

# Trabajo Práctico 2

## Enunciado

A partir del dataset resultado del TP1 vamos a construir un modelo de regresión lineal que estime el precio por m<sup>2</sup> de las propiedades.

El modelo puede construirse involucrando sólo un subconjunto del dataset limpio, con algunas características particulares. Por ejemplo, sólo departamentos en los barrios de Belgrano y Palermo ó sólo PHs del barrio de Almagro y Villa Crespo, ó sólo propiedades industriales en Pompeya. El dataset que modelen debe tener al menos 2000 registros.

## Objetivos:

- Seleccionar los datos a modelar. Justificar esa selección.
- Entrenar al menos tres modelos de regresión lineal distintos sobre estos datos.
- Evaluar la performance de los modelos y comparar ventajas y desventajas de cada uno.
- Determinar la importancia de las features incluidas en la predicción. Interpretar.

## Requisitos

Los materiales deberán ser entregados en una Jupyter notebook que satisfaga los requerimientos del proyecto. La notebook deberá estar debidamente comentada. Además, los grupos deberán crear un repositorio para el proyecto (anonimizado) en Github. Para la presentación deben armar algunos slides y grabar un video que exponga los resultados obtenidos en no más de 10 minutos.

## Material a entregar

Presentar un informe estructurado (PPT o Google Slides). Éste debe constar en una introducción (planteo del problema, la pregunta, la justificación del dataset utilizado, etc.), un desarrollo de los análisis realizados (modelos propuestos, ventajas y desventajas, visualizaciones) y una exposición de los principales resultados y conclusiones.

## Fecha de entrega

- El material deberá entregarse en la **clase 26** del curso.

## Evaluación

- Seleccionar los datos a modelar. Justificar esa selección.
- Entrenar al menos tres modelos de regresión lineal distintos sobre estos datos.
- Evaluar la performance de los modelos y comparar ventajas y desventajas de cada uno.
- Determinar la importancia de las features incluidas en la predicción. Interpretar.

|      |   |
|------|---|
| /10  | Discusión del dataset:<br>Selección de los datos a modelar. Justificación.  |
| /10  | Discusión del método:<br>Entrenamiento modelo 1 de regresión lineal y evaluación de performance                       |
| /10  | Discusión del método:<br>Entrenamiento modelo 2 de regresión lineal y evaluación de performance                       |
| /10  | Discusión del método:<br>Entrenamiento modelo 3 de regresión lineal y evaluación de performance                       |
| /10  | Discusión de los resultados: Comparación de los modelos, ventajas y desventajas                                       |
| /10  | Adecuada profundidad del trabajo  |
| /10  | Adecuada presentación de conclusiones   |
| /10  | Formato del informe (introducción/cuerpo/conclusión, gráficos)  |
| /10  | Presentación del informe (estructura del texto, storytelling, figuras)  |
| /10  | Presentación del video (tiempo adecuado, claridad en la exposición, información relevante para la audiencia definida) |
| /100 | Nota final  |