Activity 1: LA Housing Data

Alice Chang

```
m1 <- lm(log(price) ~ log(sqft) + bed + city, data = LA)
summary(m1)
```

```
##
## Call:
## lm(formula = log(price) ~ log(sqft) + bed + city, data = LA)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.26020 -0.24897 -0.01613  0.21804  1.37277
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        5.13068    0.21201  24.200   <2e-16 ***
## log(sqft)          1.20729    0.03036  39.769   <2e-16 ***
## bed               -0.03010    0.01284  -2.345   0.0191 *
## cityLong Beach    -0.88280    0.03467 -25.464   <2e-16 ***
## citySanta Monica  -0.09416    0.04022  -2.341   0.0194 *
## cityWestwood      -0.46244    0.04876  -9.484   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3704 on 1588 degrees of freedom
## Multiple R-squared:  0.8742, Adjusted R-squared:  0.8738
## F-statistic:  2206 on 5 and 1588 DF,  p-value: < 2.2e-16
```

```
m2 <- lm(log(price) ~ log(sqft) + bed + city + log(sqft):city, data = LA)
summary(m2)
```

```
##
## Call:
## lm(formula = log(price) ~ log(sqft) + bed + city + log(sqft):city,
##     data = LA)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.30385 -0.23866 -0.01576  0.21562  1.36668
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   4.43151    0.38515  11.506  < 2e-16 ***
## log(sqft)                     1.29578    0.05019  25.820  < 2e-16 ***
## bed                          -0.03794    0.01296  -2.928 0.003460 **
## cityLong Beach               -0.53386    0.37968  -1.406 0.159902
## citySanta Monica              1.75128    0.47010   3.725 0.000202 ***
## cityWestwood                  2.43192    0.90674   2.682 0.007394 **
## log(sqft):cityLong Beach     -0.03663    0.04730  -0.774 0.438807
## log(sqft):citySanta Monica   -0.24345    0.06052  -4.022 6.03e-05 ***
## log(sqft):cityWestwood       -0.38773    0.12251  -3.165 0.001581 **
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3676 on 1585 degrees of freedom
## Multiple R-squared:  0.8763, Adjusted R-squared:  0.8757
## F-statistic:  1404 on 8 and 1585 DF,  p-value: < 2.2e-16
```

```
m3 <- lm(log(price) ~ log(sqft) + bed + log(sqft):bed, data = LA)
summary(m3)
```

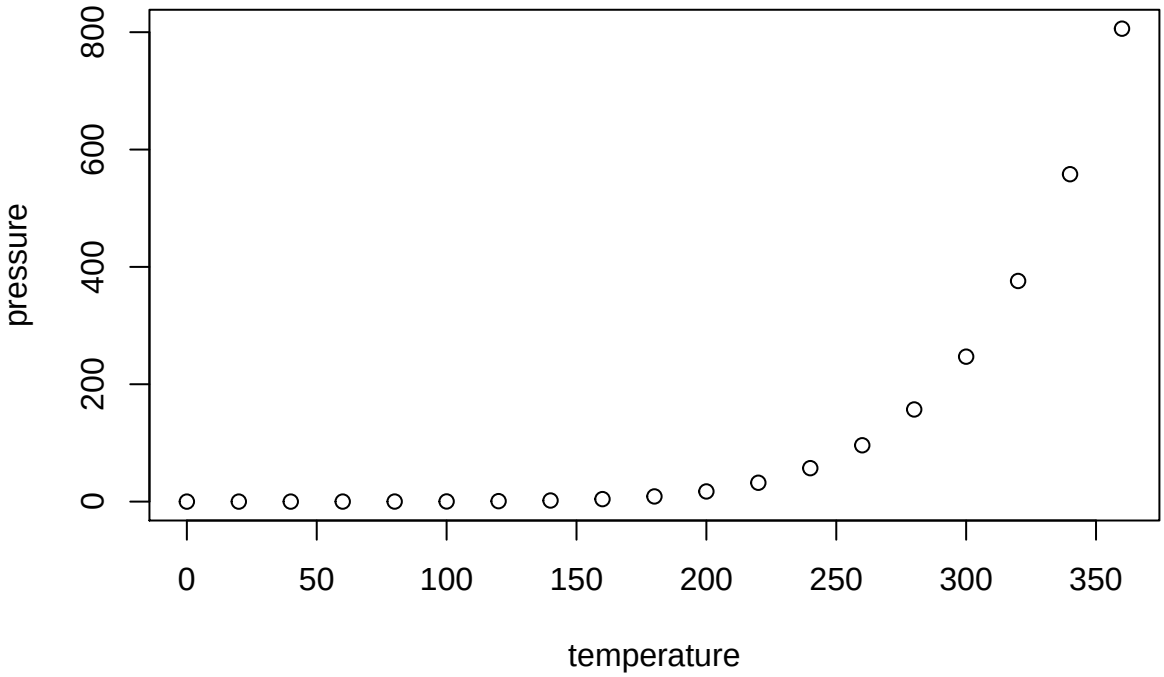
```
##
## Call:
## lm(formula = log(price) ~ log(sqft) + bed + log(sqft):bed, data = LA)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.75668 -0.32825 -0.04576  0.31841  1.85602
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.803227   0.271328  10.331  < 2e-16 ***
## log(sqft)      1.487273   0.040007  37.175  < 2e-16 ***
## bed           -0.644164   0.067255  -9.578  < 2e-16 ***
## log(sqft):bed  0.064093   0.008023   7.989 2.59e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4783 on 1590 degrees of freedom
## Multiple R-squared:  0.7899, Adjusted R-squared:  0.7895
## F-statistic:  1992 on 3 and 1590 DF,  p-value: < 2.2e-16
```

The geometry of the first model seems to be four parallel planes, while the geometry of the second model would be four planes with different slopes.

Beverly Hills appears to be the reference level for city?

For every increase in bedrooms in a home – holding all other variables constant – the log price of the house decreases by 5.13. This makes sense because by holding the square footage of a house constant and adding more rooms into the limited space, it is likely that the value of the house will go down. ## Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.