regression_analysis_prageeth

my author

2025-02-25

```
#{r setup, include=FALSE}
#knitr::opts_chunk$set(echo = FALSE) #
```

R Markdown

This is an R Markdown presentation. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document.

Slide with Bullets

- Bullet 1
- Bullet 2
- Bullet 3

Slide with R Output

```
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2.00
```

Start my work.

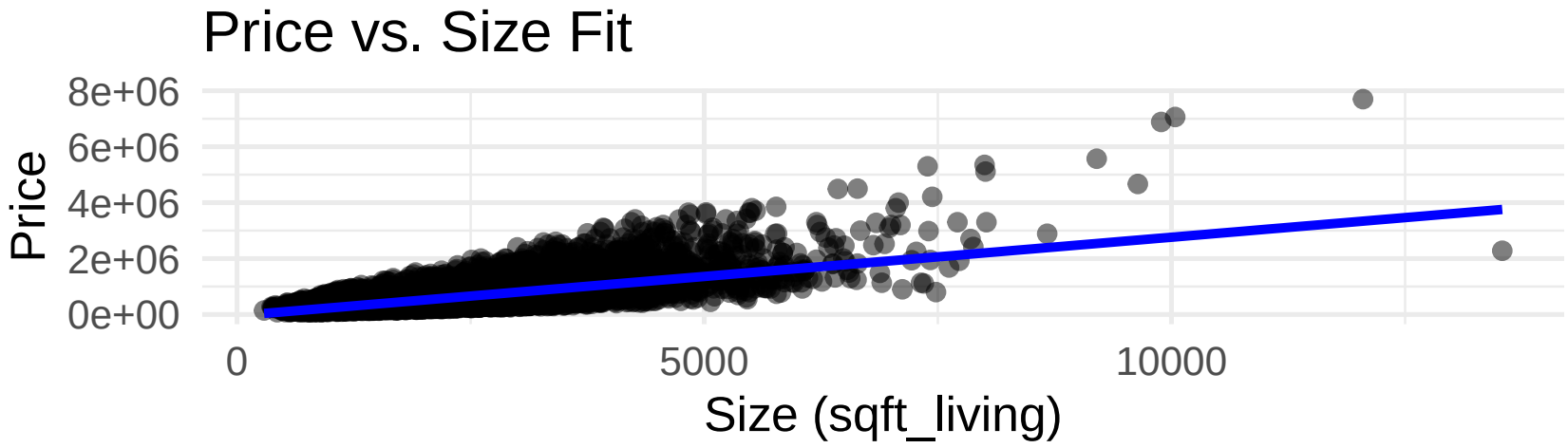
```
house_df <- read.csv("C:\\Users\\Prageeth\\Source\\MSC\\ser
features <- c("bedrooms", "bathrooms", "sqft_living", "sqft
              "condition", "grade", "sqft_above", "sqft_bas
target <- "price"
```

Simple Linear Regression

```
simple_lm <- lm ( price ~ sqft_living ,
                  data = house_df )
ggplot(house_df, aes(x = sqft_living, y = price)) +
  geom_point(alpha = 0.5) +  # Scatter plot
  geom_smooth(method = "lm", color = "blue", se = FALSE) +
  labs(title = "Price vs. Size Fit",
       x = "Size (sqft_living)",
       y = "Price") +
  theme_minimal()

simple_lm
```

output




```
##
## Call:
## lm(formula = price ~ sqft_living, data = house_df)
##
## Coefficients:
## (Intercept)  sqft_living
##    -43580.7         280.6
```

##Model Summary (Extracts Coefficients, p-values, R^2)

```
##
## Call:
## lm(formula = price ~ sqft_living + sqft_lot + bathrooms
##     data = house_df, na.action = na.omit)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1011695  -136513   -23045   100989  4782979
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5.957e+05  1.325e+04 -44.950  < 2e-16 ***
## sqft_living  2.065e+02  3.364e+00  61.373  < 2e-16 ***
## sqft_lot    -2.664e-01  4.171e-02  -6.388 1.71e-10 ***
## bathrooms   -3.944e+04  3.443e+03 -11.456  < 2e-16 ***
## grade        1.037e+05  2.285e+03  45.379  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
```

Extract RMSE

Extract R-squared

Extract P-values

Print Results

```
## RMSE: 249532.2

## R²: 0.5380018

## P-values:

##  (Intercept)  sqft_living     sqft_lot    bathrooms
## 0.000000e+00 0.000000e+00 1.711092e-10 2.689854e-30 0.0
```

Print Results

```
##
## Call:
## lm(formula = price ~ sqft_living + sqft_lot15 + bathroom
##     grade, data = house_df, na.action = na.omit)
##
## Coefficients:
## (Intercept)  sqft_living   sqft_lot15    bathrooms      b
##  -4.658e+05    2.341e+02   -7.113e-01   -2.894e+04   -4.

##
## Call:
## lm(formula = price ~ poly(sqft_living, 2) + sqft_lot15 +
##     bedrooms + grade, data = house_df)
##
## Coefficients:
##           (Intercept)  poly(sqft_living, 2)1  poly(sqft_
##             -1.824e+05              2.696e+07
##             sqft_lot15               bathrooms
##             -7.345e-01             -1.565e+04
```