

Tipo de artículo: Artículo original  
Temática: Enseñanza de las Ciencias Informáticas  
Recibido: 08/05/2015 | Aceptado: 20/05/2015

## El análisis estadístico aplicado a la gestión de la enseñanza para la toma de decisiones

### *The statistical analysis applied to teaching management for decision-making*

Yusnier Reyes Dixson <sup>1\*</sup>, Lissette Nuñez Maturel <sup>2</sup>

- <sup>1</sup> Universidad de las Ciencias Informáticas. Carretera a San Antonio de los Baños, Km. 2 ½. Torrens, La Lisa. La Habana. [ydixson@uci.cu](mailto:ydixson@uci.cu)
- <sup>2</sup> Centro Nacional de Genética Médica. Ave 31 no. 3102 Esquina 146, Reparto Cubanacán. Playa. [lissette@cngen.sld.cu](mailto:lissette@cngen.sld.cu)

\* Autor para correspondencia: [ydixson@uci.cu](mailto:ydixson@uci.cu)

---

#### Resumen

El desarrollo actual de las tecnologías para el análisis de datos, facilita a las organizaciones comprender en mayor medida sus procesos y actuar de manera más efectiva sobre los mismos. Las universidades son organizaciones donde es viable utilizar estas tecnologías para aumentar la eficiencia y calidad en el proceso de formación. En el trabajo se mostraron los resultados logrados al concebir y aplicar una propuesta procedimental para desarrollar el análisis de los datos académicos de los estudiantes. Se propuso la utilización de técnicas del análisis multivariado, para obtener información relevante sobre las características y factores que influyeron en el rendimiento académico de los estudiantes del primer año del curso académico 2011-2012 de la Facultad 4 de la Universidad de las Ciencias Informáticas. Mediante la aplicación de las técnicas propuestas, se obtuvo conocimiento concreto de qué asignaturas estuvieron interrelacionadas y qué elementos no influyeron en el rendimiento de los estudiantes en este curso, información relevante para la planificación de estudiantes y profesores. Luego de aplicar la propuesta se obtuvo un modelo de regresión lineal que puede predecir en alguna medida el rendimiento académico de los estudiantes en próximos cursos y de esta forma poder perfeccionar la gestión educativa, para lograr sea más oportuna y confiable; elementos que mucho tiene que ver con los objetivos que se persiguen en nuestras instituciones universitarias. Los resultados obtenidos constituyen una muestra de lo que se puede hacer con los volúmenes de datos que manejamos en nuestras instituciones educativas.

**Palabras clave:** análisis estadístico multivariado, educación superior, gestión del conocimiento, rendimiento académico

### **Abstract**

*The current development of technology for data analysis, enables organizations to further understand their processes and to act more effectively on them. Universities are organizations where it is feasible to use this technology as the central process that handles it is necessary to develop efficiently. In the work presented the results achieved to devise and implement a procedural proposal, based on the use of multivariate analysis techniques to obtain relevant information on the characteristics and factors that influence the academic performance of freshmen is the academic year 2011-2012 Faculty 4 University of Information Science. By applying the proposed techniques, concrete knowledge of what subjects you get are interrelated information that is relevant to the planning of the student and an important element to be considered by teachers to keep informed students as well as a way to predict academic performance of students in future courses and in this way to improve education management to be more efficient, timely and reliable elements that a lot has to do with the objectives pursued in our universities. The results obtained are an example of what to do with the large volumes of data we use in our educational institutions.*

**Keywords:** multivariate statistical analysis, higher education, knowledge management, academic performance

---

## **Introducción**

El proceso de formación en las universidades es un proceso complejo, debido a la gran variabilidad en las características de los estudiantes, que son el objeto a transformar, y a la variedad de condiciones que confluyen en el proceso para desarrollar esta transformación (Rodríguez, et al. 2010). En el caso de la formación de ingenieros en la Universidad de las Ciencias Informáticas (UCI), un primer elemento que condiciona el nivel de complejidad es la variabilidad en los conocimientos y habilidades de los estudiantes al ingresar a la universidad y adicionalmente las cualidades de los profesores, quienes tienen niveles académicos y de experiencia que en su mayoría no poseen niveles académicos y de experiencia elevados debido principalmente al porcentaje considerable de los profesores que son jóvenes y una pequeña parte acumula ya muchos años de trabajo. Otro elemento lo constituye la atención a otras tareas no directamente relacionadas con la formación (la investigación, la producción y las responsabilidades administrativas), las cuales incrementan el volumen y la complejidad del trabajo de los profesores y directivos.

Basado en estas condiciones, los directivos del proceso docente-educativo deben ser capaces de utilizar la información que aparece durante su desarrollo; integrarla, formular esquemas para la acción y ser capaces de reunir el máximo de certidumbres para confrontar la incertidumbre que se genera durante el desarrollo del proceso docente universitario. Por lo que debe crearse una inteligencia organizacional que posibilite la identificación, captación y procesamiento de

datos adecuados para la obtención de un modelo del proceso que facilite la toma de decisiones y la concepción de estrategias orientadoras (Rodríguez, et al. 2010; Monteagudo, et al. 2012). Se deben aprovechar las técnicas estadísticas multivariadas y otras para guiar el trabajo del profesor, brindarle informaciones más completas y oportunas sobre sus estudiantes, que le faciliten orientarlos de acuerdo a sus características específicas, y controlar todo el proceso de aprendizaje teniendo en cuenta esa variabilidad ya descrita. La formación del ingeniero en ciencias informáticas debe tener un enfoque de procesos ya que debe centrarse en el estudiante, estudiar cómo se produce la transformación de los conocimientos, habilidades y otras características de entrada, en características de salida; y convertir los resultados de esos estudios en conocimientos útiles para dirigir el proceso de aprendizaje en condiciones de masividad (Davenport, et al 2000; Márquez, et al. 2012). En las universidades debe crearse una inteligencia organizacional, que gestione un aprendizaje organizacional basado en la utilización de las técnicas modernas de análisis de datos, para que todos aprendan del comportamiento de los procesos que dirigen, puedan prever sus reacciones y dominarlos (Rodríguez and Heredia, 2009).

La situación descrita anteriormente, unida a la ventaja que representa para estos fines los sistemas de información y en especial el sistema de gestión universitaria con que cuenta la universidad, herramienta que posibilita registrar un gran volumen de datos viables de ser analizados; promueve el despliegue de una estrategia de análisis de datos para el descubrimiento y la gestión del conocimiento asociado al proceso de enseñanza, que contribuya a transformar los trabajos de dirección y a sugerir nuevos enfoques y acciones de los actores directos del proceso, para mejorar la productividad y eficacia de estudiantes y profesores. Este trabajo va encaminado a aprovechar con esos fines, la oportunidad que representa la exploración del gran conjunto de datos referentes al proceso docente, que semestre tras semestre se acumulan en las facultades de la UCI, considerando la efectividad mostrada por el despliegue de este tipo de estrategias para la obtención de conocimientos útiles como apoyo a la toma de decisiones (Luan, 2002).

Los objetivos de este trabajo consisten en mostrar los resultados alcanzados en la concepción de un procedimiento basado en una propuesta metodológica para desarrollar el análisis de datos como apoyo a la gestión de la enseñanza y discutir algunas ventajas que representa la información obtenida para la toma de decisiones en el ámbito académico.

## **Materiales y métodos**

Para utilizar los datos como base de creación de conocimiento, se utilizó un procedimiento que fue adaptado a partir de una propuesta metodológica cuyo objetivo es emplear los datos para generar información en apoyo a las decisiones (Davenport, 1990; Rodríguez and Heredia, 2009; Martínez-Noriegas, 2012). La propuesta constituye una primera

aproximación que se proyecta con el fin de utilizar los datos sobre el proceso docente en la gestión del mismo. En la adaptación se consideran las características de los problemas a resolver en la enseñanza de la Ingeniería en Ciencias Informáticas de la UCI. Se pretende que la aplicación continua de las ideas que se presentan, permita validar su utilidad en la Facultad 4 de la UCI, para posteriormente validar sus beneficios en otras facultades de la universidad, e incluso en otros contextos ajenos a ésta. El procedimiento es genérico, pues no se especifica con detalle el "cómo" ejecutar los diferentes pasos que se presentan. La aplicación repetida también contribuirá a incrementar la granularidad en este sentido.

En la concepción de los aspectos incluidos dentro de cada una de estas etapas, se tuvo en cuenta la propuesta realizada por (Davenport, 2000), referente a diversos factores críticos a considerar al desarrollar proyectos para convertir datos de procesos en conocimiento. Dentro de estas cuestiones destacan: garantizar las habilidades necesarias de las personas para el descubrimiento y la gestión del conocimiento, enmarcar el programa de mejora dentro de un contexto estratégico en la organización, atender las cuestiones culturales que puedan entorpecer el desarrollo del proyecto, gestionar los cambios en los roles organizativos de las personas involucradas y garantizar una correcta integración de la tecnología utilizada en todas las etapas de la gestión del conocimiento.

El procedimiento se estructuró en 4 etapas (Davenport, 1990; Rodríguez and Heredia, 2009; Martínez-Noriegas, 2012):

### **1. Caracterización del proceso a perfeccionar.**

La mejora de la gestión de un proceso está precedida por el conocimiento de las principales características de éste. En este caso, la caracterización del proceso docente es un primer paso para encontrar áreas de mejora que puedan derivarse del descubrimiento de conocimiento en los datos. Se propone incluir en esta caracterización, cuáles son los resultados del proceso, los participantes, sus responsabilidades y los recursos involucrados.

### **2. Establecer objetivos del proyecto de rediseño.**

Se recomienda la definición de los objetivos estratégicos del proceso, y en función de éstos, la definición de los objetivos del programa de mejora. Es necesario que los objetivos que se persiguen con el rediseño tributen a resolver problemas importantes en la organización o al menos, que ayuden a que ésta consiga ventajas estratégicas. Lograr la articulación de ambos objetivos incrementará el apoyo y el compromiso con el programa de rediseño.

### **3. Identificación de las actividades de gestión a rediseñar.**

Para obtener los resultados en el proceso docente, deben realizarse una serie de acciones de coordinación y gestión, como por ejemplo: la confección de los grupos académicos, la planificación del horario docente, la planificación y

ejecución de las tareas productivas, la confección y ejecución del proyecto educativo de cada brigada, del plan de actividades metodológicas del año, y además, la gestión que requiere la propia impartición de las clases. El objetivo que persigue esta etapa es identificar y caracterizar estas acciones. En este paso pueden ser identificadas una gran variedad de acciones y deberá seleccionarse sobre cuáles se centrará el rediseño. La caracterización debe comenzar por una descripción de la forma en que son realizadas las actividades relacionadas con el proceso docente. Se recomienda definir en la actividad las situaciones de decisión. Lo ideal es lograr que en cada acción se hagan explícitas las decisiones que se toman y que se conciben los resultados a través de las decisiones a tomar. Adicionalmente, deben identificarse las "necesidades de información" que tiene la actividad a mejorar, fundamentalmente la información que apoya las decisiones identificadas.

#### **4. Recopilación y análisis de datos (descubrimiento del conocimiento)**

Esta etapa tiene como objetivo aplicar técnicas de análisis de datos para descubrir regularidades, patrones y modelos en los datos que posibiliten incidir positivamente en la forma en que se desarrolla la gestión del proceso docente. Con esta etapa se pretende satisfacer las necesidades de información de la actividad sujeta a mejora, ya que precisamente las asociaciones y modelos encontrados representan la información deseada.

En este paso se define qué objetivo específico se persigue con el análisis, de forma que al ejecutarlo e interpretar los resultados, se obtenga la información que se desea; es decir, luego de identificadas las necesidades de información, es necesario analizar si parte de esta información, puede obtenerse a partir del análisis de datos. Posteriormente deben seleccionarse y recopilar los datos que permitan llegar a la información que se quiere. Esto implica seleccionar los elementos sobre los que se harán las mediciones y/o registro de las variables, y determinar las variables necesarias. Además, se deben seleccionar las fuentes de dónde se obtendrán los datos, y en el caso que los datos asociados con alguna variable no estén disponibles, corresponde especificar cómo se obtendrán. Seguidamente se deben seleccionar las técnicas y métodos matemáticos a utilizar, para lo cual será necesario analizar cada objetivo y definir si su logro requiere aplicar una técnica descriptiva, de asociación entre atributos nominales, de agrupamiento o clustering entre elementos, o de predicción (Hair, 1999; Delavari and Reza, 2008; Heredia, et al. 2012). En dependencia de la técnica, se establecerá el tamaño de muestra necesario y los supuestos básicos a comprobar.

#### **Resultados obtenidos con la aplicación del procedimiento propuesto**

Se muestra paso a paso la aplicación del procedimiento para establecer un análisis de los resultados académicos de los estudiantes del primer año de la carrera de ingeniería en ciencias informáticas de la facultad 4 en el curso 2011-2012.

### 1. *Caracterización del proceso a perfeccionar.*

En el proceso, las entradas son los estudiantes de primer año, los cuales poseen conocimientos, habilidades y valores que deben modificarse por la acción conjunta de los recursos del sistema docente educativo. Los recursos más importantes son los humanos, constituidos por los profesores, directivos (miembros de las Vicerrectorías, Decano y Vicedecanos de las Facultades) y los trabajadores de servicio. Además, los recursos materiales, donde se destacan: las instalaciones, la infraestructura tecnológica y la base material de estudio. Para que la formación del alumno sea la mejor posible, los recursos deben combinarse de la forma más óptima.

La persona encargada de lograr que las metas del proceso se cumplan, es el Vicedecano Docente, apoyado por los Jefes de Año, que actúan como responsables en cada una de las etapas (años de estudio). Los participantes trabajan en varios procesos simultáneamente: un mismo profesor trabaja en procesos de formación, y además en proyectos de investigación y productivos, en maestrías, doctorados, entre otras funciones. El proceso es interfuncional, porque en él participan profesores provenientes de los 4 departamentos docentes de la Facultad. Se vincula con varios procesos, destacando los de investigación, producción y postgrado. La responsabilidad por el resultado es colectiva, las acciones de unos interactúan con las de otros, incrementando o disminuyendo su efecto. Es un proceso donde priman las decisiones no estructuradas, basadas en el conocimiento de los actores, incluyendo al estudiante, que toma sus propias decisiones sobre la atención que le va a dar a las orientaciones recibidas. Los resultados de cada acción de los profesores dependen grandemente de las características propias del estudiante, que son muy variadas, de sus motivaciones internas y de sus habilidades para aprender; en fin, de las posibilidades de desarrollo alcanzado por la persona en el momento del aprendizaje.

### 2. *Establecer objetivos del proyecto de rediseño.*

Esta investigación será el primer paso para contribuir al logro a corto y mediano plazo de los siguientes objetivos (Heredía and Rodríguez, 2010; Heredia, 2010):

- Contribuir a la satisfacción de las necesidades y expectativas de los estudiantes en su formación integral.
- Alcanzar niveles satisfactorios de calidad con los recursos disponibles. Significa que cada estudiante rinda el máximo posible para él, que logre los mejores resultados que sea capaz. Para ello es importante que quienes lo guíen tengan una evaluación razonable de sus posibilidades y lo orienten acerca de cómo aprovecharlas y estimularlas.
- Mantener costos sociales factibles. Éste debe ser una consecuencia del objetivo anterior, al mejorar no sólo la eficiencia académica, sino al contribuir a que aquellos que logren egresar, lo hagan con la mejor preparación posible.

- Aprovechar las Tecnologías de la Información y las Comunicaciones (TICs) como agentes de aprendizaje, es decir, que actúen como mediadoras del conocimiento para el profesor al tomar sus decisiones.
- Contribuir al desarrollo de la inteligencia organizacional en la facultad, mediante la concepción de un proceso de toma de decisiones que tenga como base la utilización de los datos que se conservan sobre el proceso docente y que permita la definición de modelos que ayuden a evaluar integralmente los resultados docentes.

### *3. Identificación de las actividades de gestión a rediseñar.*

La gestión que requiere la impartición de las clases ha sido el centro de mejora de los estudios realizados hasta el momento. El profesor, como actor principal de la misma, se auxilia de la didáctica para dirigir científicamente esta actividad mediante la consideración de una serie categorías didácticas. Estas son: los objetivos, el contenido, los métodos y procedimientos de enseñanza, además de los medios de enseñanza, las formas de organización y de evaluación a utilizar (Rodríguez and Heredia, 2009). El profesor debe planificar el contenido y las evaluaciones de la asignatura, además, imparte las clases en sí, que involucra el intercambio de conocimientos entre profesor y el estudiante, y la orientación del estudio independiente. También gestiona el control del aprendizaje, lo cual incluye la ejecución de las evaluaciones planificadas y las actividades de control sistemático.

### *4. Recopilación y análisis de datos (descubrimiento del conocimiento)*

Son muchas las variables que afectan al aprendizaje y al rendimiento académico de los estudiantes. Hay variables de tipo personal, familiar y de contexto socioeconómico. A toda institución universitaria y en particular a todo profesor, debe interesarle conocer cuáles son los factores determinantes y en qué medida se puede incidir para hacer más eficaz el proceso de enseñanza-aprendizaje (Porcel, et al. 2010; Hormigo and Jordi, 2014). En la UCI los estudiantes de 1er año son los que alcanzan los resultados más deficientes en cuanto a rendimiento académico, cuestión esta que es común en la mayoría de las universidades pero en la UCI adquiere mayor relevancia debido a que solo existe una carrera, por lo que sería interesante realizar estudios sobre los factores que influyen directamente sobre el rendimiento académico de los estudiantes y de esta forma tener información relevante para el proceso de toma de decisiones que se realiza en la universidad .

Para efectuar la recopilación y el análisis de información se tomará en cuenta Rendimiento Académico a través del Promedio General como variable dependiente y como variables predictoras o independientes notas las asignaturas del primer y segundo semestre, donde se estableció una codificación que se muestra en la tabla 1 y adicionalmente se seleccionaron algunos datos de los diagnósticos aplicados a los estudiantes del 1er año cuando ingresan a la universidad. Se pretende obtener un modelo lineal de regresión múltiple que sea capaz de predecir en alguna medida el

estado académico de los estudiantes del año para poder tener un criterio preventivo sobre el tema (Ballesteros and Sánchez-Guzmán, 2012). Para ello se estableció:

*Variable dependiente:* Rendimiento Académico definido como el Promedio General (PG) obtenido por los estudiantes en el curso académico.

*Variables predictoras:* Evaluaciones de las asignaturas con mayores problemas en el año definidas por el colectivo de año y los resultados obtenidos en cursos anteriores estas son: Introducción a la Programación, Programación I, Matemática Discreta I y Matemática I. Además otras variables obtenidas de los datos almacenados en los proyectos educativos de las brigadas: estado de salud, situación económica familiar y estado civil.

Tabla 1. Codificación de asignaturas del primer y segundo semestre.

Cód.	Semestre	Nombre
EF1	Primero	Educación Física I
II	Primero	Introducción a la Informática
IP	Primero	Introducción a la Programación
M1	Primero	Matemática I
MD1	Primero	Matemática Discreta I
SN	Primero	Seguridad Nacional
AL	Segundo	Algebra Lineal
DN	Segundo	Defensa Nacional
EF2	Segundo	Educación Física II
HC	Segundo	Historia de Cuba
M2	Segundo	Matemática II
MD2	Segundo	Matemática Discreta II
PP1	Segundo	Práctica Profesional I
P1	Segundo	Programación I

Se realizará este análisis para las asignaturas del primer año de la carrera de Ingeniería en Ciencias Informáticas de la UCI. Los datos con los que se cuentan son las evaluaciones finales de las 14 asignaturas del curso académico 2011-2012 y los datos de los proyectos educativos de los 4 grupos académicos existentes en ese curso.

Los datos para realizar el análisis se resumieron en 73 estudiantes correspondientes a 4 grupos de clases. En la figura 1 se muestra una sección de los datos. El estado civil, situación económica familiar y el estado de salud general del estudiante se encuentran bien especificados en los proyectos educativos de cada brigada.

Grupo	nombre	EF1	II	IP	M1	MD1	SN	AL	DN	EF2	HC	M2	MD2	PP1	P1
4204	Ricardo Blanco	5	5	5	5	5	5	5	5	5	5	5	4	5	5
4201	Jose Alberto Fonseca	5	5	5	4	4	5	5	5	5	5	5	4	5	5
4204	Gustavo García	5	4	5	5	4	4	5	4	5	5	5	5	5	5
4203	Cesar Andrés Gonzáles Rodríguez	5	5	4	4	5	5	4	5	5	4	4	4	5	5
4201	gjimenez	5	5	5	3	5	4	4	5	5	5	4	3	5	5
4202	Ulisis Rodríguez Nespral	5	4	5	3	4	5	4	5	5	5	5	4	4	4
4201	Tania Rodríguez	5	5	3	3	4	5	5	5	5	5	4	3	5	4
4204	Julio cesar Pérez	5	5	5	4	4	5	4	4	5	4	4	4	4	4
4204	Vicente Torres	5	5	5	3	3	5	5	4	5	5	4	3	4	5
4203	Adrian Quesada Gómez	5	5	5	3	3	5	4	5	5	5	3	3	5	4
4201	Rachelys Cruz	5	5	3	3	3	5	5	5	5	4	4	5	5	3
4202	Hugo Cartaya Pardo	5	4	5	3	4	5	3	4	5	5	4	4	4	4
4203	Angel Taboada Martin	5	5	5	3	4	5	3	4	5	5	3	3	5	4
4201	Daimel Noda	5	5	5	3	3	5	5	4	5	4	3	4	5	3
4202	Omar Almaguer Guerra	5	4	5	4	5	5	3	4	4	5	4	4	3	3
4201	Frank lorenzo Morfa	5	5	4	3	5	4	3	5	5	5	4	3	4	3
4202	Luis Manuel Lugo Marquez	5	5	5	3	3	5	3	5	4	5	3	3	5	3
4203	Frank Rosendo	5	4	5	2	3	4	3	5	5	5	4	4	4	4
4203	Laura Hernández Amador	5	5	4	4	3	4	5	4	5	4	4	3	5	2
4203	Osvel Alvarez Jacomino	4	4	5	3	4	4	5	4	4	4	3	4	4	5
4201	Franly Hernández	5	5	4	3	3	5	5	4	5	4	3	4	4	3
4204	Jorge Antonio Monedero	5	4	5	3	3	4	4	5	5	4	4	4	4	3

Figura 1. Sección de los datos iniciales correspondientes a un grupo académico

Para el estado civil (EC) se tendrán tres categorías: casado, soltero con pareja y soltero.

Para el estado de salud general (ES) se tendrán dos estados: Bien y Regular.

Para definir la situación económica familiar (SEF) se tendrán cuatro categorías: Muy Buena, Buena, Regular y Mala. Se realizó una transformación de los datos y exploración de estos mediante la matriz de dispersión con el objetivo de verificar el supuesto de linealidad que deben cumplir para poder realizar un análisis. La transformación se realizó sustituyendo la nota del estudiante por un valor de 0 a 73 (cantidad de estudiantes en el análisis), en dependencia de la importancia que tenía su resultado en una asignatura determinada (Rodríguez and Heredia, 2009; Heredia, 2010). Por ejemplo si en la asignatura de EF1 los primeros 66 estudiantes tienen 5 puntos, se suman los valores correspondientes desde el 73 hasta el 8( $73 + 72 + 71 + 70 + \dots + 8$ ) y se dividen entre las cantidad de estudiantes que tienen 5 puntos (66 estudiantes). En la figura 2 se muestra una sección de los datos donde se muestra los datos después de ser transformados.

C1	C2	C3	C4	C5	C6	C7	C8	C9
Estudiante	IP	M1	MD1	P1	SEF	ES	EC	PG
1	63,5	72,5	70,5	69,5	2	1	1	60,6429
3	63,5	69,0	62,5	69,5	3	2	2	59,8214
5	63,5	72,5	62,5	69,5	3	2	2	53,6429
6	44,0	69,0	70,5	69,5	3	2	1	55,6071
7	63,5	36,5	70,5	69,5	3	2	1	52,2857
8	63,5	36,5	62,5	60,0	4	2	1	51,1429
9	21,0	36,5	62,5	60,0	2	2	1	50,7500
10	63,5	69,0	62,5	60,0	2	2	1	50,6429
11	63,5	36,5	30,5	69,5	3	1	1	47,8214
13	63,5	36,5	30,5	60,0	2	2	1	48,0357
14	21,0	36,5	30,5	32,5	3	2	1	47,1786
15	63,5	36,5	62,5	60,0	3	2	1	46,2500
16	63,5	36,5	62,5	60,0	3	2	1	44,8929
17	63,5	36,5	30,5	32,5	3	2	1	45,0357
18	63,5	69,0	70,5	32,5	4	2	1	43,2500
20	44,0	36,5	70,5	32,5	3	2	3	43,5714
21	63,5	36,5	30,5	32,5	3	2	1	41,7143
22	63,5	5,5	30,5	60,0	3	2	2	41,7143
25	44,0	69,0	30,5	8,0	3	1	1	41,8929
27	63,5	36,5	62,5	69,5	3	2	1	38,1429
28	44,0	36,5	30,5	32,5	3	1	1	41,5714

Figura 2. Sección de los datos iniciales ya transformados

Con los datos y las variables declaradas anteriormente se diseñó a partir de las técnicas de regresión un modelo general con el objetivo de identificar las dependencias entre las variables (Barahona and Aliaga, 2013), se utilizó el software estadístico Minitab 16 donde el reporte de regresión general mostró la siguiente información:

*Análisis de regresión general: PG versus IP; MD1; P1; M1; SEF; ES; EC*

*Ecuación de regresión: PG = 27,842 + 0,0954521 IP + 0,208964 MD1 + 0,0701705 P1 + 0,0612271 M1 + 0,185585 SEF - 2,88301 ES - 1,25084 EC*

Coefficientes

EE del

Término	Coef	coef.	T	P
Constante	27,8420	6,35876	4,37853	0,000
<b>IP</b>	<b>0,0955</b>	<b>0,04387</b>	<b>2,17578</b>	<b>0,033</b>

<b>MD1</b>	<b>0,2090</b>	<b>0,05680</b>	<b>3,67901</b>	<b>0,000</b>
P1	0,0702	0,04910	1,42915	0,158
M1	0,0612	0,05851	1,04639	0,299
SEF	0,1856	1,11205	0,16689	0,868
ES	-2,8830	2,31411	-1,24584	0,217
EC	-1,2508	1,39921	-0,89397	0,375

*Resumen del modelo*

S = 5,43410 R-cuad. = 57,29% **R-cuad.(ajustado) = 52 ,70%**

PRESS = 2496,71 **R-cuad.(pred.) = 44 ,45%**

Del reporte generado por el software estadístico se puede concluir que solo de las variables independientes IP y MD1 se puede tener evidencia suficiente de que serán diferentes de cero y que por tanto aportan algún tipo de información al modelo propuesto (Johnson, 1998; Johnson and Wichern, 1998). Por otro lado el modelo general tiene una media de calidad de ajuste del 52,70% que aunque no es alta puede tenerse en cuenta ya que el propósito de este trabajo no es ofrecer un modelo óptimo para el pronóstico del rendimiento académico en el primer año de la Facultad 4 de la UCI, pero si obtener un modelo del cual se tengan garantías de que es parcialmente predictor de la variable dependiente ya mencionada y que sirva dicho modelo como elemento a seguir perfeccionando haciendo estudios más profundos de las variables y quizás de incorporar nuevas variables en posteriores análisis.

Con las variables IP y MD1 que demostraron aportar información al modelo general se procedió a obtener un nuevo modelo que evidencia que el rendimiento académico de los estudiantes del primer año está condicionado por los resultados académicos en las asignaturas de Matemática Discreta 1 e Introducción a la Programación. Teniendo en cuenta solo estas variables se obtienen un nuevo modelo:

*Análisis de regresión general: PG versus IP; MD1*

*Ecuación de regresión:*  $PG = 23,2261 + 0,135027 IP + 0,253803 MD1$

*Coefficientes*

	Coef	coef.	T	P
Constante	23,2261	1,81247	12,8146	0,000
IP	0,1350	0,03897	3,4645	<b>0,001</b>
MD1	0,2538	0,04785	5,3041	<b>0,000</b>

*Resumen del modelo*

S = 5,50105 R-cuad. = 52,87% R-cuad.(ajustado) = 51 ,52%

PRESS = 2341,79 **R-cuad.(pred.) = 47 ,90%**

Ecuación de este modelo sería:  $Y=23,2261 + 0,135027 IP + 0,253803 MD1$

Este nuevo modelo aumenta la calidad predictiva a un 47.90% y se evidencia que las variables MD1 e IP son las que aportan la información relevante para el modelo de regresión lineal que permitirán predecir los resultados desde el punto de vista del rendimiento académico de los estudiantes del primer año de la facultad en análisis. A pesar de que el resultado obtenido por el modelo matemático es algo que es conocido por los profesores del año, evidencia matemáticamente la estrecha relación que existe entre las dos asignaturas y la gran influencia que tienen en el desempeño académico de los estudiantes.

Para facilitar la creación y uso del conocimiento obtenido en el análisis de datos, se puede contemplar la creación de un grupo de análisis de procesos el cual posea como un objetivo esencial la realización de actividades profesionales de estudiantes de pregrado y posgrado, de modo que siempre esté guiado y supervisado por un profesional de mayor nivel científico y académico en el área de análisis de datos. Este grupo no solo debería de encargarse del procesamiento de los datos, sino que una parte importante de las tareas que deberían realizar serán a través de la participación en su interpretación y en la formalización de esas interpretaciones que se deriven de los análisis a realizar. La utilización del conocimiento obtenido en la ejecución del proceso de formación, está precedida por acciones que propicien que cada participante comprenda y comparta los objetivos y métodos que se emplearán.

El etiquetado y descripción de las interpretaciones derivadas del análisis; deben ser realizadas bajo la dirección de los respectivos jefes de año y por los especialistas del grupo, con el asesoramiento de los profesores con más experiencias de cada año académico. Por las características de la facultad, en todos los años se puede contar con profesores con amplia experiencia en la docencia del año y con conocimientos matemáticos suficientes para la tarea. Estos resultados se llevan a los colectivos de año y de disciplina, a cuyos integrantes corresponde explicar y enriquecer el análisis de causas subyacentes, convirtiéndolo en conocimiento profesional compartido.

Para lograr el éxito del análisis de datos como vía de descubrimiento de conocimiento útil, los directivos del proceso de formación, requieren desarrollar sus habilidades de persuasión, desplegando todo el atractivo del trabajo en el equipo y minimizando sus inconvenientes. El conocimiento descubierto podrá usarse para simular nuevos entornos y evaluar los resultados más probables; la confrontación posterior con la realidad permitirá el ajuste progresivo de los modelos matemáticos encontrados y de los métodos como base para las decisiones.

## **Conclusiones**

Se han identificado las dependencias entre los resultados de grupos de asignaturas con el rendimiento académico de los estudiantes valorado como su promedio general de un curso. Para lograr una correcta creación del conocimiento a

partir de los datos y para garantizar la oportuna utilización del mismo, se recomienda un procedimiento, el cual brinda la posibilidad de planificar y ejecutar las acciones educativas con mayor atención en las diferencias individuales de los alumnos; además de dotar la gestión del proceso de formación de un carácter más preventivo. En principio los resultados del análisis estadístico pueden ser empleados por el profesor al impartir la clase, además por el jefe de colectivo de año. No se pretende invadir la didáctica, sino apoyarla con información valiosa extraída de los datos.

Es necesario tener en cuenta que la propuesta se aplica solamente en la Facultad 4 de Ingeniería en Ciencias Informáticas de las UCI, de manera que es importante interpretar con cautela los resultados, especialmente los resultados estadísticos, evitando generalizaciones apresuradas que no sean válidas.

El análisis del rendimiento académico que se realiza está centrado en la evaluación de los conocimientos y habilidades de los estudiantes según sus calificaciones, pero realmente la caracterización del aprendizaje abarca, además, la valoración del desarrollo del pensamiento y la formación de valores lograda por el estudiante. Por lo tanto, deben conjugarse sus resultados con otras evaluaciones del proceso docente para poder juzgar la calidad del aprendizaje. Como se ha comentado, ésta es una primera aplicación de la propuesta, pero en futuras aplicaciones será necesario explicitarla aún más, especificando con más detalle qué hacer en cada paso.

## Referencias

- BARAHONA, P. and ALIAGA, V. Variables predictoras del rendimiento académico de los alumnos de primer año de las carreras de Humanidades de la Universidad de Atacama. *Revista Internacional de Investigación en Ciencias Sociales*, 2013, 9(2): p. 207-220
- BALLESTEROS, A. and SÁNCHEZ-GUZMÁN, D. Minería de datos aplicada en la gestión educativa. En: XVII Simposio de Tecnología Avanzada. MEXICO D. F: CICATA-IPN, 2012, p. 29
- DAVENPORT, T y SHORT, J. 1990. The New Industrial Engineering: Information Technology and Business Process Redesign. s.l.: Sloan Management Review, 1990. págs. 11-27. Vol. 31. ISSN 1532-9194.
- DAVENPORT, T., y otros. 2000. Data to Knowledge to Results: Building an Analytic. [En línea] 2000. [Citado el: 14 de noviembre de 2013.] <http://www.accenture.com/isc>
- DELAVARI, N. and REZA, M. 2008. Data Mining Application in Higher Learning Institutions. s.l. : Informatics in Education, 2008. págs. 31- 54. ISSN 1648-5831.
- HAIR, J., y otros. 1999. Análisis Multivariante. Madrid: Prentice Hall Iberia, 1999. ISBN 84-8322-035-0.

HEREDIA, J. 2010. Análisis de datos en apoyo a la productividad en el proceso de formación de ingenieros.

La Habana: Instituto Superior Politécnico José Antonio Echeverría, CUJAE, 2010.

HEREDIA, J. y RODRÍGUEZ, A. 2010. Modelo basado en el análisis de datos como apoyo a la gestión de la enseñanza. La Habana: s.n., 2010. ISBN 978-959-261-317-1.

HEREDIA, J., RODRÍGUEZ, A., & VILALTA, J. A. El análisis de datos en apoyo a la gestión de la enseñanza en la carrera Ingeniería Industrial. Ingeniería Industrial, 2012, 33(1): p. 19–30.

HORMIGO, I., and JORDI C. Uso de analítica para dar soporte a la toma de decisiones docentes. En: Jornadas sobre la enseñanza universitaria de la informática. Aplicación de las TIC al proceso de enseñanza-aprendizaje. Oviedo: Universidad de Oviedo, 2014, p. 83–90.

JOHNSON, D. 1998. Métodos multivariados aplicados al análisis de datos. México, D. F: InternationalThomson Editores, 1998. ISBN 968-7529-90-3.

JOHNSON, R. and WICHERN, D. 1998. Applied Multivariate Statistical Analysis. New Jersey: Prentice Hall, 1998. ISBN 0-13-121973-1.

LUAN, J. 2002. Data Mining and Knowledge Management in Higher Education -Potential Applications. Annual Forum for. [En línea] 2002. [Citado el: 21 de noviembre de 2013.] <http://www.air2002.org>

MARTÍNEZ-NORIEGAS, H. Metodología de procesamiento de los datos educacionales para facilitar el aumento de la promoción en el primer año de la UCI. Tesis de Maestría Tecnologías para el Apoyo a la Toma de Decisiones, Instituto Superior Politécnico José Antonio Echeverría, La Habana, 2012.

MÁRQUEZ, C., ROMERO, C. and VENTURA, S. Predicción del fracaso escolar mediante técnicas de minería de datos. IEEE-RITA, 2012, 7(3): p. 109-117

MONTEAGUDO, Y., HERNÁNDEZ, R., & GONZÁLEZ, D. Y. Análisis multivariado del aprendizaje de ingeniería informática en la municipalización. Investigación de Operaciones, 2012, 33 (3): p. 226–238.

PORCEL, E., DAPOZO, G., and VICTORIA, M. Predicción del rendimiento académico de alumnos de primer año de la facena (UNNE) en función de su caracterización socioeducativa. Revista Electrónica de Investigación Educativa, 2010, 12 (2): 1-21

RODRÍGUEZ, A. and HEREDIA, J. 2009. Rediseño de procesos de gestión de la enseñanza basado en tecnologías informativas. [En línea] 2009. [Citado el: 20 de noviembre de 2013.] <http://www.dtic.co.cu/novenasemanatecnologica>

RODRÍGUEZ, A., VIÑA, S., CASARES, R. and HEREDIA, J. Gestión de Comportamientos Basada en Inteligencia Organizacional. Revista de Inteligencia Empresarial, 2010, 1(Especial): p. 10-15