

Trabajo Práctico N° 1

El conjunto de datos con el que vamos a trabajar se encuentra en el siguiente link:

<http://www.stat.berkeley.edu/users/statlabs/labs.html>

STAT LABS: Data

Data Descriptions

- [Birth weight I](#)
- [Video games](#)
- [Radon](#)
- [CMV DNA](#)
- [HIV](#)
- [Crab molting](#)

Maternal Smoking and Infant Health I:

[Birth weight I](#)

Variable	Description
bwt	Birth weight in ounces
smoke	Smoking status of mother 0=not now, 1=yes now, 9=unknown

Video Game Survey:

[Video games](#)

En la columna **Data Descriptions**, buscamos el link al conjunto de datos *Birth weight II*.

Stat Labs Data Page - Microsoft Internet Explorer

Archivo Edición Ver Favoritos Herramientas Ayuda

Atrás Búsqueda Favoritos Ir Vínculos

Dirección <http://www.stat.berkeley.edu/users/statlabs/labs.html#babies>

Maternal Smoking and Infant Health II:

[Birth weight II](#)

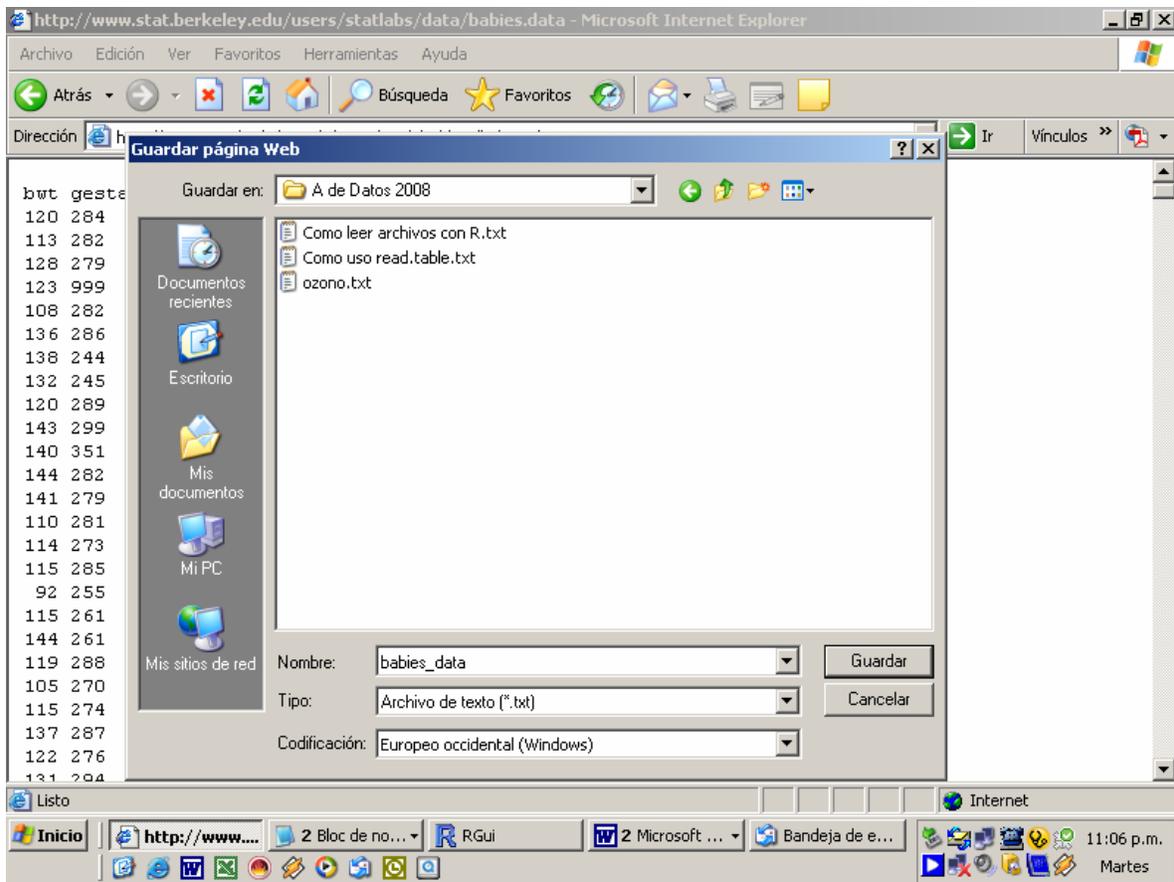
Variable	Description
bwt	Birth weight in ounces (999 unknown)
gestation	Length of pregnancy in days (999 unknown)
parity	0= first born, 9=unknown
age	mother's age in years
height	mother's height in inches (99 unknown)
weight	Mother's prepregnancy weight in pounds (999 unknown)
smoke	Smoking status of mother 0=not now, 1=yes now, 9=unknown

[Additional data on the parents in the Child Health and Development Study](#)
Link to a [description](#) of the data.

Down Syndrome in Genetically Altered Mice:

Inicio Stat Labs Data Pa... 2 Bloc de notas RGui Documento1 - Micro... 10:51 p.m. Martes

Hacemos doble clic en el nombre del archivo y se abre un archivo de texto.
Lo guardamos con la extensión `.txt` (elegimos guardarlo como archivo de texto) en nuestra carpeta de trabajo.



Luego accedemos desde R o S-plus como con cualquier otro archivo de texto.

Para obtener una descripción del conjunto de datos hacemos doble clic en [description](#). Aparece la descripción de más variables que corresponden al archivo extendido.

Trabajaremos con 7 variables.

"bwt" "gestation" "parity" "age" "height" "weight" "smoke"

Antes de empezar a trabajar, reemplace los datos desconocidos o faltantes con NA. Use el help para ver que hace *replace*

- 1- Halle las principales características que presentan los datos para las variables age, height y weight.
- 2 - Describa el comportamiento de los datos para la variable bwt de acuerdo a si la madre es fumadora o no y compare y de acuerdo a si es primeriza o no y compare.
- 3- Idem 2, según las semanas de gestación, tomando 3 categorías: menos de 240 días, entre 240 y 300 días, y más de 300 días. Compare los 3 grupos.