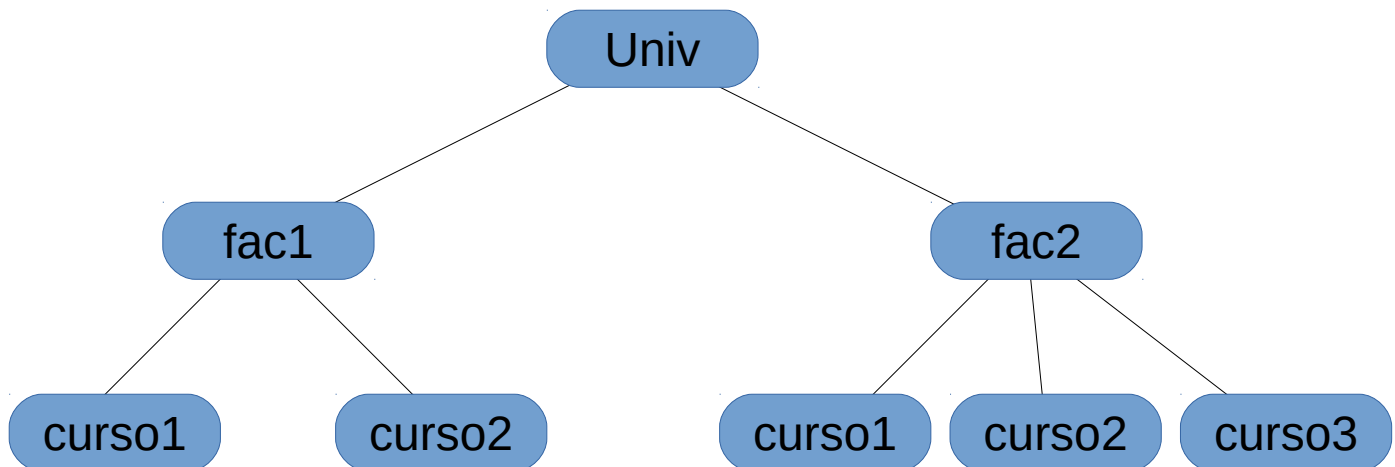


Taller 5

Tablas de símbolos - Árboles de búsqueda

Existen muchas situaciones en las que se desean totalizar estadísticas por grupos de ADTs que cumplen ciertas características. E.g. estudiantes por facultad y curso, estudiantes por facultad o total de estudiantes. Situaciones como esta se pueden calcular y representar fácilmente mediante tablas de símbolos anidadas, donde cada nivel de anidación corresponde con un criterio adicional de selección, por ejemplo:



Así, cada nodo representa las estadísticas correspondientes a cada nivel de agrupación y la raíz del árbol a las estadísticas total de la colección.

Ejercicio a desarrollar

Reutilizando la colección de [canciones](#) utilizada en el taller 4 implementar un árbol de agrupación de datos con las siguientes características:

- El nivel inferior del árbol agrupa las canciones por artista, calculando el número de canciones y el promedio de *hotness* de cada artista.
- El segundo nivel agrupa las canciones por género (atributo *terms* del dataset), de nuevo totalizando el número de canciones y el promedio de *hotness*.
- El nivel raíz contiene el número total de canciones y el *hotness* promedio de toda la colección.

Características del programa:

1. Implementar el árbol de agrupamiento utilizando árboles de búsqueda binarios (las llaves del árbol de primer nivel son los géneros y los del segundo nivel son la combinación género-artista).

El programa principal debe:

2. Leer todas las canciones de la colección. Agregar los datos de cada canción a los nodos correspondientes del árbol.
3. Imprimir el las estadísticas de la colección haciendo un recorrido en pre-orden del árbol (se reportan los totales de un nodo y luego los de sus nodos hijo).

Condiciones del taller

Grupos de trabajo de máximo 3 personas

Fecha de entrega: Mayo 25 de 2019.

La nota del taller reemplaza un quiz.