

Problema de Entrenamiento 1

December 21, 2021

Autor: Harvey Rodriguez Gil

Problemática:

Cálculo del precio de los inmuebles usados de tipo apartamento para la ciudad de Medellín

Descripción:

El mundo del sector inmobiliario cada vez se mueve más a nivel de transacciones (venta y compra) tanto de propiedades nuevas como de usadas. Para el caso de las propiedades nuevas, el precio de estas viene fijado normalmente por la constructora a partir de un estudio de mercado que les permita identificar el mejor precio. Para el caso de las propiedades usadas, es la persona interesada en comprar o vender quien está a cargo de la labor de identificar si el valor propuesto está bien definido. Para esto, se podría generar un algoritmo el cual a partir de ciertas variables de interés genere el valor de la propiedad y no quede sesgado por sentimiento por parte del vendedor y/o comprador. Por esta razón, se plantea la importancia de poder predecir el precio del inmueble a partir de sus características.

Descripción de variables de interés:

- **BBDD_ID:** Identificador de la propiedad en la base de datos.
- **PRICE [cuantitativa - continua]:** Valor comercial de la propiedad
- **CONSTRUCTION_DATE:** Fecha de construcción de la propiedad
- **NEIGHBORHOOD [cualitativa]:** Barrio donde se encuentra localizada la propiedad.
- **PARKING_LOTS [cuantitativa - discreta]:** Parqueaderos asociados a la propiedad.
- **AREA [cuantitativa - continua]:** Área de la propiedad
- **ADMINISTRATION_PRICE [cuantitativa - continua]:** Valor de la administración de la propiedad
- **STRATUM [cuantitativa - discreta]:** Estrato de la zona donde se encuentra la propiedad
- **ELEVATOR [cualitativa]:** Tiene elevador la unidad/edificio donde se encuentra la propiedad
- **GREEN_ZONES [cualitativa]:** Tiene zonas verdes la unidad/edificio donde se encuentra la propiedad
- **GYM [cualitativa]:** Tiene gimnasio la unidad/edificio donde se encuentra la propiedad
- **POOL [cualitativa]:** Tiene piscina la unidad/edificio donde se encuentra la propiedad
- **TURCO [cualitativa]:** Tiene baño turco la unidad/edificio donde se encuentra la propiedad
- **VISITOR_PARKING_LOTS [cualitativa]:** Tiene parqueadero de visitantes la unidad/edificio donde se encuentra la propiedad
- **YARD [cualitativa]:** Tiene jardín la unidad/edificio donde se encuentra la propiedad
- **KIDS_PARK [cualitativa]:** Tiene juegos infantiles la unidad/edificio donde se encuentra la propiedad

- **BBQ [cualitativa]:** Tiene zona de bbq la unidad/edificio donde se encuentra la propiedad
- **SOCIAL_ROOM [cualitativa]:** Tiene salón social la unidad/edificio donde se encuentra la propiedad
- **SPORTS_FIELDS [cualitativa]:** Tiene canchas de deportes la unidad/edificio donde se encuentra la propiedad
- **SOLARIUM [cualitativa]:** Tiene solario la unidad/edificio donde se encuentra la propiedad
- **STORAGE_ROOM [cuantitativa - discreta]:** Cuarto útil asociado a la propiedad
- **BEDROOM [cuantitativa - discreta]:** Cantidad de dormitorios que tiene la propiedad
- **BATHROOM [cuantitativa - discreta]:** Cantidad de baños que tiene la propiedad
- **HALF_BATHROOM [cuantitativa - discreta]:** Tiene baño social la propiedad
- **VIEW_TYPE [cualitativa]:** Tipo de vista que tiene el apartamento: Exterior, interior.
- **PARKING_LOT_TYPES [cualitativa]:** Tipo de parqueadero que tiene la propiedad: lineal, sencillo
- **KITCHEN_FEATURES [cualitativa]:** Características de la cocina que tiene la propiedad
- **BALCONIES [cuantitativa - discreta]:** Tiene balcón la propiedad
- **LIVING_ROOMS [cuantitativa - discreta]:** Tiene sala de estar la propiedad
- **DINNING_ROOMS [cuantitativa - discreta]:** Tiene comedor la propiedad
- **LAUNDRY_AREAS [cuantitativa - discreta]:** Tiene area de ropas las propiedad
- **SERVICE_ROOMS [cuantitativa - discreta]:** Tiene habitación para la persona del servicio
- **DRESSING_ROOMS [cuantitativa - discreta]:** Tiene vestier la propiedad
- **SERVICE_BATHROOMS [cuantitativa - discreta]:** Tiene baño de servicio la propiedad
- **STUDYING_ROOMS [cuantitativa - discreta]:** Tiene estudio la propiedad

```
[ ]: import pandas as pd
from IPython.display import display

df = pd.read_csv('precio_propiedades.csv')
pd.options.display.max_columns = None
df.head()
```

```
[ ]:  BBDD_ID      PRICE      CONSTRUCTION_DATE      NEIGHBORHOOD  PARKING_LOTS  \
0    327727  420000000  2007-01-01 5:00:00      Manila      0
1    507188  270000000  2011-01-01 5:00:00  Loma de los Bernal      1
2    507344  450000000  2009-01-01 5:00:00      Estadio      1
3    420492  350000000  1995-01-01 5:00:00    Las Acacias      0
4     97630  510000000  2009-01-02 0:00:00  Conquistadores      2

      AREA  ADMINISTRATION_PRICE  STRATUM  STORAGE_ROOM  BEDROOM  BATHROOM  \
0   155.0             0.0         2.0           NaN         3         2
1    67.0          301000.0         4.0           NaN         3         2
2    84.0          360000.0         5.0           NaN         2         2
3   118.0           190.0         5.0           NaN         4         3
4   113.0          620000.0         5.0           NaN         3         3
```

	HALF_BATHROOM	VIEW_TYPE	PARKING_LOT_TYPES	KITCHEN_FEATURES	BALCONIES	\
0	NaN	NaN	NaN	NaN	NaN	
1	NaN	NaN	NaN	Integral	1.0	
2	NaN	NaN	Cubierto	Integral	1.0	
3	NaN	NaN	Cubierto	NaN	1.0	
4	NaN	NaN	NaN	Integral	1.0	

	LIVING_ROOMS	DINNING_ROOMS	LAUNDRY_AREAS	SERVICE_ROOMS	DRESSING_ROOMS	\
0	1.0	1.0	NaN	NaN	NaN	
1	1.0	1.0	1	NaN	NaN	
2	1.0	1.0	1	NaN	NaN	
3	1.0	1.0	1	1.0	NaN	
4	0.0	NaN	1	NaN	NaN	

	SERVICE_BATHROOMS	STUDYING_ROOMS	GREEN_ZONES	GYM	POOL	TURCO	\
0	NaN	NaN	NO	NO	NO	NO	
1	NaN	NaN	YES	NO	NO	NO	
2	NaN	1.0	NO	NO	NO	NO	
3	1.0	NaN	NO	NO	NO	NO	
4	NaN	1.0	NO	NO	NO	NO	

	VISITOR_PARKING_LOTS	YARD	KIDS_PARK	BBQ	SOCIAL_ROOM	SPORTS_FIELD	SOLARIUM
0	NO	NO	NO	NO	NO	NO	NO
1	YES	NO	YES	NO	YES	NO	NO
2	YES	NO	NO	NO	NO	NO	NO
3	NO	NO	NO	NO	NO	NO	NO
4	NO	NO	NO	NO	YES	NO	NO