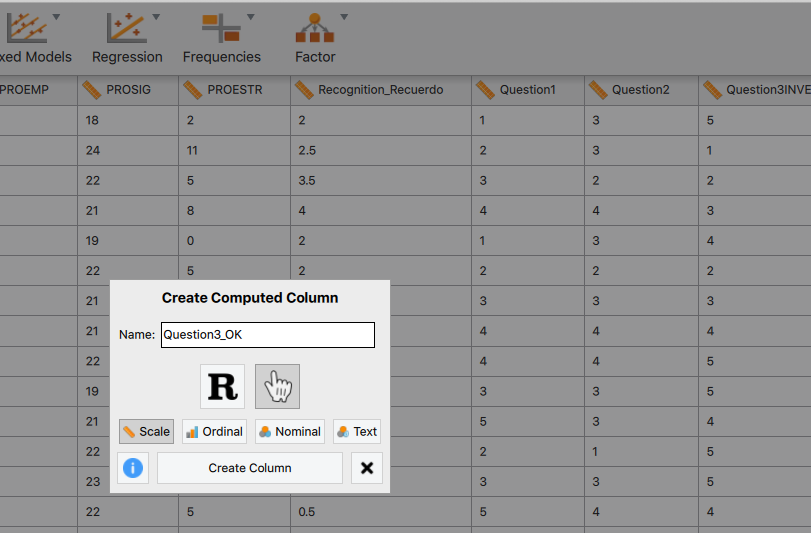
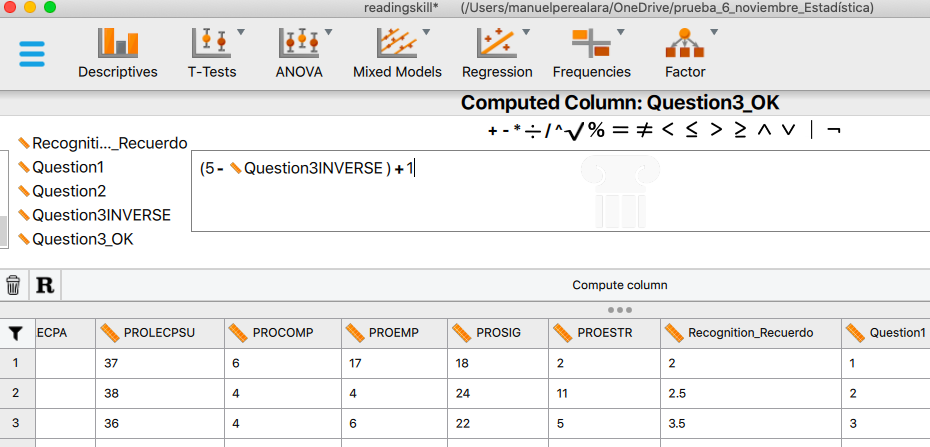
**EJERCICIO 4, 2023, Estadística**

Tenemos la base adjunta (<https://www.uv.es/mperea/readingskill.sav>) que contiene datos de habilidad lectora mediante el test PROLEC en estudiantes de secundaria, así como el número promedio de palabras aprendidas de manera incidental promediado a través de varias clases. Tenemos también las respuestas (en una escala Likert de 1-5) en un test de extroversión: las dos primeras, a más puntuación más extroversión (v.g., “mis amigos/as me consideran una persona sociable”; “prefiero ir de fiesta a leer un libro”) y la tercera es inversa (más puntuación menos extroversión (v.g., “me considero una persona callada”).

1. *¿Observas diferencias en extroversión entre chicos y chicas? (Has de sumar las 3 preguntas; la tercera pregunta debes codificarla en el orden correcto: a más puntuación, más extraversión) Recuerda poner las preguntas de escala Likert como cuantitativas, si no lo están.*

Primero, debemos pasar poner las tres preguntas tipo Likert como cuantitativas (“scale”) y poner la cuestión 3 en el orden correcto.

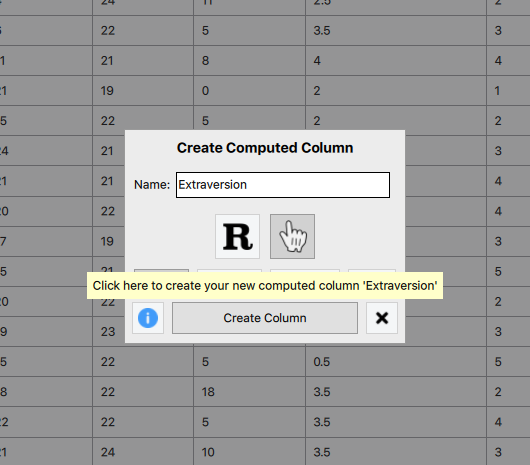


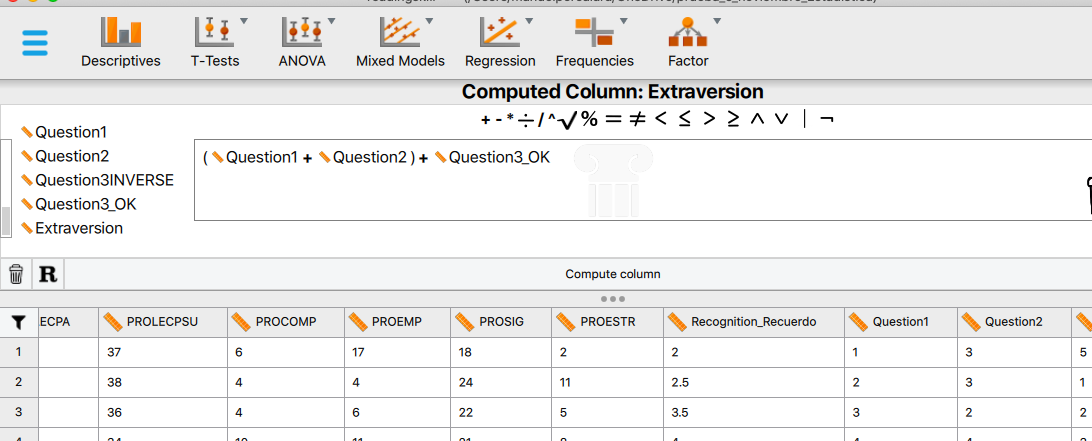


La fórmula es 5 - PREGUNTA\_ORDEN\_INVERSO + 1

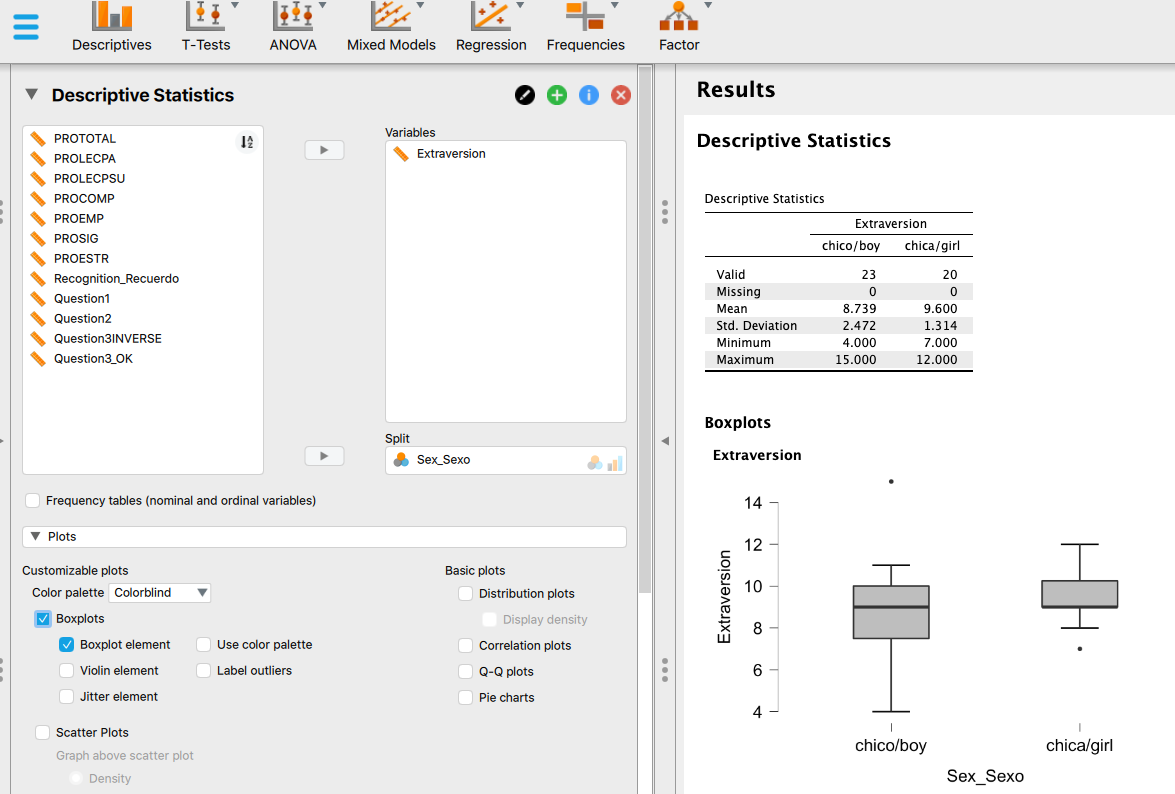
(el 5 es el valor máximo de la escala Likert; si fuera una escala Likert de 1-10 entonces el número sería 10 en lugar de 5, lógicamente).

Y ahora es sumarlas, en la nueva variable “extraversión”





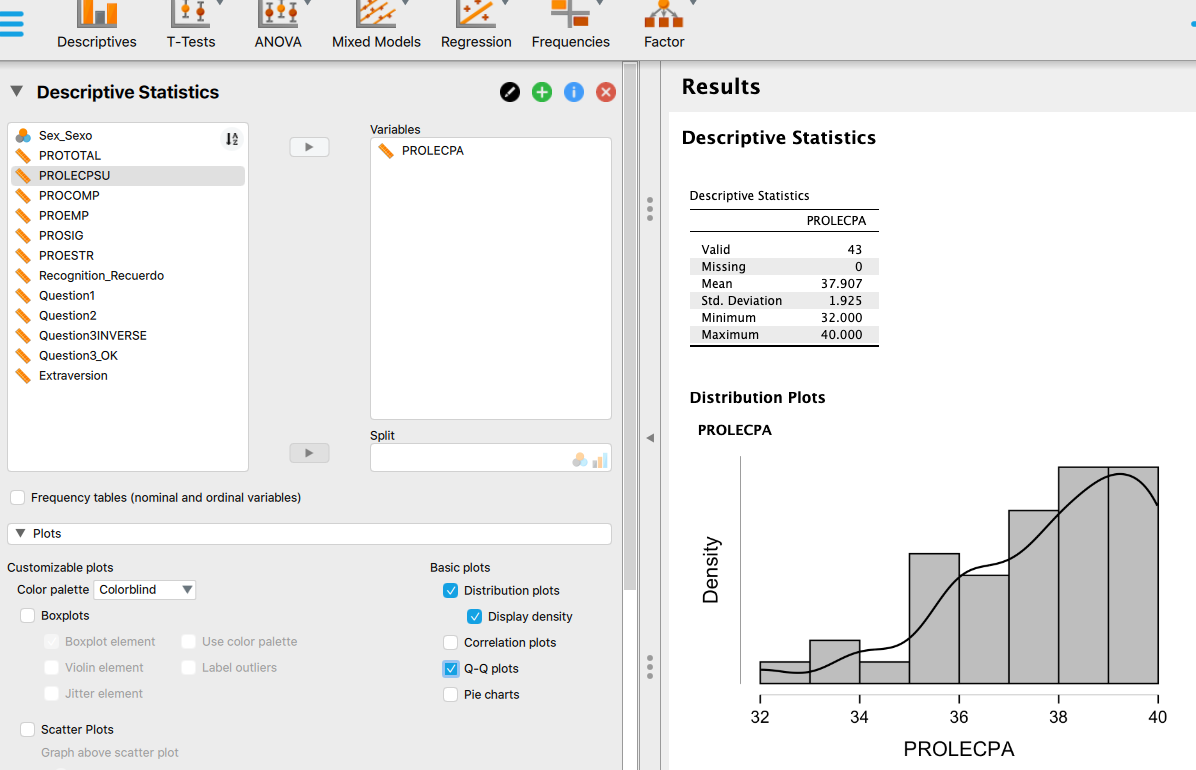
Ahora, habéis de mirar el diagrama de caja y las medias de chicos y chicas. Donde extraversión es la variable dependiente y género la variable independiente.



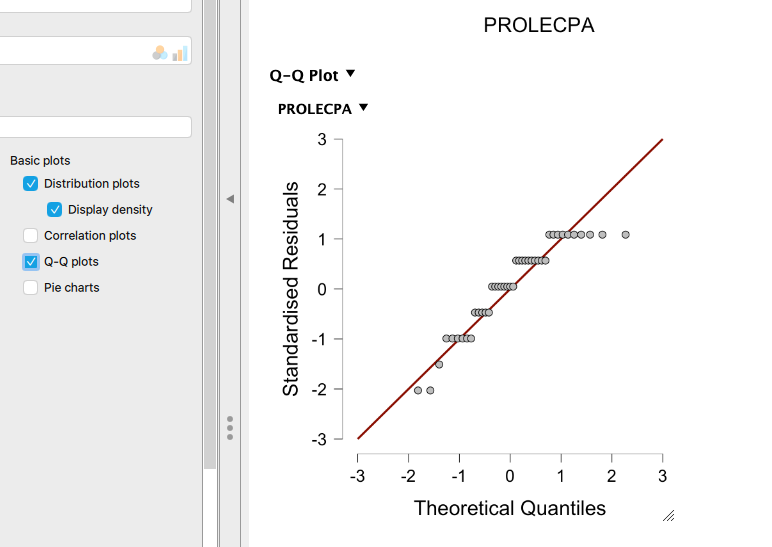
Se aprecia que, en la muestra, las chicas muestran niveles algo mayores de extroversión que los chicos (M = 9.6 vs. 8.7, respectivamente). Además, parece haber menor variabilidad en el grupo de chicas que en el de chicos (SD = 1.3 vs. 2.5, respectivamente).

*2. Globalmente, ¿el test PROLECPA (medida de lectura de palabras en el test PROLEC) sigue aproximadamente una distribución normal? ¿Qué asimetría tiene? ¿Y la curtosis?*

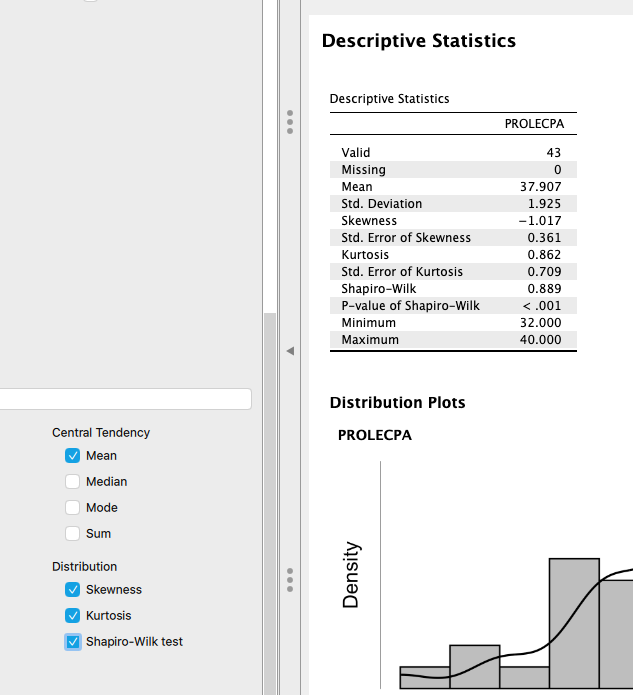
La mejor estrategia para empezar es mirar el histograma (suavizado):



Que se aprecia una clara asimetría negativa, así que no parece que la distribución sea normal. Además, en el Q-Q Plot, los puntos tampoco parecen acercarse a la recta.



Si queremos obtener índices (recordar que la distribución normal es simétrica y su forma es “mesocúrtica”, podemos calcular el índice de asimetría y el de curtosis. Y también aplicar la prueba de Shapiro-Wilks (si el valor de probabilidad p que se obtiene es menor de 0.05 entonces es que muy poco probable que realmente los datos siguen la distribución normal).

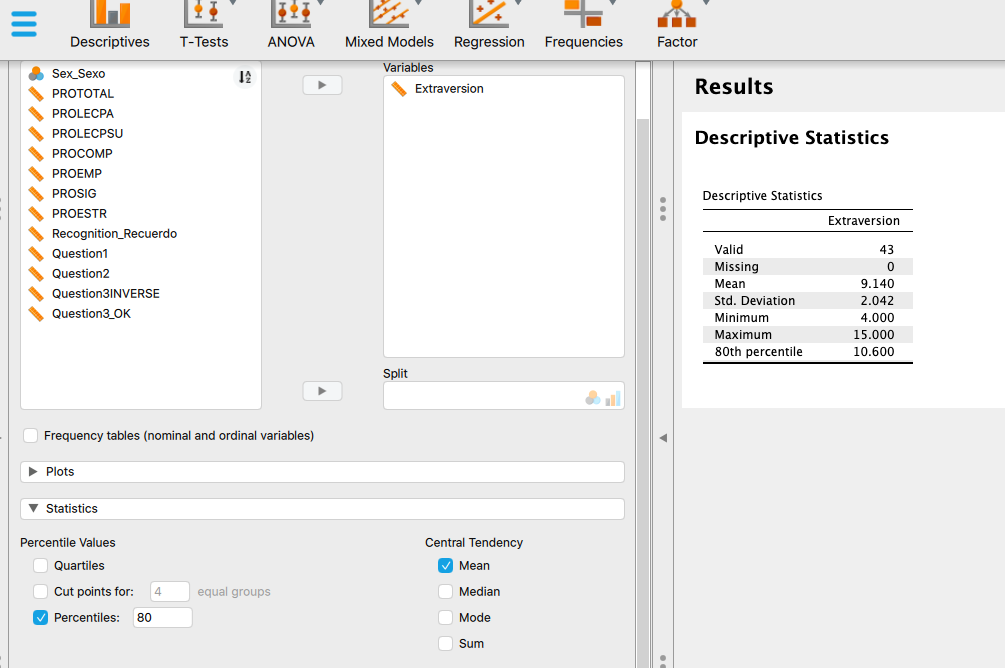


Se aprecia que el índice de asimetría es negativo, como se esperaba, -1.017, y que la distribución parece leptocúrtica (índice de curtosis = 0.86). Y la prueba de Shapiro-Wilk muestra un valor de p muy pequeño….

Por tanto, tanto los estadísticos como los gráficos claramente muestran que la distribución de PROLECPA no sigue la distribución normal.

*3. Si queremos seleccionar al 20% de personas más extravertidas para cursos de comunicación, ¿qué punto de corte elegiremos?*

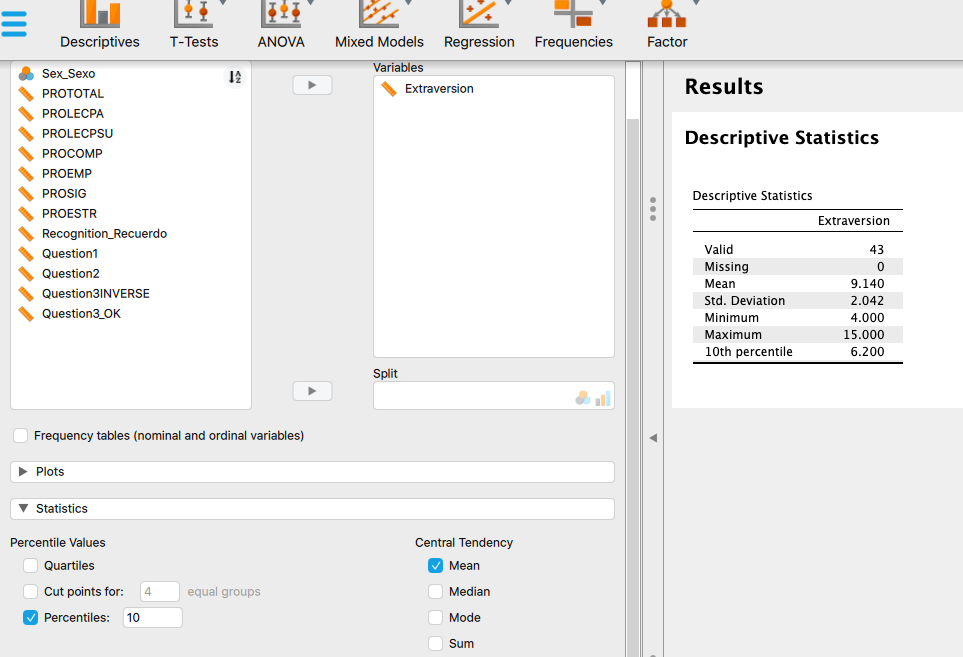
Pues el Percentil 80, que es la puntuación que deja el 20% de datos por arriba.



Ese valor es 10.6

*4. Si queremos seleccionar al 10% personas más introvertidas para un curso de habilidades sociales, ¿qué punto de corte elegiremos?*

El Percentil 10, que es la puntuación que deja el 10% de datos por abajo.



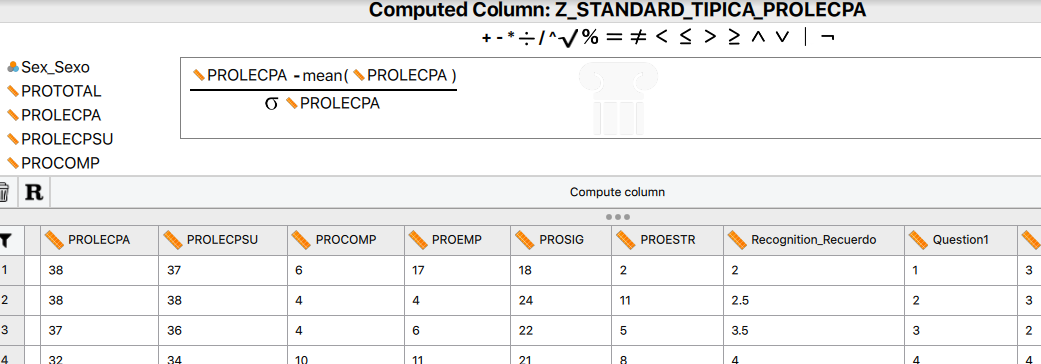
El resultado es 6.2

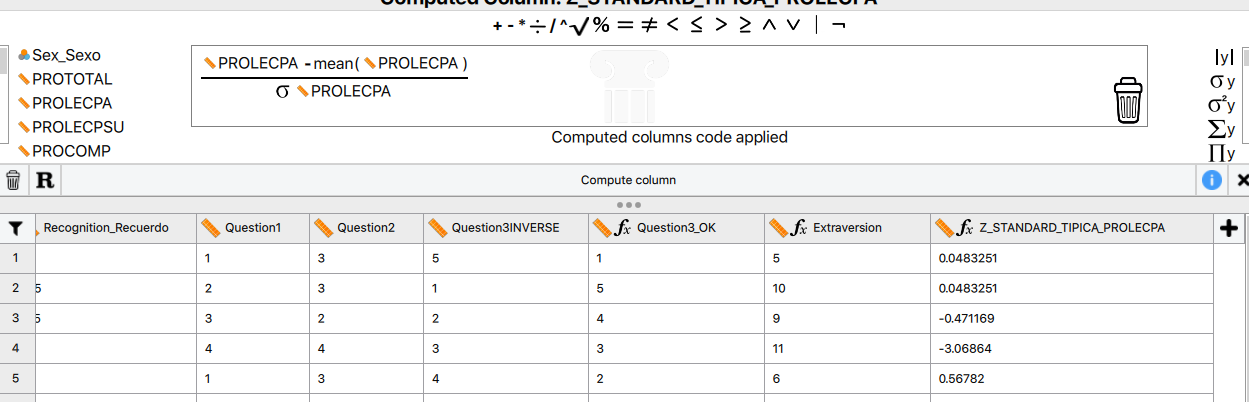
*5. ¿Qué puntuación típica corresponde en PROLECPA a la persona de la segunda fila en el archivo y qué significa?*

Es hacer

ccale(PROLECPA) como nueva columna, indicando “R” en lugar de la mano

O alternativamente (lleva más tiempo, con la mano en lugar de R) calcular (Xi-M)/s como nueva columna





El valor es .048, lo que quiere decir que la persona segunda está 0.048 desviaciones típicas por encima de la media. (O sea, que está solamente muy poco por encima de la media del grupo.)