**30 de septiembre. Estadística-I**

Un grupo de investigadores quiere conocer la actitud hacia una marca después de ver un anuncio. En todos los anuncios aparecía un/a modelo, pero había cuatro versiones del anuncio que se diferenciaban solo en el color de ojos del (de la) modelo (color azul, color marrón, color verde, indefinido [el color de ojos no se podía ver]). Los participantes fueron asignados al azar a una de las cuatro versiones del anuncio. Después de ver el anuncio, los participantes tenían que responder a 10 preguntas de la encuesta sobre actitudes hacia la marca (en una escala Likert de 7 puntos). Los promedios más altos correspondían a actitudes más positivas; la columna que tenemos es el promedio de estas 10 preguntas por participante.

*¿Qué es la variable independiente? ¿Y cuántos niveles?*

El color de ojos del (de la) modelo, que tiene 4 niveles (azul, marrón, verde, indefinido)

*¿Cuál es la variable dependiente? ¿Es cuantitativa?*

Es la actitud hacia la marca. Sí, se la trata como cuantitativa (es un promedio de 10 preguntas con una escala tipo Likert).

*¿Tenemos un experimento o un cuasi-experimento? ¿Por qué?*

Es un experimento. La asignación de los participantes a los grupos fue al azar.

*¿Los investigadores han utilizado un diseño entre-sujetos?*

Sí, cada participante fue asignado a un ÚNICO grupo.

*¿Cuántas filas necesitamos para analizar los datos? ¿Y columnas?*

Filas: Tantas como participantes.

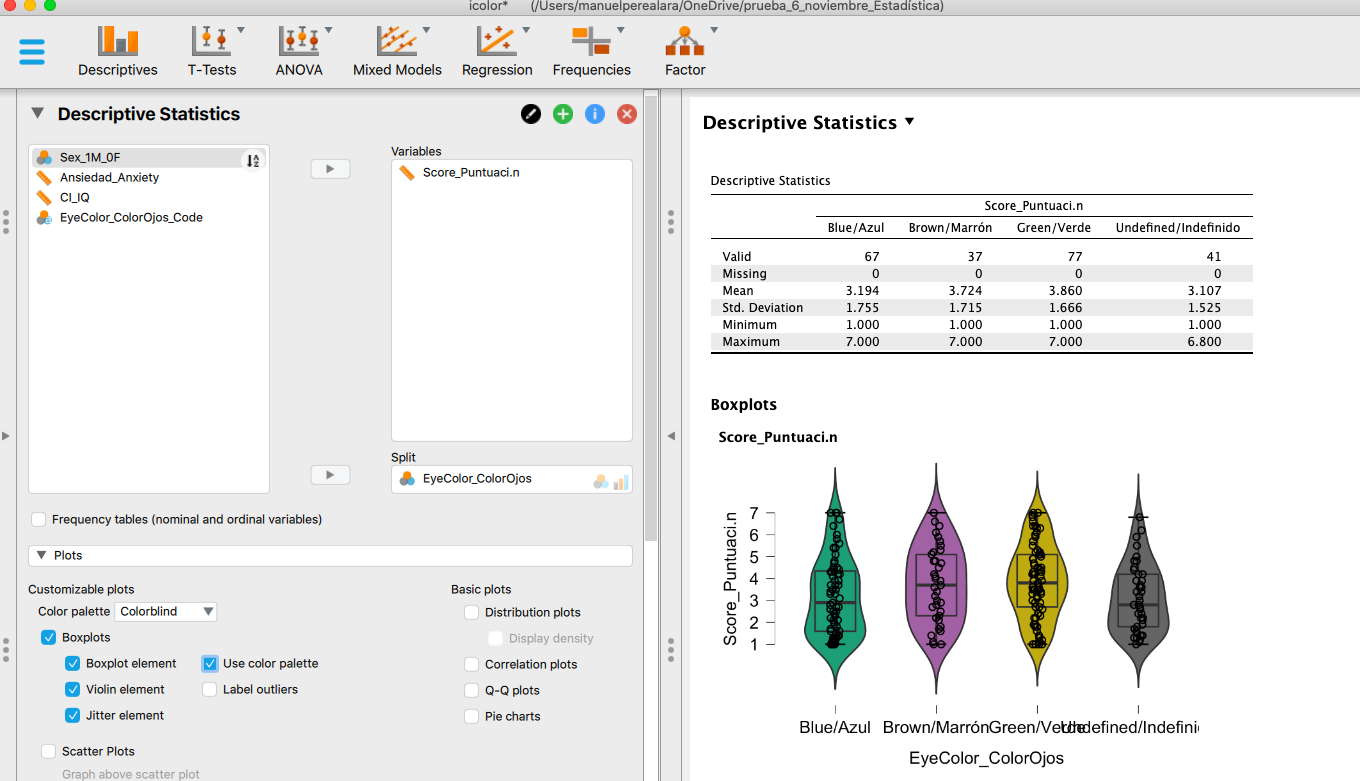
Columnas: dos, una para indicar el grupo (1,2,3,4) and otra para indicar la variable dependiente.

(En realidad, se podría indicar que había además las 10 columnas con las respuestas a cada una de las preguntas, para así tener en el mismo fichero todos los datos, los originales, y la columna promedio que es en realidad con la que vamos a operar.)

Los datos están disponibles en <https://www.uv.es/mperea/icolor.csv>. Guarde el archivo y ábrelo con JASP. (También añadimos algunas otras variables ...)

*Elige el gráfico que consideres más adecuado y luego interpreta los resultados; también puedes observar (y emplear) las medias por condición. ¿Se te ocurre en un nuevo experimento en este tema?*

Lo más práctico es un diagrama de caja, y además es lo más compacto: en un solo gráfico tenemos las distribuciones de las cuatro condiciones. Podríamos haber también hecho cuatro histogramas, pero es más complicado de ver todo como un conjunto. Y, lógicamente, hemos de prestar atención también a los índices de tendencia central como la media, que JASP proporciona



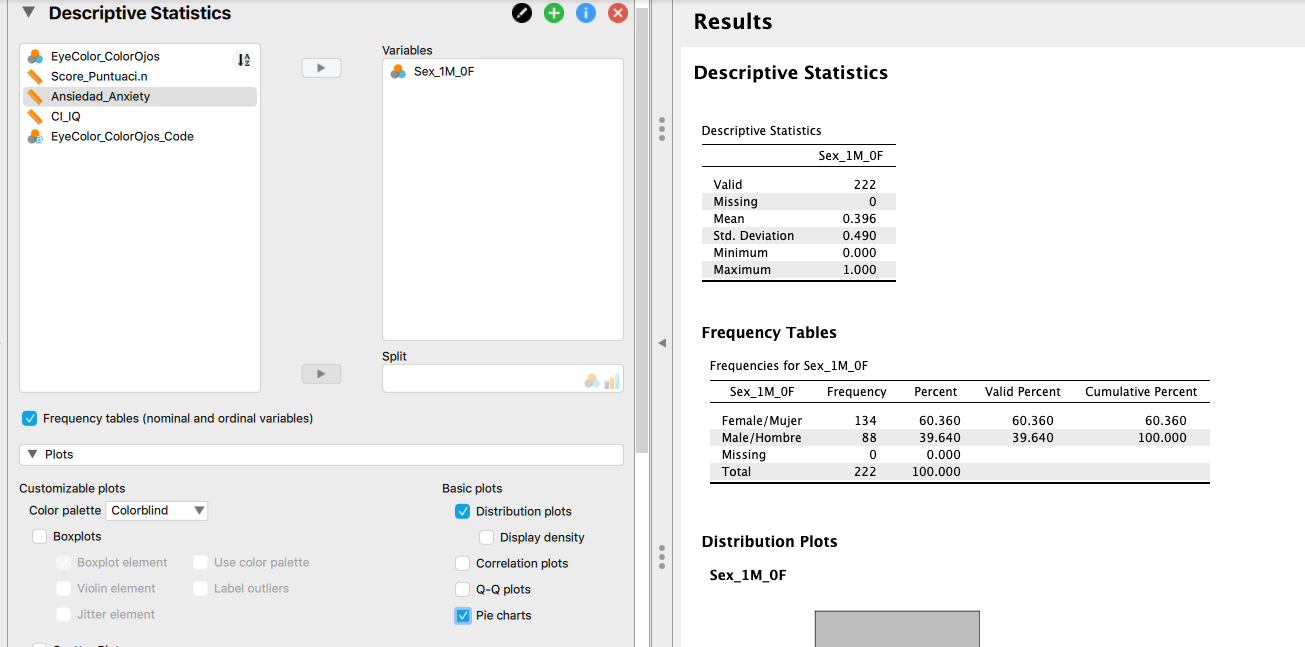
Se aprecia que los participantes asignan puntuaciones más altas al anuncio cuando el/la modelo tiene ojos marrones (M = 3.72) o verdes (M = 3.86) que cuando tiene ojos azules (M = 3.19) o no se aprecian el color de los ojos (M = 3.11).

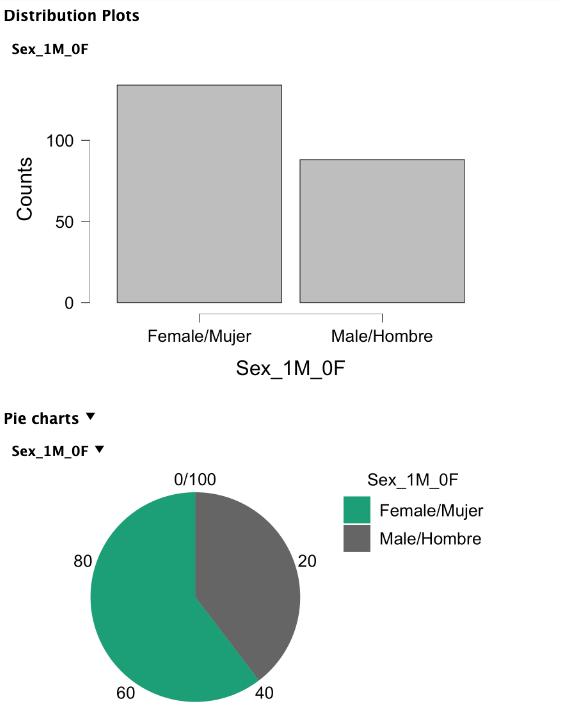
¿Por qué? A la hora de interpretar los datos hemos de tener en cuenta un elemento clave. El experimento se realizó con personas de Texas, en los que más del 90% de participantes eran de origen hispano. Es posible que las personas de origen hispano pudieran considerar al (a la) modelo de ojos azules como “de otro grupo” y por tanto que los paricipantes no eran el objeto del anuncio (es decir, podrían pensar que se dirige más al público anglosajón). Lógicamente, para probar esta hipótesis, habría que realizar el mismo experimento en otro estado de los EE.UU. donde haya una muestra mucho mayor de personas cuyo origen fuera el norte de Europa (Michigan o Wisconsin, seguramente).

Por otra parte, se podría haber hecho el experimento con un diseño “intra-sujetos”. En este caso, cada persona habría de evaluar CUATRO anuncios (uno de cada color de ojos).

*¿Se equilibró el número de hombres y mujeres en el conjunto de participantes? ¿Cuál es la escala de medida? ¿Qué gráfico prefieres? ¿Optarías simplemente por la tabla de frecuencias?*

Podemos realizar bien un diagrama de barras (es la opción de “Distribution Plot”) o un diagrama de sectores. En realidad, con la tabla de frecuencias ya bastaría. Las tres opciones dan la misma información.

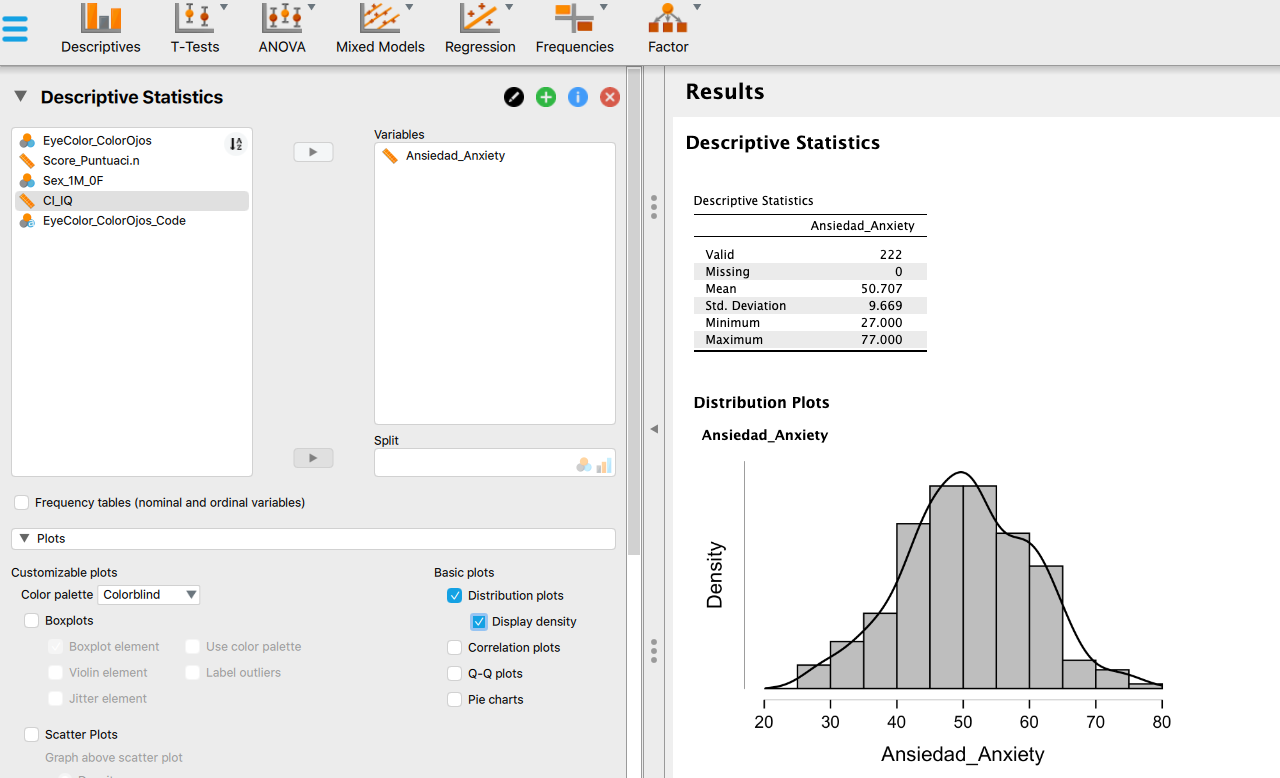




Se aprecia que la muestra no está del todo balanceada: 60.4% de mujeres vs. 39.6% de hombres.

*También obtuvimos una medida de ansiedad de los participantes. ¿Crees que la "ansiedad" siguió una distribución normal?*

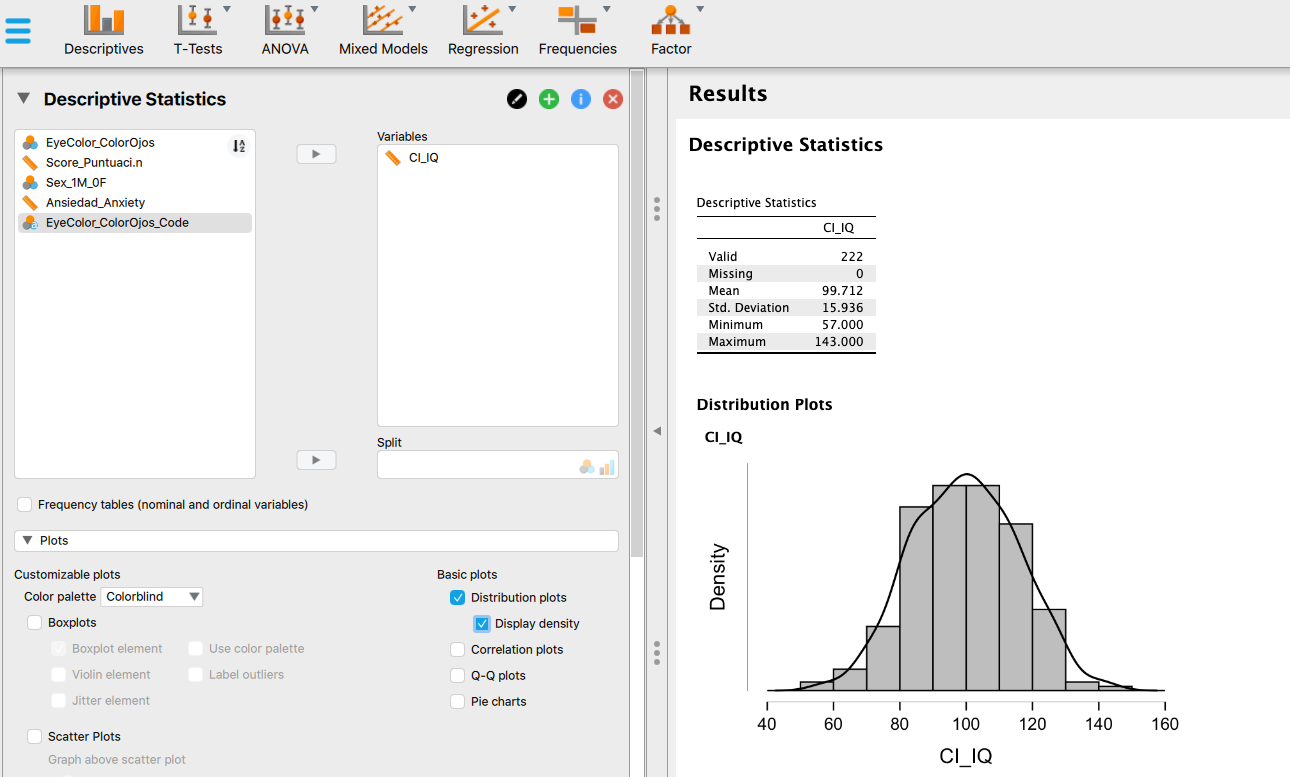
Aquí es hacer un histograma y suavizarlo. ¿Cómo? Pues de nuevo “Distribution Plot” (al ser una variable cuantitativa, JASP hace un histograma) y pulsar la opción de “display density” para tener la curva suavizada.



Sí, parece una distribución aproximadamente normal—en temas posteriores comentaremos otras maneras de ver de manera más exacta la “normalidad” de la curva.

*Idem para el coeficiente intelectual.*

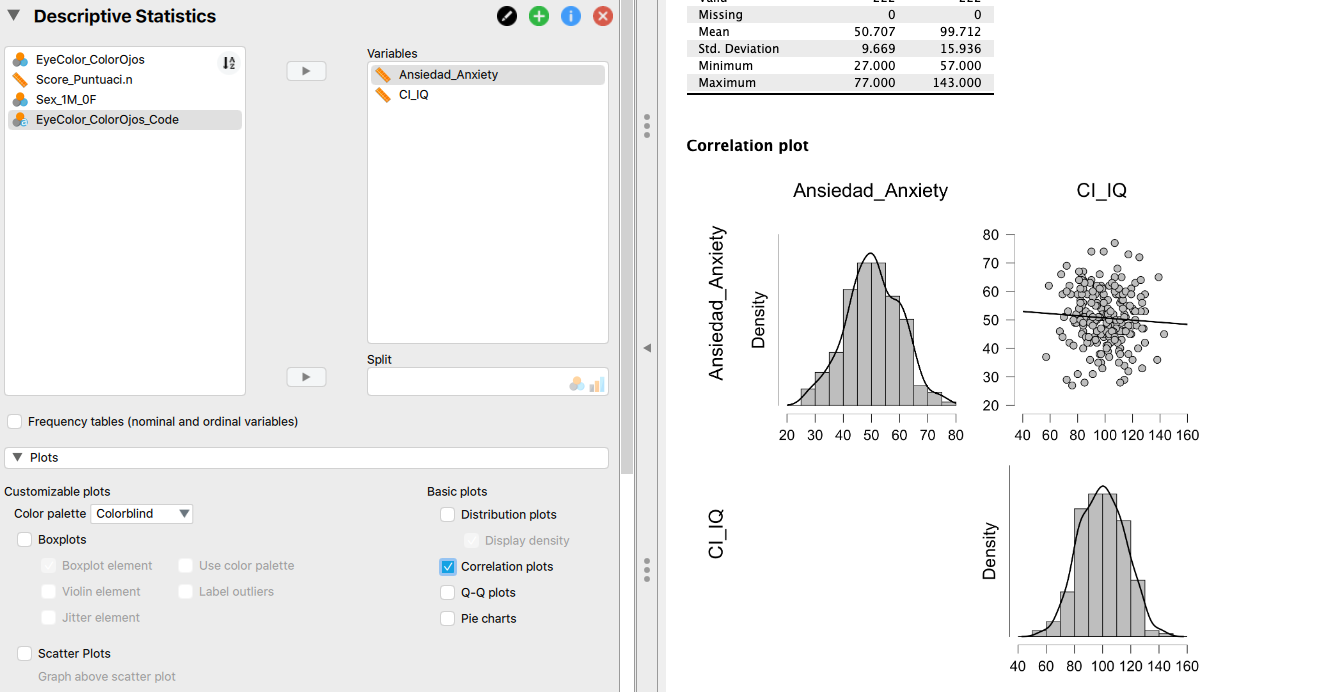
Empleamos los mismos pasos que antes. Y sí, parece que la inteligencia sigue aproximadamente una distribución normal.



*¿Están relacionadas la ansiedad y el coeficiente intelectual? Elige el gráfico adecuado.*

Queremos ver si ambas variables están relacionadas. En este caso, es el “correlation plot”. Observar que el diagrama de dispersión de acompaña de otros dos gráficos, que son simplemente los histogramas de cada una de las dos variables.

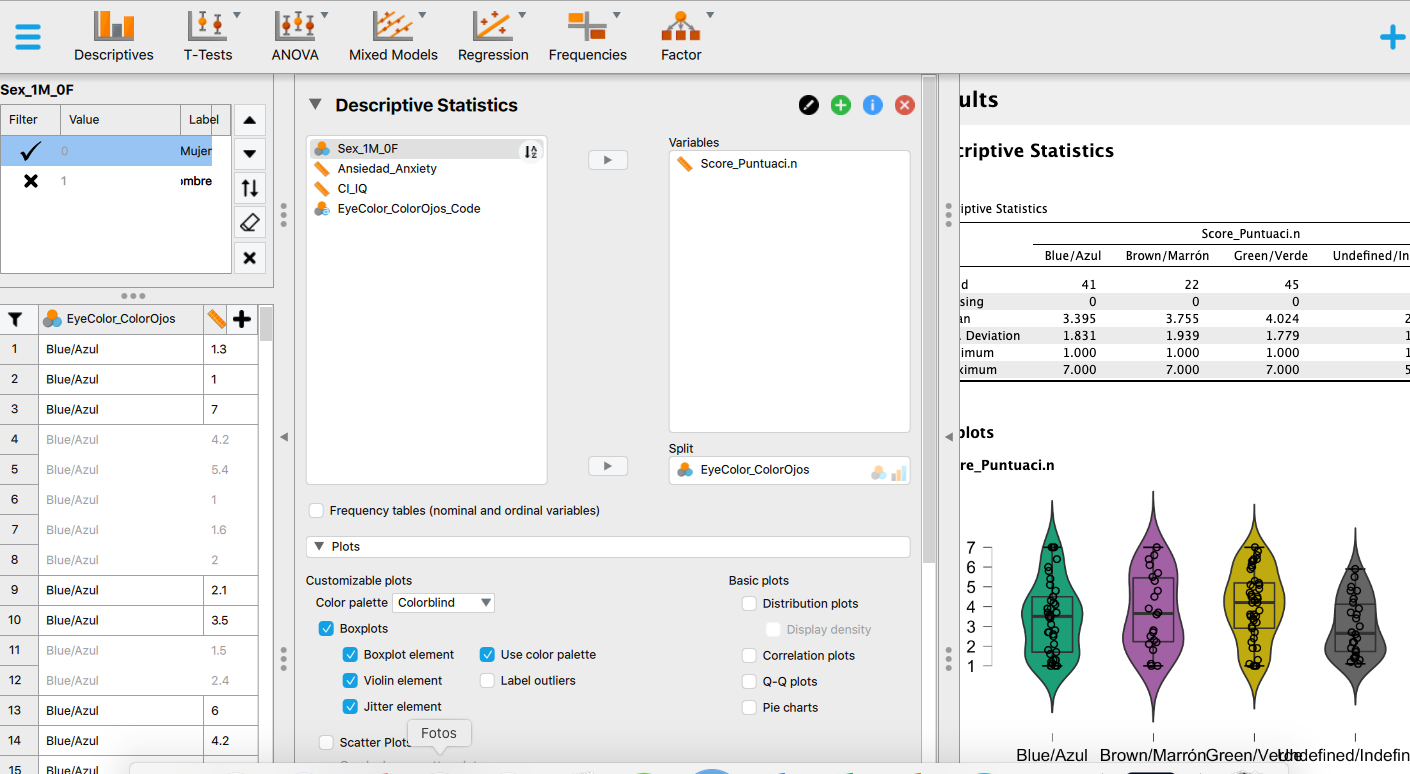
Claramente, no se aprecia relación alguna entre ambas Ansiedad e Inteligencia. Hay una nueva de puntos.



*Selecciona solo a las mujeres de la muestra y re-haz los análisis (el del anuncio) nuevamente para ver si el patrón es similar al análisis de toda la muestra.*

Es filtrar en la columna de Sexo, y elegir solamente las mujeres en el análisis del anuncio.

Se aprecia más o menos el mismo patrón de datos que en la muestra completa.



Referencias (no en formato APA…)

Moore, D. S., McCabe, G. P., and Craig, B. A. (2012). Introduction to the Practice of Statistics (7th ed.). New York: Freeman.

Simpson, P. M., Sturges, D. L., and Tanguma, J. (2008). The eyes have it, or do they? The effects of model eye color and eye gaze on consumer as response. The Journal of Applied Business and Economics, 8: 60-72.