





INSTITUTO POLITECNICO NACIONAL

Unidad Profesional Interdisciplinaria de Ingeniería y Ciencias Sociales y Administrativas.

**Almacenamiento de Datos y su Administración**

***Fabián Ramírez Galindo***

***Tarea 3***

***Introducción al lenguaje de programación R***

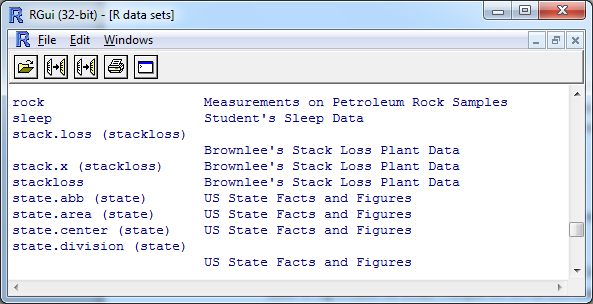
**09/11/2015**

Desarrollo

Ejecute data() en la consola de R y revise los conjuntos de datos disponibles. De éstos, ubique el nombre del conjunto de datos que corresponde a la descripción “Measurements on Petroleum Rock Samples“.

**La ejecución de la instrucción data nos mostrara el conjunto de datos disponibles, en uso nuestro caso haremos del conjunto de datos rock.**

**> data()**



Sobre este conjunto de datos, indique los comandos o funciones que le permiten hacer, saber o determinar:

* Los nombres de las columnas del conjunto de datos. Consulte la ayuda en línea para saber el significado de éstas. Indique los comandos usados.

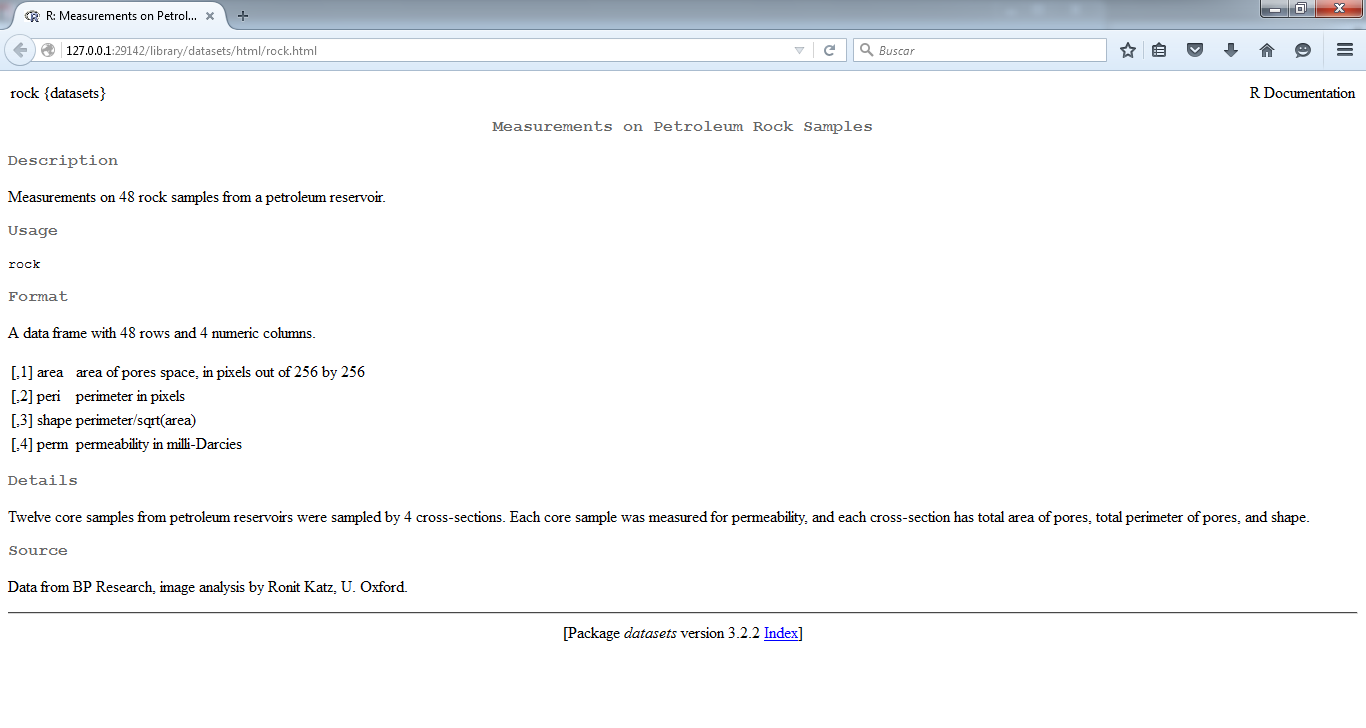
**Para obtener el nombre de las columnas se tendrá que ejecutar el comando names y pasar como parámetro el nombre del conjunto de datos.**

**Se tendrá que ejecutar el comando help proporcionando el conjunto de datos rock para saber el nombre de las columnas y el significado de estas.**

**> names(rock)**

**[1] "area" "peri" "shape" "perm"**

**> help(rock)**



* Las dimensiones del conjunto de datos. Describa el significado de lo que muestra la consola.

**Para obtener las dimensiones del conjunto de datos se tendrá que ejecutar la instrucción dim proporcionándole el nombre del conjunto de datos. La ejecución de la instrucción nos mostrara 2 valores, donde el primer valor corresponde al número de observaciones y el segundo valor al número de columnas.**

**> dim(rock)**

**[1] 48 4**

* El número de observaciones del conjunto de datos, esto es, sólo mostrar la cantidad de observaciones que compone al conjunto de datos (uno de los valores del punto previo).

**El número de observaciones del conjunto de datos es 48.**

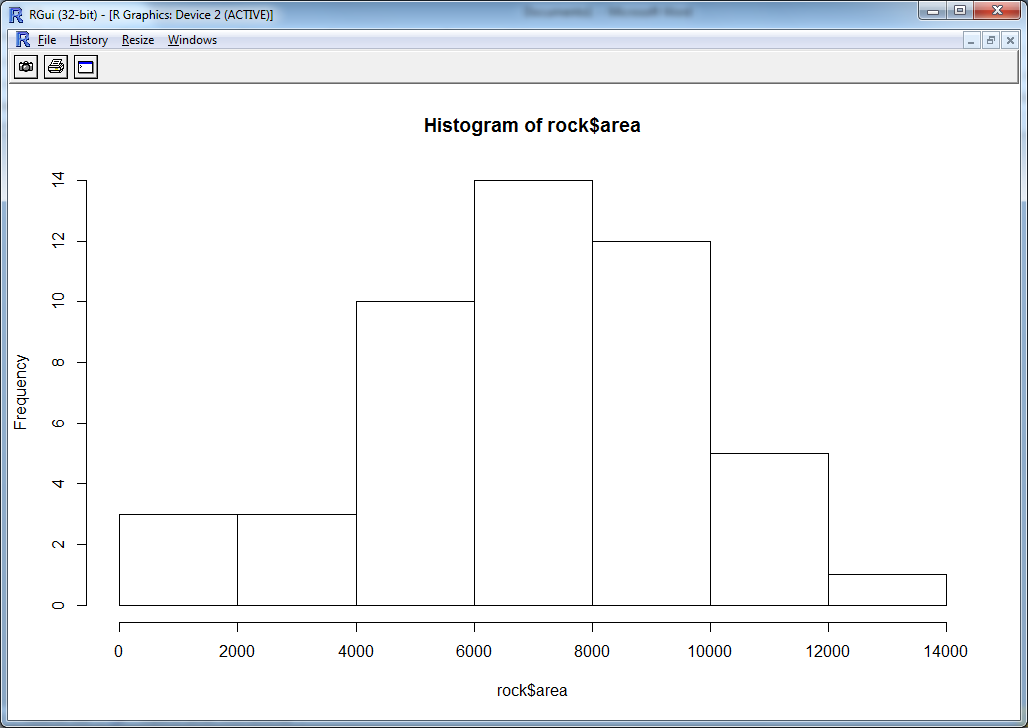
* El número de categorías o clasificaciones que componen cada observación (en otras palabras, el otro dato no mostrado en el punto previo).

**El número de categorías son 4.**

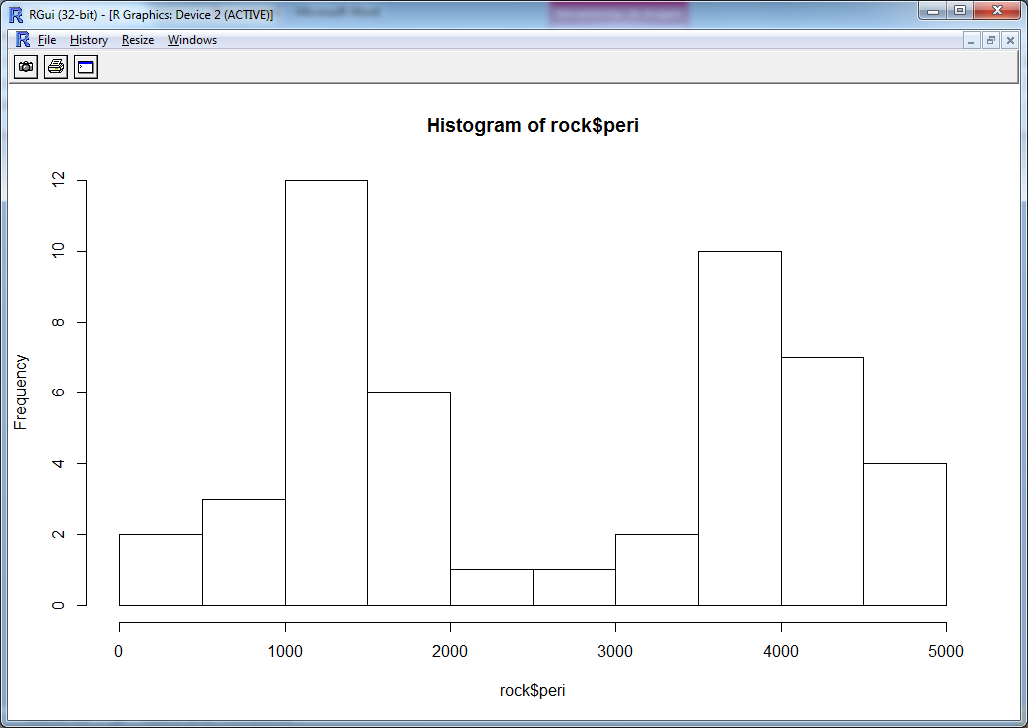
* Haga un histograma de cada una de las columnas (tipo de observación) del conjunto de datos.

**Para mostrar el histograma de cada una de las columnas será necesario ejecutar la instrucción hist pasándole como parámetro el nombre de la columna.**

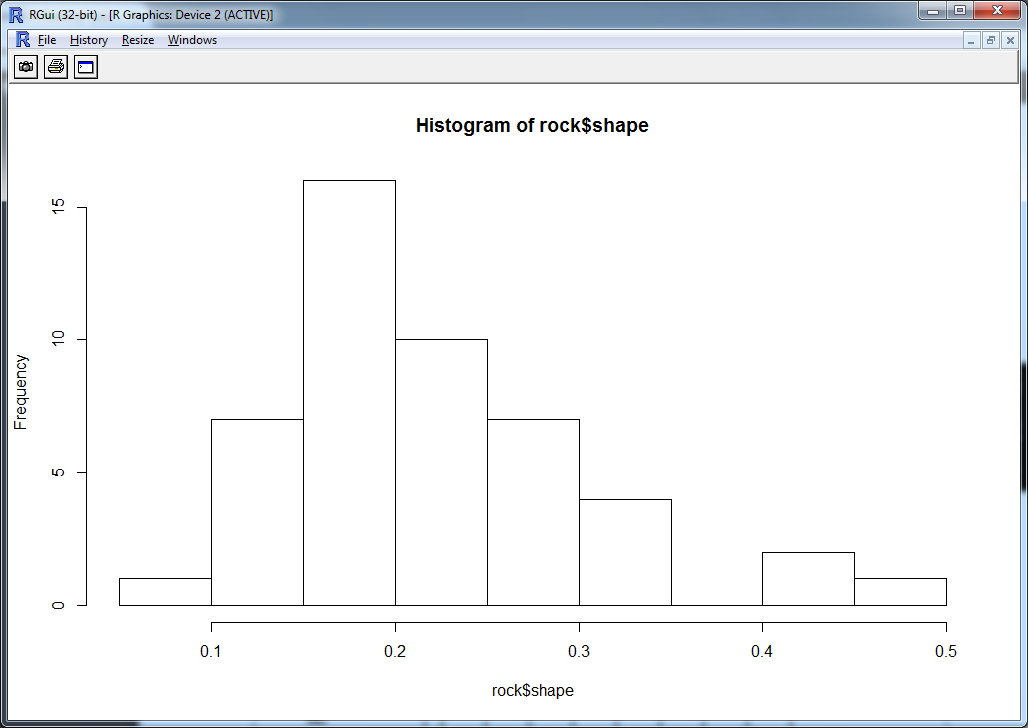
**> hist(rock$area)**



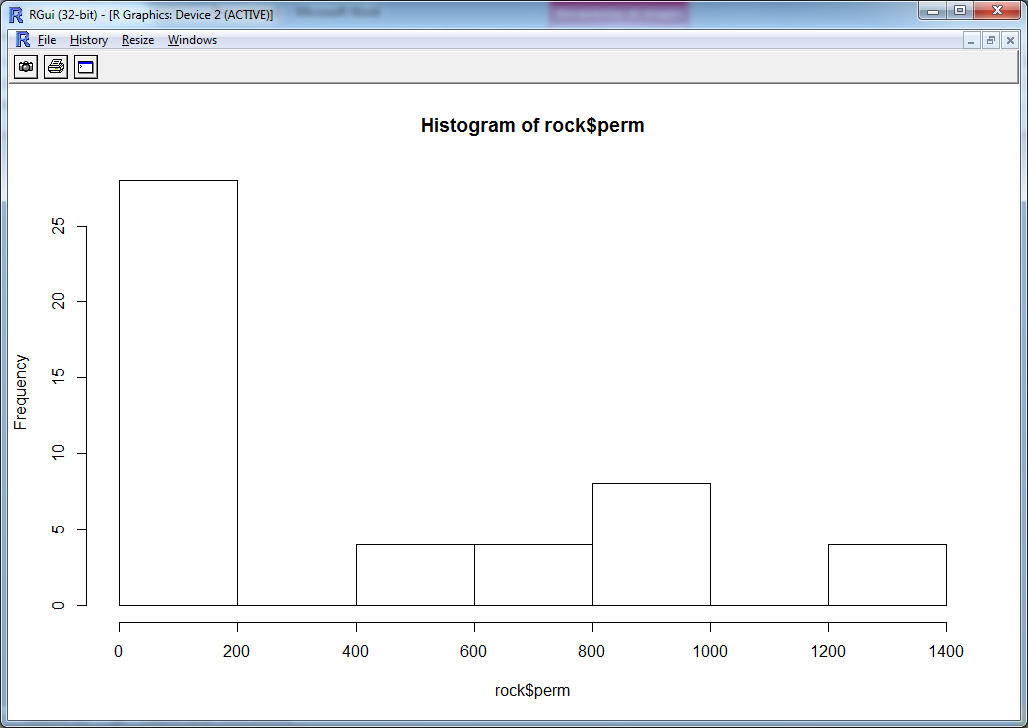
**> hist(rock$peri)**



**> hist(rock$shape)**



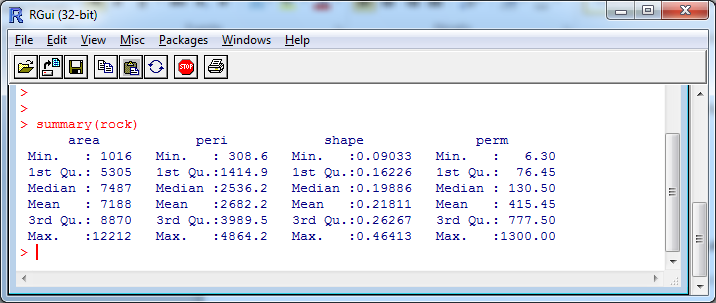
**> hist(rock$perm)**



* Obtenga un resumen del conjunto de datos (tip: utilice la función summary()).

**Para obtener el resumen del conjunto de datos se tendrá que ejecutar el comando summary y proporcionarle el nombre del conjunto de datos.**

**> summary(rock)**



* Obtenga una gráfica que permita apreciar la relación del área de las rocas de la muestra contra su permeabilidad.

**Para generar la gráfica ejecutamos la instrucción plot, en este caso se le proporcionaron los parámetros área y permeabilidad, los parámetros main, xlab y ylab se utiliza para asignar un texto a la gráfica, al eje” x” y al eje “y”.**

**> plot(rock$area,rock$perm,main="Rocas",xlab="Área de las rocas", ylab="Permeabilidad")**

