



Universidad de Antioquia  
Especialización en Analítica y Ciencia de Datos  
Estadística y Análisis Exploratorio  
Trabajo 2  
Profesor Javier Lozano

## Trabajo 2

La base de datos con la que le corresponde trabajar, se obtiene como una muestra aleatoria que corresponde a los resultados obtenidos por los jóvenes de todo el país en las pruebas Saber 11, del año 2022. La información aparece agregada por Colegio y solo abarca los colegios de Antioquia. Dicha base contiene las variables: NATURALEZA (Colegio Oficial o No-Oficial), JORNADA (Asociado a los diferentes tipos de jornada de estudio), PROMLECT (Promedio obtenido por el colegio en Lectura Crítica) y PROMMATE (Promedio obtenido por el colegio en Matemáticas).

1. Construir en Python un IC para al 95 % de confianza necesario para responder: ¿Puede afirmarse que el resultado medio obtenido por los colegios en Lectura es superior a 45 puntos? Justifique su respuesta.
2. ¿Se puede afirmar que el resultado medio obtenido por los colegios en Matemáticas es mayor en los Privados que en los Públicos? Justifique su respuesta mediante una prueba de hipótesis con una significancia del 4%.
3. Analice si hay una relación lineal entre las variables Promedio obtenido por el colegio en Lectura Crítica y Promedio obtenido por el colegio en Matemáticas, mediante una regresión lineal con todos sus respectivos componentes.