

Inteligencia Empresarial

Práctica 2

**Cubo multidimensional y operaciones OLAP**

Miguel Oleo Blanco

Tabla de contenidos

[Introducción 3](#_Toc95043876)

[Adquisición de los datos y creación del cubo 3](#_Toc95043877)

[Operaciones OLAP 3](#_Toc95043878)

[Slice 4](#_Toc95043879)

[Roll-accross 4](#_Toc95043880)

[Pivot + roll-up 4](#_Toc95043881)

[Ejemplo 2 5](#_Toc95043882)

[Operaciones OLAP 6](#_Toc95043883)

# Introducción

En esta práctica vamos a crear un cubo de cuatro dimensiones a través de datos de Internet. Una vez tengamos el cubo, procedemos a realizar operaciones OLAP y estudiar como funciona todo esto en R.

# Adquisición de los datos y creación del cubo

Para conseguir los datos con los que vamos a trabajar, empleo una librería de R llamada arules. Esta librería al cargarla nos proporciona varias tablas con datos que podemos emplear para nuestros estudios. En concreto voy a usar los datos que se llaman AdultUCI. Para leer la librería y estos datos se hace de la siguiente forma:

Text

Description automatically generated

También he cargado la librería data.table. Esta librería la voy a emplear para crear las tablas. En concreto, la tabla AdultUCI viene con muchas variables y, para simular que creamos un cubo con datos de varias tables, voy a dividir estas variables en varias tablas que más tarde juntaremos. En concreto, creamos las tablas con data.table y luego para juntarla usaremos un merge con left/right join y otras operaciones. Es importante advertir que como estamos trabajando con muchas observaciones, este método no funciona, por lo que muestro como debería de ser el proceso, pero para generar la tabla conjunto, lo hago de forma manual:

Text

Description automatically generated with medium confidence

Po último creamos el cubo de 4 dimensiones, indicando el valor a guardar (en mi caso la variable fnlwgt) y las dimensiones (age, sex, race e income):



# Operaciones OLAP

Antes de empezar con las operaciones OLAP, podemos observar información sobre el cubo. Por ejemplo, con dimnames podemos observar las dimensiones del cubo y sus posibles valores. Esto es muy importante para comprobar que la creación del cubo es correcto antes de empezar a operar con el:

A picture containing calendar

Description automatically generated

## Slice

Con esta operación obtenemos un conjunto de valores (es una selección):

A picture containing chart

Description automatically generated

Table

Description automatically generated

## Roll-accross

Esta operación es como drill-up pero esta no se hace sobre una jerarquía. Esta operación trabaja yendo de lo específico a lo general, quitando atributos. A continuación, se muestra el ejemplo:

Table

Description automatically generated with medium confidence

## Pivot + roll-up

Esta operación es muy parecida a la anterior añadiendo el pivot. Esta operación hace una rotación sobre las variables, por lo que se muestran es orden distinto:

A picture containing text

Description automatically generated

Text

Description automatically generated

# Ejemplo 2

En este otro ejemplo voy a crear a mano las tablas para un ejemplo algo más real (aunque con muchos menos datos. Estas tablas son las siguientes:

* State\_table: Contiene la key en forma de iniciales de la ciudad, el nombre de la ciudad y el nombre del país.
* Month\_table: Contiene como key el número del mes. La descripción son las iniciales del mes. Por último, el cuatrimestre.
* Prod\_table: Contiene tres marcas de SUV de fabricantes Alemanes y sus respectivos precios.
* Sales\_fact: Esta tabla es autogenerada a partir de las otras tres tablas. Esta contiene el mes, año, producto, unidades vendidas y euros generados.

A continuación, se muestran los 6 primeros datos de la tabla sales\_fact:

Text, table

Description automatically generated

A partir de esta tabla procedo a crear dos cubos con los que haré operaciones OLAP. EL primer cubo contiene los euros generados según el producto, el mes, el año y la localización. El segundo es otro ejemplo igual que el primero, pero con las dimensiones en otro orden. También se muestran datos sobre los cubos:

Text

Description automatically generated with medium confidence

Table

Description automatically generatedA screenshot of a computer

Description automatically generated with medium confidence

## Operaciones OLAP

Graphical user interface, text, application, email

Description automatically generatedGraphical user interface, text, email

Description automatically generatedText

Description automatically generatedGraphical user interface, text

Description automatically generated with medium confidenceText

Description automatically generatedA continuación, se van a mostrar las operaciones OLAP que se llevan a cabo sobre estos cubos. No voy a entrar en detalle sobre las operaciones, ya que han sido explicadas en el primer ejemplo

Text

Description automatically generated