**Ejercicio 3, 2021. Estadística. Respuestas**

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.

*1. ¿Observas diferencias en habilidad lectora (empleando la medida global del test: PROTOTAL) entre chicos y chicas?*

*2. Globalmente, ¿el test PROLEC (medida global PROTOTAL) sigue aproximadamente una distribución normal? ¿Qué asimetría tiene? ¿Y la curtosis?*

*3. Si queremos seleccionar al 15% de mejores lectores para un curso de escritura creativa, ¿qué punto de corte elegiremos?*

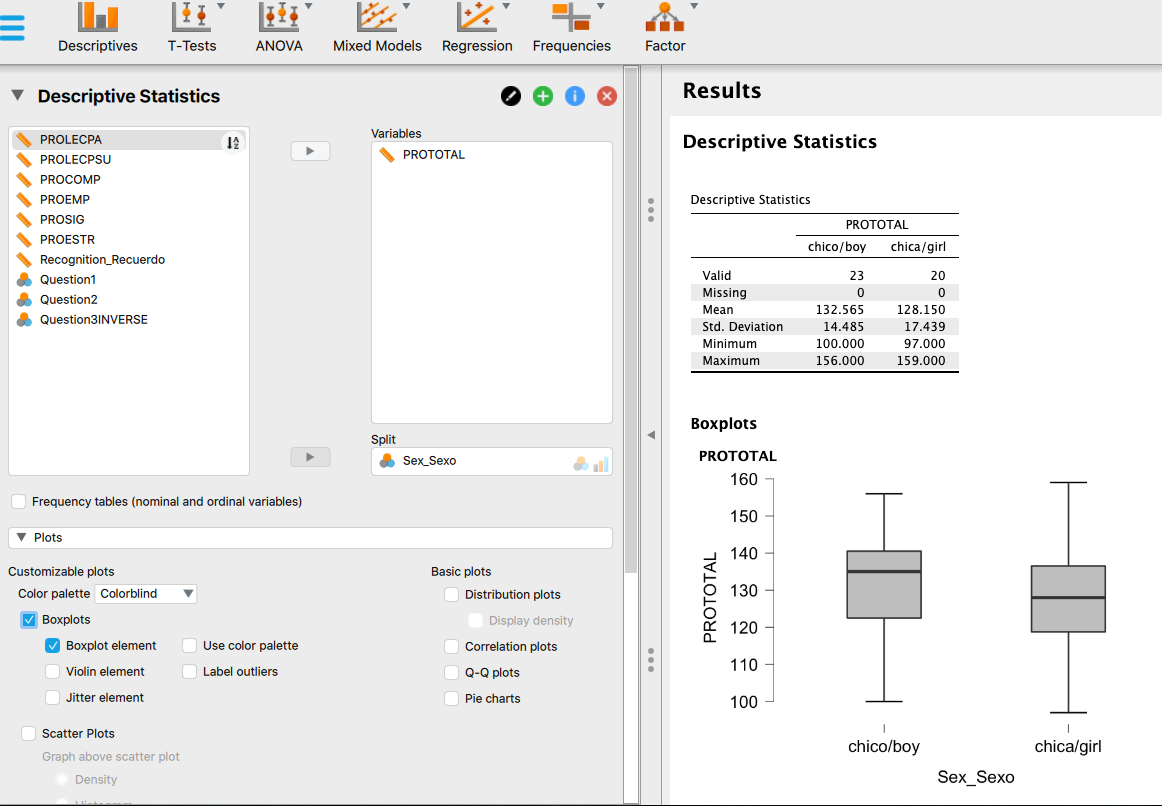
*4. Si queremos seleccionar al 10% de peores lectores para un programa de mejora de habilidad lectora, ¿qué punto de corte elegiremos?*

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

**Respuestas**

*1. ¿Observas diferencias en habilidad lectora (empleando la medida global del test: PROTOTAL) entre chicos y chicas?*

La variable independiente sería sexo y la variable dependiente la habilidad lectora en el PROLEC (PROTOTAL).



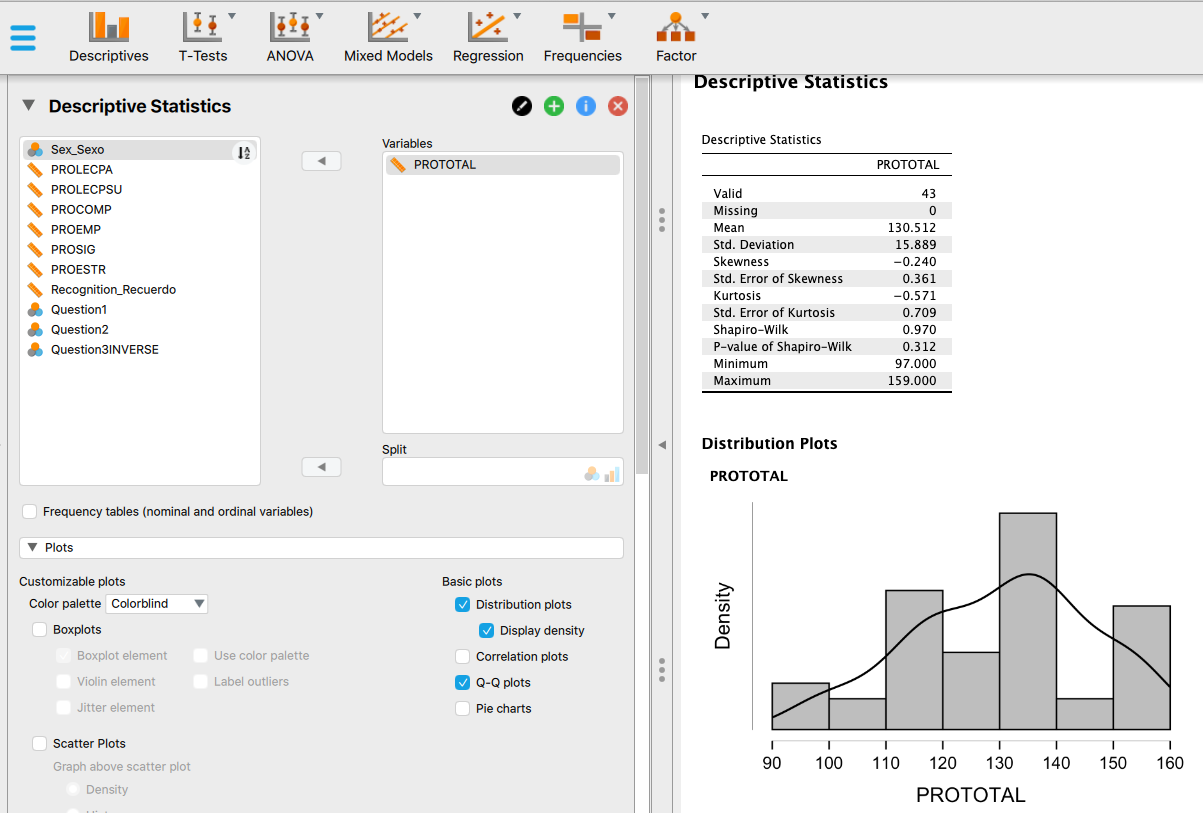
Los chicos tuvieron, en promedio, puntuaciones en habilidad lectora algo mayores que las chicas en esta muestra (M = 132.6 vs. 128.2, respectivamente).

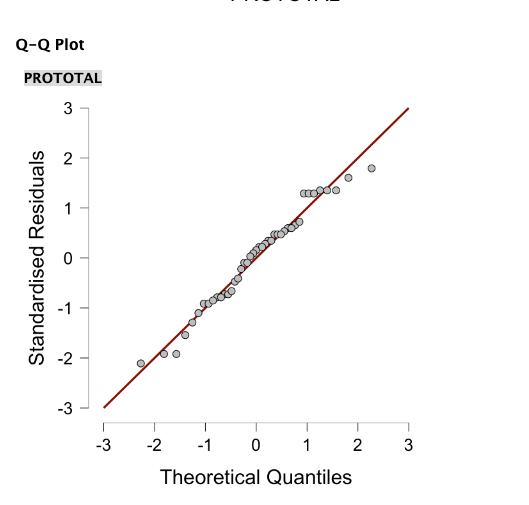
(Hemos empleado la media; con la mediana las diferencias también estaban, como se aprecia en el diagrama de cajas)

*2. Globalmente, ¿el test PROLEC (medida global PROTOTAL) sigue aproximadamente una distribución normal? ¿Qué asimetría tiene? ¿Y la curtosis?*

Lo más práctico es ver el histograma, en particular la línea suavizada, en la que se aprecia una ligera asimetría negativa, que de hecho se confirma en el índice que es -.240. La curtosis es -.571 que sugiere que la distribución es algo platicúrtica.

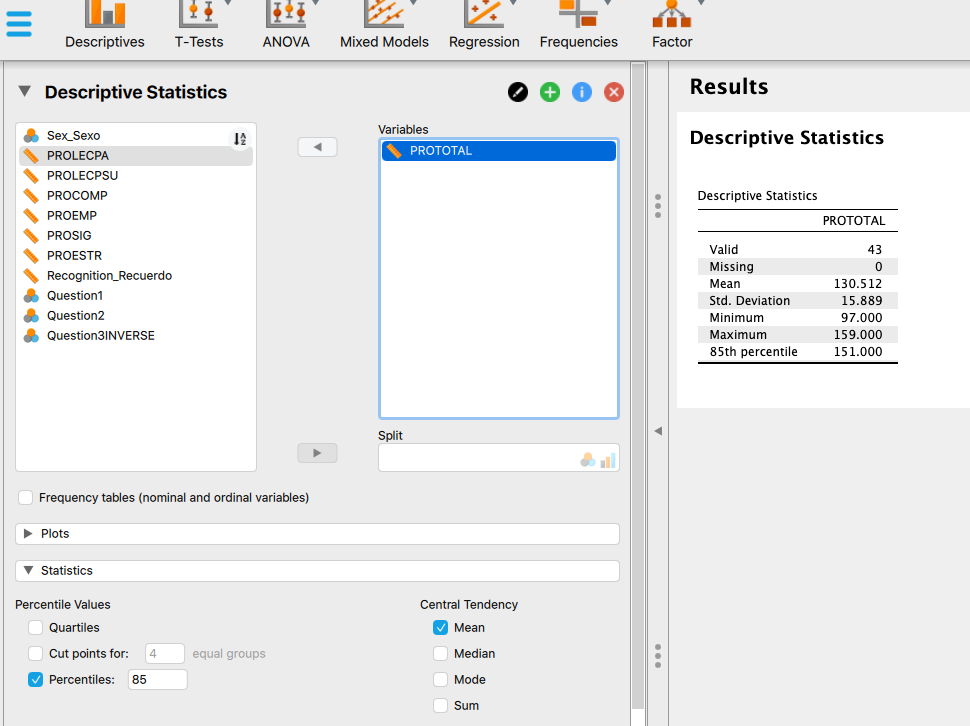
No obstante, el Q-Q plot muestra que los puntos no están muy lejos de la recta, lo que sugiere que la distribución del PROLEC en realidad no se aleja demasiado de la distribución normal. De hecho, si se hiciera una prueba de Shapiro-Wilks de normalidad, el valor de probabilidad p es mayor de 0.05 lo que sugiere que la distribución de datos en PROTOTAL no se aparta significativamente de la distribución normal.





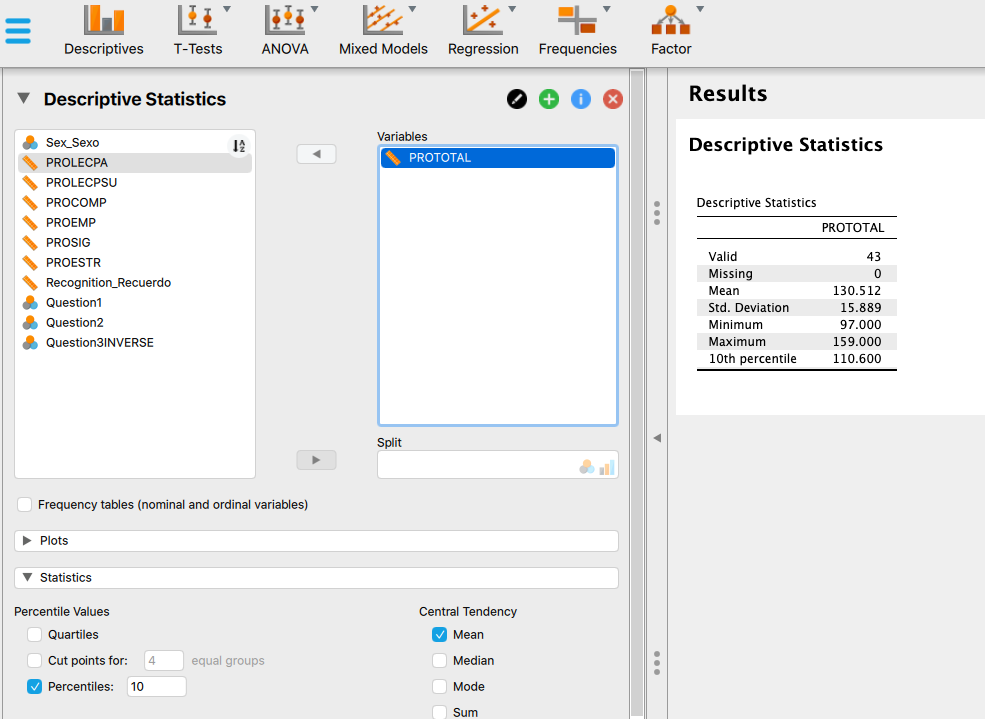
*3. Si queremos seleccionar al 15% de mejores lectores para un curso de escritura creativa, ¿qué punto de corte elegiremos?*

Es elegir el Percentil 85, dado que debajo de ese valor estará el 85% de las personas y por consiguiente, encima de ese valor estará el 15% de las personas. En nuestro caso es 151.



*4. Si queremos seleccionar al 10% de peores lectores para un programa de mejora de habilidad lectora, ¿qué punto de corte elegiremos?*

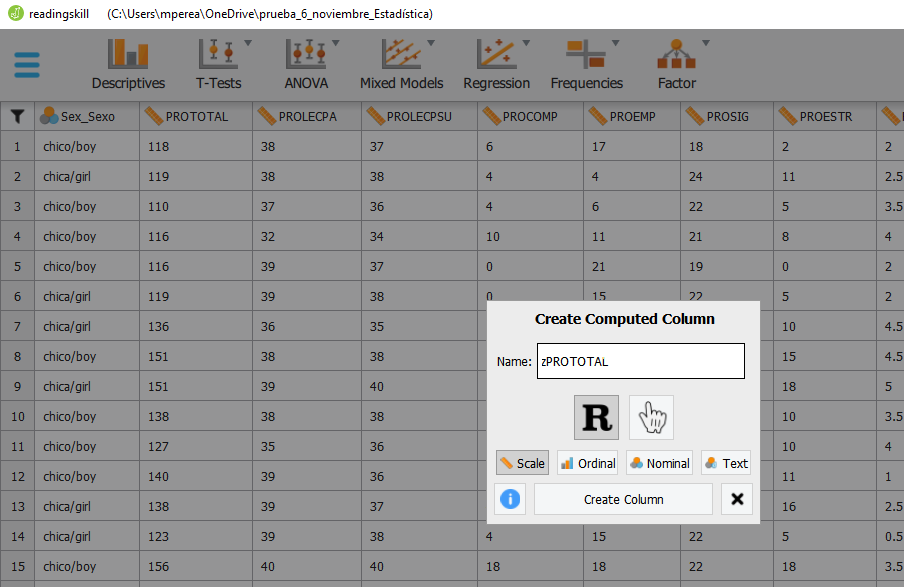
Es elegir el Percentil 10, dado que debajo de ese valor estará el 10% de las personas En nuestro caso es 110.6.



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

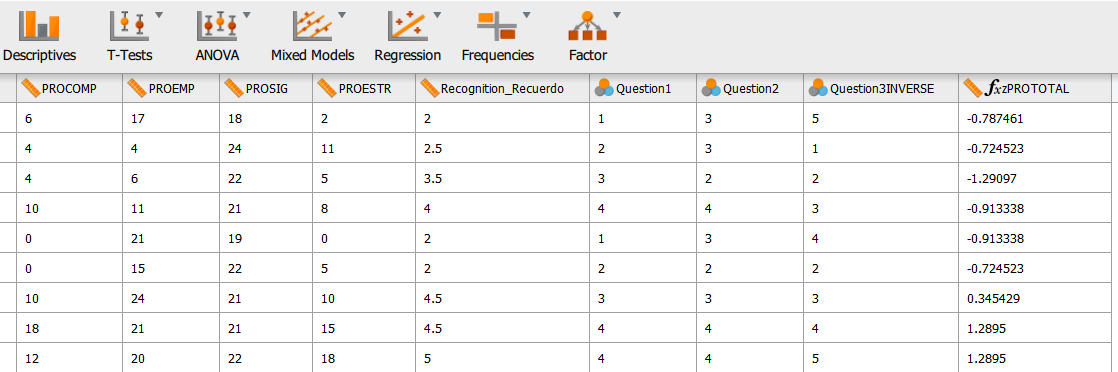
Para conseguir puntuaciones típicas, es seguir la fórmula

(Xi-M)/s



Y la fórmula para pasar a puntuaciones típicas es, dentro de R para crear una nueva columna:

scale(VARIABLE\_A\_TIPIFICAR)



El valor de z para la primera persona es -0.78. Esto significa que esta persona está 0.78 desviaciones típicas debajo de la media en habilidad lectora. Es una persona con habilidad lectora menor que la media, sin ser excesivamente baja.