

# Desafio 2

## Tasador de Propiedades

### Grupo 4

Barrionuevo, Nicolas  
Carrillo, Lucas  
Carrascosa, Alejandro

Introducción

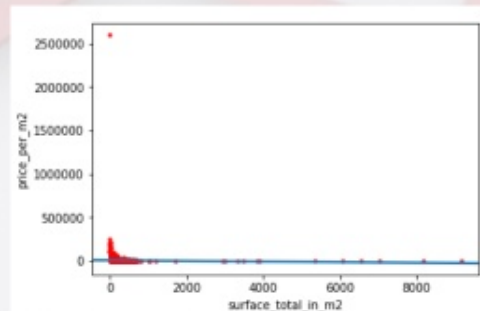
Primer  
Modelo

Modelo  
Final?

Segundo  
Modelo

Conclusiones

- Decidimos usar los campos que el tasador de properatti online usa y para reducir un poco los datos nos concentramos en la información de Capital Federal
- Creamos un primer modelo con datos crudos y las principales variables categóricas



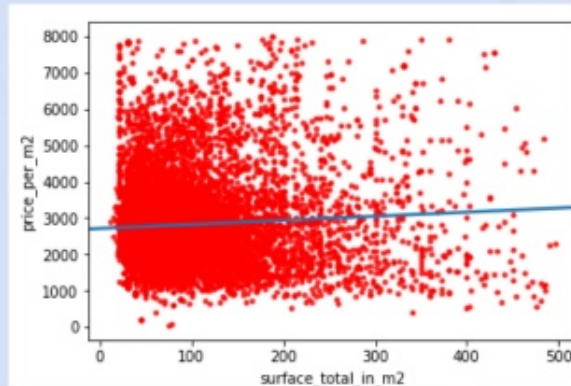
### Score del Modelo

El score del modelo para este primer modelo fue 0,0017

**Limpieza**

## Limpieza

Despues de tan terrible score, supimos que algo habia mal por lo tanto limpiamos outliers a groso modo para tener una distribución mas normal



# Desafío 2

## Tasador de Propiedades

### Grupo 4

Barrionuevo, Nicolas  
Carrillo, Lucas  
Carrascosa, Alejandro

Introducción

Primer  
Modelo

Modelo  
Final?

Segundo  
Modelo

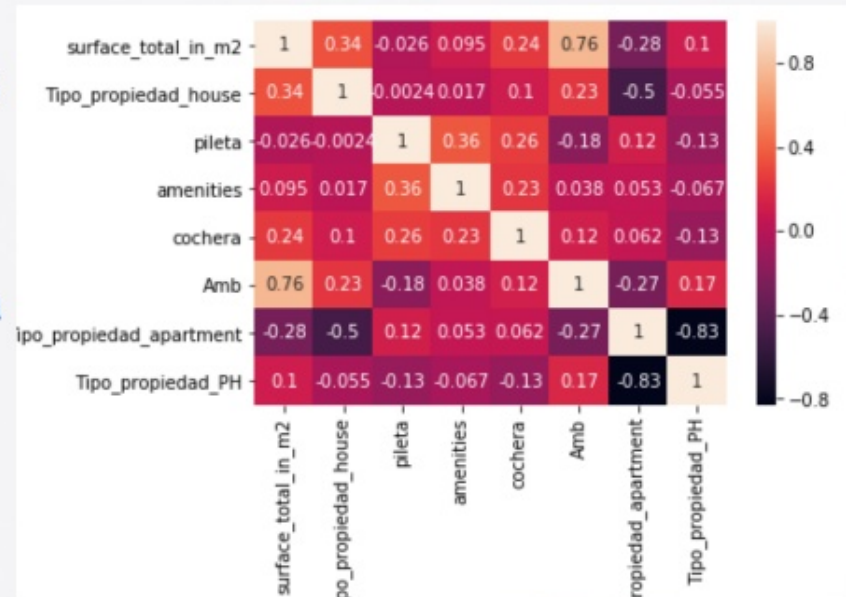
Conclusiones



# Primer Modelo

- Con el dataset limpio creamos nuestro primer modelo usando solo algunas de las variables categoricas ej: Ambientes, Superficie, Tipo de Propiedad. Con esas variables obtuvimos un score de 0,25
- Convertimos en variables dummies la cantidad de ambientes, si tiene cochera o no, si tiene amenities o no, etc
- Viendo que obtuvimos un mejor score con mas variables, creamos un nuevo modelo utilizando todas incluyendo todas las dummies el cual el resultado fue 0,51

Heatmap de variables usadas



# Desafío 2

## Tasador de Propiedades

### Grupo 4

Barrionuevo, Nicolas  
Carrillo, Lucas  
Carrascosa, Alejandro

Introducción

Primer  
Modelo

Modelo  
Final?

Segundo  
Modelo

Conclusiones

## Segundo Modelo

- Para mejorar el modelo intentamos usar Ridge y Lasso
- Luego de normalizar las variables categóricas (No dummies)
- Usamos StandardScaler para tratar las variables categóricas (Sin dummies) obtuvimos 2 conclusiones
  - Normalizado y pocas variables Score:0,20
  - Normalizado y todas las variables Score: 64450,04 Horrible

# Desafío 2

## Tasador de Propiedades

### Grupo 4

Barrionuevo, Nicolas  
Carrillo, Lucas  
Carrascosa, Alejandro

Introducción

Primer  
Modelo

Modelo  
Final?

Segundo  
Modelo

Conclusiones



## No hubo modelo 3

Hicimos varias pruebas y no pudimos mejorar los modelos que teníamos y nos quedamos con el ridge que fue el mejor

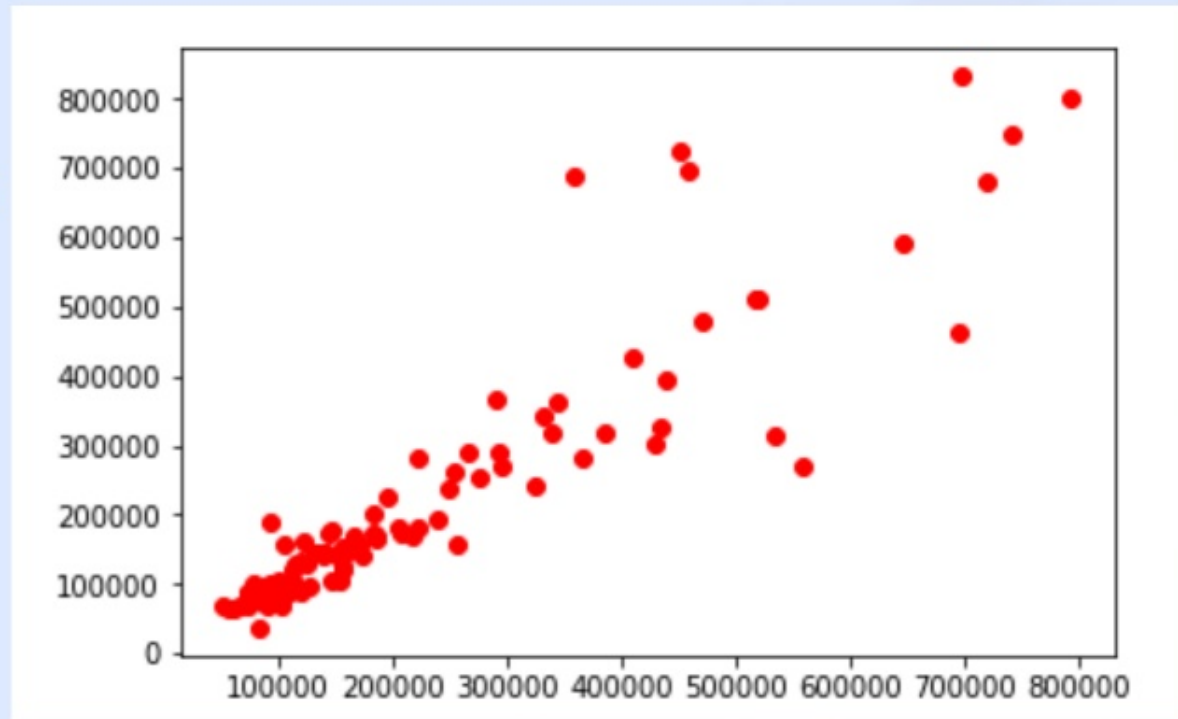
**Modelo  
Ridge**

**Distribución**

# Ridge

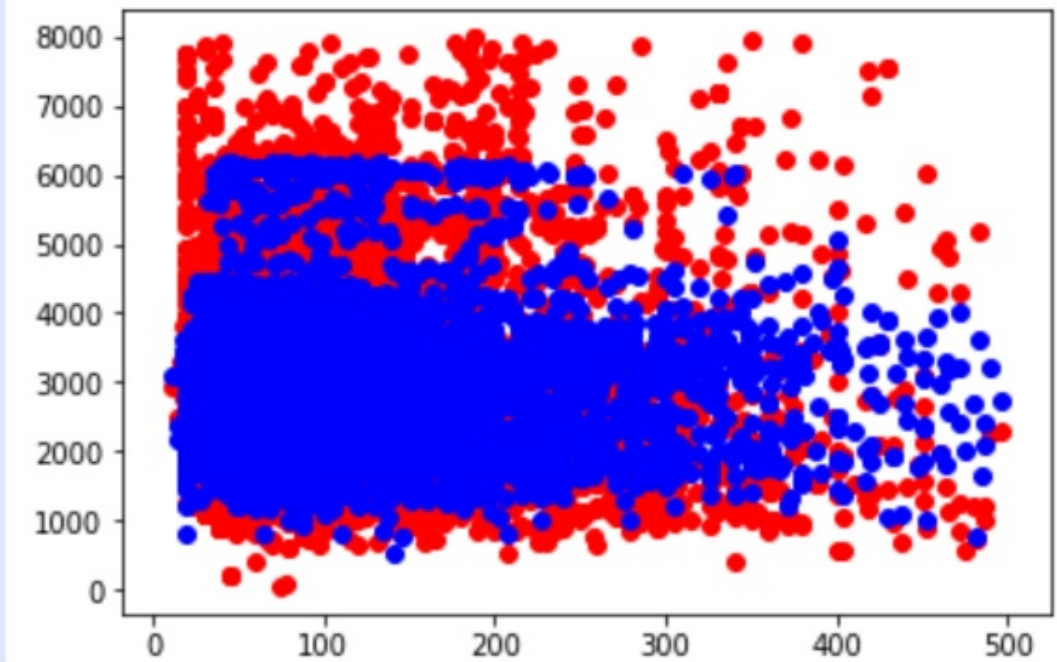
Score: 0,5105

La imagen muestra la  
distribucion entre Precio  
Calculado y Precio Estimado



# Distribución

La imagen muestra la  
distribucion de Superficie Total y  
el precio x M2  
Rojo : Actual  
Azul: Estimado



# Desafio 2

## Tasador de Propiedades

### Grupo 4

Barrionuevo, Nicolas  
Carrillo, Lucas  
Carrascosa, Alejandro

Introducción

Primer  
Modelo

Modelo  
Final?

Segundo  
Modelo

Conclusiones



## Que descubrimos?

- Mayor numero de variables mejor score
- Los outliers impactan mucho en los modelos

surface_total_in_m2	price_per_m2	precio	id	precio_m2_pred	precio_pred	valuation	dif_precio
120.0	652.173913	78260.889565	41909	2277.180523	273261.662776	Subvaluada	-1625.006610
30.0	1233.333333	37000.000000	26178	2759.592755	82787.782638	Subvaluada	-1526.259421
49.0	1472.448980	72150.000000	10791	2761.855712	135330.929896	Subvaluada	-1289.406733
70.0	2642.857143	185000.000000	56518	3917.577300	274230.410973	Subvaluada	-1274.720157
133.0	2804.511278	373000.000000	117155	3923.329250	521802.790215	Subvaluada	-1118.817972

# Desafío 2

## Tasador de Propiedades

### Grupo 4

Barrionuevo, Nicolas  
Carrillo, Lucas  
Carrascosa, Alejandro

Introducción

Primer  
Modelo

Modelo  
Final?

Segundo  
Modelo

Conclusiones