**Semana 1**

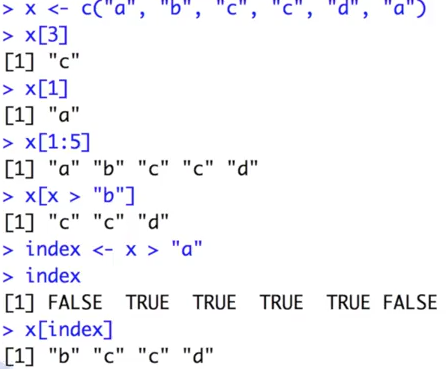
**05. Lección de Swirl 3: Obtener Subconjuntos de Datos**

**Lección 4 – Subconjuntos de Datos**

En esta lección aprenderemos a sacar un subconjunto de datos de objetos o variables ya disponibles en memoria. Este tipo de objetos será de forma de matrices o listas y de data frames. Hay 3 operadores básicos para hacer una extracción de un conjunto de datos. Estos son:

* Corchetes (“[]”); Extraen elementos de la misma clase al objeto que estamos extrayendo
* Dobles Corchetes: Extrae elementos de una lista o dataframe. La clase regresada pued no ser una lista o dataframe.
* El símbolo de dolas (“$”); Extrae elementos de un dataframe.

El corchete puede usarse para pedir el elemento de una posición (empezando desde la posición 1, no 0 como en otros lenguajes), el conjunto de una secuencia de posiciones o pasar un vector lógico y regresa los lugares en los que el valor del vector lógico son mayores que uno dado.



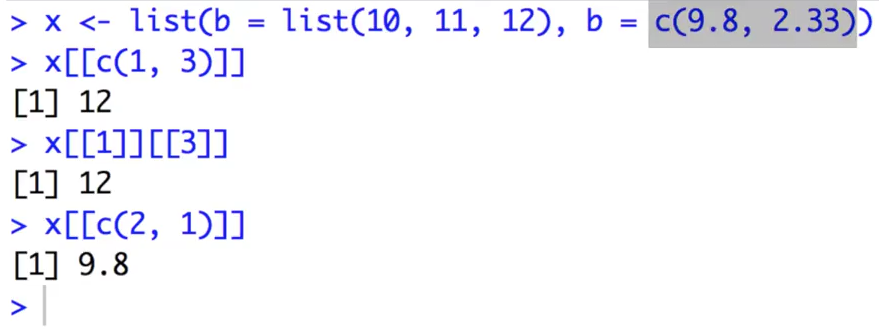
Para hacer un subconjunto de datos de una matriz, podemos usar los operadores con corchetes recordando que la matriz tiene dos dimensiones. Podemos entonces pedir las siguientes opciones: *matrix[x,y]; matrix[,y]; matrix[x,]*.Por ejemplo, pedir una posición exacta mezclando ambos valores.

Una cosa importante recordar es que al pedir el elemento de una matriz, esta la devuelve como un vector, perdiendo su condición de matriz, pero si usamos el parámetro *drop=FALSE*, R nos devuelve también una matriz con un renglón y una columna.

Para extraer elementos de lista es mediante corchetes, dobles corchetes o mediante el valor dólar y el valor exacto (sólo puede ser usada con nombres literales). Cuando hemos nombrado a los elementos de nuestra lista. Si mandamos a llamar usando solo corchetes solo nos devuelve el elemento de una lista exacta.

Con el operador de doble corchete podemos utilizar índices computados, es decir, que podemos generar un índice y meterlo en el doble corchete, extrayendo los elementos que deseemos.

Los elementos anidados de una lista son listas dentro de otras.



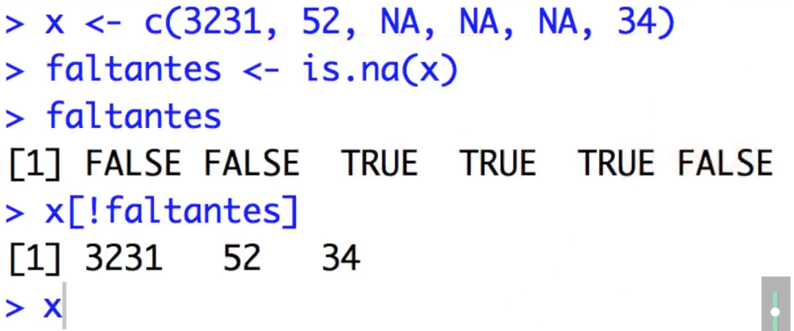
Si quisiéramos tomar el valor 12, es decir, el primer valor de la segunda lista dentro de x, se usan dobles corchetes.

SI creamos una lista con un único elemento y este tiene un nombre complejo, se puede hacer un emparejado parcial, que es buscar aquel con el mismo valor más cercano. Esto se puede ver en el siguiente ejemplo:

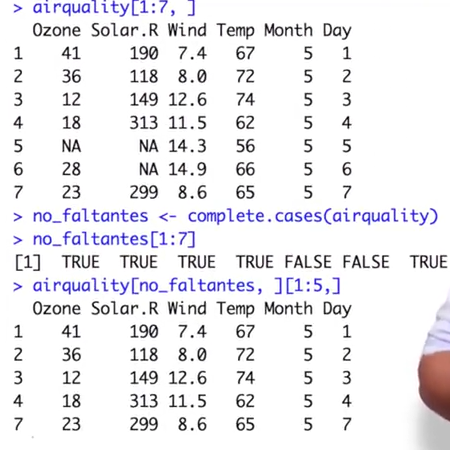


Lo mismo pasa con el símbolo de dólar. Si utilizamos el parámetro *exact*, al intentar hacer uso del emparejado parcial no va a funcionar.

Cuando nos encontamos con valores NA, querremos eliminarlos y quedarnos con los valores diferentes a NA. Se genera un vector de valores lógicos con *is.na()*, devolviendo un vector con valores verdaderos donde haya NA. Si queremos lo contrario tenemos que negar este vector usando !.



Si tenemos un data frame, y tenemos NA en algunas columnas, para eliminar las filas con NA se usa con *complete.cases()*, en la que se genera un vector lógico y se llama ese vector lógico.



**Tarea RStudio**

Se pueden usar enteros negativos para obtener un vector que consiste en todos los elementos excepto los elementos especificados, como por ejemplo *mi\_vector[-9:-15].* Con el parámetro *rep* indicas que quieres solo las posiciones impares.