

# TALLER EVALUADO 1 – REGRESIÓN LINEAL

## Instrucciones

- La evaluación puede ser realizada de forma individual o en parejas de dos personas, entregando un solo archivo.
- La solución debe ser entregada a más tardar el día sábado 31 de julio a las 23:59hrs, **no se aceptarán pruebas fuera de este plazo.**
- El formato de archivo **debe** tener de nombre el **Apellido\_Nombre\_sección**, por ejemplo, Muñoz\_Diego\_Johnson, o en caso de ser pareja, **Muñoz\_Diego\_Johnson-Paredes\_Camila\_Turing**
- Los integrantes del grupo pueden ser de diferentes secciones
- **Solo un** integrante del grupo debe subir la evaluación al buzón
- En caso de realizar la solución en Rmarkdown, entregar solo el archivo HTML/PDF el cual deberá tener el código utilizado seguido de la solución de cada pregunta.
- Cada pregunta tiene una puntuación de 1p.

# REGRESIÓN LINEAL

Se le contratan sus servicios como Data Science en un centro meteorológico para poder crear un modelo que pueda estimar la diferencia de temperatura que habrá en un día (temperatura máxima menos la mínima), utilizando la información del día anterior, para ello se dispone de 40000 observaciones y un conjunto de 18 variables.

VARIABLE		DESCRIPCIÓN
Lluvia		Cantidad de lluvia registrada ese día en mm.
Koppen		Clasificación koppen de la zona de medición (Temperate, Subtropical, Grassland, Tropical, Desert)
Estación		Estación del Año
Evaporacion		Evaporación (mm) en 24 horas
Sol		Número de horas de sol brillante en el día
Dir9am		Dirección del viento a las 9am
Dir3pm		Dirección del viento a las 3pm
Hum9am		Porcentaje de humedad a las 9am
Hum3pm		Porcentaje de humedad a las 3pm
Pres9am		Presión atmosférica (hpa) a nivel del mar a las 9am
Pre3pm		Presión atmosférica (hpa) a nivel del mar a las 3pm
Nub9am		Fracción del cielo cubierto por nubes a las 9am. Se mide en "octavos", de manera que un valor 0 indica cielo totalmente despejado y 8, cielo totalmente cubierto.
Nub3pm		Fracción del cielo cubierto por nubes a las 3pm. Se mide en "octavos", de manera que un valor 0 indica cielo totalmente despejado y 8, cielo totalmente cubierto.
Temp9am		Temperatura en grados Celsius a las 9am
Temp3pm		Temperatura en grados Celsius a las 3pm
LluviaHoy		Variable indicadora que toma el valor 1 si la precipitación en mm. en las últimas 24 hrs. excede 1 mm. y 0 si no.
VelRafaga		La velocidad (km / h) de la ráfaga de viento más fuerte en las 24 horas hasta la medianoche
Target		Diferencia de temperatura el día siguiente de la medición

# REGRESIÓN LINEAL

**Pregunta 1:** Cargue los datos **lluvia.csv**, verificando que estos hayan sido leídos correctamente, codifique como factor las variables que corresponda, además, genere un resumen exploratorio de los datos.

**Pregunta 2:** Para las variables continuas, genere un análisis gráfico de dichas variables con el target y un análisis de correlación lineal

**Pregunta 3:** Realice un modelo de regresión lineal simple para modelar la diferencia de temperatura utilizando la covariable Temp3pm, además, genere el gráfico del target vs Temp3pm y añada la curva ajustada. ¿Es esta variable significativa?, ¿Cuál es la variabilidad explicada por el modelo?

**Pregunta 4:** Ajuste un modelo de regresión lineal utilizando la metodología forward y backward, ¿Los modelos obtenidos son diferentes?, ¿Son estos modelos significativos?, ¿Cuál es el porcentaje de variabilidad explicada por el mejor de estos modelos?

**Pregunta 5:** Realice un estudio de residuos con los principales supuestos del mejor modelo de la pregunta 4, por medio de test estadísticos y gráficos. ¿Es un modelo adecuado?

**Pregunta 6:** Se dispone de un conjunto de datos **Datos extras.csv** en la cual se tiene información de 5000 tomas de muestras de una central meteorológica, utilizando el mejor modelo de la pregunta 4, genere la estimación de diferencias de temperaturas para dichas observaciones

**DESARROLLE LOS EJERCICIOS EN SU SCRIPT (.R o .RMD), GUARDELO CON SU APELLIDO\_NOMBRE\_SECCIÓN Y CARGUELO EN EL BUZÓN CORRESPONDIENTE A SU SECCIÓN HASTA EL DÍA SÁBADO 31 DE JULIO A LAS 23:59 HRS**