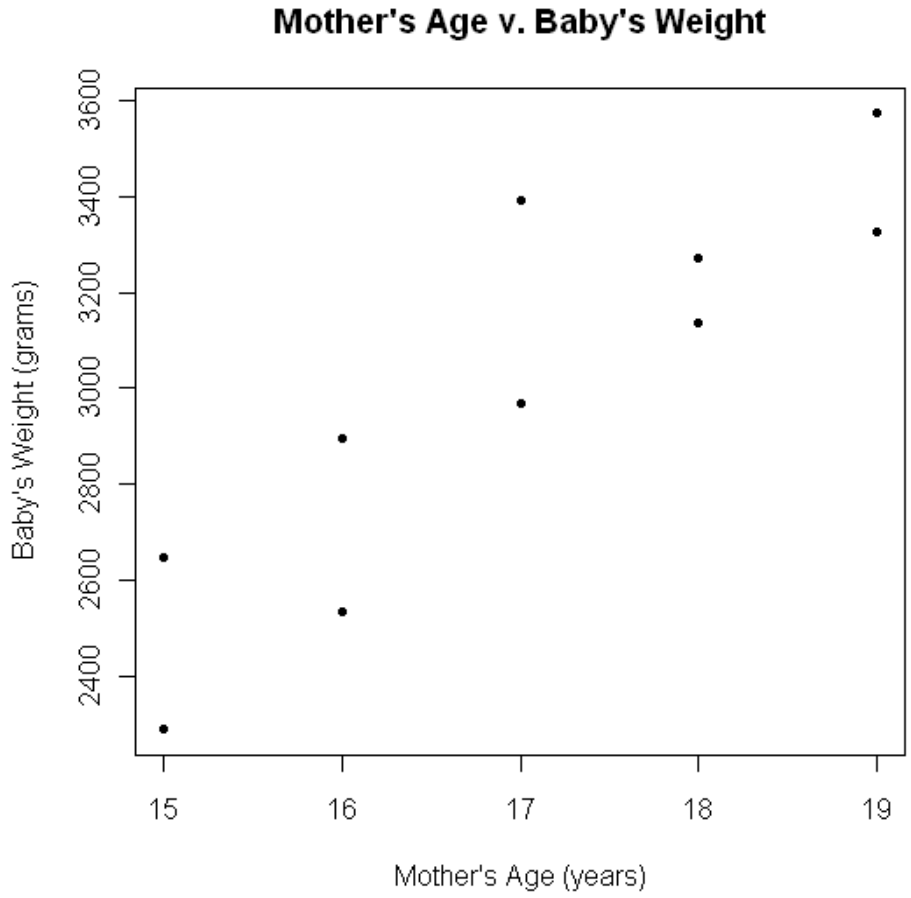
Simple Linear Regression

The following data are from a study in the *Journal of School Health.* For 10 teenage mothers, the age of the mother and the birth weight of the baby are recorded:

Age(years)	15	17	18	15	16	19	17	16	18	19
Weight(grams)	2289	3393	3271	2648	2897	3327	2970	2535	3138	3573

Below are a scatter plot of the data and the output from R.




```
Coefficients:
                Estimate   Std. Error      t value    Pr(>|t|)
(Intercept)     -1163.45   783.14          -1.486     0.175682
age             245.15     45.91           5.340      0.000694

Residual standard error: 205.3 on 8 degrees of freedom
F-statistic: 28.52 on 1 and 8 DF, p-value: 0.0006941

Analysis of Variance Table
Response: weight
           Df         Sum Sq     Mean Sq    F value    Pr(>F)
age        1          1201970    1201970    28.515     0.0006941
Residuals  8          337212     42152
```

Simple Linear Regression

1. What is the estimated regression line?

2. Interpret the slope.

3. Interpret the y-intercept. Is this interpretation useful or reasonable?

4. What is the coefficient of determination? Interpret it.

5. What is the coefficient of correlation? Interpret it.

6. What is the estimated variance of the error term, ε?

7. Calculate the first two residuals.

8. What is the estimated average baby weight if the mother is 17 years old?

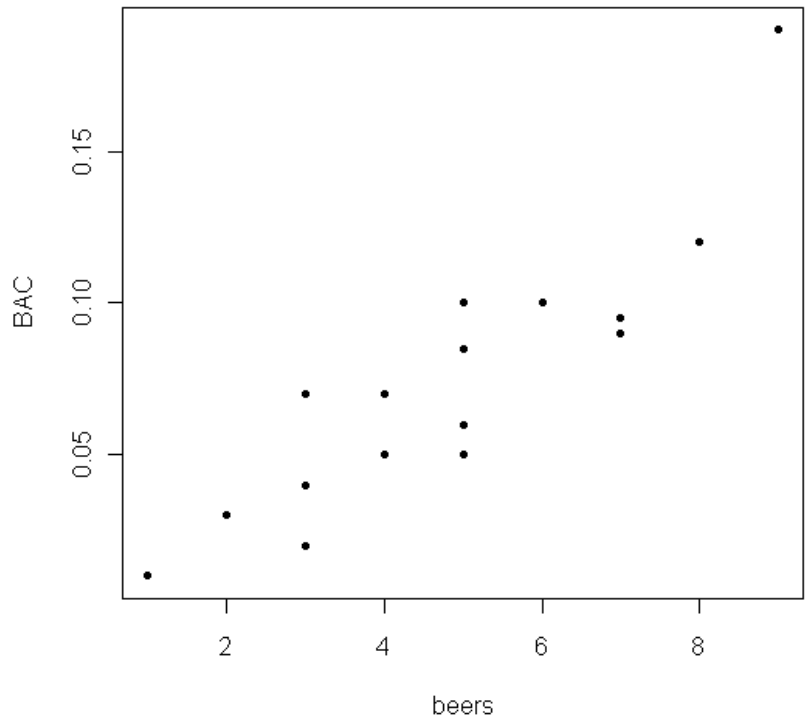
9. What is the estimated average mother's age for baby born with a weight of 3500 grams?

10. What is the estimated average baby weight if the mother is 30 years old?

Simple Linear Regression

In a study of alcohol consumption and related blood alcohol content, 16 student volunteers at Ohio State University drank a randomly assigned number of cans of beer. Thirty minutes later, a police officer measured their percent blood alcohol content (BAC):

# of Beers	BAC
5	0.10
2	0.03
9	0.19
8	0.12
3	0.04
7	0.095
3	0.07
5	0.06
3	0.02
5	0.05
4	0.07
6	0.10
5	0.085
7	0.09
1	0.01
4	0.05




```
Coefficients:
            Estimate    Std. Error  t value     Pr(>|t|)
(Intercept) -0.012701   0.012638    -1.005      0.332
beers       0.017964    0.002402    7.480       2.97e-06

Residual standard error: 0.02044 on 14 degrees of freedom
F-statistic: 55.94 on 1 and 14 DF, p-value: 2.969e-06

Analysis of Variance Table
Response: BAC
           Df    Sum Sq     Mean Sq    F value    Pr(>F)
Beers      1     0.0233753  0.0233753  55.944     2.969e-06
Residuals  14    0.0058497  0.0004178
```

11. What is the estimated regression line?

12. Interpret the slope.

13. Interpret y-intercept. Is this interpretation useful or reasonable?

14. What is the coefficient of determination? Interpret it.

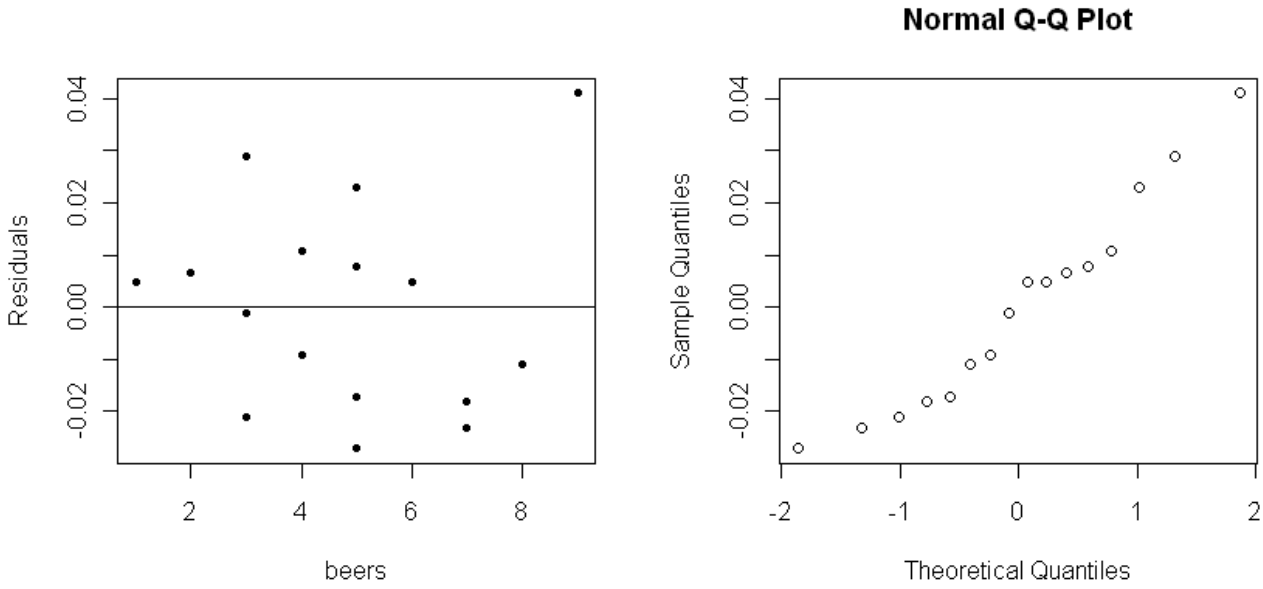
15. What is the coefficient of correlation? Interpret it.

16. What is the estimated average BAC for a student who drinks 7 beers?

17. What are the hypotheses for testing whether number of beers is associated with BAC?

18. What is the test statistic associated with the hypotheses in question 17?

19. What is the p-value associated with questions 17 and 18?



20. Are the assumptions of the simple linear regression model justified? Explain why or why not.

21. Do you see any other possible problems with the model?