

Tarea 2

En esta tarea utilizaremos los datos **Hitters** de la librería **ISLR**. Estos datos contiene información del salario y medidas de rendimiento de 322 jugadores de baseball profesionales en 1986 a 1988. El objetivo principal es determinar los factores que afectan el salario.

Las variables que se utilizarán son:

- **Salary**: salario del jugador anual en 1987 (miles de dolares).
- **Hits**: número de aciertos al bate en 1986.
- **CHits**: número de aciertos al bate durante toda la carrera.
- **Runs**: número de carreras en 1986.
- **CRuns**: número de carreras durante toda la carrera.
- **HmRun**: número de jonrones en 1986.
- **CHmRun**: número de jonrones durante toda la carrera.
- **RBI**: número de carreras impulsadas en 1986.
- **CRBI**: número de carreras impulsadas durante toda la carrera.

1. Ajuste un modelo para el salario en función de las demás variables (expresé claramente el modelo). Evalúe los supuestos. Si no se cumple alguno, haga transformaciones para corregirlo.
2. Evalúe si hay problemas de multicolinealidad usando los indicadores vistos en clase.
3. Utilice la regresión de ridge para corregir los problema de multicolinealidad. Considere varios valores de k y seleccione un valor óptimo.
4. Utilice la regresión por componentes principales para corregir el problema de multicolinealidad.
5. Compare los modelos por mínimos cuadrados ordinarios, regresión de ridge, y regresión por componentes principales. ¿cuáles son las covariables que tienen un efecto significativo sobre el salario del jugador?

Para la entrega, tenga en cuenta lo siguiente:

- El reporte no debe exceder 6 páginas (texto en una sola columna). No incluir códigos o salidas de R (cree sus propias tablas).
- En el reporte, todas las tablas y figuras que incluyan deben estar enumeradas y referenciadas en texto (por ejemplo, “como se observa en la Figura 1, ...” o “la Tabla 2 presenta ...”). Solo incluya tablas y figura que sean relevantes (es decir, todas las tablas y figuras debe tener una referencia en el texto).
- Fecha de entrega: 14 de abril de 2023 al empezar la clase (en físico y a través del campus virtual).