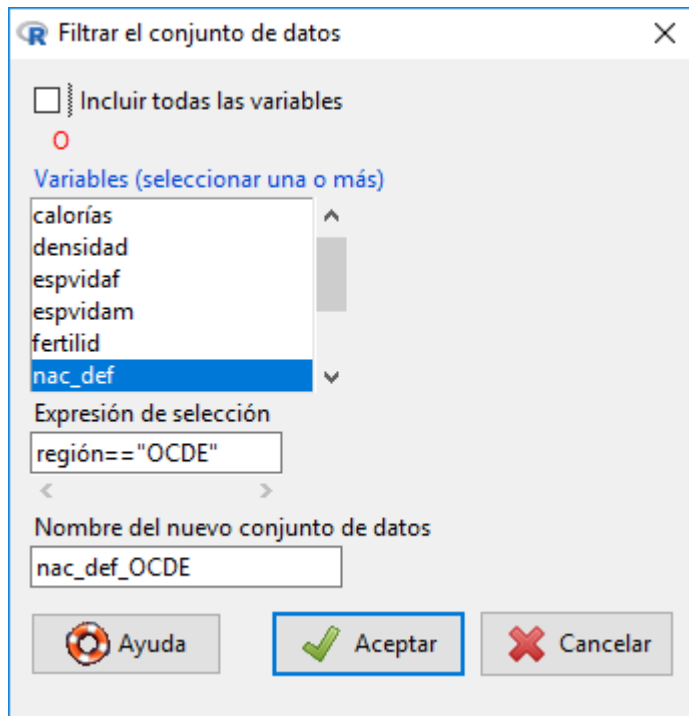


## PRÁCTICA 6

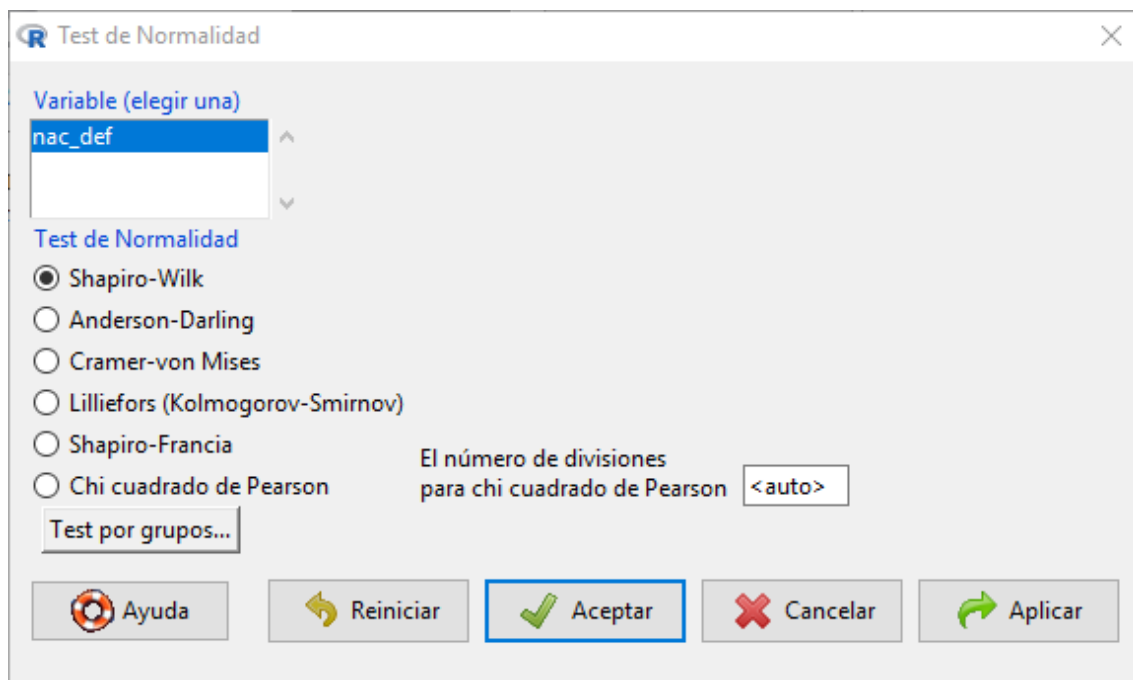
REGLA PARA RESOLVER C.H (CONTRASTE DE HIPÓTESIS) CON R: Se rechaza  $H_0$  si el p-valor obtenido es  $\leq \alpha$

**Ejemplo 1:** Estudiar, para los datos del archivo *Mundo95*, la normalidad de los datos correspondiente a la tasa de nacimientos/defunciones en la región OCDE. Plantear y resolver un contraste de hipótesis adecuado con nivel de significación ( $\alpha = 0.05$ ).

Datos → Conjunto de datos activo → Filtrar el conjunto de datos activos...



Estadísticos → Resúmenes → Test de normalidad



Shapiro-Wilk normality test

```
data:  nac_def
W = 0.89222, p-value = 0.0248
```

Como p-valor es  $< \alpha \rightarrow$  Se rechaza  $H_0$

## Gráficas → Gráfica de comparación de cuantiles

Gráfica de comparación de cuantiles (QQ)

Datos Opciones

Variable (elegir una)

nac\_def

Ayuda Reiniciar Aceptar Cancelar Aplicar

Gráfica de comparación de cuantiles (QQ)

Datos Opciones

Opciones gráfica

Distribución

☒ Normal

☐ t  $gl =$

☐ Chi-cuadrado  $gl =$

☐ F  $gl$  del numerador =   $gl$  del denominador =

☐ Otro Especificar:  Parámetros:

Identificar Observaciones

☒ Automáticamente

☐ Interactivamente con el ratón

☐ No identificar

Número de puntos a identificar

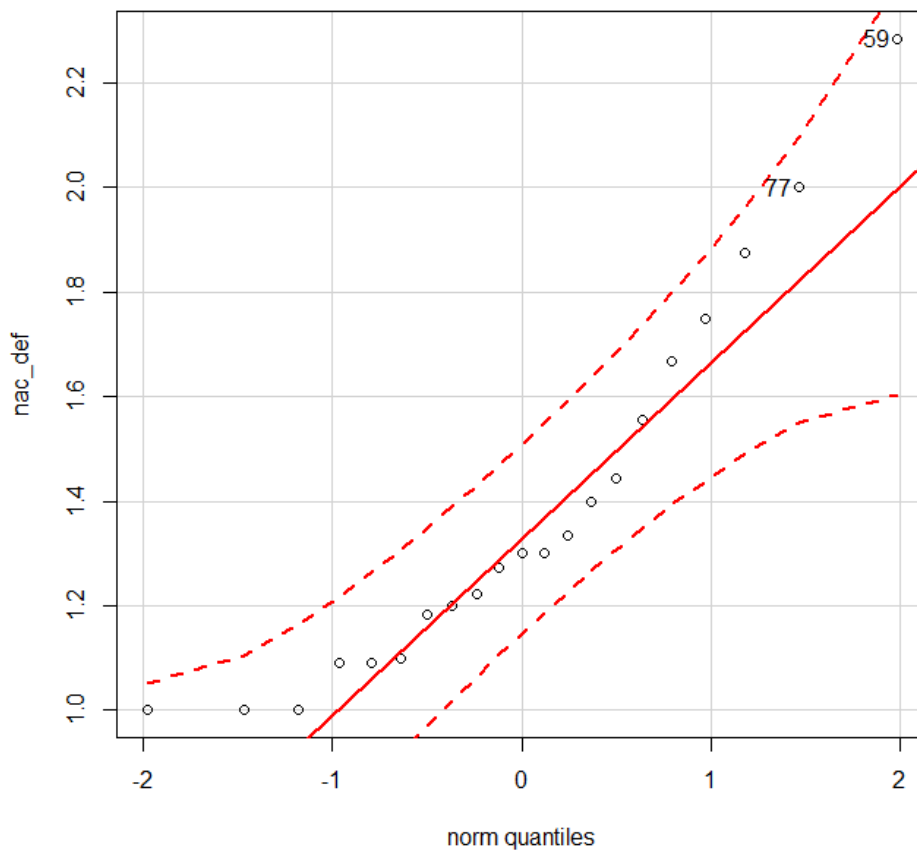
Dibujar etiquetas

Etiqueta del eje x

Etiqueta del eje y

Título del gráfico

Ayuda Reiniciar Aceptar Cancelar Aplicar



**Ejemplo 2:** : Un proceso químico debe producir cada día una cantidad media de 800 toneladas de un producto. El encargado de vigilar el buen funcionamiento del proceso sospecha que se está produciendo menos de esa cantidad y por ello ha observado las producciones diarias durante los últimos 5 días, que son: 802, 795, 752, 810, 783. ¿Hay evidencia para afirmar, al 5 % de significación, que la sospecha es correcta?

Datos → Nuevo conjunto de datos → Nombre: EJEMPLO2

Editor de datos: EJEMPLO2

Fichero Editar Ayuda

Añadir fila Añadir columna

	rowname	Producción
1	1	802
2	2	795
3	3	752
4	4	810
5	5	783

Ayuda Aceptar Cancelar

Estadísticos → Resúmenes → Test de normalidad

Test de Normalidad

Variable (elegir una)  
Producción

Test de Normalidad

☒ Shapiro-Wilk  
☐ Anderson-Darling  
☐ Cramer-von Mises  
☐ Lilliefors (Kolmogorov-Smirnov)  
☐ Shapiro-Francia  
☐ Chi cuadrado de Pearson

El número de divisiones para chi cuadrado de Pearson: <auto>

Test por grupos...

Ayuda Reiniciar Aceptar Cancelar Aplicar

Shapiro-Wilk normality test

```
data: Producción
W = 0.90797, p-value = 0.4555
```

Estadísticos → Medias → Test t para una muestra

Como vamos a hacer un **contraste**, nos da igual lo que pongamos en el nivel de confianza

Test t para una muestra

Variable (elegir una)  
Producción

Hipótesis alternativa

☐ Media poblacional != mu0  
☒ Media poblacional < mu0  
☐ Media poblacional > mu0

Hipótesis nula: mu = 800  
Nivel de confianza: .95

Ayuda Reiniciar Aceptar Cancelar Aplicar

One Sample t-test

```
data: Producción
t = -1.146, df = 4, p-value = 0.1578
alternative hypothesis: true mean is less than 800
95 percent confidence interval:
 -Inf 809.9791
sample estimates:
mean of x
 788.4
```

Como p-valor = 0,1578 >  $\alpha$  (0,05) → No se rechaza  $H_0$

Ahora pasamos a hacer un **intervalo de confianza**

Estadísticos → Medias → Test t para una muestra

## One Sample t-test

```
data: Producción
t = -1.146, df = 4, p-value = 0.3157
alternative hypothesis: true mean is not equal to 800
95 percent confidence interval:
 760.2961 816.5039
sample estimates:
mean of x
 788.4
```

**Ejemplo 4:** Se llevó a cabo un estudio para determinar el grado en el cual el alcohol entorpece la habilidad de pensamiento para llevar a cabo determinada tarea. Se seleccionaron al azar diez personas de distintas características y se les pidió que participaran en el experimento. Inicialmente, cada persona llevó a cabo la tarea sin nada de alcohol en su organismo y se midió el tiempo (en minutos) en realizarla. Posteriormente, la tarea volvió a llevarse a cabo después de que cada persona había consumido una cantidad suficiente de alcohol para tener un contenido en su organismo del 0.1 %. Los datos recogidos fueron los siguientes:

Sujeto	Tiempo antes	Tiempo después
1	28	29
2	22	35
3	55	57
4	45	51
5	32	36
6	35	58
7	40	51
8	25	34
9	37	48
10	20	30

¿Puede concluirse, al 5 %, que el consumo de alcohol incrementa el tiempo necesario para realizar la tarea?

Datos → Nuevo conjunto de datos → Nombre: EJEMPLO4

Editor de datos: EJEMPLO4

Fichero Editar Ayuda

Añadir fila Añadir columna

		1	2
	rowname	tiempo_antes	tiempo_despues
1	1	28	29
2	2	22	35
3	3	55	57
4	4	45	51
5	5	32	36
6	6	35	58
7	7	40	51
8	8	25	34
9	9	37	48
10	10	20	30

Ayuda Aceptar Cancelar

Estadísticos → Medias → Test t para datos relacionados

Test t para datos emparejados

Datos Opciones

Primera variable (elegir una) Segunda variable (elegir una)

tiempo\_antes tiempo\_antes

tiempo\_despues tiempo\_despues

Ayuda Reiniciar Aceptar Cancelar Aplicar

Test t para datos emparejados

Datos Opciones

Hipótesis alternativa Nivel de confianza

☐ Bilateral ☒ Diferencia < 0 ☐ Diferencia > 0

.95

Ayuda Reiniciar Aceptar Cancelar Aplicar

**PARA RESOLVER EL CONTRASTE, SIEMPRE SE COGE EL “p-value”**

Paired t-test

```
data: tiempo_antes and tiempo_despues
t = -4.4508, df = 9, p-value = 0.0007992
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf -5.293262
sample estimates:
mean of the differences
-9
```

Respuesta: Sí, se puede

Media muestral = -9

**Ejemplo 5:** Dividimos aleatoriamente un conjunto de pacientes afectados por cierto virus en dos grupos a los que se administra respectivamente un placebo y un tratamiento en estudio para combatir dicha afección vírica. Después de tratar a los pacientes durante 2 meses se mide la concentración de virus de cada uno de ellos. Los resultados se muestran en la siguiente tabla:

Sujeto	Administración	Nivel	Sujeto	Administración	Nivel
1	Placebo	23	2	Placebo	25
3	Placebo	26	4	Placebo	24
5	Placebo	26	6	Placebo	24
7	Placebo	22	8	Placebo	25
9	Placebo	27	10	Placebo	25
11	Tratamiento	22	12	Tratamiento	23
13	Tratamiento	22	14	Tratamiento	24
15	Tratamiento	19	16	Tratamiento	20
17	Tratamiento	21	28	Tratamiento	23
19	Tratamiento	22	20	Tratamiento	23

¿Puede decirse, al 5 % de significación, que el tratamiento hace el efecto deseado en la infección, es decir, que disminuye el nivel de virus?

Datos → Nuevo conjunto de datos → Nombre: EJEMPLO5

Editor de datos: EJEMPLO5

Fichero Editar Ayuda

Añadir fila Añadir columna

	1	2
rowname	Nivel	Administración
1	23	p
2	25	p
3	26	p
4	24	p
5	26	p
6	24	p
7	22	p
8	25	p
9	27	p
10	25	p
11	22	t
12	23	t
13	22	t
14	24	t
15	19	t
16	20	t
17	21	t
18	23	t
19	22	t
20	23	t

Ayuda Aceptar Cancelar

Estadísticos → Varianzas → Test F para dos varianzas

Test F de dos varianzas

Datos Opciones

Grupos (elegir uno)

Administración

Variable explicada (elegir una)

Nivel

Ayuda Reiniciar Aceptar Cancelar Aplicar



**Test F de dos varianzas**

**Datos** Opciones

Cociente: p / t

Hipótesis alternativa

☒ Bilateral

☐ Cociente < 1

☐ Cociente > 1

Nivel de confianza: .95

Ayuda Reiniciar Aceptar Cancelar Aplicar

F test to compare two variances

```
data: Nivel by Administración
F = 0.96172, num df = 9, denom df = 9, p-value = 0.9546
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.2388783 3.8718891
sample estimates:
ratio of variances
 0.9617225
```

Como  $p\text{-valor} = 0,9546 > \alpha \rightarrow$  No se rechaza  $H_0 \rightarrow \sigma x^2 = \sigma y^2$

Estadísticos  $\rightarrow$  Medias  $\rightarrow$  Test t para muestra independientes

**Test t para muestras independientes**

**Datos** Opciones

Grupos (elegir uno)

Administración

Variable explicada (elegir una)

Nivel

Ayuda Reiniciar Aceptar Cancelar Aplicar

**Test t para muestras independientes**

**Datos** Opciones

Diferencia: p - t

Hipótesis alternativa

☐ Bilateral

☒ Diferencia < 0

☐ Diferencia > 0

Nivel de confianza

.95

¿Suponer varianzas iguales?

☒ Sí

☐ No

Ayuda Reiniciar Aceptar Cancelar Aplicar

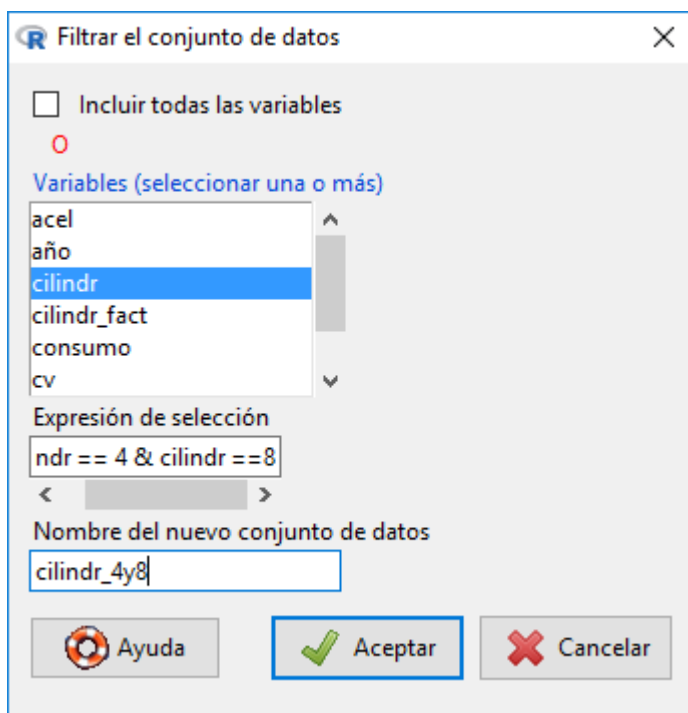
```
data: Nivel by Administración
t = 4.1485, df = 18, p-value = 0.9997
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf 3.970404
sample estimates:
mean in group p mean in group t
      24.7          21.9
```

**Ejercicio 6:** Cargar el fichero de datos Rcars. Plantear y resolver contrastes de hipótesis adecuados para estudiar si, para un nivel de significación del 10 %, existen evidencias significativas para afirmar que el consumo medio de los coches con 8 cilindros es superior al consumo medio de los coches con 4 cilindros.

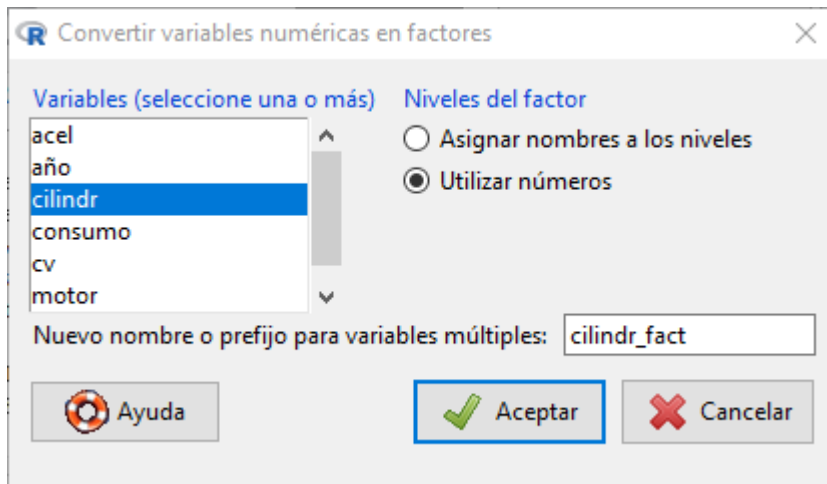
**Nota:** Los menús que permiten realizar contrastes sobre las diferencias de medias y los cocientes de varianzas de dos poblaciones únicamente se activan cuando en el conjunto de datos existe una variable categórica con, exactamente, dos modalidades. Esta variable es la que permite definir qué casos corresponden a cada una de las poblaciones. Por ello, para realizar el ejercicio, será necesario realizar los siguientes pasos:

- Filtrar el conjunto de datos activo para construir un nuevo conjunto de datos que contenga, únicamente, los casos correspondientes a los coches que tienen 4 u 8 cilindros.
- Convertir la variable *cilindr* del nuevo conjunto de datos en una variable cualitativa mediante el menú Datos→Modificar variables del conjunto de datos activo→Convertir variable numérica en factor.
- Realizar los contrastes de hipótesis necesarios para resolver el problema.

Datos → Conjunto de datos activo → Filtrar el conjunto de datos activo...

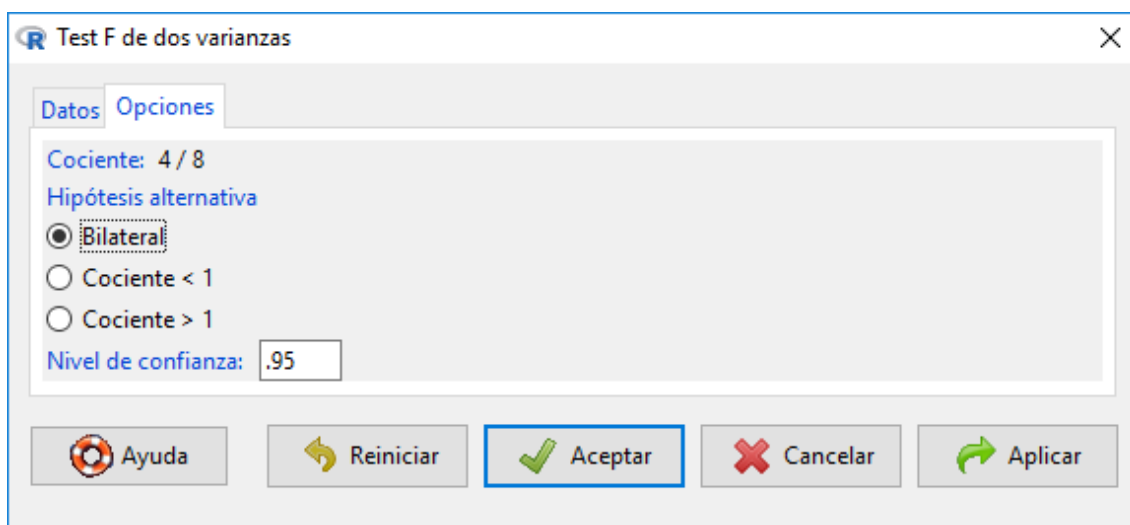
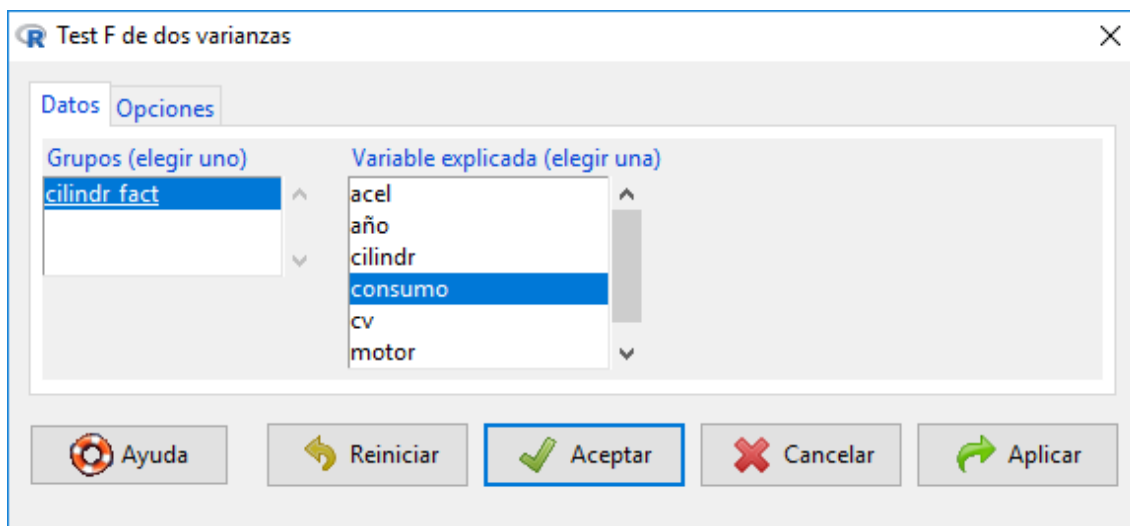


Datos → Modificar Conjunto de datos activo → Convertir variable numérica a factor



Ahora podremos hacer la varianza:

Estadísticos → Varianza → Test F para dos varianzas



Estadísticos → Varianza → Test t para muestras independientes

**Test t para muestras independientes**

**Datos** **Opciones**

Grupos (elegir uno): cilindr\_fact

Variable explicada (elegir una): consumo

ayuda Reiniciar Aceptar Cancelar Aplicar

**Test t para muestras independientes**

**Datos** **Opciones**

Diferencia: 4 - 8

Hipótesis alternativa: ☒ Bilateral ☐ Diferencia < 0 ☐ Diferencia > 0

Nivel de confianza: .95

¿Suponer varianzas iguales? ☐ Sí ☒ No

ayuda Reiniciar Aceptar Cancelar Aplicar

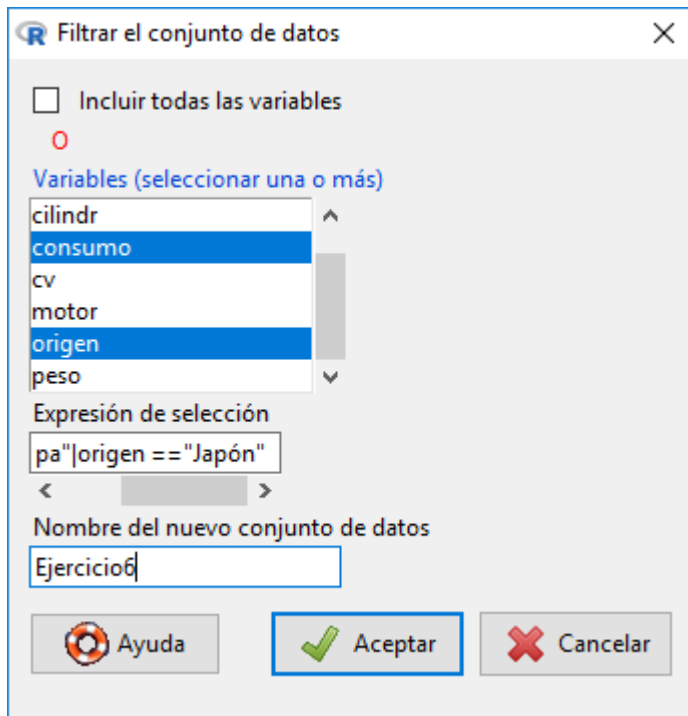
Welch Two Sample t-test

```
data: consumo by cilindr_fact
t = -26.847, df = 137.48, p-value < 2.2e-16
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf -7.405345
sample estimates:
mean in group 4 mean in group 8
 8.323529      16.215686
```

**Ejercicio 7:** Cargar el fichero de datos Rcars. Plantear y resolver contrastes de hipótesis adecuados para estudiar si, para un nivel de significación del 5%, existen evidencias significativas para afirmar que el consumo medio de los coches con europeos es distinto del consumo medio de los coches japoneses.

**Nota:** En este caso, como paso previo a la resolución del problema, tenemos que construir un nuevo conjunto de datos que contenga únicamente los casos correspondientes a los coches europeos y japoneses. El primer paso consistiría pues en filtrar el conjunto de datos original para seleccionar estos casos. Si una vez hecho esto solicitamos un resumen del conjunto de datos activo observamos que, en el nuevo conjunto de datos, la variable origen sigue presentando tres modalidades: Europa, Japón y E.E.U.U., si bien la modalidad E.E.U.U. tiene frecuencia 0.

Datos → Conjunto de datos activo → Filtrar el conjunto de datos activos



Filtrar el conjunto de datos

☐ Incluir todas las variables

Variables (seleccionar una o más)

cilindr  
consumo  
cv  
motor  
origen  
peso

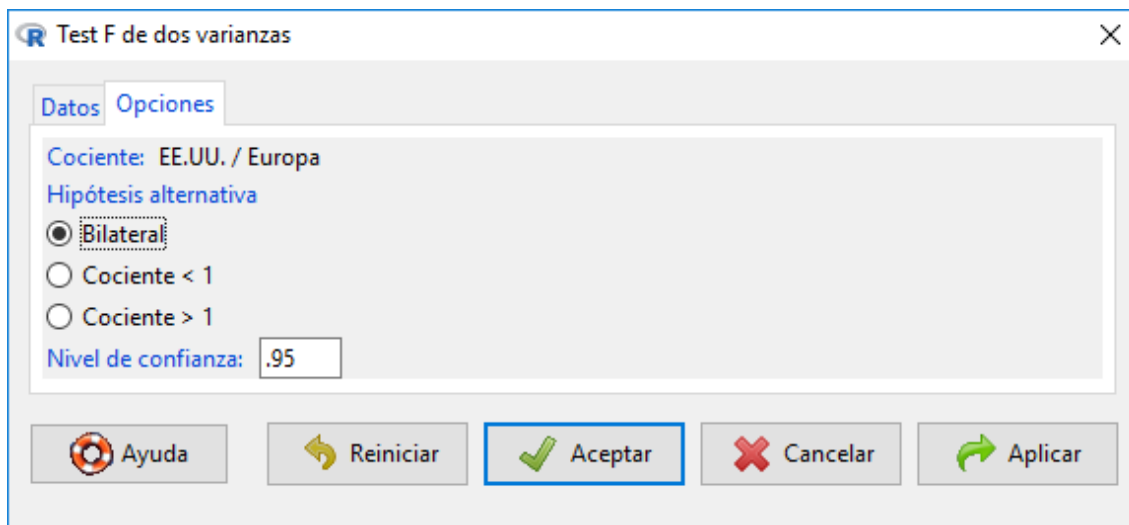
Expresión de selección  
pa|origen == "Japón"

Nombre del nuevo conjunto de datos  
Ejercicio6

Ayuda Aceptar Cancelar

//MIRAR EN CASA, NO ES ASÍ

Estadísticos → Varianza → Test F para dos varianzas



Test F de dos varianzas

Datos Opciones

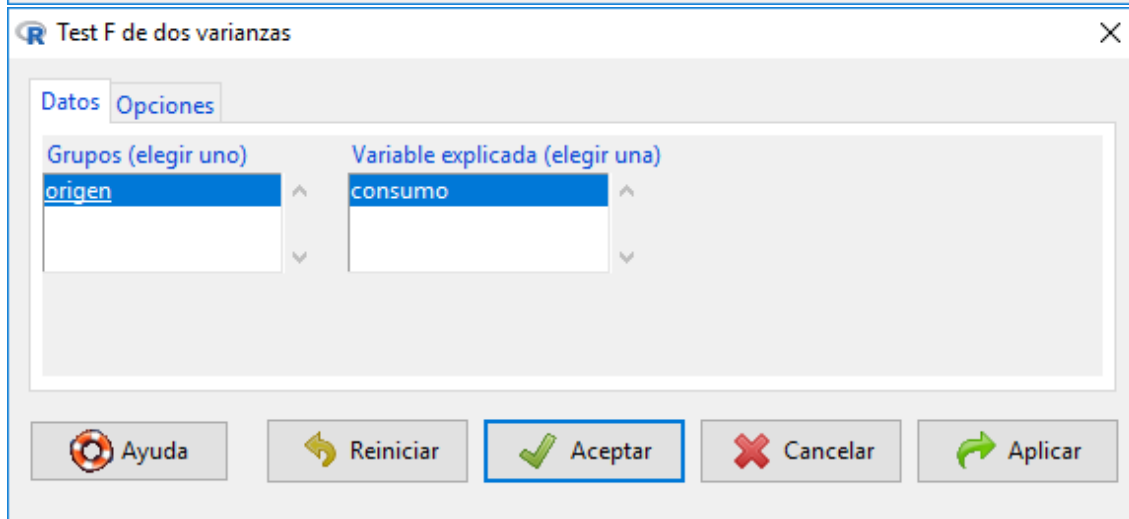
Cociente: EE.UU. / Europa

Hipótesis alternativa

☒ Bilateral  
☐ Cociente < 1  
☐ Cociente > 1

Nivel de confianza: .95

Ayuda Reiniciar Aceptar Cancelar Aplicar



Test F de dos varianzas

Datos Opciones

Grupos (elegir uno)

origen

Variable explicada (elegir una)

consumo

Ayuda Reiniciar Aceptar Cancelar Aplicar

