Stat 3301: Homework 1

Jane Weissberg (weissberg.11)

2024-08-26

Question 1

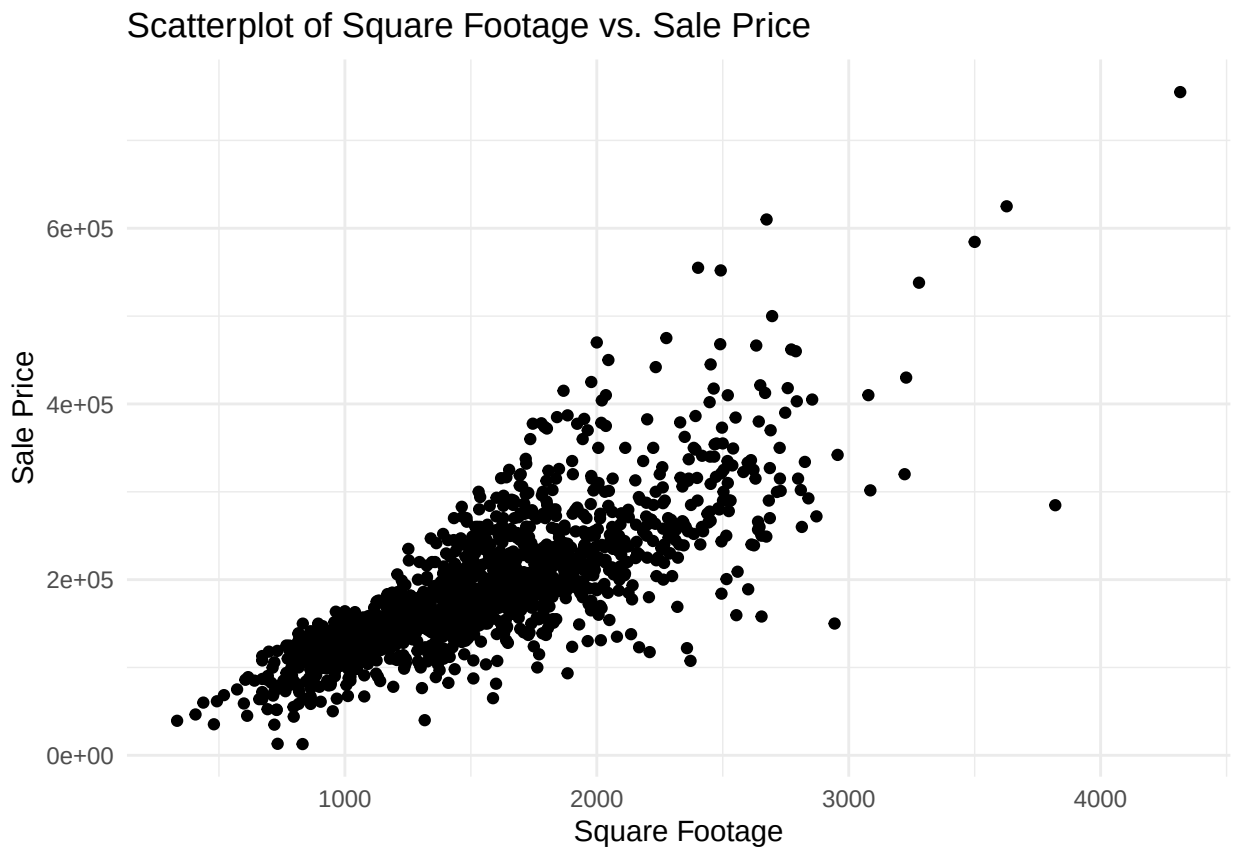
Import the data set ames_real_estate.csv (which is available on Carmen) and use it to do the following: Find the smallest and largest observed sale prices separately for each neighborhood. Make a scatterplot with the square footage of the house on the x-axis and the sale price of the house on the y-axis. Write a sentence describing how the relationship between SalePrice and SqFt.

Solution to Question 1

```
price_summary <- ames_data %>%
  group_by(Neighborhood) %>%
  summarize(
    MinPrice = min(SalePrice, na.rm = TRUE),
    MaxPrice = max(SalePrice, na.rm = TRUE)
  )
print(price_summary)
```

```
## # A tibble: 21 x 3
##    Neighborhood MinPrice MaxPrice
##    <chr>           <dbl>    <dbl>
##  1 Blmngtn        159895   159895
##  2 BrkSide         39300   223500
##  3 ClearCr        143000   328000
##  4 CollgCr        110000   475000
##  5 Crawfor         90350   335000
##  6 Edwards         58500   320000
##  7 Gilbert        115000   377500
##  8 IDOTRR          13100   212300
##  9 Mitchel         81500   300000
## 10 NAmes           68000   301600
## # i 11 more rows
```

```
ggplot(ames_data, aes(x=SqFt,y=SalePrice))+geom_point()+labs(x = "Square Footage", y = "Sale Price", ti
```



There is a positive relationship between square footage and sales price, indicating that as the size of the house increases, so does the price.

Question 2

Use the data set Heights from the libraryalr4 to calculate the average height for daughters whose mothers are at least 65 inches tall. (The purpose of this question is to make sure you have the package installed correctly.)

Solution to Question 2

```
heights_data <- Heights
average_height <- heights_data %>%
  filter(mheight >= 65) %>%
  summarize(AvgHeight = mean(dheight, na.rm = TRUE))
print(average_height)
```

```
##   AvgHeight
## 1  65.70102
```