

STATISTICAL ESSENTIALS

Laboratorio Practico N°04:

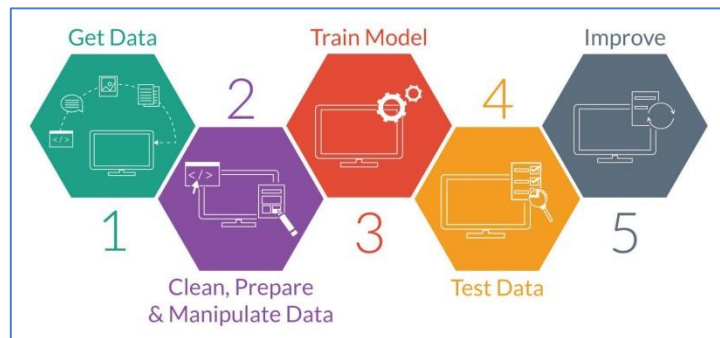
Se entrega el siguiente set de datos pertenecientes al sector inmobiliario, **bostonvivienda.csv**, el cual es un dataset que cuenta con las principales características de los hogares de un estado del país de EEUU. La empresa que está haciendo un estudio de los principales lugares de inversión de bienes raíces, necesita saber cuál es el precio medio de las casas para su posterior venta y para esto le pide a usted ayudarlo a elaborar a través de la analítica predictiva una solución a su problema de negocio.

Debido a que usted está llevando un programa de especialización en Estadística requieren que los ayude a identificar los **PRECIOS MEDIOS DE LAS CASAS**.

- Las variables que se disponibilizan son :
 - 1) **crim**: tasa de delincuencia per cápita por ciudad.
 - 2) **zn**: proporción de suelo residencial dividido en zonas para lotes de más de 25,000 pies cuadrados.
 - 3) **indus**: proporción de acres de negocios no minoristas por la ciudad.
 - 4) **nox**: concentración de óxidos de nitrógeno (partes por 10 millones).
 - 5) **rm**: número promedio de habitaciones por vivienda.
 - 6) **edad**: proporción de unidades ocupadas por sus propietarios construidas antes de 1940.
 - 7) **dis**: media ponderada de las distancias a cinco centros de empleo de Boston.
 - 8) **rad**: Índice de la accesibilidad a las autopistas radiales.
 - 9) **impuesto**: tasa de impuestos a la propiedad por el valor total por \ \$ 10.000.
 - 10) **ptratio**: proporción de alumnos por profesor por ciudad.
 - 11) **negro**: $1000 (B_k - 0,63)^2$, donde B_k es la proporción de negros por la ciudad.
 - 12) **lstat**: estatus más bajo de la población (por ciento).
 - 13) **medv**: valor mediano de las viviendas ocupadas por sus propietarios en \ \$ 1000s.

Entregables del Laboratorio:

- 1.- Definir el problema de la naturaleza que se tiene a continuación, además de los objetivos de negocio bien definidos.
- 2.- ¿Qué tipo de variables se utilizan en el problema de negocio?
- 3.- Construya mediante técnicas de dependencia de respuesta numérica una solución analítica para determinar el precio medio de las casas de la localidad en estudio.



- 4.- De los drivers o features o variables independientes seleccionados cuáles de éstos son los más importantes para determinar los precios de las casas.
- 5.- ¿Cuáles son las principales métricas de negocio o ratios esperados para la gestión con la solución analítica que estamos construyendo?

