

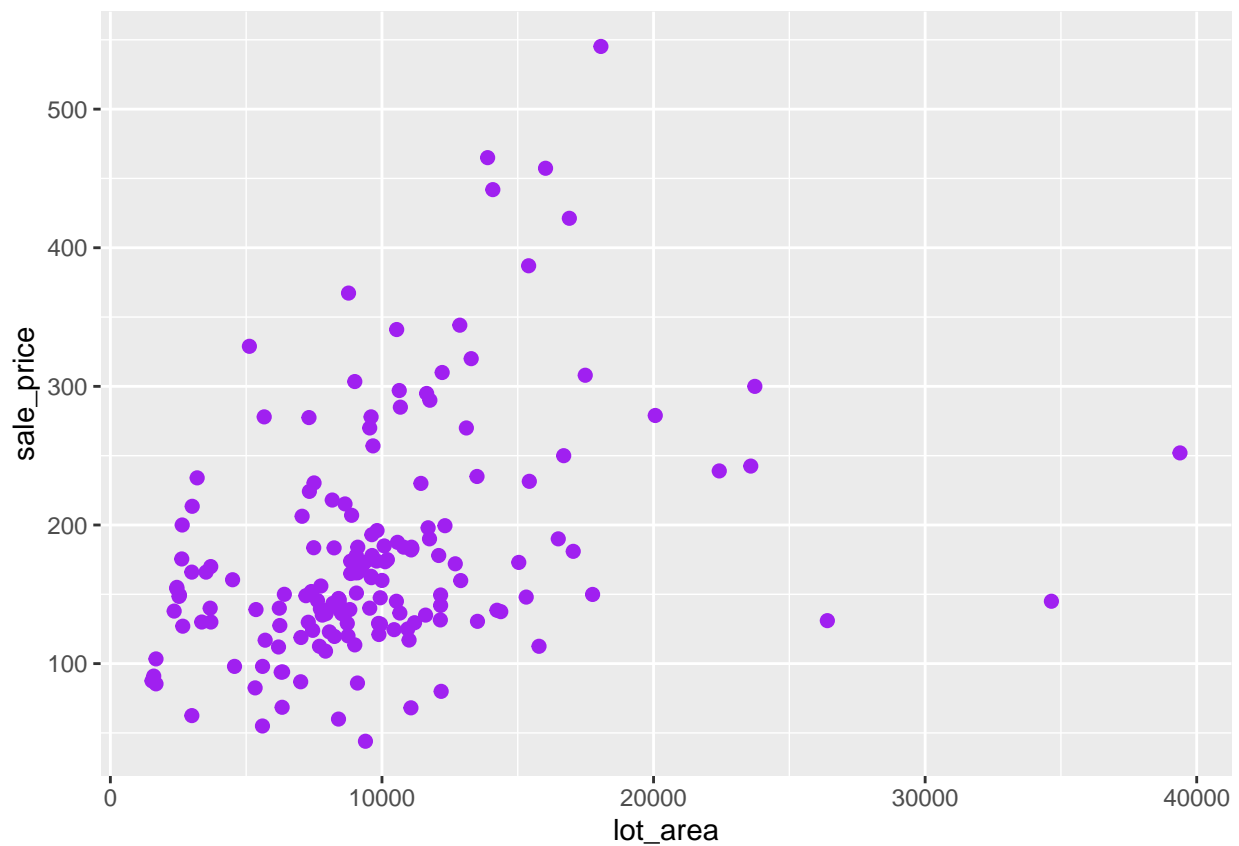
# HW4

Andrew Shao

2024-10-02

## 5.2.1 Q1

```
library(ggplot2)
library(r02pro)
ggplot(data = sahp) +
  geom_point(mapping = aes(x = lot_area,
                           y = sale_price),
            color = 'purple',
            size = 2)
```



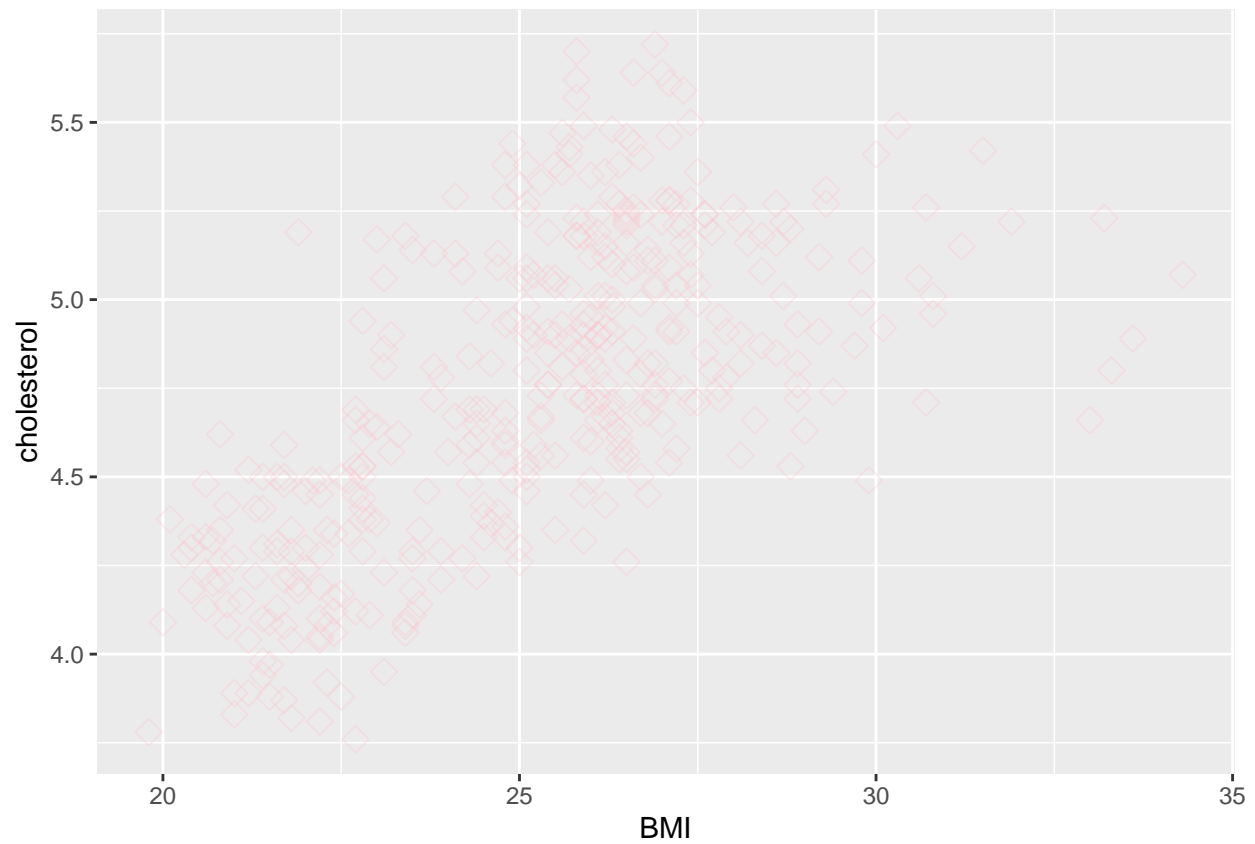
## 5.2.1 Q2

```
ggplot(data = gm2004) +
  geom_point(mapping = aes(x = BMI,
```

```

        y = cholesterol),
  color = 'pink',
  size = 3,
  alpha = 0.3,
  shape = 5)

```

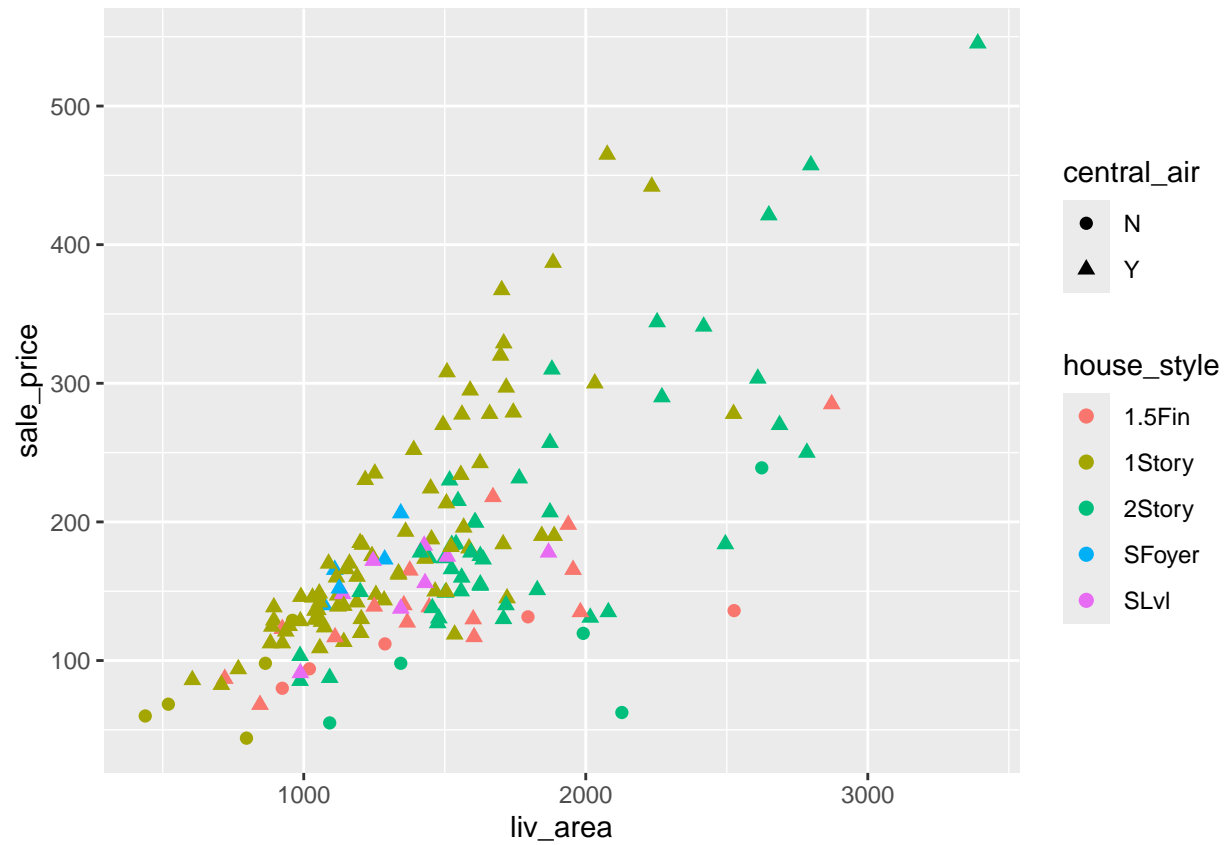


### 5.3.5 Q1

```

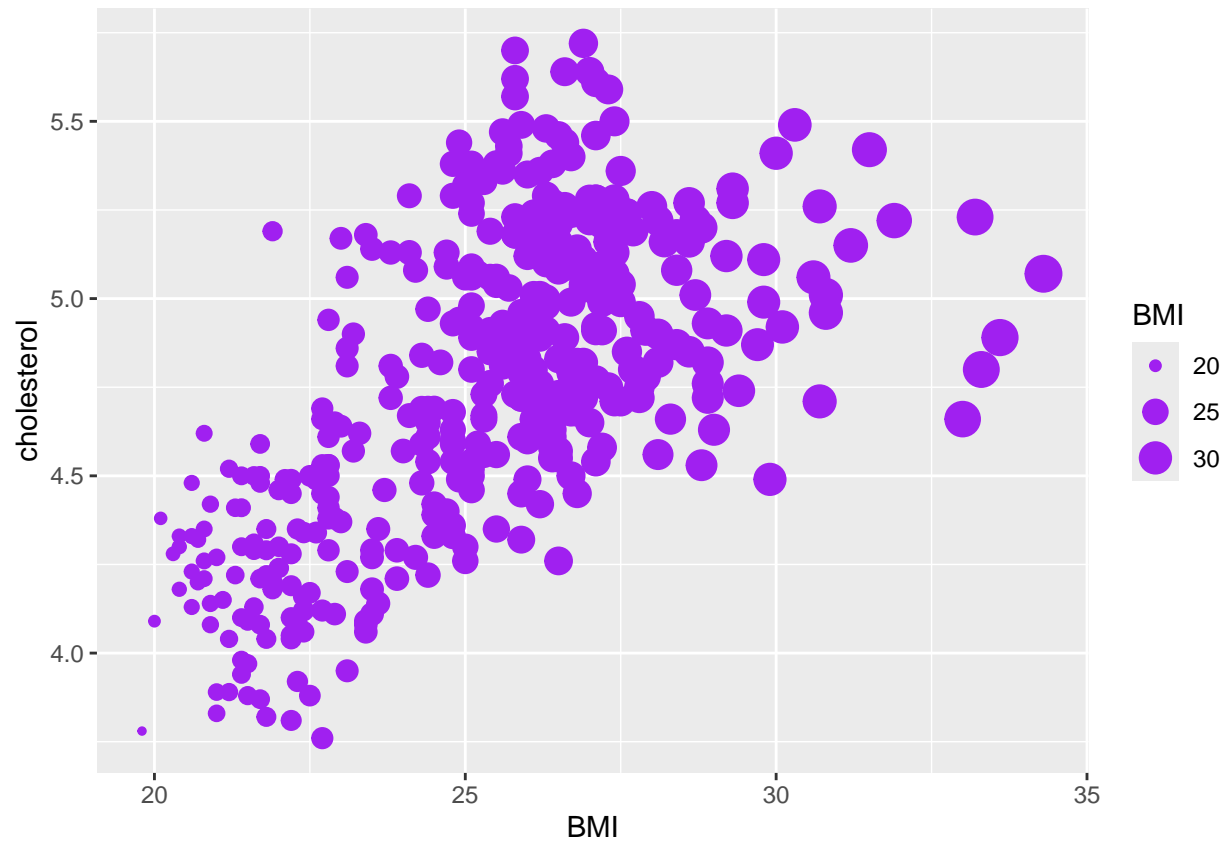
ggplot(data = sahp) +
  geom_point(mapping = aes(x = liv_area,
    y = sale_price,
    color = house_style,
    shape = central_air),
    size = 2)

```



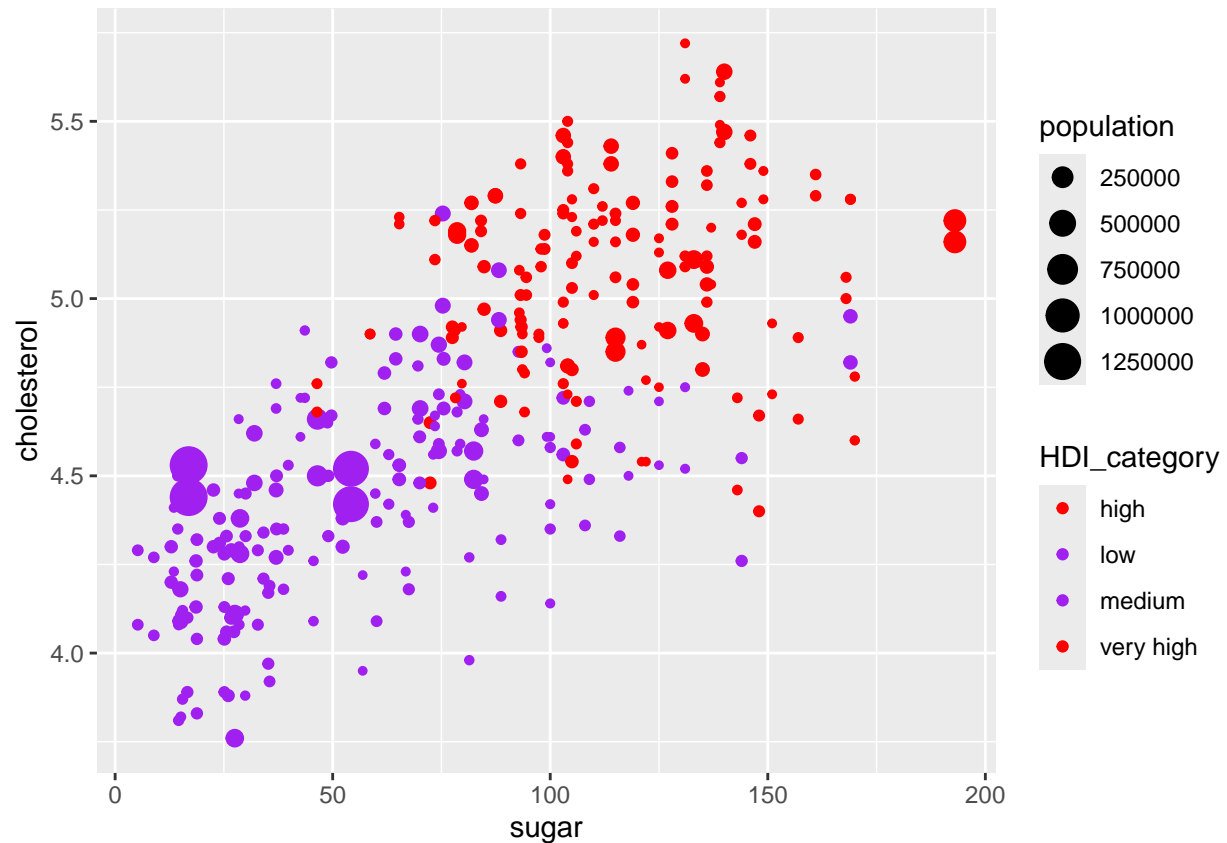
### 5.3.5 Q2

```
ggplot(data = gm2004) +
  geom_point(mapping = aes(x = BMI,
                           y = cholesterol,
                           size = BMI),
             color = 'purple')
```



### 5.3.5 Q3

```
ggplot(data = remove_missing(gm2004, vars = 'HDI_category')) +
  geom_point(mapping = aes(x = sugar,
                           y = cholesterol,
                           color = HDI_category,
                           size = population)) +
  scale_color_manual(values = c('low' = 'purple', 'medium' = 'purple', 'high' = 'red', 'very high' = 'red'))
```



High HDI is associated with high cholesterol and sugar.

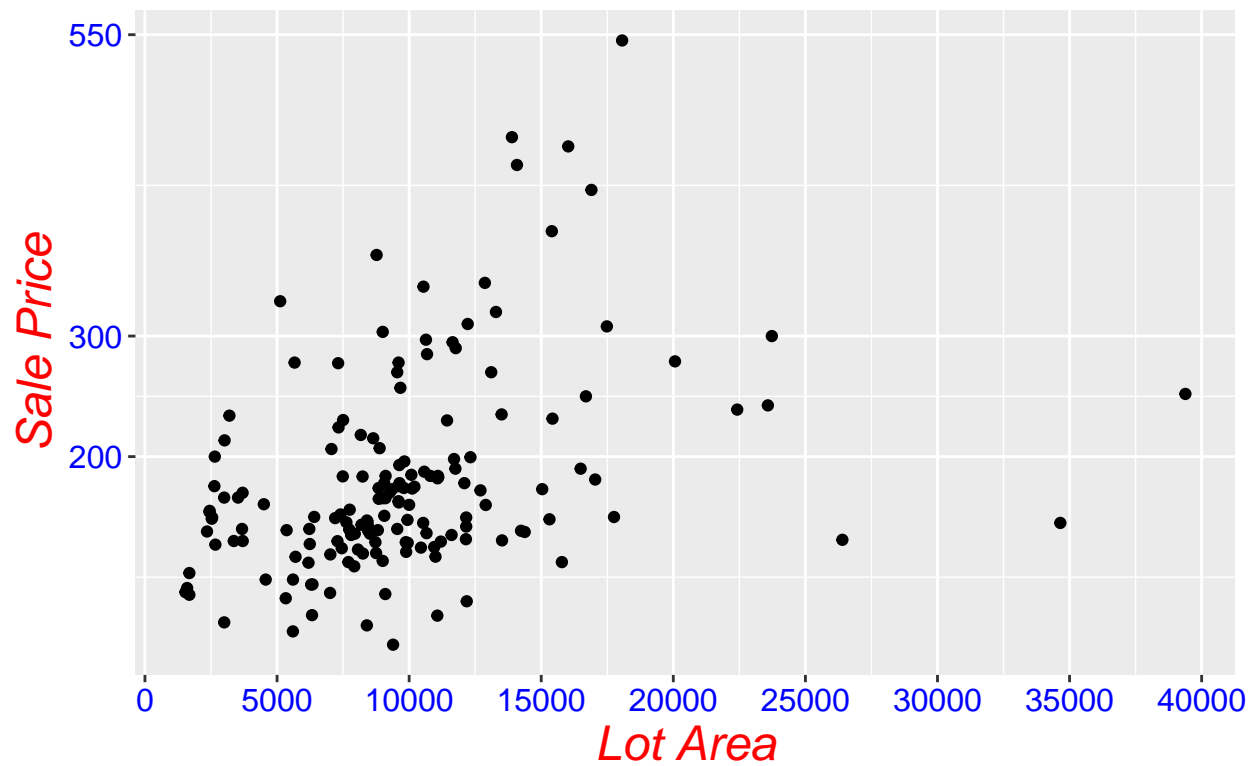
#### 5.4.6 Q1

```
mytheme <- theme(axis.title = element_text(size = 18,
                                             color = "red",
                                             face = "italic"),
                  axis.text = element_text(size = 12,
                                             color = "blue"),
                  plot.title = element_text(size = 24,
                                             color = "magenta",
                                             face = "bold",
                                             hjust = 0.5))

myplot <- ggplot(data = sahp) +
  geom_point(mapping = aes(x = lot_area,
                           y = sale_price)) +
  scale_x_continuous(breaks = seq(0, 40000, 5000)) +
  scale_y_continuous(breaks = c(0, 200, 300, 550)) +
  xlab('Lot Area') +
  ylab('Sale Price') +
  ggtitle('Sale Price vs Lot Area') +
  mytheme

myplot
```

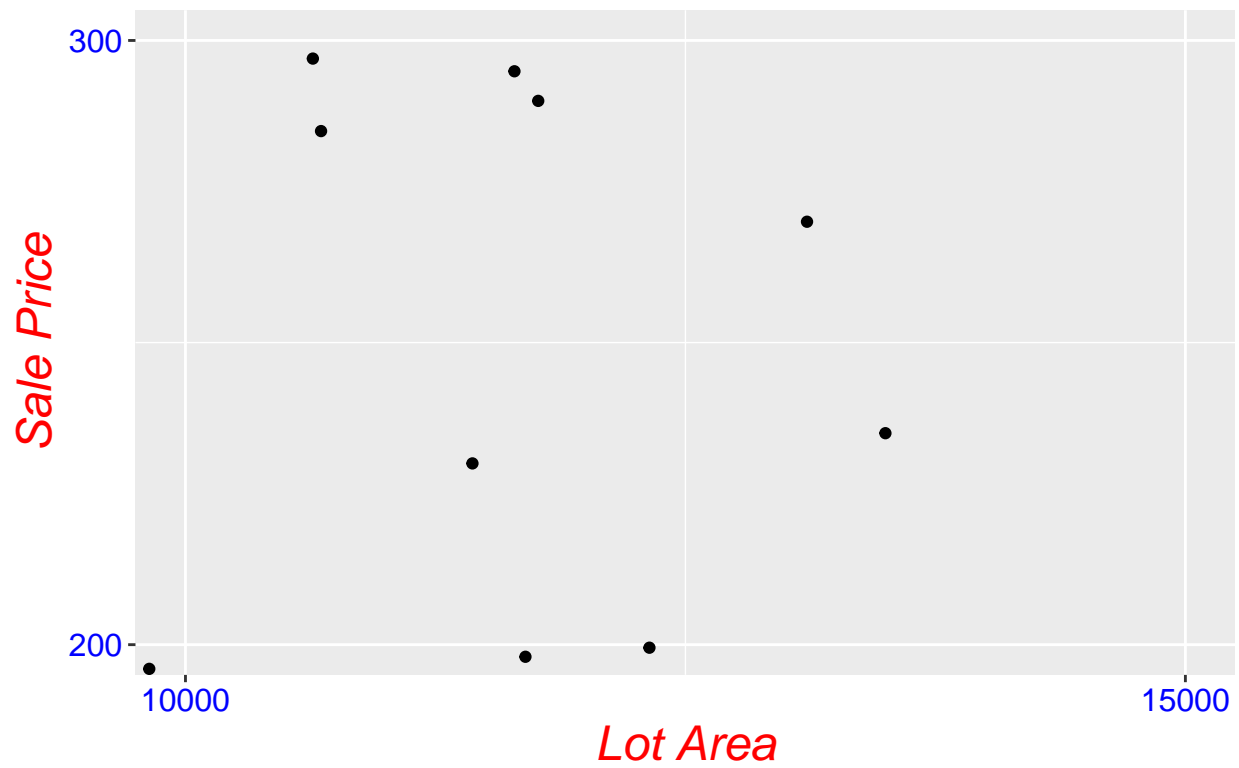
## Sale Price vs Lot Area



### 5.4.6 Q2

```
myplot +  
  coord_cartesian(xlim = c(10000, 15000),  
                  ylim = c(200, 300))
```

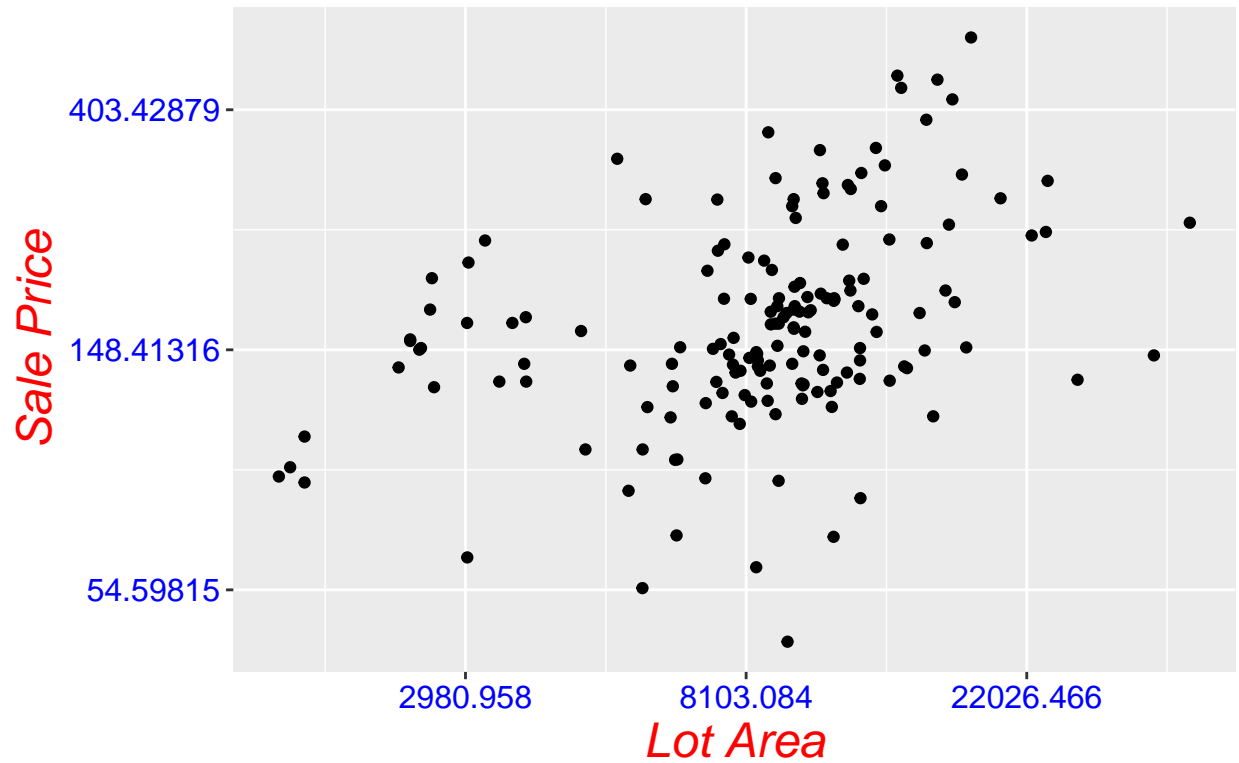
## Sale Price vs Lot Area



### 5.4.6 Q3

```
myplot +  
  scale_x_continuous(trans = 'log') +  
  scale_y_continuous(trans = 'log')
```

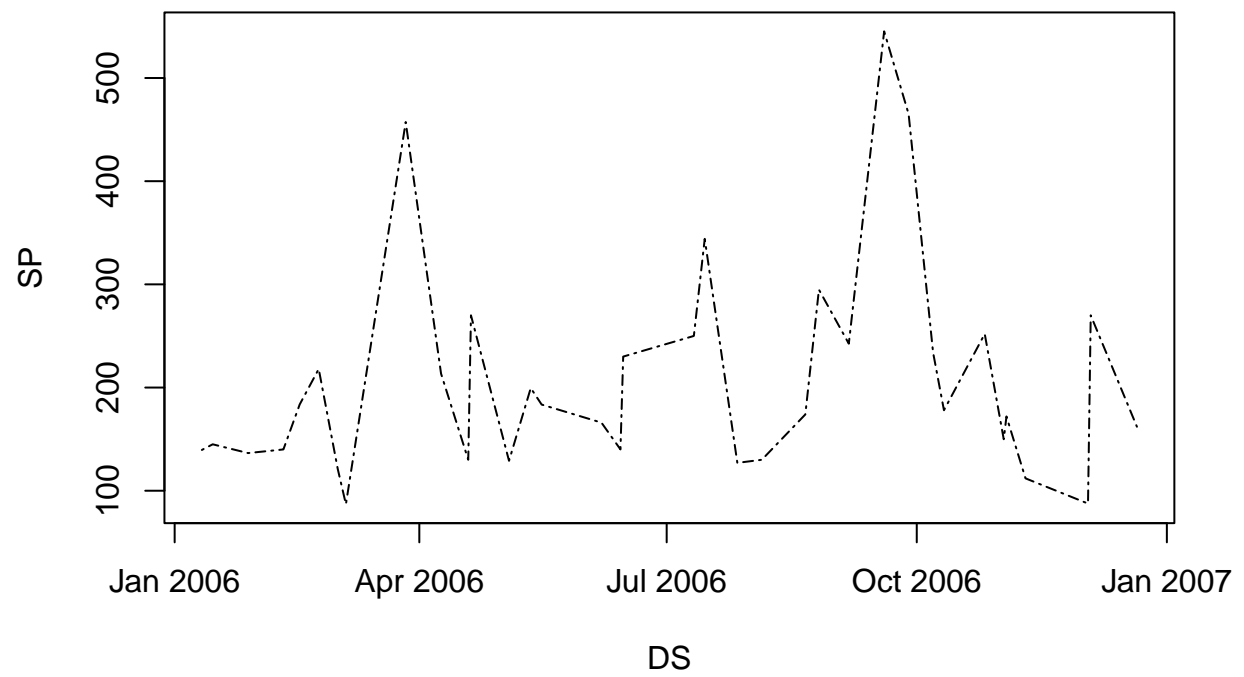
## Sale Price vs Lot Area



### 5.5.5 Q1

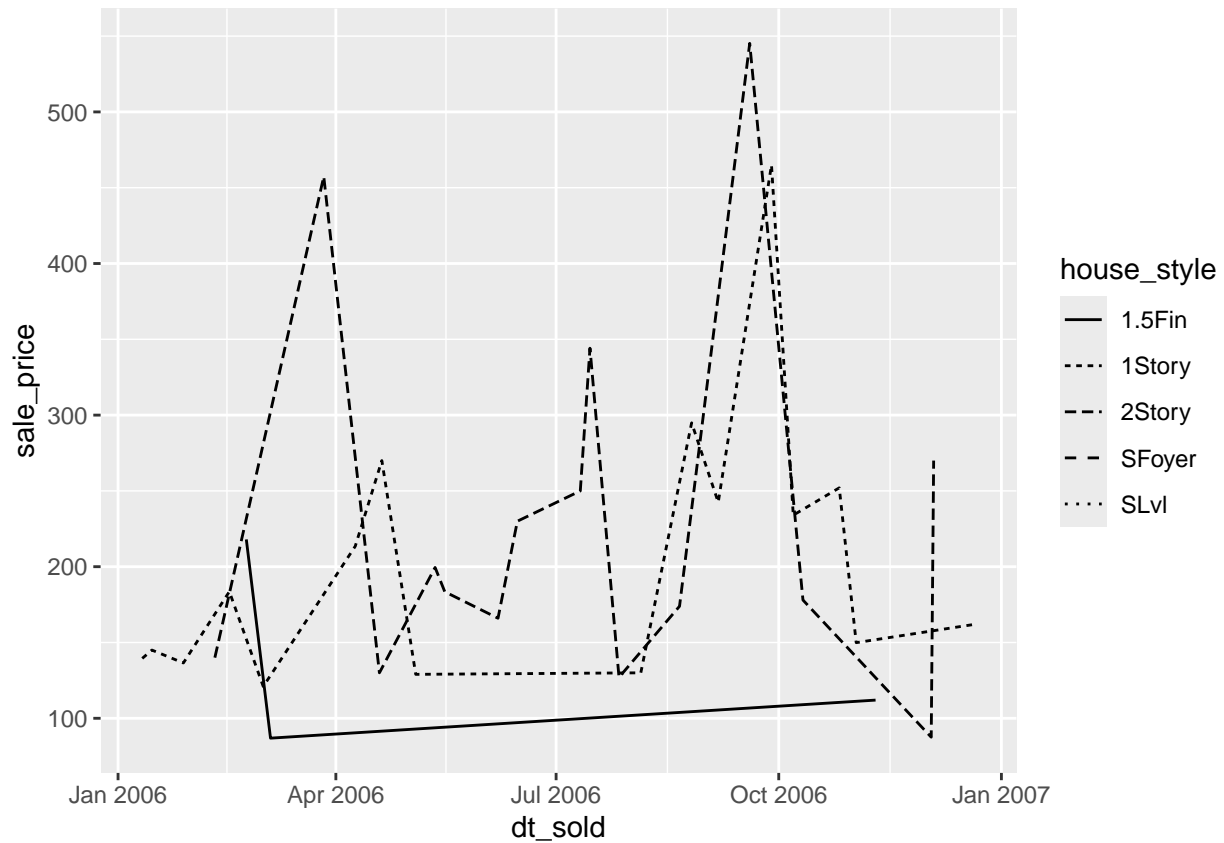
```
sahp_2006 <- sahp[format(sahp$dt_sold, "%Y") < 2007, ] #all houses sold before 2007
plot(sahp_2006$dt_sold[order(sahp_2006$dt_sold)],
     sahp_2006$sale_price[order(sahp_2006$dt_sold)],
     type = 'l',
     xlab = 'DS',
     ylab = 'SP',
     lty = 'twodash')
```





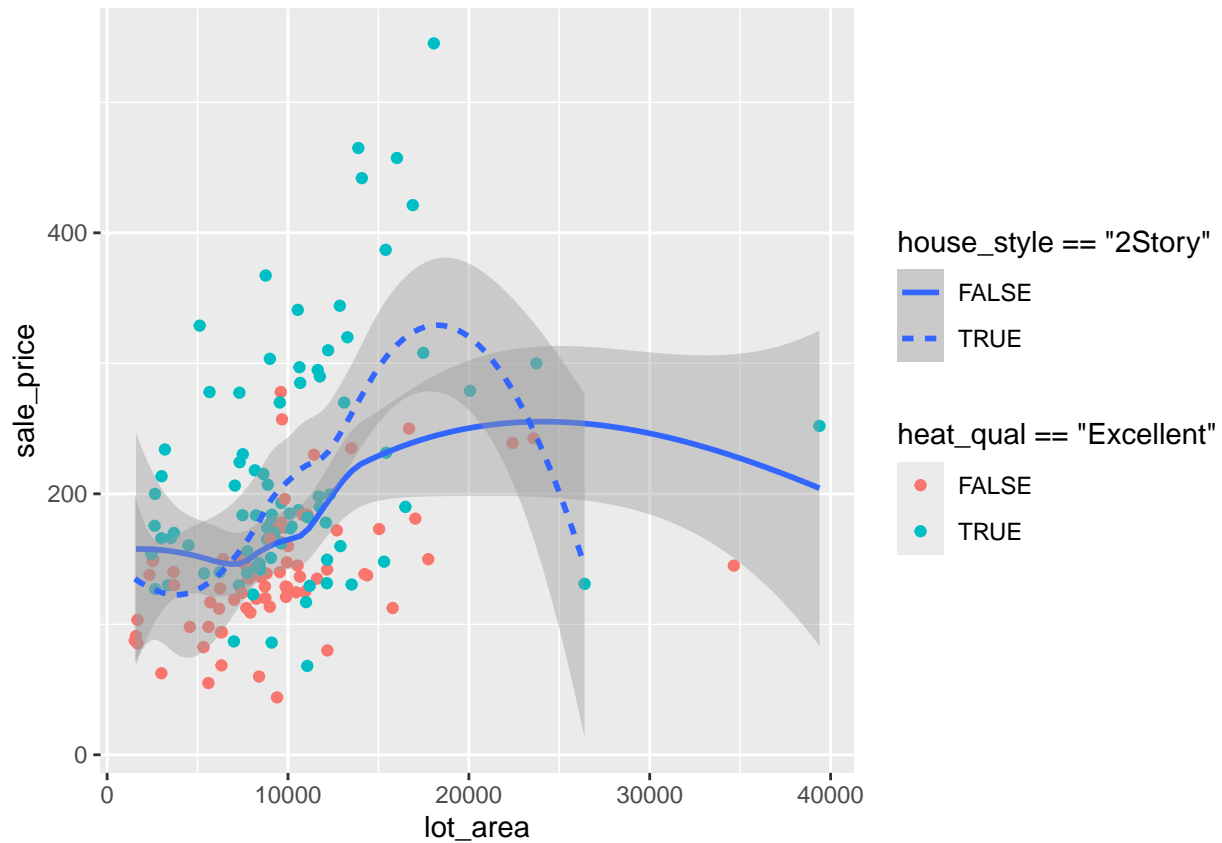
### 5.5.5 Q2

```
ggplot(data = sahp_2006) +  
  geom_line(mapping = aes(x = dt_sold,  
                          y = sale_price,  
                          linetype = house_style))
```



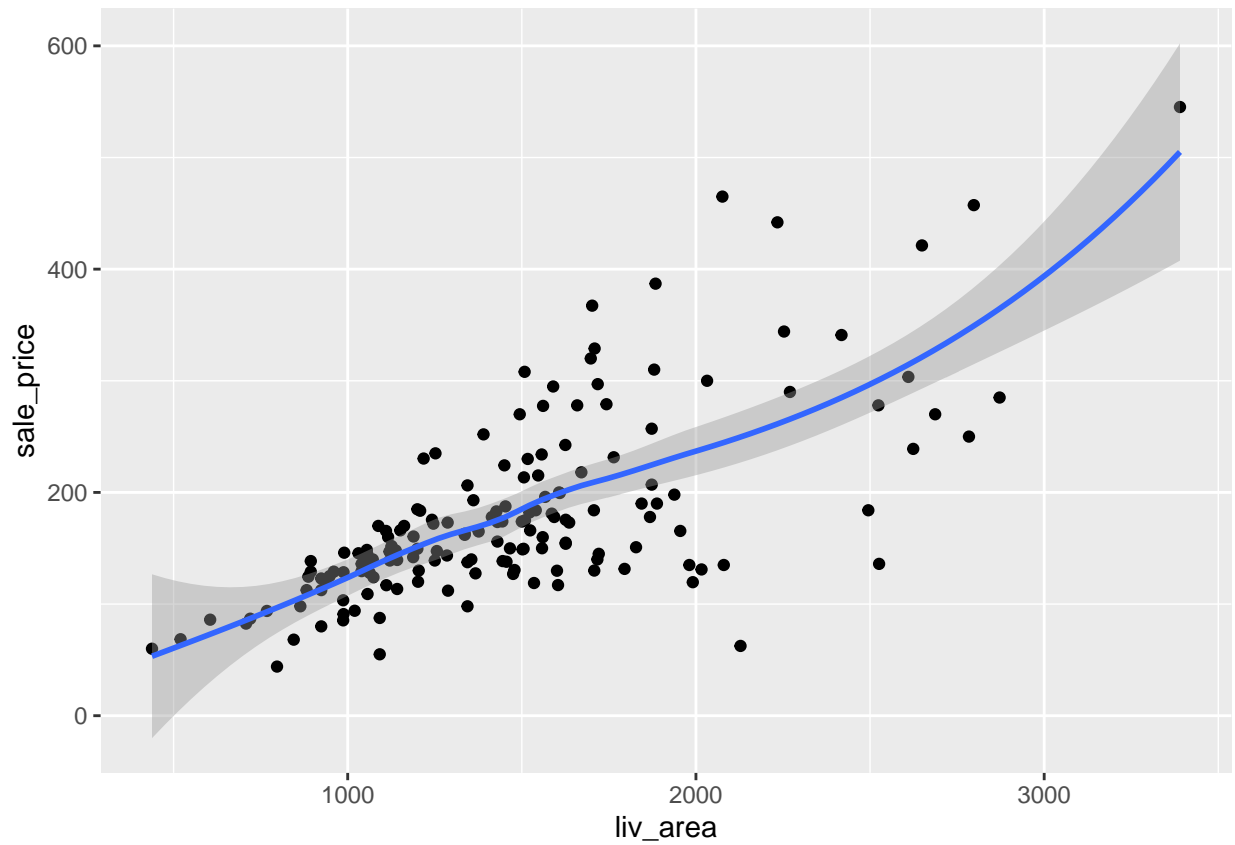
### 5.7.3 Q1

```
ggplot(data = sahp,
       mapping = aes(x = lot_area,
                     y = sale_price)) +
  geom_point(mapping = aes(colour = heat_qual == 'Excellent')) +
  geom_smooth(mapping = aes(linetype = house_style == '2Story'))
```



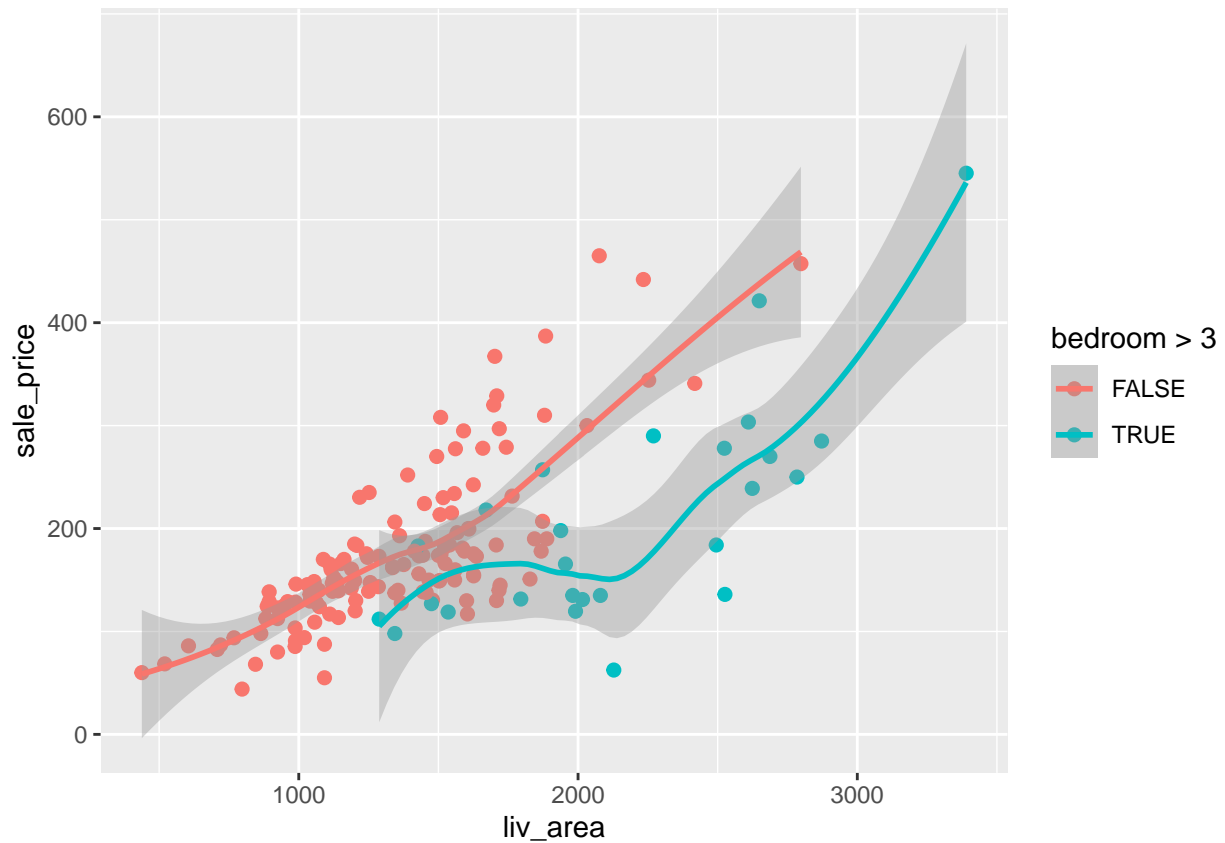
### 5.7.3 Q2

```
new_plot <- ggplot(data = sahp,
                    mapping = aes(x = liv_area,
                                  y = sale_price)) +
  geom_point() +
  geom_smooth()
new_plot
```



### 5.7.3 Q3

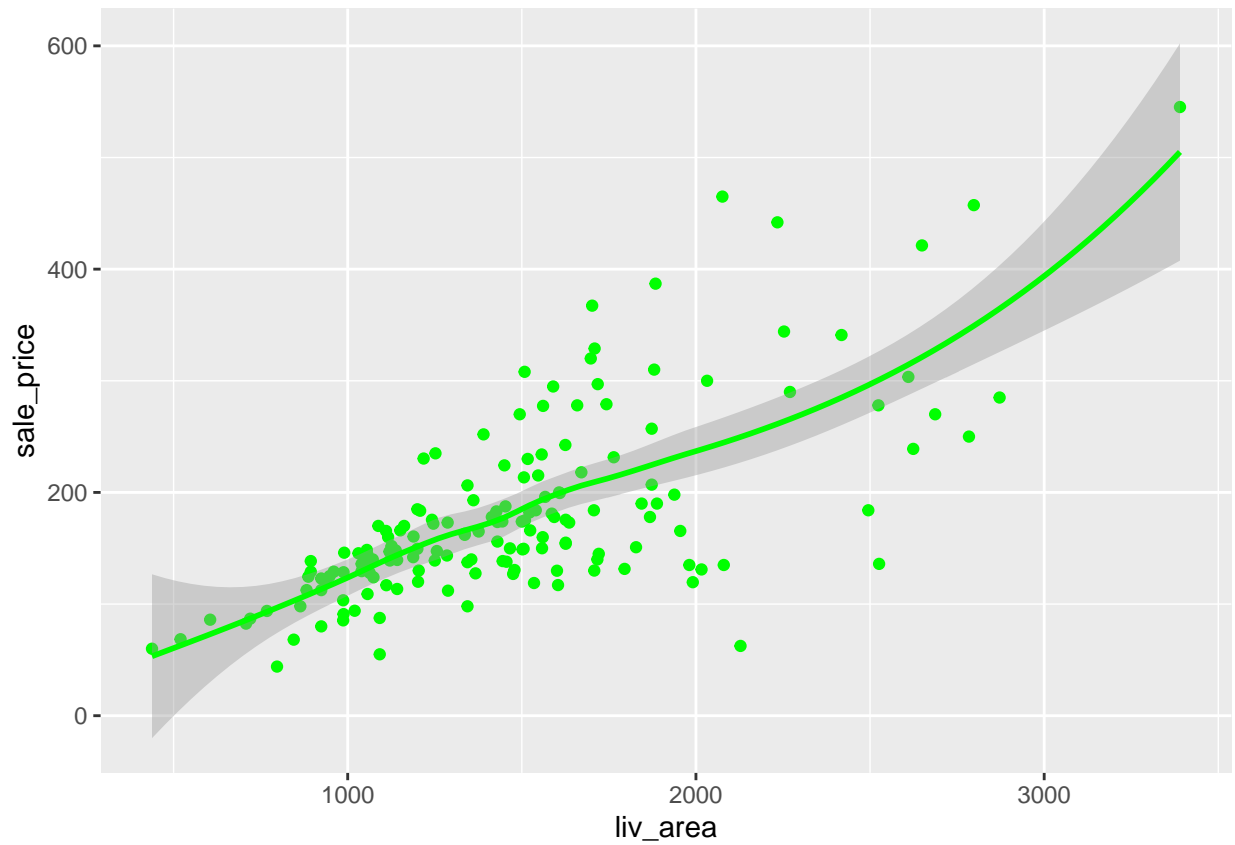
```
ggplot(data = sahp,
        mapping = aes(x = liv_area,
                       y = sale_price)) +
  geom_point(mapping = aes(color = bedroom > 3),
             size = 2) +
  geom_smooth(mapping = aes(color = bedroom > 3))
```



### 5.7.3 Q4

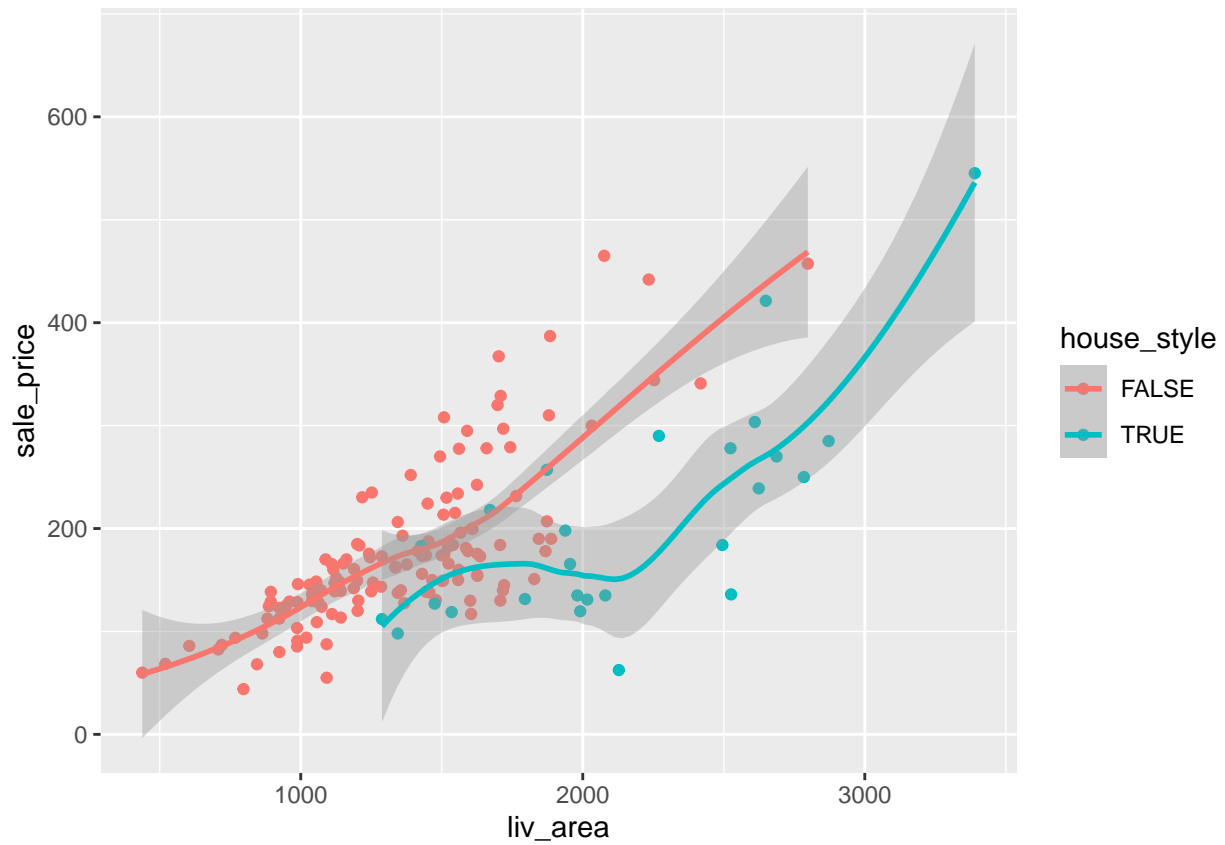
No, there are no global constant-valued aesthetics.

```
ggplot(data = sahp,  
       mapping = aes(x = liv_area,  
                     y = sale_price)) +  
  geom_point(color = "green") +  
  geom_smooth(color = "green")
```



### 5.7.3 Q5

```
ggplot(data = sahp,  
       mapping = aes(x = liv_area,  
                     y = sale_price,  
                     color = house_style)) +  
  geom_point(mapping = aes(color = bedroom > 3)) +  
  geom_smooth(mapping = aes(color = bedroom > 3))
```



The global aesthetic mapping has been overwritten by the local geom aesthetic mapping. The name remains `house_style` because it is an artifact of `ggplot` implementation.