



ESTADÍSTICA APLICADA A CIENCIAS DE LA SALUD: ANÁLISIS ESTADÍSTICO DESCRIPTIVO

Juan Antonio Valera Calero PT, PhD

Ibai López de Uralde Villanueva PT, PhD

Gustavo Plaza Manzano PT, PhD

Departamento de Radiología, Rehabilitación y Fisioterapia

Universidad Complutense de Madrid

GENERALIDADES DEL ANÁLISIS ESTADÍSTICO DESCRIPTIVO

El análisis estadístico descriptivo tiene como objetivo proporcionar la tendencia central y la dispersión de una característica numérica (por ejemplo, la altura de una población) o la frecuencia de una característica cualitativa (por ejemplo, el número de personas o el porcentaje de la población con los ojos verdes). En función del tipo de variable, se realizará un proceso u otro para informar de los datos.

En cualquier caso, los análisis descriptivos consistirán en 4 procedimientos estructurados:

- Preparar la base de datos.
- Análisis de la normalidad.
- Análisis descriptivo.
- Construcción de gráficos y tablas.

PREPARAR LA BASE DE DATOS

En primer lugar, al abrir el programa estadístico, pincharemos en la pestaña “Vista de variables” e introduciremos el nombre de las variables que queramos estudiar. Es importante recordar que el nombre de las variables no permite el uso de la barra espaciadora.

En el momento que se han introducido el nombre de las variables, se habilitarán varias opciones incluyendo el tipo de variable (por lo general, siempre numérico), la anchura (número de dígitos visibles al introducir los datos de cada participante) o el número de decimales que permitiremos en cada variable.

Las etiquetas de valor pueden asignarse en la casilla de valores. Las **variables categóricas** o **cuantitativas** son aquellas que analizan cualidades o atributos. Para poder analizarlas en el programa estadístico SPSS, se asociará un número a cada una de las categorías de esa variable. Por ejemplo, en la variable categórica “color de ojos”, asociaremos en las etiquetas de valor en la vista de variables el número 1 a la característica “Azul”, el 2 a la característica “Verdes” y el 3 a la característica “Marrón”. En el caso de las variables **cuantitativas** no será necesario la inclusión de etiquetas de valor, dado que estas variables adoptan valores numéricos. Pueden ser divididas en variables de tipo **discreto** (adoptan valores numéricos en unidades que no pueden ser fraccionadas, como por ejemplo el número de dientes) o **continuas** (adoptan valores numéricos en unidades que sí pueden ser fraccionadas en decimales, por ejemplo, la altura o el peso).

Como medida opcional para facilitar la visualización del tipo de variable, las casillas de “medida” te permite diferenciar con símbolos aquellas variables que son categóricas (Nominal) de aquellas que son cuantitativas (Ordinal o Escala).

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
1	Edad	Númérico	8	2		Ninguno	Ninguno	8	Derecha	Escala	Entrada
2	Altura	Númérico	8	2		Ninguno	Ninguno	8	Derecha	Escala	Entrada
3	ColorOjos	Númérico	8	2		Ninguno	Ninguno	8	Derecha	Nominal	Entrada
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											

Etiquetas de valor

Etiquetas de valor

Valor: Ortografía...

Etiqueta:

1,00 = "Azul"

2,00 = "Verde"

3,00 = "Marrón"

Añadir Cambiar Eliminar

Ayuda Cancelar Aceptar

Vista de datos **Vista de variables**

Fuente: Elaboración propia

Una vez que la base de datos está estructurada, podemos proceder a incluir los datos de cada participante en la pestaña “Vista de datos”. Para ello, podemos 1) introducir los valores manualmente a medida que realizamos la investigación o podemos 2) copiar y pegar los datos que hayamos introducido en otro formato (por ejemplo, Excel).

Análisis estadístico

Edad	Altura	Color de ojos
25	1,50	Azul
23	2,10	Verde
27	1,79	Azul
28	1,56	Verde
26	1,60	Marrón
24	1,98	Marrón
19	1,88	Marrón
30	1,85	Azul
32	1,90	Azul
28	1,92	Azul
24	1,56	Marrón
25	1,98	Marrón
27	1,96	Marrón
27	1,97	Marrón
21	1,98	Marrón
28	2,00	Verde
27	1,55	Verde
25	1,85	Marrón
25	1,99	Marrón
26	1,98	Marrón
27	1,61	Marrón

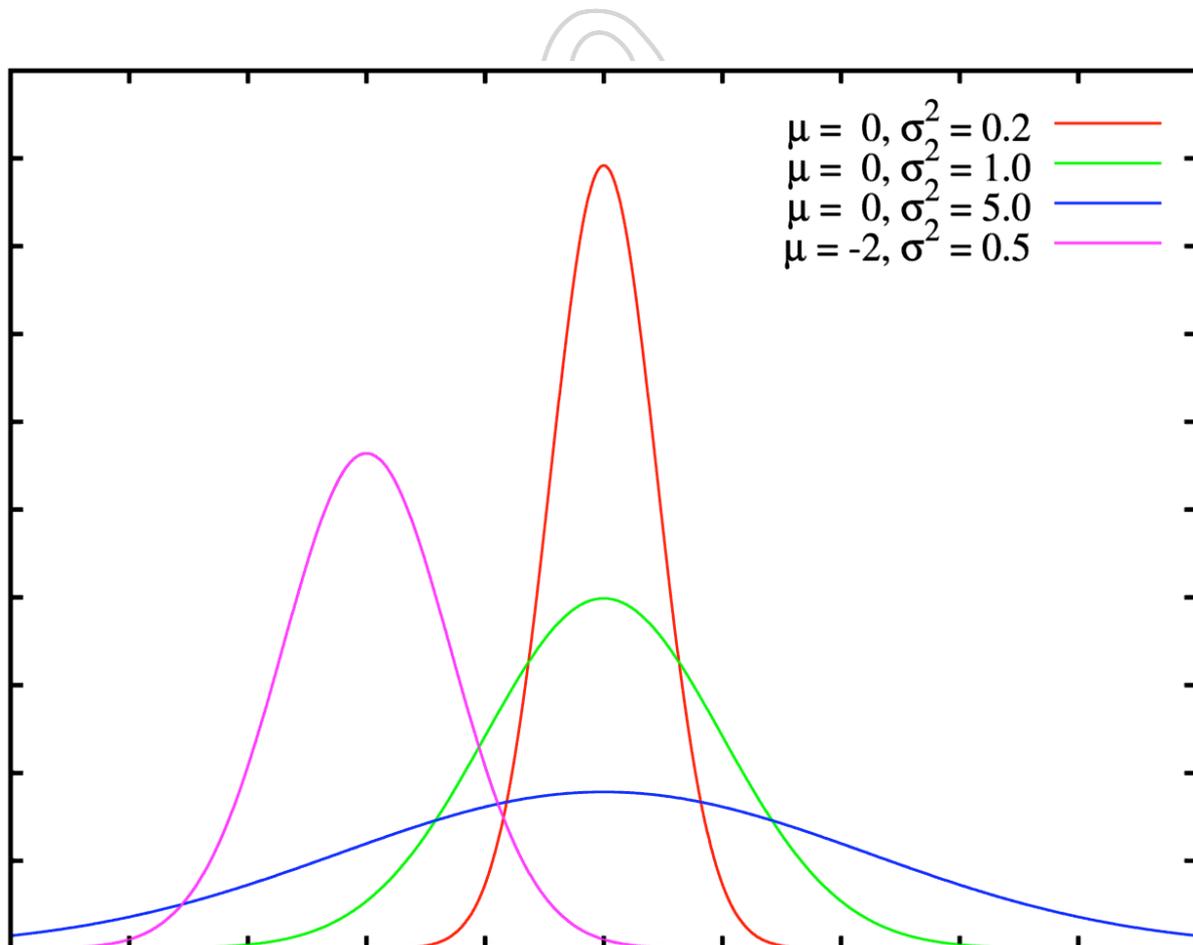
Fuente: Elaboración propia

UNIVERSIDAD COMPLUTENSE MADRID

ANÁLISIS DE LA NORMALIDAD

En estadística, la normalidad hace referencia a una distribución de probabilidad determinada (distribución normal, se visualiza en la campana de Gauss de la figura). Siempre que tengamos variables cuantitativas, es importante verificar su normalidad para que los análisis descriptivos sean los adecuados. Si la distribución es **normal** se utilizarán la **media** como medida de tendencia central y la **desviación standard** como medida de dispersión, mientras que si la distribución es **no normal** se utilizarán la **mediana** como medida de tendencia central y el **rango intercuartil** (cuartiles 1 y 3) como medida de dispersión.

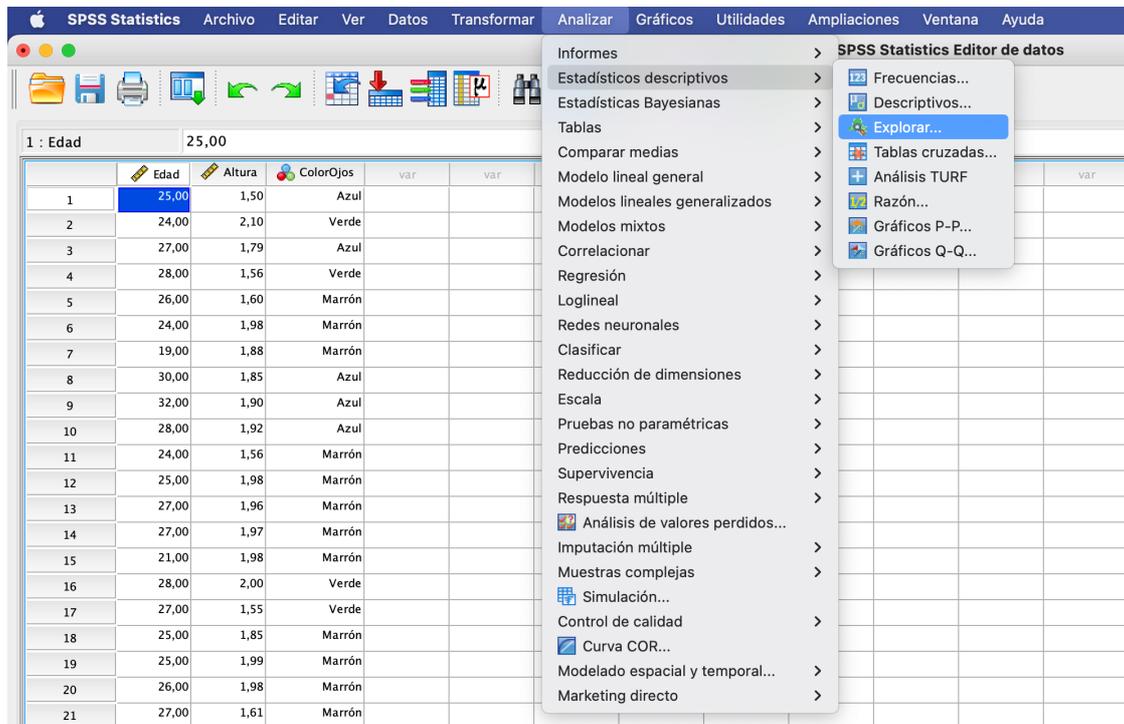
Como ejemplo, desde el punto de vista estadístico, lo “normal” en las puntuaciones de los exámenes de los estudiantes es encontrar un mayor número de estudiantes que obtienen notas intermedias entre las mínimas y máximas obtenidas. En este caso, la línea que delimita ambas mitades de la curva es definida como la media, siendo la anchura de ambos lados de la gráfica la desviación standard. Las 4 curvas dibujadas abajo siguen una distribución normal dado que cumplen con las características de una campana de Gauss. Las diferencias entre ellas son la curtosis (por ejemplo, la clase roja ha obtenido una media de 5/10, estando la mayoría de las notas comprendidas entre el 4 y el 6 mientras que la clase azul, a pesar de mantener una media de 5/10 al igual que la anterior curva, la mayoría de notas se comprenden entre el 1 y el 9).



Fuente: https://es.wikipedia.org/wiki/Funci3n_gaussiana

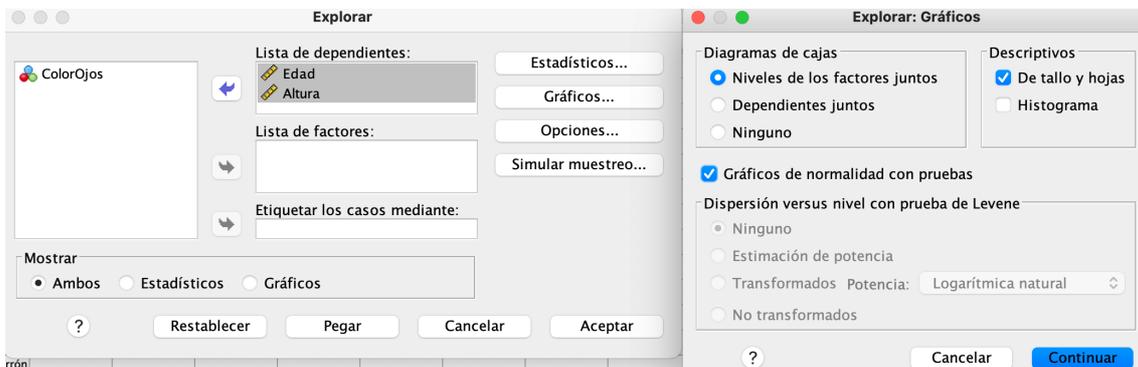
Para verificar la normalidad, podemos utilizar el programa estadístico para realizar las pruebas correspondientes. Si la significancia asintótica bilateral obtenida en las pruebas es <0.05 se considerará “no normal” y si el valor es >0.05 se considerará “normal”.

Los dos test estadísticos utilizados con este propósito serán el test de **Kolmogorov-Smirnoff con corrección de Lilliefors** para muestras de hasta 50 sujetos y el test de **Shapiro-Wilk** para muestras mayores de 50 sujetos. El acceso a los resultados de ambos test es idéntico: Analizar – Descriptivos – Explorar.



Fuente: Elaboración propia

Tras introducir las variables cuantitativas y marcar (dentro de la opción de “Gráficos”) las pruebas de normalidad, aparecerá una ventana nueva de resultados con el contenido de ambos tests. En este caso, al ser menos de 50 individuos tomaremos como referencia los valores obtenidos en la prueba de Kolmogorov-Smirnov.



Fuente: Elaboración propia

Pruebas de normalidad

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Edad	,149	21	,200*	,949	21	,322
Altura	,201	21	,027	,853	21	,005

*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

Fuente: Elaboración propia

Los resultados muestran una distribución normal de la edad ($0.200 > 0.05$) y no normal de la altura ($0.027 < 0.05$).

ANÁLISIS ESTADÍSTICO DESCRIPTIVO

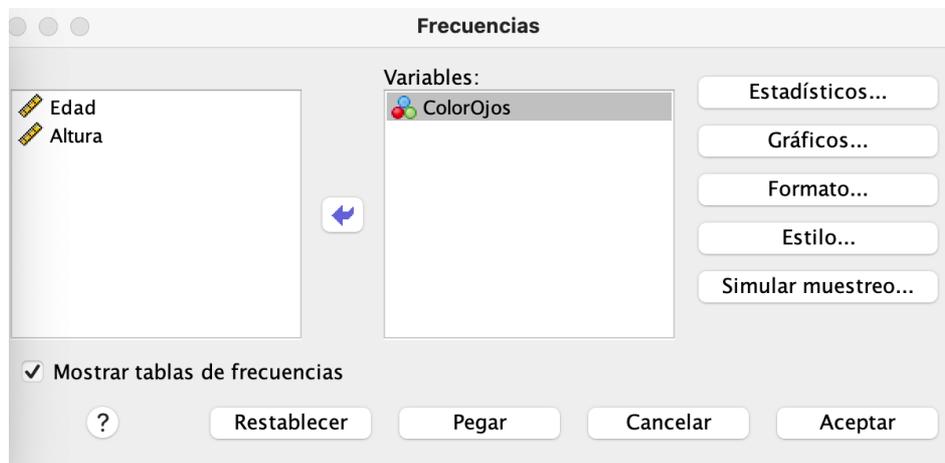
El análisis descriptivo de las **variables categóricas** consistirá en aportar la frecuencia de cada categoría (es decir, cuantas veces se repite el 1, 2 y 3 para la variable color de pelo), así como el porcentaje de representación con respecto al total de la muestra estudiada. Para adquirir los datos y poder transferirlos a una tabla, será necesario pinchar en Analizar – Descriptivos – Frecuencias.

The screenshot shows the SPSS Statistics interface. The main window displays a data table with the following columns: Edad, Altura, ColorOjos, and var. The data rows are numbered 1 to 21. The 'Analyze' menu is open, and the 'Frecuencias...' option is highlighted. Other options in the menu include Estadísticos descriptivos, Estadísticas Bayesianas, Tablas, Comparar medias, Modelo lineal general, Modelos lineales generalizados, Modelos mixtos, Correlacionar, Regresión, Loglineal, Redes neuronales, Clasificar, Reducción de dimensiones, Escala, Pruebas no paramétricas, Predicciones, Supervivencia, Respuesta múltiple, Análisis de valores perdidos..., Imputación múltiple, Muestras complejas, Simulación..., Control de calidad, Curva COR..., Modelado espacial y temporal..., and Marketing directo.

	Edad	Altura	ColorOjos	var
1	25,00	1,50	Azul	
2	23,00	2,10	Verde	
3	27,00	1,79	Azul	
4	28,00	1,56	Verde	
5	26,00	1,60	Marrón	
6	24,00	1,98	Marrón	
7	19,00	1,88	Marrón	
8	30,00	1,85	Azul	
9	32,00	1,90	Azul	
10	28,00	1,92	Azul	
11	24,00	1,56	Marrón	
12	25,00	1,98	Marrón	
13	27,00	1,96	Marrón	
14	27,00	1,97	Marrón	
15	21,00	1,98	Marrón	
16	28,00	2,00	Verde	
17	27,00	1,55	Verde	
18	25,00	1,53	Marrón	
19	25,00	1,99	Marrón	
20	26,00	1,98	Marrón	
21	27,00	1,61	Marrón	

Fuente: Elaboración propia

Introduciremos las variables categóricas para conocer su frecuencia. Aunque pueden introducirse también las variables cuantitativas, esto carece de sentido dado que no aporta información útil saber el número de personas que pesa 41, 42, 43 kg, etc.



Fuente: Elaboración propia

De nuevo, se abrirá en la ventana de resultados la frecuencia y porcentaje para los valores obtenidos en cada categoría de esa misma variable

Estadísticos

ColorOjos		
N	Válido	21
	Perdidos	0

UNI

		ColorOjos			
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Azul	5	23,8	23,8	23,8
	Verde	4	19,0	19,0	42,9
	Marrón	12	57,1	57,1	100,0
	Total	21	100,0	100,0	

INSE

Fuente: Elaboración propia

En el caso de las variables cuantitativas, proporcionaremos la tendencia central (media o mediana) y dispersión (desviación standard o rango intercuartil) en función de lo obtenido en las pruebas de normalidad. El procedimiento descrito aporta todos los valores, por lo que seleccionaremos sólo los necesarios durante la construcción de las tablas. Para ello iremos a Analizar – Descriptivos – Frecuencias.

	Edad	Altura	ColorOjos	var
1	25,00	1,50	Azul	
2	23,00	2,10	Verde	
3	27,00	1,79	Azul	
4	28,00	1,56	Verde	
5	26,00	1,60	Marrón	
6	24,00	1,98	Marrón	
7	19,00	1,88	Marrón	
8	30,00	1,85	Azul	
9	32,00	1,90	Azul	
10	28,00	1,92	Azul	
11	24,00	1,56	Marrón	
12	25,00	1,98	Marrón	
13	27,00	1,96	Marrón	
14	27,00	1,97	Marrón	
15	21,00	1,98	Marrón	
16	28,00	2,00	Verde	
17	27,00	1,55	Verde	
18	25,00	1,53	Marrón	
19	25,00	1,99	Marrón	
20	26,00	1,98	Marrón	
21	27,00	1,61	Marrón	

Fuente: Elaboración propia

Se introducirán las variables cuantitativas, se desmarcará la opción de mostrar tablas de frecuencias y en estadísticos habilitaremos los cuartiles, desviación standard, media y mediana.

The first screenshot shows the 'Frecuencias' dialog box with 'ColorOjos' in the 'Mostrar' list and 'Edad' and 'Altura' in the 'Variables:' list. The 'Mostrar tablas de frecuencias' checkbox is unchecked. The second screenshot shows the 'Frecuencias: Estadísticos' dialog box with 'Cuartiles' checked under 'Valores percentiles', 'Desviación estándar' checked under 'Dispersión', and 'Media' and 'Mediana' checked under 'Tendencia central'.

Fuente: Elaboración propia

Una vez aceptemos, obtendremos los datos para su selección.

Estadísticos

		Edad	Altura
N	Válido	21	21
	Perdidos	0	0
Media		25,9048	1,8186
Mediana		26,0000	1,9000
Desv. Desviación		2,86190	,19913
Percentiles	25	24,5000	1,5800
	50	26,0000	1,9000
	75	27,5000	1,9800

Fuente: Elaboración propia

CONSTRUCCIÓN DE TABLAS Y GRÁFICOS

Los datos obtenidos en el programa estadístico deben ser filtrados, dado que no todos los valores reportados tienen que ser tenidos en cuenta y solo cargarían las tablas de información poco útil.

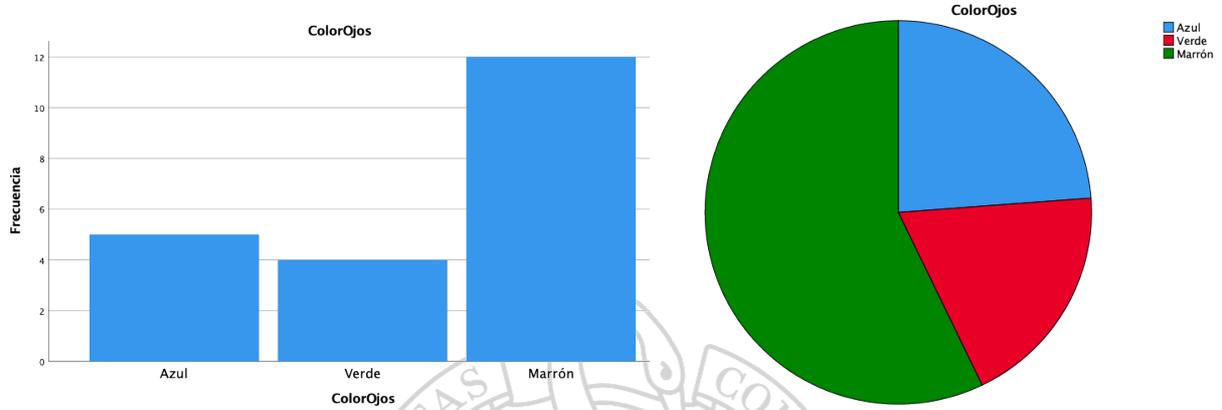
En cuanto a los gráficos, el programa estadístico ofrece la opción de generar múltiples gráficos. Para ello, iremos a Gráficos - Cuadro de diálogos antiguos – (Tipo de gráfico deseado).

The screenshot shows the SPSS Statistics interface. The main window displays a data table with columns: Edad, Altura, ColorOjos, and three empty 'var' columns. The 'Edad' column is selected, and the value '25,00' is highlighted. The 'Gráficos' menu is open, showing options like 'Generador de gráficos...', 'Selector de plantillas de tablero...', 'Comparar subgrupos', and 'Gráfico de Weibull...'. A sub-menu 'Cuadros de diálogo antiguos' is also visible, listing various chart types such as 'Barras...', 'Barras 3D...', 'Líneas...', 'Áreas...', 'Circular...', 'Máximos y mínimos...', 'Diagramas de cajas...', 'Barras de error...', 'Pirámide de población...', 'Dispersión/Puntos...', and 'Histograma...'.

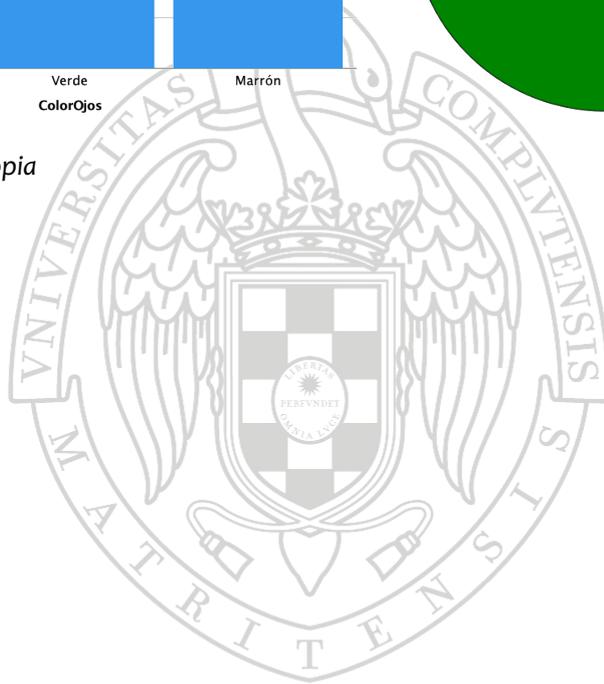
Fuente: Elaboración propia



Por ejemplo, estos son los gráficos de barra y circulares obtenidos para la frecuencia y porcentaje (respectivamente) del color de ojos de la muestra analizada.



Fuente: Elaboración propia



UNIVERSIDAD COMPLUTENSE
MADRID