

linear_regression

ManuelRojas

2024-05-05

Cargamos los datos, eliminamos valores perdidos y detectamos que la columna ocean_proximity es categorica y la convertimos en binaria con dummies menos 1 variable para evitar los problemas de colinealidad.

```
HouseDF <- read.csv("C:/Users/Manuel/Desktop/42/IA/housing.csv", sep =
",", header = TRUE)
HouseDF <- HouseDF[complete.cases(HouseDF), ]
dummies <- model.matrix(~ ocean_proximity - 1, data = HouseDF)
HouseDF <- cbind(HouseDF, dummies)
```

Creamos el modelo de regresión para comprobar los supuestos. Aparecen NA probablemente a que los nombres de las variables dummies contienen espacios o caracteres especiales.

```
model <- lm(median_house_value ~ ., data = HouseDF)
summary(model)

##
## Call:
## lm(formula = median_house_value ~ ., data = HouseDF)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -556980  -42683  -10497   28765   779052
##
## Coefficients: (5 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -2.270e+06  8.801e+04  -25.791  < 2e-16 ***
## longitude     -2.681e+04  1.020e+03  -26.296  < 2e-16 ***
## latitude      -2.548e+04  1.005e+03  -25.363  < 2e-16 ***
## housing_median_age  1.073e+03  4.389e+01   24.439  < 2e-16 ***
## total_rooms    -6.193e+00  7.915e-01   -7.825  5.32e-15 ***
## total_bedrooms  1.006e+02  6.869e+00   14.640  < 2e-16 ***
## population     -3.797e+01  1.076e+00  -35.282  < 2e-16 ***
## households      4.962e+01  7.451e+00    6.659  2.83e-11 ***
## median_income   3.926e+04  3.380e+02  116.151  < 2e-16 ***
## ocean_proximityINLAND -3.928e+04  1.744e+03  -22.522  < 2e-16 ***
## ocean_proximityISLAND  1.529e+05  3.074e+04    4.974  6.62e-07 ***
## ocean_proximityNEAR BAY -3.954e+03  1.913e+03   -2.067  0.03879 *
## ocean_proximityNEAR OCEAN  4.278e+03  1.570e+03    2.726  0.00642 **
## `ocean_proximity<1H OCEAN`      NA         NA      NA      NA
## ocean_proximityINLAND            NA         NA      NA      NA
```

```
## ocean_proximityISLAND      NA      NA      NA      NA
## `ocean_proximityNEAR BAY`   NA      NA      NA      NA
## `ocean_proximityNEAR OCEAN` NA      NA      NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 68660 on 20420 degrees of freedom
## Multiple R-squared:  0.6465, Adjusted R-squared:  0.6463
## F-statistic: 3112 on 12 and 20420 DF, p-value: < 2.2e-16
```

Vamos analizar los 4 supuestos Linealidad, normalidad, homocedasticidad e independencia.

Linealidad - Debe tender a 0 y es practicamente 0. Podemos comprobar que tenemos linealidad

```
mean(model$residuals)

## [1] 1.5954e-11
```

Normalidad Dado que el valor p es extremadamente pequeño (menor que cualquier nivel de significancia comúnmente utilizado, como 0.05), tenemos suficiente evidencia para rechazar la hipótesis nula de que los datos provienen de una distribución normal.

```
tamano_muestra <- 500 indices_muestra <- sample(1:length(sresid),
tamano_muestra) muestra <- sresid[indices_muestra]
shapiro.test(muestra)
```

Homocedasticidad

```
library(car)

## Loading required package: carData

breusch_pagan_test <- ncvTest(model)
print(breusch_pagan_test)

## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 1044.492, Df = 1, p = < 2.22e-16
```

El test de Breusch-Pagan muestra una varianza no constante en los residuos del modelo, lo que sugiere heterocedasticidad

```
model <- lm(median_house_value ~ ., data = HouseDF)
summary(model)

##
## Call:
## lm(formula = median_house_value ~ ., data = HouseDF)
##
## Residuals:
```

```
##      Min      1Q  Median      3Q      Max
## -556980 -42683 -10497  28765  779052
##
## Coefficients: (5 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -2.270e+06  8.801e+04 -25.791 < 2e-16 ***
## longitude      -2.681e+04  1.020e+03 -26.296 < 2e-16 ***
## latitude       -2.548e+04  1.005e+03 -25.363 < 2e-16 ***
## housing_median_age  1.073e+03  4.389e+01  24.439 < 2e-16 ***
## total_rooms     -6.193e+00  7.915e-01  -7.825 5.32e-15 ***
## total_bedrooms  1.006e+02  6.869e+00  14.640 < 2e-16 ***
## population     -3.797e+01  1.076e+00 -35.282 < 2e-16 ***
## households      4.962e+01  7.451e+00   6.659 2.83e-11 ***
## median_income   3.926e+04  3.380e+02 116.151 < 2e-16 ***
## ocean_proximityINLAND -3.928e+04  1.744e+03 -22.522 < 2e-16 ***
## ocean_proximityISLAND  1.529e+05  3.074e+04   4.974 6.62e-07 ***
## ocean_proximityNEAR BAY -3.954e+03  1.913e+03  -2.067 0.03879 *
## ocean_proximityNEAR OCEAN  4.278e+03  1.570e+03   2.726 0.00642 **
## `ocean_proximity<1H OCEAN`      NA         NA      NA      NA
## ocean_proximityINLAND            NA         NA      NA      NA
## ocean_proximityISLAND            NA         NA      NA      NA
## `ocean_proximityNEAR BAY`        NA         NA      NA      NA
## `ocean_proximityNEAR OCEAN`      NA         NA      NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 68660 on 20420 degrees of freedom
## Multiple R-squared:  0.6465, Adjusted R-squared:  0.6463
## F-statistic: 3112 on 12 and 20420 DF, p-value: < 2.2e-16

r_squared <- summary(model)$r.squared

cat("El coeficiente de determinación (R^2) es:", r_squared, "\n")

## El coeficiente de determinación (R^2) es: 0.6464638

correlation_matrix <- cor(HouseDF[, -which(names(HouseDF) ==
"ocean_proximity")])

print(correlation_matrix)

##              longitude      latitude housing_median_age
## longitude      1.000000000 -0.92461611 -0.10935655
## latitude      -0.924616113  1.000000000  0.01189907
## housing_median_age -0.109356549  0.01189907  1.00000000
## total_rooms      0.045480167 -0.03666681 -0.36062830
## total_bedrooms   0.069608022 -0.06698283 -0.32045104
## population       0.100270301 -0.10899734 -0.29578730
## households       0.056512772 -0.07177419 -0.30276797
## median_income    -0.015550150 -0.07962632 -0.11827772
## median_house_value -0.045398219 -0.14463821  0.10643205
```

## ocean_proximity<1H OCEAN	0.320831061	-0.44692798	0.04555326
## ocean_proximityINLAND	-0.055337446	0.35108357	-0.23696771
## ocean_proximityISLAND	0.009500707	-0.01666228	0.01710531
## ocean_proximityNEAR BAY	-0.474713712	0.35878451	0.25614946
## ocean_proximityNEAR OCEAN	0.046184552	-0.16134231	0.02079651
##	total_rooms	total_bedrooms	population
households			
## longitude	0.045480167	0.0696080218	0.100270301
0.056512772			
## latitude	-0.036666806	-0.0669828347	-0.108997344 -
0.071774185			
## housing_median_age	-0.360628300	-0.3204510418	-0.295787297 -
0.302767973			
## total_rooms	1.000000000	0.9303795047	0.857281251
0.918991534			
## total_bedrooms	0.930379505	1.0000000000	0.877746743
0.979728271			
## population	0.857281251	0.8777467432	1.000000000
0.907185900			
## households	0.918991534	0.9797282708	0.907185900
1.000000000			
## median_income	0.197881519	-0.0077228495	0.005086624
0.013433892			
## median_house_value	0.133294135	0.0496861802	-0.025299732
0.064893549			
## ocean_proximity<1H OCEAN	-0.003776840	0.0183142338	0.073449507
0.041883141			
## ocean_proximityINLAND	0.026477500	-0.0064628874	-0.019601810 -
0.038264635			
## ocean_proximityISLAND	-0.007602616	-0.0043614668	-0.010450530 -
0.009119262			
## ocean_proximityNEAR BAY	-0.023647215	-0.0198734892	-0.061479653 -
0.011279544			
## ocean_proximityNEAR OCEAN	-0.008674313	0.0006793635	-0.023498264
0.002105848			
##	median_income	median_house_value	
## longitude	-0.015550150	-0.04539822	
## latitude	-0.079626319	-0.14463821	
## housing_median_age	-0.118277723	0.10643205	
## total_rooms	0.197881519	0.13329413	
## total_bedrooms	-0.007722850	0.04968618	
## population	0.005086624	-0.02529973	
## households	0.013433892	0.06489355	
## median_income	1.000000000	0.68835548	
## median_house_value	0.688355475	1.00000000	
## ocean_proximity<1H OCEAN	0.168715326	0.25761447	
## ocean_proximityINLAND	-0.237535952	-0.48478695	
## ocean_proximityISLAND	-0.009281442	0.02352509	
## ocean_proximityNEAR BAY	0.056676699	0.16052561	
## ocean_proximityNEAR OCEAN	0.027351350	0.14037756	

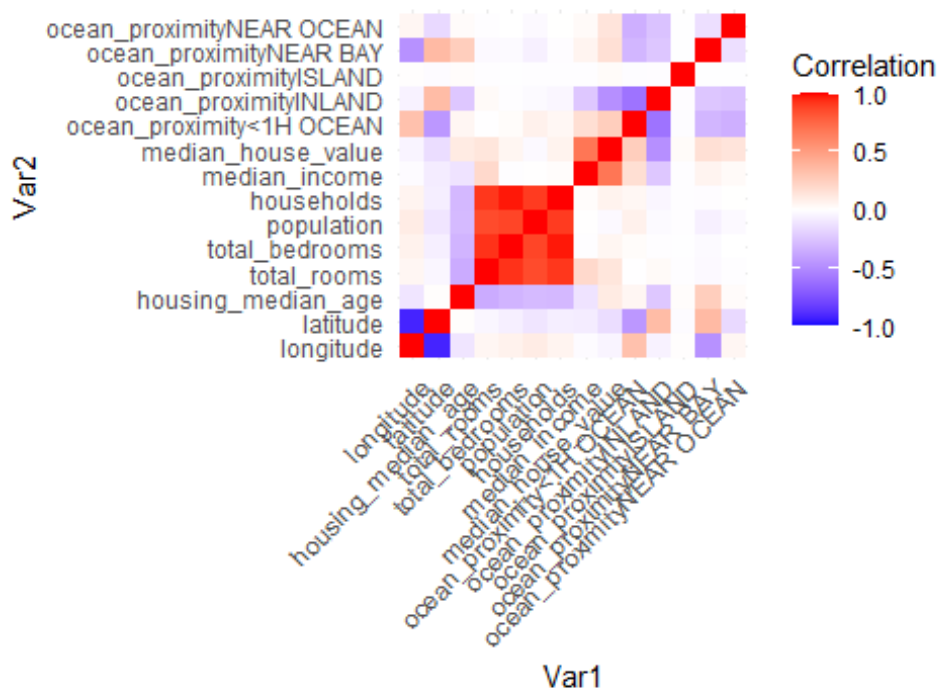
##	ocean_proximity<1H OCEAN	
ocean_proximityINLAND		
## longitude	0.32083106	-
0.055337446		
## latitude	-0.44692798	
0.351083573		
## housing_median_age	0.04555326	-
0.236967709		
## total_rooms	-0.00377684	
0.026477500		
## total_bedrooms	0.01831423	-
0.006462887		
## population	0.07344951	-
0.019601810		
## households	0.04188314	-
0.038264635		
## median_income	0.16871533	-
0.237535952		
## median_house_value	0.25761447	-
0.484786953		
## ocean_proximity<1H OCEAN	1.00000000	-
0.607778068		
## ocean_proximityINLAND	-0.60777807	
1.000000000		
## ocean_proximityISLAND	-0.01392768	-
0.010680964		
## ocean_proximityNEAR BAY	-0.31472106	-
0.241355664		
## ocean_proximityNEAR OCEAN	-0.34201752	-
0.262288983		
##	ocean_proximityISLAND ocean_proximityNEAR	
BAY		
## longitude	0.009500707	-
0.474713712		
## latitude	-0.016662279	
0.358784505		
## housing_median_age	0.017105315	
0.256149460		
## total_rooms	-0.007602616	-
0.023647215		
## total_bedrooms	-0.004361467	-
0.019873489		
## population	-0.010450530	-
0.061479653		
## households	-0.009119262	-
0.011279544		
## median_income	-0.009281442	
0.056676699		
## median_house_value	0.023525092	
0.160525612		

```

## ocean_proximity<1H OCEAN          -0.013927679      -
0.314721059
## ocean_proximityINLAND              -0.010680964      -
0.241355664
## ocean_proximityISLAND              1.000000000      -
0.005530842
## ocean_proximityNEAR BAY            -0.005530842
1.000000000
## ocean_proximityNEAR OCEAN          -0.006010544      -
0.135819094
##                                ocean_proximityNEAR OCEAN
## longitude                          0.0461845517
## latitude                          -0.1613423071
## housing_median_age                 0.0207965056
## total_rooms                       -0.0086743127
## total_bedrooms                    0.0006793635
## population                        -0.0234982639
## households                         0.0021058479
## median_income                      0.0273513502
## median_house_value                 0.1403775589
## ocean_proximity<1H OCEAN          -0.3420175233
## ocean_proximityINLAND              -0.2622889832
## ocean_proximityISLAND              -0.0060105438
## ocean_proximityNEAR BAY            -0.1358190938
## ocean_proximityNEAR OCEAN          1.0000000000

library(ggplot2)
library(reshape2)
ggplot(data = melt(correlation_matrix), aes(x = Var1, y = Var2, fill =
value)) +
  geom_tile() +
  scale_fill_gradient2(low = "blue", high = "red", mid = "white",
midpoint = 0,
                        limit = c(-1, 1), space = "Lab",
name="Correlation") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, vjust = 1, size = 10,
hjust = 1)) +
  coord_fixed()

```



Mapa de calor

de correlaciones entre variables: Este gráfico muestra la fuerza y la dirección de la relación lineal entre todas las parejas de variables en tus datos. Cada celda en el mapa de calor representa el coeficiente de correlación entre dos variables. Los colores más claros indican una correlación positiva (más cercana a 1), mientras que los colores más oscuros indican una correlación negativa (más cercana a -1).

longitud y latitud: Tienen una fuerte correlación negativa de aproximadamente -0.92, lo que indica que las ubicaciones más al oeste tienen latitudes más altas y viceversa.

housing_median_age y median_house_value: Muestran una correlación positiva modesta de alrededor de 0.11, sugiriendo que los precios de las viviendas tienden a ser ligeramente más altos en áreas con mayor edad media de vivienda.

total_rooms, total_bedrooms, population y households: Estas variables están altamente correlacionadas, lo que sugiere que áreas con más habitaciones probablemente también tengan más dormitorios, población y hogares.

median_income y median_house_value: Existe una correlación positiva fuerte de aproximadamente 0.69, indicando que los precios de las viviendas tienden a ser más altos en áreas con ingresos medios más altos.

Eliminamos los NA

```
HouseDFnoNA <- na.omit(HouseDF)

model <- lm(median_house_value ~ ., data = HouseDF)
summary(model)

##
## Call:
```

```
## lm(formula = median_house_value ~ ., data = HouseDF)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -556980  -42683  -10497   28765  779052
##
## Coefficients: (5 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -2.270e+06  8.801e+04 -25.791 < 2e-16 ***
## longitude     -2.681e+04  1.020e+03 -26.296 < 2e-16 ***
## latitude      -2.548e+04  1.005e+03 -25.363 < 2e-16 ***
## housing_median_age  1.073e+03  4.389e+01  24.439 < 2e-16 ***
## total_rooms    -6.193e+00  7.915e-01  -7.825 5.32e-15 ***
## total_bedrooms  1.006e+02  6.869e+00  14.640 < 2e-16 ***
## population     -3.797e+01  1.076e+00 -35.282 < 2e-16 ***
## households      4.962e+01  7.451e+00   6.659 2.83e-11 ***
## median_income   3.926e+04  3.380e+02  116.151 < 2e-16 ***
## ocean_proximityINLAND -3.928e+04  1.744e+03 -22.522 < 2e-16 ***
## ocean_proximityISLAND  1.529e+05  3.074e+04   4.974 6.62e-07 ***
## ocean_proximityNEAR BAY -3.954e+03  1.913e+03  -2.067 0.03879 *
## ocean_proximityNEAR OCEAN  4.278e+03  1.570e+03   2.726 0.00642 **
## `ocean_proximity<1H OCEAN`      NA         NA      NA      NA
## ocean_proximityINLAND            NA         NA      NA      NA
## ocean_proximityISLAND            NA         NA      NA      NA
## `ocean_proximityNEAR BAY`        NA         NA      NA      NA
## `ocean_proximityNEAR OCEAN`      NA         NA      NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 68660 on 20420 degrees of freedom
## Multiple R-squared:  0.6465, Adjusted R-squared:  0.6463
## F-statistic: 3112 on 12 and 20420 DF, p-value: < 2.2e-16
```

Aún eliminando los NA no hay diferencia, continuamos con un 64%

Normalizamos el

```
library(caret)

## Loading required package: lattice

numeric_cols <- sapply(HouseDF, is.numeric)
HouseDF_normalized <- HouseDF
HouseDF_normalized[, numeric_cols] <- scale(HouseDF_normalized[,
numeric_cols])
summary(HouseDF_normalized)

##      longitude      latitude  housing_median_age  total_rooms
## Min.      :-2.3854   Min.      :-1.4479   Min.      :-2.19453   Min.      :-
1.2056
## 1st Qu.: -1.1127   1st Qu.: -0.7973   1st Qu.: -0.84445   1st Qu.: -
```



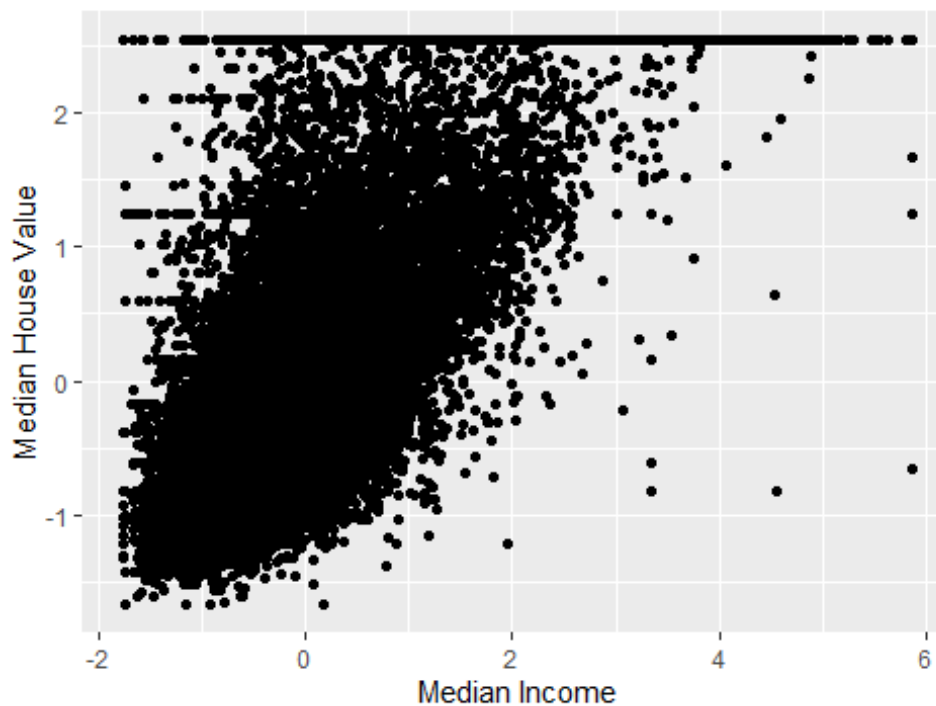
```

0.5430
## Median : 0.5394 Median :-0.6428 Median : 0.02914 Median :-
0.2332
## Mean : 0.0000 Mean : 0.0000 Mean : 0.00000 Mean :
0.0000
## 3rd Qu.: 0.7790 3rd Qu.: 0.9768 3rd Qu.: 0.66447 3rd Qu.:
0.2318
## Max. : 2.6256 Max. : 2.9568 Max. : 1.85572 Max.
:16.7867
## total_bedrooms population households median_income
## Min. :-1.2741 Min. :-1.2548 Min. :-1.3038 Min. :-1.7750
## 1st Qu.: -0.5740 1st Qu.: -0.5630 1st Qu.: -0.5740 1st Qu.: -0.6884
## Median : -0.2441 Median : -0.2285 Median : -0.2366 Median : -0.1762
## Mean : 0.0000 Mean : 0.0000 Mean : 0.0000 Mean : 0.0000
## 3rd Qu.: 0.2590 3rd Qu.: 0.2621 3rd Qu.: 0.2735 3rd Qu.: 0.4596
## Max. :14.0184 Max. :30.2301 Max. :14.6026 Max. : 5.8595
## median_house_value ocean_proximity ocean_proximity<1H OCEAN
## Min. :-1.6621 Length:20433 Min. :-0.8902
## 1st Qu.: -0.7568 Class :character 1st Qu.: -0.8902
## Median : -0.2353 Mode :character Median : -0.8902
## Mean : 0.0000 Mean : 0.0000
## 3rd Qu.: 0.5010 3rd Qu.: 1.1233
## Max. : 2.5394 Max. : 1.1233
## ocean_proximityINLAND ocean_proximityISLAND ocean_proximityNEAR BAY
## Min. :-0.6827 Min. :-0.01564 Min. :-0.3535
## 1st Qu.: -0.6827 1st Qu.: -0.01564 1st Qu.: -0.3535
## Median : -0.6827 Median : -0.01564 Median : -0.3535
## Mean : 0.0000 Mean : 0.00000 Mean : 0.0000
## 3rd Qu.: 1.4647 3rd Qu.: -0.01564 3rd Qu.: -0.3535
## Max. : 1.4647 Max. :63.91713 Max. : 2.8286
## ocean_proximityNEAR OCEAN
## Min. :-0.3842
## 1st Qu.: -0.3842
## Median : -0.3842
## Mean : 0.0000
## 3rd Qu.: -0.3842
## Max. : 2.6028

library(ggplot2)
ggplot(data = HouseDF_normalized, aes(x = median_income, y =
median_house_value)) +
  geom_point() +
  labs(x = "Median Income", y = "Median House Value") +
  ggtitle("Scatter Plot of Median Income vs. Median House Value")

```

Scatter Plot of Median Income vs. Median House Value



```
model <- lm(median_house_value ~ ., data = HouseDF_normalized)
summary(model)
```

```
##
## Call:
## lm(formula = median_house_value ~ ., data = HouseDF_normalized)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.8250 -0.3698 -0.0909  0.2492  6.7488
##
## Coefficients: (5 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.106906   0.007280  14.685 < 2e-16 ***
## longitude     -0.465384   0.017698 -26.296 < 2e-16 ***
## latitude      -0.471594   0.018594 -25.363 < 2e-16 ***
## housing_median_age  0.116991  0.004787  24.439 < 2e-16 ***
## total_rooms    -0.117242  0.014983  -7.825 5.32e-15 ***
## total_bedrooms  0.367070  0.025073  14.640 < 2e-16 ***
## population     -0.372735  0.010564 -35.282 < 2e-16 ***
## households      0.164322  0.024677   6.659 2.83e-11 ***
## median_income   0.645947  0.005561 116.151 < 2e-16 ***
## ocean_proximityINLAND -0.340313  0.015110 -22.522 < 2e-16 ***
## ocean_proximityISLAND  1.324564  0.266312   4.974 6.62e-07 ***
## ocean_proximityNEAR BAY -0.034253  0.016575  -2.067 0.03879 *
## ocean_proximityNEAR OCEAN  0.037061  0.013597   2.726 0.00642 **
## `ocean_proximity<1H OCEAN`      NA         NA      NA      NA
```

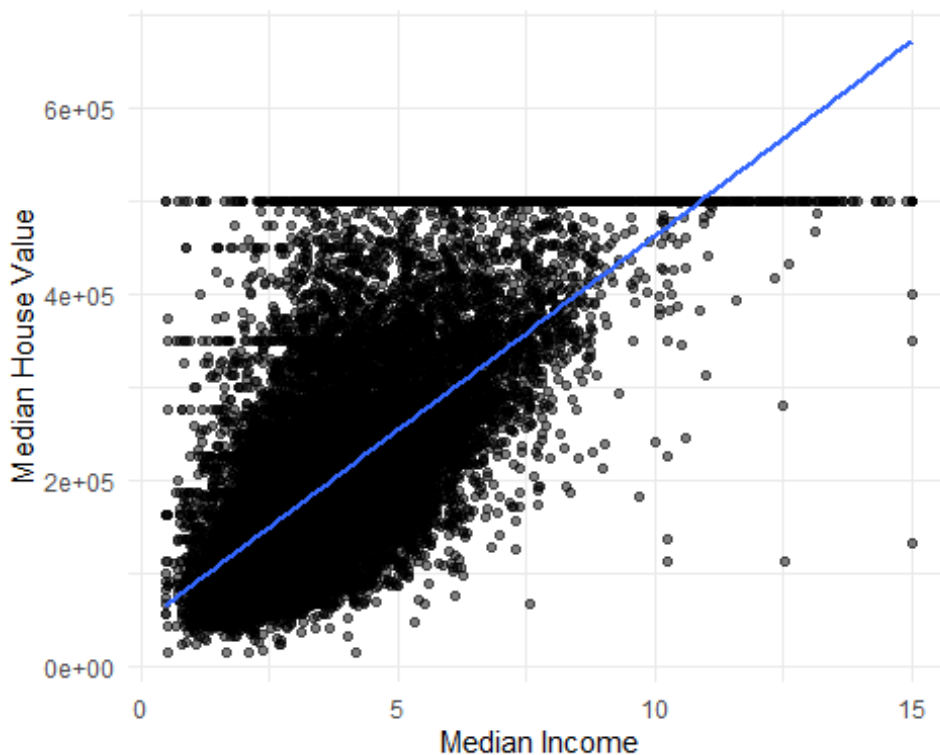
```
## ocean_proximityINLAND      NA      NA      NA      NA
## ocean_proximityISLAND      NA      NA      NA      NA
## `ocean_proximityNEAR BAY`   NA      NA      NA      NA
## `ocean_proximityNEAR OCEAN` NA      NA      NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5948 on 20420 degrees of freedom
## Multiple R-squared:  0.6465, Adjusted R-squared:  0.6463
## F-statistic: 3112 on 12 and 20420 DF, p-value: < 2.2e-16
```

Normalizando el modelo seguimos con el mismo 64,65%.

Visualización de datos

```
# Gráfico de dispersión de ingresos medios vs precios de viviendas
ggplot(data = HouseDF, aes(x = median_income, y = median_house_value)) +
  geom_point(alpha = 0.5) +
  geom_smooth(method = "lm", se = FALSE) +
  labs(x = "Median Income", y = "Median House Value") +
  theme_minimal()

## `geom_smooth()` using formula = 'y ~ x'
```

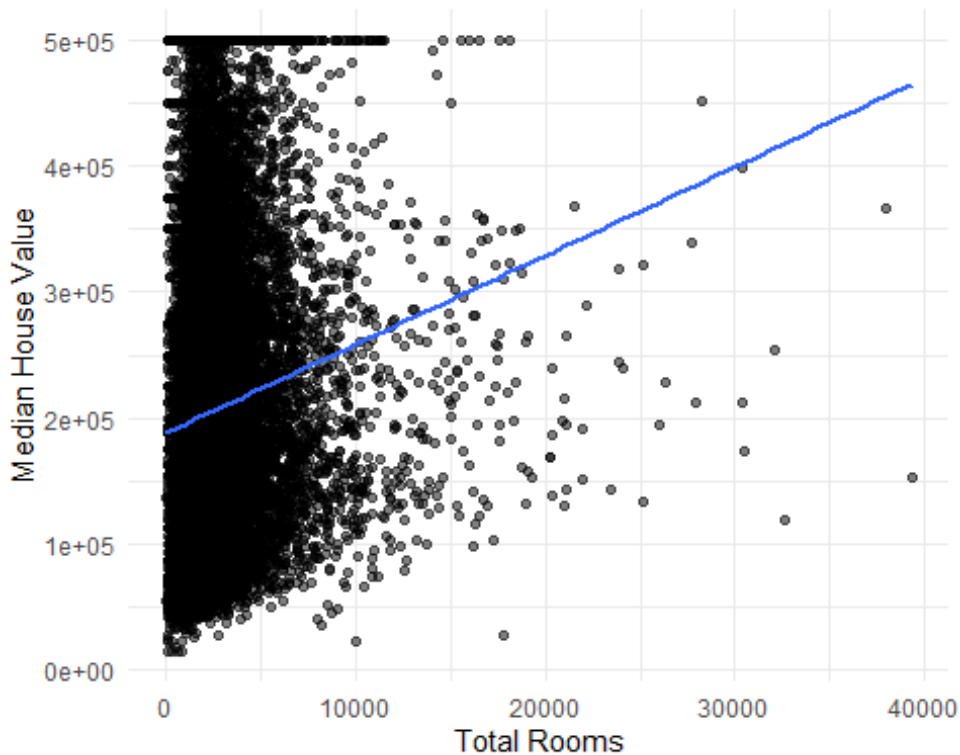


Gráficos de dispersión entre precios de viviendas y características significativas: Estos gráficos muestran la relación entre los precios de las viviendas y las características más significativas, como los ingresos medios y el número de habitaciones. Cada punto

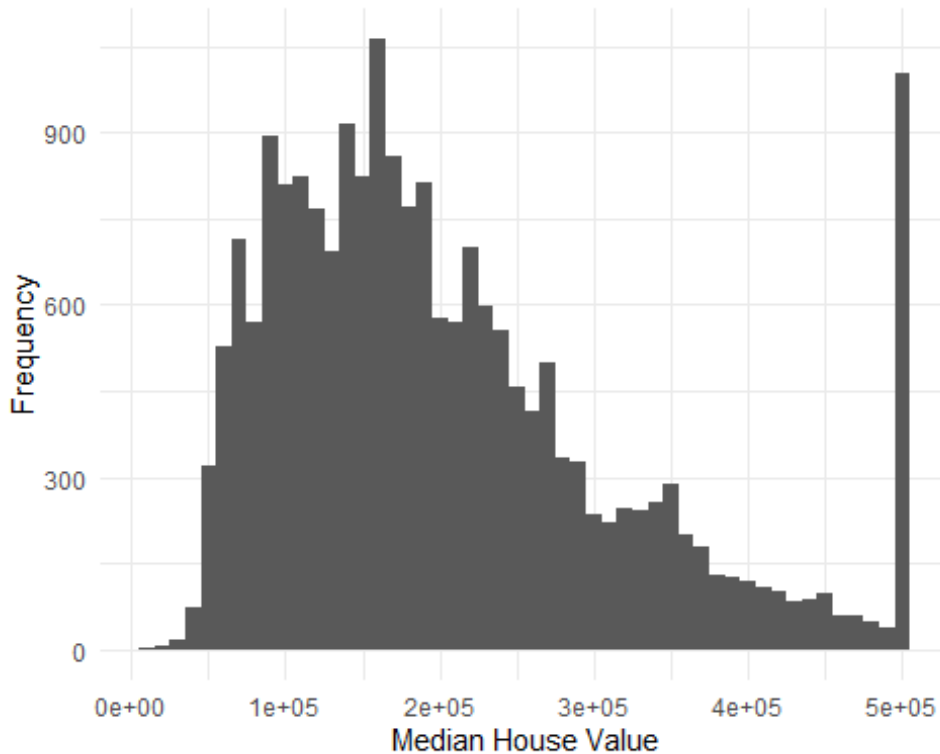
en el gráfico representa una observación y la posición en el eje x representa una característica significativa (por ejemplo, ingresos medios) y la posición en el eje y representa el precio de la vivienda

```
# Gráfico de dispersión de número de habitaciones vs precios de viviendas
ggplot(data = HouseDF, aes(x = total_rooms, y = median_house_value)) +
  geom_point(alpha = 0.5) +
  geom_smooth(method = "lm", se = FALSE) +
  labs(x = "Total Rooms", y = "Median House Value") +
  theme_minimal()

## `geom_smooth()` using formula = 'y ~ x'
```



```
# Histograma de precios de viviendas
ggplot(data = HouseDF, aes(x = median_house_value)) +
  geom_histogram(binwidth = 10000) +
  labs(x = "Median House Value", y = "Frequency") +
  theme_minimal()
```



Histograma de

la distribución de precios de viviendas: Este gráfico muestra la distribución de los precios de las viviendas en forma de un histograma. En el eje x se encuentran los precios de las viviendas, divididos en intervalos (bins), y en el eje y se muestra la frecuencia o la cantidad de observaciones en cada intervalo.

```
residuals <- residuals(model)

# Gráfico de residuales vs valores ajustados
ggplot(data = data.frame(residuals = residuals, fitted = fitted(model)),
  aes(x = fitted, y = residuals)) +
  geom_point() +
  geom_hline(yintercept = 0, linetype = "dashed", color = "red") +
  labs(x = "Fitted Values", y = "Residuals") +
  theme_minimal()
```

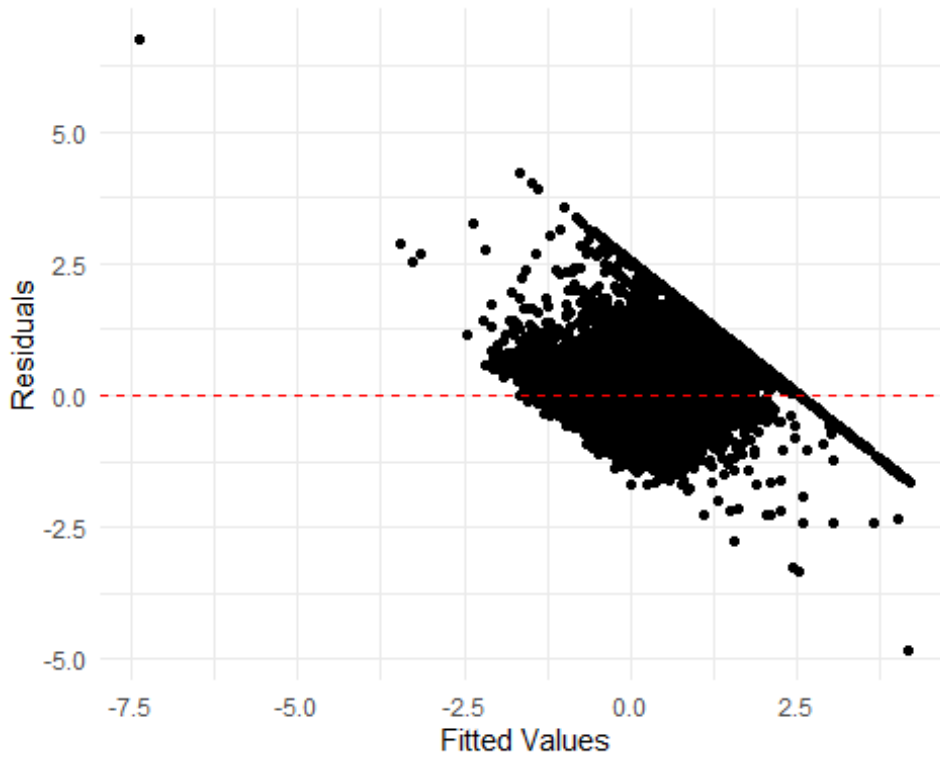


Gráfico de

residuales vs valores ajustados: Muestra los residuales en el eje y y los valores ajustados por el modelo en el eje x. Te permite evaluar si los residuales tienen una relación sistemática con los valores ajustados

```
# Gráfico de residuales vs valores observados
ggplot(data = data.frame(residuals = residuals, observed =
HouseDF$median_house_value), aes(x = observed, y = residuals)) +
  geom_point() +
  geom_hline(yintercept = 0, linetype = "dashed", color = "red") +
  labs(x = "Observed Values", y = "Residuals") +
  theme_minimal()
```

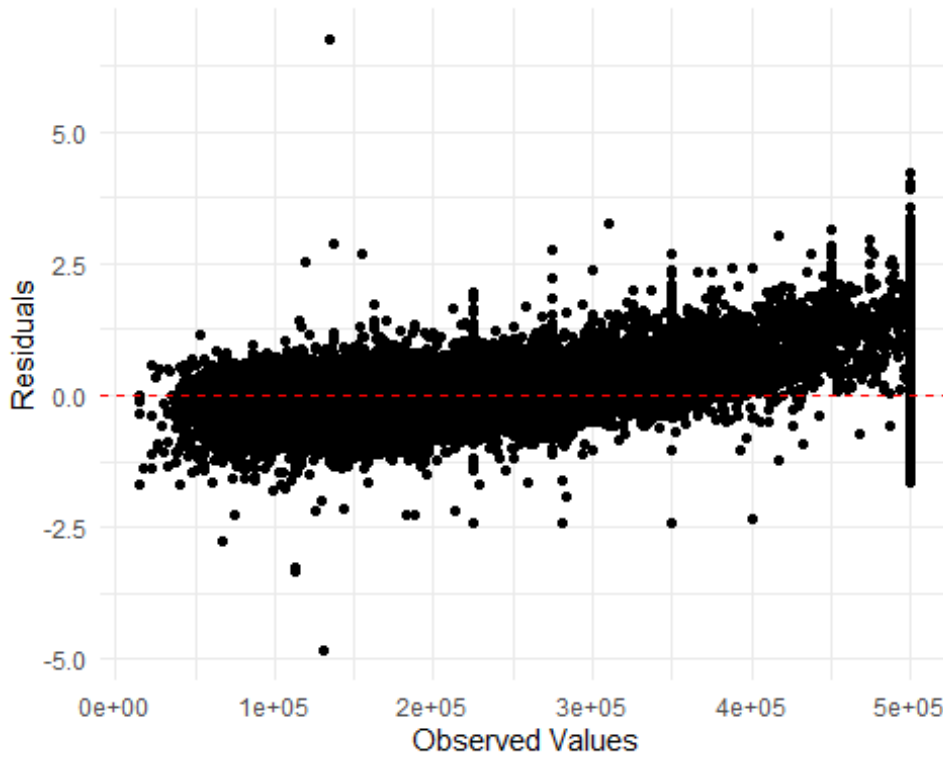


Gráfico de
residuales vs variables predictoras: Muestra los residuales en el eje y y una variable
predictora en el eje x. Puedes utilizar esto para evaluar si los residuales tienen una
relación sistemática con las variables predictoras y detectar posibles problemas de
especificación del modelo.