***PRÁCTICA 1: GESTIÓN DE DATOS Y ESTADÍSTICA UNIVARIANTE***

Atanasov Angelov, Daniel – [daniel.atanasov24@estudiantes.uva.es](mailto:daniel.atanasov24@estudiantes.uva.es)

Sanz Tomé, Raúl – [raul.sanz24@estudiantes.uva.es](mailto:raul.sanz24@estudiantes.uva.es)

**Resumen:**

En esta práctica de estadística con Python, analizamos un conjunto de datos mediante resumen estadístico y tablas de frecuencias. Visualizamos la distribución con histogramas, diagramas de barras y sectores, y exploramos la dispersión con gráficos de caja y bigotes y de dispersión. El polígono de frecuencias relativas permitió identificar tendencias. Los resultados revelaron la variabilidad y la tendencia central de los datos, evidenciando posibles valores atípicos. Concluimos que la combinación de estas herramientas facilita la interpretación de los datos, permitiendo detectar patrones clave y características de la distribución para una mejor toma de decisiones.

**1. Introducción**

En esta práctica hemos trabajado con un conjunto de datos estadísticos de una sola variable, con el propósito de aprender a gestionarlos, manejar ficheros y analizar descriptivamente los resultados obtenidos.

**2. Metodología**

Los procesos que hemos seguido han sido los siguientes: comenzamos empleamos las bibliotecas necesarias para el manejo de esos datos, como numpy, pandas, matplotlib, entre otras.

Después hemos elaborado la tabla de frecuencias relacionada con los datos y, a continuación, hemos hecho el resumen estadístico con el que obtenemos los valores de la media, moda, mediana, y más.

Para finalizar, hemos realizado las correspondientes representaciones gráficas de los datos. Estas representaciones incluyen gráficos de dispersión, gráficos de caja y bigotes, histogramas, diagramas de barras y diagramas de sectores.

**3. Resultados**

Para empezar, hemos realizado tablas de frecuencias las cuales muestran de forma ordenada los datos y sus correspondientes frecuencias absolutas y relativas, tanto acumuladas como no. Junto a esto, hemos generado un gráfico de dispersión y otro de caja y bigotes. En el primero hemos visto como se dispersan los datos de cada una de las variantes, en concreto, las estaturas de las mujeres y de los hombres. Ambos poseen una mayor concentración en torno a los valores centrales o intermedios. Sin embargo, la distribución de las mujeres nos muestra que estas tienen más valores extremos y que, además, están más alejados de los valores centrales. En cuanto al segundo gráfico, observamos que el rango intercuartílico de la estatura de las mujeres se encuentra entre 1,62 y 1,82; mientras que en el caso de los hombres está entre 1,68 y 1, 85, aproximadamente en ambos casos. Además, observamos que la mediana en el caso de los hombres es menor que en el de las mujeres y, como veíamos en el gráfico anterior, los datos extremos en el caso de las mujeres llegan a estar más alejados de la zona central.