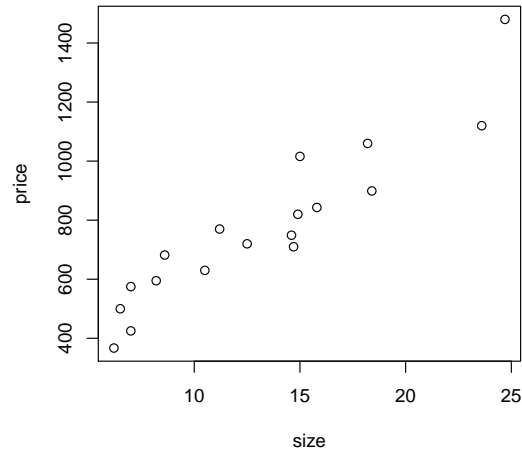


## Least Squares Fit and ANOVA Table

1. Here is a scatterplot of the sizes (in 100 ft<sup>2</sup>) and prices (in \$1000) for  $n = 18$  apartments in the Village.



Here is the Minitab output for the least squares regression fit to the housing data. Some of the entries have been redacted (replaced by question marks).

The regression equation is  
 $\text{price} = 182 + 45.0 \text{ size}$

Predictor	Coef	SE Coef	T	P
Constant	182.27	62.43	2.92	0.010
Size(100sqft)	44.950	4.370	10.29	0.000

S = 101.4    R-Sq = 86.9%    R-Sq(adj) = ?????

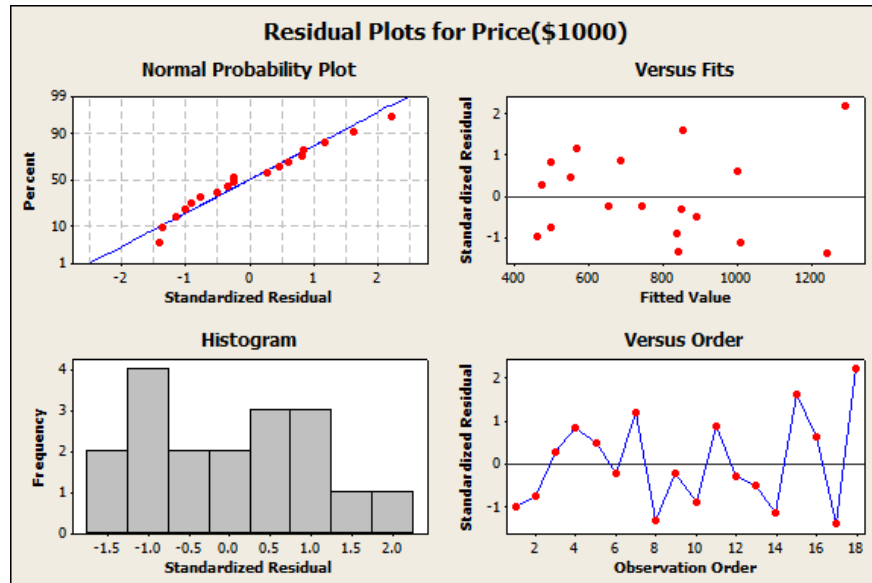
### Analysis of Variance

	DF	SS	MS	F	P
Regression	1	1087556.8	1087556.84	??????	?????
Residual Error	16	164431.4	10276.96		
Total	17	1251988.3			

- (a) Explain the meaning of the non-redacted values in the first 6 lines of the output (ending before the “Analysis of Variance” table).
- (b) Explain the meaning of the values in the Analysis of Variance table.

## Model Assumptions

2. Here are plots of the residuals from the least squares fit to the housing data.



Do the plots indicate any potential violations in assumptions? Specifically, answer the following questions.

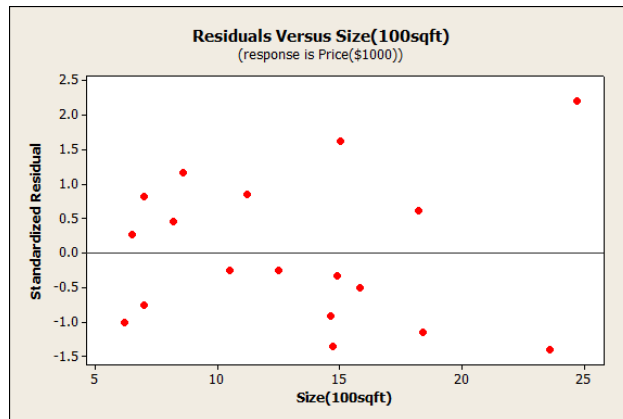
(a) Do the residual errors look approximately normal?

(b) Does the error variance look constant?

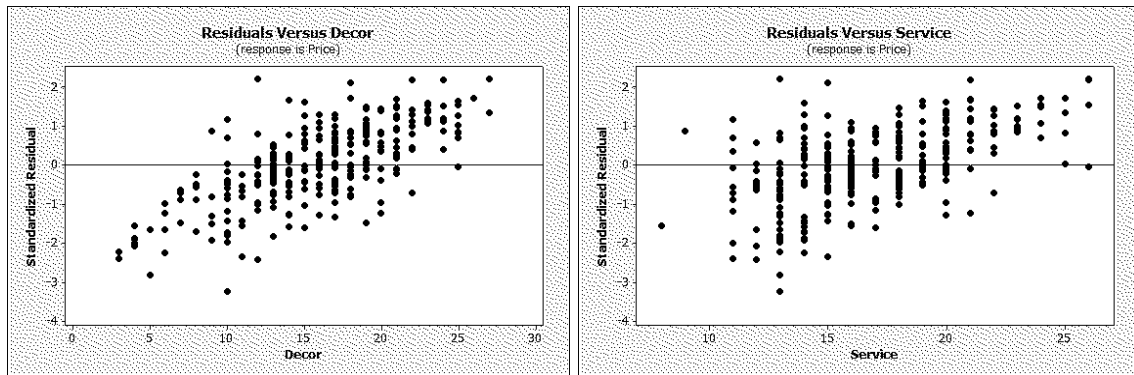
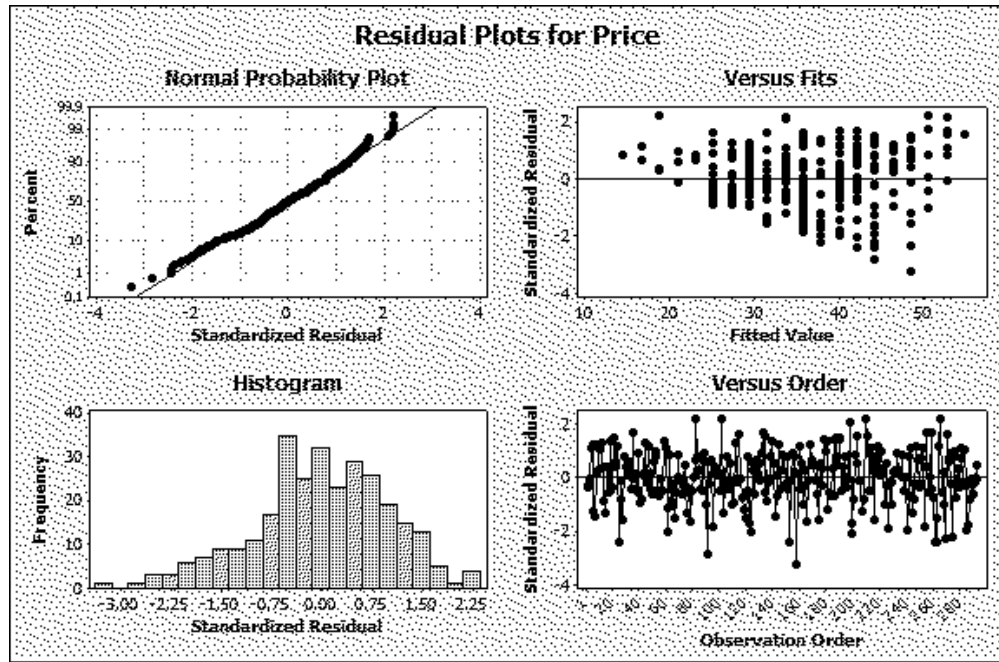
(c) Is there any apparent dependence in the residuals?

3. What is a “standardized” residual?

4. Here is a plot of the residuals versus Size ( $x$ ). Why is this plot nearly identical to the plot of residuals versus fits?



5. Here are some plots of the residuals from the fit of Price to Food for the Zagat data:



Use the plots to assess whether or not the four regression assumptions hold.