

Simulador de tareas docentes para la asignatura de Procesamiento de Datos en las Geociencias para la carrera de Geofísica

Oswaldo Rodríguez Morán¹

¹ Centro de Investigaciones del Petróleo, Churrucá No. 481 entre Washington y Vía Blanca, Cerro, C. P. 12000, La Habana, Cuba, Email: ormoran@ceinpet.cupet.cu.

RESUMEN

Los ejemplos de aplicaciones geológicas-geofísicas que se utilizan para el aprendizaje de los estudiantes de la carrera de Ingeniería Geofísica de la Facultad de Civil de la Universidad Tecnológica de La Habana (CUJAE), en particular para los que cursan la asignatura de Procesamiento de Datos en las Geociencias, no abundan. En la actualidad parecen haberse agotado las posibles variantes de tareas geólogo-geofísicas de corte docente para los estudiantes. El problema es entonces, que los estudiantes coleccionan el historial de las tareas docentes de los años anteriores, incluso sus soluciones, quedando sin efecto el esperado impacto, de estas tareas en el proceso docente-educativo. Por esta razón, el objetivo general que se propone es el de simular diferentes tareas geólogo-geofísicas con objetivos docentes para la asignatura Procesamiento de Datos en las Geociencias de la carrera de Ingeniería Geofísica. Se utilizaron los documentos referidos a los objetivos instructivos de la asignatura, además del plan de secuencia de clases con los objetivos de cada tema de la asignatura y se instrumentaron los métodos estadísticos clásicos univariados y multivariados. Para cada tema de la asignatura se generaron, de modo conceptual, variantes de tareas geólogo-geofísicas, las cuales podrán incrementarse en otras versiones futuras. Como resultado, se hizo un análisis de los fundamentos y exigencias de cada una de las tareas geólogo-geofísicas de la asignatura con objetivos docentes y, el diseño y puesta a punto de un algoritmo para generar las tareas geológicas-geofísicas propuestas. Así, en cada curso, cada estudiante podrá tener una variante única de tarea para solucionar, lo cual contribuye al proceso docente-educativo.

Palabras claves: Simulación, Procesamiento de datos, Geociencias, Proceso docente-educativo.

ABSTRACT

Examples of geological-geophysical applications that are used for the learning of students of the Geophysical Engineering career of the Faculty of Civil of the Technological University of Havana (CUJAE), in particular for those who study the subject of Processing of Data in the Geosciences, they do not abound. At present, it seems to have been exhausted the possible variants of geological-geophysical tasks of teaching for students. So we are faced with the problem of students collecting the history of the tasks of previous years, including their solutions, leaving out the expected impact of these tasks on the teaching-learning process. For this reason, the general objective we propose is to simulate different geological-geophysical tasks with teaching objectives for the subject Data Processing in the Geosciences of the Geophysical Engineering career. Were used the documents relating to the instructional objectives for the subject, in addition to the sequence of classes with the objectives of each topic of the course and were used too, the classical statistical methods univariate and multivariate analyzes. For each theme of the subject, variants of geological-geophysical tasks were conceptually generated, which could be increased in other future versions. As a result, an analysis was made of the fundamentals and demands of each of the geological and geophysical tasks of the subject with teaching objectives and, the design and development of an algorithm to generate the geological-geophysical tasks proposals. Thus, in each course, each student can have a

unique variant of task to solve, which contributes to the teaching-learning process.

Keywords: Simulation, Data processing, Geosciences, Teaching-learning process, Simulator teaching tasks.

RESUMO

Os exemplos de aplicações geológico-geofísicas que são utilizadas para a aprendizagem dos alunos da carreira de Engenharia Geofísica da Faculdade de Ciências da Universidade Tecnológica de Havana (CUJAE), em particular para aqueles que estudam o assunto do Processamento de Dados Em Geociências, eles não abundam. No presente, parece ter esgotado as possíveis variantes de tarefas geológico-geofísicas de ensino para estudantes. Então, somos confrontados com o problema dos alunos que coletam a história das tarefas de ensino de anos anteriores, incluindo suas soluções, deixando o impacto esperado dessas tarefas no processo de ensino educacional sem efeito. Por esta razão, o objetivo geral que propomos é simular diferentes tarefas geológico-geofísicas com objetivos de ensino para o assunto Processamento de Dados nas Geociências da carreira de Engenharia Geofísica. Utilizamos os documentos relativos aos objetivos de instrução para o sujeito, para além da seqüência de aulas com os objetivos de cada tópico do curso e foram os métodos estatísticos clássicos análises univariada e multivariada. Para cada tópico do assunto, as variantes de tarefas geológico-geofísicas foram geradas conceitualmente, o que poderia ser aumentado em outras versões futuras. Como resultado, foi feita uma análise dos fundamentos e requisitos de cada uma das tarefas geológicas e geofísicas do sujeito com objetivos de ensino e o projeto e desenvolvimento de um algoritmo para gerar as tarefas propostas geológico-geofísicas. Assim, em cada curso, cada aluno pode ter uma variante única de tarefa a resolver, o que contribui para o processo de ensino educacional.

Palavras-chave: Simulação, Processamento de Dados, Geociências, Processo de ensino educacional, Tarefas de ensino simulador.

INTRODUCCIÓN

Entre los objetivos fundamentales de la asignatura de Procesamiento de Datos en la Geociencias, impartida en el tercer año de la carrera de Ingeniería Geofísica, está el de crear las habilidades necesarias para el manejo y el pro-

cesamiento de los datos geólogo-geofísicos, en la solución de las tareas geológicas propuestas. Sin embargo, la cantidad de ejemplos prácticos de procesamiento de datos, derivados en lo fundamental de casos reales, son relativamente escasos para los requerimientos de los estudiantes en cada curso. Esto hace que, en general, en los últimos cursos, se hayan repetido, de modo inevitable, los casos de ejemplo de procesamiento de tareas docentes, enfrentándonos a la realidad, de que todos los estudiantes, al comienzo de los cursos, copian todas las tareas de la asignatura y, por supuesto, tareas docentes ya resueltas. Esta problemática crea un incumplimiento sistemático en los objetivos de la asignatura. Sobre la base de esta problemática, se hará la pregunta científica siguiente: ¿Cómo poder generar tareas geólogo-geofísicas docentes que sean únicas para cada estudiante y que cumplan con los objetivos de la asignatura de Procesamiento de Datos en la Geociencias? Teniendo identificado el objeto de estudio como las tareas geólogo-geofísicas, se establecerá el objetivo general de este trabajo como el de simular tareas geólogo-geofísicas con objetivos docentes para la asignatura Procesamiento de Datos en la Geociencias, de la carrera de Ingeniería Geofísica.

Los objetivos específicos, los cuales tributarán al general serán: (1) Analizar los fundamentos y exigencias de cada una de las tareas geólogo-geofísicas con objetivos docentes y (2) Diseñar algoritmos generadores sobre la plataforma de *Microsoft Excel* para la creación de las tareas docentes propuestas.

A partir de este problema se llevó a cabo el diseño de una serie de algoritmos para generar de modo automático las tareas docentes y laboratorios de la asignatura, incluso, una para cada estudiante y no descansar su solución en el trabajo por equipos, como de forma tradicional se ha hecho. A estos algoritmos se les debe exigir las siguientes características:

- Que los datos “observados”, generados de modo automático, respondan a una supuesta realidad geológica-geofísica determinada.
- Que estos datos representen una data única, lo que implica una solución siempre única en general.
- Que el algoritmo generador de tareas docentes para los estudiantes, pueda también ofrecer las soluciones correspondientes de cada tarea. De esta forma se le facilitará, al

profesor el chequeo de las mismas, las cuales tendrán características individuales.

En estas actividades docentes de tareas y laboratorios, se entregará una guía de acorde al procedimiento de cada una de ellas. Se deberá tener un extraordinario cuidado al elaborar la guía con la que el estudiante debe trabajar de manera independiente para llevar a cabo la preparación individual y la solución de las tareas. Las preguntas o aspectos del contenido seguirá un orden lógico que ayude al estudiante a realizar su trabajo y debe irse concretando en la medida que el curso académico avance. Las prácticas se organizarán y planificarán de manera sistemática de manera que el estudiante pueda vencer de forma progresiva los problemas más complejos.

MATERIALES Y MÉTODOS

Los elementos de los cuales partimos para establecer los fundamentos de este trabajo son los objetivos instructivos de la asignatura Procesamiento de Datos en la Geociencias, de la carrera de Ingeniería Geofísica de la Facultad de Civil de la Universidad Tecnológica de la Habana, sobre los cuales se sustenta el proceso docente-educativo de esta asignatura (Objetivos y Secuencia de Actividades de la asignatura Procesamiento de Datos en la Geociencias, 2015), ellos son los siguientes:

- 1) Conocer el papel de los métodos estadísticos en las Geociencias y su importancia durante el procesamiento y análisis de información geólogo-geofísica.
- 2) Conocer los fundamentos teóricos del análisis estadístico multivariado y el análisis exploratorio de una matriz de datos.
- 3) Conocer las diferentes técnicas estadísticas multivariadas que pueden ser empleadas durante el procesamiento y análisis de información geólogo-geofísica.
- 4) Dominar las diferencias entre señal útil y ruido, así como la importancia de estos conceptos en la actividad de la Geofísica Aplicada.
- 5) Conocer técnicas para la detección de señales útiles débiles tanto unidimensionales como bidimensionales.
- 6) Describir las técnicas geo estadísticas para el estudio de las variables regionalizadas.
- 7) Conocer las posibilidades de las técnicas

estadísticas de reconocimiento de patrones durante el procesamiento y análisis de información geólogo-geofísica.

- 8) Describir las técnicas de reconocimiento de patrones utilizadas durante la solución de tareas en las investigaciones geólogo-geofísicas.
- 9) Describir técnicas estadísticas avanzadas utilizadas en evaluación de la efectividad geológica de los atributos existentes en una matriz digital de datos geólogo-geofísicos.

No puede perderse de vista que el producto final de este trabajo, es decir, las tareas geólogo-geofísicas diseñadas para ser resueltas por los estudiantes, deberá cumplir las exigencias de estos objetivos instructivos.

Los métodos utilizados para lograr el objetivo propuesto fueron: los métodos estadísticos (Krumbein, W., F. Graybill, 1961; Dixon W. and Massey F., 1966; Cramer, H., 1968; Davis, J.C., 1986; Spiegel. M. R., 1971; Freud, J. E., 1977; Ostle, B., 1977) y de simulación, la programación de celdas dinámicas en Hojas *Excel* y la programación de macros en lenguaje de programación en *Visual Basic* en la plataforma *Microsoft Excel* para la automatización de este proceso de creación de tareas geológicas-geofísicas docentes (Alfonso, 1979-82; Alfonso, 1989).

Entre los documentos utilizados para la formulación de las bases de este trabajo, fueron:

- Plan de secuencia de clases con los objetivos de cada tema de la asignatura. A partir de este plan se identificaron los siguientes de procesamiento que se imparten en la asignatura, estos son: (1) Análisis Exploratorio de Datos; (2) Regionalización de variables por el método del Inverso de la Distancia; (3) Análisis estructural; (4) Regionalización de variables por el método *Kriging*; (5) Análisis de Componentes Principales y Análisis de Factores; (6) Análisis de Clasificación Supervisada; y (7) Análisis de Clasificación no Supervisada (Objetivos y Secuencia de Actividades de la asignatura Procesamiento de Datos en la Geociencias, 2015).
- Recopilación de los diferentes rangos y valores medios de las propiedades petrofísicas de las rocas. Se confeccionó una tabla, la cual posee 56 variantes de: Tipo de roca,

Nombre de roca, Propiedad, Valor mínimo, Valor Medio y Valor Máximo (REFERENCIA). Esta tabla muestra los valores más probables y los rangos de las propiedades físicas de densidad [g/cm^3] y la porosidad [%] para las rocas ígnea, metamórfica, para minerales, minerales no metálicos y depósitos no consolidados.

- Propuesta de los rangos de valores de campos potenciales para una región simulada. A partir de los objetivos de la asignatura y del plan de secuencia de clases, se identifican los siguientes procesamientos de datos, que podrán ser aplicados a diferentes tareas geólogo- geofísica: (1) Análisis Exploratorio de Datos, (2) Métodos de Interpolación, (3) Análisis Estructural, (4) Regionalización de variables por el método *Kriging*, (5) Análisis de Componentes Principales y Análisis de Factores, (6) Análisis de Clasificación Supervisada y (7) Análisis de Clasificación no Supervisada.

RESULTADOS Y DISCUSIÓN

El contenido de las preguntas para cada tarea, fue concebido a partir de la definición de cinco sectores textuales. El primero es un sector textual variable, significa que para cada tarea docente generada, variará este sector textual y se denominará con las siglas: STV1 (Sector Textual Variable 1) y se identificará en los ejemplos mostrados con fuente de color rojo y en negritas. El segundo es otro sector textual variable, en el que se propone el tipo de atributo a procesar y se denominará con las siglas: STV2 (Sector Textual Variable 2) y que se identificará en los ejemplos mostrados con fuente de color azul y en negritas. El tercer sector es numérico variable y define la variabilidad de los datos simulados para esta tarea y sus siglas son: SNV (Sector Numérico Variable), se identificará en los ejemplos mostrados con fuente de color verde y en negritas. El cuarto es el sector textual invariante, en el que se expresan las diferentes preguntas y aspectos solicitados para que sean respondidos por los estudiantes y que no poseen cambio textual entre las tareas docentes y sus siglas son: STI (Sector Textual Invariante) y el color de la fuente para su identificación será de color negro. El quinto, representa la Tabla de Objeto-Propiedades asociada al problema, la cual es generada de forma aleatoria,

en general con datos normalizados o tipificados. La identificación de la tabla se mostrará en color negro y en negritas.

A continuación se expone un ejemplo de la nomenclatura anterior:

En un levantamiento geológico se tomaron 15 muestras georreferenciadas y se determinó en el laboratorio el valor de la densidad de cada una de ellas. En la **Tabla 2-1** se ofrecen las coordenadas Lambert y el valor del atributo correspondiente (densidad [g/cm^3]). (a) Determine el valor del atributo (densidad [g/cm^3]) en la ubicación espacial: (350412.9, 334891.3), por el método de Interpolación del Inverso de la Distancia. (b) Haga un mapa en el que se muestre la ubicación, tanto los valores de base del cálculo, como el punto donde se necesita estimar el valor del atributo. (c) Expresé los cálculos en una tabla apropiada, donde se muestre el procedimiento de la estimación.

A continuación se expondrá el contenido estructural y el alcance de cada una de las 7 tareas geólogo-geofísicas simuladas:

1) Análisis Exploratorio de Datos. Constará de cuatro subpreguntas:

- 1) Pregunta 1.1- Tratará sobre una serie de datos de una propiedad física (densidad, porosidad, etc.) entregados al estudiante y tiene como objetivos: (1) Calcular los cuatro momentos estadísticos de los datos; (2) Hacer un análisis exploratorio de los datos de la Tabla 1.1, (3) Hacer una propuesta analítica de los posibles tipos de rocas (ígnea, metamórfica o sedimentaria), mineral o mineral no metálico recolectados; y (4) Construir el histograma y el gráfico de frecuencia acumulativa de los datos.

Como es necesario garantizar la máxima variedad de los datos entregados a cada estudiante, es que se seleccionará de forma aleatoria el tipo de roca (ígnea, metamórfica o sedimentaria), mineral o mineral no metálico de una tabla creada (REFERENCIA), después tomar el rango correspondiente al objeto seleccionado y generar los datos entregados al estudiante a partir de una función acumulativa de frecuencia modelo (Figura 1). En este caso, puede apreciarse en la Figura 1 cómo si se generan números aleatorios correspondientes a la ordenada de

la función acumulativa de frecuencia modelo, pueden obtenerse valores aleatorios dentro del rango propuesto por la selección del objeto de la tarea. Estos valores de la propiedad física resultante será el conjunto de datos simulados para esta tarea. Como ejemplo final, una de las variantes para esta Tarea 1.1, es:

A una cantidad de 211 muestras listadas en la Tabla 1.1 (Columna A), se le han determinado en el laboratorio su densidad

[g/cm³]. (1) Calcule los cuatro momentos estadísticos de los datos; (2) Haga un Análisis Exploratorio de los datos de la Tabla I. (3) Haga una propuesta analítica de los posibles tipos de rocas (ígneas, metamórficas o sedimentaria), mineral o mineral no metálico recolectados, a partir de las referencias entregadas. (4) Construya el histograma y el gráfico de frecuencia acumulativa de los datos de la Tabla 1.1.

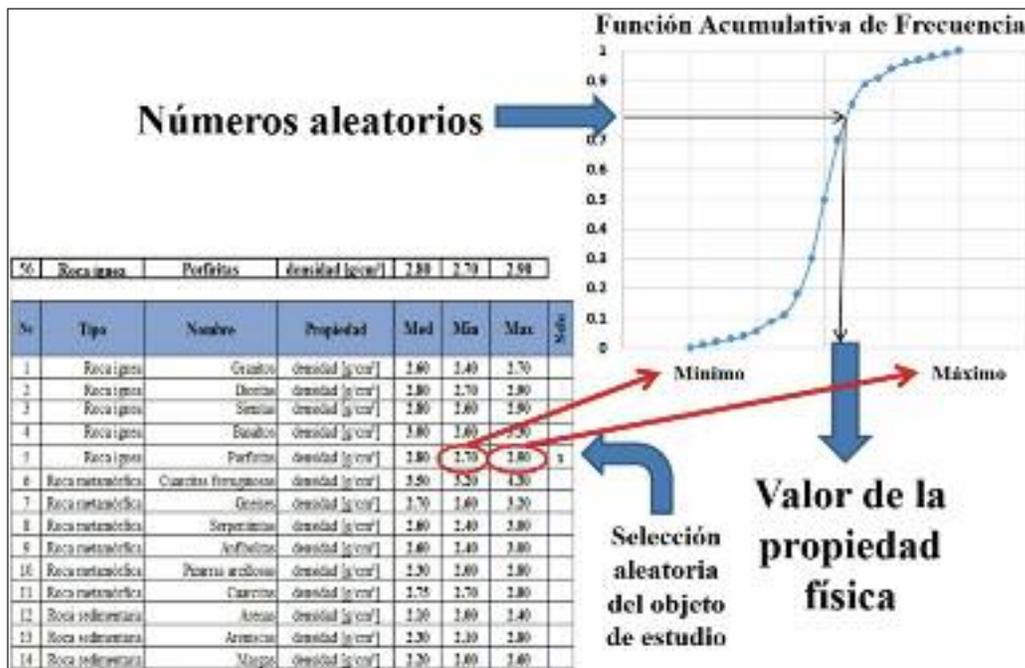


Figura 1. Proceso de selección del objeto de la tarea docente, generación de números aleatorios y de los datos simulados para la tarea.

2) Pregunta 1.2.- Se refiere a una serie de datos graficados en un histograma, a los cuales se le pide al estudiante que realice un Análisis Exploratorio de Datos a partir de datos agrupados. Los datos son tomados de la tabla de 56 variantes de: Tipo de roca, Nombre de roca, Propiedad, Valor mínimo, Valor Medio y Valor Máximo, con todos los presupuestos de aleatoriedad en la selección. En este problema se creará por simulación dos distribuciones con cierto grado de mezcla, donde el estudiante tendrá la oportunidad de apreciar si existen dos poblaciones, decir cuál sería la distribución principal y cuál la "contaminación", además decir cuál de ella en apariencia podría tener distribución normal o no (siempre adicionando la necesidad de hacer

pruebas de normalidad para esa distribución de datos), si las distribuciones son asimétricas, decir si es hacia la derecha o la izquierda y de modo conclusivo, si existen fluctuaciones estadísticas o *outlier*. En la Figura 2 puede apreciarse, cómo al seleccionar de modo aleatorio las condiciones deseadas, por ejemplo, si la distribución principal tiene distribución normal, o asimétrica a la derecha o asimétrica a la izquierda. Además si la contaminación está situada en valores menores (izquierda) o sobre valores mayores (derecha) y si la fluctuación estadística está en el extremo derecho o izquierdo. Una realización de las posibles variantes de esta tarea es:

En la Tabla 1.2 se muestra los datos de las

frecuencias absolutas de datos agrupados, los cuales se presentan graficados en el Histograma de abajo. Estos datos son los valores de densidad [g/cm³] de 285 muestras de Epidota (Mineral no metálico). Haga un Análisis Exploratorio valorativo de estos

datos verificando: (1) La existencia de la distribución principal respecto al tipo de roca del problema. (2) La existencia de una contaminación en los datos y sus características. (3) La posible existencia de alguna fluctuación y sus características.

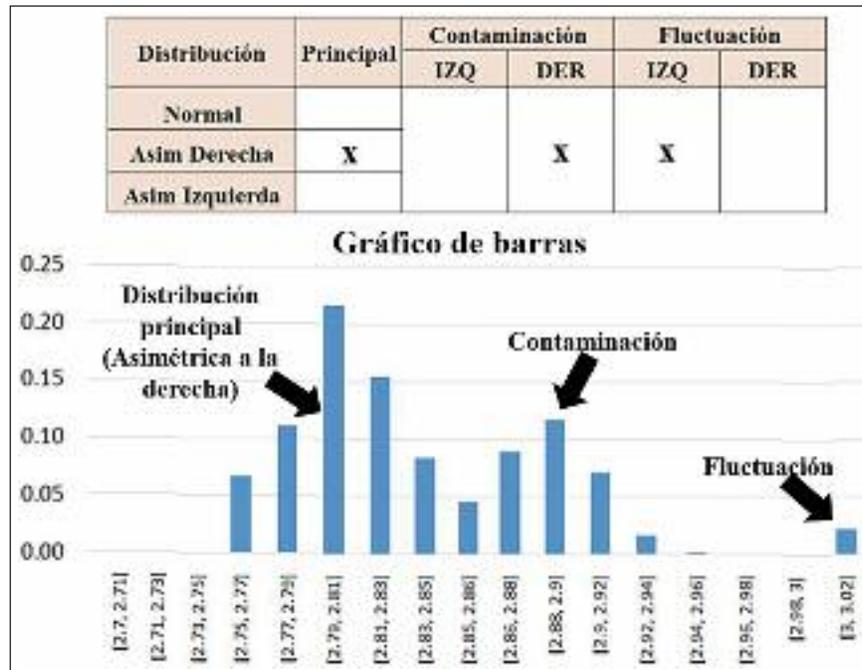


Figura 2. Proceso de creación de los datos para el segundo inciso de la tarea 1, basada en el Análisis Exploratorio de Datos agrupados.

3) Pregunta 1.3- Se le entrega al estudiante un gráfico 2D establecido entre las variables expresadas como V1 (variable independiente) y V2 (variable dependiente), las cuales han sido normalizadas en un rango de 0 a 1 y se pide hacer el gráfico 2D de dependencia y un Análisis Exploratorio de Datos de este gráfico. En este análisis el estudiante será capaz de decir cuántos grupos se aprecian en el gráfico, la característica de distribución de cada grupo y por último, si existen fluctuaciones e identificarlas en el gráfico. En la Figura 3 se aprecia una de las variantes. A continuación una ejemplificación de la tarea:

En la Tabla 1.3, se encuentran relacionados dos atributos (V1 y V2) de 38 muestras de un levantamiento geológico. Estos atributos se han normalizados en un rango de 0 a 1. Haga un Análisis Exploratorio de Datos precisando lo siguiente: (1) ¿Cuántos grupos se

pueden apreciar en el gráfico?; (2) ¿Qué característica posee cada grupo identificado?, entre ellos: el tipo de dependencia, valor aproximado de la pendiente, el centro aproximado del grupo; (3) A su consideración, identifique las fluctuaciones que puedan existir; (4) Construya el gráfico correspondiente entre V1 y V2.

4) Pregunta 1.4- Se le presenta al estudiante de una matriz de correlación de 6 muestras geológicas a partir de sus atributos y se le pide que diga cuáles parejas de objetos están correlacionados por encima (o por debajo) de un valor de similitud determinado. Un ejemplo de la tarea es el siguiente: Se dispone de una matriz de correlación de 6 muestras geológicas a partir de sus atributos (Tabla 1.4). ¿Cuáles parejas de muestras se correlacionan por encima de un valor de correlación de 0.5?

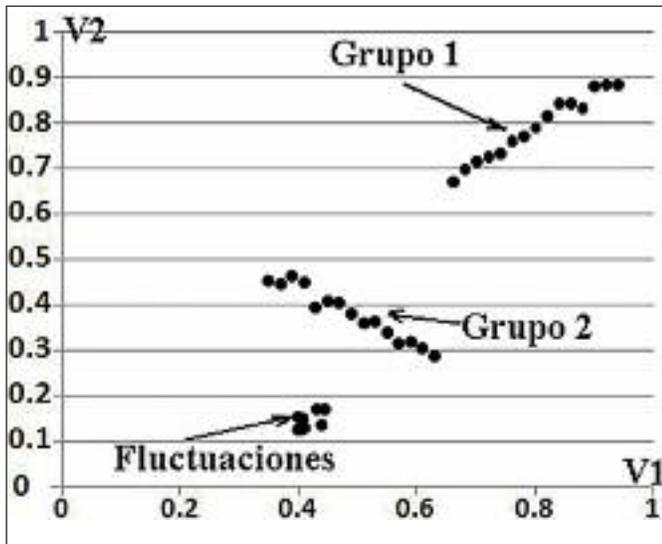


Figura 3. Ejemplo de datos 2D donde el estudiante deberá hacer un Análisis Exploratorio de Datos.

2) Regionalización de variables por el método del Inverso de la Distancia.

Esta segunda tarea, propone aplicar el método de interpolación por el método Inverso de la Distancia, para determinar el valor de un atributo en una ubicación espacial donde se desconoce. Se ofrece al estudiante una cantidad de tríadas de valores en el ámbito de 10 km² (coordenadas variables) y un atributo para cada tríada (profundidad [m], densidad [g/cm³], altura [m] o Bouguer [mGal]). Por último se pide estimar el valor del atributo del problema en una ubicación [x₀, y₀] dada.

En la Figura 4 se muestra un ejemplo de simulación en que propone estimar el valor del atributo en una ubicación espacial donde se desconoce, además, esta tarea se expresará de la siguiente forma:

Se ha realizado un levantamiento topográfico en un área determinada, con el fin de elaborar un Modelo Digital de Elevación. En la Tabla 2 se ofrecen las coordenadas Lamberts (x,y) y el valor del atributo correspondiente (Altura del relieve [metros]) de 10 tríadas de valores. (a) Determine el valor del atributo en la ubicación espacial: (350336.5, 335473.4), por el método de Interpolación del Inverso de la Distancia. (b) Haga un mapa en el que se muestre la ubicación, tanto los valores de base del cálculo, como el punto donde se necesita estimar el valor del atributo. (c) Exprese los cálculos en una tabla apropiada, donde se muestre el procedimiento de la estimación.

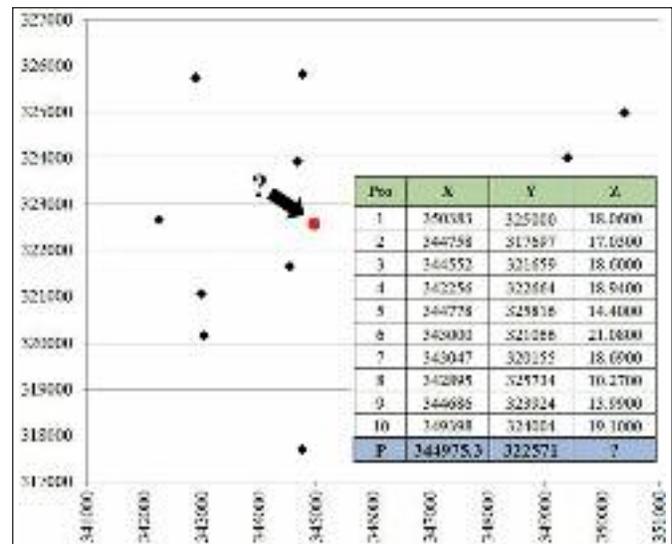


Figura 4. Ejemplo de simulación en el que se propone estimar el valor del atributo en una ubicación espacial donde se desconoce este.

3) Análisis estructural.

En esta tarea, al estudiante se le provee de una distribución espacial discreta en 2D del atributo de profundidad y se le solicita hacer el análisis estructural de la distribución espacial del atributo dado. Para ello tendrá que: (1) Calcular y graficar el semivariograma isotrópico de los datos de referencia y representar en el mismo gráfico su propuesta del semivariograma Teórico (Modelo esférico) informando el valor del Efecto *Nugget*, el Alcance y la Meseta, (2) Calcular y graficar el semivariograma anisotrópico de los datos de referencia en la dirección Norte-Sur y representar en el mismo gráfico su propuesta del semivariograma Teórico (Modelo esférico) informando el valor del Efecto *Nugget*, El alcance y la Meseta; (3) Calcular y graficar el semivariograma anisotrópico de los datos de referencia en la dirección Este-Oeste y representar en el mismo gráfico su propuesta del semivariograma Teórico (Modelo esférico) informando el valor del Efecto *Nugget*, El alcance y la Meseta; (4) Al final se realizará un análisis de los resultados en los incisos (1), (2) y (3) y proponer una clasificación del tipo de anisotropía del atributo estudiado, si es que existe. En la Figura 5 se muestra un ejemplo del resultado de la simulación para la tarea del Análisis Estructural. En esta figura puede apreciarse los variogramas experimentales isotrópicos y los anisotrópicos N-S y E-O con sus respectivas propuestas de variogramas teóricos. La pregunta es enfocada de la siguiente forma:

Se dispone de un conjunto de tríadas de valores en la Tabla 3 mostrada. Estos valores representan la distribución espacial discreta en 2D del atributo de profundidad [metros]. Se desea hacer el análisis estructural de la distribución espacial del atributo. (1) Calcule y grafique el semivariograma isotrópico de los datos de referencia. Represente en el mismo gráfico su propuesta del semivariograma Teórico (Modelo esférico) informando el valor del Efecto *Nugget*, el Alcance y la Meseta. (2) Calcule y grafique el semivariograma anisotrópico de los datos de referencia en la dirección Norte-Sur.

Represente en el mismo gráfico su propuesta del semivariograma Teórico (Modelo Esférico) informando el valor del Efecto *Nugget*, El alcance y la Meseta. (3) Calcule y grafique el semivariograma anisotrópico de los datos de referencia en la dirección Este-Oeste. Represente en el mismo gráfico su propuesta del semivariograma Teórico (Modelo Esférico) informando el valor del Efecto *Nugget*, El alcance y la Meseta. (4) Realice un análisis de los resultados en los incisos (1), (2) y (3) y proponga una clasificación del tipo de anisotropía del atributo estudiado, si es que existe.

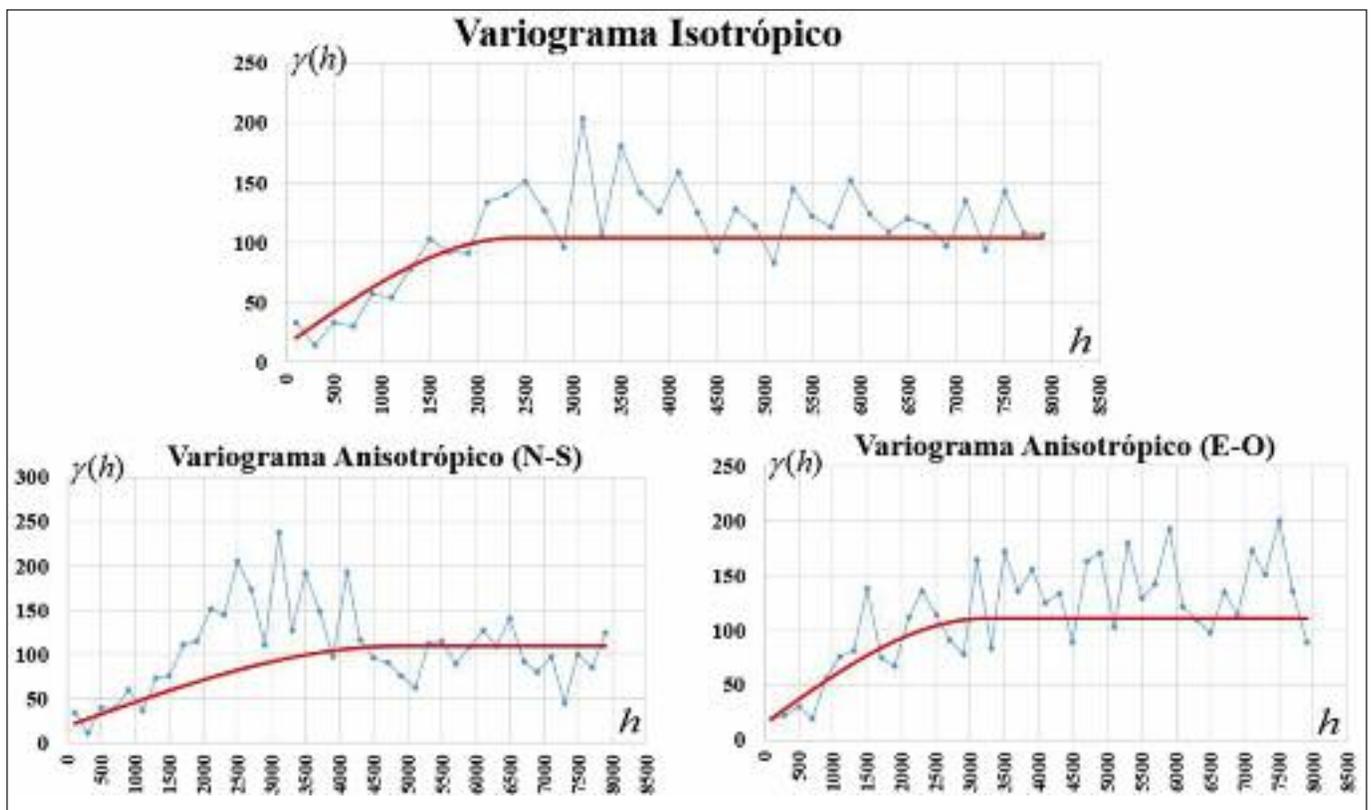


Figura 5. Ejemplo del resultado de la simulación para la tarea del Análisis Estructural.

4) Regionalización de variables por el método *Kriging*.

Se generará de forma automática una tabla de 100 tríadas de valores con datos de altura del relieve [m] en un área de ubicación variable. Se pedirá al estudiante que: (1) Regionalice y haga el mapa de la variable por el método del Inverso de la Distancia en una red de 50 x 50 m. (2) Regionalice y haga el mapa de la variable por el método *Kriging* en una red de 50 x 50 m. (3) Haga una discusión

sobre los resultados de cada una de las regionalizaciones (Figura 6). La pregunta pertinente quedaría estructurada de la siguiente forma:

En una zona de área cuadrada, limitada por las coordenadas Lambert del punto inferior izquierdo: (367000, 356000) y el punto superior derecho: (378000, 367000) se ha realizado un levantamiento geoquímico (Contenido de Oro [Au]). La cantidad de puntos en el área son de

100 triadas de valores (x, y, z). (1) En el ámbito del área señalado haga una regionalización del atributo por el método Inverso de la Distancia. (2) Realice un comentario sobre la característica de la distribución areal del atributo. (3) Haga una regionalización del atributo

por el método *Kriging* y realice un comentario sobre el resultado de las regionalizaciones efectuadas por ambos métodos. Nota: Utilice todos los puntos de la Tabla para la regionalización y en todos los casos, realice la interpolación con una red de 50 x 50 puntos.

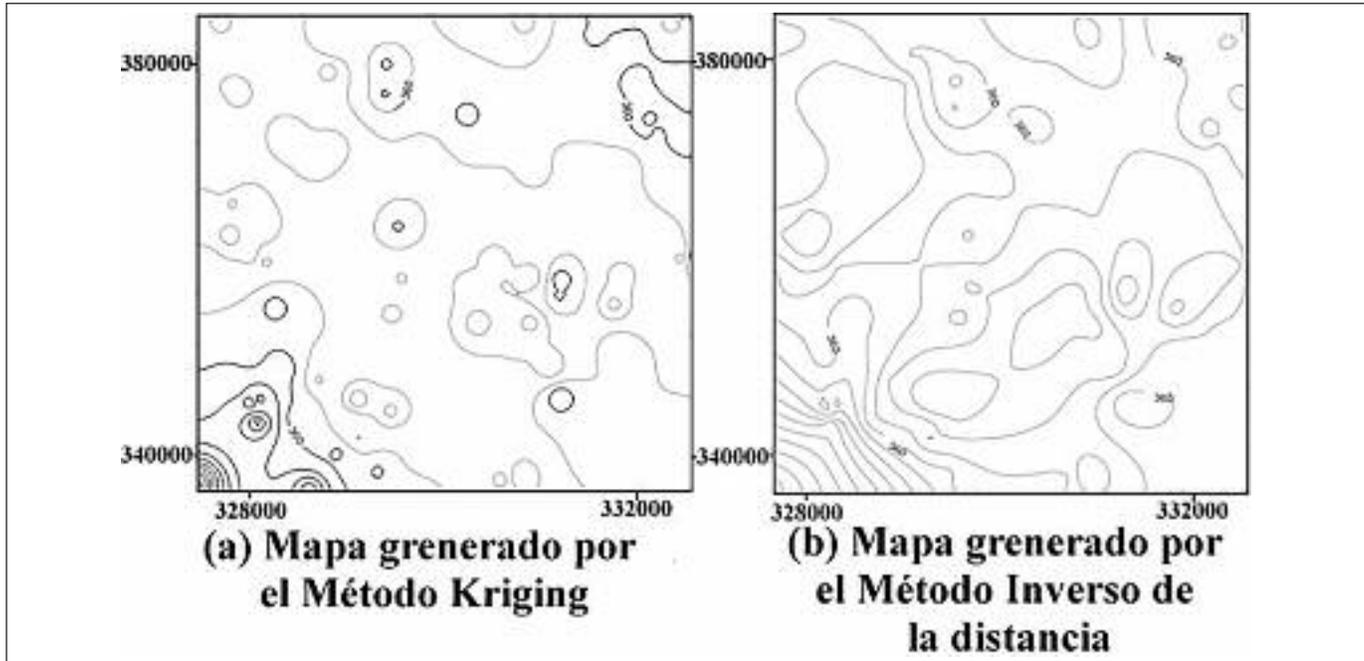


Figura 6. Mapas generados por la Tabla 4 con el objetivo de comparar dos métodos de regionalización de variables.

5) Análisis de Componentes Principales y Análisis de Factores.

Se generará de modo automático una Tabla de Objetos-Propiedades (TOP) con datos normalizados, en los que aparecen una serie de casos y entre 7 y 12 atributos. De forma Aleatoria se seleccionan dos parejas de atributos, los cuáles se generarán de forma dependiente cada pareja, para estructurar el problema. El estudiante debe ser capaz en esta tarea, aplicando el Análisis de las Componentes Principales y el de Análisis de Factores, de: (1) Mostrar en tablas convenientes el resultado de calcular los valores y vectores propios del sistema; (2) de construir el gráfico de composición de cada uno de los valores propios del sistema; (3) saber decir el por ciento de explicación que el sistema aportaría, si se toman los tres primeros valores propios; (4) Hacer una propuesta de la posible reducción dimensional del sistema y explicar por qué; (5) Construir el gráfico de los dos primeros factores (F1 vs F2) y hacer una discusión del grado de dependencia entre los atributos para esos factores. En la Figura 7

se aprecia una variante simulada, en la que se permite ver la tabla de Objeto-Propiedades de 30 casos y 8 atributos, además, las variables escogidas de modo aleatorio para representar el caso de variables dependientes (la variable 3 dependiente de la 4 y la variable 5 dependiente de la 6), y al final, la tabla de los autovalores, el gráfico correspondiente y el gráfico 2D de los Factores F1 y F2. La pregunta de esta tarea sería:

En el estudio ingeniero geológico para la construcción de un edificio, se tomaron 30 muestras de rocas, a las cuáles se les fueron determinadas las propiedades físico-mecánicas en el laboratorio de mecánica de suelos, en total 8 atributos (V1, V2, V3, V4, V5, V6, V7, V8). En la Tabla 5 de Objeto-Propiedades entregada para este problema se muestran, para cada objeto, los atributos que han sido normalizados entre valores de 0 y 1. Aplique el método de las Componentes Principales para: (1) Mostrar en tablas convenientes el re-

sultado de calcular los valores y vectores propios del sistema. (2) Construir el gráfico de composición de cada uno de los valores propios del sistema. (3) Responder: ¿Qué porcentaje de explicación del sistema aportaría tomar los tres primeros valores propios? (4)

Hacer una propuesta de la posible reducción dimensional del sistema y explique por qué. (5) Construir el gráfico de los dos primeros factores (F1 vs F2) y hacer una discusión del grado de dependencia entre los atributos para esos factores.

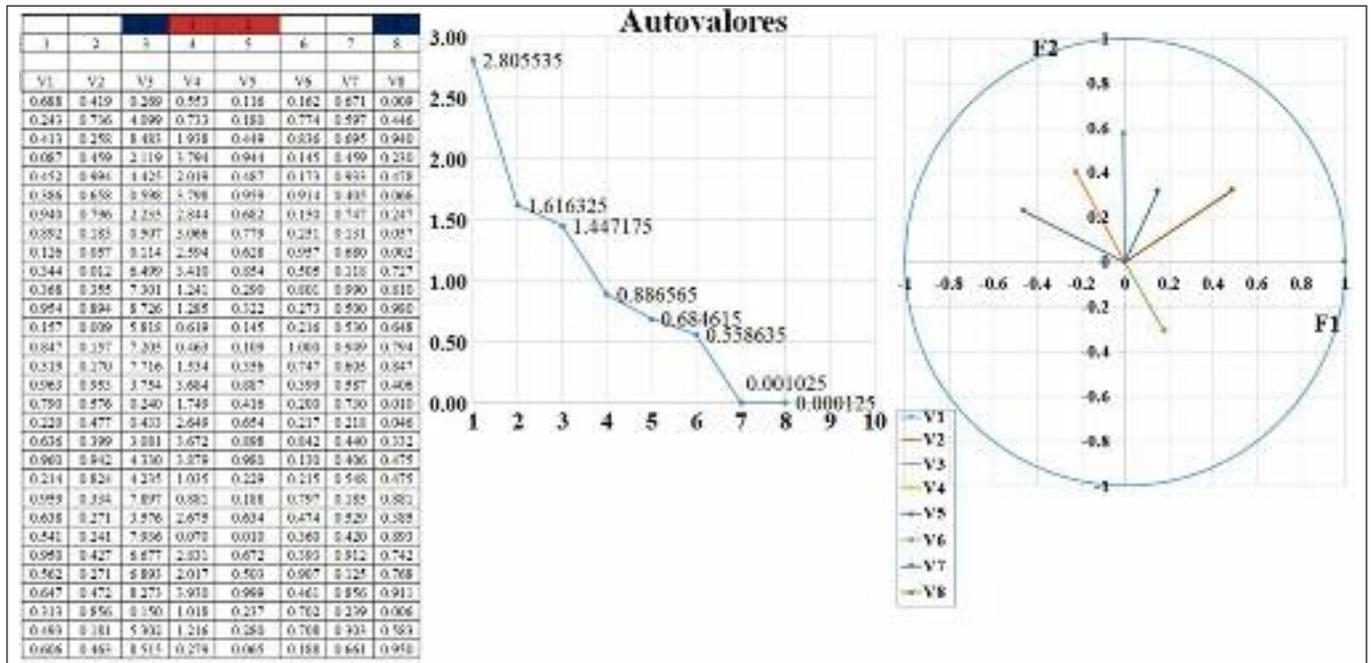


Figura 7. Variante simulada para la tarea de Análisis de Componentes Principales y Análisis de Factores.

6) Análisis de Clasificación Supervisada.

Para esta tarea se genera una Tabla de Objeto-Propiedades de datos tipificados de 3 a 8 atributos y se enuncia el problema de una tarea geólogo-geofísico determinada, la cual podrá poseer entre 5 y 8 grupos posibles de trabajo, con una cantidad entre 50 y 180 de objetos o casos. Se pedirá al estudiante hacer una clasificación a partir de los patrones mostrados utilizando la herramienta gráfica del dendograma, por lo que hay que exponer de forma explícita los casos y la clasificación correspondiente. En la Figura 8 se muestra un ejemplo de gráfico de dendograma, de una TOP propuesta para una clasificación supervisada. Los estudiantes podrán –de modo visual– realizar la clasificación de los elementos por clasificar de la TOP entregada. Un ejemplo de todas las posibles variantes sería:

Se quiere hacer una clasificación supervisada a partir de 123 objetos y 3 atributos mostrados en la TOP de la Tabla 6, cuyos datos están tipificados. Estos objetos representan los posibles tipos de relieve que se desean clasificar. Se han defi-

nido en esta TOP, 8 patrones, los cuales son: LLANO, SEMI LLANO, ONDULADO, COLINOSO, LOMAS BAJAS, LOMAS MEDIAS, LOMAS ALTAS, y MONTAÑOSO. Realice la clasificación solicitada a partir de los patrones mostrados utilizando la herramienta gráfica del dendograma.

7) Análisis de Clasificación no Supervisada.

Situándose en una región dividida en celdas cuadradas de 500 x 500 metros, limitada por las coordenadas del punto inferior izquierdo $[X_i, Y_i]$ y por las coordenadas del punto superior derecho $[X_f, Y_f]$ de coordenadas Lambert, donde a cada celda se le asocia k atributos geólogo-geofísicos los cuales han sido tipificados, los estudiantes tendrán que: (1) Hacer una clasificación no supervisada de las celdas del área a partir de sus atributos; (2) Construir el mapa de la región y hacer la zonación correspondiente al resultado de la clasificación; (3) Al final se hará una interpretación de los resultados desde el punto de vista

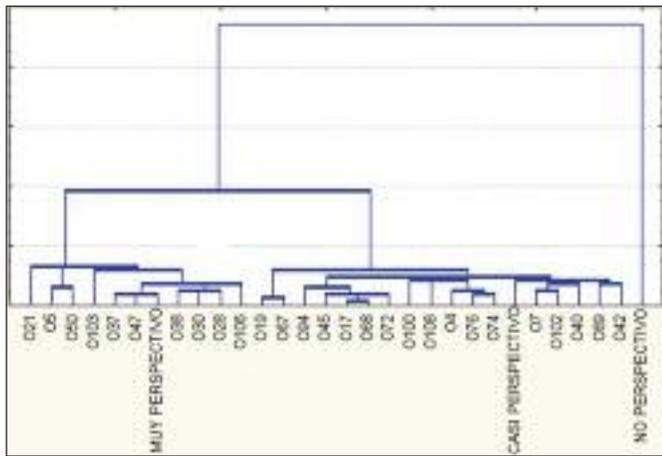


Figura 8. Dendrograma que representaría la clasificación supervisada de la en la Tarea 6.

cuantitativo-espacial. En la Figura 9 se muestra una parte de la TOP de esta tarea y el resultado de la clasificación espacial no supervisada de la región simulada. Una variante de pregunta para esta tarea es:

Se ha dividido en celdas cuadradas de 500 x 500 metros una región, limitada por las coordenadas del punto inferior izquierdo [438000, 325000] y por las coordenadas del punto superior derecho [442500, 329500] de coordenadas Lambert. Cada celda posee 7 atributos geológico-geofísicos los cuales han sido tipificados. (1) Haga una clasificación no supervisada de las celdas del área a partir de sus atributos, para ello se entrega la Tabla de Objeto-Propiedades Tabla 7. Utilice el dendrograma para realizar la clasificación a partir de la formación de grupos más estable que posea el sistema. (2) Construya el mapa de la región y la zonación correspondiente al resultado de la clasificación. (3) Construya una tabla con las columnas: GRUPO, #CELDAS, AREA, para calcular el área que representa cada grupo obtenido. (4) Haga una interpretación de los resultados desde el punto de vista cualitativo-espacial.

Este sistema de simulación de tareas docentes para la solución de problemas supuestos de la Geociencias, se diseñó en la plataforma *Microsoft Excel*. Se utilizaron tanto el procedimiento de cálculo de celdas dinámicas, como la programación de macros en *Visual Basic*, para dar solución a cada uno de las tareas de la asignatura.

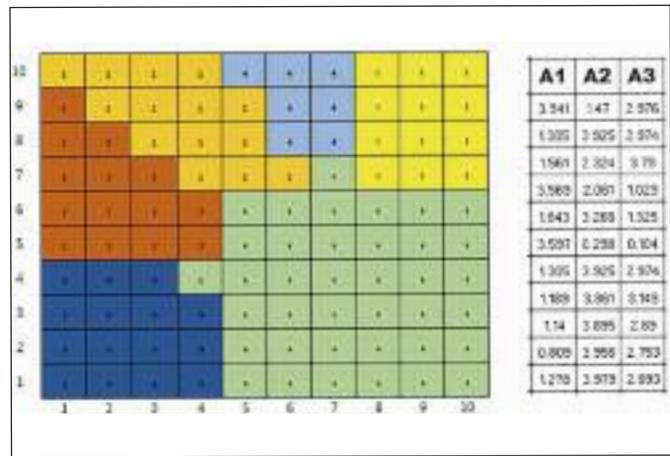


Figura 9. Resultado de la clasificación no supervisada de datos espaciales correspondiente a la Tarea 7. A la derecha se muestra una parte de la Tabla de Objeto-Propiedades suministrada para esta tarea.

En la hoja principal del programa (DatosGenerales) se introducen los nombres de los estudiantes que representan la matrícula del curso vigente (Figura 10). En esta hoja también se debe introducir el curso a que hace referencia la corrida, por ejemplo: 2016-17. De forma informativa aparece también el enunciado de cada una de las siete tareas docentes y sus siglas utilizadas por el programa. Aparece una imagen de una PC que representa un hipervínculo a la ejecución del programa (Figura 9).

En la ejecución de este programa se crean de modo automático dos carpetas en la raíz donde se esté ejecutando el simulador. Una carpeta contendrá las tareas docentes en un fichero en *Excel* (*.xls), para cada uno de los estudiantes listados en la hoja principal (DatosGenerales), con el nombre de carpeta: PDG_2016_2017_Estudiantes y la otra PDG_2016_2017_Profesor, donde estarán las respuestas a las tareas docentes de cada uno de los estudiantes en un fichero en *Excel* (*.xls) (Figura 11).

CONCLUSIONES

Se logró generar de forma automática tareas docentes que cumplen con los objetivos de la asignatura de Procesamiento de Datos en la Geociencias de la carrera de Ingeniería Geofísica, como respuesta a los limitados casos reales de tareas geológico-geofísicas existentes:

- a) Se analizaron los fundamentos y exigencias de cada una de las tareas geológico-geofísicas con objetivos docentes, de forma tal que cumplen con los objetivos docentes trazados por el Departamento de Geociencias de la Facultad de Civil de la Universidad Tecnológica de La Habana.

SIMULADOR DE TAREAS SOBRE EL PROCESAMIENTO DE DATOS EN LA GEOCIENCIA

14 MATRÍCULA		Curso: 2016-2017
Nº	Nombre y apellidos	
1	Baez Gutiérrez Rigoberto	
2	Bejerado Kindelan Aliss Maria	
3	Carbonell López Yanetka	
4	Díaz Méndez Ana Laura	
5	Dopico Acosta Diana	
6	Estrada Ramos Yinet	
7	Ferrada Labrador Carlos E.	
8	Fuentes Rodríguez Daisy Rocío	
9	Pandora Uffo Dayana E.	
10	García Martínez Ariel	
11	González Trujillo Omar	
12	Garciga Acosta Anabel	
13	Guerra Errasti Claudia	
14	Herbavarría Govin Elizabeth	

Tarea	Tema	Siglas	Sele
1	Análisis Exploratorio de Datos	AED	
2	Métodos de Interpolación	MIT	
3	Análisis Estructural	AET	
4	Regionalización de variables	RGV	
5	Análisis de Componentes Principales y Análisis de Factores	ACP-AF	
6	Clasificación Supervisada	CS	
7	Clasificación no Supervisada	CNS	

Figura 10. Hoja principal (DatosGenerales) donde se introduce la matrícula del curso vigente.

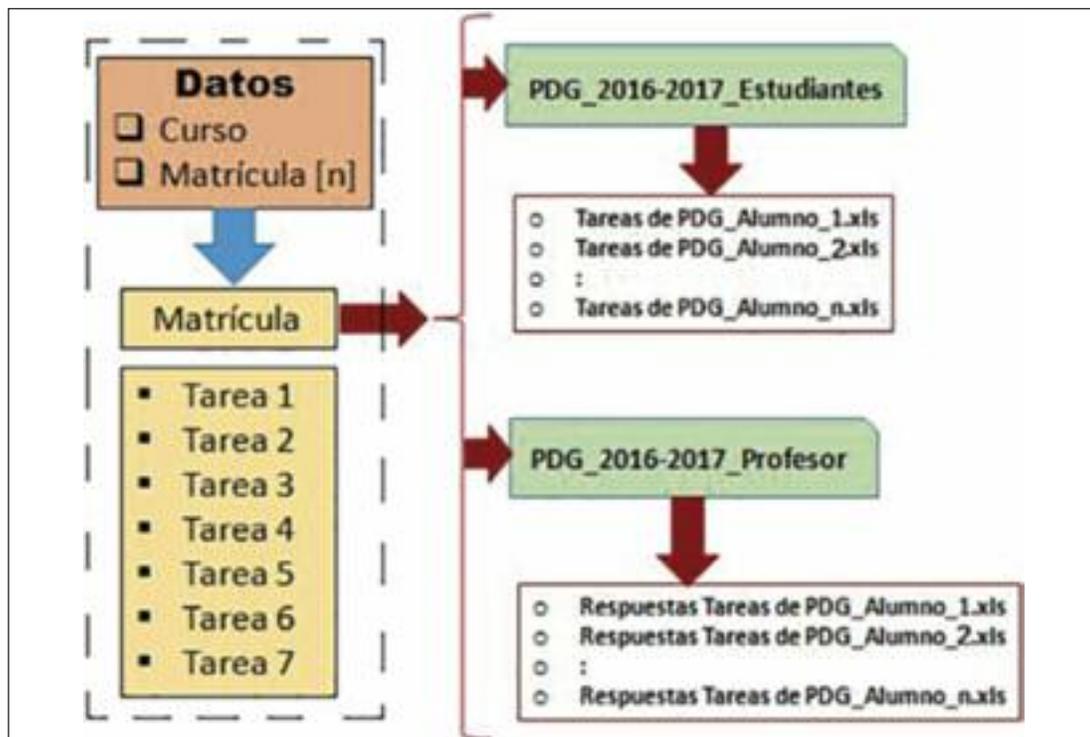


Figura 11. Esquema general del simulador de tareas docentes.

- b) Se diseñaron los algoritmos pertinentes para simular cada una de las tareas docentes y las respuestas correspondientes, ésta última, como herramienta de chequeo de estas tareas para el profesor de la asignatura.
- c) Se diseñaron los algoritmos generadores sobre la plata-

forma de Microsoft Excel para la creación de las tareas docentes propuestas, con la opción de aumentar la variedad de tareas geólogo-geofísicas y rangos para la generación de los diferentes tipos de datos.

REFERENCIAS BIBLIOGRÁFICAS

- Alfonso, J. A.**, 1979-82, Tratamiento numérico de datos geológicos, Conferencia de curso de postgrado. Facultad de Ingeniería Civil. ISPJAE. Ciudad de La Habana. Inédito.
- Alfonso, J. A.**, 1989, Estadísticas en las ciencias geológicas. Editorial ISPJAE. Tomo I, pp. 477 y Tomo II, 308 pp.
- Cramer, H.**, 1968, Elementos de la teoría de las probabilidades. Editora Aguilar, Madrid. Davis, J. C., 1986, Statistics and data analysis in geology. John Willey and Sons. Inc. New York, 550 pp.
- Dixon, W., Massey, F.**, 1966, Introducción al análisis estadístico. Editora del Castillo, Madrid.
- Freud, J. E.**, 1977, Estadística elemental moderna. Editora Pueblo y Educación, 466 pp. Krumbain, W., Graybill, F., 1961, An Introduction to Statistical Models in Geology. McGraw-Hill, New York.
- Ostle, B.**, 1977, Estadística Aplicada. Editora Científico-Técnica. Ciudad de La Habana, 629 pp.
- Objetivos y Secuencia de Actividades de la asignatura Procesamiento de Datos en la Geociencias**, 2015, Carrera de Ingeniería Geofísica Aplicada de la Facultad de Civil de la Universidad Tecnológica de La Habana.
- Spiegel. M. R.**, 1971, Teoría y problemas de estadística. Editora Ciencia y Técnica. Instituto Cubano del Libro. La Habana, 359 pp.

