

The R logo is a large, stylized graphic in the background, composed of several blue and yellow rectangular blocks arranged in a 3D-like structure.

# ANÁLISIS DE DATOS CON R



contáctenos: [enei@inei.gob.pe](mailto:enei@inei.gob.pe) / 433-3127



Pasaje Hernán Velarde 285 Lima.

Entre la cuadra 01 y 02 de la Av. Arequipa.

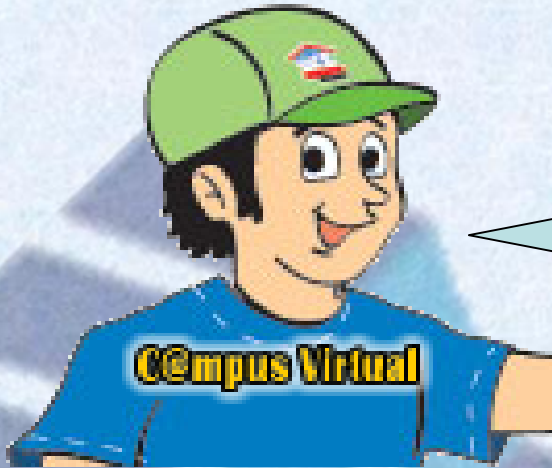
Correo: [enei@inei.gob.pe](mailto:enei@inei.gob.pe) / [campusvirtual@inei.gob.pe](mailto:campusvirtual@inei.gob.pe)

Teléfonos: 433-3127 - 332-4650

## Centro Andino de Formación y Capacitación en Estadística

Cursos Especializados en Estadística e Informática

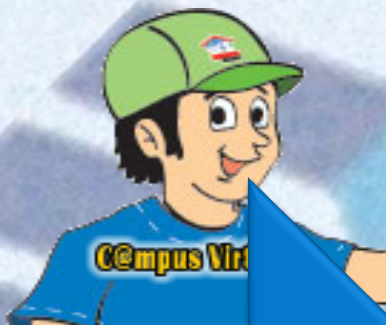




**Estimado alumno,  
buen día. Cualquier  
consulta no dudes en  
comentarme o  
avisarme.**

**Este curso es  
netamente práctico y  
se que lograremos  
objetivos importantes.**

**Recuerda siempre nuestro correo  
para cualquier consulta**



**campusvirtual@inei.gob.pe**





# **LECTURA DE DATOS DE FICHEROS EXTERNOS**

- La función `read.table()`
- La función `data()`
- La librería `foreign`





# read.table()

## La función read.table()

La función `read.table` permite leer datos desde ficheros en formato ASCII. Devuelve como resultado un `data.frame`, por tanto, se supone que cada línea contiene los datos para un individuo.

#El fichero `thuesen.txt` tiene el siguiente aspecto:

```
"blood.glucose" "short.velocity"
```

```
"1" 15.3 1.76
```

```
"2" 10.8 1.34
```

```
...
```

```
"24" 9.5 1.7
```

```
> thuesen<-
```

```
read.table(file="thuesen.txt",header=T)
```

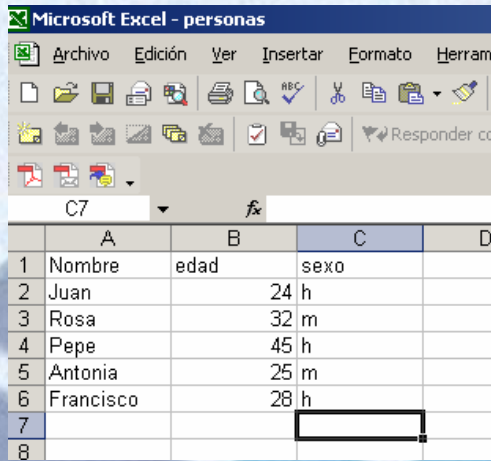
```
> thuesen
```

	blood.glucose	short.velocity
1	15.3	1.76
2	10.8	1.34
...		
24	9.5	1.70

# read.table()

## La función read.table()

El fichero EXCEL personas.xls tiene el siguiente aspecto:



	A	B	C	D
1	Nombre	edad	sexo	
2	Juan	24	h	
3	Rosa	32	m	
4	Pepe	45	h	
5	Antonia	25	m	
6	Francisco	28	h	
7				
8				

```
> #Guardamos el fichero EXCEL como un fichero ASCII delimitado por tabulaciones
```

```
> personas<-read.table(file="personas.txt",header=T)
```

```
> personas
```

	Nombre	edad	sexo
1	Juan	24	h
2	Rosa	32	m
3	Pepe	45	h
4	Antonia	25	m
5	Francisco	28	h



# data()

## La función data()

La función data() permite cargar en el workspace conjuntos de datos que ya existen en formato R y que forman parte de alguna de las librerías.

```
#para ver los datasets actualmente disponibles
```

```
> data()
```

```
> library(ISwR)
```

```
> library(MASS)
```

```
> data()
```

```
> ?airquality
```

```
> data(airquality)
```

New York Air Quality Measurements

Description:

Daily air quality measurements in New York, May to September 1973.

Usage:

```
data(airquality)
```

Format:

A data frame with 154 observations on 6 variables.

'[,1]'	'Ozone'	numeric	Ozone (ppb)
'[,2]'	'Solar.R'	numeric	Solar R (lang)
'[,3]'	'Wind'	numeric	Wind (mph)
'[,4]'	'Temp'	numeric	Temperature (degrees F)
'[,5]'	'Month'	numeric	Month (1-12)
'[,6]'	'Day'	numeric	Day of month (1-31)

```
> airquality
```

	Ozone	Solar.R	Wind	Temp	Month	Day
1	41	190	7.4	67	5	1
2	36	118	8.0	72	5	2
...						
153	20	223	11.5	68	9	30

# Librería foreign

## La librería foreign

Contiene funciones para importar datos que se encuentran en otros formatos.

```
#importamos el archivo coches.sav (SPSS)
> library(foreign)
> coches<-read.spss("coches.sav",to.data.frame=T)
> coches
```

	CONSUMO	MOTOR	CV	PESO	ACEL	A.O	ORIGEN	CILINDR	FILTER..
1	13	5031	130	1168	12.0	70	EE.UU.	8 cilindros	No Seleccionado
2	16	5735	165	1231	11.5	70	EE.UU.	8 cilindros	No Seleccionado
3	13	5211	150	1145	11.0	70	EE.UU.	8 cilindros	No Seleccionado
4	15	4982	150	1144	12.0	70	EE.UU.	8 cilindros	No Seleccionado

Existen otras funciones para leer datos en otros formatos: `read.dta()` para STATA, `read.mtp()` para MINITAB Export File, `read.ssd()` para SAS.



# write.table()

## La función write.table()

Permite escribir los datos en un fichero ASCII.

```
#guardamos los datos de los coches en un fichero texto  
de
```

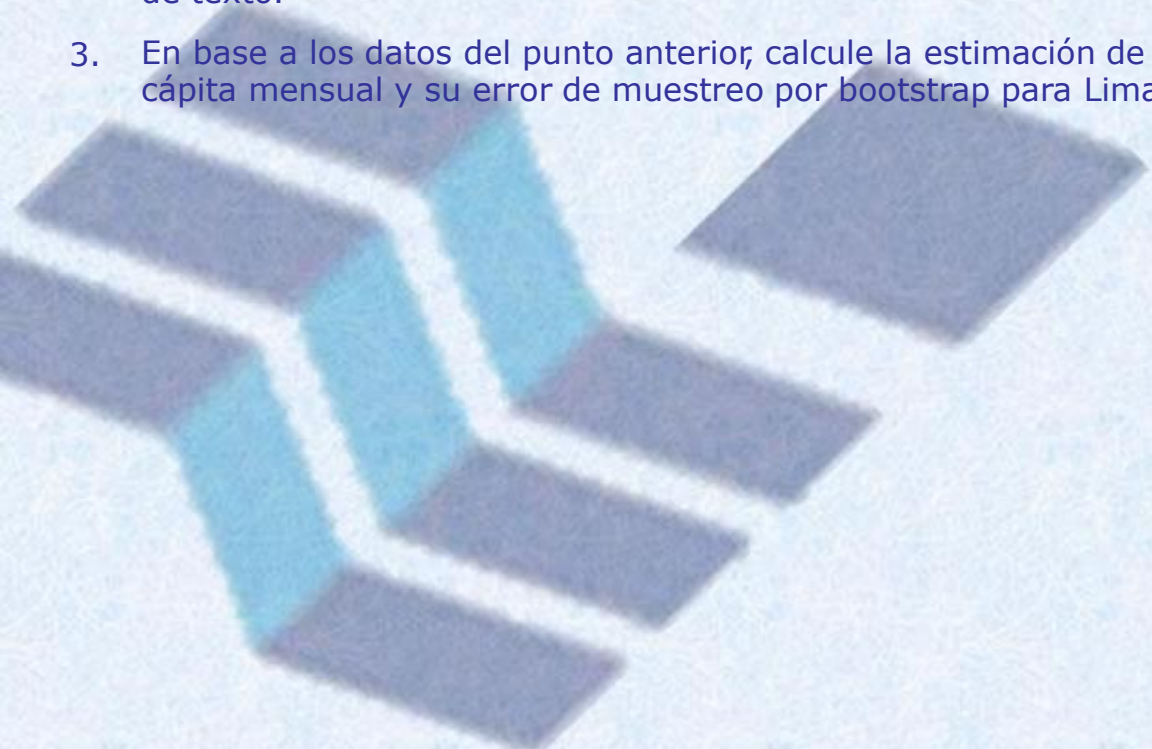
```
> write.table(coches, file="coches.txt")
```

El fichero generado tiene el siguiente aspecto:

```
"CONSUMO" "MOTOR" "CV" "PESO" "ACEL" "A.O" "ORIGEN" "CILINDR" "FILTER.."
"1" 13 5031 130 1168 12.0 70 "EE.UU." "8 cilindros" "No Seleccionado"
"2" 16 5735 165 1231 11.5 70 "EE.UU." "8 cilindros" "No Seleccionado"
"3" 13 5211 150 1145 11.0 70 "EE.UU." "8 cilindros" "No Seleccionado"
"4" 15 4982 150 1144 12.0 70 "EE.UU." "8 cilindros" "No Seleccionado"
```

# Ejercicios propuestos

1. Descarga el archivo SUMARIA de la ENAHO 2013 (página web del INEI) y carga los datos al entorno de R. Graba los datos en formato texto.
2. Con los datos cargados de la SUMARIA ENAHO 2013, crea una variable que contenga el gasto mensual per cápita del hogar y un factor de expansión de personas. Graba el archivo en documento de texto.
3. En base a los datos del punto anterior, calcule la estimación de la media aritmética del gasto per cápita mensual y su error de muestreo por bootstrap para Lima Metropolitana.





## Comunicación constante con la Escuela del INEI

**Correo de la Escuela del INEI**  
**[enei@inei.gob.pe](mailto:enei@inei.gob.pe)**

**Área de Educación Virtual**  
**[campusvirtual@inei.gob.pe](mailto:campusvirtual@inei.gob.pe)**

