Review OLS Example: Solutions

1. `saleprice` $= 115017 + 73.77\times$ `sqfeet`.

2. OLS minimizes $\sum_{i=1}^{N} {u_i}^2 = \sum_{i=1}^{N} (y_i - \beta_1 - \beta_2 x_i)^2$
gives $\widehat{\beta}_2 = \sum_{i=1}^{N}(x_i - \bar{x})(x_i - \bar{y}) / \sum_{i=1}^{N}(x_i - \bar{x})^2$ and $\widehat{\beta}_1 = \bar{y} - \widehat{\beta}_2 \bar{x}$.

3. By \$73.77.

4. $R^2 = 0.6175$. This is quite high for cross-section data.
This is fraction of variation explained by model: $R^2 = \sum_{i=1}^{N}(\widehat{y}_i - \bar{y})^2 / \sum_{i=1}^{N}(y_i - \bar{y})^2$.

5. Yes as $p = 0.00 < 0.05$ or $t = 6.60 > t_{.025;27} =$ `invttail(27,.025)`$= 2.05$.

6. That $y_i = \beta_1 + \beta_2 x_i + u_i$ where u_i are iid $\mathcal{N}[0, \sigma^2]$.

7. $t = (\widehat{\beta}_2 - 50)/s_{\widehat{\beta}_2} = (73.77 - 50)/11.175 = 2.13$.
$p =$ `ttail(27,2.13)`$\times 2 = .021 \times 2 = .042$.
Reject $H_0 : \beta_{sqfeet} = 50$ against $H_a : \beta_{sqfeet} \neq 50$ as $p < .05$.

8. By \$68.37.

9. Yes as $p = 0.00 < 0.05$ or $t = 4.40 > t_{.025;27} =$ `invttail(22,.025)`$= 2.07$.

10. No. All have $p > 0.05$ using two-sided tests.

11. $R^2 = 0.6506$. This is not a big improvement on 0.6175 with just `sqfeet` as regressor.

12. Yes. $F = 6.83$ has $p = 0.000 < 0.05$.

13. Yes using R^2. No using $\bar{R}^2$ which falls from 0.6033 to 0.5552.
We could also do F-test on the five extra regressors (but this will lead to non-rejection of H_0 since $\bar{R}^2$ did not increase).