

ECON_203_HW3_Ancel_Charles

Charles Ancel

2024-02-09

```
library(readxl)

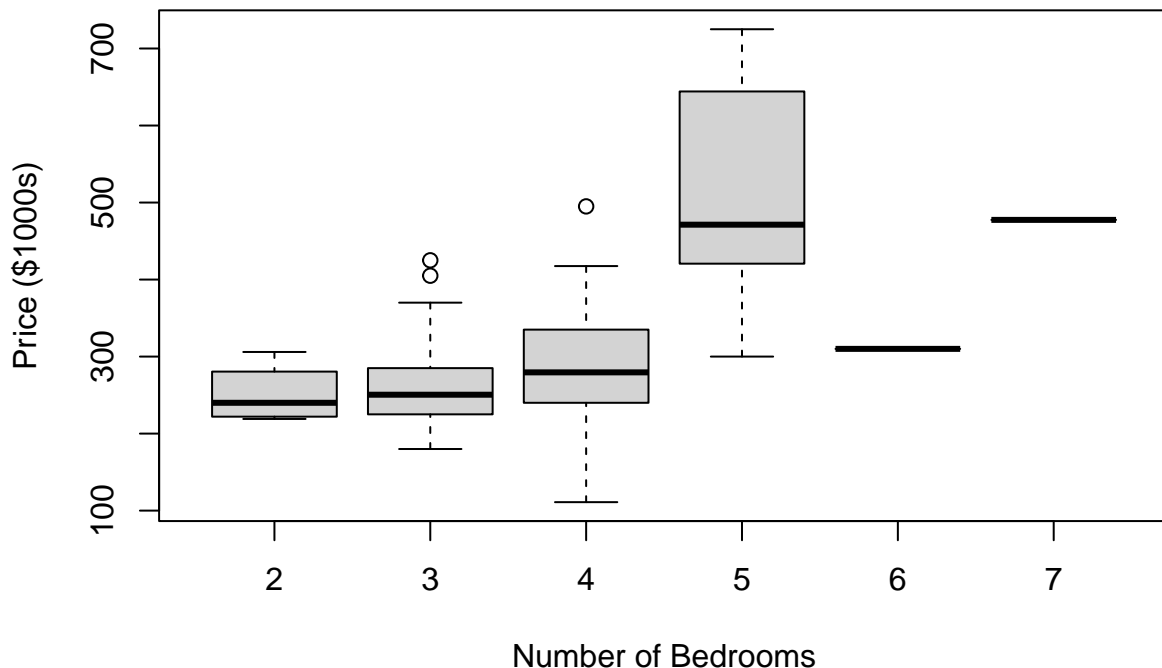
# Load the dataset
housing_data <- read_excel("housing.xls")
```

Task A: Box Plots Analysis

1. House Prices by Number of Bedrooms

```
boxplot(price ~ bdrms, data=housing_data, main="House Prices by Number of Bedrooms",
        xlab="Number of Bedrooms", ylab="Price ($1000s)")
```

House Prices by Number of Bedrooms

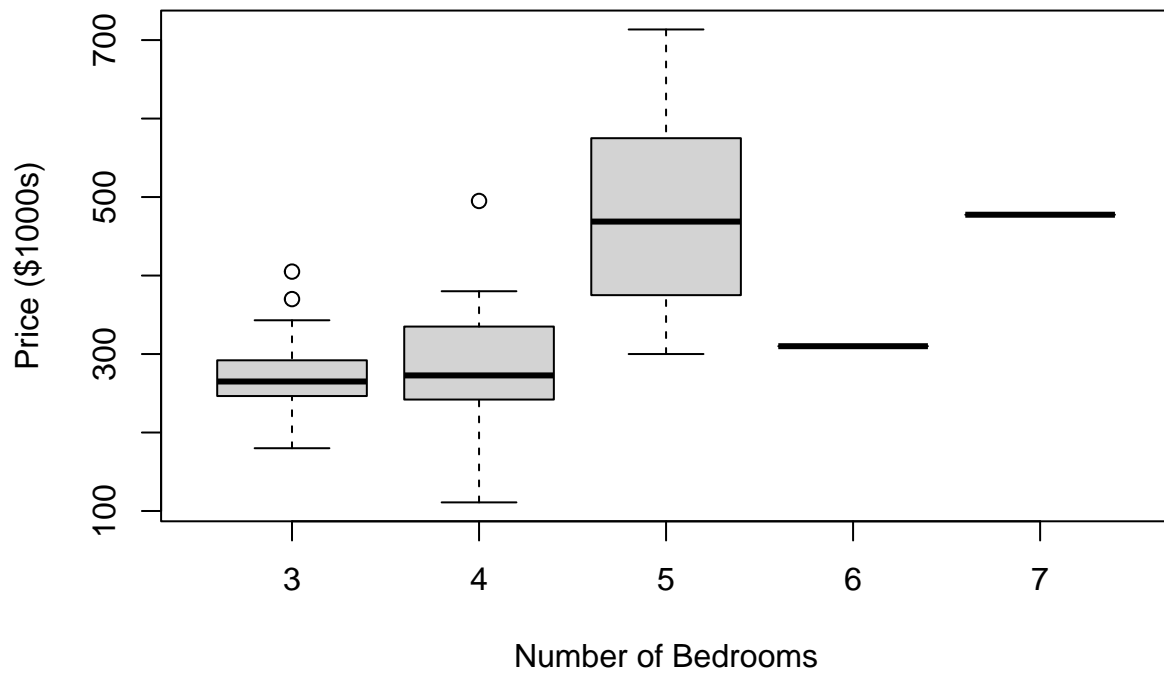


2. House Prices for Colonial vs. Non-Colonial Homes

```
boxplot(price[colonial == 1] ~ bdrms[colonial == 1], data=housing_data,
        main="Colonial House Prices by Number of Bedrooms",
        xlab="Number of Bedrooms", ylab="Price ($1000s)")
```

Colonial Houses

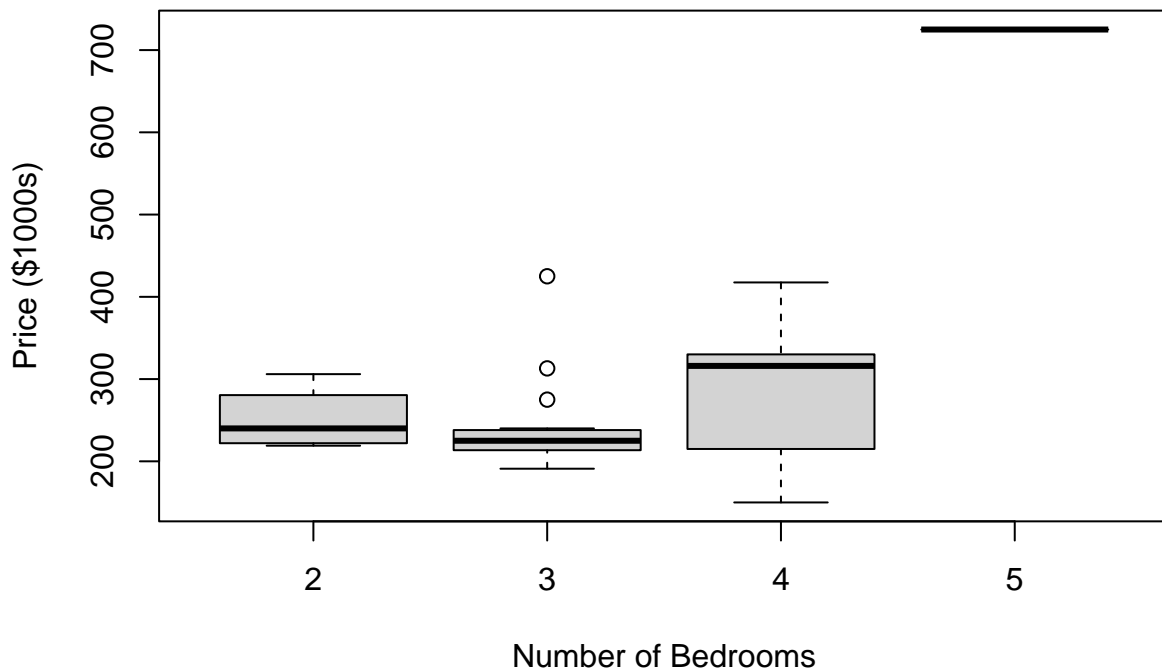
Colonial House Prices by Number of Bedrooms



```
boxplot(price[colonial == 0] ~ bdrms[colonial == 0], data=housing_data,  
        main="Non-Colonial House Prices by Number of Bedrooms",  
        xlab="Number of Bedrooms", ylab="Price ($1000s)")
```

Non-Colonial Houses

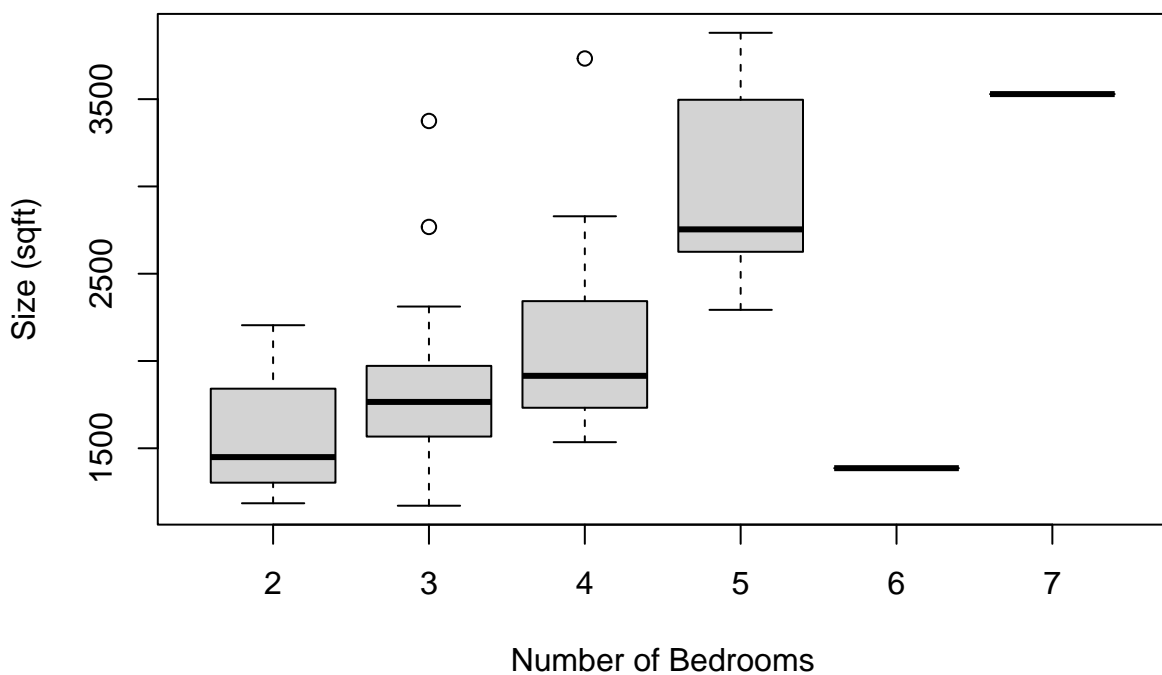
Non-Colonial House Prices by Number of Bedrooms



3. House Size by Number of Bedrooms

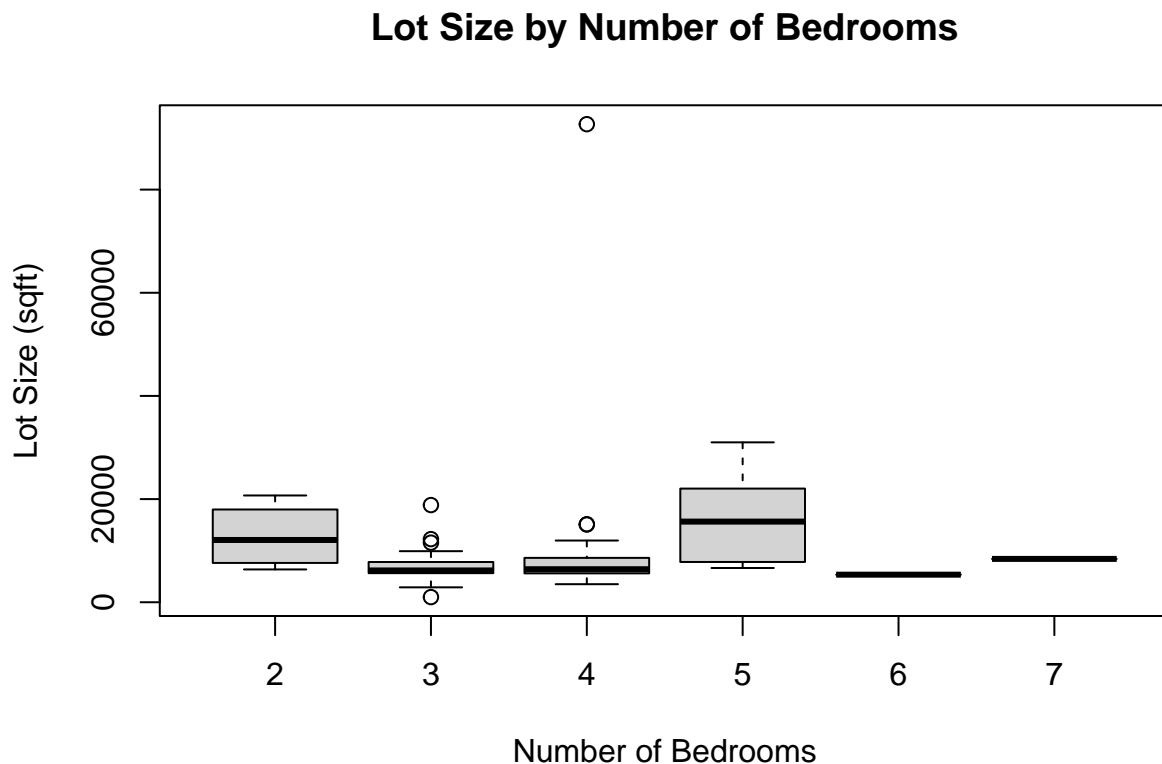
```
boxplot(sqrft ~ bdrms, data=housing_data, main="House Size by Number of Bedrooms",  
        xlab="Number of Bedrooms", ylab="Size (sqft)")
```

House Size by Number of Bedrooms



4. Lot Size by Number of Bedrooms

```
boxplot(lotsize ~ bdrms, data=housing_data, main="Lot Size by Number of Bedrooms",  
        xlab="Number of Bedrooms", ylab="Lot Size (sqft)")
```



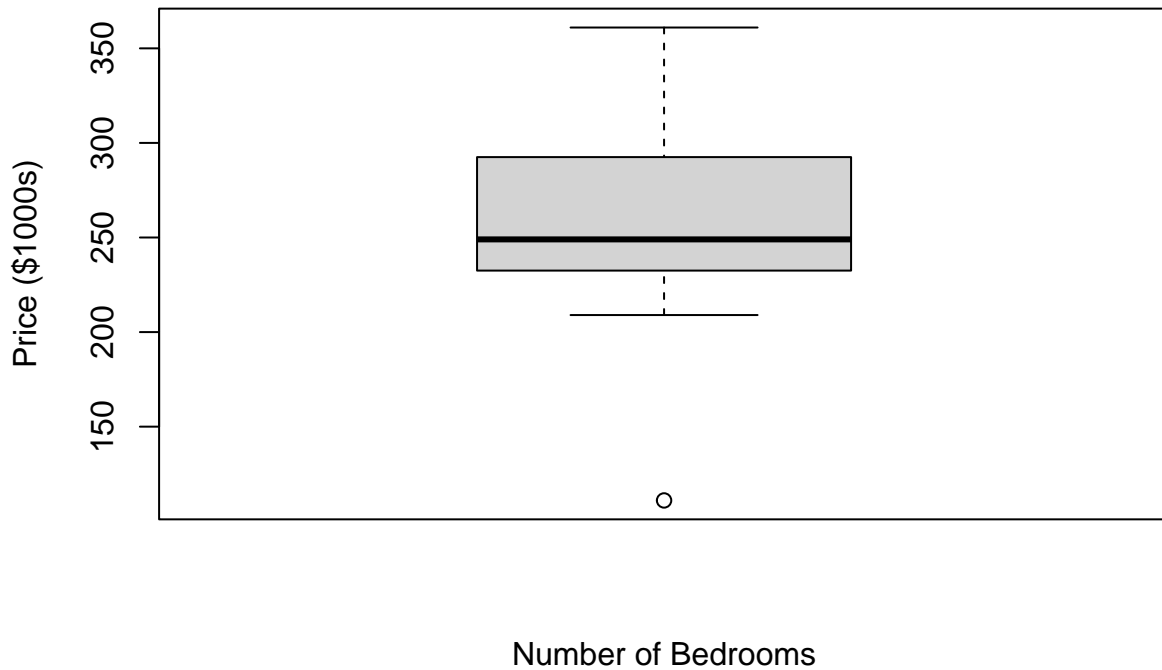
Task B: Price Range Estimation

Box Plots for Specific House Features

Condition 1: Lotsize ≤ 9000 and Sqrft ≤ 2500 .

```
boxplot(price[colonial == 1 & bdrms == 4 & lotsize <= 9000 & sqrft <= 2500]  
        ~ bdrms[colonial == 1 & bdrms == 4 & lotsize <= 9000 & sqrft <= 2500],  
        data=housing_data, main="Colonial 4-Bedroom Houses with Specified Conditions",  
        xlab="Number of Bedrooms", ylab="Price ($1000s)")
```

Colonial 4-Bedroom Houses with Specified Conditions



Condition 2: Lotsize > 9000 and Sqrft > 2500.

```
boxplot(price[colonial == 1 & bdrms == 4 & lotsize > 9000 & sqrft > 2500]  
~ bdrms[colonial == 1 & bdrms == 4 & lotsize > 9000 & sqrft > 2500],  
data=housing_data, main="Colonial 4-Bedroom Houses Exceeding Specified Conditions",  
xlab="Number of Bedrooms", ylab="Price ($1000s)")
```

Colonial 4-Bedroom Houses Exceeding Specified Conditions

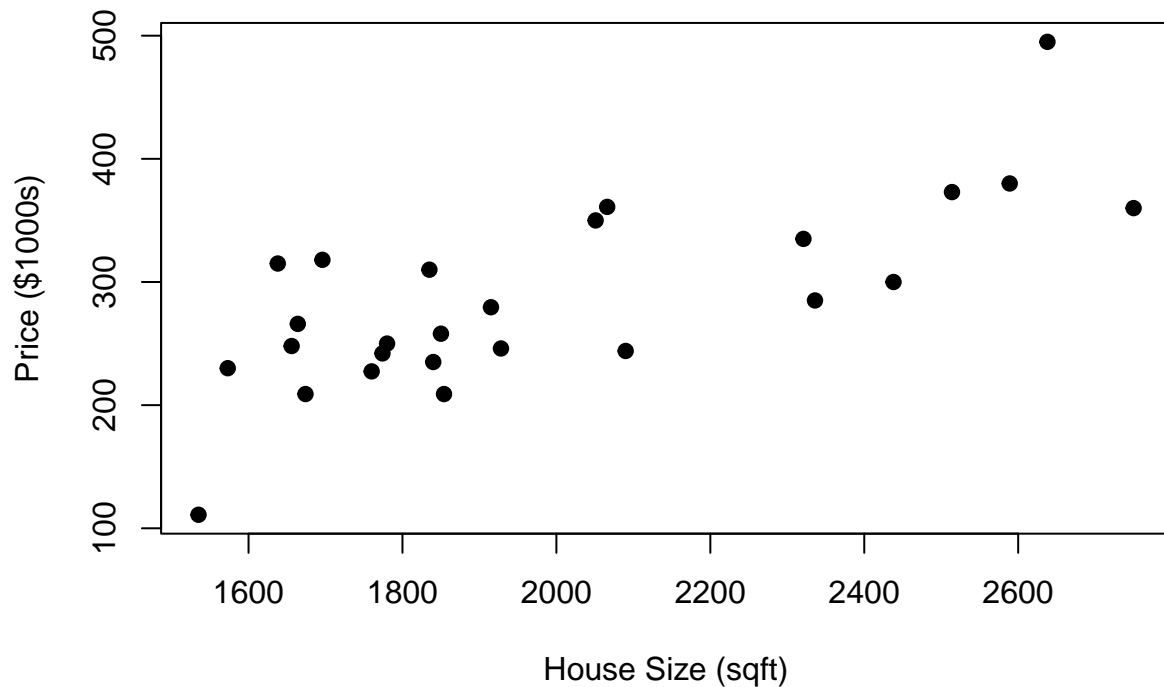


Scatter Plots for House and Lot Sizes

House Size vs. Price

```
plot(housing_data$sqft[housing_data$colonial == 1 & housing_data$bdrms == 4],  
     housing_data$price[housing_data$colonial == 1 & housing_data$bdrms == 4],  
     main = "House Size vs. Price for Colonial Four-Bedroom Houses",  
     xlab = "House Size (sqft)",  
     ylab = "Price ($1000s)",  
     pch = 19)
```

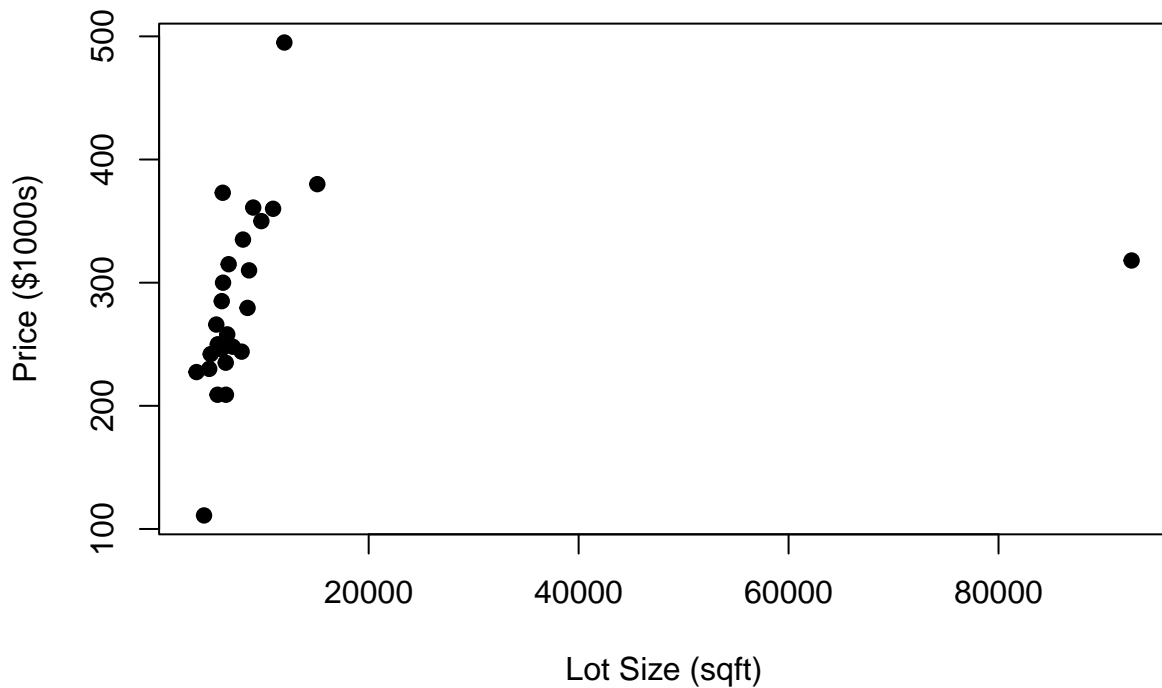
House Size vs. Price for Colonial Four-Bedroom Houses



Lot Size vs. Price

```
plot(housing_data$lotsize[housing_data$colonial == 1 & housing_data$bdrms == 4],  
     housing_data$price[housing_data$colonial == 1 & housing_data$bdrms == 4],  
     main = "Lot Size vs. Price for Colonial Four-Bedroom Houses",  
     xlab = "Lot Size (sqft)",  
     ylab = "Price ($1000s)",  
     pch = 19)
```

Lot Size vs. Price for Colonial Four-Bedroom Houses



Question 2

Task 2(a):

We are to determine which statement is true regarding the house prices and sizes:

1. **Houses with five bedrooms are the ones with the lowest prices.**
 - From the first box plot (House Prices by Number of Bedrooms), it does not appear that houses with five bedrooms have the lowest prices; in fact, they are priced in the mid-range compared to others.
2. **Houses with six bedrooms are the most expensive ones.**
 - The box plot for houses with six bedrooms shows a high price range, but it's not clear if they are the most expensive since the plot for seven bedrooms is not fully shown.
3. **Among the colonial houses, the cheapest house has four bedrooms.**
 - In the second box plot (Colonial House Prices by Number of Bedrooms), we can see that the colonial houses with four bedrooms do not appear to be the cheapest. The plot indicates that the price range for four-bedroom colonial houses includes both lower-end and higher-end prices within its distribution.
4. **Among the non-colonial houses, the most expensive one has four bedrooms.**
 - The third box plot (Non-Colonial House Prices by Number of Bedrooms) shows that the box plot for four-bedroom non-colonial houses does not reach the highest price point compared to those with two and three bedrooms.
5. **The largest houses are the ones with six bedrooms.**
 - The fourth box plot (House Size by Number of Bedrooms) indicates that houses with six bedrooms have a larger size on average compared to other categories, which supports this statement as being true.

Based on the box plots provided, the most likely true statement is:

5. **The largest houses are the ones with six bedrooms.**

Task 2(b):

This question is related to the adequate price range for a client's colonial four-bedroom house with a lot size ≤ 9000 sqft and house size ≤ 2500 sqft:

1. **Below 200,000**

- Looking at the seventh box plot (Colonial 4-Bedroom Houses with Specified Conditions), the price does not seem to go below \$200,000.

2. **Between 200,000 and 250,000**

- This range is more consistent with the lower whisker of the seventh box plot, indicating that it could be a reasonable lower bound for pricing.

3. **Between 300,000 and 400,000**

- The central tendency (median) of the box plot for colonial four-bedroom houses with specified conditions seems to be in this range.

4. **Between 400,000 and 500,000**

- The box plot does not extend this high, so it is unlikely to be the correct price range.

5. **Above 500,000**

- This is not supported by the seventh box plot.

Based on the seventh box plot provided, the most likely correct answer seems to be:

3. Between 300,000 and 400,000

This is the range that includes the median price for colonial four-bedroom houses under the specified conditions, making it a reasonable estimate for the client's house price range.