Linear Regression (Review)

1. Here is the Minitab output that result from fitting a regression model to the housing data ($n = 18$). Some of the values have been replaced by question marks.

```
Analysis of Variance

Source            DF    Adj SS    Adj MS    F-Value   P-Value
Regression         ?   1087557   ???????    ??????     ?????
  Size(100sqft)    ?   ???????   ???????    ??????     ?????
Error             ??    164431     ?????
  Lack-of-Fit     ??    ??????     ?????      ????     ?????
  Pure Error       ?     ?????     ?????
Total             ??   1251988
```

```
Model Summary

      S    R-sq   R-sq(adj)   R-sq(pred)
101.375  86.87%     ??????       ??????
```

```
Coefficients

Term             Coef   SE Coef   T-Value   P-Value    VIF
Constant        182.3      62.4      2.92     0.010
Size(100sqft)   44.95      4.37     10.29     0.000   ????
```
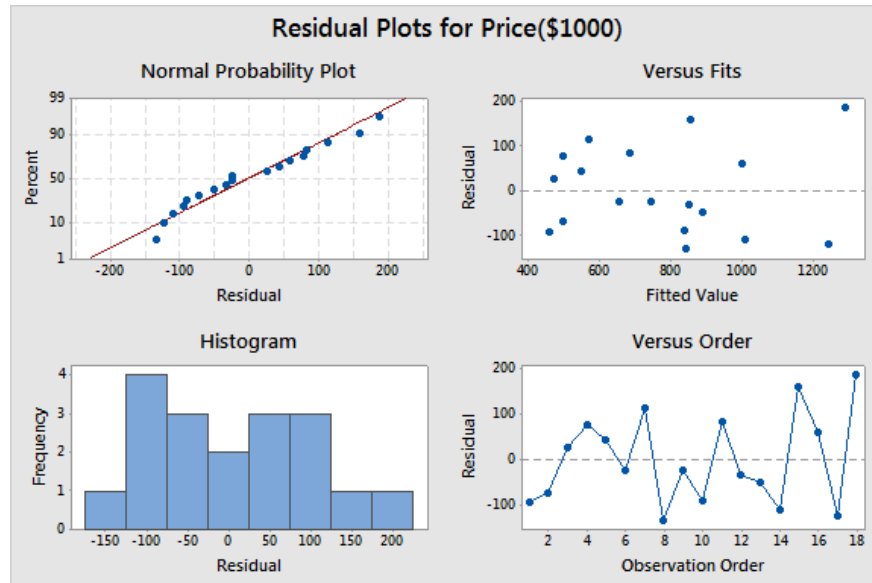
```
Regression Equation

Price($1000) = 182.3 + 44.95 Size(100sqft)
```

Explain the meanings and uses of all of the non-missing numbers in the regression output.

Model Assumptions

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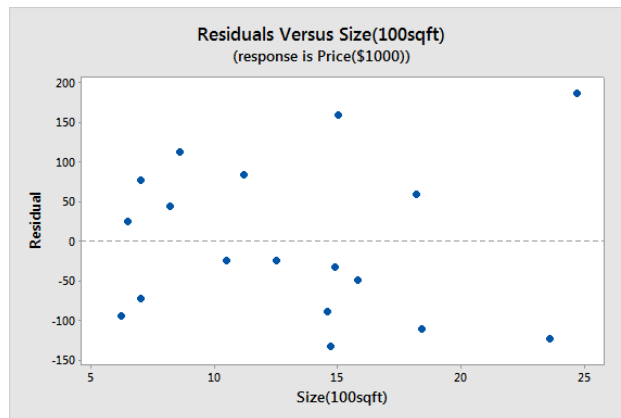

Do the plots indicate any potential violations in assumptions? Specifically, answer the following questions.

(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. Here is a plot of the residuals versus Size (x).

Residuals Versus Size(100sqft)
(response is Price($1000))

(a) Why is this plot nearly identical to the plot of residuals versus fits?

(b) Does the plot of residuals versus fit always look like the plot of residuals versus x?

4. Here are some plots of the residuals from the fit of Price to Food for the Zagat data:
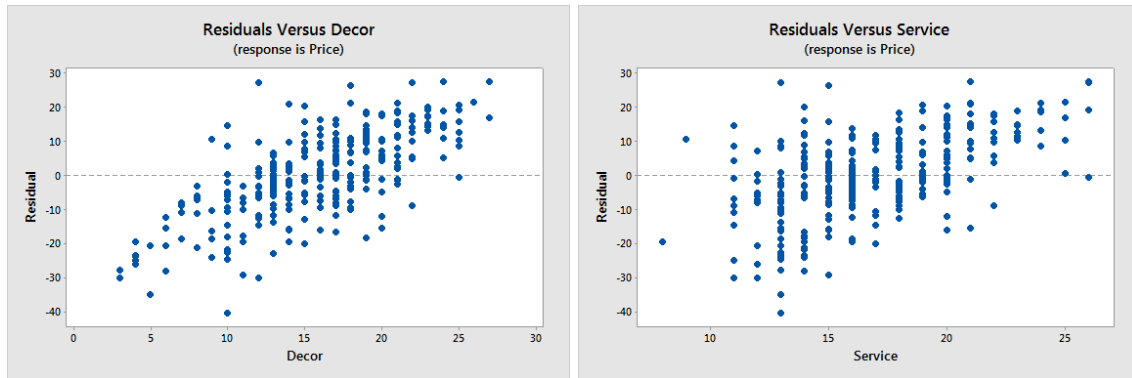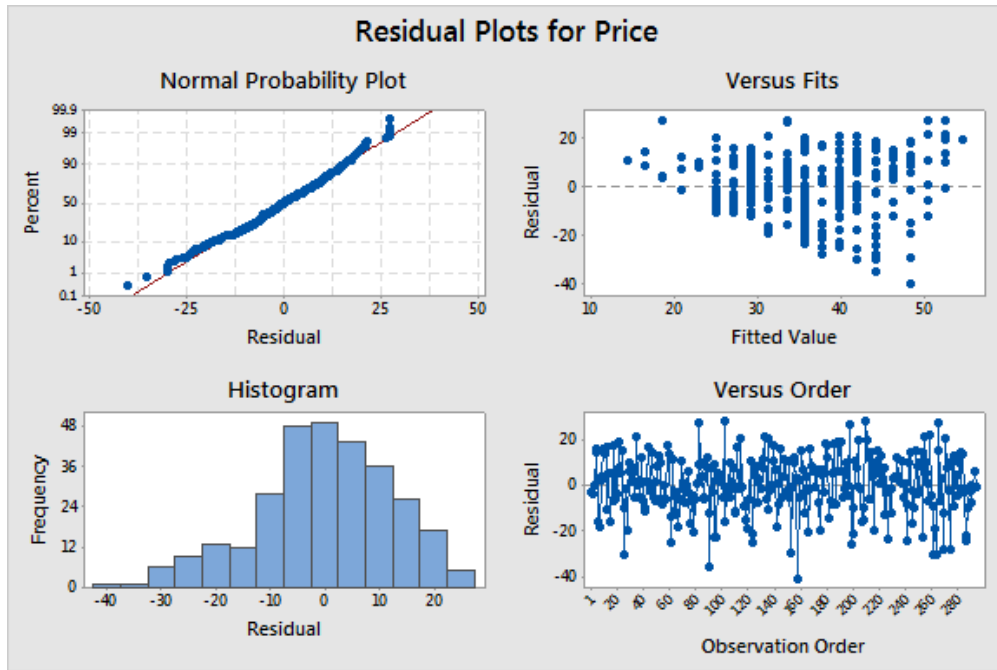
Use the plots to assess whether or not the four regression assumptions hold.