VISUALBITOOL: Una Herramienta para la Visualización de Datos en Inteligencia de Negocios



MANUAL DEL USUARIO

Grupo de Investigación GRIAS- KDD

Departamento de Sistemas

Facultad de Ingeniería

Universidad de Nariño

San Juan de Pasto

2015

Contenido

Pág.

1. INTRODUCCIÓN A VISUALBITOOL 5
1.2 REQUERIMIENTOS 5
1.3 INSTALACIÓN. 5
2. ENTORNO DE LA HERRAMIENTA CON EJEMPLO PASO A PASO 6
2.1 ¿CÓMO INICIAR VISUALBITOOL? 6
2.2 CARGAR ARCHIVO DE DATOS. 6
2.3 APLICAR CALIDAD DE DATOS. 7
2.4 DISTRIBUCIÓN DE DATOS. 11
2.5 FILTROS DE ATRIBUTOS. 15
3. APLICAR PATRONES DE MINERIA. 20
3.1 VISUALIZACIÓN DE UN ÁRBOL DE DECISIÓN. 20
3.2 VISUALIZACIÓN DE LA TÉCNICA DE AGRUPAMIENTO. 24
3.3 VISUALIZACIÓN DE REGLAS DE ASOCIACIÓN 30
4. ANÁLISIS DE DATOS OLAP 35
4.1 SELECCIONAR EL CUBO 35
4.2 REALIZAR LA CONEXIÓN CON POSTGRESQL 37
4.3 ANÁLISIS OLAP 39

INDICE DE FIGURAS

Pág.

Figura 1. Inicio de VisualBITool. 6
Figura 2. Explorador de archivos. 7
Figura 3. Mensaje de éxito. 7
Figura 4. Opción Atributos. 8
Figura 5. Lista de atributos con sus respectivas visualizaciones 8
Figura 6. Gráfico de tabla datos. 9
Figura 7. Gráfico de Área. 10
Figura 8. Gráfico de barras. 10
Figura 9. Gráfico de barras. 11
Figura 10. Opción Datos. 12
Figura 11. Interfaz para el proceso de distribución de datos 12
Figura 12. Gráfico de dispersión. 13
Figura 13. Gráfico de dispersión. 14
Figura 14. Datos de una Instancia. 14
Figura 15. Menú principal de VisualBITool 15
Figura 16. Interfaz para el módulo de Filtros 15
Figura 17. Interfaz para la eliminación de atributos 16
Figura 18. Nuevo archivo con la eliminación de los atributos Na y K 17
Figura 19. Interfaz para la discretización de atributos. 17
Figura 20. Selección de los atributos que serán discretizados 18
Figura 21. Mensaje de éxito al discretizar los atributos 18
Figura 22. Nuevo archivo con los atributos discretizados 19
Figura 23**.** Opción Árboles de decisión. 20
Figura 24. Interfaz de parámetros para los Árboles de decisión. 21
Figura 25. Tabla para editar los parámetros del árbol de decisión 21
Figura 26. Gráficas para representar un árbol de decisión 22
Figura 27. Gráfico Árbol Jerárquico. 23
Figura 28. Árbol de Sangría. 24
Figura 29. Opción agrupamiento. 25
Figura 30. Interfaz agrupamiento. 25
Figura 31. Interfaz de usuario para la tarea de agrupamiento. 26
Figura 32. Gráfico de Burbujas. 27
Figura 33. Información general de una instancia. 27
Figura 34. Información detallada de una instancia. 28
Figura 35. Gráfico Mapa de árbol 28
Figura 36. Expansión de grupo. 29
Figura 37. Tabla de grupos. 30
Figura 38. Eliminar Atributos. 31
Figura 39. Opción Reglas de asociación. 32

Figura 40. Interfaz para la tarea de Reglas de asociación. 32
Figura 41. Interfaz para seleccionar Reglas de asociación. 33
Figura 42. Gráfico de red. 34
Figura 43. Interfaz para seleccionar un archivo. 35
Figura 44. Navegador de archivos 36
Figura 45. Archivo XML Schema_uden_cantidad seleccionado para ser subido 36
Figura 46. Mensaje de éxito al subir el archivo 37
Figura 47. Menú principal de VisualBITool 37
Figura 48. Gestor de pestañas de VisualBITool 38
Figura 49. Ventana para editar los parámetros de la conexión 38
Figura 50. Mensaje de éxito en la conexión 39
Figura 51. Menú principal de VisualBITool 40
Figura 52. Gestor de pestañas de VisualBITool 40
Figura 53. Interfaz para análisis OLAP de VisualBITool 41
Figura 54. Ventana de gestión de un cubo OLAP 41
Figura 55. Ventana de selección de las partes de un cubo OLAP 42
Figura 56. Imagen de un cubo con las dimensiones seleccionadas 43
Figura 57. Imagen de un cubo con las tres dimensiones seleccionadas 43
Figura 58. Nivel y filtros de una dimensión 44
Figura 59. Ventana con los datos de un nivel para realizar un filtro 44
Figura 60. Resultado de una consulta 45
Figura 61. Reporte en Excel de la consulta realizada 46
Figura 62. Estructura de una consulta 46

1. INTRODUCCIÓN A VISUALBITOOL

VISUALBITOOL es un software de apoyo a la toma de decisiones, mediante la visualización de datos. Esta herramienta permite generar diferentes tipos de gráficos dependiendo de la tarea que se desea realizar. Las tareas que aplica este software son: Calidad de Datos, Minería de Datos como árboles de decisión, agrupamiento, reglas de asociación y también el análisis OLAP, que es la forma de visualizar los datos desde diferentes perspectivas mediante un cubo.

1.2 REQUERIMIENTOS

Para el correcto funcionamiento de la herramienta se necesita tener instalado el siguiente software:

- PostgreSQL 8.3 o superior
- La Máquina Virtual de Java jre6 o superior
- Tomcat superior a 6.0.20 o inferior a 7.
- Archivo VisualBiTool.war

1.3 INSTALACIÓN.

Para esto es necesario el archivo VisualBITool.war y seguir los siguientes pasos:

- Localizar el directorio webapps en los directorios de instalación de Tomcat y copiar el archivo **VisualBITool.war**
- Ejecutar Tomcat.
- Abrir un navegador y colocar la siguiente URL: http:// localhost: 8080/ VisualBITool/index.jsp
- Finalmente se ejecuta la aplicación.

2. ENTORNO DE LA HERRAMIENTA CON EJEMPLO PASO A PASO

2.1 ¿CÓMO INICIAR VISUALBITOOL?

Para iniciar simplemente se ejecuta en el navegador la siguiente URL: http://localhost:8080/VisualBiTool/index.jsp , como se indica en la figura 1.



Figura 1. Inicio de VisualBITool.

2.2 CARGAR ARCHIVO DE DATOS.

Se utilizará un archivo de datos de nombre ***Drug*** y extensión ***ARFF***, el cual posee información de personas que se les aplica un determinado fármaco dependiendo de ciertos atributos.

Se inicia haciendo clic sobre el botón '***Seleccionar archivo***', se abre un explorador de archivos y se busca el archivo anteriormente mencionado. Ver figura 2

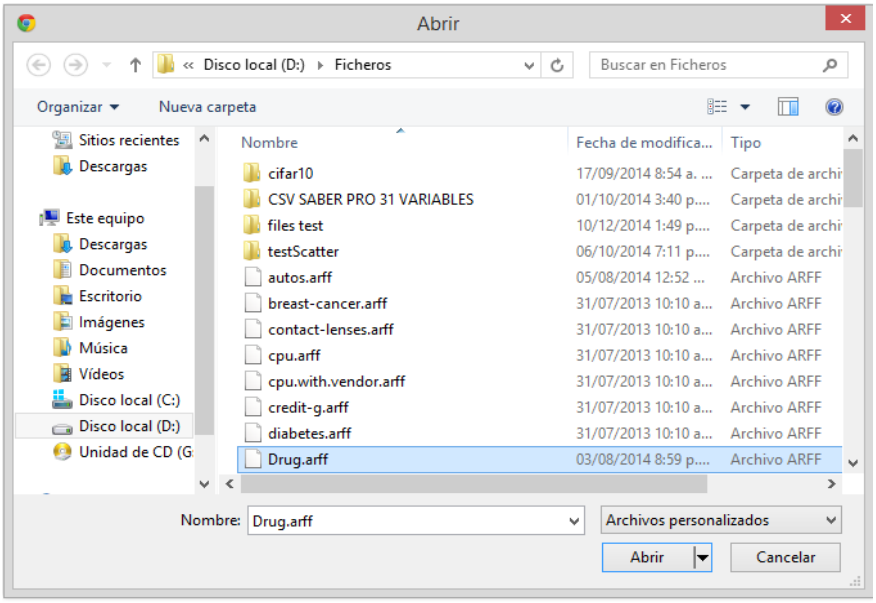


Figura 2. Explorador de archivos.

Se hace *clic* sobre el la imagen que representa una carpeta de archivo como se ve en la figura 1, para subir el archivo seleccionado y finalmente debe aparecer un mensaje de éxito. Ver figura 3

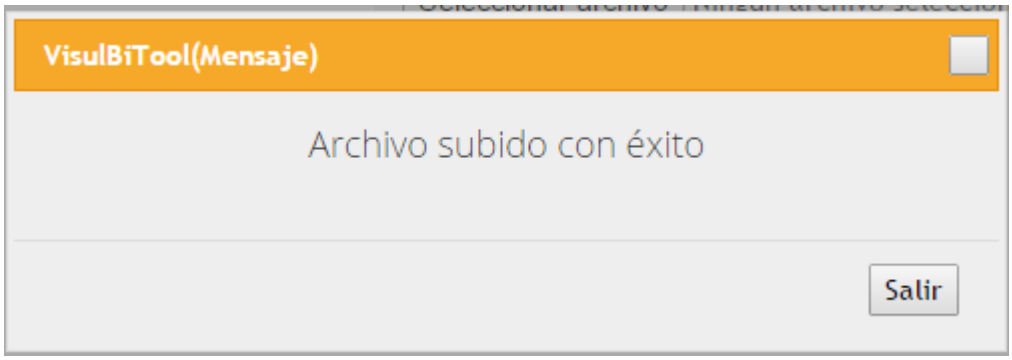


Figura 3. Mensaje de éxito.

Nota: Para realizar la tarea de Calidad de datos y Minería de datos solo se utiliza formato de archivos ARFF y CSV.

2.3 APLICAR CALIDAD DE DATOS.

En el menú Principal se busca la opción '***Calidad de Datos***' y se hace *clic* sobre la opción '***Atributos***', ver figura 4, y se obtendrá una lista de atributos con sus respectivas visualizaciones, como se muestra en la figura 5.

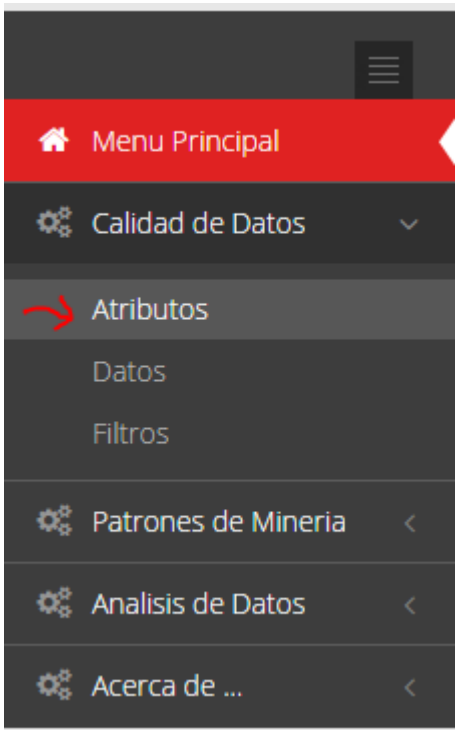


Figura 4. Opción Atributos.



Figura 5. Lista de atributos con sus respectivas visualizaciones

Nota: Si un atributo es de tipo numérico, se generan cuatro tipos de gráficos. (Tabla, área, barras y circulo) y si es categórico o nominal se generan tres (Tabla, barras y circulo).

Ahora al hacer *clic* sobre la imagen que representa una '***Tabla de datos***',ver figura 5, se genera una tabla con los respectivos valores de cada atributo. Si el atributo es numérico, se obtendrá rangos de la distribución de sus valores y variables estadísticas como el valor máximo, mínimo, la media y la desviación estándar y si es nominal sus diferentes valores categóricos, y además se permite descargar estos resultados. Ver figura 6

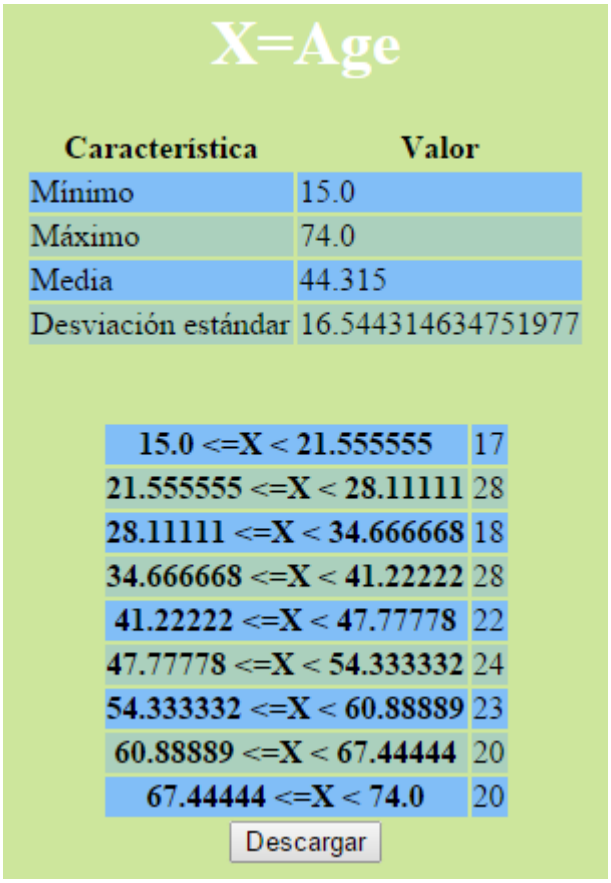


Figura 6. Gráfico de tabla datos.

Y al hacer clic sobre las restantes tres imágenes, ver figura 5, se obtendrá otros gráficos que representan la distribución de los valores de un atributo. Ver figuras 7,8 y 9.

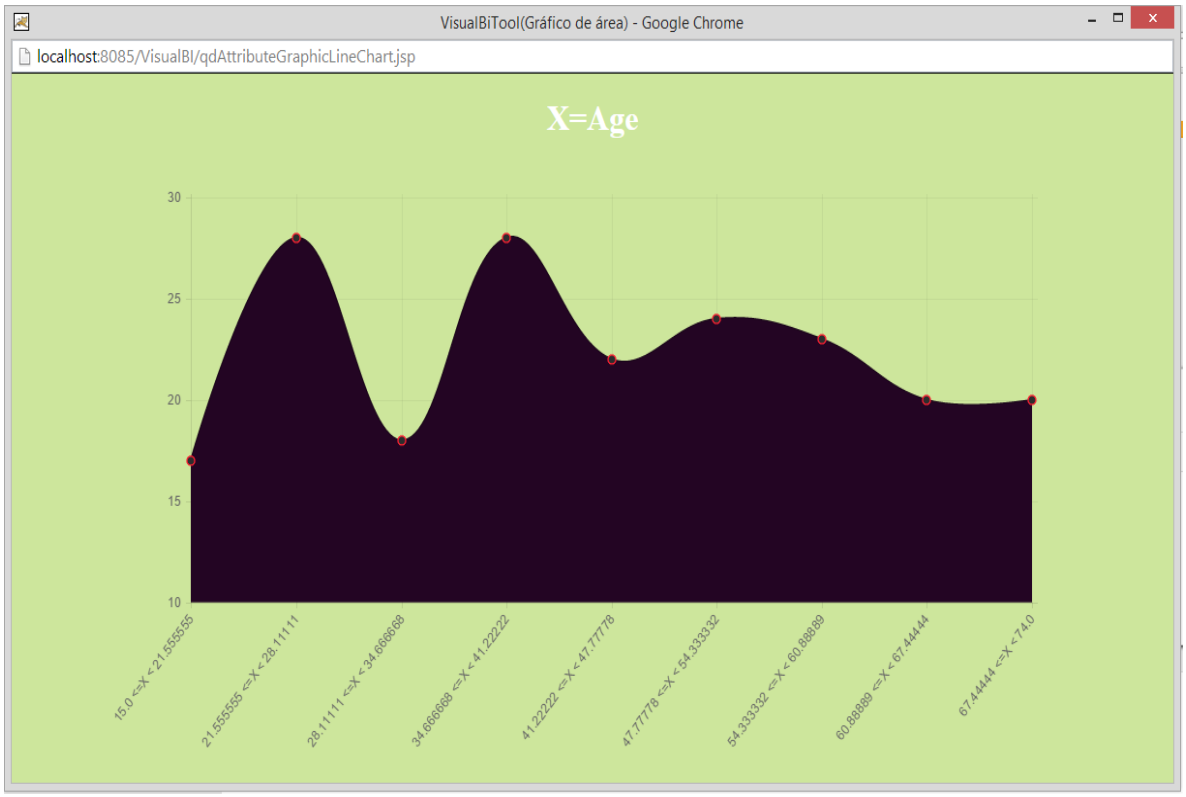


Figura 7. Gráfico de Área.

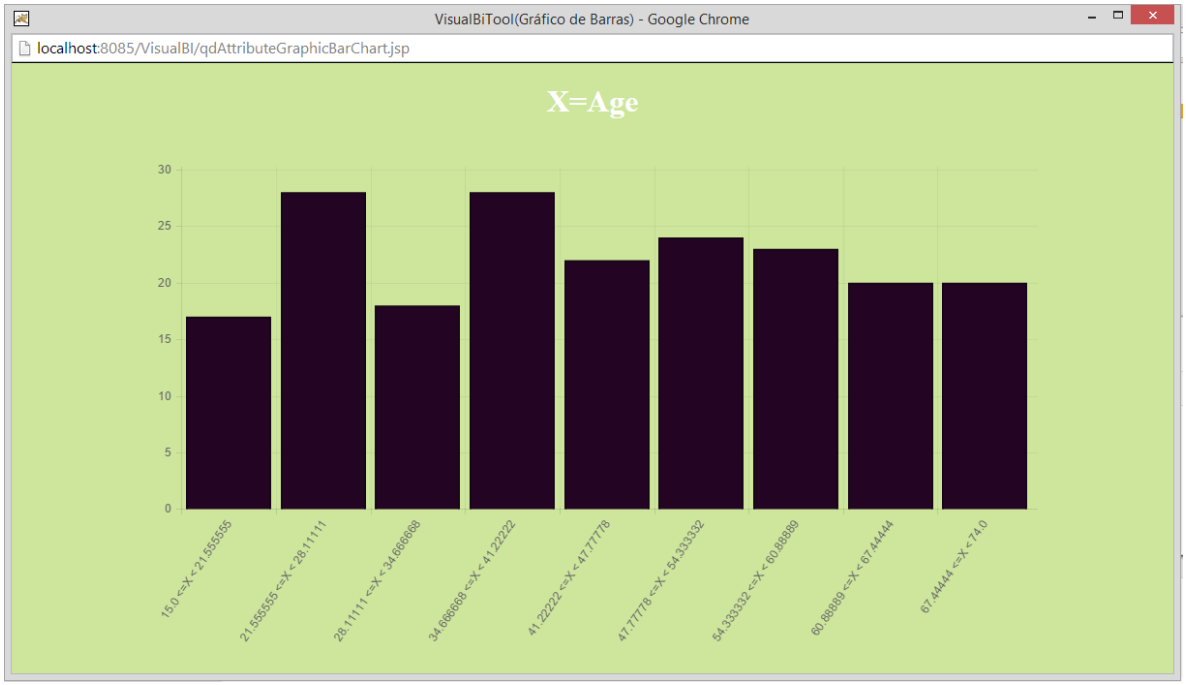


Figura 8. Gráfico de barras.

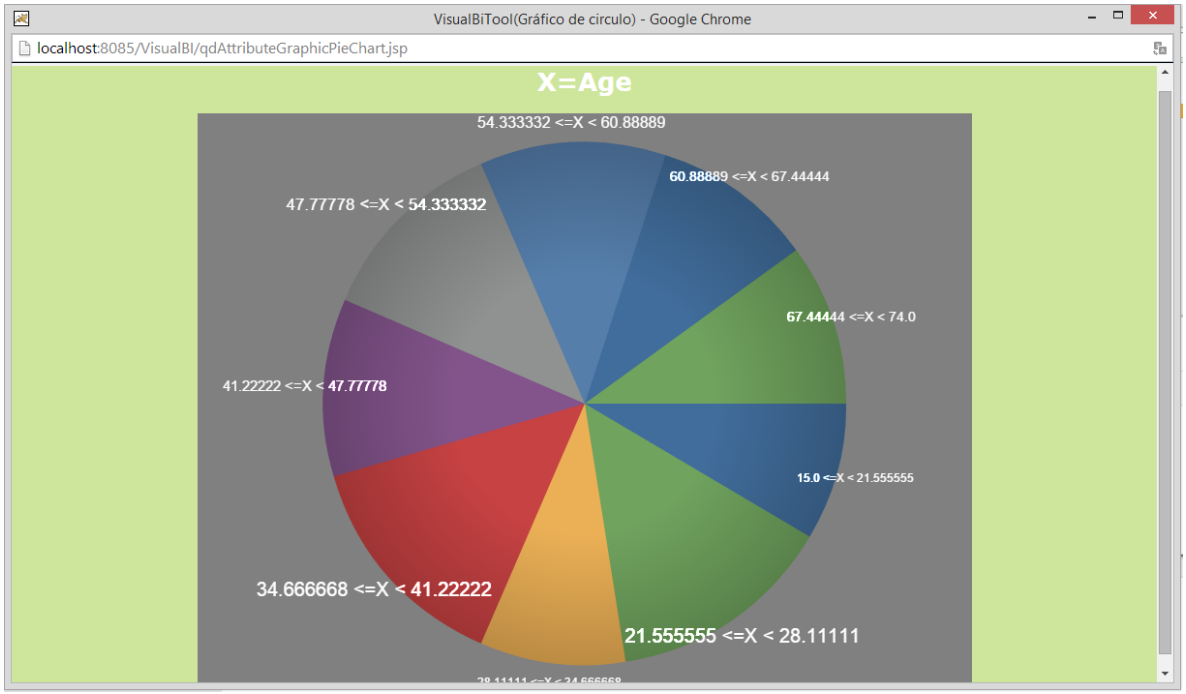


Figura 9. Gráfico de barras.

2.4 DISTRIBUCIÓN DE DATOS.

Ahora, se retorna al menú principal en la opción '**Calidad de datos**', ver figura 10, y se selecciona la opción '**Datos**'. Ver figura 11.

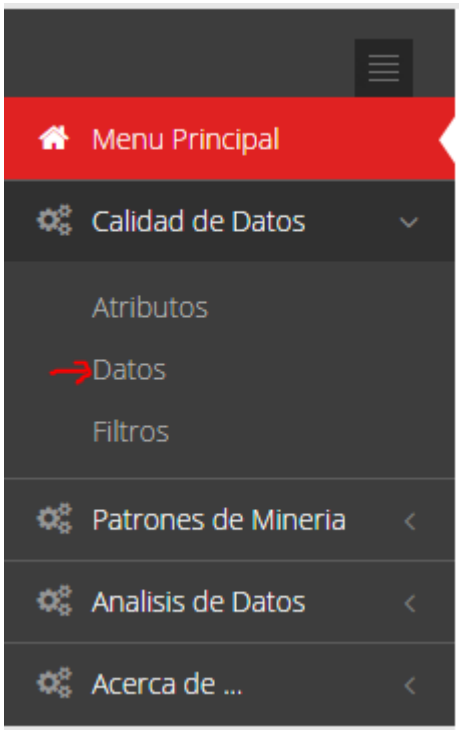


Figura 10. Opción Datos.



Figura 11. Interfaz para el proceso de distribución de datos

En esta sección se generará un gráfico de dispersión, con el fin de visualizar los datos (instancias) teniendo en cuenta el cruce de valores de tres atributos seleccionados.

Los campos '***Límite Inferior***' y '***Límite superior***' permiten fijar rangos de instancias para prevenir la sobrecarga de datos en un gráfico de dispersión. Los campos '***Atributo 1***', '***Atributo 2***' y '**Atributo 3**' permiten seleccionar los atributos a cruzar Ver figura 12.

En este ejemplo se seleccionará los atributos edad (Age), sodio (Na), fármaco (Drug) y se hace clic en el botón '***procesar***'. Ver Figura 12

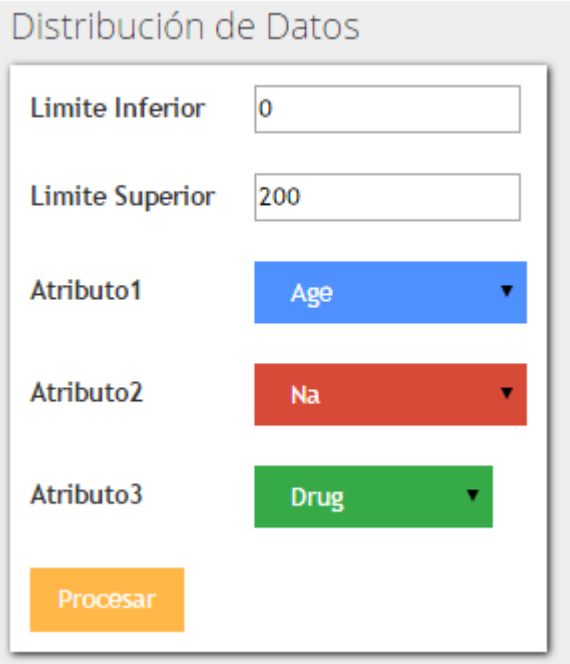


Figura 12. Gráfico de dispersión.

Al hacer clic en la imagen '**gráfico de dispersión**', ver figura 11,se obtiene el gráfico resultante del procesamiento de los atributos. Ver Figura 13.

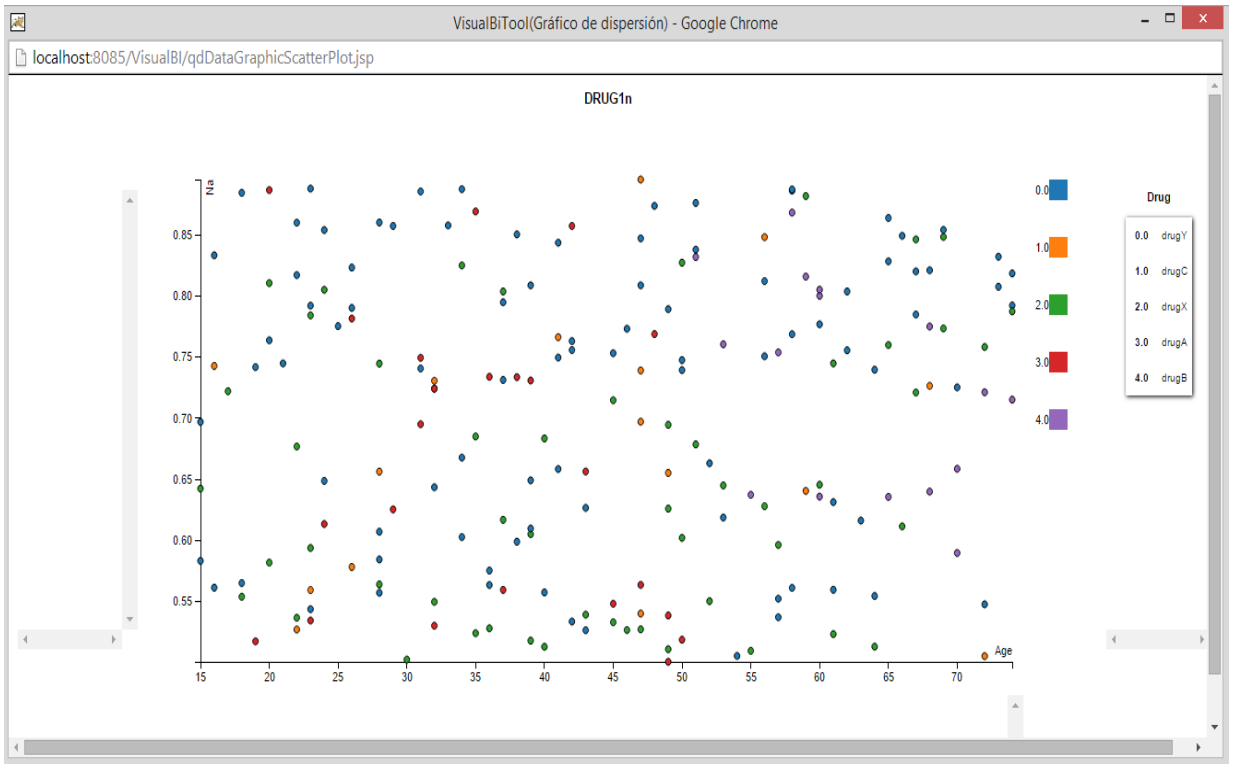


Figura 13. Gráfico de dispersión.

Al hacer clic sobre cada punto, se visualiza detalladamente mediante una tabla los valores de una instancia. Ver Figura 14



Figura 14. Datos de una Instancia.

2.5 FILTROS DE ATRIBUTOS.

Se selecciona del Menú Principal la opción ***Calidad de Datos/Filtros,*** como se ve en la figura 15.

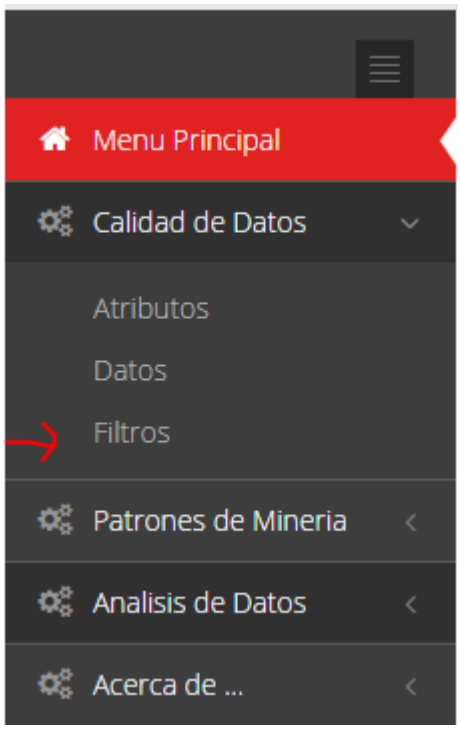


Figura 15. Menú principal de VisualBITool

La figura 16 muestra la interfaz para el proceso de filtrado.



Figura 16. Interfaz para el módulo de Filtros

En la tarea de filtrado, se selecciona los atributos necesarios para posteriormente ser aplicados en algunas técnicas de minería de datos que más adelante se trataran.se presentan tres botones **Todo, Ninguno** Y **Eliminar**, como se ve en la figura 17, El botón '***Todo***' permite chequear todas las opciones, El botón '***Ninguno***' permite des chequear las opciones seleccionadas, El botón '***Eliminar***' suprime los atributos seleccionados, ver figura 17.



Figura 17. Interfaz para la eliminación de atributos

Al seleccionar los atributos deseados y realizar alguna tarea con alguno de los tres botones, este caso se eliminan los atributos sodio (Na) y potasio (K), ver figura 17, se genera un nuevo archivo que se puede descargar con la imagen que representa la figura de una flecha, ver figura 17. El nuevo archivo se puede ver en la figura 18. Nótese que los atributos sodio (Na) y potasio (K) ya no se encuentran.

```
@relation DRUG1n

@attribute Age numeric
@attribute Sex {F,M}
@attribute BP {HIGH,LOW,NORMAL}
@attribute Cholesterol {HIGH,NORMAL}
@attribute Drug {drugY,drugC,drugX,drugA,drugB}

@data
23,F,HIGH,HIGH,drugY
47,M,LOW,HIGH,drugC
47,M,LOW,HIGH,drugC
28,F,NORMAL,HIGH,drugX
61,F,LOW,HIGH,drugY
22,F,NORMAL,HIGH,drugX
49,F,NORMAL,HIGH,drugY
41,M,LOW,HIGH,drugC
60,M,NORMAL,HIGH,drugY
43,M,LOW,NORMAL,drugY
47,F,LOW,HIGH,drugC
```

Figura 18. Nuevo archivo con la eliminación de los atributos Na y K

Ahora en la parte de discretización como se ve en la figura 19. Se indica los atributos que son de tipo numérico en este caso edad (Age), sodio (Na) y potasio (K), que se deben discretizar.



Figura 19. Interfaz para la discretización de atributos

En la ventana de discretización se presenta un campo denominado “Número de Bins”, ver figura 19, el cual es el número de intervalos con los cuales se pretende discretizar un atributo de tipo numérico, en este caso el número será cinco y discretizará los atributos edad (Age) y sodio (Na). Ver figura 20, A continuación se da clic en el botón discretizar. Ver figura 21.



Figura 20. Selección de los atributos que serán discretizados

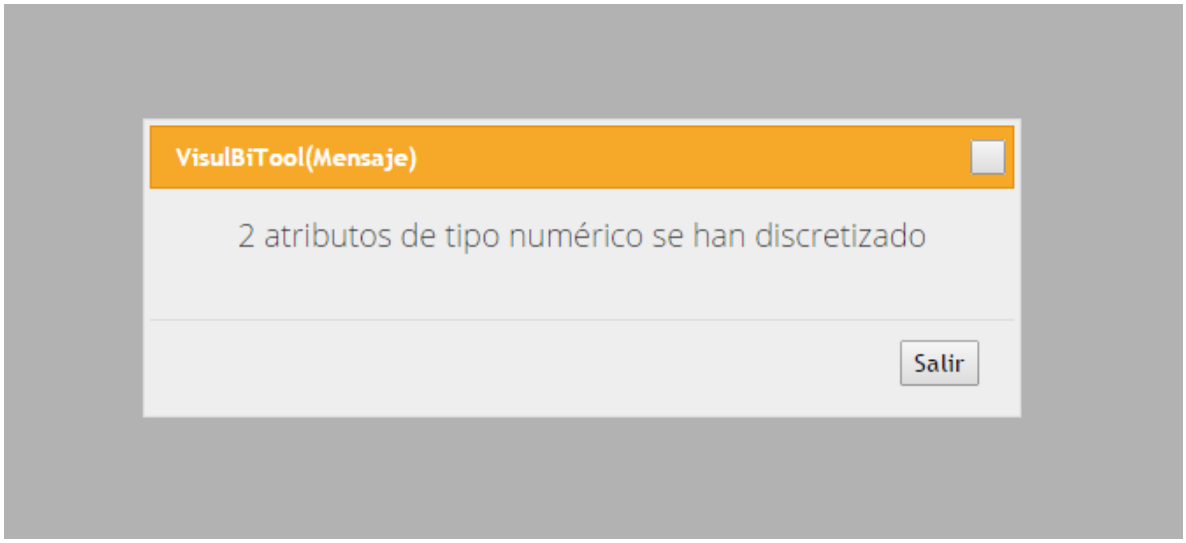


Figura 21. Mensaje de éxito al discretizar los atributos

Ahora se puede descargar el nuevo archivo, que se ve en la figura 22, con los dos atributos discretizados (nótese que los atributos se discretizaron en cinco intervalos), dando clic en la imagen que representa una fecha. Ver figura 20.

```
@relation 'DRUG1n-weka.filters.unsupervised.attribute.Discretize-unset-class-temporarily-B5-M-1.0-R1,5'

@attribute Age {'-inf-26.8','26.8-38.6','38.6-50.4','50.4-62.2','62.2-inf'}
@attribute Sex {F,M}
@attribute BP {HIGH,LOW,NORMAL}
@attribute Cholesterol {HIGH,NORMAL}
@attribute Na {'-inf-0.579346','0.579346-0.658524','0.658524-0.737701','0.737701-0.816879','0.816879-inf'}
@attribute K numeric
@attribute Drug {drugY,drugC,drugX,drugA,drugB}

@data
'-inf-26.8',F,HIGH,HIGH,'0.737701-0.816879',0.031258,drugY
'38.6-50.4',M,LOW,HIGH,'0.737701-0.816879',0.056468,drugC
'38.6-50.4',M,LOW,HIGH,'0.658524-0.737701',0.068944,drugC
'26.8-38.6',F,NORMAL,HIGH,'-inf-0.579346',0.072289,drugX
'50.4-62.2',F,LOW,HIGH,'-inf-0.579346',0.030998,drugY
'-inf-26.8',F,NORMAL,HIGH,'0.658524-0.737701',0.078647,drugX
'38.6-50.4',F,NORMAL,HIGH,'0.737701-0.816879',0.048518,drugY
'38.6-50.4',M,LOW,HIGH,'0.737701-0.816879',0.069461,drugC
'50.4-62.2',M,NORMAL,HIGH,'0.737701-0.816879',0.05123,drugY
'38.6-50.4',M,LOW,NORMAL,'-inf-0.579346',0.027164,drugY
'38.6-50.4',F,LOW,HIGH,'0.816879-inf',0.076147,drugC
'26.8-38.6',F,HIGH,NORMAL,'0.658524-0.737701',0.034782,drugY
'38.6-50.4',M,LOW,HIGH,'0.579346-0.658524',0.040746,drugY
'62.2-inf',F,LOW,HIGH,'0.737701-0.816879',0.037851,drugY
'38.6-50.4',F,NORMAL,HIGH,'0.816879-inf',0.065166,drugX
'-inf-26.8',F,HIGH,NORMAL,'0.816879-inf',0.053742,drugY
'62.2-inf',M,LOW,NORMAL,'0.816879-inf',0.074111,drugX
'38.6-50.4',M,HIGH,HIGH,'0.579346-0.658524',0.046979,drugA
'-inf-26.8',M,LOW,HIGH,'-inf-0.579346',0.076609,drugC
'26.8-38.6',F,HIGH,NORMAL,'0.579346-0.658524',0.024773,drugY
'50.4-62.2',M,LOW,NORMAL,'-inf-0.579346',0.028061,drugY
'62.2-inf',M,NORMAL,HIGH,'0.579346-0.658524',0.023773,drugY
'38.6-50.4',M,LOW,NORMAL,'0.737701-0.816879',0.026472,drugY
'38.6-50.4',F,LOW,HIGH,'0.816879-inf',0.058155,drugY
'26.8-38.6',F,LOW,HIGH,'0.816879-inf',0.025634,drugY
'26.8-38.6',F,HIGH,NORMAL,'-inf-0.579346',0.029604,drugY
```

Figura 22. Nuevo archivo con los atributos discretizados

3. APLICAR PATRONES DE MINERÍA.

3.1 VISUALIZACIÓN DE UN ÁRBOL DE DECISIÓN.

En el menú principal y en el módulo de '***Patrones de minería***' se selecciona la opción '***árboles de decisión***'. Ver figura 23

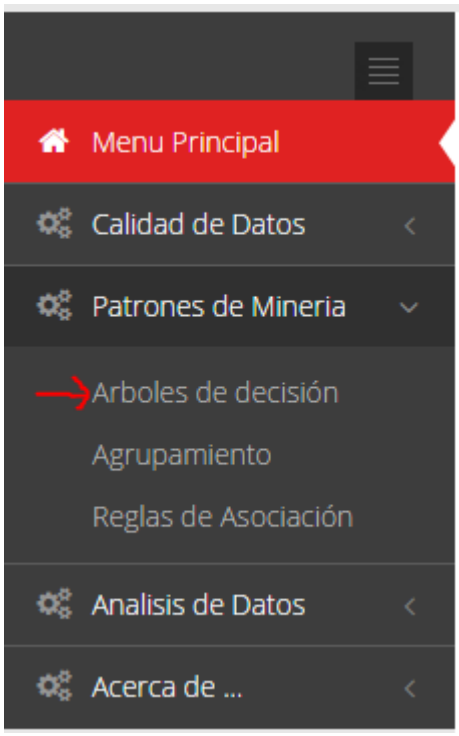


Figura 23. Opción Árboles de decisión.

En la figura 24 se presentan unos parámetros para la construcción del árbol de decisión éstos parámetros son:

Factor de confianza: Nivel de precisión del árbol de decisión.

Número mínimo de objetos: Cantidad mínima de instancias con las que se creara el árbol.

Número de pliegues: Cantidad de ramificaciones del árbol para reducir el error en el proceso de poda del árbol y evitar el sobre ajuste de información.

Semilla: Número aleatorio con el cual deseo iniciar el proceso de poda para reducir su error.

Atributo clase: Atributo con el cual concluyo una regla de clasificación en un árbol de decisión. Son los valores que se ubicaran en las hojas del árbol.

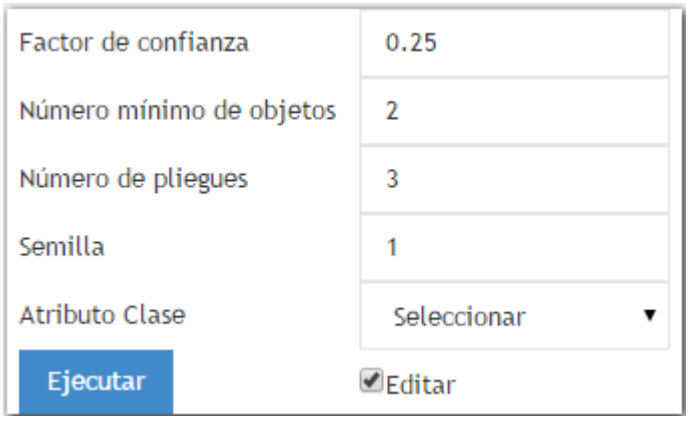


Figura 24. Interfaz de parámetros para los Árboles de decisión.

Para la construcción de un árbol de decisión simplemente se presiona el botón '***Ejecutar***' y si es necesario personalizar algunos de sus parámetros, se chequea la opción '***Editar***' para poder modificar los valores por defecto. Ver Figura 25

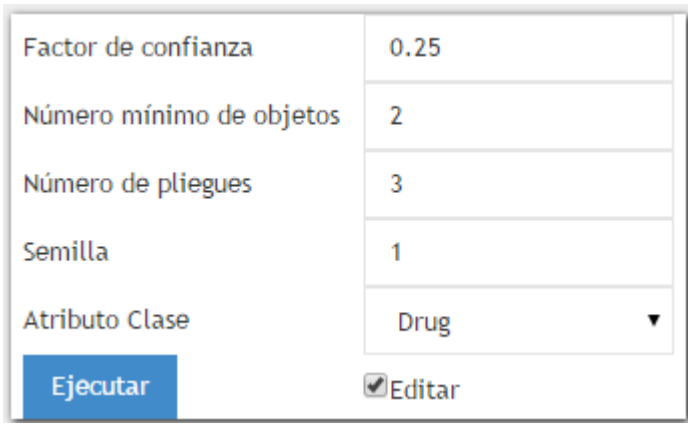


Figura 25. Tabla para editar los parámetros del árbol de decisión

La figura 26 muestra la interfaz de usuario para la tarea de clasificación por árboles de decisión con las diferentes gráficas para representar un árbol de decisión.

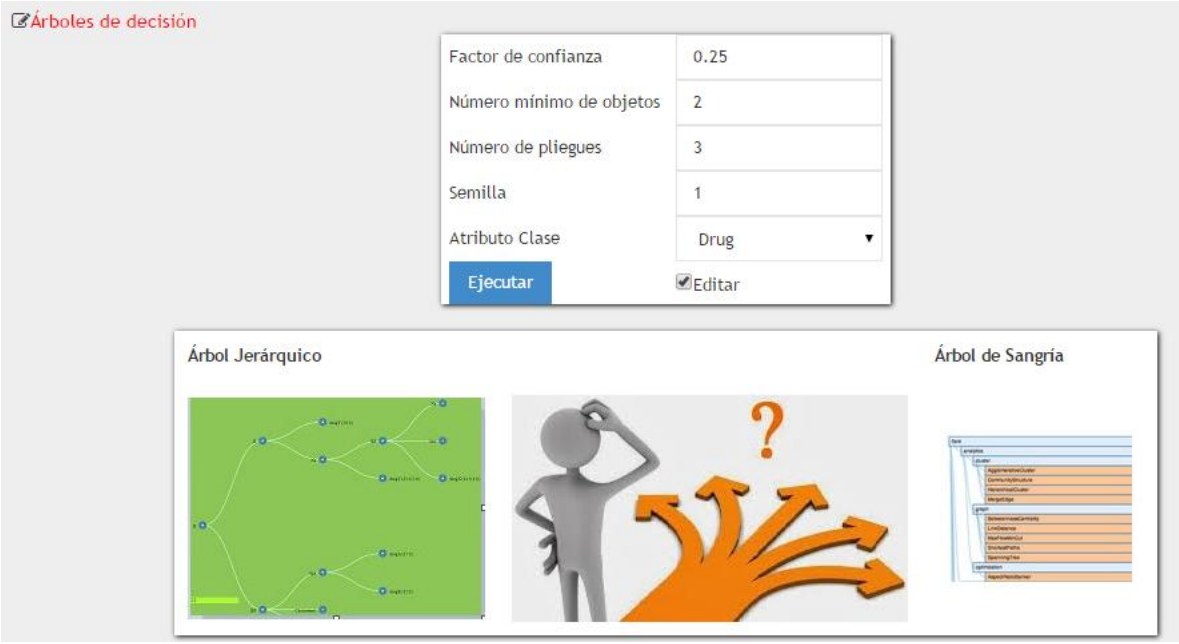


Figura 26. Gráficas para representar un árbol de decisión

Finalmente se genera el árbol de decisión con dos formas de Visualización '***Árbol Jerárquico***' y '***Árbol de Sangría***'. Para observar el árbol generado simplemente se debe hacer '***clic***' sobre cada imagen. Ver figura 26.

La figura 27 muestra un árbol jerárquico donde cada nodo tiene dos eventos. El primero evento se refiere a deslizar el '**Mouse**' sobre el nodo y se visualizara una ventana de color '***Amarillo verdoso***' que contiene un valor determinado del nodo padre. El segundo evento se aplica a aquellos nodos que poseen un punto rojo en su centro, que al presionar sobre él, se abren más ramas.

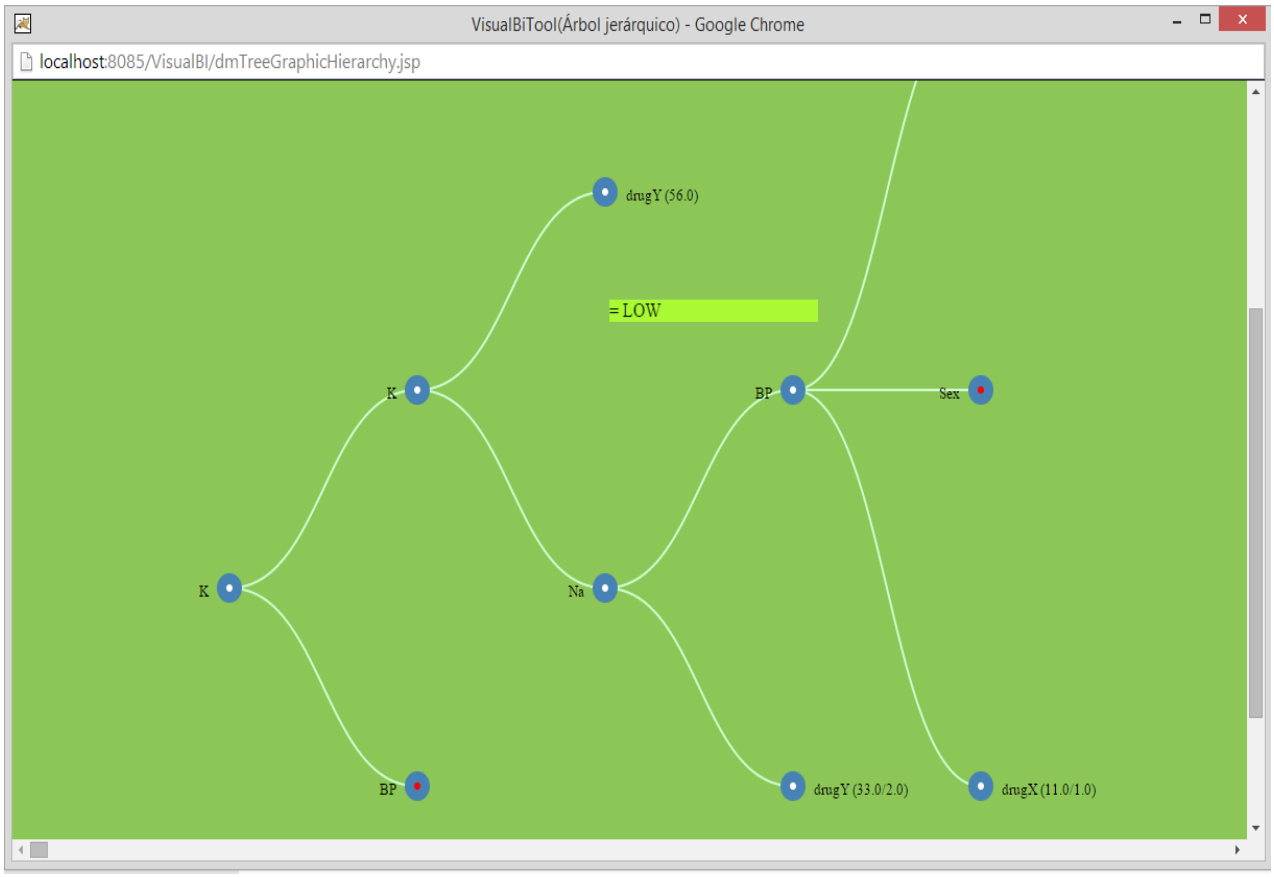


Figura 27. Gráfico Árbol Jerárquico.

La figura 28 muestra un árbol de sangría, donde los valores de cada nodo se encuentran a la izquierda y los nodos hijos se encuentran a la derecha. Al presionar sobre cada barra de color '***azul***' se despliega más ramas del nodo, hasta que finalmente se llegue a las barras de color '***rosa***' que son las hojas del árbol.

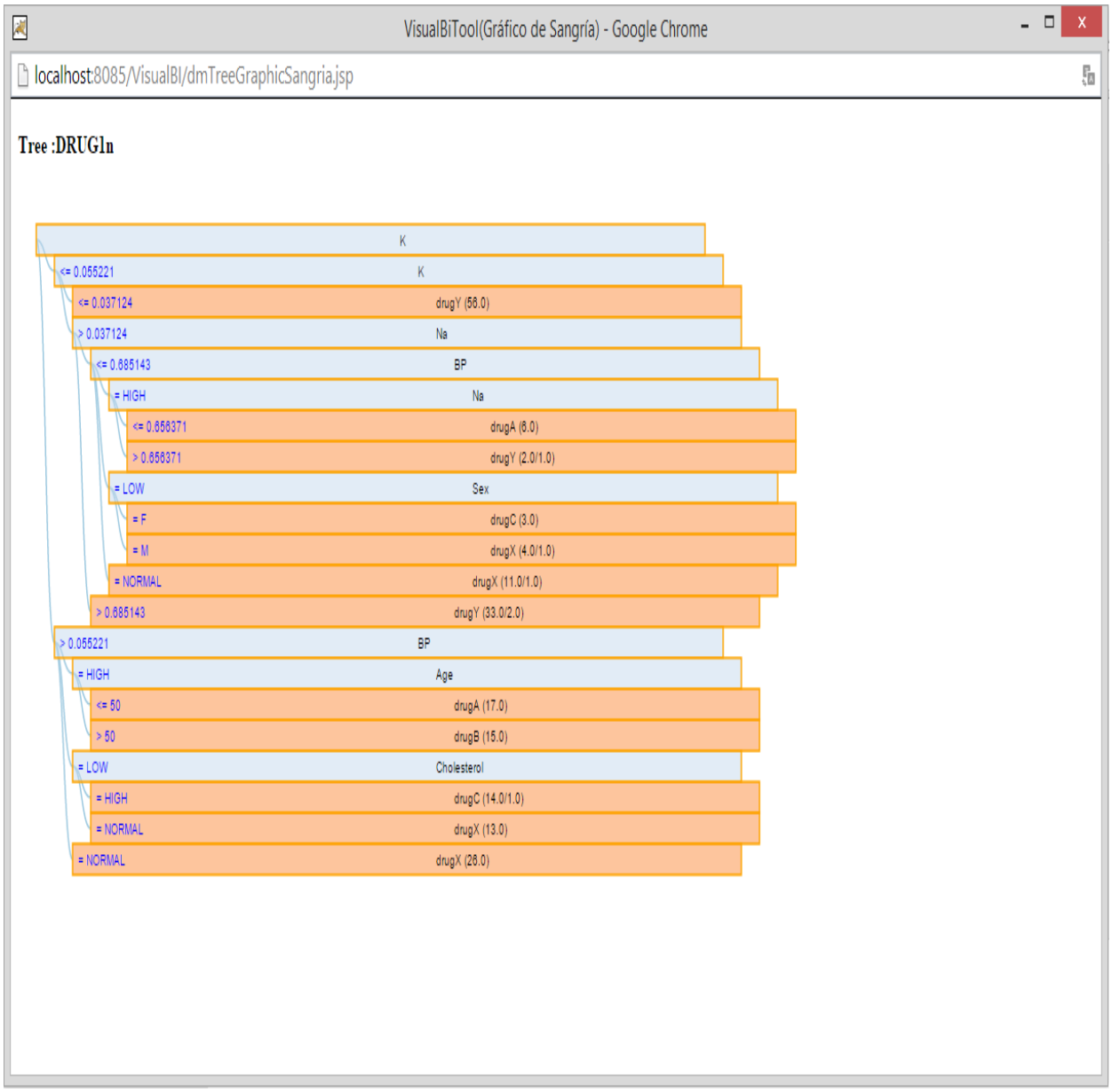


Figura 28. Árbol de Sangría.

3.2 VISUALIZACIÓN DE LA TÉCNICA DE AGRUPAMIENTO.

En el menú principal en el módulo de '***Patrones de minería***' se selecciona la opción '***Agrupamiento***'. Ver figura 29.

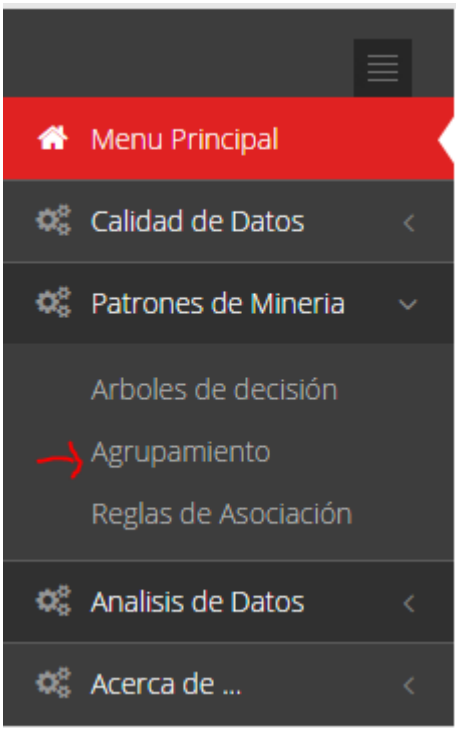


Figura 29. Opción agrupamiento.

En la figura 30 muestra un campo de tipo numérico, en el cual se debe ingresar la cantidad de grupos que se desean ver, en este caso se ingresan seis.

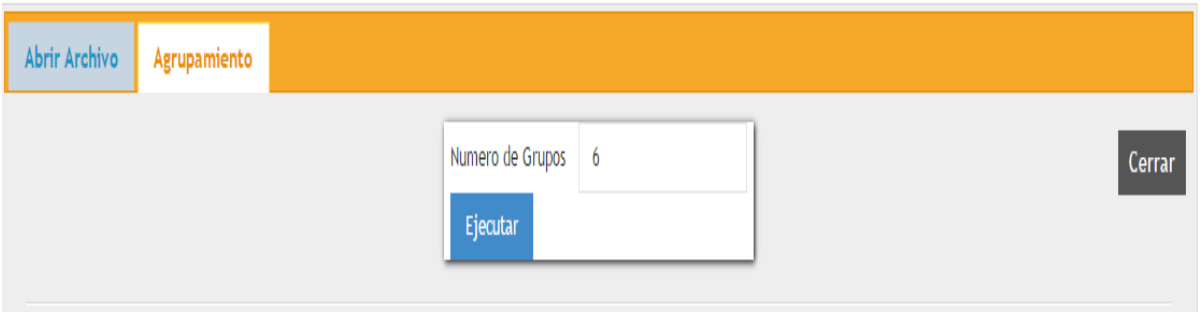


Figura 30. Interfaz agrupamiento.

Luego se da clic en el botón '***Ejecutar***', ver figura 30, para poder ingresar a la interfaz de usuario para la tarea de agrupamiento. Ver figura 31.



Figura 31. Interfaz de usuario para la tarea de agrupamiento.

Antes de presionar sobre cada imagen y efectuar la visualización, se debe ingresar los parámetros correspondientes a un intervalo de instancias a visualizar el límite inferior y superior además del campo con el número de instancias que se desean ver en cada grupo (cluster) y luego se da clic sobre el botón '***Procesar Instancias***'. Ver figura 31.

Y ahora ya se puede presionar sobre cada imagen '***Gráfico de Burbujas'*** y '***Mapa de árbol***'. Ver figura 31.

La figura 32 permite efectuar dos eventos sobre cada punto que representa a cada instancia, los cuales son:

El evento '***Deslizar el mouse***' que muestra una ventana con información que corresponde al **índice**, **grupo y distancia** de la instancia con respecto a su centro o centroide. (Ver Figura 33)

El evento ´***Clic***´ que se efectúa sobre un punto del grupo (cluster) permite visualizar detalladamente los valores de una instancia y centro del grupo (centroide) a través de una tabla (Ver figura 34).

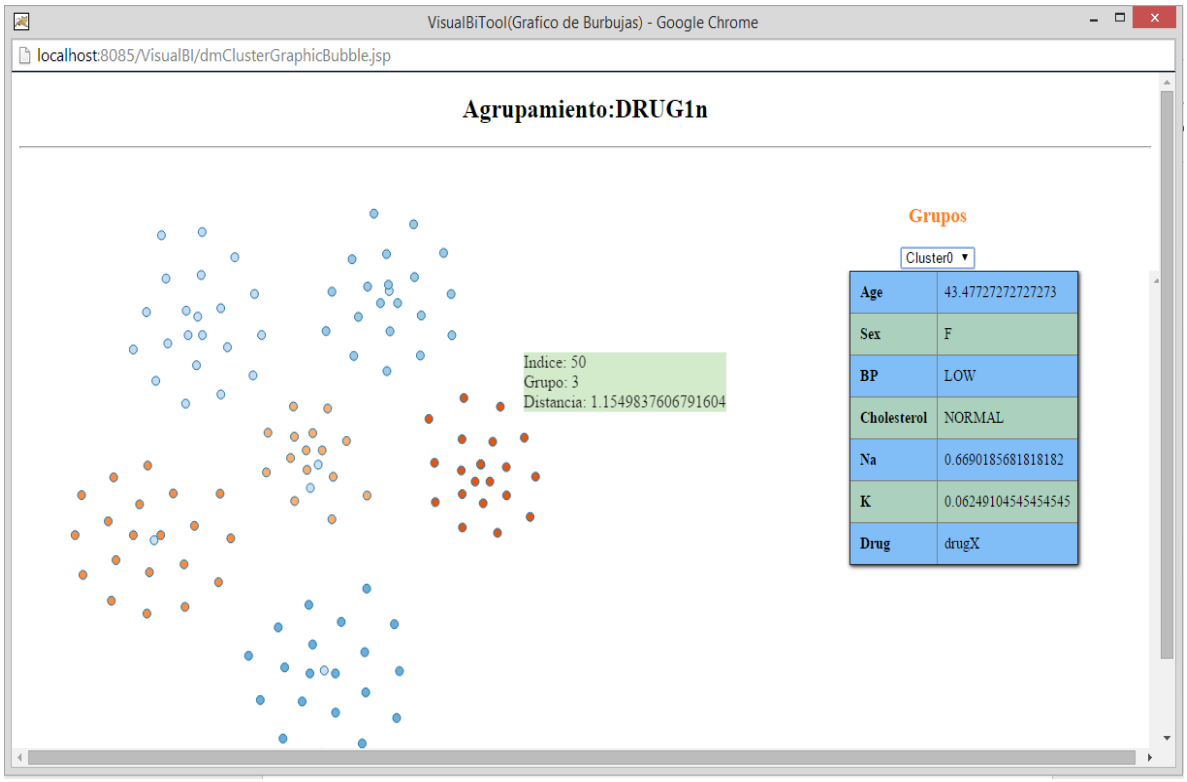


Figura 32. Gráfico de Burbujas.

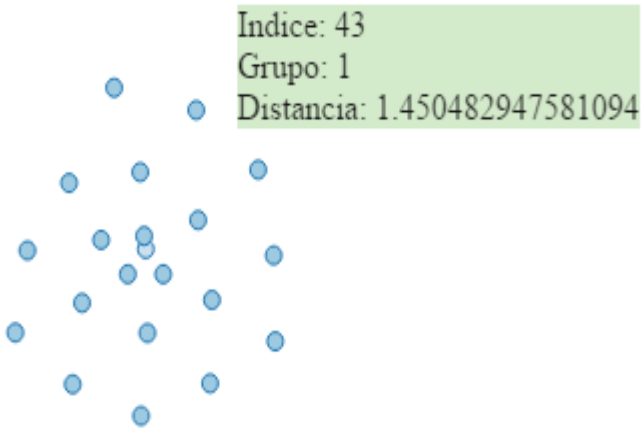


Figura 33. Información general de una instancia.



Figura 34. Información detallada de una instancia.

La figura 35 tiene la misma funcionalidad que la figura 32, de representar los grupos formados, solo que esta tiene un evento adicional el cual es hacer '**doble clic**' sobre el número de instancia para poder ver sus valores. Permitiendo expandir solo el grupo al cual pertenece. En este caso se aplicará a la instancia trece que se encuentra en el grupo 2. Ver Figura 36.

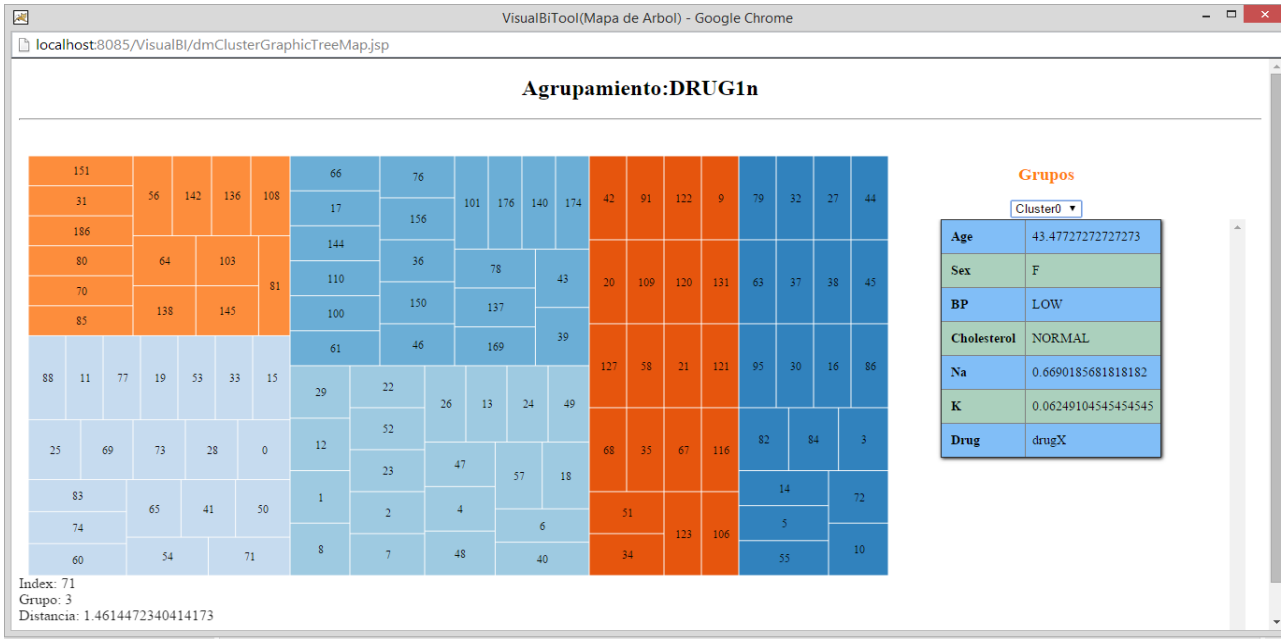


Figura 35. Gráfico Mapa de árbol

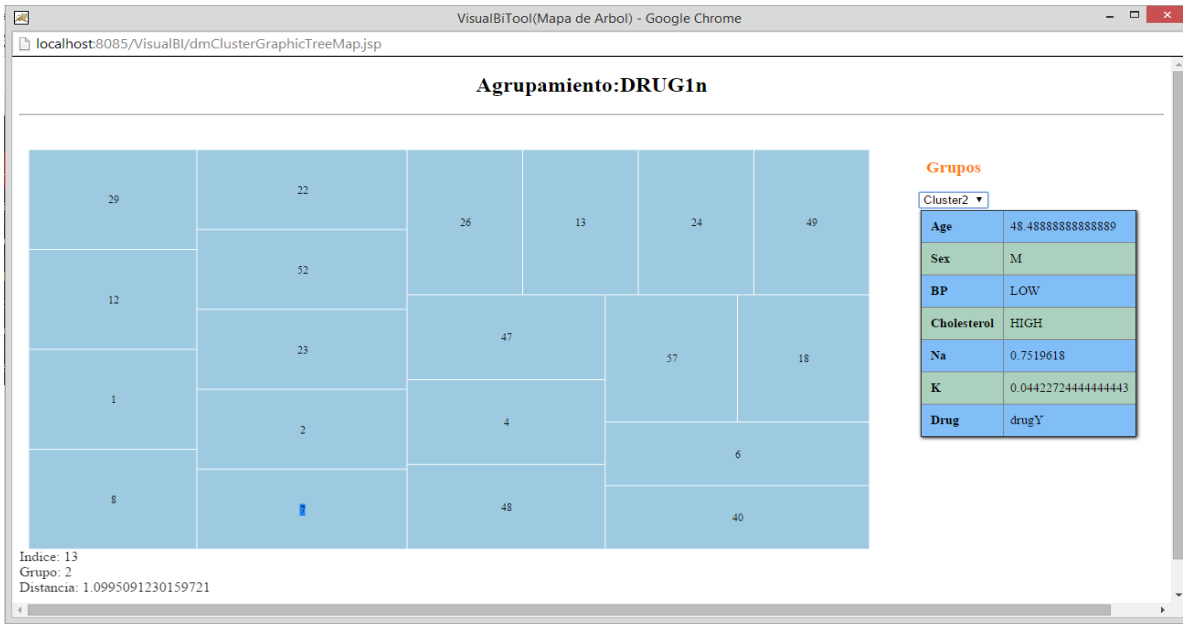
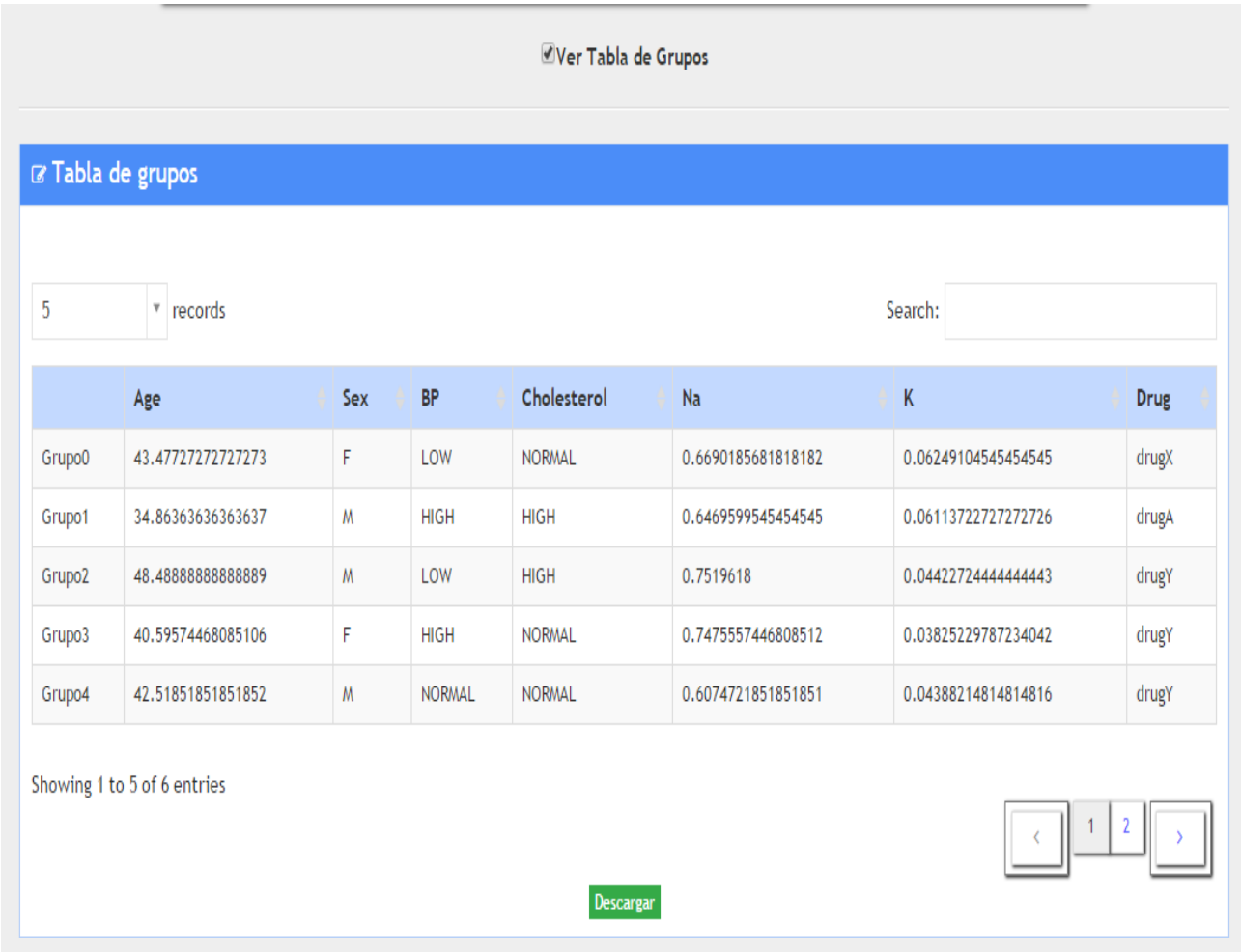


Figura 36. Expansión de grupo.

La figura 35 hereda los dos primeros eventos vistos en la figura 32.

Nota: El gráfico de burbujas y mapa de árbol distribuye las instancias en cada grupo dependiendo de la distancia euclidiana de una instancia con respecto al centro de su grupo (centroide). Las organiza de mayor a menor o viceversa.

En la parte inferior de las imágenes de visualización de grupos hay una opción de chequeo, la cual permite desplegar una tabla con los valores correspondientes a los grupos generados. Ver figura 37



	Age	Sex	BP	Cholesterol	Na	K	Drug
Grupo0	43.47727272727273	F	LOW	NORMAL	0.6690185681818182	0.06249104545454545	drugX
Grupo1	34.86363636363637	M	HIGH	HIGH	0.6469599545454545	0.06113722727272726	drugA
Grupo2	48.48888888888889	M	LOW	HIGH	0.7519618	0.04422724444444443	drugY
Grupo3	40.59574468085106	F	HIGH	NORMAL	0.7475557446808512	0.03825229787234042	drugY
Grupo4	42.51851851851852	M	NORMAL	NORMAL	0.6074721851851851	0.04388214814814816	drugY

Figura 37. Tabla de grupos.

Y finalmente al presionar sobre el botón '***Descargar***', que se encuentra en la parte inferior de la figura 37, permite almacenar esta información en el disco duro u otro dispositivo de almacenamiento.

3.3 VISUALIZACIÓN DE REGLAS DE ASOCIACIÓN

Antes de seleccionar la opción '***Reglas de asociación'***, verificar que los atributos sean los adecuados, o sea que no sean numéricos. En este caso se discretizará los atributos que son de tipo numérico edad (Age), sodio (Na) y potasio (K). Ver Figura 38



Figura 38. Eliminar Atributos.

En el menú principal y en el módulo de '***Patrones de minería***' se selecciona la opción '***Reglas de asociación*** ', Ver figuras 39, y así poder ingresar a la interfaz de usuario para esta tarea de minería de datos. Ver figura 40.

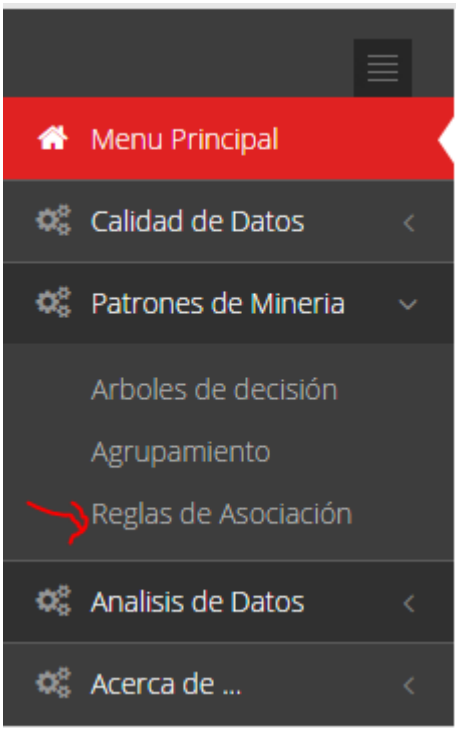


Figura 39. Opción Reglas de asociación.

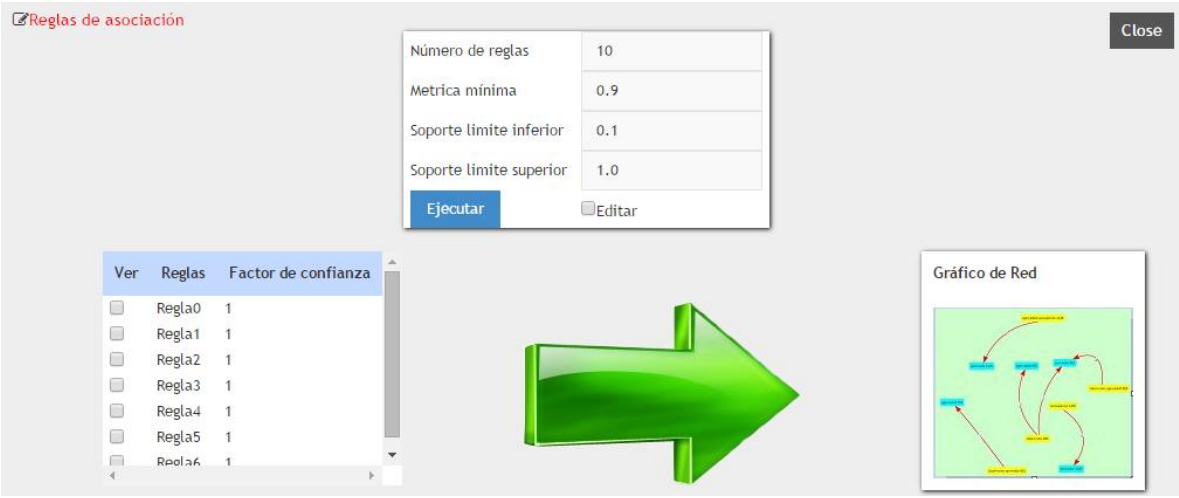


Figura 40. Interfaz para la tarea de Reglas de asociación.

En la figura 40 se presenta una tabla con los parámetros de una regla de asociación, los parámetros son:

Número de reglas: Cantidad de reglas que serán representadas.

Métrica mínima: Establecer el tipo de métricas mediante el cual se alinean las reglas.La confianzaes la proporción de los ejemplos cubiertos por la premisa de qu e también están cubiertos por la consecuencia.

La elevación es dividida por la proporción de todos los ejemplos que están cubiertos por la consecuencia de confianza. Esta es una medida de la importancia de la asociación independiente de apoyo.

Soporte límite inferior: Soporte mínimo de confianza para la búsqueda de reglas de asociación.

Soporte límite superior: Soporte máximo de confianza para la búsqueda de reglas de asociación.

Para la generación de reglas de asociación se presiona el botón '***Ejecutar***' y si es necesario personalizar algunos de sus parámetros se chequea la opción '***Editar***' para modificar los valores por defecto. Ver Figura 41

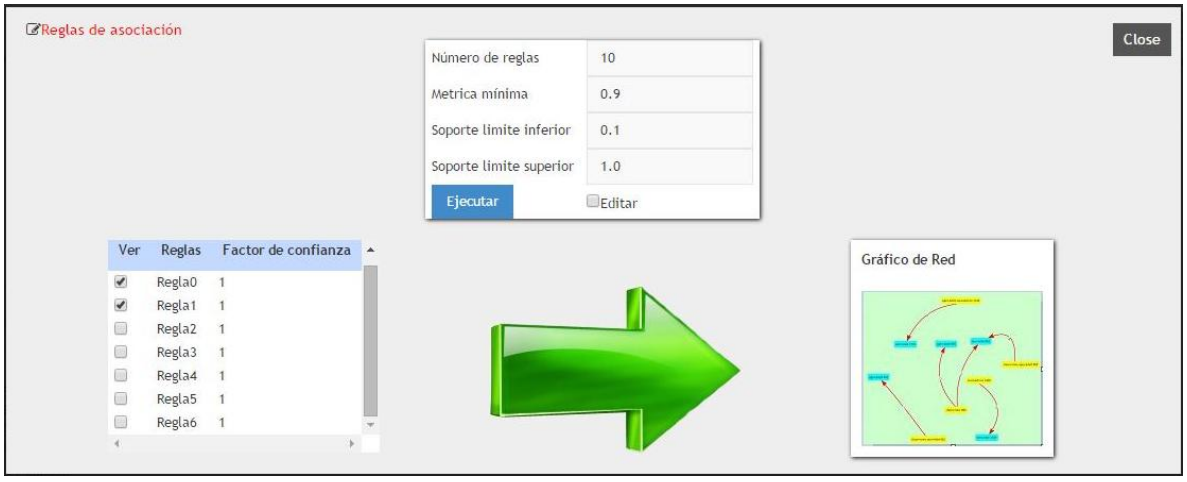


Figura 41. Interfaz para seleccionar Reglas de asociación.

En la figura 41, se crea una tabla con las reglas generadas. Para visualizarlas mediante un gráfico de red es necesario chequear por lo menos una y luego presionar sobre la imagen '***Gráfico de Red***'. Con esta gráfica se puede observar la estructura de cada regla. Ver figura 42

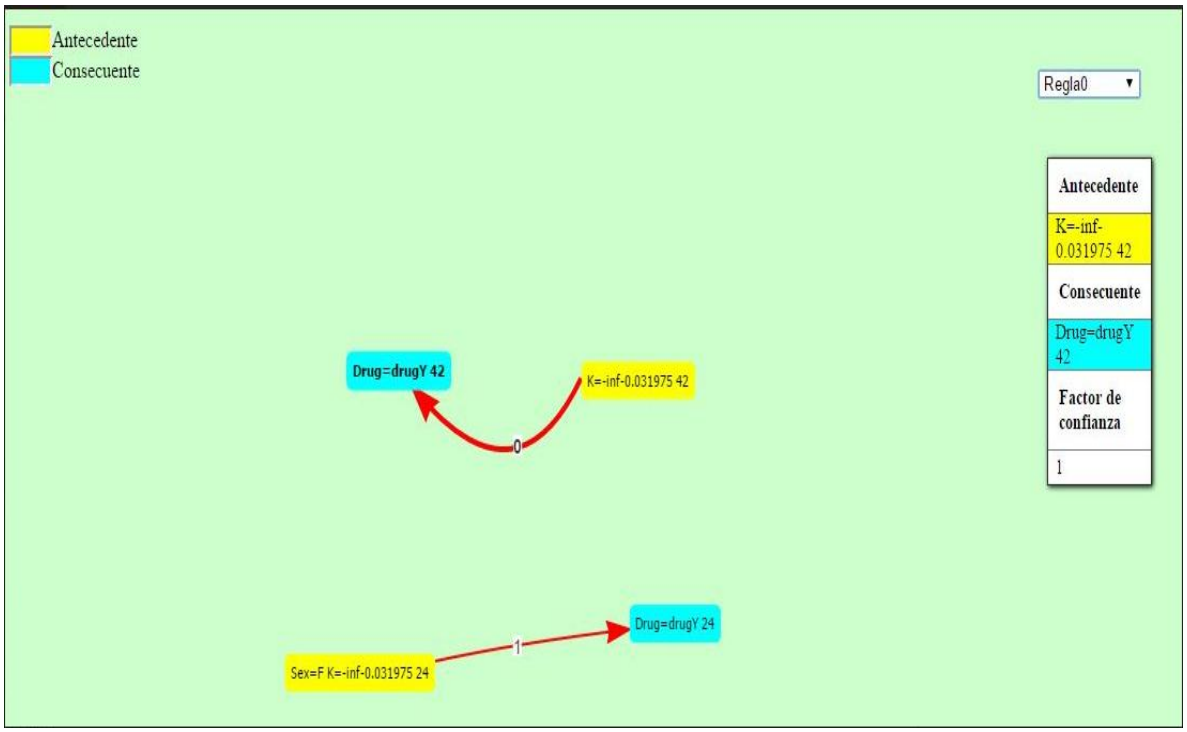


Figura 42. Gráfico de red.

4. ANÁLISIS DE DATOS OLAP

4.1 SELECCIONAR EL CUBO

El formato del archivo que será leído por el programa es XML, para seleccionar el archivo en el equipo, se da clic en el botón seleccionar archivo como se muestra en la figura 43.



Figura 43. Interfaz para seleccionar un archivo.

Se abre una ventana en la cual se puede buscar el archivo con el que se desea trabajar como se ve en la figura 44.

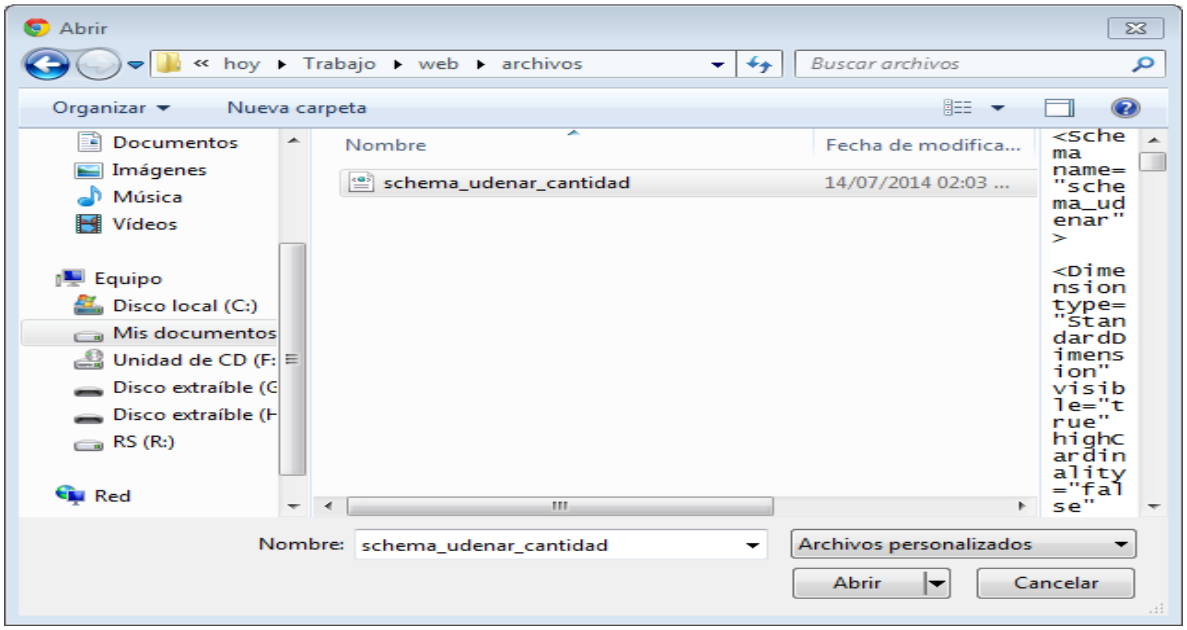


Figura 44. Navegador de archivos

Se hace clic en el botón abrir, y el archivo aparecera frente al botón seleccionar archivo como se ve en la figura 45. y se da clic en la figura que representa una carpeta de archivo con una flecha, para que el archivo quede listo para ser navegado por la herramienta.



Figura 45. Archivo XML Schema_uden_cantidad seleccionado para ser subido

Si el archivo se ha subido sin ningún problema aparecerá un mensaje de éxito, ver figura 46, y se puede continuar con el proceso.

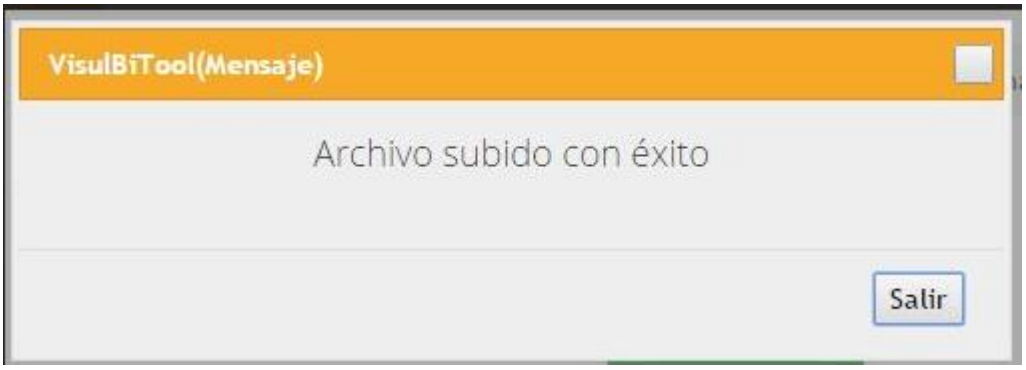


Figura 46. Mensaje de éxito al subir el archivo

4.2 REALIZAR LA CONEXIÓN CON POSTGRESQL

Luego en el menú principal/análisis de datos/ se selecciona la opción conexión OLAP como se muestra en la figura 47.

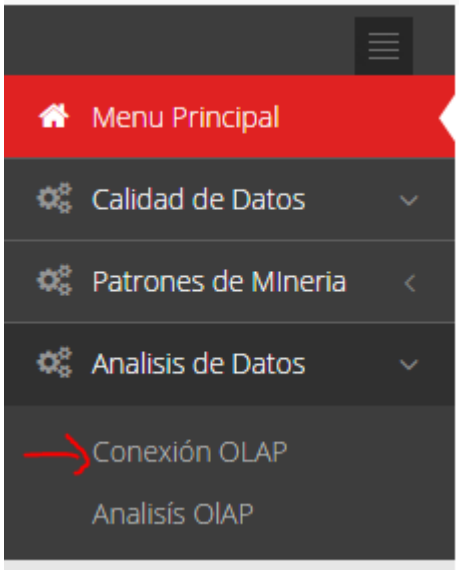


Figura 47. Menú principal de VisualBITool

En la parte superior se abre una pestaña , la cual se debe seleccionar para ingresar a la interfaz de usuario para el análisis OLAP. Ver figura 48.

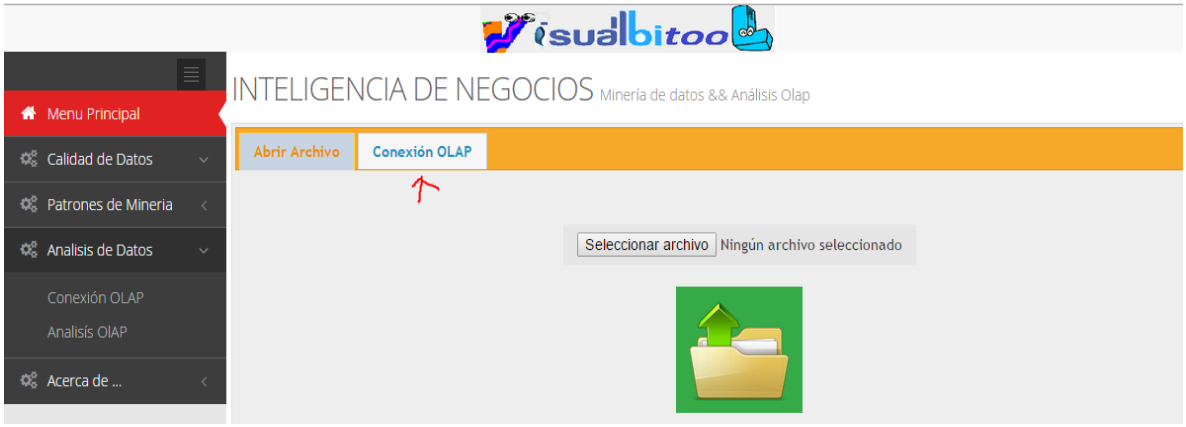


Figura 48. Gestor de pestañas de VisualBITool

Luego se abre una ventana en la cual se puede editar los parámetros de conexión con la bodega de datos en PostgreSQL. Ver figura 49.



Figura 49. Ventana para editar los parámetros de la conexión

Los parametros de la figura 49, que se deben editar de acuerdo al tipo de conexión que se dese realizar son:

Esquema: Es el nombre del archivo que contiene el esquema del cubo OLAP, por defecto aparece el nombre del archivo XML que se subio anteriormente.

Servidor: Es el localhost si se trabaja en el mismo equipo o la direccion IP del equipo que sera el servidor.

Base de datos: Es el nombre de la base de datos con la cual se realizara la conexión.

Usuario: Es el usuario que se creó en el gestor de la base de datos en este caso en PostgreSQL.

Contraseña: Es la clave con la cual el usuario ingresa al gestor de la base de datos en PostgreSQL

Al editar los parámetros de acuerdo a nuestra necesidad, se da clic en conectar, ver figura 49, y si todo está bien aparecerá un mensaje de conexión exitosa. Ver figura 50.

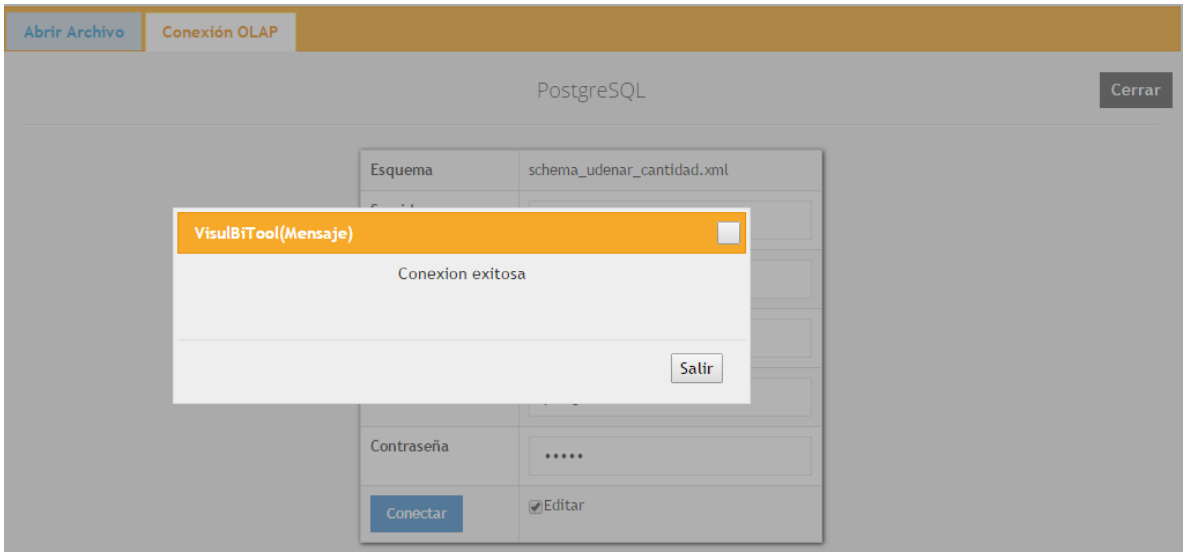


Figura 50. Mensaje de éxito en la conexión

4.3 ANÁLISIS OLAP

En el menú principal seleccionar Análisis de Datos/Análisis OLAP como se ve en la figura 51.

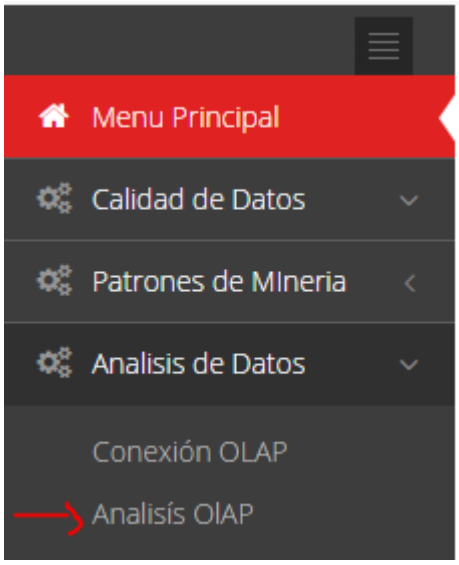


Figura 51. Menú principal de VisualBITool

En la parte superior se abre una pestaña , la cual se selecciona como se muestra en la figura 52.



Figura 52. Gestor de pestañas de VisualBITool

Se abre una ventana en la cual se puede observar la estructura de un cubo y una tabla con la cual se puede navegar sobre el cubo OLAP. Ver figura 53.

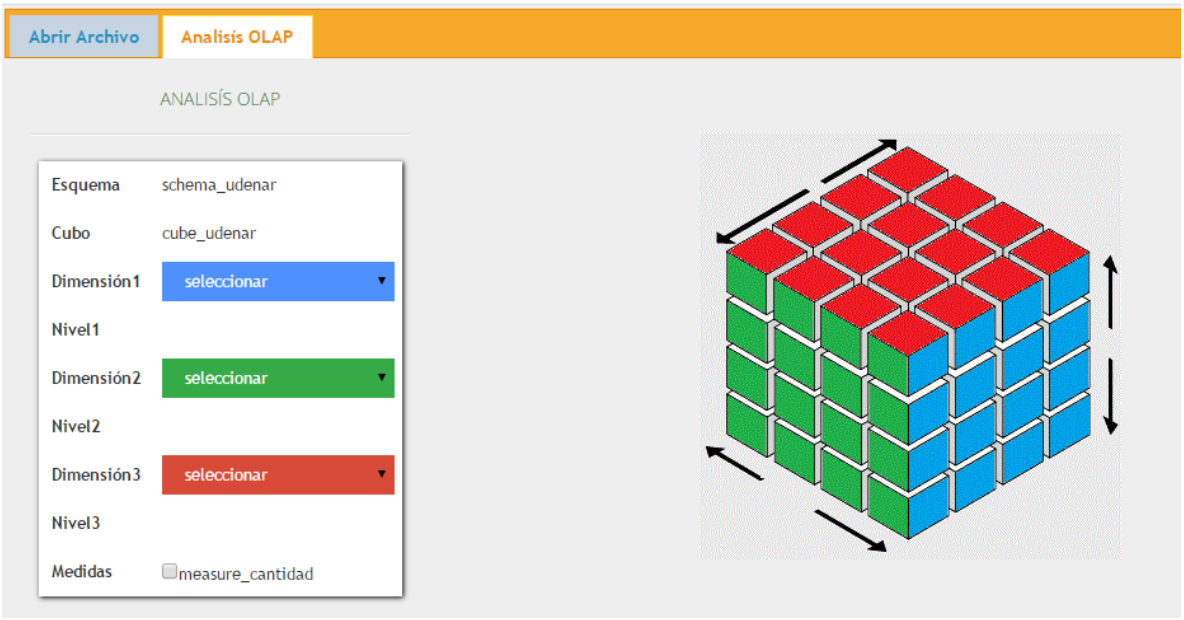


Figura 53. Interfaz para análisis OLAP de VisualBITool

La ventana que se a abierto se divide en dos partes análisis OLAP y la imagen de un cubo. Ver figura 53.

La primera es la de análisis OLAP, tiene diferentes opciónes que permite manipular las propiedades del cubo que se desea analizar. Ver figura 54.

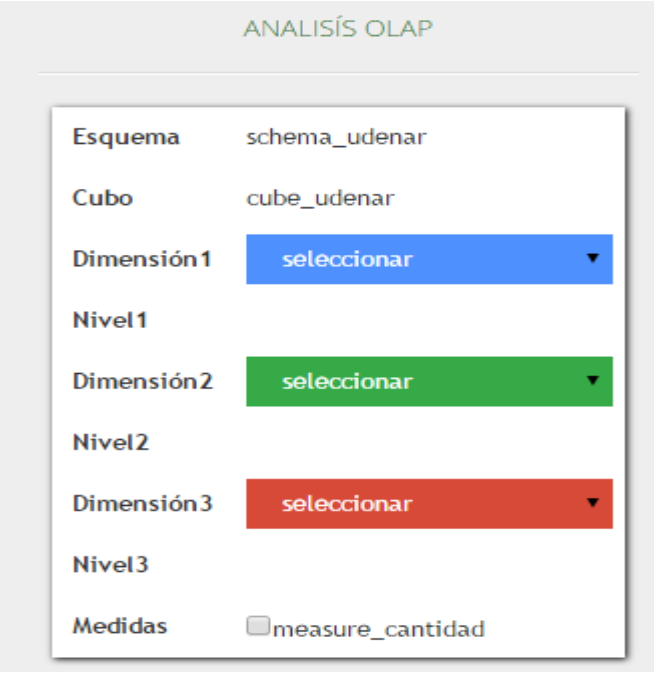


Figura 54. Ventana de gestión de un cubo OLAP

Los diferentes parametros que aparecen en la figura 54 son:

Esquema: Es el nombre del esquema del cubo

Cubo: Es el nombre del cubo

Los parametros de Dimension permiten seleccionar tres dimensiones del cubo que es objeto del análisis asi:

Dimensión1: Permite seleccionar el nombre de la primera dimension

Nivel1: Permite seleccionar un nivel de la priemera dimension

Dimensión2: Permite seleccionar el nombre de la segunda dimension

Nivel2: Permite seleccionar un nivel de la segunda dimension

Dimensión3: Permite seleccionar el nombre de la tersera dimension

Nivel3: Permite seleccionar un nivel de la tersera dimension

Medidas: Permite seleccionar las medidas del cubo con las que se desea realizaer el análisis

La figura 55 muestra un ejemplo de como quedan seleccionados los diferentes parametros.

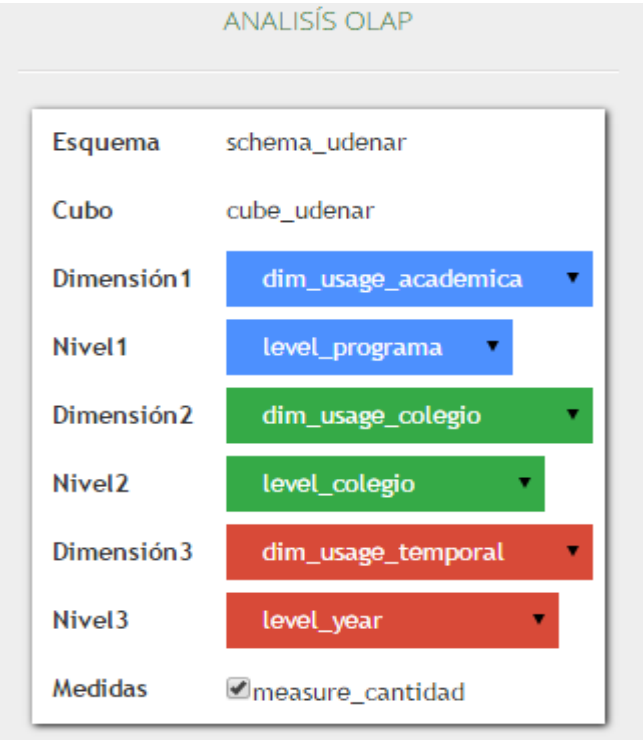


Figura 55. Ventana de selección de las partes de un cubo OLAP

La imagen del cubo que se presenta en la parte izquierda de la ventana de la figura 56 , permite ubicar las dimensiones seleccionadas en cada lado , para dar una idea de que se esta trabajando la información organizada en una forma multidimensional.

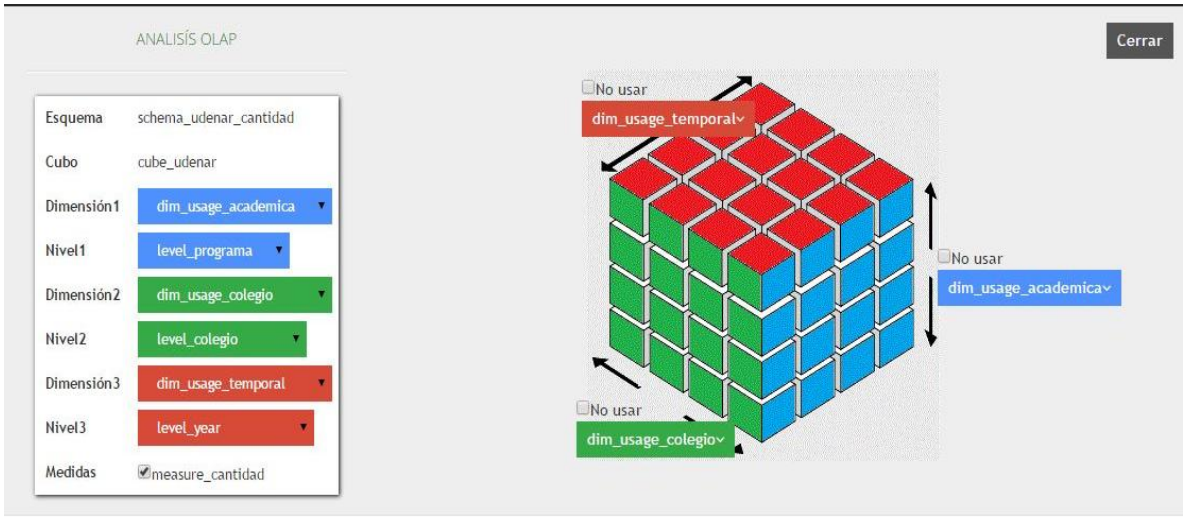


Figura 56. Imagen de un cubo con las dimensiones seleccionadas

Se puede observar que las dimensiones seleccionadas se colocan a los lados del cubo, además el cubo permite realizar las diferentes consultas que se deseen asi:

La figura 57 muestra al cubo y las diferentes opciónes que presenta

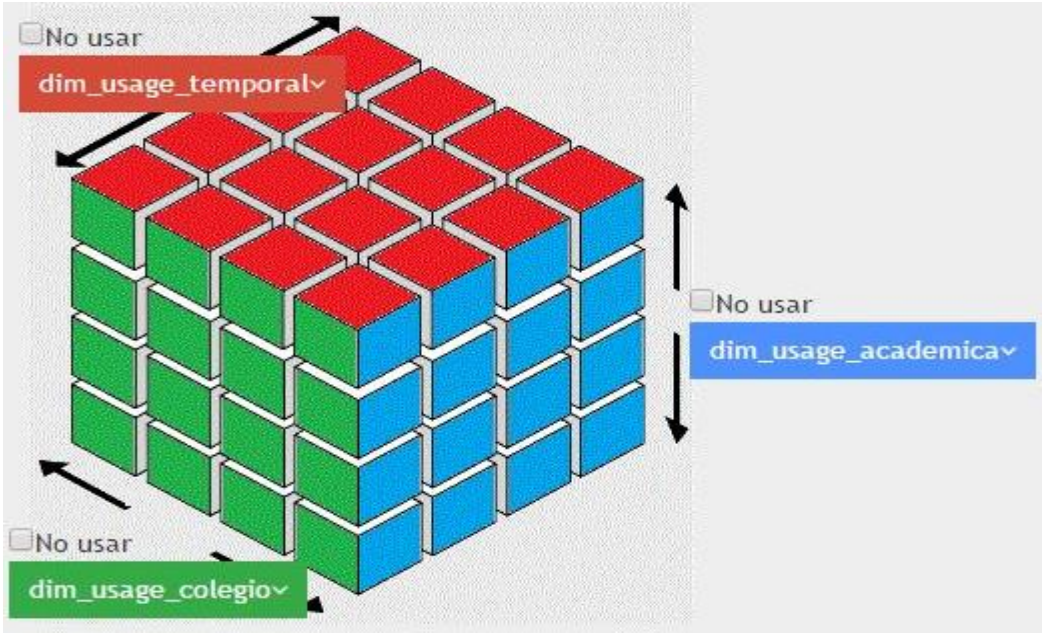


Figura 57. Imagen de un cubo con las tres dimensiones seleccionadas

Cada dimension de la figura 58, en su parte superior tienen una opción para ser seleccionada “**No usar”**, esta opción permite decidir si se desea trabajar en las consultas con las tres,con dos o con una dimension.

Adema cada dimension al dar clic sobre ella se desprende un menú, el cual permite observar el nivel con el que se esta trabajando y una opción de filtrado.ver figura 58.



Figura 58. Nivel y filtros de una dimensión

La opción Filtro de la figura 58, permite realizar filtros sobre las diferentes consultas si no se desea mirar todos los datos de las dimensiones, permitiendo asi facilitar el análisis de la información

Al dar clic en Filtro de la figura 58, se presenta una ventana con todos los datos que tiene el nivel de la dimension seleccionada, en el cual se puede seleccionar los datos que intervendran en el resultado de la consulta. En este caso solo se desea hacer la consulta por el colegio ATENEO HORIZONTE. Ver figura 59.



Figura 59. Ventana con los datos de un nivel para realizar un filtro

Para ejecutar la consulta , se da clic en la imagen que representa un cubo.ver figura 57.

La figura 60 muestra el resultado de la consulta en la cual se deseaba conocer la cantidad de estudiantes por año en cada programa que sean del colegio ATENEO HORIZONTE.

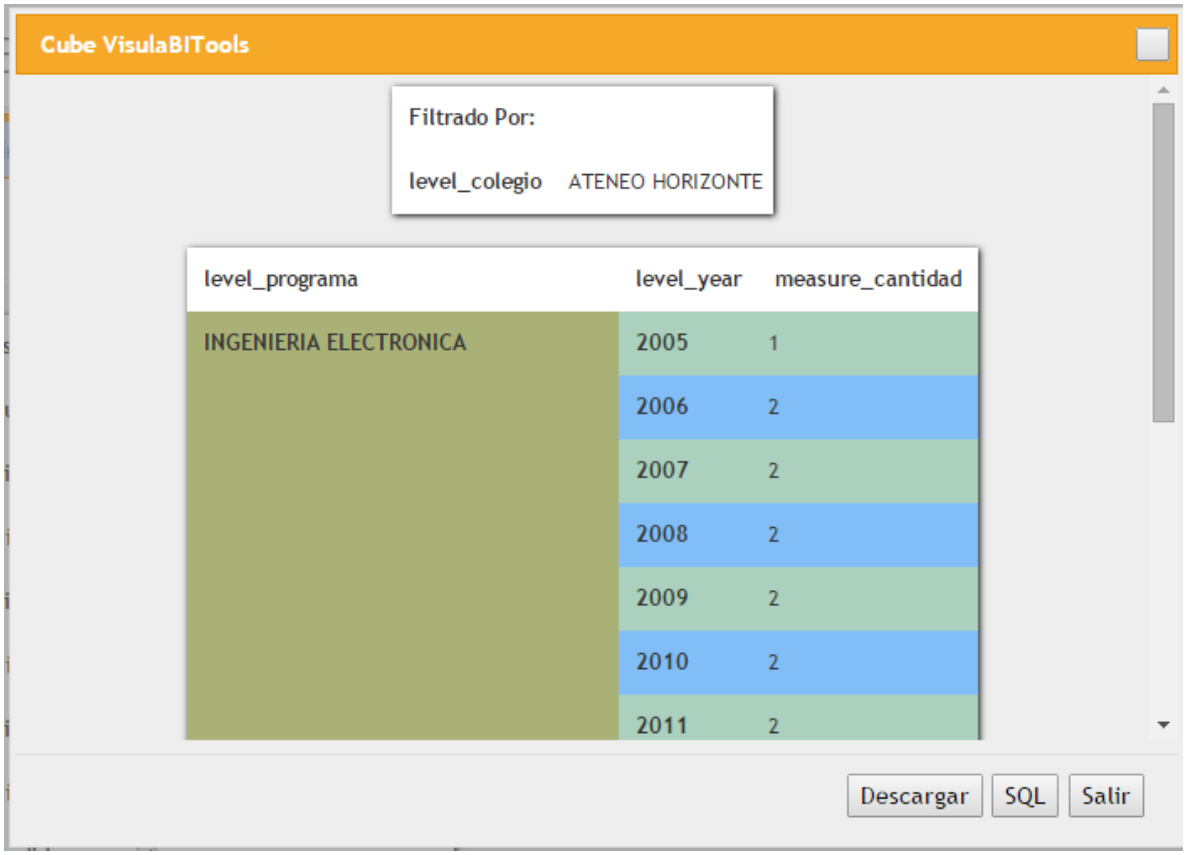
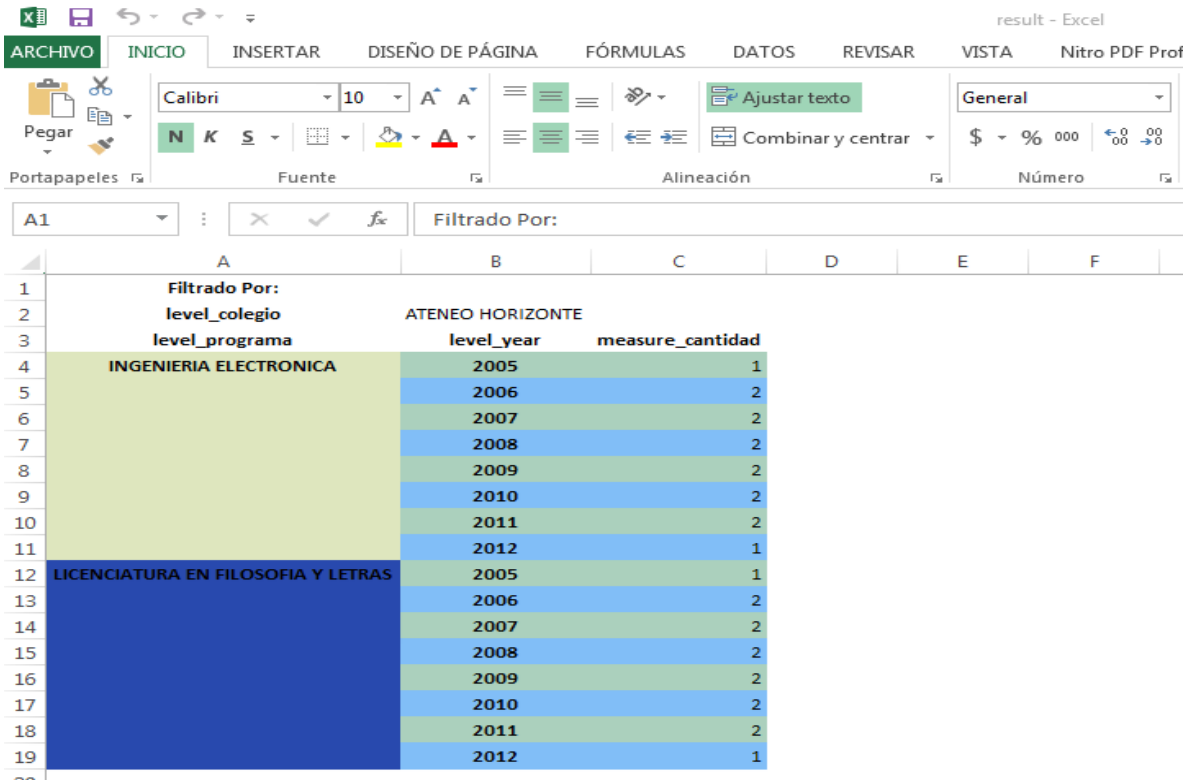


Figura 60. Resultado de una consulta

La ventana de la figura 60, presenta tres botones Descargar, SQL y Salir.

Al dar clic en Descargar, la información de la consulta que se muestra en la figura 60, se exporta a un archivo Excel como se ve en la figura 61.



result - Excel

ARCHIVO | INICIO | INSERTAR | DISEÑO DE PÁGINA | FÓRMULAS | DATOS | REVISAR | VISTA | Nitro PDF Prof

Pegar | Portapapeles | Calibri | 10 | Fuente | Ajustar texto | Combinar y centrar | Alineación | General | Número

A1 | Filtrado Por:

	A	B	C	D	E	F
1	Filtrado Por:					
2	level_colegio	ATENEO HORIZONTE				
3	level_programa	level_year	measure_cantidad			
4	INGENIERIA ELECTRONICA	2005	1			
5		2006	2			
6		2007	2			
7		2008	2			
8		2009	2			
9		2010	2			
10		2011	2			
11		2012	1			
12	LICENCIATURA EN FILOSOFIA Y LETRAS	2005	1			
13		2006	2			
14		2007	2			
15		2008	2			
16		2009	2			
17		2010	2			
18		2011	2			
19		2012	1			

Figura 61. Reporte en Excel de la consulta realizada

Al dar clic en el botón SQL, que se ve en la figura 60, se abre una ventana mostrando la estructura SQL de la consulta. Ver figura 62.

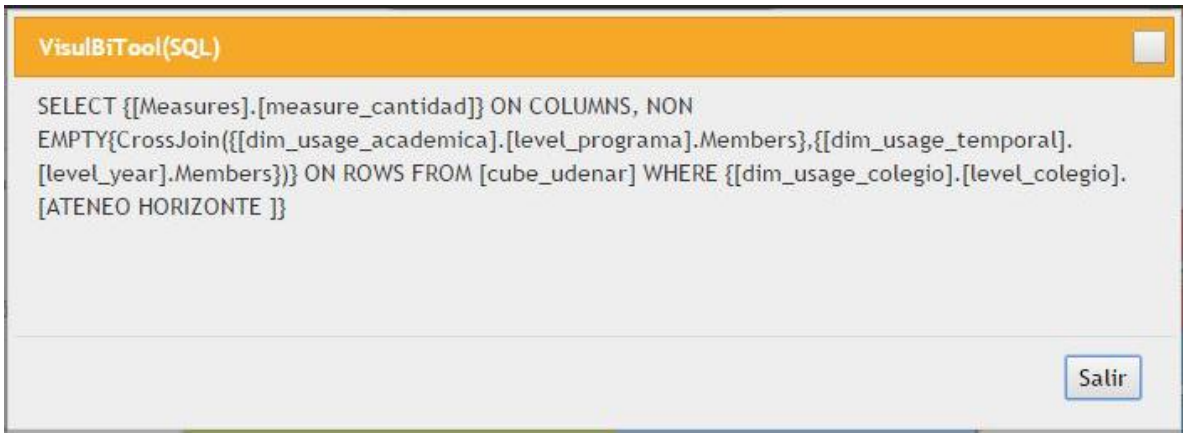


Figura 62. Estructura de una consulta

Al dar clic en el botón salir que se ve en la figura 60 se cierra la ventana con el resultado de la consulta.