7.5 Urban/rural ratings of counties. Refer to the *Professional Geographer* (February 2000) study of urban and rural counties in the western United States, Exercise 4.16 (p.190). Recall that six independent variables—total county population (x_1), population density (x_2), population concentration (x_3), population growth (x_4), proportion of county land in farms (x_5), and 5-year change in agricultural land base (x_6)—were used to model the urban/rural rating (y) of a county. Prior to running the multiple regression analysis, the researchers were concerned about possible multicollinearity in the data. The correlation matrix (shown on the next page) is a table of correlations between all pairs of the independent variables.

INDEPENDENT VARIABLE	x_1	x_2	x_3	x_4	x_5
x_1 Total population					
x_2 Population density	.20				
x_3 Population concentration	.45	.43			
x_4 Population growth	−.05	−.14	−.01		
x_5 Farm land	−.16	−.15	−.07	−.20	
x_6 Agricultural change	−.12	−.12	−.22	−.06	−.06

Source: Berry, K. A., et al. "Interpreting what is rural and urban for western U.S. counties," *Professional Geographer*, Vol. 52, No. 1, Feb. 2000 (Table 2).

(a) Based on the correlation matrix, is there any evidence of extreme multicollinearity?

Ans: No because of no strong correlation observed from the table above

(b) Refer to the multiple regression results in the table given in Exercise 4.16 (p. 190). Based on the reported tests, is there any evidence of extreme multicollinearity?

INDEPENDENT VARIABLE	β ESTIMATE	p-VALUE
x_1: Total population	0.110	0.045
x_2: Population density	0.065	0.230
x_3: Population concentration	0.540	0.000
x_4: Population growth	−0.009	0.860
x_5: Farm land	−0.150	0.003
x_6: Agricultural change	−0.027	0.580

Overall model: $R^2 = .44$ $\quad R_a^2 = .43$
$F = 32.47$ $\quad p\text{-value} < .001$

Source: Berry, K. A., et al. "Interpreting what is rural and urban for western U.S. counties," *Professional Geographer*, Vol. 52, No. 1, Feb. 2000 (Table 2).

Ans: no because we did not observe nonsignificant *t*-tests for all (or nearly all) the individual parameters.

7.10 FDA investigation of a meat-processing plant. A particular meat-processing plant slaughters steers and cuts and wraps the beef for its customers. Suppose a complaint has been filed with the Food and Drug Administration (FDA) against the processing plant. The complaint alleges that the consumer does not get all the beef from the steer he purchases. In particular, one consumer purchased a 300-pound steer but received only 150 pounds of cut and wrapped beef. To settle the complaint, the FDA collected data on the live weights and dressed weights of nine steers processed by a reputable meat-processing plant (not the firm in question). The results are listed in the table.

STEERS

LIVE WEIGHT x, pounds	DRESSED WEIGHT y, pounds
420	280
380	250
480	310
340	210
450	290
460	280
430	270
370	240
390	250

(a) Fit the model $E(y) = \beta_0 + \beta_1 x$ to the data.

```
> x<-c(420,380,480,340,450,460,430,370,390)
> y<-c(280,250,310,210,290,280,270,240,250)
> fit<-lm(y~x)
> fit

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)            x
      5.711        0.626

> summary(fit)

Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-13.656  -4.877   2.603   3.824  11.382

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.71059   26.31520   0.217    0.834
x            0.62597    0.06331   9.887 2.31e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.303 on 7 degrees of freedom
Multiple R-squared:  0.9332,    Adjusted R-squared:  0.9236
F-statistic: 97.75 on 1 and 7 DF,  p-value: 2.306e-05
```

Ans: $E(y) = 5.711 + 0.626x$

(b) Construct a 95% prediction interval for the dressed weight y of a 300-pound steer.

```
> predict(fit,newdata=data.frame(x=300),interval="prediction")
       fit      lwr      upr
1 193.5013 166.7388 220.2637
```

Ans: 95% PI (166.7388, 220.2637)

(c) Would you recommend that the FDA use the interval obtained in part (b) to determine whether the dressed weight of 150 pounds is a reasonable amount to receive from a 300-pound steer? Explain.

Ans: No, extrapolation (300 is not in the range of x's observed)