clase 1 (mayor o igual a 70), reiterando que la función de clasificación se encuentra en el umbral de aprobados y reprobados, enfocando a este último como elemento para realizar tutoría académica.

En segunda instancia, el análisis proporciona una serie de funciones de clasificación, que se identifican a partir de que son diferentes a 0, como se puede ver de manera parcial en la Figura 3.31, en donde se resaltan con color verde.

El total de variables o atributos significativos se muestra en la Tabla 3.9, en donde no sólo se puede observar el número y etiqueta del atributo, sino también su descripción.

Con el afán de separar los atributos derivados de una respuesta concreta por parte del encuestado, de los que son resultantes del promedio o diagnóstico de CENEVAL, se diferenciaron por tonalidades del color verde.

		0	1	
2824	Funciones de clasificación:			
2825				
2826				
2827	Intersección	-1238288392.075	-1667.213	
2828	ANO_NAC-1993	0.000	0.000	
2829	ANO_NAC-1992	0.000	0.000	
2830	1 ANO_NAC-1990	-1448753380.094	-1671733.970	AÑO DE NACIMIENTO
2831	ANO_NAC-1984	0.000	0.000	
2832	ANO_NAC-1991	0.000	0.000	
2833	SEXO-1	0.000	0.000	
2834	SEXO-2	0.000	0.000	
2835	EDO_PROC-16	0.000	0.000	
2836	EDO_PROC-21	0.000	0.000	
2837	2 REG_PROC-1	3627928.935	4205.000	¿EN QUÉ RÉGIMEN DE ESCUELA ESTUDIO EL ÚLTIMO AÑO DE BACHILLERATO?
2838	REG_PROC-2	0.000	0.000	
2839	MOD_BAC-1	0.000	0.000	
2840	MOD_BAC-2	0.000	0.000	
2841	3 MOD_BAC-3	515241395.787	594528.989	¿EN QUÉ MODALIDAD OBTUVO SU CERTIFICADO DEL BACHILLERATO?
2842	4 PROM_BAC-7	-275702.995	-319.000	¿CUÁL FUE SU PROMEDIO GENERAL EN EL BACHILLERATO?
2843	PROM_BAC-2	0.000	0.000	
2844	PROM_BAC-5	0.000	0.000	
2845	PROM_BAC-6	0.000	0.000	
2846	PROM_BAC-3	0.000	0.000	

Figura 3.31 Identificación de funciones clasificadoras general.

Tabla 3.9 Atributos derivados de las funciones clasificadoras en general.

No.	Atributo	0	1	Descripción del Atributo
1	ANO_NAC-1990	-1448753380.094	-1671733.970	AÑO DE NACIMIENTO
2	REG_PROC-1	3627928.935	4205.000	¿EN QUÉ RÉGIMEN DE ESCUELA ESTUDIO EL ÚLTIMO AÑO DE BACHILLERATO?
3	MOD_BAC-3	515241395.787	594528.989	¿EN QUÉ MODALIDAD OBTUVO SU CERTIFICADO DEL BACHILLERATO?
4	PROM_BAC-7	-275702.995	-319.000	¿CUÁL FUE SU PROMEDIO GENERAL EN EL BACHILLERATO?
5	BEC_BHD-1	338561508.946	390687.993	¿RECIBIÓ BECA EN EL BACHILLERATO POR HABILIDAD DEPORTIVA O ARTÍSTICA?
6	CUA_LIB-2	499553.991	580.000	EN LOS ÚLTIMOS SEIS MESES, ¿CUÁNTOS LIBROS HA LEÍDO (NO CUENTE LOS QUE LE DEJARON LEER EN LA ESCUELA)?
7	EST_ALCA-3	-890908.984	-1015.000	¿CUÁL ES EL NIVEL MÁXIMO DE ESTUDIOS QUE LE GUSTARÍA ALCANZAR?
8	BIEN_PC-3	-3151511.436	-36365.999	¿CUÁNTAS COMPUTADORAS HAY EN SU CASA?
9	LI_INTER-4	-730886.987	-841.000	HABILIDAD PARA LEER INFORMACIÓN EN INTERNET EN INGLÉS
10	HAB_PTEX-3	293276.995	348.000	HABILIDAD EN LA COMPUTADORA PARA CREAR Y EDITAR UN DOCUMENTO UTILIZANDO UN PROCESADOR DE TEXTO
11	PRLM-80	20308612.637	23432.000	CALIFICACIÓN DE RAZONAMIENTO LÓGICO MATEMÁTICO EN PORCENTAJE DE ACIERTOS
12	PMAT-80	-94298515.314	-108807.998	CALIFICACIÓN DE MATEMÁTICAS (SELECCIÓN) EN PORCENTAJE DE ACIERTOS
13	PESP-65	-36559226.346	-42194.999	CALIFICACIÓN DE ESPAÑOL EN PORCENTAJE DE ACIERTOS
14	PDD_MG_FIS-65	-1288295.977	-1479.000	CALIFICACIÓN DE FÍSICA EN PORCENTAJE DE ACIERTOS
15	PDD_MG_MAT-70	-174372704.882	-201201.996	CALIFICACIÓN DE MATEMÁTICAS "DIAGNÓSTICO" EN ÍNDICE DIAGNÓSTICO
16	PDD_MG_QUI-65	365188119.470	421398.992	CALIFICACIÓN DE QUÍMICA EN ÍNDICE DIAGNÓSTICO
17	PDD_MG_ING-75	-92492337.346	-106719.998	CALIFICACIÓN DE INGLÉS EN ÍNDICE DIAGNÓSTICO
18	DDD_MG_CAL-1	7086729.873	8178.000	DICTAMEN DE CÁLCULO

---

Con el fin de sacar el Perfil del Aspirante, en esta ocasión no se tomarán en cuenta las funciones clasificadoras de las calificaciones de las cinco materias (1T1 ... 1T6), dado que por su condición de no inscritos, no las generan todavía.

Sin embargo, en su momento se tuvieron que tomar en cuenta para la distribución de datos, que sirvió para corroborar en el modelo que los datos sean tomados en cuenta, que existe una conexión hacia lo que todavía no se sabe que vaya a pasar (calificaciones de primer semestre), pero evidentemente si está tocando esos puntos, significa que sí están correlacionados, por lo tanto se sabe que existe solución y si tiene solución, entonces se pueden tomar estos atributos para ser considerados en la detección de necesidades de tutoría.

Posteriormente se hizo el mismo análisis por cada una de las materias del primer semestre de la carrera, para determinar el Perfil del Aspirante pero por cada disciplina y obtener atributos más específicos para la toma de decisiones, lo cual puede estar enfocado a la estructura de temas que se pueden ofertar en el curso propedéutico o ya estando inscrito el alumno, ofrecerle asesorías técnicas dirigidas o invitarlo a participar en círculos de estudios, en un sentido preventivo.

Para lograr cada uno de estos análisis de forma aislada, se adaptaron las bases de datos, de tal suerte que se tomarán para cada caso las mismas variables independientes, sólo que incluyera el registro de calificaciones de la materia a tratar; se utilizó como variable dependiente el promedio de las cinco materias.

En total se hicieron un Análisis Discriminante Lineal Global como se muestra en la Figura 3.31 que incluye la Tabla 3.9 y cinco Análisis Discriminantes Lineales independientes por Materia, como se muestra de la Figura 3.32 a la Figura 3.36, con sus respectivas Tablas de Atributos, de la Tabla 3.10 a la Tabla 3.14.

		0	1
2257	<b>Funciones de clasificación (1T1 Cálculo Diferencial e Integral)</b>		
2258			
2259			
2260	Intersección	-264075053.042	-597298.202
2271	REG_PROC-2	0.000	0.000
2272	1 MOD_BAC-1	227968.996	10846.000
2273	MOD_BAC-2	0.000	0.000
2274	MOD_BAC-3	0.000	0.000
2275	2 PROM_BAC-7	-540095.989	-25664.999
2276	PROM_BAC-2	0.000	0.000
2277	PROM_BAC-5	0.000	0.000
2278	PROM_BAC-6	0.000	0.000
2279	PROM_BAC-3	0.000	0.000
2280	3 PROM_BAC-4	-139663.997	-6641.000
2281	PROM_BAC-8	0.000	0.000
2282	4 BEC_BDAC-2	26211562.486	1246593.976
2283	BEC_BDAC-1	0.000	0.000
2284	BEC_BNE-2	0.000	0.000
2285	BEC_BNE-1	0.000	0.000
2286	BEC_BHD-2	0.000	0.000
2287	BEC_BHD-1	0.000	0.000
2288	CUA_LIB-6	0.000	0.000
2289	CUA_LIB-3	0.000	0.000

Figura 3.32 Identificación de funciones clasificadoras de la materia de Cálculo Diferencial e Integral.

Tabla 3.10 Atributos derivados de las funciones clasificadoras de la materia de Cálculo Diferencial e Integral

No.	Atributo	0	1	Descripción del Atributo
1	MOD_BAC-1	227968.996	10846.000	¿EN QUÉ MODALIDAD OBTUVO SU CERTIFICADO DEL BACHILLERATO?
2	PROM_BAC-7	-540095.989	-25664.999	¿CUÁL FUE SU PROMEDIO GENERAL EN EL BACHILLERATO?
3	BEC_BDAC-2	26211562.486	1246593.976	¿RECIBIÓ BECA EN EL BACHILLERATO POR DESEMPEÑO ACADÉMICO?
4	HRS TRAB-5	64513775.736	3068141.940	MIENTRAS ESTUDIABA EL BACHILLERATO, ¿CUÁNTAS HORAS A LA SEMANA DEDICABA USTED A TRABAJAR?
5	ESCO_MAD-7	651684768.228	30992734.393	¿CUÁL ES EL MÁXIMO NIVEL DE ESTUDIOS ALCANZADO POR SU MADRE (AUNQUE HAYA FALLECIDO)?
6	ESCO_PAD-4	-147296101.113	-7005065.863	¿CUÁL ES EL MÁXIMO NIVEL DE ESTUDIOS ALCANZADO POR SU PADRE (AUNQUE HAYA FALLECIDO)?
7	BIEN_PC-5	210519985.874	10011872.804	¿CUÁNTAS COMPUTADORAS HAY EN SU CASA?
8	HAB_PTEX-3	4842390.905	230288.995	HABILIDAD EN LA COMPUTADORA PARA CREAR Y EDITAR UN DOCUMENTO UTILIZANDO UN PROCESADOR DE TEXTO
9	PRLM-80	-71384658.601	-3394913.933	CALIFICACIÓN DE RAZONAMIENTO LÓGICO MATEMÁTICO EN PORCENTAJE DE ACIERTOS
10	PMAT-75	-1328895.974	-63190.999	CALIFICACIÓN DE MATEMÁTICAS (SELECCIÓN) EN PORCENTAJE DE ACIERTOS
11	PRV-80	-148512738.089	-7062949.862	CALIFICACIÓN DE RAZONAMIENTO VERBAL EN PORCENTAJE DE ACIERTOS
12	PESP-70	-215585.996	-10237.000	CALIFICACIÓN DE ESPAÑOL EN PORCENTAJE DE ACIERTOS
13	PTIC-80	-3172193.938	-150857.997	CALIFICACIÓN DE TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN EN PORCENTAJE DE ACIERTOS
14	PDD_MG_CAL-60	10293085.798	489519.990	CALIFICACIÓN DE CÁLCULO EN PORCENTAJE DE ACIERTOS
15	PDD_MG_FIS-50	2075645.959	98715.998	CALIFICACIÓN DE FÍSICA EN PORCENTAJE DE ACIERTOS
16	PDD_MG_MAT-60	9753859.809	463912.991	CALIFICACIÓN DE MATEMÁTICAS "DIAGNÓSTICO" EN PORCENTAJE DE ACIERTOS
17	PDD_MG_QUI-85	142348498.210	6769817.867	CALIFICACIÓN DE QUÍMICA EN PORCENTAJE DE ACIERTOS
18	PDD_MG_ING-90	-42576987.166	-2024837.960	CALIFICACIÓN DE INGLÉS EN PORCENTAJE DE ACIERTOS
19	DDD_MG_MAT-3	-324799.994	-15457.000	DICTAMEN DE MATEMÁTICAS "DIAGNÓSTICO"
20	DDD_MG_QUI-2	-796455.984	-37873.999	DICTAMEN DE QUÍMICA

		0	1
2258	<b>Funciones de clasificación (1T2 Fundamentos de Programación)</b>		
2259		0	1
2260	Intersección	-89047993.692	-21213.135
2265	ANO_NAC-1991	0.000	0.000
2266	1 SEXO-1	60523.000	986.000
2267	SEXO-2	0.000	0.000
2268	EDO_PROC-16	0.000	0.000
2269	EDO_PROC-21	0.000	0.000
2270	REG_PROC-1	0.000	0.000
2271	REG_PROC-2	0.000	0.000
2272	2 MOD_BAC-1	-68962.000	-1015.000
2273	MOD_BAC-2	0.000	0.000
2274	MOD_BAC-3	0.000	0.000
2275	PROM_BAC-7	0.000	0.000
2276	3 PROM_BAC-2	78578631.971	1210692.000
2277	PROM_BAC-5	0.000	0.000
2278	PROM_BAC-6	0.000	0.000
2279	PROM_BAC-3	0.000	0.000
2280	PROM_BAC-4	0.000	0.000
2281	PROM_BAC-8	0.000	0.000
2282	BEC_BDAC-2	0.000	0.000
2283	BEC_BDAC-1	0.000	0.000

Figura 3.33 Identificación de funciones clasificadoras de la materia de Fundamentos de Programación

Tabla 3.11 Atributos derivados de las funciones clasificadoras de la materia de Fundamentos de Programación

No.	Atributo	0	1	Descripción del Atributo
1	SEXO-1	60523.000	986.000	SEXO
2	MOD_BAC-1	-68962.000	-1015.000	¿EN QUÉ MODALIDAD OBTUVO SU CERTIFICADO DEL BACHILLERATO?
3	PROM_BAC-2	78578631.971	1210692.000	¿CUÁL FUE SU PROMEDIO GENERAL EN EL BACHILLERATO?
4	BEC_BNE-1	-76792.000	-1189.000	¿RECIBIÓ BECA EN EL BACHILLERATO POR NECESIDAD ECONÓMICA?
5	CUA_LIB-2	-1078684.000	-16588.000	EN LOS ÚLTIMOS SEIS MESES, ¿CUÁNTOS LIBROS HA LEÍDO (NO CUENTE LOS QUE LE DEJARON LEER EN LA ESCUELA)?
6	ESCO_PAD-6	-281793.000	-4321.000	¿CUÁL ES EL MÁXIMO NIVEL DE ESTUDIOS ALCANZADO POR SU PADRE (AUNQUE HAYA FALLECIDO)?
7	LI_INTER-3	13360415.995	205871.000	HABILIDAD PARA LEER INFORMACIÓN EN INTERNET EN INGLÉS
8	LI_NOTI-1	186819420.930	2878336.999	HABILIDAD PARA LEER NOTICIAS EN INGLÉS
9	HAB_PTEX-3	3703966.999	57072.000	HABILIDAD EN LA COMPUTADORA PARA CREAR Y EDITAR UN DOCUMENTO UTILIZANDO UN PROCESADOR DE TEXTO
10	PMAT-65	24810021.991	382191.000	CALIFICACIÓN DE MATEMÁTICAS (SELECCIÓN) EN PORCENTAJE DE ACIERTOS
11	PRV-65	82496212.969	1270925.000	CALIFICACIÓN DE RAZONAMIENTO VERBAL EN PORCENTAJE DE ACIERTOS
12	PESP-90	6056301.998	93380.000	CALIFICACIÓN DE ESPAÑOL EN PORCENTAJE DE ACIERTOS
13	PTIC-95	-20437459.992	-314882.000	CALIFICACIÓN DE TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN EN PORCENTAJE DE ACIERTOS
14	PDD_MG_CAL-60	1458873.999	22475.000	CALIFICACIÓN DE CÁLCULO EN PORCENTAJE DE ACIERTOS
15	PDD_MG_FIS-90	14727330.994	227012.000	CALIFICACIÓN DE FÍSICA EN PORCENTAJE DE ACIERTOS
16	PDD_MG_MAT-50	-85176798.968	-1312250.000	CALIFICACIÓN DE MATEMÁTICAS "DIAGNÓSTICO" EN PORCENTAJE DE ACIERTOS
17	PDD_MG QUI-80	7417532.997	114289.000	CALIFICACIÓN DE QUÍMICA EN PORCENTAJE DE ACIERTOS
18	PDD_MG_ING-40	197625429.926	3044825.999	CALIFICACIÓN DE INGLÉS EN PORCENTAJE DE ACIERTOS
19	DDD_MG_MAT-2	128093.000	1972.000	DICTAMEN DE MATEMÁTICAS "DIAGNÓSTICO"
20	DDD_MG_ING-2	2664171.999	41122.000	DICTAMEN DE INGLÉS

	A	B	C	D	E	F	G	H	I	J	K	L
2257		Funciones de clasificación (1T3 Matemáticas Discretas)										
2258												
2259			0	1								
2260		Intersección	-1694255747.764	-2875625.161								
2261		ANO_NAC-1993	0.000	0.000								
2262		ANO_NAC-1992	0.000	0.000								
2263	1	ANO_NAC-1990	-351712641.122	14489821.129								
2275		PROM_BAC-7	0.000	0.000								
2276	2	PROM_BAC-2	1899189745.858	-78242609.695								
2277		PROM_BAC-5	0.000	0.000								
2278	3	PROM_BAC-6	-51636617.458	2127324.019								
2279		PROM_BAC-3	0.000	0.000								
2280		PROM_BAC-4	0.000	0.000								
2281		PROM_BAC-8	0.000	0.000								
2282		BEC_BDAC-2	0.000	0.000								
2283		BEC_BDAC-1	0.000	0.000								
2284		BEC_BNE-2	0.000	0.000								
2285	4	BEC_BNE-1	-32551891.289	1341076.012								
2286		BEC_BHD-2	0.000	0.000								
2287		BEC_BHD-1	0.000	0.000								
2288		CUA_LIB-6	0.000	0.000								
2289		CUA_LIB-3	0.000	0.000								
2290		CUA_LIB-2	0.000	0.000								

Figura 3.34 Identificación de funciones clasificadoras de la materia de Matemáticas Discretas

Tabla 3.12 Atributos derivados de las funciones clasificadoras de la materia de Matemáticas Discretas

No.	Atributo	0	1	Descripción del Atributo
1	ANO_NAC-1990	-351712641.122	14489821.129	AÑO DE NACIMIENTO
2	PROM_BAC-6	-51636617.458	2127324.019	¿CUÁL FUE SU PROMEDIO GENERAL EN EL BACHILLERATO?
3	BEC_BNE-1	-32551891.289	1341076.012	¿RECIBIÓ BECA EN EL BACHILLERATO POR NECESIDAD ECONÓMICA?
4	HRS_TRAB-1	-139635378.239	5752701.051	MIENTRAS ESTUDIABA EL BACHILLERATO, ¿CUÁNTAS HORAS A LA SEMANA DEDICABA USTED A TRABAJAR?
5	ESCO_MAD-5	-6007843.053	247515.002	¿CUÁL ES EL MÁXIMO NIVEL DE ESTUDIOS ALCANZADO POR SU MADRE (AUNQUE HAYA FALLECIDO)?
6	ESCO_PAD-4	296658083.633	-12221673.108	¿CUÁL ES EL MÁXIMO NIVEL DE ESTUDIOS ALCANZADO POR SU PADRE (AUNQUE HAYA FALLECIDO)?
7	BIEN_PC-2	-356584.003	14703.000	¿CUÁNTAS COMPUTADORAS HAY EN SU CASA?
8	LI_INTER-3	-109409518.971	4507441.040	HABILIDAD PARA LEER INFORMACIÓN EN INTERNET EN INGLÉS
9	PRLM-75	-588352.005	24244.000	CALIFICACIÓN DE RAZONAMIENTO LÓGICO MATEMÁTICO EN PORCENTAJE DE ACIERTOS
10	PRV-75	-334892.003	13804.000	CALIFICACIÓN DE RAZONAMIENTO VERBAL EN PORCENTAJE DE ACIERTOS
11	PESP-65	600422878.330	-24736159.220	CALIFICACIÓN DE ESPAÑOL EN PORCENTAJE DE ACIERTOS
12	PTIC-85	-1186381571.531	48876426.434	CALIFICACIÓN DE TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN EN PORCENTAJE DE ACIERTOS
13	PDD_MG_MAT-85	-9682230.086	398895.004	CALIFICACIÓN DE MATEMÁTICAS "DIAGNÓSTICO" EN PORCENTAJE DE ACIERTOS
14	PDD_MG QUI-45	2898427413.728	-119409154.060	CALIFICACIÓN DE QUÍMICA EN PORCENTAJE DE ACIERTOS
15	PDD_MG ING-65	-132900360.180	5475229.049	CALIFICACIÓN DE INGLÉS EN PORCENTAJE DE ACIERTOS
16	DDD_MG FIS-3	-29834214.265	1229136.011	DICTAMEN DE FÍSICA

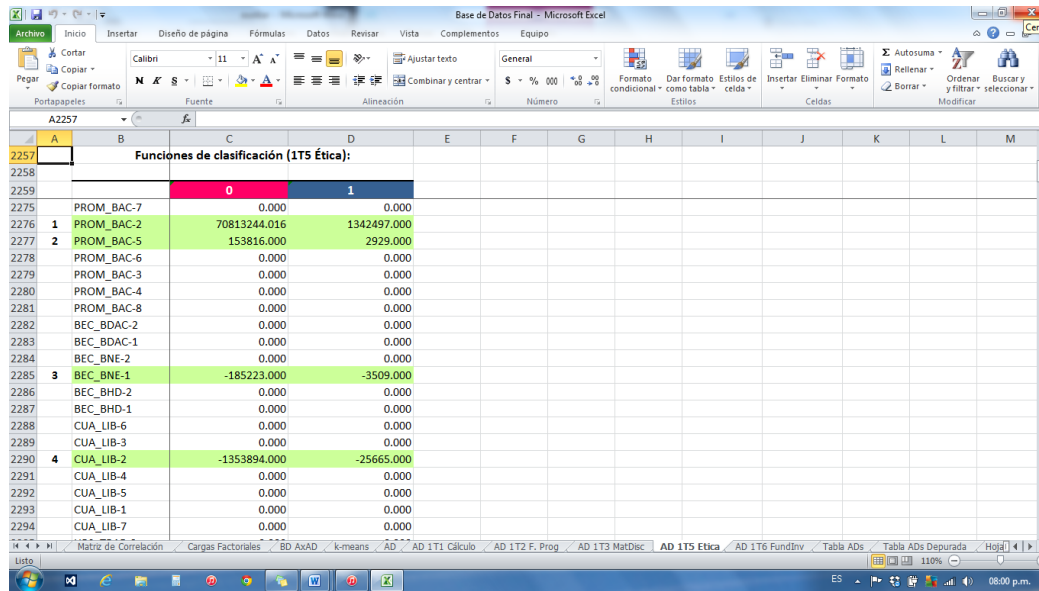


Figura 3.35 Identificación de funciones clasificadoras de la materia de Ética

Tabla 3.13 Atributos derivados de las funciones clasificadoras de la materia de Ética.

No.	Atributo	0	1	Descripción del Atributo
1	PROM_BAC-5	153816.000	2929.000	¿CUÁL FUE SU PROMEDIO GENERAL EN EL BACHILLERATO?
2	BEC_BNE-1	-185223.000	-3509.000	¿RECIBIÓ BECA EN EL BACHILLERATO POR NECESIDAD ECONÓMICA?
3	CUA_LIB-2	-1353894.000	-25665.000	EN LOS ÚLTIMOS SEIS MESES, ¿CUÁNTOS LIBROS HA LEÍDO (NO CUENTE LOS QUE LE DEJARON LEER EN LA ESCUELA)?
4	EST_ALCA-1	836766.000	15863.000	¿CUÁL ES EL NIVEL MÁXIMO DE ESTUDIOS QUE LE GUSTARÍA ALCANZAR?
5	LI_INTER-3	11929817.003	226171.000	HABILIDAD PARA LEER INFORMACIÓN EN INTERNET EN INGLÉS
6	LI_TAÇA-2	73486.000	1392.000	HABILIDAD PARA LEER TEXTOS ACADÉMICOS (LIBROS, REVISTAS) EN INGLÉS
7	LI_NOTI-1	170292582.039	3228425.001	HABILIDAD PARA LEER NOTICIAS EN INGLÉS
8	HAB_PTEX-3	4122495.001	78184.000	HABILIDAD EN LA COMPUTADORA PARA CREAR Y EDITAR UN DOCUMENTO UTILIZANDO UN PROCESADOR DE TEXTO
9	PMAT-75	123598.000	2349.000	CALIFICACIÓN DE MATEMÁTICAS (SELECCIÓN) EN PORCENTAJE DE ACIERTOS
10	PRV-65	75339941.017	1428308.000	CALIFICACIÓN DE RAZONAMIENTO VERBAL EN PORCENTAJE DE ACIERTOS
11	PESP-90	5401366.001	102428.000	CALIFICACIÓN DE ESPAÑOL EN PORCENTAJE DE ACIERTOS
12	PTIC-95	-18445334.004	-349682.000	CALIFICACIÓN DE TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN EN PORCENTAJE DE ACIERTOS
13	PDD_MG_CAL-60	1568697.000	29754.000	CALIFICACIÓN DE CÁLCULO EN PORCENTAJE DE ACIERTOS
14	PDD_MG_FIS-90	14058214.003	266568.000	CALIFICACIÓN DE FÍSICA EN PORCENTAJE DE ACIERTOS
15	PDD_MG_MAT-75	314737.000	5974.000	CALIFICACIÓN DE MATEMÁTICAS "DIAGNÓSTICO" EN PORCENTAJE DE ACIERTOS
16	PDD_MG QUI-80	6498639.001	123192.000	CALIFICACIÓN DE QUÍMICA EN PORCENTAJE DE ACIERTOS
17	PDD_MG_ING-40	180451340.042	3421043.001	CALIFICACIÓN DE INGLÉS EN PORCENTAJE DE ACIERTOS
18	DDD_MG_CAL-2	73805.000	1392.000	DICTAMEN DE CÁLCULO
19	DDD_MG QUI-2	331180.000	6293.000	DICTAMEN DE QUÍMICA
20	DDD_MG_ING-2	2649063.001	50257.000	DICTAMEN DE INGLÉS

		0	1	
2284	<b>Funciones de clasificación (175 Fundamentos de Investigación)</b>			
2285				
2286		0	1	
2287	Intersección	-4160987844.543	-16211533.875	
2288	ANO_NAC-1993	0.000	0.000	
2289	ANO_NAC-1992	0.000	0.000	
2290	1 ANO_NAC-1990	-1632644382.354	-101907261.791	
2292	ANO_NAC-1991	0.000	0.000	
2293	2 SEXO-1	580811.919	36249.995	
2294	SEXO-2	0.000	0.000	
2295	EDO_PROC-16	0.000	0.000	
2296	EDO_PROC-21	0.000	0.000	
2297	REG_PROC-1	0.000	0.000	
2298	REG_PROC-2	0.000	0.000	
2299	3 MOD_BAC-1	484883122.391	30265700.780	
2300	MOD_BAC-2	0.000	0.000	
2301	MOD_BAC-3	0.000	0.000	
2302	PROM_BAC-7	0.000	0.000	
2303	PROM_BAC-2	0.000	0.000	
2304	PROM_BAC-5	0.000	0.000	
2305	4 PROM_BAC-6	-51182846.863	-3194755.555	
2306	PROM_BAC-3	0.000	0.000	
2307	PROM_BAC-4	0.000	0.000	

Figura 3.36 Identificación de funciones clasificadoras de la materia de Fundamentos de la Investigación

Tabla 3.14 Atributos derivados de las funciones clasificadoras de la materia de Fundamentos de la Investigación

No.	Atributo	0	1	Descripción del Atributo
1	ANO_NAC-1990	-1632644382.354	-101907261.791	AÑO DE NACIMIENTO
2	SEXO-1	580811.919	36249.995	SEXO
3	MOD_BAC-1	484883122.391	30265700.780	¿EN QUÉ MODALIDAD OBTUVO SU CERTIFICADO DEL BACHILLERATO?
4	PROM_BAC-6	-51182846.863	-3194755.555	¿CUÁL FUE SU PROMEDIO GENERAL EN EL BACHILLERATO?
5	BEC_BNE-2	17992206.491	1123053.843	¿RECIBIÓ BECA EN EL BACHILLERATO POR NECESIDAD ECONÓMICA?
6	HRS TRAB-4	9208890.716	574808.920	MIENTRAS ESTUDIABA EL BACHILLERATO, ¿CUÁNTAS HORAS A LA SEMANA DEDICABA USTED A TRABAJAR?
7	ESCO_PAD-8	4740115865.068	295871066.746	¿CUÁL ES EL MÁXIMO NIVEL DE ESTUDIOS ALCANZADO POR SU PADRE (AUNQUE HAYA FALLECIDO)?
8	UJ_INTER-3	517784690.803	32319364.494	HABILIDAD PARA LEER INFORMACIÓN EN INTERNET EN INGLÉS
9	UJ_TACA-4	636587060.238	39734837.460	HABILIDAD PARA LEER TEXTOS ACADÉMICOS (LIBROS, REVISTAS) EN INGLÉS
10	UJ_NOTI-1	8455871426.966	527803087.406	HABILIDAD PARA LEER NOTICIAS EN INGLÉS
11	HAB_PRES-3	1650911.770	103065.986	HABILIDAD PARA UTILIZAR PROGRAMAS PARA HACER PRESENTACIONES
12	PMAT-75	-2479644.654	-154772.978	CALIFICACIÓN DE MATEMÁTICAS (SELECCIÓN) EN PORCENTAJE DE ACIERTOS
13	PESP-70	-169985187.298	-10610228.521	CALIFICACIÓN DE ESPAÑOL EN PORCENTAJE DE ACIERTOS
14	PTIC-85	486598443.152	30372768.765	CALIFICACIÓN DE TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN EN PORCENTAJE DE ACIERTOS
15	PDD_MG_CAL-65	2689862637.942	167897274.589	CALIFICACIÓN DE CÁLCULO EN PORCENTAJE DE ACIERTOS
16	PDD_MG_FIS-45	283421194.482	17690751.533	CALIFICACIÓN DE FÍSICA EN PORCENTAJE DE ACIERTOS
17	PDD_MG_MAT-100	-28672818.002	-1789705.750	CALIFICACIÓN DE MATEMÁTICAS "DIAGNÓSTICO" EN PORCENTAJE DE ACIERTOS
18	PDD_MG QUI-45	4794817252.441	299285439.270	CALIFICACIÓN DE QUÍMICA EN PORCENTAJE DE ACIERTOS
19	PDD_MG_ING-40	10083406300.033	629391378.242	CALIFICACIÓN DE INGLÉS EN PORCENTAJE DE ACIERTOS
20	DDD_MG_CAL-3	-4093262.429	-255489.964	DICTAMEN DE CÁLCULO
21	DDD_MG_MAT-2	17516432.558	1093357.848	DICTAMEN DE MATEMÁTICAS "DIAGNÓSTICO"
22	DDD_MG QUI-2	15866883.788	990407.862	DICTAMEN DE QUÍMICA



---

### 3.5 Fase 4. Evaluación e interpretación.

Medir la calidad de los patrones descubiertos por un algoritmo de minería de datos no es un problema trivial, ya que esta medida puede considerar varios criterios, algunos de ellos bastante subjetivos. Idealmente, los patrones descubiertos deben tener tres cualidades: ser precisos, comprensibles (inteligibles), e interesantes (útiles y novedosos). Según las aplicaciones puede ser interesante mejorar algún criterio y sacrificar ligeramente otro.

Normalmente, incluso después de la evaluación y validación pertinentes los modelos aprendidos no se pueden utilizar directamente, sino que necesitan una fase de refinamiento, que permita concretar cuál es el conocimiento que aportan y, en esa situación cómo se pueden utilizar en la toma de decisión final. (Hernández Orallo, Ramírez Quintana, & Ferri Ramírez, 2004).

En esta fase, se realiza una interpretación sobre los atributos generados por el Análisis de Componentes Principales (PCA) y el de Análisis Discriminatorio Lineal (LDA).

El procedimiento comparativo será que en primera instancia se observan los componentes principales generados por el PCA (F1 ... F3) y sus coincidencias; posteriormente los atributos iguales, que resultan del PCA y del LDA y por último las variables independientes, que aunque no coincidan en los dos estudios, son importantes para la toma de decisiones. Esto se puede apreciar en la Tabla 3.15.

Las celdas de color anaranjado, resaltan el atributo **PROM\_BAC** que tiene la característica de coincidir entre componentes principales y además entre PCA y LDA.

Hay que resaltar que el atributo **PROM\_BAC – 7**, genera información significativa ya que se debe tomar en cuenta para el Perfil del Aspirante que para ingresar a la carrera de Ingeniería en Tecnologías de Información y Comunicaciones (ITIC), debe tener como Promedio General de Bachillerato, una evaluación que se encuentre dentro del rango de 9.0 a 9.4.

**Tabla 3.15 Tabla comparativa de atributos del Análisis de Componentes Principales vs Análisis Discriminante Lineal**

		Atributos independientes (PCA)		Atributos que coinciden		Atributos independientes (LDA)	
Análisis de Componentes Principales (PCA)	F1	PROM_BAC	¿CUÁL FUE SU PROMEDIO GENERAL EN EL BACHILLERATO?	PROM_BAC-7			
		PRLM	CALIFICACIÓN DE RAZONAMIENTO LÓGICO MATEMÁTICO EN PORCENTAJE DE ACIERTOS	PRLM-80			
		PMAT	CALIFICACIÓN DE MATEMÁTICAS (SELECCIÓN) EN PORCENTAJE DE ACIERTOS	PMAT-80			
		PESP	CALIFICACIÓN DE ESPAÑOL EN PORCENTAJE DE ACIERTOS	PESP-65			
		PDD_MG_FIS	CALIFICACIÓN DE FÍSICA EN PORCENTAJE DE ACIERTOS	PDD_MG_FIS-65			
		PDD_MG_MAT	CALIFICACIÓN DE MATEMÁTICAS "DIAGNÓSTICO" EN ÍNDICE DIAGNÓSTICO	PDD_MG_MAT-70			
		PDD_MG_QUI	CALIFICACIÓN DE QUÍMICA EN ÍNDICE DIAGNÓSTICO	PDD_MG_QUI-65			
		PDD_MG_ING	CALIFICACIÓN DE INGLÉS EN ÍNDICE DIAGNÓSTICO	PDD_MG_ING-75			
		DDD_MG_CAL	DICTAMEN DE CÁLCULO	DDD_MG_CAL-1			
		¿CUÁL ES EL MÁXIMO NIVEL DE ESTUDIOS ALCANZADO POR SU PADRE (AUNQUE HAYA FALLECIDO)?	ESCO_PAD		REG_PROC-1	¿EN QUÉ RÉGIMEN DE ESCUELA ESTUDIO EL ÚLTIMO AÑO DE BACHILLERATO?	
	CALIFICACIÓN DE CÁLCULO EN PORCENTAJE DE ACIERTOS	PDD_MG_CAL		CUA_LIB-2	EN LOS ÚLTIMOS SEIS MESES, ¿CUÁNTOS LIBROS HA LEÍDO (NO CUENTE LOS QUE LE DEJARON LEER EN LA ESCUELA)?		
	DICTAMEN DE FÍSICA	DDD_MG_FIS		EST_ALCA-3	¿CUÁL ES EL NIVEL MÁXIMO DE ESTUDIOS QUE LE GUSTARÍA ALCANZAR?		
	DICTAMEN DE MATEMÁTICAS "DIAGNÓSTICO"	DDD_MG_MAT					
DICTAMEN DE QUÍMICA	DDD_MG_QUI						
DICTAMEN DE INGLÉS	DDD_MG_ING						
F2		LI_INTER	HABILIDAD PARA LEER INFORMACIÓN EN INTERNET EN INGLÉS	LI_INTER-4			
	¿CUÁL ES EL MÁXIMO NIVEL DE ESTUDIOS ALCANZADO POR SU MADRE (AUNQUE HAYA FALLECIDO)?	ESCO_MAD					
	HABILIDAD PARA LEER TEXTOS ACADÉMICOS (LIBROS, REVISTAS) EN INGLÉS	LI_TACA					
	HABILIDAD PARA LEER NOTICIAS EN INGLÉS	LI_NOTI					
	HABILIDAD PARA UTILIZAR PROGRAMAS PARA HACER PRESENTACIONES	HAB_PRES					
	HABILIDAD PARA EMPLEAR FUNCIONES BÁSICAS EN HOJA DE CÁLCULO (CAPTURA DE DATOS, FORMATO, ORDENAMIENTOS Y USO DE FÓRMULAS)	HAB_FBAS					
	HABILIDAD PARA BAJAR PROGRAMAS DE INTERNET	HAB_BAJ					
F3		ANO_NAC	AÑO DE NACIMIENTO	ANO_NAC-1990			
	¿CUÁL FUE SU PROMEDIO GENERAL EN EL BACHILLERATO?	PROM_BAC		MOD_BAC-3	¿EN QUÉ MODALIDAD OBTUVO SU CERTIFICADO DEL BACHILLERATO?		
	¿RECIBIÓ BECA EN EL BACHILLERATO POR DESEMPEÑO ACADÉMICO?	BEC_BDAC		BEC_BHD-1	¿RECIBIÓ BECA EN EL BACHILLERATO POR HABILIDAD DEPORTIVA O ARTÍSTICA?		
	¿RECIBIÓ BECA EN EL BACHILLERATO POR NECESIDAD ECONÓMICA?	BEC_BNE					
	INTERNET (Bienes y servicios con los que cuenta en casa)	SER_INTE		BIEN_PC-3	¿CUÁNTAS COMPUTADORAS HAY EN SU CASA?		
		HAB_PTEX	HABILIDAD EN LA COMPUTADORA PARA CREAR Y EDITAR UN DOCUMENTO UTILIZANDO UN PROCESADOR DE TEXTO	HAB_PTEX-3			
	HABILIDAD PARA UTILIZAR PROGRAMAS PARA HACER PRESENTACIONES	HAB_PRES					
	HABILIDAD PARA BAJAR PROGRAMAS DE INTERNET	HAB_BAJ					
CALIFICACIÓN DE TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN EN PORCENTAJE DE ACIERTOS	PTIC						

Análisis Discriminante (LDA)

---

En el caso del atributo **HAB\_PRES**, coincide entre los componentes principales F2 y F3, lo que significa que el Aspirante a ITIC debe tener habilidad para utilizar programas para hacer presentaciones electrónicas.

En el caso del atributo **HAB\_BAJ**, coincide entre los componentes principales F2 y F3, lo que significa que el Aspirante a ITIC debe tener habilidad para bajar programas de Internet.

Los doce atributos coincidentes en color amarillo, lo que significa que aparecen en los dos análisis (PCA y LDA), se pueden interpretar como la lista de requisitos necesarios que el Aspirante a la carrera de ITIC debe cubrir para su ingreso:

- ▶ Su promedio de bachillerato debe ser entre 9.0 y 9.4. **PROM\_BAC – 7**
- ▶ Su calificación de razonamiento lógico matemático en porcentaje de aciertos debe ser mínimo de 80 puntos (rango: 0 – 100). **PRLM-80**
- ▶ Su calificación de matemáticas (selección) en porcentaje de aciertos, debe ser mínimo de 80 puntos (rango: 0 – 100). **PMAT-80**
- ▶ Su calificación de español en porcentaje de aciertos, debe ser mínimo de 65 puntos (rango: 0 – 100). **PESP-65**
- ▶ Su calificación de física en índice diagnóstico debe ser mínimo de 70 puntos (rango: 0 – 100). **PDD\_MG\_FIS-65**
- ▶ Su calificación de matemáticas "diagnóstico" en índice diagnóstico debe ser mínimo de 70 puntos (rango: 0 – 100). **PDD\_MG\_MAT-70**
- ▶ Su calificación de química en índice diagnóstico debe ser mínimo de 65 puntos (rango: 0 – 100). **PDD\_MG QUI-65**
- ▶ Su calificación de inglés en índice diagnóstico debe ser mínimo de 75 puntos (rango: 0 – 100). **PDD\_MG\_ING-75**
- ▶ El dictamen de cálculo mínimo debe estar en la categoría de Elemental. **DDD\_MG\_CAL-1**

- ▶ El Aspirante deberá ser Muy Hábil (4), para leer información en Internet en inglés. **LI\_INTER-4**
- ▶ El año de nacimiento del Aspirante deberá ser en 1990, es decir que si el Examen fue en el 2011, los que ingresen deberán tener 21 años de edad. **ANO\_NAC-1990**
- ▶ Un nivel moderado de habilidad en la computadora para crear y editar un documento utilizando un procesador de texto. **HAB\_PTEX-3**

Los atributos independientes en cada uno de los análisis, se consideran también significativos para tomarse en cuenta en el Perfil del Aspirante, los cuales pueden servir como un examen diagnóstico secundario o como un filtro de selección, cuando el Aspirante se encuentre en igualdad de circunstancias que otro, a pesar de que los dos hayan cubierto los requisitos antes planteados. Para esto se describen los atributos que se iluminan en la Tabla 3.15, de color gris, pero que corresponden a cada análisis.

A continuación se enlistan los catorce requisitos que se deben considerar, derivados de los atributos del Análisis de Componentes Principales (PCA), los cuales se muestran en un orden de ideas lógico, por categorías; sin embargo se muestra la función clasificadora de cada atributo:

- ▶ El máximo nivel de estudios alcanzados del padre (aunque haya fallecido). **(F1)ESCO\_PAD**.
- ▶ El máximo nivel de estudios alcanzado por su madre (aunque haya fallecido). **(F2)ESCO\_MAD**
- ▶ Si recibió beca en el bachillerato por desempeño académico. **(F3)BEC\_BDAC**
- ▶ Si recibió beca en el bachillerato por necesidad económica. **(F3)BEC\_BNE**
- ▶ Si cuenta con servicio de Internet. **(F3)SER\_INTE**
- ▶ El nivel de habilidad para leer textos académicos (libros, revistas) en inglés. **(F2)LI\_TACA**
- ▶ El nivel de habilidad para leer noticias en inglés. **(F2)LI\_NOTI**

- 
- ▶ El nivel de habilidad para emplear funciones básicas en hoja de cálculo (captura de datos, formato, ordenamientos y uso de fórmulas) .  
(F2)HAB\_FBAS
  - ▶ El resultado de la calificación de tecnologías de información y comunicación en porcentaje de aciertos. (F3)PTIC
  - ▶ El resultado de la calificación de cálculo en porcentaje de aciertos.  
(F1)PDD\_MG\_CAL
  - ▶ El resultado del dictamen de física. (F1)DDD\_MG\_FIS
  - ▶ El resultado del dictamen de matemáticas "diagnóstico". (F1)DDD\_MG\_MAT
  - ▶ El resultado del dictamen de química. (F1)DDD\_MG\_QUI
  - ▶ El resultado del dictamen de inglés. (F1)DDD\_MG\_ING

Es importante resaltar que estas variables, no determinan cual es el valor óptimo para considerarlo, se puede suponer que se espera el máximo valor óptimo, sin embargo, al realizar el estudio del Análisis Discriminante Lineal (LDA), las variables resultantes sí especifican el valor constante que se debe considerar, como se muestra en los seis requisitos que a continuación se enlistan:

- ▶ El régimen de escuela en la que estudió el Aspirante, en el último año de bachillerato es preferentemente Público. REG\_PROC-1
- ▶ El Aspirante debió haber leído por lo menos un libro, que no tengan que ver con los que le dejaron leer en la escuela. CUA\_LIB-2
- ▶ Las expectativas sobre el nivel máximo de estudios que el Aspirante pretende alcanzar es Posgrado (especialidad, maestría, doctorado) EST\_ALCA-3
- ▶ La modalidad que el Aspirante debe tener para obtener su certificado de Bachillerato, es el de Profesional Técnico Bachiller. MOD\_BAC-3
- ▶ El Aspirante debió haber recibido una beca por habilidad deportiva o artística. BEC\_BHD-1
- ▶ Preferentemente, el Aspirante debe tener como mínimo dos computadoras en su casa. BIEN\_PC-3

---

### 3.6 Fase 5. Difusión, uso y monitorización.

Una vez construido y validado el modelo puede usarse principalmente para dos finalidades: para que un analista recomiende acciones basadas en el modelo y en sus resultados o bien para aplicar el modelo a diferentes conjuntos de datos. También puede incorporarse a otras aplicaciones.

Tanto en el caso de una aplicación manual o automática del modelo, es necesario su difusión, es decir que se distribuya y se comunique a los posibles usuarios, ya sea por causas habituales dentro de la organización, reuniones, intranet, etc. El nuevo conocimiento extraído debe integrar el *know – how* de la organización.

También es importante medir lo bien que el modelo evoluciona. Aún y cuando el modelo funcione bien debemos continuamente comprobar las prestaciones del mismo. Esto se debe principalmente a que los patrones pueden cambiar. Por lo tanto el modelo debe ser monitoreado, lo que significa que de tiempo en tiempo el modelo tendrá que ser reevaluado, reentrenado y posiblemente reconstruido completamente. (Hernández Orallo, Ramírez Quintana, & Ferri Ramírez, 2004)

Los resultados obtenidos y descritos en la Fase 4, permiten tener un instrumento de evaluación para que el ITM, pueda elegir mejor a sus Aspirantes y así disminuir los niveles de reprobación y deserción en el primer semestre de la carrera de ITIC.

Es importante aclarar que los resultados de estos análisis, sólo funcionan no sólo para la base de datos generada por el EXANI – II de CENEVAL, sino que sólo para esa generación de Aspirantes que aplicaron dicho examen en Junio del 2011 y que además les interesa ingresar a la carrera de ITIC.

---

Podría parecer que es limitado, sin embargo no lo es, a partir de que este Modelo Inteligente, obtiene resultados diferentes para cada circunstancia, es decir, por cada generación de estudiantes y por cada carrera que aspiren a ingresar.

Para ejecutar esta Fase 5 de difusión, uso y monitorización, se recomienda en primer lugar crear un instrumento de selección de Aspirantes, basado en los atributos significativos que resultaron de los Análisis de Minería de Datos, sobre la información generada por el EXANI – II, cada vez que se aplique.

Como resultado de esto, se elige a los Aspirantes que cumplan con estos atributos en sus valores o rangos más óptimos, para que puedan ser aceptados y se inscriban al primer cuatrimestre.

En el caso de los que no cubran los requisitos planteados a su vez pueden considerarse como candidatos a cursar un propedéutico, cuyos temas no sólo serán de formación del Modelo Institucional, sino que también se puede agotar temas específicos, relacionados con los atributos significativos. Por ejemplo en este caso práctico se identificaron variables que tienen que ver con las áreas de Razonamiento Lógico Matemático, Matemáticas, Español, Física, Química, Inglés y Cálculo. En este análisis no se consideró como relevante el Razonamiento Verbal.

También existen variables significativas, respecto a los elegidos por el Modelo Inteligente, que pueden servir de referencia para la toma de decisiones, a nivel de Tutor y para las áreas encargadas de hacer estudios socioeconómicos del Alumno o las que se encargan de generar estudios sicométricos o psicológicos.

Se demuestra con esta tesis que a partir de la aplicación de Minería de Datos al EXANI – II, se puede obtener información valiosa para generar el Perfil del Aspirante para ingresar a una Institución de Educación Superior, con características muy particulares, que favorezcan la calidad de los estudiantes.

---

## IV. RESULTADOS Y DISCUSIÓN

*“El más importante y principal negocio público, es la buena educación de la juventud”. Platón*

En este capítulo se podrá apreciar a detalle las fases del Modelo Inteligente de Tutorías (MIT), propuesta que es inspirada en la necesidad que tienen las Instituciones de Educación Superior (IES), en formar individuos de manera integral, que abarque no sólo los conocimientos tecnológicos que brinda por naturaleza cualquier programa educativo, sino que también exista la atención personalizada por parte de un tutor académico, en función de las necesidades propias del estudiante y del período y circunstancia en que se encuentra cursando su carrera profesional.

El MIT aplica algoritmos de minería de datos, a través de insumos cíclicos de información, para generar perfiles idóneos de tutores y alumnos, para que a través de un sistema experto, se generen las condiciones óptimas en la gestión y seguimiento del trabajo tutorial, en IES. Se compone de cuatro fases:

- ▶ Primera Fase: Inicio  
→ Captación y Primer Semestre (S1).
- ▶ Segunda Fase: Formación  
→ Del Segundo Semestre (S2) al Tercer Semestre (S3).
- ▶ Tercera Fase: Seguimiento  
→ Del Cuarto Semestre (S4) al Sexto Semestre (S6).
- ▶ Cuarta Fase: Cierre  
→ Del Séptimo Semestre (S7) hasta titularse.

Se identifican dos tipos de perfiles en el proceso de enseñanza – aprendizaje, el primero es el perfil del tutor, que a su vez se clasifica en tres tipos, que se definen a continuación:



- ▶ Tutor para Formación. (Primera y Segunda Fase = 3 semestres)
- ▶ Tutor para Seguimiento. (Tercera Fase = 3 semestres)
- ▶ Tutor para Cierre. (Cuarta Fase = 3 semestres hasta Titulación)

El segundo es el perfil del alumno, que de manera paralela con los perfiles de tutor, es decir, en los mismos períodos de tiempo, se identifican cuatro tipos:

- ▶ Aspirante
- ▶ Alumno en Formación
- ▶ Alumno en Seguimiento
- ▶ Alumno en Cierre

En lo subsecuente, se describe cada una de las Fases del MIT, para lo cual se presenta en la Figura 4.1, la nomenclatura utilizada para su entendimiento:

**Nomenclatura:**

**CENEVAL:** Centro Nacional de Evaluación para la Educación Superior.

**EXANI-II:** Examen Nacional de Ingreso a la Educación Superior.

**CPNMS:** Competencias Profesionales a Nivel Medio Superior.

**PAS:** Perfil del Aspirante.

**GALFi:** Grupo de Alumnos en Formación inicial.

**PALF:** Perfil del Alumno en Formación.

**GALF:** Grupo de Alumnos en Formación.

**PALS:** Perfil del Alumno en Seguimiento.

**GALS:** Grupo de Alumnos en Seguimiento.

**PALC:** Perfil del Alumno en Cierre.

**GALC:** Grupo de Alumnos en Cierre.

**PTF:** Perfil del Tutor para Formación.

**PTS:** Perfil del Tutor para Seguimiento.

**PTC:** Perfil del Tutor para Cierre.

**PTA:** Procedimiento de Tutorías Académicas en el Sistema Nacional de Institutos Tecnológicos.

Figura 4.1 Nomenclatura utilizada en el MIT

#### 4.1 Fase 1.- De Aspirantes a Alumnos de Nuevo Ingreso.

Se efectúa a partir de que los Aspirantes atienden la oferta educativa que el ITM ofrece, y como requisito de ingreso, acuden a realizar el EXANI – II, de CENEVAL, el cual se presenta como uno de los principales insumos de esta etapa, considerando tanto su Examen de Selección como su Examen Diagnóstico; aunado a esto se toma en cuenta algún instrumento de evaluación de las Competencias Profesionales a Nivel Medio Superior, así como un Examen Psicopedagógico.

Los insumos del Aspirante son ingresados al MIT, el cual entrega como resultado el Perfil del Aspirante (PAS) y a partir de esto se forma un Grupo de Alumnos en Formación Inicial (GALFi), como se muestra en la Figura 4.2.



Figura 4.2 Fase de Formación de Grupo de Alumnos en Formación Inicial del MIT.

Al mismo tiempo se ingresan al MIT los insumos del Docente, que son su Trayectoria Académica en la Institución, su Experiencia Laboral Docente a Nivel Superior, su Experiencia Laboral Tecnológica, la aplicación de un Instrumento de Evaluación de Competencias Profesionales y un Examen Psicopedagógico.

Los insumos del Docente son ingresados al MIT y se tendrá como resultado el Perfil del Tutor para Formación (PTF), como se muestra en la Figura 4.3.



Figura 4.3 Fase de Perfil del Tutor para Formación del MIT.

En el primer semestre de la carrera, al GALFi se le asignará un PTF, para que a través del Procedimiento para las Tutorías Académicas (PTA), sirva de acompañamiento académico e inculque al alumno en formación, los valores institucionales y el modelo educativo del ITM, como se muestra en la Figura 4.4.

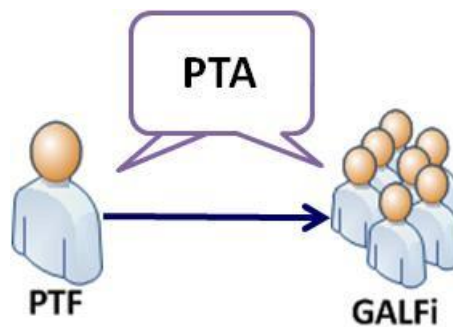


Figura 4.4 Procedimiento para las Tutorías Académicas (PTA).

#### 4.2 Fase 2.- Alumnos en Formación Inicial con Tutor para Formación.

A partir del segundo semestre, el alumno tiene la libertad de escoger sus materias, lo que trae como consecuencia que el GALFi que se formó, no se conserve físicamente.

Por lo anterior, se crea un Perfil de Alumno en Formación (PALF), a partir de los insumos que ingresa al MIT, como se observa en la Figura 4.5, basado en su trayectoria dentro del ITM.

Los insumos son Materias a las que se inscribe, Entrevistas con el tutor, Examen sicopedagógico, Examen Socioeconómico, Kardex (Historial Académico), Bitácora de Tutorías y Bitácora de Asesorías.

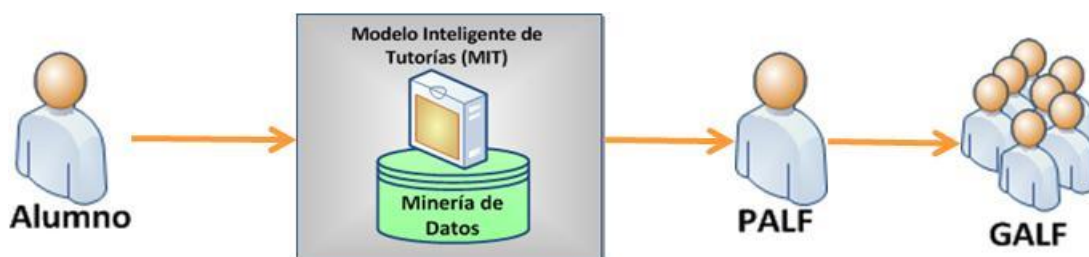


Figura 4.5 Fase de Formación de Grupo de Alumnos en Formación (GALF) del MIT.

Durante el segundo y tercer semestre, como se aprecia en la Figura 4.6, el Grupo de Alumnos en Formación (GALF), conserva al Docente que tenga el Perfil de Tutor en Formación (PTF), el cual seguirá aplicando el PTA.

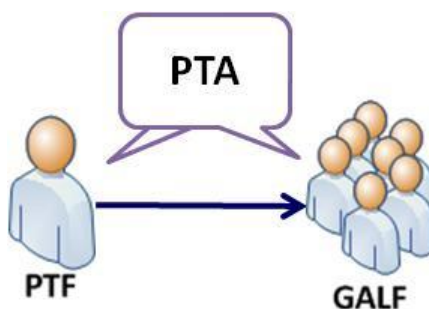


Figura 4.6 Procedimiento para las Tutorías Académicas para Grupo de Alumnos en Formación.

El MIT considera la circunstancia de que el alumno, a partir del segundo semestre, puede cambiar de grupo en función de las materias a las que se inscribe, ya que puede contener un GALF virtual, con información obtenida a partir de la aplicación de minería de datos, que gestiona el conocimiento para la toma de decisiones del PTF.

### 4.3 Fase 3.- Alumnos en Seguimiento con Tutor para Seguimiento.

Del cuarto al sexto semestre, se considera que el alumno ya tiene cierto grado de madurez académico a partir de que se ha mantenido e inicia con las materias de especialidad, por lo que es importante volver a ingresar los insumos actualizados, tanto del Alumno como del Docente, con el objeto de que el MIT, determine un Perfil de Alumno en Seguimiento (PALS) y un Perfil de Tutor para Seguimiento (PTS), respectivamente. Esto se muestra en la Figura 4.7.

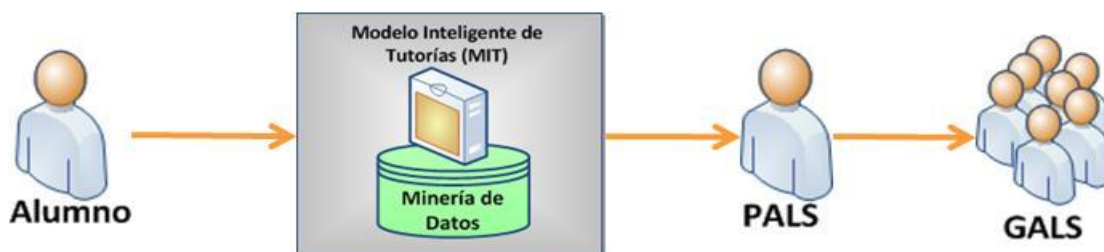


Figura 4.7 Fase de Formación de Grupo de Alumnos en Seguimiento (GALS) del MIT.

Se considera como una ventaja, que el Docente, con el perfil idóneo para este fin, sea una persona diferente al anterior, la cual asuma la responsabilidad del Grupo de Alumnos en Seguimiento (GALS), se puede observar en la Figura 4.8, que es necesario volver a ingresar los insumos actualizados del Docente al MIT, para obtener un Perfil de Tutor para Seguimiento (PTS).



Figura 4.8. Fase de Perfil del Tutor para Seguimiento del MIT.

La función principal del PTS, es brindarle apoyo a cada uno de los alumnos del GALS, aplicando en todo momento el PTA, como se observa en la Figura 4.9, para que continúe con su carrera durante la etapa más difícil, porque para entonces ya ha invertido tiempo, dinero y esfuerzo, además que el grado de dificultad de las materias aumenta.

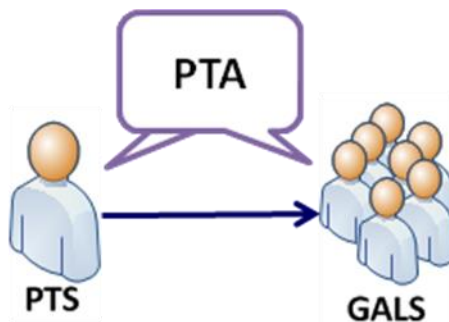


Figura 4.9 Procedimiento para las Tutorías Académicas para Grupo de Alumnos en Seguimiento.

#### 4.4 Fase 4.- Alumnos en Cierre con Tutor para Cierre.

Del séptimo al noveno semestre, se considera que el alumno ya está más que involucrado con su carrera y tiene la madurez académica para terminar, sin embargo, es importante que cuente con un guía, conforme al PTA, para la culminación de su grado a nivel superior.

De igual forma, como se aprecia en la Figura 4.10, así como en las etapas anteriores, es necesario que exista una retroalimentación al modelo, a partir de los insumos actualizados, en cada caso, para generar el nuevo conocimiento y crear los perfiles idóneos de PALC para formar el grupo virtual y del PTC, para que tenga las cualidades que se requieren para guiar al alumno en su proceso de culminación de su carrera.

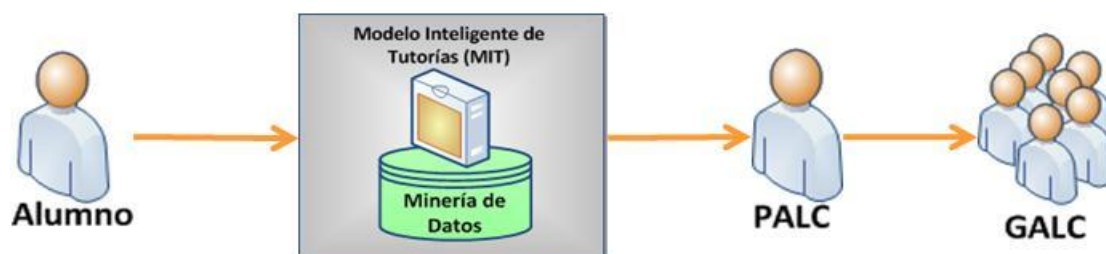


Figura 4.10 Fase de Formación de Grupo de Alumnos en Cierre (GALC) del MIT.

En esta última etapa, se sugiere que a través del MIT se obtenga el Perfil de Alumno en Cierre (PALC), para formar un Grupo de Alumnos en Cierre (GALC) y asignarle un Perfil de Tutor para Cierre (PTC), como se muestra en la Figura 4.11, el cual no sólo será responsable de que el GALC termine sus créditos y se convierta en un grupo de jóvenes egresados, sino que además garantice un seguimiento hasta que cada integrante de su grupo obtenga el grado superior a través de su titulación como se muestra en la Figura 4.12.

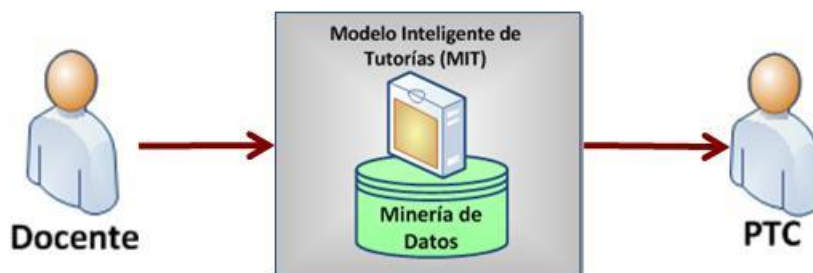


Figura 4.11 Fase de Perfil del Tutor para Cierre del MIT.

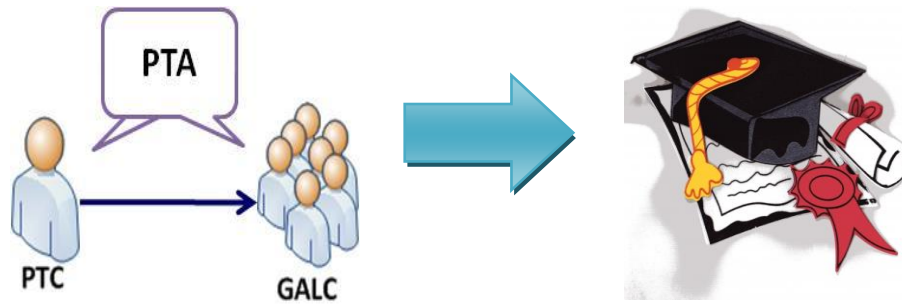


Figura 4.12 Propósito del MIT para que el estudiante obtenga el grado de Licenciatura.

En la Figura 4.13, se muestra completo el Modelo Inteligente de Tutorías Académicas en Instituciones de Educación Superior, donde se incluye todo el proceso que se ha discutido a lo largo de este capítulo.

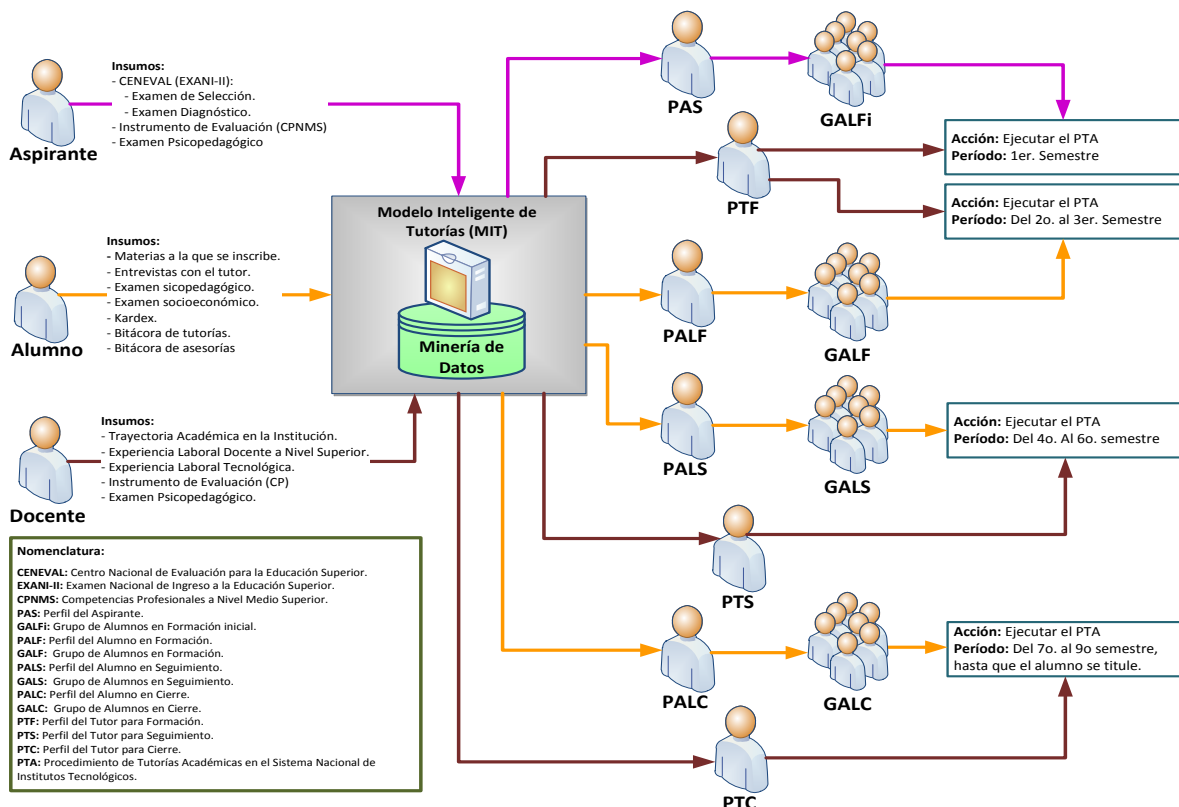


Figura 4.13 Modelo Inteligente de Tutorías Académicas en Instituciones de Educación Superior.



---

El alcance de este modelo para alcanzar el grado de maestría, fue la obtención del Perfil del Aspirante, sólo tomando en consideración el Examen de Ingreso (EXANI – II), elaborado por CENEVAL y que fue aplicado en Junio del 2011, en el Instituto Tecnológico de Morelia. Dicho examen generó una base de datos en Excel en la cual se aplicaron las técnicas de Minería de Datos planteadas en esta tesis y se obtuvo como resultado un instrumento de selección de Aspirantes, tomando en cuenta los atributos significativos de dichos análisis.

Sin embargo, como se puede apreciar en el MIT, no sólo se sugiere como insumo el EXANI – II, sino también la aplicación de un Instrumento de Evaluación de Competencias Profesionales y un Examen Psicopedagógico, cuya información deberá estar contenida en una base de datos a la que se le pueda aplicar Minería de Datos.

Con el MIT se espera que cada una de las personas que ingresan al ITM, genere un almacén de datos, durante toda su estancia académica, desde que es aspirante hasta que obtiene el grado a nivel superior, dicho almacén de datos será sujeto a la aplicación de análisis estadísticos descriptivos e inferenciales, en cada semestre, con el afán de obtener información valiosa para la generación de los perfiles de alumnos correspondientes a cada nivel académico.

El modelo se vuelve inteligente a partir de que las bases de datos van aumentando en información a través del tiempo y los resultados de la aplicación de Minería de Datos son diferentes para cada circunstancia. Se espera que los atributos significativos del mismo alumno con el perfil de Formación (PALF) comparada con el perfil de Seguimiento (PALS) y con el de Cierre (PALC), sean diferentes, ya que son situaciones diferentes que involucran a la misma persona, por lo tanto los instrumentos de tutoría y asesoría serán adaptados para cada caso.

---

Se podría pensar que quizá en este caso el tomador de decisiones, tendrá que esperar el resultado particular de cada alumno, para ejecutar la tutoría, sin embargo, a partir de que el MIT también pretende generar grupos con características similares, a partir de la aplicación de Minería de Datos, como es el caso de los ya denominados GALFi, GALF, GALS y GALC, se aplicaría la labor tutorial de manera grupal y no necesariamente para cada estudiante.

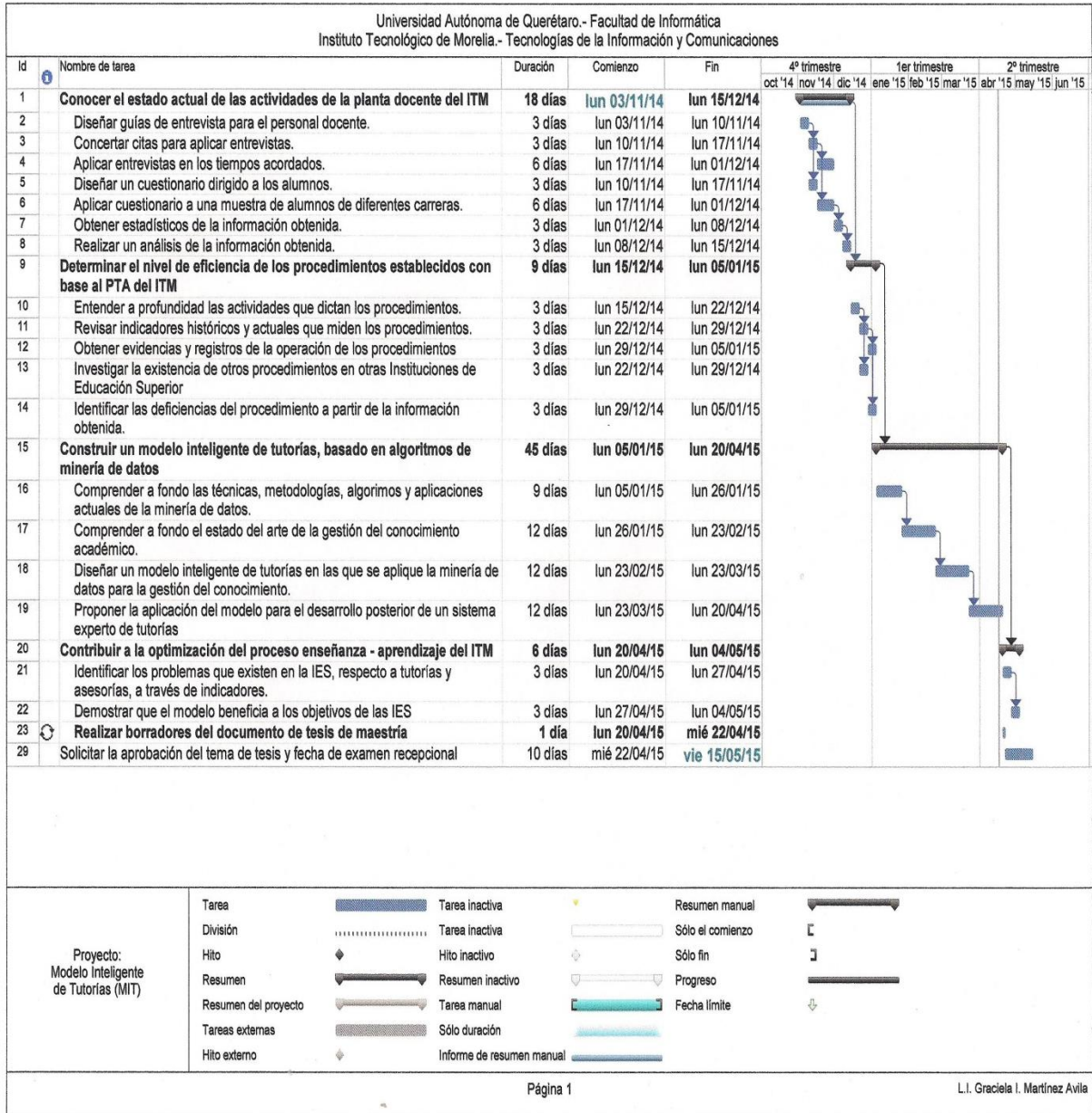
El MIT también refuerza esa interacción del Tutor con el Alumno, a partir de que se debe generar otro almacén de datos con los insumos correspondientes a la persona que ostenta dicha labor, creándose el PTA, PTF, PTS y el PTC y que al igual que el Aspirante, Alumno o Grupo, van cambiando a través del tiempo, adicionando más información y generando atributos adaptados a cada circunstancia.

Las características del MIT son que en primera instancia es cíclico porque se realiza de manera permanente en las mismas unidades de tiempo y es incremental, porque las bases de datos crecen conforme se va retroalimentando con todos los insumos descritos y por último es finito, porque termina cuando la persona alcanza su nivel de Licenciatura, el MIT para ese caso en particular se termina; lo mismo sucede con los grupos generados, sin embargo, en el caso de los perfiles de los Tutores puede durar más tiempo, aunque en algún momento dejará de laborar por alguna razón en la IES y ahí es cuando se finiquita su participación en el Modelo Inteligente de Tutorías.

A futuro se sugiere trabajar y demostrar cada una de las fases del MIT, aplicando técnicas y herramientas de Minería de Datos para que los resultados de esos análisis contribuyan en la toma de decisiones para preservar con calidad la matrícula de las Instituciones de Educación Superior.

# APÉNDICE

## ANEXO A.1 Cronograma de actividades.





Todos los análisis de Minería de Datos, para el Capítulo III. Metodología, fueron realizados en un archivo de Microsoft Excel, utilizando como herramienta estadística XLSTAT, por lo que fue necesario incluir a esta tesis, un disco compacto, en donde se encuentra el archivo en mención, cuyas Hojas se están etiquetadas de manera consecutiva, las cuales se enlistan a continuación:

<b>Anexo</b>	<b>Descripción</b>
<b>A.3</b> R065304302RGI_G	Base de datos original, generada por el EXANI – II de CENEVAL, aplicado en Junio del 2011.
<b>A.4</b> DD lyT	Diccionario de Datos del EXANI – II de CENEVAL.
<b>A.5</b> Desc. Atributos	Descripción de los 134 atributos del EXANI – II de CENEVAL.
<b>A.6</b> Secciones	Descripción de atributos clasificados por secciones del EXANI – II de CENEVAL.
<b>A.7</b> Elecc. Atributos	Elección manual de los atributos significativos, en función del interés por crear el perfil del Aspirante.
<b>A.8</b> Materias	Las materias y profesores titulares del primer cuatrimestre de la carrera de Ingeniería en Tecnologías de la Información y Comunicaciones, en Septiembre del 2011.
<b>A.9</b> BDITIC	Base de Datos depurada con sólo los registros de la carrera de Ingeniería en Tecnologías de la Información y Comunicaciones, a través de filtros.
<b>A.10</b> Alumnos 1T4	Identificación de aspirantes que se encuentran en la Base de Datos del EXANI – II, comparados con las actas de evaluación de la materia de Introducción a las TIC's (1T4).
<b>A.11</b> BD AxED	Base de Datos para el Análisis con Estadística Descriptiva.

<b>Anexo</b>	<b>Descripción</b>
<b>A.12</b> Desc	Resultados de aplicar XLSTAT en la opción de Estadísticos descriptivos.
<b>A.13</b> ACP	Resultados de aplicar XLSTAT en la opción de Análisis de Componentes Principales (ACP).
<b>A.14</b> Cargas Factoriales	Cargas Factoriales de los tres Componentes Principales (F1, F2 y F3).
<b>A.15</b> BD AxAD	Base de Datos para el Análisis con k – means y Análisis Discriminante Lineal.
<b>A.16</b> k – means	Resultados de aplicar XLSTAT en la opción de Clusters de k-means.
<b>A.17</b> AD	Resultados de aplicar XLSTAT en la opción de Análisis Discriminante (AD).
<b>A.18</b> AD 1T1 Calculo	Resultados de aplicar XLSTAT en la opción de Análisis Discriminante (AD), tomando en cuenta como variable dependiente las evaluaciones de la materia de Cálculo.
<b>A.19</b> AD 1T2 F. Prog	Resultados de aplicar XLSTAT en la opción de Análisis Discriminante (AD), tomando en cuenta como variable dependiente las evaluaciones de la materia de Fundamentos de Programación.
<b>A.20</b> AD AD 1T3 MatDisc	Resultados de aplicar XLSTAT en la opción de Análisis Discriminante (AD), tomando en cuenta como variable dependiente las evaluaciones de la materia de Matemáticas Discretas.
<b>A.21</b> AD 1T5 Etica	Resultados de aplicar XLSTAT en la opción de Análisis Discriminante (AD), tomando en cuenta como variable dependiente las evaluaciones de la materia de Ética.
<b>A.22</b> AD 1T6 FundInv	Resultados de aplicar XLSTAT en la opción de Análisis Discriminante (AD), tomando en cuenta como variable dependiente las evaluaciones de la materia de Fundamentos de Investigación.

---

<b>Anexo</b>	<b>Descripción</b>
<b>A.23</b> Tabla ADs	Tabla de todos los atributos que resultaron del Análisis Discriminante tanto de manera Global, como de cada una de las materias.
<b>A.24</b> Tabla ADs Depurada	Tabla de los atributos no repetitivos que resultaron del Análisis Discriminante tanto de manera Global, como de cada una de las materias.
<b>A.25</b> Evaluación	Tabla comparativa de atributos del Análisis de Componentes Principales vs Análisis Discriminante Lineal.

---

## LITERATURA CITADA

- ANUIES. (2000). *Programas Institucionales de Tutoría. Una propuesta de la ANUIES para su organización y funcionamiento en las instituciones de educación superior*. México: Serie Investigaciones.
- CENEVAL. (2014). *Centro Nacional de Evaluación para la Educación Superior, A.C.* Recuperado el 9 de Septiembre de 2014, de <http://www.ceneval.edu.mx/ceneval-web/content.do?page=1702>
- CENEVAL. (2014). *Centro Nacional de Evaluación para la Educación Superior, A.C.* Recuperado el 11 de Noviembre de 2014, de <http://www.ceneval.edu.mx/ceneval-web/content.do?page=1702>
- CIEES. (7 de Octubre de 2013). *Comités Interinstitucionales para la Evaluación de la Educación Superior, A.C.* Recuperado el 11 de Noviembre de 2014, de <http://www.ciees.edu.mx/index.php/que-son-los-ciees/presentacion>
- Clark, P., & Boswell, R. (2000). *Data Mining. Practical Machine Learning Tools and Techniques with Java Implementation*. Morgan Kaufmann Publishers.
- Computación, D. d. (2005). *Programa de Innovación y Desarrollo Departamental 2005 - 2010*. Morelia, Michoacán, México: Instituto Tecnológico de Morelia.
- COPAES. (2014). *Consejo para la Acreditación de la Educación Superior, A. C.* Recuperado el 1 de Noviembre de 2014, de <http://www.copaes.org/>
- DGEST. (Agosto de 2006). *Programa Nacional de Tutoría*. Recuperado el Noviembre de 2012, de Dirección General de Educación Superior Tecnológica.- Coordinación Sectorial de Normatividad Académica:  
<http://148.208.235.3:4000/~tutorias/tutoria/PROGRAMA%20NACIONAL%20DE%20TUTORIAS%20DGEST.pdf>



- 
- DGEST. (2006). *Programa Nacional de Tutorías*. México, D.F.: Coordinación Sectorial de Normatividad Académica de la Dirección General de Educación Superior Tecnológica.
- DGEST. (Octubre de 2011). *Lineamiento para la Operación del Programa de Tutoría Versión 1.0*. Recuperado el 2012 de Octubre de 11, de Planes de Estudio 2009 - 2010:  
[http://www.itcj.edu.mx/doc/Lineamientos/Lineamiento\\_Operacion\\_Tutoria.pdf](http://www.itcj.edu.mx/doc/Lineamientos/Lineamiento_Operacion_Tutoria.pdf)
- DGEST. (2014). *Dirección General de Educación Superior Tecnológicos*. Recuperado el 11 de Noviembre de 2014, de <http://www.dgit.gob.mx/informacion/institutos-tecnologicos-de-mexico>
- Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery: An Overview. *Advances in Knowledge Discovery and Data Mining*, 1-34.
- García, C. D. (2 de Noviembre de 2012). Reconocen al ITM con premio SEP-ANUIES. *Sol de Morelia*, pág. 13.
- Hernández Orallo, J., Ramírez Quintana, M. J., & Ferri Ramírez, C. (2004). *Introducción a la minería de datos*. Madrid, España: Pearson Educación, S.A. de C.V.
- ITM. (2009). *Programa Institucional de Innovación y Desarrollo 2007 - 2012 del Instituto Tecnológico de Morelia*. Morelia, Michoacán, México: Digital Electronic Pro, S. de R.L. de C.V.
- ITM. (2014). *Instituto Tecnológico de Morelia*. Recuperado el 2014 de Noviembre de 11, de <http://itmorelia.edu.mx/>
- Pérez, L. C. (2006). *Data Mining: Soluciones con Enterprise Miner*. España: Alfaomega Grupo Editor.- Colección Alfaomega Ra – Ma.
- Poder Ejecutivo. (1 de Enero de 2013). *Plan Nacional de Desarrollo 2013 - 2014*. Recuperado el 3 de Noviembre de 2014, de <http://pnd.gob.mx/>

- Poder Ejecutivo, d. I. (2007). *Plan Nacional de Desarrollo 2007-2012 de los Estados Unidos Mexicanos*. Ciudad de México: Talleres de Impresión de Estampillas y Valores (TIEV) de la SHCP.
- PSE, P. S.-2. (13 de Noviembre de 2013). *Diario Oficial*. Recuperado el 11 de Noviembre de 2014, de [http://normatecainterna.sep.gob.mx/work/models/normateca/Resource/253/1/images/programa\\_sectorial\\_educacion\\_2013\\_2018.pdf](http://normatecainterna.sep.gob.mx/work/models/normateca/Resource/253/1/images/programa_sectorial_educacion_2013_2018.pdf)
- SNEST. (2012). *Manual del Tutor*. Recuperado el Marzo de 2013, de [www.ittux.edu.mx/module-Descargas-prep\\_hand\\_out-lid-260.html](http://www.ittux.edu.mx/module-Descargas-prep_hand_out-lid-260.html)
- SNIT. (Febrero de 2013). *Subsecretaría de Educación Superior. - Tecnológico Nacional de México*. Recuperado el 13 de Noviembre de 2014, de Normateca de la dirección de docencia: [http://www.tecnm.mx/images/areas/docencia01/documentos/MANUAL\\_DEL\\_TUTOR.pdf](http://www.tecnm.mx/images/areas/docencia01/documentos/MANUAL_DEL_TUTOR.pdf)
- TNM. (2014). *Tecnológico Nacional de México*. Recuperado el 8 de Noviembre de 2014, de Subsecretaría de Educación Superior: <http://www.tecnm.mx/informacion/institutos-tecnologicos-de-mexico>
- UADY. (2006). *Universidad Autónoma de Yucatán*. Recuperado el 21 de Octubre de 2014, de Facultad de Matemáticas: <https://www.google.com.mx/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0CBsQFjAA&url=https%3A%2F%2Fintranet.matematicas.uady.mx%2Fportal%2Ftutorias%2FCAT%2FDOCUMENTOS%2FFMAT%2FCURSO%2520Y%2520MATERIALES%2FMATERIALES%2520PRIMER%2520CURSO.do>
- UAEH. (2013). *Investigación de la Tutoría Académica*. Recuperado el 20 de Noviembre de 2014, de Red de Tutorías. Región Centro - Sur de la ANUIES: [http://www.uaeh.edu.mx/campus/icsa/descargas/libro\\_tutorias\\_red\\_final\\_ver\\_14.pdf](http://www.uaeh.edu.mx/campus/icsa/descargas/libro_tutorias_red_final_ver_14.pdf)
- UDG. (2004). *La tutoría académica y la calidad de la educación*. Recuperado el 3 de Diciembre de 2014, de Universidad de Guadalajara: <http://148.202.105.12/tutoria/pdfc/C2.pdf>

---

UNESCO. (9 de Octubre de 1998). *Declaración Mundial sobre la Educación Superior en el Siglo XXI: Visión y Acción*. Recuperado el 1 de Abril de 2013, de [http://www.unesco.org/education/educprog/wche/declaration\\_spa.htm#marco](http://www.unesco.org/education/educprog/wche/declaration_spa.htm#marco)

XLSTAT. (1995 - 2014). *XLSTAT*. Recuperado el 29 de Octubre de 2014, de <http://www.xlstat.com/es/productos-soluciones.html>