**MANEJANDO DATOS**

**Introducción**

Uno de los desafíos en el cual nos enfrentamos al trabajar con conjuntos de datos masivos es la recopilación, preparación y gestión de datos de una variedad de fuentes distintas.

Si bien cubriremos más adelante la preparación de datos, la limpieza de datos y la administración de datos en profundidad; Trabajaremos estas tareas antes de desarrollar algoritmos de aprendizaje automático en el mundo real, en este momento destacaremos **la funcionalidad básica para capturar datos de R** únicamente.

**Guardar, cargar y eliminar datos en R**

Cuando haya pasado mucho tiempo obteniendo un marco de datos en el formulario deseado, no debería tener que volver a crear su trabajo cada vez que reinicie su sesión R.

Para guardar una estructura de datos en un archivo que puede recargarse más tarde o transferirse a otro sistema, use la función **save ()**.

La función save () escribe una o más estructuras de datos R en la ubicación especificada por el parámetro de archivo.

Los archivos de datos R tienen una extensión **RData**.

Suponga que tiene tres objetos llamados "x", "y", y "z" que desea guardar en un archivo permanente.

x **<-** c**(**1**:**50**)**

y **<-** c**(**"Lunes", "Martes", "Miercoles", "Jueves", "Viernes"**)**

z **<-** factor**(**c**(**"p", "q", "p", "r", "q"**))**

**SAVE()**

Independientemente de si son vectores, factores, listas o marcos de datos, podríamos guardarlos en un archivo llamado mydata.RData usando el siguiente comando:

save**(**x, y, z, file **=** "mydata.RData"**)**

**LOAD()**

El comando load () puede recrear cualquier estructura de datos que se haya guardado en un archivo .RData. Para cargar el archivo mydata.RData que guardamos en el código anterior, simplemente escriba:

load**(**"mydata.RData"**)**

**Cuidado**

Todas las estructuras de datos almacenadas en el archivo que está importando con el comando load () se agregarán a su espacio de trabajo, **van a sobrescribir cualquier trabajo que tenga abierto**.

Si necesita finalizar su sesión R a toda prisa, el comando **save.image ()** escribirá toda su sesión en un archivo llamado simplemente .RData. De forma predeterminada, R buscará este archivo la próxima vez que inicie R, y su sesión se volverá a crear tal como la dejó.

**LS()**

Después de trabajar en una sesión R durante algún tiempo, es posible que haya acumulado varias estructuras de datos. La función de listado ls () devuelve un vector de todas las estructuras de datos actualmente en la memoria.

**RM()**

R se eliminará automáticamente de su memoria al salir de la sesión, pero para estructuras de datos grandes, es posible que desee liberar la memoria antes. La función rm() o remove() puede usarse para este propósito. Por ejemplo, para eliminar los objetos m y subject1, simplemente escriba:

Pueden utilizar Ctl + Alt + R para correr todo el ejemplo.

rm**(**m, subject1**)**

La función rm () también se puede suministrar con un vector de caracteres de los nombres de objetos que se eliminarán. Esto funciona con la función ls () para borrar toda la sesión R:

rm**(**list **=** ls**())**

**¡Tenga mucho cuidado al ejecutar el comando anterior, ya que no se le pedirá que guarde sus objetos!**