

---

---

# Rentas Inmuebles INDIA

— ¿Cómo predecir el precio de la  
renta de inmuebles? —

---

---

Elaborado por: Gilbert Barreto

# INDICE

- 01 | Contexto y Audiencia
- 02 | Objetivos
- 03 | Hipótesis/Preguntas de Interés
- 04 | Análisis Exploratorio de Datos
- 05 | Aplicación de algoritmos de ML
- 06 | Optimización de Modelos
- 07 | Conclusiones



# Contexto y Audiencia

## Contexto

De forma general, la búsqueda de vivienda ha cambiado significativamente en los últimos años. Por razones económicas, la modernización y la ampliación de trabajos remotos se ha observado un aumento en el interés de renta de inmuebles. Esto puede significar un problema para personas con limitación presupuestaria y un reto para agentes inmobiliarios en la búsqueda de potenciales clientes.

Ideal pudiera ser, tener una herramienta que permita, en base a las características principales de un inmueble y su ubicación, estimar el precio de un inmueble.

## Esto permitiría:

1. Optimizar la búsqueda de inmuebles, otorgando una referencia o base de negociación.
2. Facilitar al vendedor a la hora de establecer un precio competitivo.

## Audiencia y Limitaciones

Compradores y vendedores del sector inmobiliario son el principal target de este estudio, específicamente aquellos que buscan inmuebles arrendados.

La muestra está centrada en 5 ciudades de la India. Bajo este contexto el estudio está limitado a las siguientes localidades:

Kolkata, Mumbai, Bangalore, Delhi, Chennai y Hyderabad.

# Preguntas de interés

## Pregunta Principal

- ¿Cómo se comporta la Renta de Inmuebles en la India?

## Preguntas secundarias

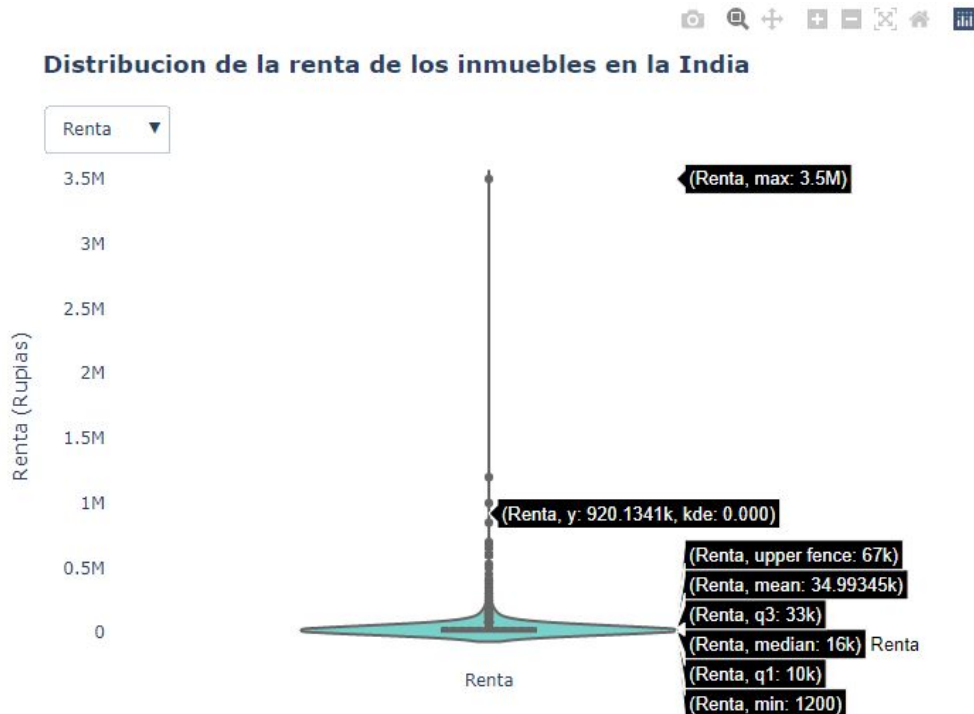
- ¿Cual es la Renta promedio de los inmuebles en la India?
- ¿Cómo son las rentas promedio por ciudad ?
- ¿Cuáles son las características de los inmuebles que influyen más sobre la renta?
- ¿Es el tamaño un factor determinante del precio?
- ¿Qué otras características pueden influir significativamente sobre el precio?

# ANÁLISIS EXPLORATORIO (EDA)

# ¿Cuál es la renta promedio de los alquileres en la India?

- La **renta de los alquileres** en la India se encuentran entre los **1.200** y **3,5 millones** de rupias.
- Existe una asimetría muy significativa de los datos ya que la mayoría de las rentas se encuentran entre los **10.000** y **33.000** rupias. Las rentas por encima de estos valores se podrían asumir como correspondientes a alquileres de lujo.
- Dada la **gran asimetría** observada, se considera la **mediana** un valor más representativo de los datos que la media. La mediana de la renta se encuentra en los **16.000 rupias**.

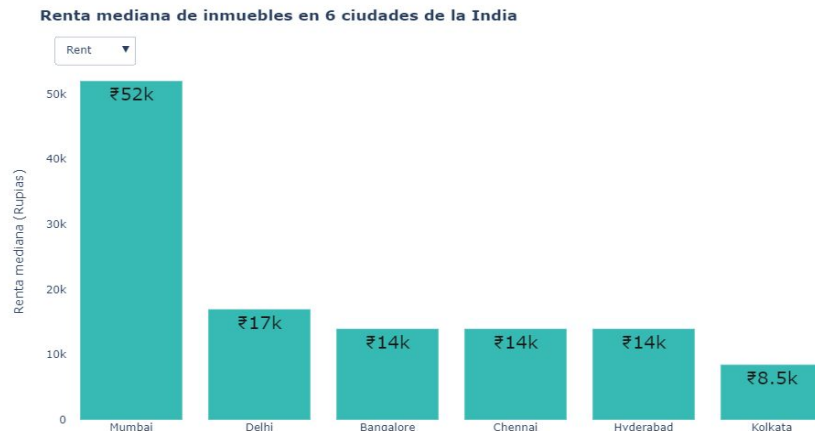
🏠 Entendiendo esta realidad. ¿Será este comportamiento común en todas las ciudades de la India?



# ¿Cómo son las rentas promedio por ciudad ?

La muestra de datos contiene información de inmuebles de 5 ciudades de la India: Mumbai, Delhi, Bangalore, Chennai, Hyderabad y Kolkata.

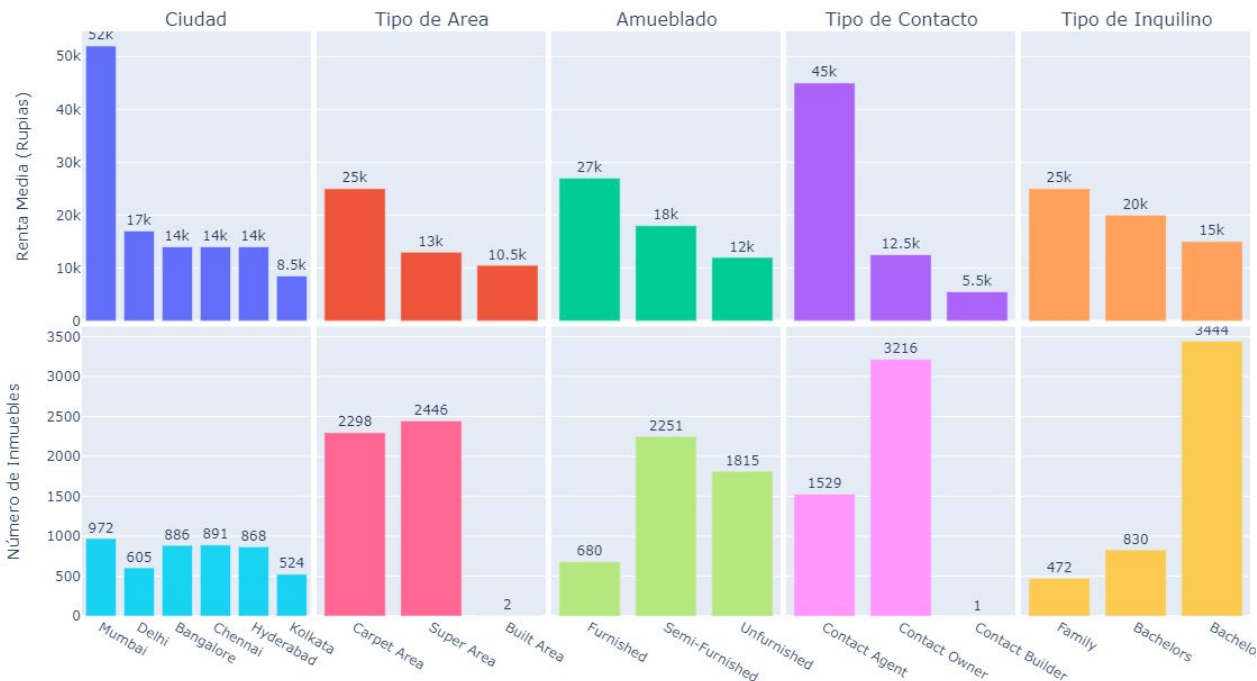
- La mediana de la renta en **Mumbai** es la más alta con una cifra de **₹52.000**, seguido de **Delhi (₹17.000)** en segundo lugar.
- **Bangalore, Chennai** y **Hyderabad** ocupan un tercer lugar con unos **₹14.000**.
- Finalizando con **Kolkata** con una renta mediana de **₹8.500**.
- **Delhi** y **Mumbai** son las únicas ciudades por encima de la mediana de la renta muestral, es decir superan el valor de los **₹16.000**.



¿Como el comportamiento de la renta agrupado en otras características?

# ¿Cuáles son las características de los inmuebles que influyen más sobre la renta?

Análisis de la Renta de Inmuebles en la India por Característica



Tal como muestra el gráfico, los factores que tienden a encarecer los precios de las rentas en la India, adicionales a la ciudad, son:

1. **Amueblado:** Mientras **más amueblado sea el inmueble** será siempre más costoso. No obstante, este tipo representa la minoría. El 85% de inmuebles de la muestra son semi-amueblados o no-amueblados.
2. **Tipo de Contacto:** Los inmuebles que establecen contacto con **agencias inmobiliarias** son significativamente más caros que aquellos que establecen otro tipo de contacto. Este tipo de inmuebles representan aproximadamente un 25% de la muestra.
3. **Tipo de Inquilino:** Aquellos alquileres, cuyos dueños **son indiferentes en su preferencia de inquilino** son los que tienden a ser más baratos.

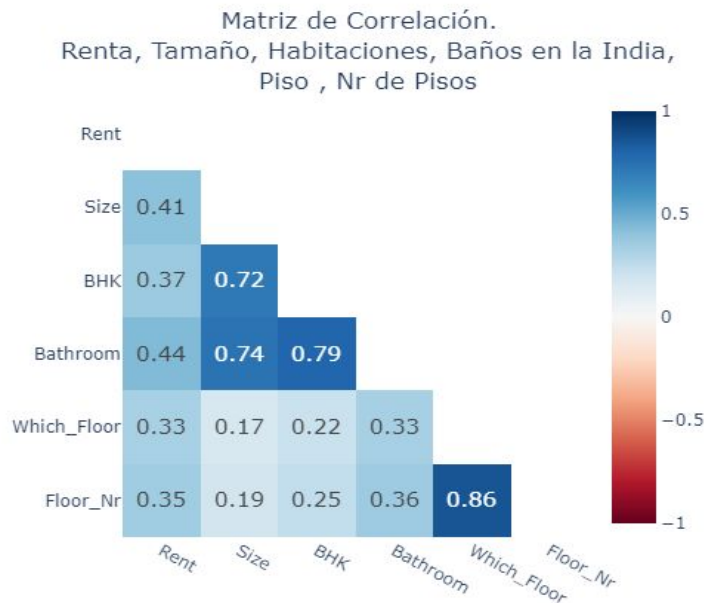
🏠 Nos hace falta analizar el tamaño y ... ¿cuáles otros?



# Es el tamaño un factor determinante del precio

Para entender cómo se relaciona las variables restantes utilizamos una matriz de correlación donde podemos observar:

- **Todas las variables** tienen una correlación débil con respecto a la renta, entre el 33% y el 44%.
- Se puede esperar un aumento del tamaño de los inmuebles en la medida que aumente el número de habitaciones o baños.
  - Por el contrario, estas variables guardan muy poca relación con el tamaño de los inmuebles y el piso del alquiler (Which\_Floor, Floor\_Nr).



# APLICACIÓN DE ALGORITMOS DE ML

# ALGORITMOS EMPLEADOS

- Regresión Lineal Múltiple (RLM)
- Regresión de Árboles de Decisión (DT)
- Regresión de Random Forest (RFR) -> **Modelo elegido (MVP)**
- Regresión de Support Vector Machine (SVM\_R).



## Resultados:

Métricas	MSE	MAE	MAPE	R2
RLM	58,900,724.7429	5,394.4828	0.3231	0.6918
DT	80,516,960.6421	6,111.1483	0.3622	0.5787
RFR (MVP)	47,377,779.9594	4,766.0439	0.2892	0.7521
SVM_R	211,553,708.9196	9,715.4446	0.5432	-0.1068

# OPTIMIZACIÓN DE MODELOS

# OPTIMIZACIÓN CON FEATURE ENGINEERING

## 1ra Optimización

- Se añaden las variables Latitud, Longitud
- Se crea una variable agrupando Latitud y Longitud en Clusters por ciudad.

## 2da Optimización

- Se eliminan las variables con menor Feature Importance



### Resultados:

Métricas	RFR(MVP)	RFR (Nuevas Variables)	RFR(Eliminando variables por Feature Importance)
<b>MSE</b>	47,377,779.9594	38,007,387.3638	38,502,020.8959
<b>MAE</b>	4,766.0439	4,156.3990	4,200.9349
<b>MAPE</b>	0.2892	0.2576	0.2613
<b>R2</b>	0.7521	0.7996	0.7970

# OPTIMIZACIÓN DE HIPERPARÁMETROS



	RFR(MVP)	RFR (Nuevas Variables)	RFR(Eliminando variables por Feature Importance)	RFR_Hiperpar ametros_1	RFR_Hiperpar ametros_2
<b>MSE</b>	47,947,699.5469	38,007,387.3638	38,502,020.8959	38,419,958.0099	38,124,934.5789
<b>MAE</b>	4,799.4751	4,156.3990	4,200.9349	4,176.9219	4,172.4314
<b>MAPE</b>	0.2909	0.2576	0.2613	0.2568	0.2586
<b>R2</b>	0.7491	0.7996	0.7970	0.7974	0.7990



Se utiliza el algoritmo Halving GridSearchCV con Cross-Validation de 5. La segunda iteración muestra mejores resultados.

# CONCLUSIONES

# CONCLUSIONES

 Se realizó el EDA con los siguientes resultados:

- La ubicación, la persona contacto y el amueblado son variables que determinan significativamente el precio de la renta.
- La renta es muy dispersa. El 10% de los datos son outliers.

 Se aplicaron modelos de ML determinando la Regresión de Random Forest como el modelo MVP.

 Se optimizó el modelo:

- Agregando y eliminando variables (Feature Engineering)
- Se aplicó el algoritmo Halving GridSearch con CrossValidation de 5.

 Las métricas del modelo final tienen un  $R^2$  de 0,79 y un MAE de 4.172.



---

---

**MUCHAS GRACIAS !**

---

---