

# houses

Chris Parrish

January 18, 2016

houses

references:

- Cannon, et al., Stat2, chapter 04, example 4.1

Import the data.

```
data <- read.csv("Houses.csv", header=TRUE)
head(data, 4)
```

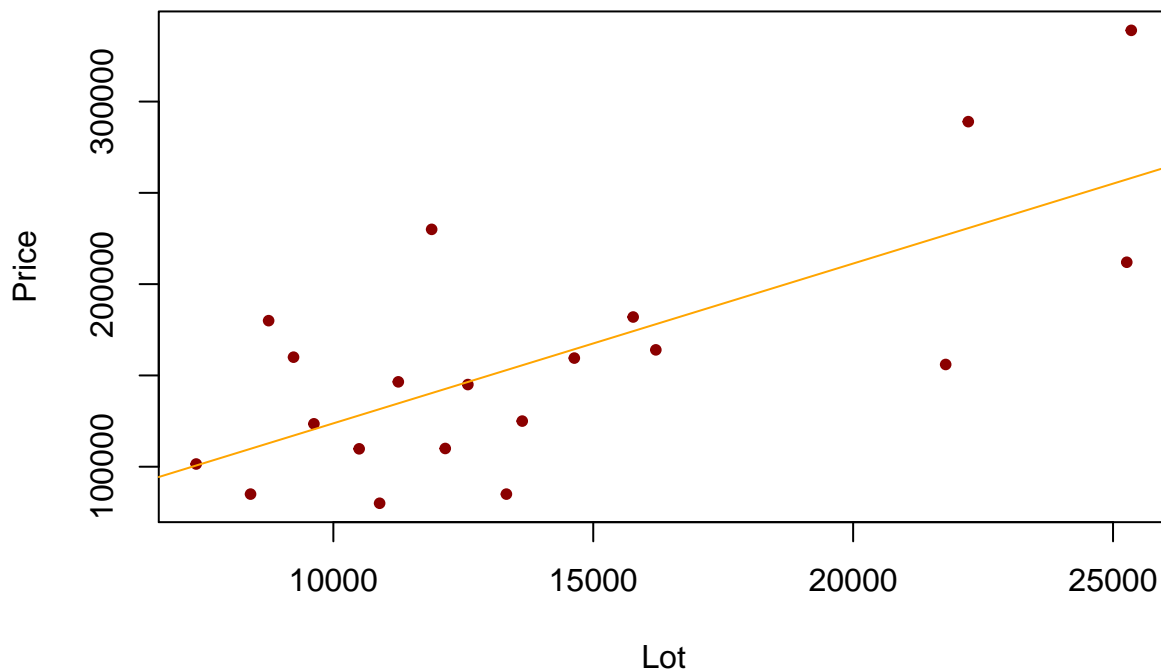
```
##   Price Size  Lot
## 1 212000 4148 25264
## 2 230000 2501 11891
## 3 339000 4374 25351
## 4 289000 2398 22215
```

```
dim(data)
```

```
## [1] 20 3
```

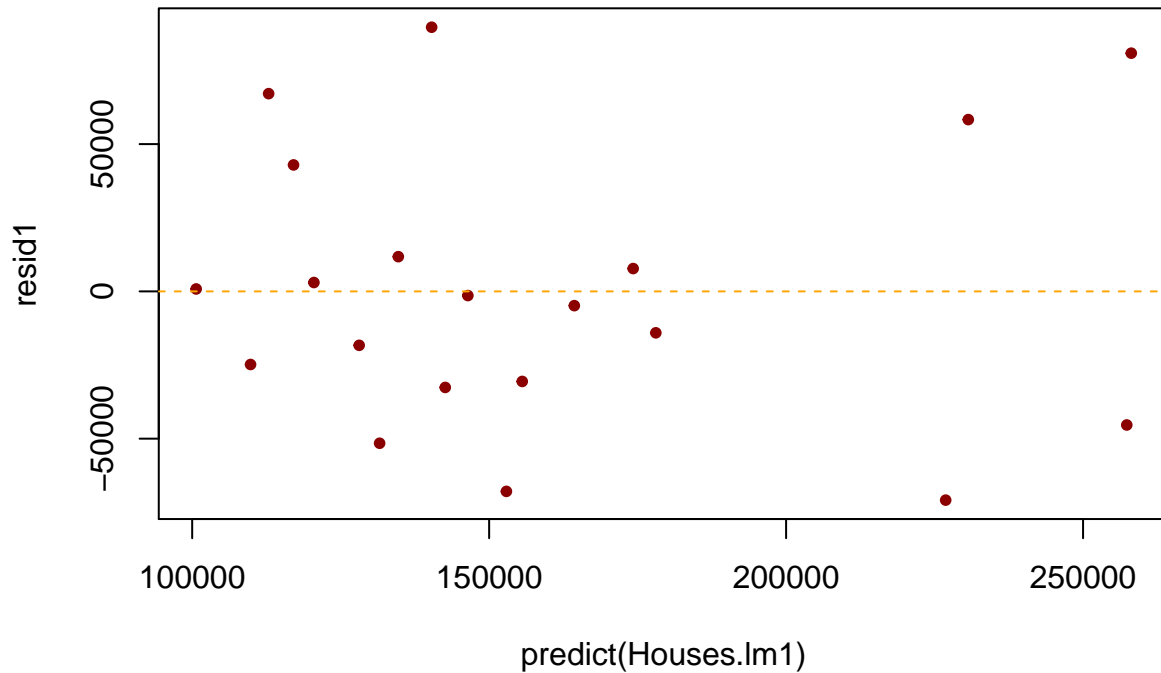
Predict *Price* by *Lot* alone.

```
plot(Price ~ Lot, data=data,
     pch=20, col="darkred")
Houses.lm1 <- lm(Price ~ Lot, data=data)
abline(Houses.lm1, col="orange")
```



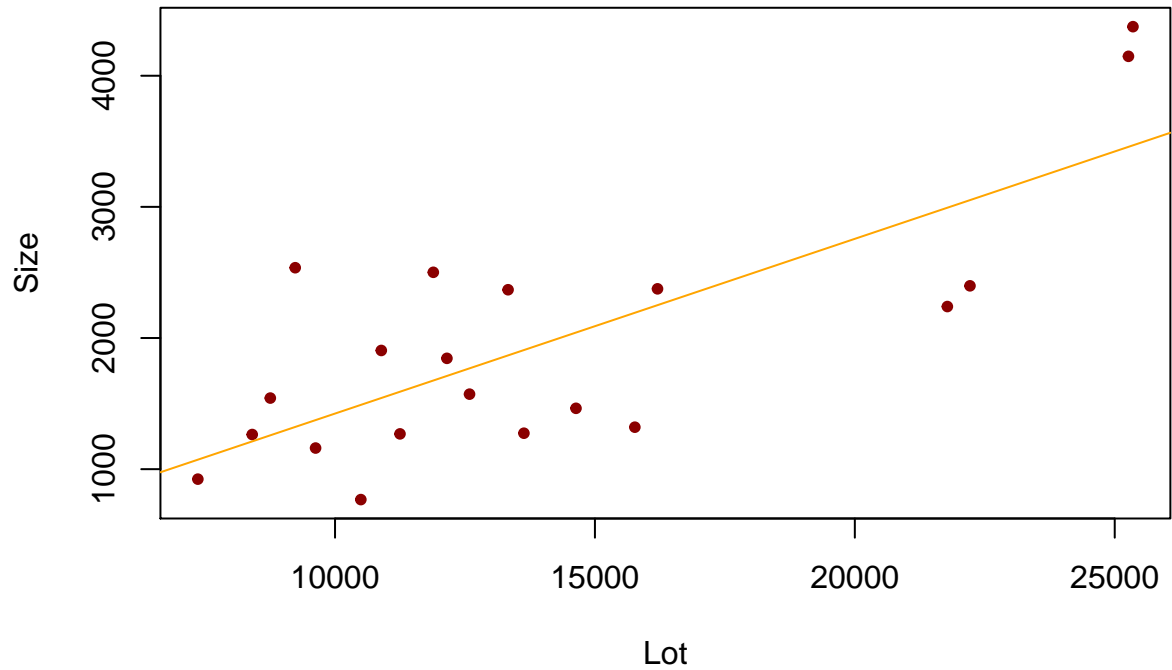
Residuals for  $Price \sim Lot$ ,  $data = data$ .

```
resid1 <- resid(Houses.lm1)
plot(predict(Houses.lm1), resid1,
      pch=20, col="darkred")
abline(h=0, col="orange", lty="dashed")
```



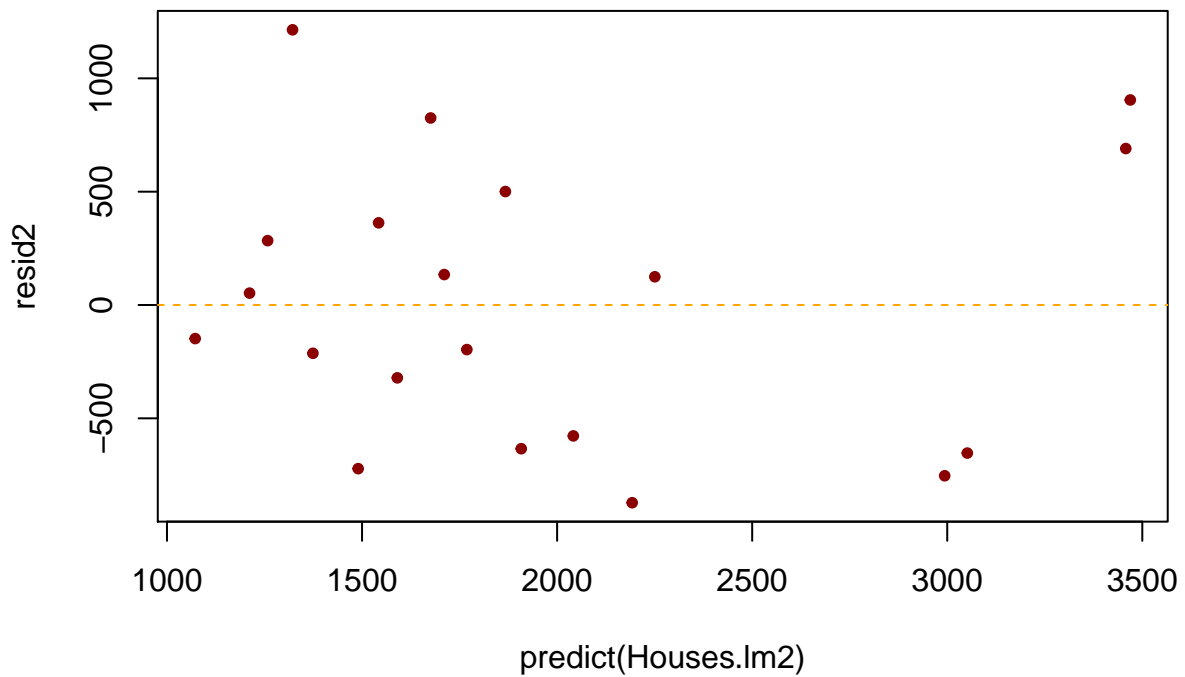
Predict *Size* by *Lot* alone.

```
plot(Size ~ Lot, data=data,
      pch=20, col="darkred")
Houses.lm2 <- lm(Size ~ Lot, data=data)
abline(Houses.lm2, col="orange")
```



Residuals for  $Size \sim Lot$ ,  $data = data$ .

```
resid2 <- resid(Houses.lm2)
plot(predict(Houses.lm2), resid2,
      pch=20, col="darkred")
abline(h=0, col="orange", lty="dashed")
```



Model:  $resid1 \sim resid2$

```
Houses.lm3 <- lm(resid1 ~ resid2)
```

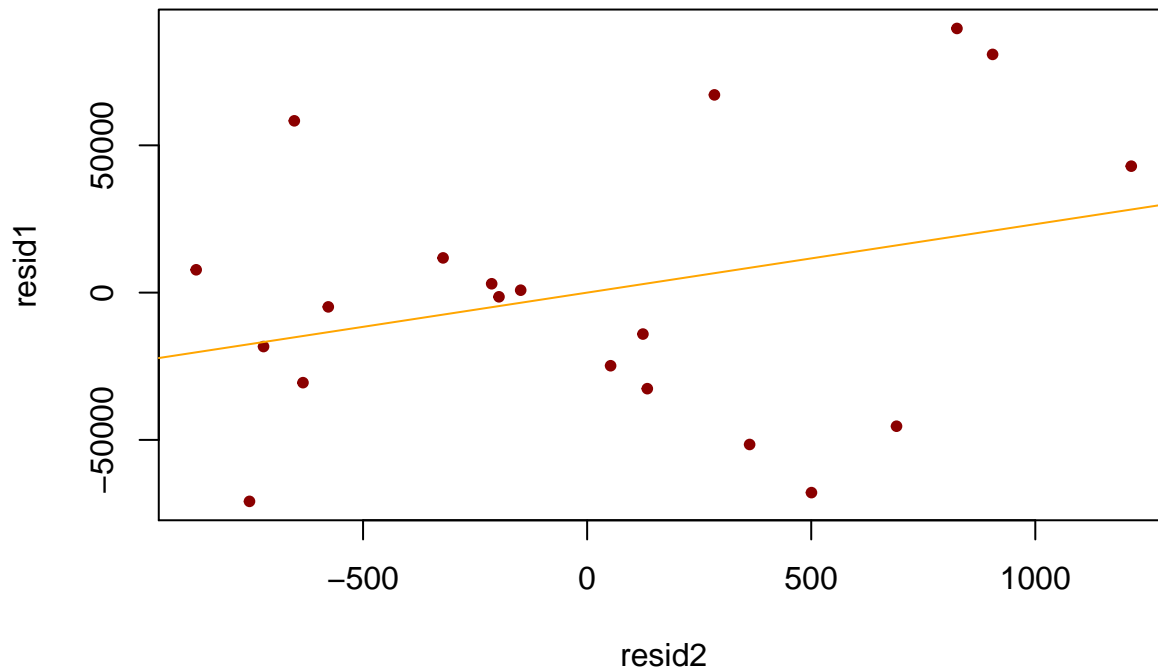
$$\widehat{resid1} = 0 + 23.232 resid2$$

```
coef(Houses.lm3)
```

```
## (Intercept)      resid2  
## 1.771783e-12  2.323240e+01
```

Added variable plot:  $resid1 \sim resid2$

```
plot(resid1 ~ resid2,  
     pch=20, col="darkred")  
abline(Houses.lm3, col="orange")
```



```
cor(resid1, resid2)
```

```
## [1] 0.3033374
```

Multiple regression.

```
Houses.lm4 <- lm(Price ~ Lot + Size, data=data)
```

Conclusion:

“Each additional square foot of *Size* corresponds to an additional \$23.23 of *Price* while controlling for *Lot* being in the model.” (Cannon, p.168)

```
options(show.signif.stars=FALSE)
summary(Houses.lm4)
```

```
##
## Call:
## lm(formula = Price ~ Lot + Size, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -79532 -28464   3713  21450  73507
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 34121.649  29716.458   1.148  0.2668
## Lot           5.657     3.075   1.839  0.0834
## Size          23.232    17.700   1.313  0.2068
##
## Residual standard error: 47400 on 17 degrees of freedom
## Multiple R-squared:  0.5571, Adjusted R-squared:  0.505
## F-statistic: 10.69 on 2 and 17 DF,  p-value: 0.000985
```