*Noviembre 2019, Ejemplos*

**Enunciado 1**: Queremos examinar hasta qué punto el grado de ingenio que asignamos a una frase difiere cuando la frase la creó alguien de quien tenemos buena opinión o no, y a la vez analizar si este efecto difiere en universitarios y no-universitarios. El experimentador selecciona 10 frases ocurrentes, en cinco de ellas indica que era de una persona con mala prensa (Hitler, …) y en otras cinco indica que eran de una persona con buena prensa (Dalai Lama, …). Participan 40 personas entre 20 y 25 años (20 universitarios, 20 no-universitarios), que ven todas las frases, en orden aleatorio, e indican el grado de ingenio de cada una de las frases en una escala de 1 a 10. Lo que se analiza en el estudio es, para cada participante, el promedio de ingenio en las cinco frases de las personas con buena prensa y el promedio de ingenio de las cinco frases de las personas con mala prensa.

1--¿Cuál es/son la/s variable/s independiente/s?

2--¿Cuál es/son la/s variable/s dependiente/s?

3--¿Qué tipo de diseño se ha empleado, entre-sujetos, intra-sujetos o mixto? Justifícalo.

4--¿Es un diseño experimental? Justifícalo

5--¿Cómo organizarías los datos con SPSS respecto a las filas y columnas? Justifícalo.

6--¿Cómo estamos considerando la variable “grado de ingenio” en el estudio?

7--Pensemos que el grado de ingenio aparece como más alto para las personas con buena prensa que para los de mala prensa. Y que otro investigador dice que necesitamos una línea base que permita saber si la diferencia se debe a que hay un sesgo a favor de las personas con buena prensa, un sesgo en contra de las personas con mala prensa, o que en realidad se dan ambos sesgos. ¿Qué nuevo estudio realizarías para contestar dicha pregunta?

8— Pensemos que el grado de ingenio aparece como más alto para las personas con buena prensa que para los de mala prensa. Damos una charla con esos datos, y una investigadora indica un posible problema: Quizás las cinco frases que pusimos como de las personas de mala prensa eran en realidad menos ingeniosas que las otras cinco. ¿Cómo solucionarías dicho problema en un nuevo estudio?

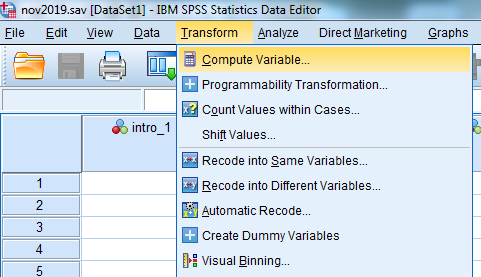
**Enunciado 2**. Tenemos un conjunto de datos con diversas variables: introversión (4 preguntas, la tercera es inversa; cada pregunta puntuaba de 1 a 10), ansiedad, maquiavelismo, edad y si son de ciudad grande (más de 50,000 habitantes; “1” en el Excel) o no (“2” en el Excel). Los datos están en <http://www.uv.es/mperea/nov2019.sav>

(soluciones al final del archivo)

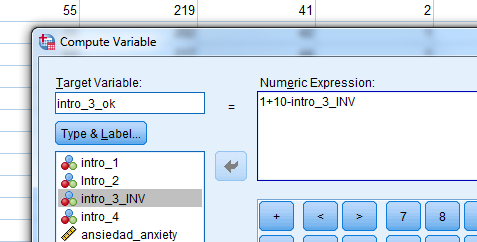
1. Crea la variable “introversión\_total” como la suma de las cuatro puntuaciones de introversión. Recuerda que ahora mismo la tercera es inversa. Una vez hecho esto, ¿hay diferencias en introversión entre las personas de ciudad grande y las que no? Efectúa el diagrama de caja e indica los estadísticos que consideres pertinentes.
2. Dicotomiza la variable edad en “menores de 40 años” (valor 1) y “40 o más” (valor 2). Una vez hecho esto, ¿hay diferencias en ansiedad entre el grupo de más jóvenes y los más mayores? Efectúa el diagrama de caja e indica los estadísticos que consideres pertinentes.
3. Queremos detectar a las personas más ansiosas, especialmente al 5% de personas más ansiosas. ¿Qué puntuación en ansiedad constituirá el punto de corte?
4. ¿Es similar el número de personas de ciudad grande con las que no en la muestra? Realiza el gráfico correspondiente. Puedes obtener también la tabla de frecuencias.
5. ¿La distribución de maquiavelismo en la muestra sigue la distribución normal? En caso que no, qué tipo de transformación harías?
6. ¿Qué puntuación T corresponde a la persona de la fila primera en Ansiedad? ¿Qué quiere decir?
7. Queremos seleccionar solamente a las personas de más de 40 años. ¿Encuentras diferencias en maquiavelismo entre las personas de ciudad grande y las que no? Efectúa el diagrama de caja e indica los estadísticos que consideres pertinentes.

Enunciado 2. Soluciones

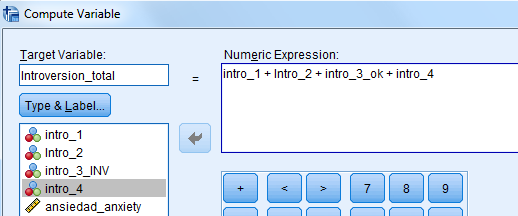
Pregunta 1. En primer lugar hay que hacer el ítem tercero de Introversión en el orden correcto. Que es ir al menú de Transformar y de ahí Computar Variable:



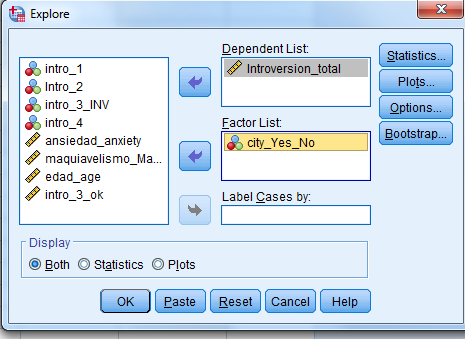
E indicar la fórmula. Dado que es una escala de 1-10, entonces el valor nuevo (correcto) del ítem 3 será: 1+10-ValorAntiguoItem3



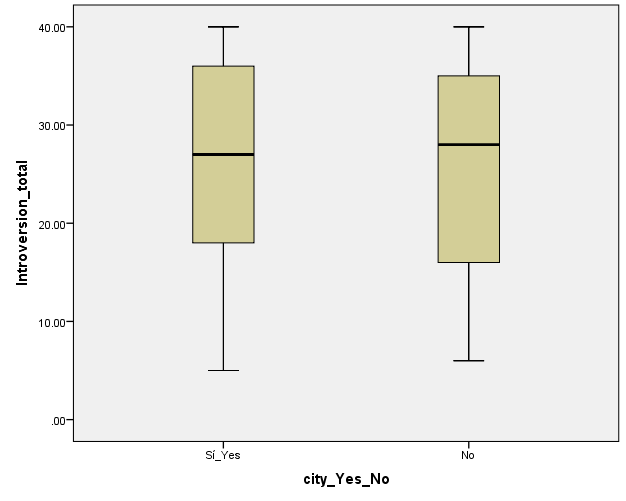
Y ahora es hacer la suma de las cuatro preguntas, como se pide en la pregunta:

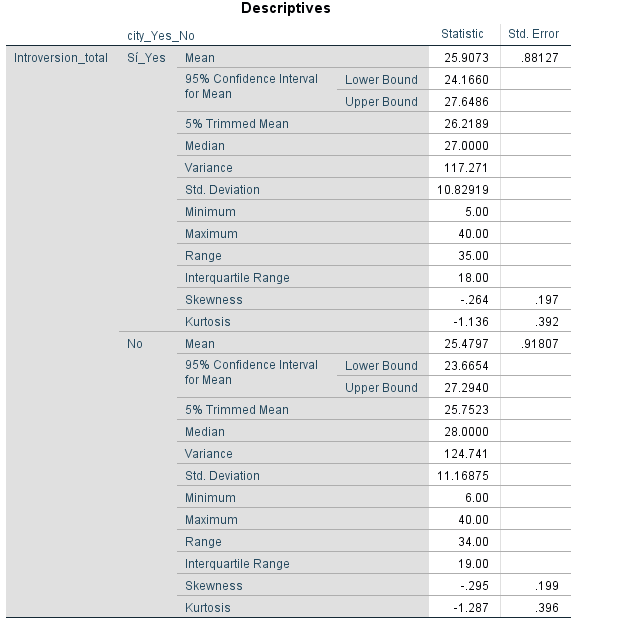


Y ahora es ver si el nivel de introversión difiere entre las personas que viven en ciudad grande o no, mediante un diagrama de caja yendo a Analizar/Descriptivos/Explorar.

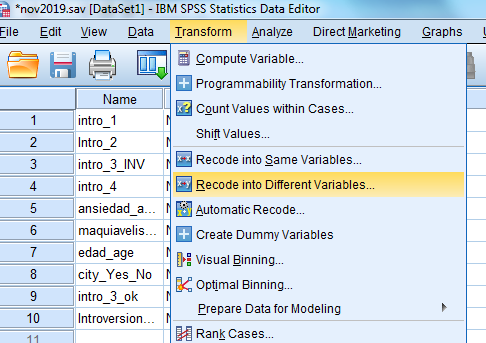


Como se aprecia en el diagrama, los niveles de introversión son muy similares para las personas que viven en ciudades grandes y las que no. La media de introversión en ambos grupos es muy parecida: 25.9 vs. 25.5, respectivamente. (Pensad que no hay motivo por el que las personas que vivan en ciudad grande tengan diferencias en introversión con los que no.)

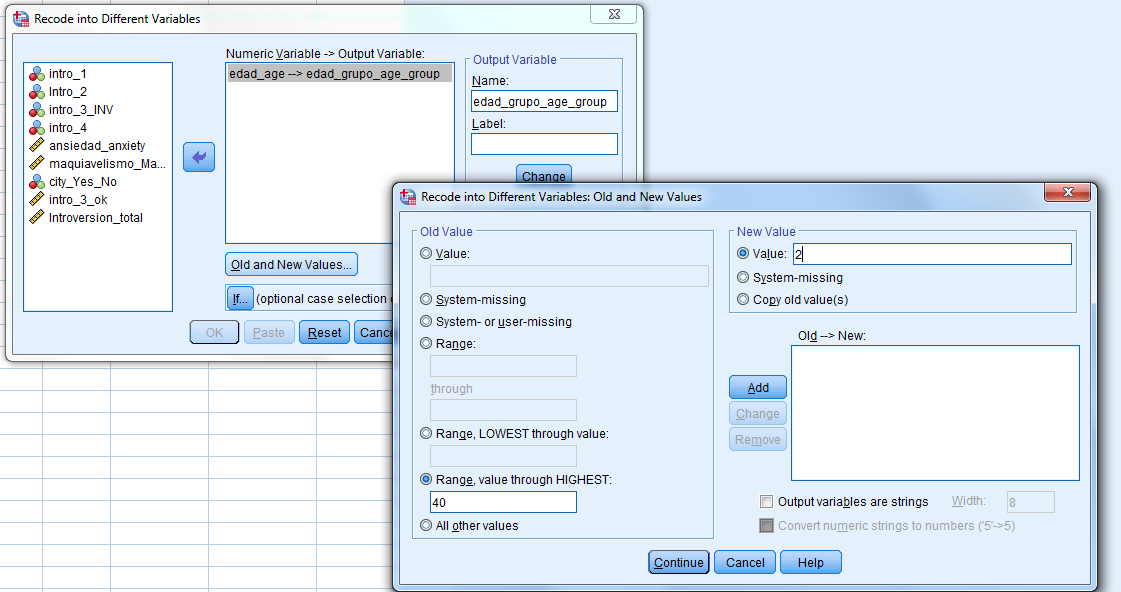


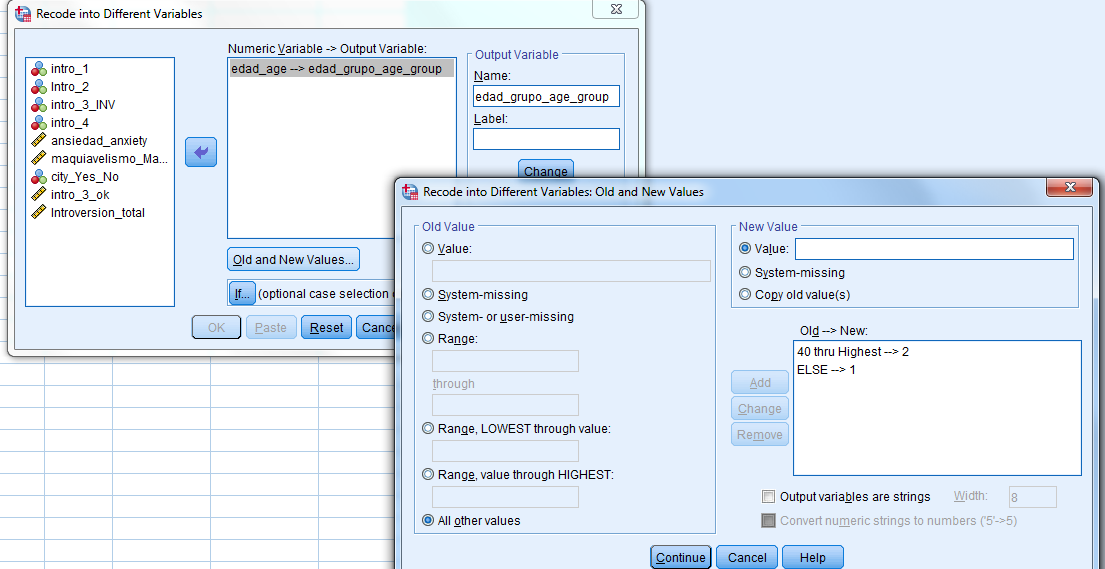


2. Hemos de dicotomizar primero la variable “edad” en dos grupos. Que es ir a Transformar y de ahí a Recodificar en Variables Diferentes:

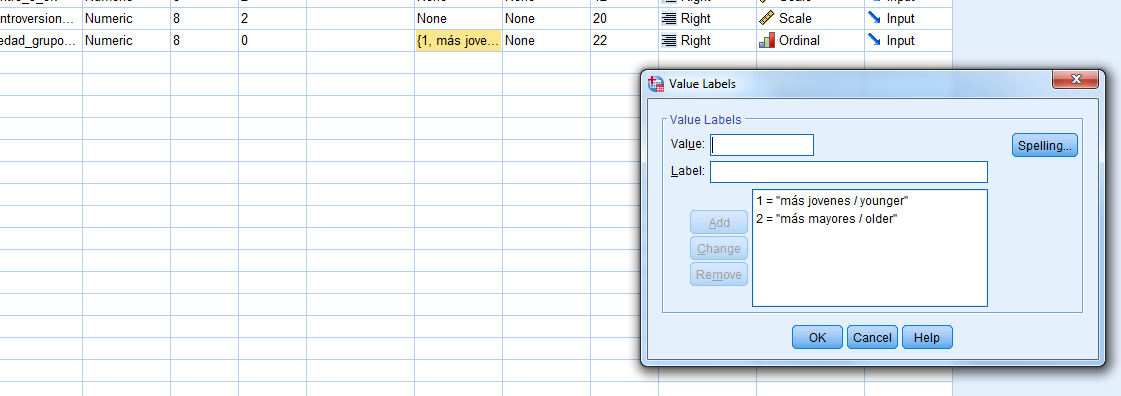


Ahora hay que elegir un nombre para la nueva variable “edad\_grupos”, y los de 40 años o más serán 2, y el resto 1:



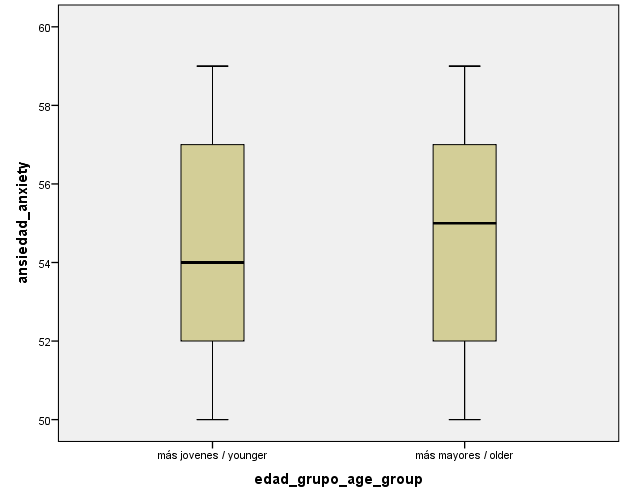


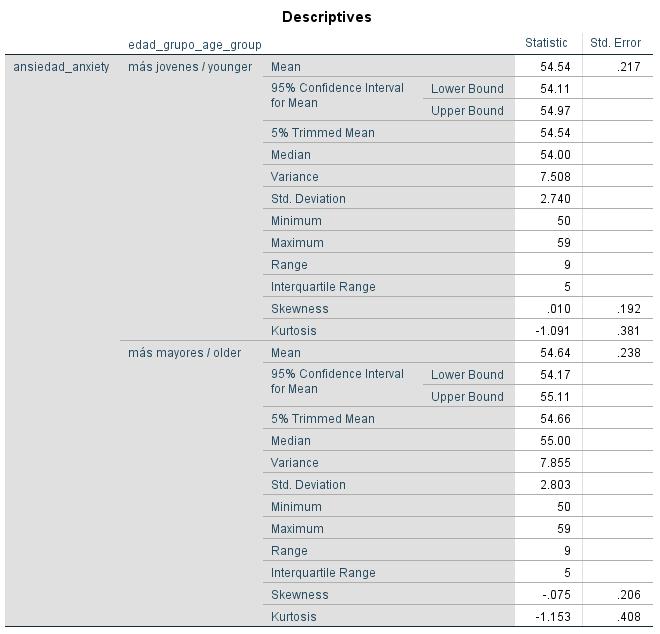
Y ya lo tenemos. Podéis añadir en “Vista de Variable” que 1 son los más jóvenes y 2 los más mayores para esta nueva variable, y que es “Ordinal”.



Y ahora es ver si hay diferencias en Ansiedad entre ambos grupos en Analizar/Explorar:

Como se ve en el diagrama de caja, no se aprecian diferencias en ansiedad entre ambos grupos: Las medias son virtualmente iguales para los más jóvenes y para los más mayores (54.5 vs. 54.6, respectivamente).



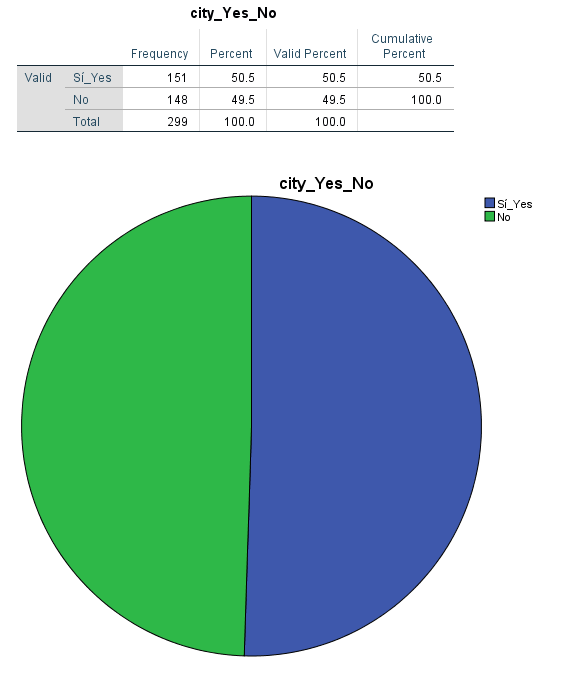


3. Lo que queremos es saber qué puntuación deja por encima el 5% de datos, que es paralelo a decir qué puntuación deja por debajo el 95% de los datos. Es decir, el punto de corte es el Percentil 95. Es ir a Analizar-Descriptivos-Frecuencias, y de ahí es clic en botón de Estadísticos. Y elegir el Percentil 95.

La respuesta es 59, que es el punto de corte que se pide.

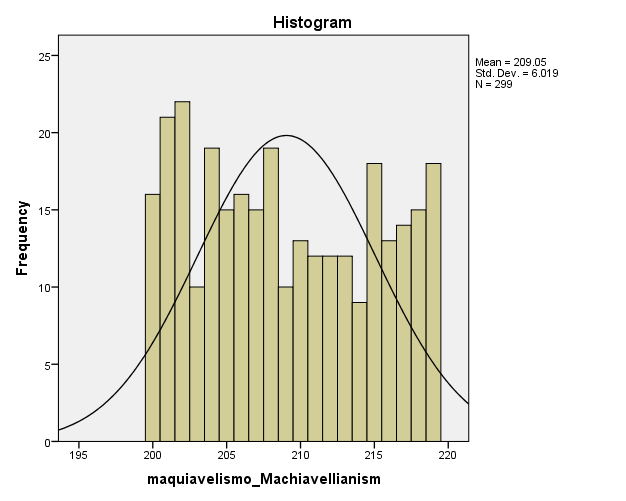
4. Aquí se pide si el número de personas que viven en una ciudad grande en la muestra es similar a los que no. Esto se puede ver fácilmente en Diagrama Circular o en el Diagrama de Barras. También en la Tabla de frecuencias. Es ir a Analizar/Descriptivos/Frecuencias e ir a Gráficos. Elegir uno de los dos que indicamos antes. Por defecto SPSS indica la tabla de frecuencias.

Como se aprecia tanto en la tabla de frecuencias como en el gráfico circular, pues prácticamente la mitad de las personas vive en ciudad grande y el resto no.



5. Para ver si la distribución de una variable cuantitativa sigue una distribución normal, lo mejor es obtener el Histograma (Analizar/Descriptivos/Frecuencias y elegir Histograma en y la Opción de Mostrar curva normal dentro del botón “Gráficos).

Lo que sale es:



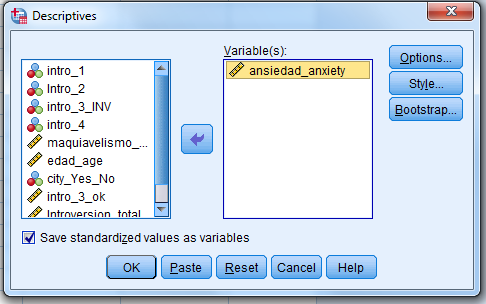
La distribución de los datos claramente no se ajusta a una distribución normal. Ante la pregunta de si se puede hacer alguna transformación, los casos que hemos visto en el Temario era si la distribución tenía asimetría positiva vs. negativa. Pero en este caso no se aprecia una forma definida. Por tanto, no se puede hacer mucho...

Pregunta 6. Las puntuaciones T son escalas derivadas de la distribución normal. En concreto: Ti=50+10\*zi

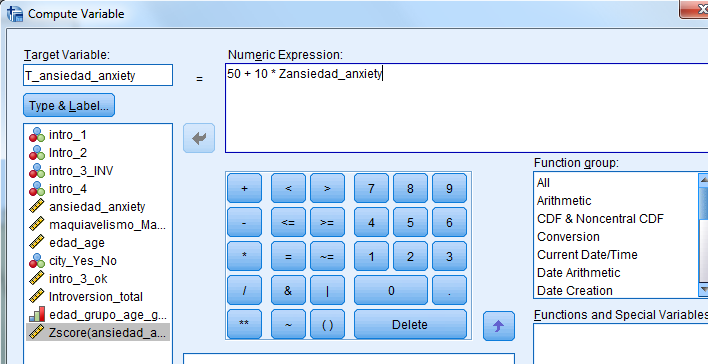
La media de las puntuaciones T es 50 y la desviación típica es 10.

Un ejemplo de interpretación: Una puntuación T de 60 es que la persona está UNA DESVIACIÓN TÍPICA (10 puntos sobre la media que es 50) ENCIMA DE LA MEDIA.

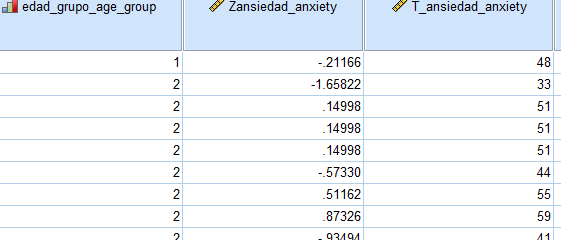
En el caso que nos ocupa hemos primero de tener las puntuaciones típicas de Ansiedad para todos los participantes. Que es ir a Analizar/Descriptivos/Descriptivos e indicar la opción de puntuaciones tipificadas:



Y ahora hemos de aplicar la fórmula de arriba para obtener las puntuaciones T (empleando el menú de Transformar y de ahí a Calcular Variable):



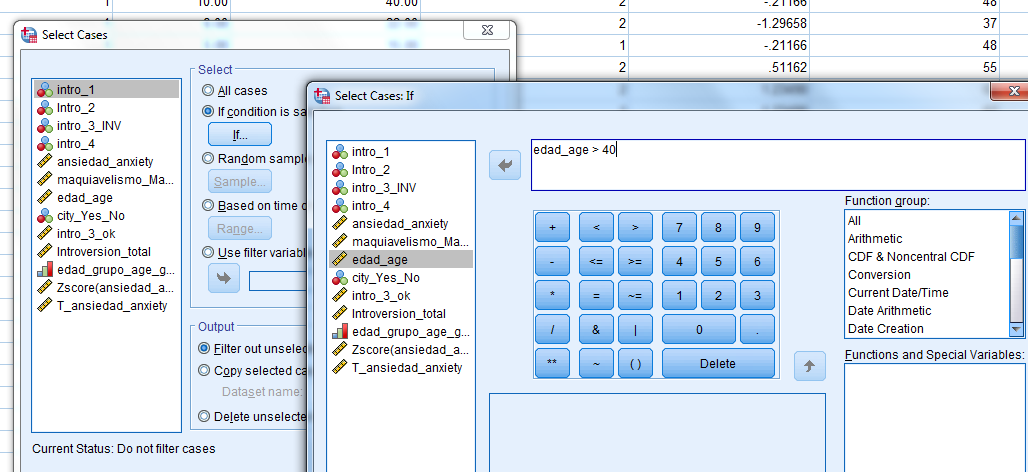
Y ahora es mirar la puntuación T de la persona en la primera fila (recordad quitar decimales en las puntuaciones T vía “Vista de Variable” porque no se suelen emplear)



La puntuación T de la persona de la primera fila es 48, que está muy ligeramente por debajo (2 unidades) de la media (que es 50). Por tanto, el grado de ansiedad de esta persona entra dentro de lo habitual, lo normal.

Aunque no se pregunta, observad que el grado de ansiedad de la persona 2 está casi dos desviaciones típicas debajo de la media (33 puntos, cuando la media es 50), es decir, esta persona es MUY POCO ansiosa.

Pregunta 7. Es seleccionar las personas de más de 40 años en Seleccionar Casos en menú de Datos. Y después es obtener el Diagrama de Caja con Maquiavelismo como variable dependiente y Ciudad (grande, pequeña) como Variable Independiente:



Como se observa en el Diagrama de Caja, para los mayores de 40 años, los niveles de maquiavelismo parecen ser muy similares independientemente que vivan en una ciudad grande o no, como se deduce del solapamiento de ambas cajas. Únicamente se aprecia unos muy ligeros valores de maquiavelismo en las personas que no viven en ciudad grande respecto a los de ciudad grande (medias: 209.7 vs. 208.3, respectivamente).

