

## MIA—V1E4-002 | Data, Big Data and Open Data

Actividad Nro.1: Análisis de datos recolectados - Google Colab

Nombre:

Fecha:

1. A partir del dataset **ventas-por-factura.xlsx** graficar el **Monto** total de ventas mensual de todos los registros (sin las devoluciones) y concluir la gráfica en un par de palabras cuál es el mes con mayor Monto y menor monto ventas, además indicar el número de ventas (**Cantidad**):
  - a. Gráfico:
  - b. Conclusión:
2. A partir de laboratorio **S01\_P01\_Analisis\_de\_datos\_recolectados\_prac\_v2.ipynb** y el dataset **S01\_P01\_Captura-de-información\_MIAV1E4\_v01.csv** colectado en un grupo de personas, obtén los siguientes resultados:
  - a. Generar una gráfica de histograma de frecuencia con los campos "**Entorno\_de\_preferencia**" y en un párrafo interpreta los resultados.
  - b. A partir del campo "**Peso**" calcular las estadísticas "**promedio**", "**mediana**", "**min**", "**max**". Interpretar los resultados y comentar sobre la consistencia de los datos, ¿son correcto los datos?
  - c. Realizar un diagrama de dispersión con el campo "**Edad**" y "**Peso**" y obtén una interpretación de los resultados.

**Extra (opcional):**

- d. Realiza un histograma de frecuencia por un rango de edades ( 20-25, 26-30,31-35, 36-40, mayor a 41). Realiza una interpretación de los resultados
- e. ¿Genera un Gráfico de Cajas y calcula los cortes externos (outlayer) para el campo “¿**Altura**”, cuantos datos están fuera de los extremos (outlayers)? ¿Y qué utilidad podemos observar a la capacidad de identificar los datos extremos o anómalos?

Plazo de entrega, hasta el día jueves 14 de agosto del 2025 a las 23.59