



# Reto técnico Data Science

Salvador Barcenás Valladolid

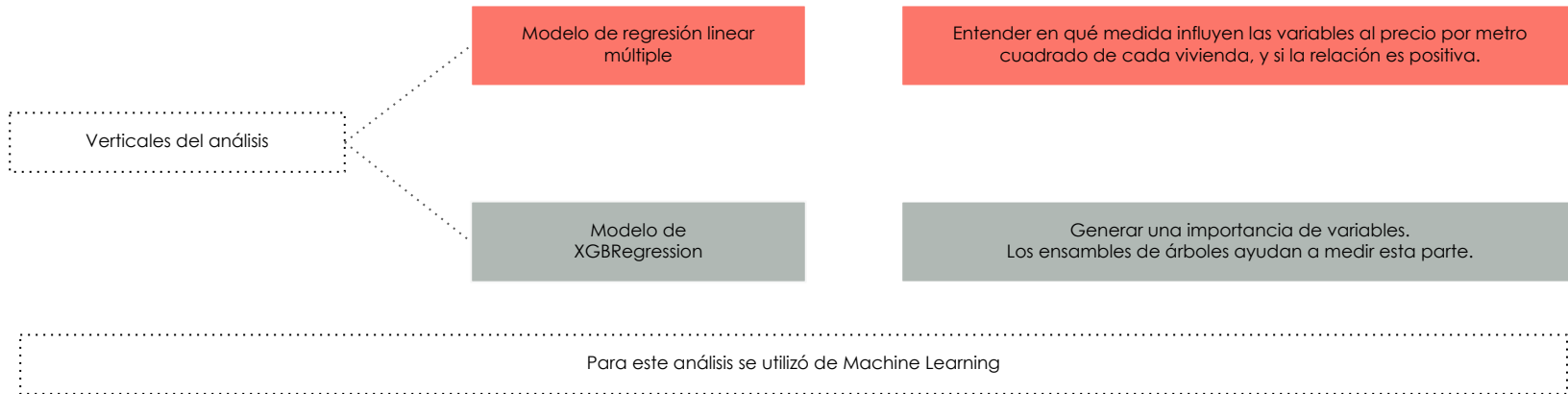
# Precio por metro cuadrado

**Objetivo:** Entender los factores que influyen en el precio por metro cuadrado de cada vivienda.

- Generar un modelo de regresión lineal múltiple
- Realizar un modelo XGBoost para entender la importancia de variables

**Fuentes externas:** Para realizar este proyecto, se utilizaron las siguientes fuentes externas

- *DENUE*
- *Censo poblacional 2020*
- *Índice de Rezago Social del CONEVAL*
- *Boletín estadístico de la incidencia delictiva*
- *Marco geoestadístico a nivel AGEB INEGI*



# Hallazgos

## Ranking de importancia de variables

Variable	Ranking importancia	Relación	Beta
Número de estacionamientos	1	Positiva	5,824
Distancia al metro más cercano	2	Positiva	3.47
Comodidades	3	Positiva	274
Número de baños	4	Positiva	3,052
Número de establecimientos de alimentos y bebidas	5	Positiva	5.58
Número de de servicios inmobiliarios y de alquiler	6	Negativa	-532
Número de pisos	7	Positiva	2,160
El departamento es de tipo loft	8	Positiva	19,421

\* Los universos de establecimientos son en el AGEB donde se encuentra el departamento

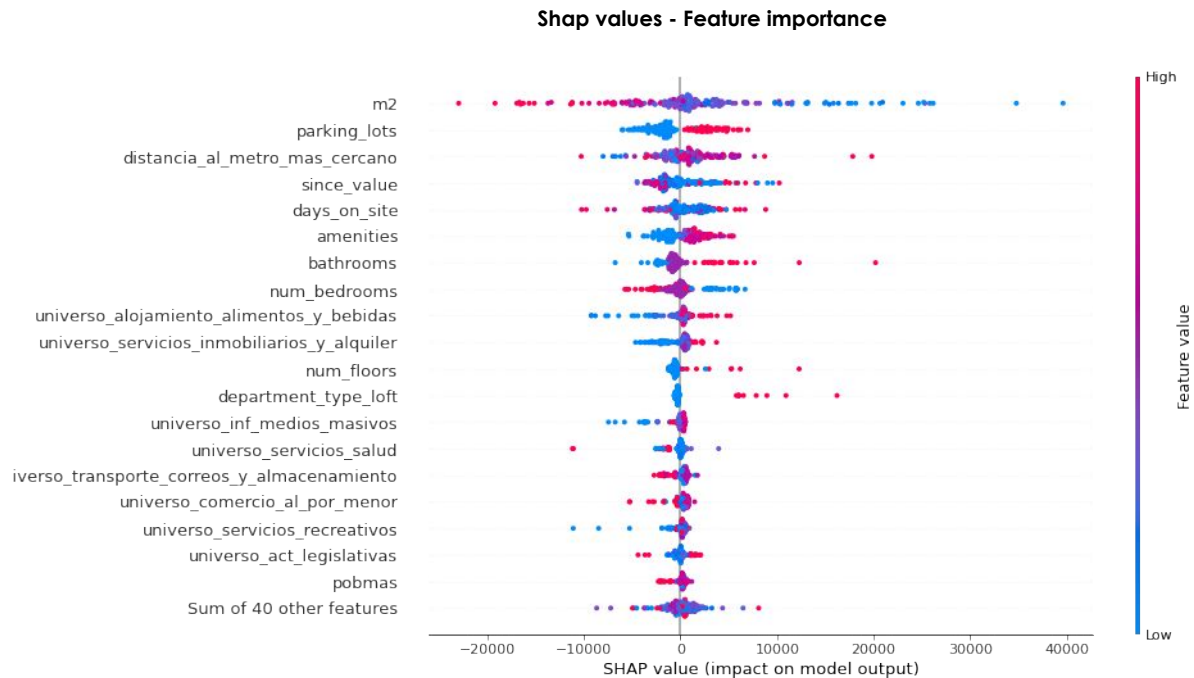
# Hallazgos

## Ranking de importancia de variables

Variable	Ranking importancia	Relación	Beta
Número de establecimientos de medios masivos	9	Positiva	914
Número de establecimientos de servicios de salud	10	Positiva	20.5
Número de servicios de transporte y almacenamiento	11	Positiva	-1,592
Número de establecimientos de comercio al por menor	12	Positiva	11.93
Número de establecimientos de servicios recreativo	13	Positiva	1,425
Número de de servicios legislativos	14	Negativa	-479
Población masculina en el AGEB	15	Positiva	-13.91

# Hallazgos

Shap values



# Conclusiones

## Influencia positiva

- Por cada metro de cercanía del departamento a un transporte público (metro), incrementa el precio del departamento en \$3.47
- Por cada estacionamiento, incrementa el precio del departamento en \$5,842
- Por cada unidad de comodidades, incrementa el precio del departamento en \$274
- Por cada baño, incrementa el precio del departamento en \$3,052
- Por cada establecimiento de alimentos y bebidas en el AGEB, incrementa el precio del departamento en \$5.58
- Por cada piso, incrementa el precio del departamento en \$2,160
- Si el departamento es de tipo loft, incrementa el precio del departamento en \$19,421

## Influencia negativa

- Por cada servicio inmobiliario y de alquiler en el AGEB, disminuye el precio del departamento en -\$532
- Por cada servicio legislativo en el AGEB, disminuye el precio del departamento en -\$479
- Por cada delito en el patrimonio en el municipio, disminuye el precio del departamento en -\$16.35
- Por cada delito contra la vida e integridad corporal en el municipio, disminuye el precio del departamento en -\$16.35

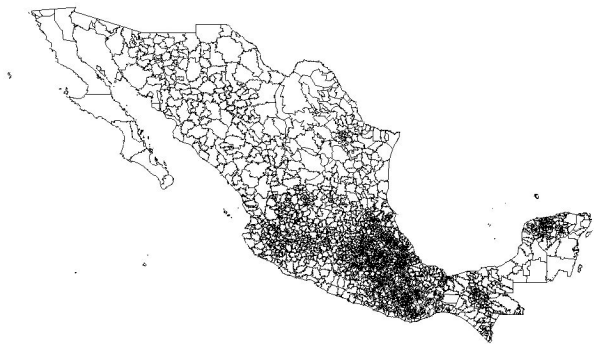
\* Estos incrementos/decrementos son en promedio y suponen que todas las demás variables permanecen fijas

\* Los incrementos/decrementos son promedio y están en pesos mexicanos

# Anexos

## Cruces de información

Para poder realizar los cruces de información se desarrolló un programa que consiste en mapear cada departamento con su llave de AGEB



(Latitud,Longitud) del departamento

Permite mapear cada coordenada a un polígono AGEB y por tanto obtener su llave

# Anexos

## Métricas de los modelos

### Modelo de regresión lineal múltiple

Métrica	Valor
R squared	29.83
Mean absolute error	13,909
Mean square error	399,041,539
Root mean squared error	19,976

### Modelo de XGBRegressor

Métrica	Valor
R squared	37.02
Mean absolute error	10,757
Mean square error	305,536,379
Root mean squared error	17,479

El XGBoost Regresor tiene un mejor performance, sin embargo el MRLM permite mayor interpretabilidad