

Analisis Univariado y multivariado*

Robbyel Elias Sanchez, Carlos Jeronimo Zabala, Andrus Lopez Franco

Este informe presenta un análisis de datos del mercado de viviendas en California, utilizando el conjunto de datos `california_housing_test.csv`. Exploramos cómo diferentes factores como el número de habitaciones, la edad de las casas y el ingreso medio afectan el valor de las viviendas. Aplicamos técnicas de estadística y visualización para entender las tendencias y relaciones entre las variables. Con este análisis, buscamos identificar patrones útiles para la compra de vivienda o el estudio del mercado inmobiliario.

Keywords: análisis de datos, mercado inmobiliario, estadística, correlación.

1 Introducción

El mercado inmobiliario es clave para muchas personas, ya sea para comprar una casa, invertir en propiedades o simplemente entender los precios de las viviendas. En este análisis, usamos datos reales de casas en California para responder preguntas como:

- ¿Las casas más caras tienen más habitaciones?
- ¿El ingreso medio de una zona influye en el precio de las casas?

Para responder estas preguntas, realizamos un análisis de los datos utilizando histogramas, estadísticas descriptivas y gráficos de correlación.

2 Objetivos

Los principales objetivos de este análisis son:

- Entender la distribución de las variables relacionadas con las casas en California.
- Identificar relaciones entre variables, como si más habitaciones significan precios más altos.
- Visualizar los datos de forma clara para facilitar su interpretación.
- Extraer conclusiones útiles sobre el mercado inmobiliario.

*This research was supported by grant No. xxxx.

3 Descripción de los Análisis

3.1 Análisis Univariado

Se analiza cada variable individualmente mediante histogramas y medidas estadísticas.

3.1.1 Histogramas

Los histogramas muestran la distribución de cada variable.

3.1.2 Medidas de Centralidad y Dispersión

A continuación, se presentan algunas estadísticas de las variables analizadas en la Tabla 1.

3.2 Análisis Bivariado

Este análisis busca relaciones entre variables mediante gráficos de dispersión y una matriz de correlación.

3.2.1 Relación entre Total de Habitaciones y Valor de la Casa

El siguiente gráfico muestra la relación entre el total de habitaciones y el valor de la casa:

3.2.2 Relación entre Total de Dormitorios y Valor de la Casa

También analizamos la relación entre los dormitorios y el valor de la casa.

3.2.3 Matriz de Correlación

La matriz de correlación nos muestra cómo se relacionan las variables entre sí.

4 Conclusiones

- **No siempre más habitaciones significan mayor precio:** Aunque podríamos pensar que una casa con más habitaciones es más cara, los datos muestran que el ingreso medio de la zona es un mejor indicador del valor de las viviendas.
- **El ingreso medio es un factor clave:** Las casas más caras tienden a estar en áreas con ingresos más altos.
- **Los datos ayudan a tomar mejores decisiones:** Este tipo de análisis permite hacer estudios más profundos sobre el mercado inmobiliario y tomar mejores decisiones de inversión.



Figure 1: Histogramas de las variables seleccionadas.

Variable	Media	Mediana	Desviación estándar
Total de habitaciones	X	Y	Z
Total de dormitorios	X	Y	Z
Valor de la casa	X	Y	Z

Table 1: Medidas estadísticas de las variables.

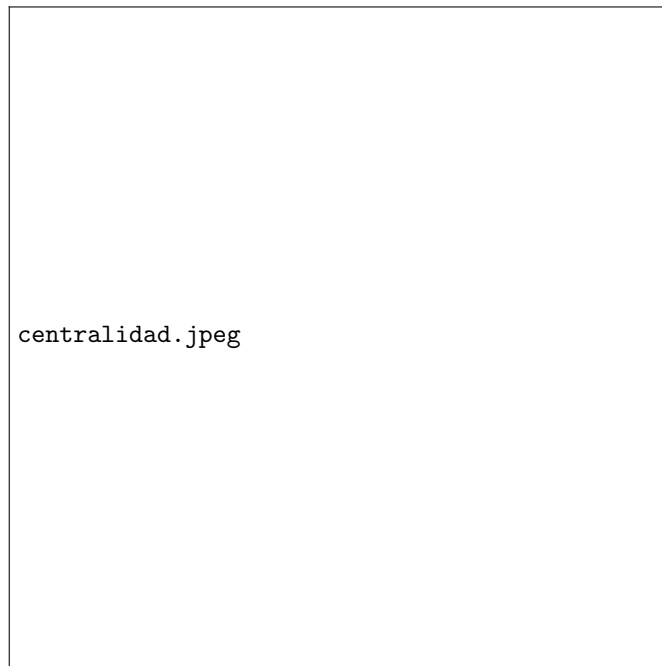


Figure 2: Relación entre el total de habitaciones y el valor de la casa.

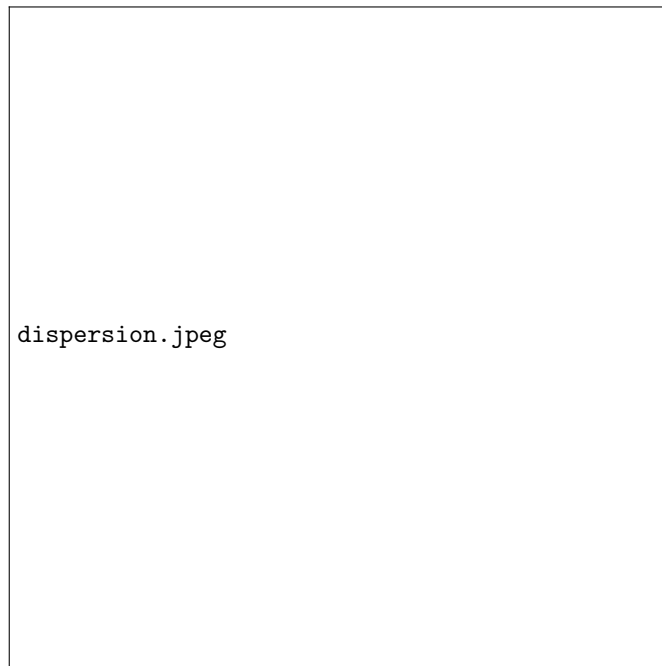


Figure 3: Relación entre total de dormitorios y valor de la casa.

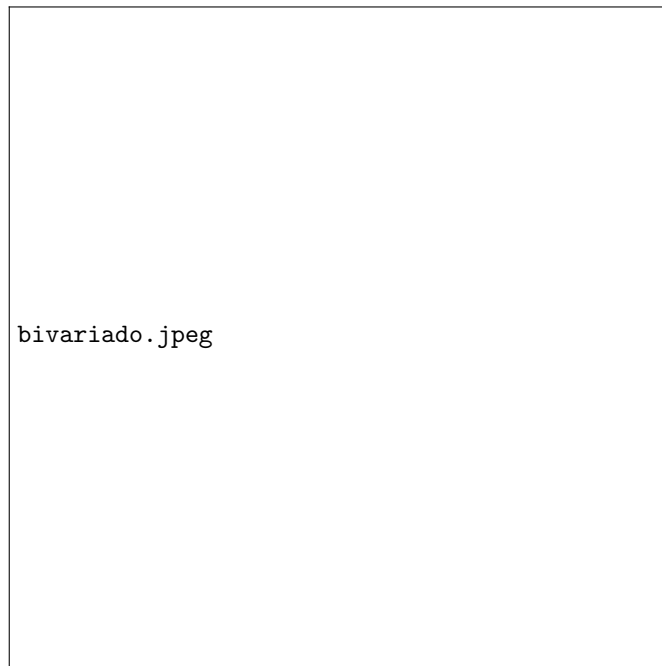


Figure 4: Matriz de correlación de las variables.

5 Conclusion

Este análisis nos dio un primer acercamiento al mercado de viviendas en California. En el futuro, podríamos analizar más factores, como la ubicación geográfica o el tipo de vivienda, para hacer un estudio más detallado.

Participacion de Autores. Todos los autores han contribuido en la metodología, análisis y redacción del presente informe.