



**CURSO: MINERÍA DE DATOS**  
**MAESTRÍA EN EXPLOTACIÓN DE DATOS Y DESCUBRIMIENTO DE CONOCIMIENTO**

LABORATORIO I: Conceptos de Estadística Descriptiva

**INTRODUCCIÓN:**

Esta práctica inicial de laboratorio tiene como objetivo realizar una primera aproximación al Lenguaje R, utilizando el enfoque de análisis exploratorio de datos sobre un dataset, a efectos de repasar conceptos fundamentales de estadística descriptiva.

**CONSIGNAS:**

A partir del dataset *MPI\_subnational.csv*<sup>1</sup> (Multidimensional Poverty Measures), se solicita trabajar sobre las siguientes consignas:

1. **Exploración de datos.** Explore y explique en qué consiste el dataset utilizando herramientas de exploración de datos.
  - a. Releve las características de los atributos.
  - b. Represente gráficamente la cantidad de ciudades agrupados por Región.
2. **Medidas de posición.** Calcule las medidas de posición para los atributos numéricos y agrupe los cálculos de acuerdo al País (*atributo Country*).
  - a. Ordene los resultados de la media observada para el MPI resultante y concluya respecto a los valores observados en términos de la pobreza. Help(order).
  - b. Grafique las variables y observe su comportamiento (graph : barplot, pie & hist).
3. **Medidas de dispersión.** Calcular el desvío estándar, la varianza y el rango para cada una de las variables.
  - a. Realice diagramas de cajas y scatterplot's. Documente las conclusiones.
4. **Medidas de asociación.** Calcular el coeficiente de correlación de todas las variables y explique el resultado. ¿Qué tipo de gráficos describen mejor esta relación entre las variables?

Referencias sugeridas:

An Introduction to R: <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>

---

<sup>1</sup> Disponible en: <https://www.kaggle.com/ophi/mpi/data>