

Ejercicios de la clase_2:

- 1) Construir una carpeta Ejercicio2
- 2) Bajar desde Paideia:
SudAmerica.csv,
Baseball.csv,
Students.csv,
- 3) Abrir Rstudio en la carpeta misma con Rmarkdown
- 4) cargar los archivos
- 5) revisando el conjunto de comandos en Clase_2.R correr la regresión de IDH con Tajo Urbano
- 6) realizar una función que tome en entrada regresor y criterio y salen los resultados. Se haga de manera que salgan también pesos, pendientes, y apalancamiento del regresor.
- 7) aplicar a SudAmerica para averiguar que todo funciona y luego sobre Baseball: modelar WinningPerc utilizando BattingAvg
Students: modelar final con midterm
- 8) sobre los tres archivos, correr con *lm()* los modelos, el modelo sin intercepta (se escribe “criterio ~ regresor -1”), el modelo con intercepta pero tirando la unidad con mayor apalancamiento y el tirando la unidad con mayor pendiente.
- 9) representar en gráfico las cuatro rectas conseguidas.
- 10) Comentar las diferencias encontradas.
- 11) guardar los archivos .Rmd y .html y enviar-los a sergio@camiz.it antes de las 23:59 del 21 de Abril.
- 12) identificar la base de datos para el examen final, describir el objetivo del trabajo y la estructura de los datos (2 páginas) enviar a sergio@camiz.it

antes de las 23:59 del 21 de Abril. Los datos tienen que observar el siguiente:

- Numero de observaciones: entre 40 y 500 (por el caso, una muestra)

- número de variables entre 5 y 10

- la variable respuesta siempre cuantitativa

- las variables explicativas cuantitativas sino una/dos cualitativas

- Evitar series de tiempo para garantizar la independencia de las observaciones.