

Cápsula 0006

«Seleccionar casos en SPSS»

1. Los datos

Los datos que estamos usando para este material provienen de un estudio hipotético que examina cuánto tiempo les toma a las personas quedarse dormidas durante una lección de estadística.

	Duration	Sex	Height	var	var	var	var	var	var
1	3.86	Female	67						
2	6.66	Male	61						
3	10.21	Female	66						
4	6.21	Female	64						
5	11.82	Female	73						
6	5.81	Female	65						
7	7.49	Male	65						
8	7.71	Male	65						
9	5.93	Male	65						
10	10.72	Female	72						
11	7.43	Male	70						
12	6.99	Female	74						
13	7.10	Female	69						
14	10.39	Male	64						
15	5.09	Male	71						
16	3.24	Male	71						
17	7.87	Female	68						
18	6.26	Female	62						
19	5.99	Female	65						
20	6.17	Male	56						

Las dos variables que nos interesan aquí son el sexo, ya sea masculino o femenino, y la duración, que es la cantidad de minutos que transcurren desde el inicio de una lección de estadística antes de que un sujeto se duerma.

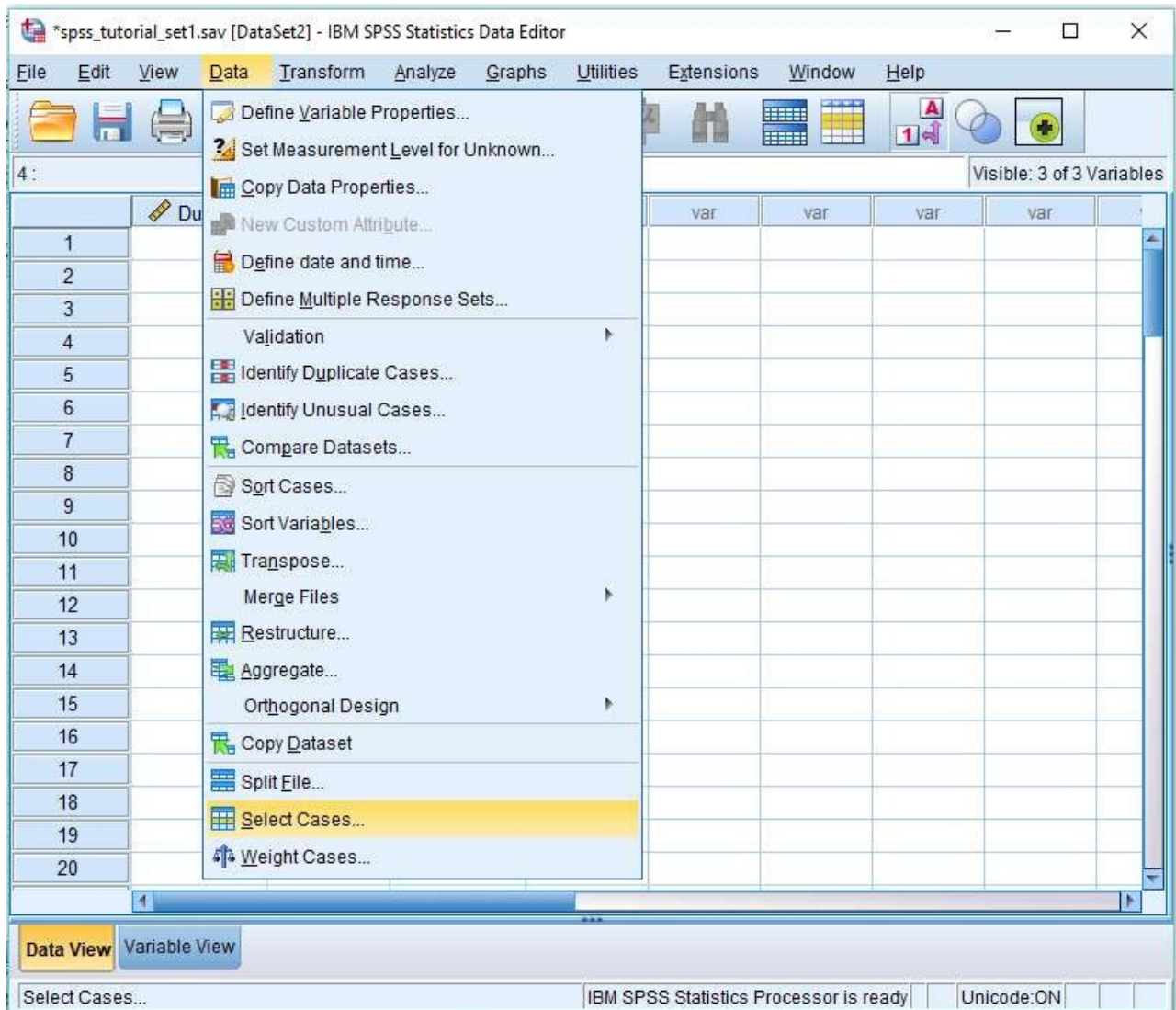
Imagine que ya sabemos que, en la población en su conjunto, la cantidad de tiempo promedio que toma una mujer para conciliar el sueño son 8,15 minutos. Queremos comparar esto con el tiempo promedio de las mujeres de nuestra muestra. Pero el problema es que nuestra muestra contiene datos tanto para hombres como para mujeres, y cualquier prueba que realicemos se basará en esa base. La pregunta es

¿cómo seleccionamos solo casos femeninos, excluyendo así a los hombres de las pruebas que realizamos?

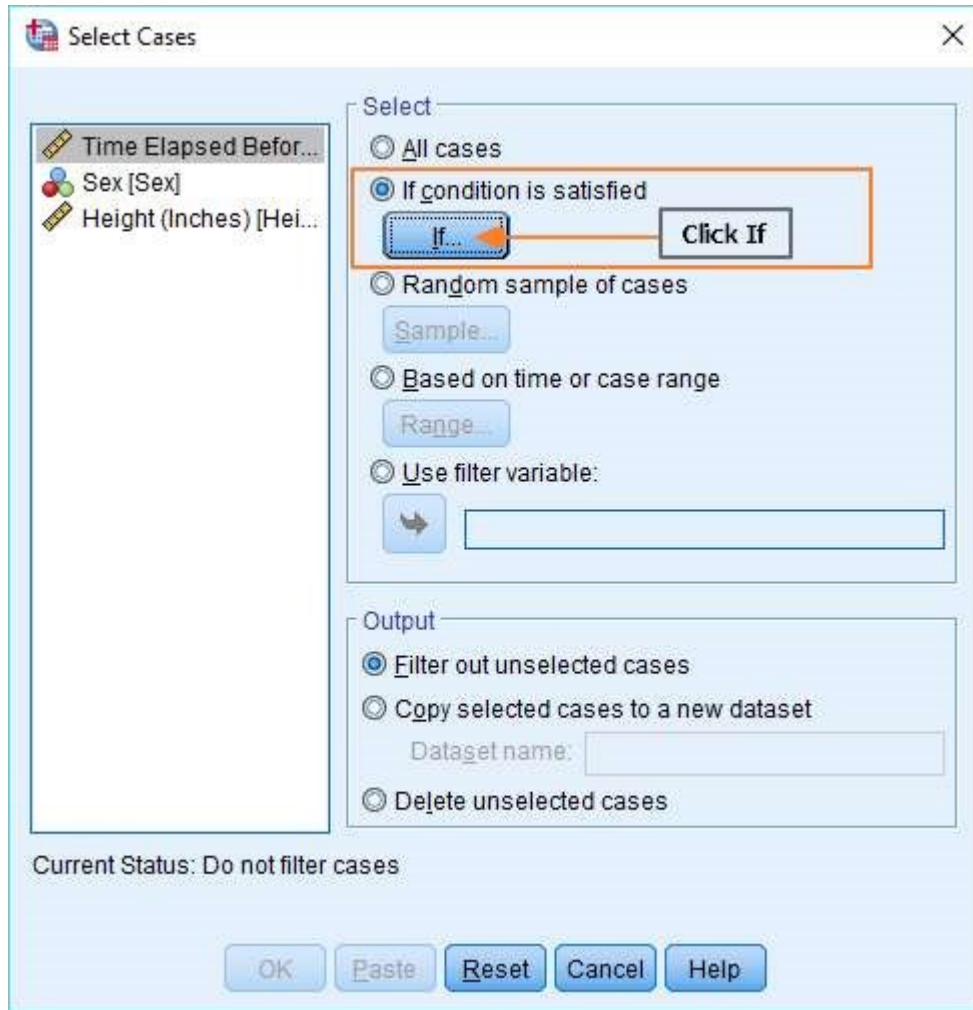
2. Seleccionar casos

Aquí es donde la funcionalidad de casos seleccionados resulta útil.

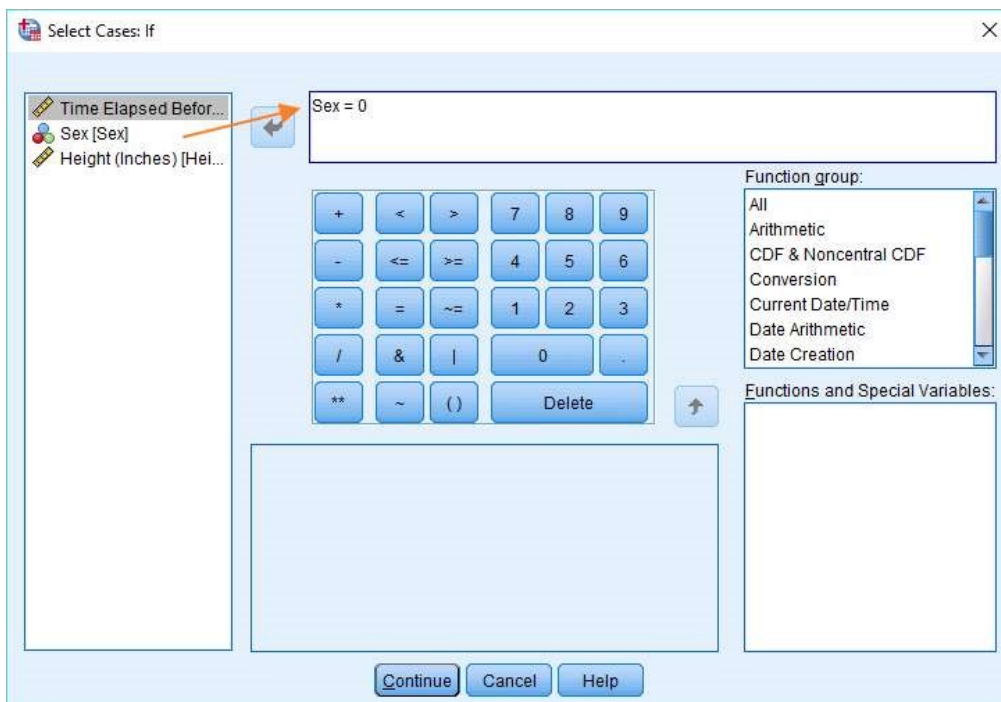
Para comenzar, haga clic en Datos -> Seleccionar casos.



Esto abrirá el cuadro de diálogo Seleccionar casos. Esto proporciona una serie de opciones diferentes para seleccionar casos. Nos centraremos en la opción «Si se cumple la condición», que debe seleccionar.



Una vez que lo haya seleccionado, debe hacer clic en el botón Si... (como se muestra arriba). Aparecerá el cuadro de diálogo Seleccionar casos: Si. Aquí es donde haces el trabajo de seleccionar casos exclusivos para mujeres.



La idea aquí es construir una expresión en el cuadro de texto en la parte superior que funcione para seleccionar casos. Aquí puede ver que tenemos "Sexo = 0", que le dice a SPSS que solo debe seleccionar los casos donde el valor de la variable Sexo es 0 (Mujer = 0, Hombre = 1).

Obviamente, es posible construir expresiones mucho más complejas que esta simple prueba de equivalencia. Por ejemplo, podría indicarle a SPSS que seleccione los casos en los que el sexo es femenino y la altura es superior a 68 pulgadas («Sexo = 0 y altura > 68»), o donde la duración es superior a 8 minutos o la altura es inferior a 60 pulgadas («Duración > 8 | Altura < 60 »).

Una vez que haya configurado la expresión, como se indicó anteriormente, presione el botón Continuar y luego haga clic en Aceptar en el cuadro de diálogo Seleccionar casos. SPSS ahora seleccionará los casos según sus instrucciones.

3. Casos seleccionados

Si echa un vistazo a la Vista de datos, verá que las cosas han cambiado para indicar que SPSS ahora está funcionando con un subconjunto del conjunto de datos original.

The screenshot shows the IBM SPSS Statistics Data Editor window. The title bar reads '*spss_tutorial_set1.sav [DataSet2] - IBM SPSS Statistics Data Editor'. The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Extensions, Window, and Help. The toolbar contains various icons for file operations and data manipulation. The main data grid shows 20 rows of data with columns for Duration, Sex, Height, and filter_\$. The 'filter_\$' column is highlighted in orange, and the 'Data View' tab is selected. The status bar at the bottom indicates 'IBM SPSS Statistics Processor is ready' and 'Unicode:ON Filter On'.

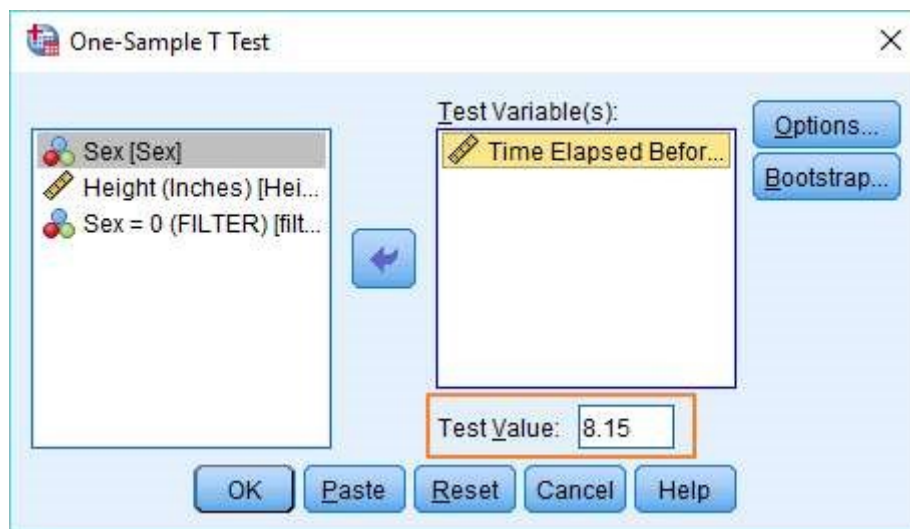
	Duration	Sex	Height	filter_\$	var	var	var	var
1	5.50	Female	66	Selected				
2	5.50	Male	66	Not Selected				
3	5.51	Female	67	Selected				
4	5.51	Female	68	Selected				
5	6.26	Female	62	Selected				
6	6.26	Female	70	Selected				
7	7.11	Male	61	Not Selected				
8	7.11	Male	64	Not Selected				
9	2.09	Female	75	Selected				
10	2.94	Male	65	Not Selected				
11	3.24	Male	71	Not Selected				
12	3.86	Female	67	Selected				
13	3.98	Male	68	Not Selected				
14	4.15	Male	71	Not Selected				
15	4.18	Female	71	Selected				
16	4.33	Female	67	Selected				
17	4.41	Female	65	Selected				
18	4.70	Female	65	Selected				
19	4.88	Male	73	Not Selected				
20	5.09	Male	71	Not Selected				

Como puede ver, SPSS ha tachado los casos de la izquierda que no están seleccionados. También ha introducido una nueva variable de filtro que especifica si se ha seleccionado un caso o no. Finalmente, en la parte inferior derecha, dice Filter On, que le indica que cualquier prueba o análisis que ejecute estará en un subconjunto de los datos, es decir, solo en los casos seleccionados.

4. Prueba T de una muestra

Veamos esto ejecutando una prueba t de una muestra para comparar la cantidad promedio de tiempo que las mujeres de la población general tardan en quedarse dormidas en una lección de estadística con el promedio de las mujeres de nuestra muestra.

Haga clic en Analizar -> Comparar medias -> Prueba T de una muestra, y luego configure la prueba de esta manera.

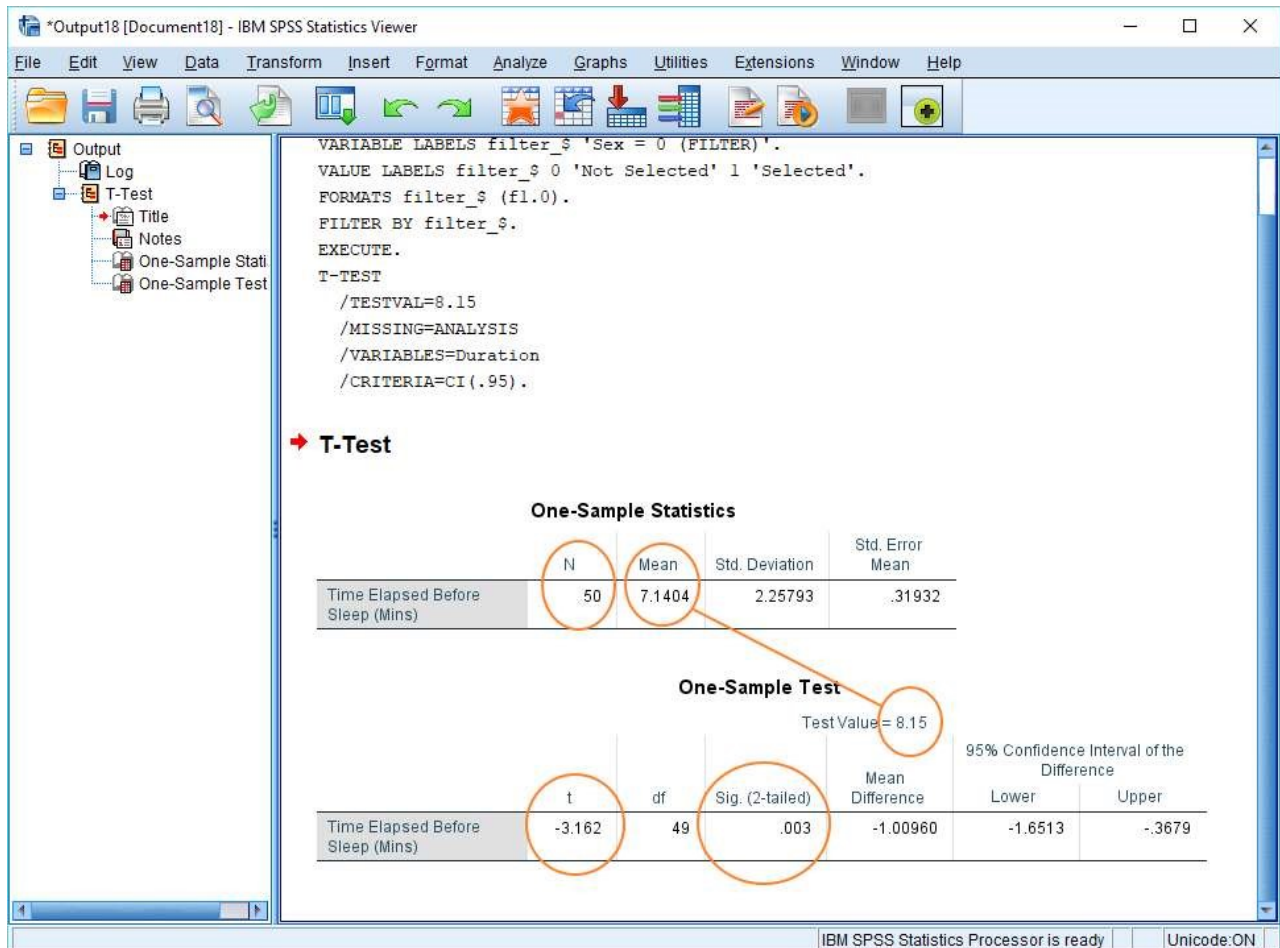


Puede ver que tenemos Duración como nuestra variable de prueba, y la estamos comparando con una media poblacional de 8.15 minutos (la cantidad promedio de tiempo que les toma a las mujeres de la población general quedarse dormidas en una lección de estadística).

Presiona OK para ejecutar la prueba.

5. El resultado

Este es el resultado.



El valor de N aquí es 50, lo que le indica inmediatamente que los casos seleccionados han funcionado. Nuestro conjunto de datos contiene 100 casos, de los cuales 50 son mujeres.

En términos del resultado, podemos ver que las mujeres de nuestra muestra se duermen en promedio 1 minuto más rápido que las mujeres de la población general. Esta es una diferencia significativa, con un valor de -3,1 y un *pag*-valor de .003.

NOTA FINAL:

Para devolver un conjunto de datos a su estado no filtrado volviendo al cuadro de diálogo Seleccionar casos (Datos -> Seleccionar casos) y eligiendo Todos los casos (la primera opción disponible). Esto no eliminará la nueva variable de filtro, pero la dejará inactiva. También notará que «Filtro activado» ya no se mostrará en la parte inferior derecha de la Vista de datos.