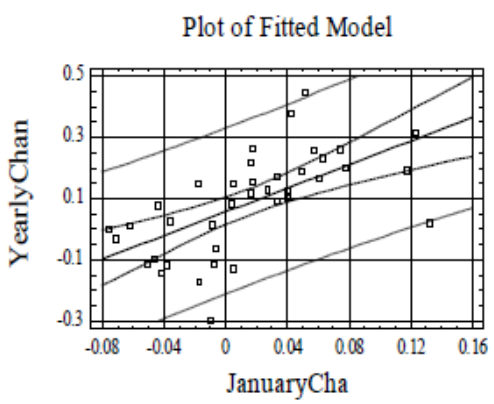
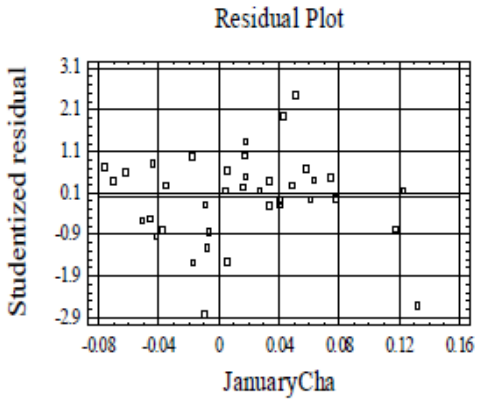
Exercise 1

It has been suggested that January is a good indicator of the behaviour of the stock market during the entire year. We ran a regression on changes in stocks in the S&P500 Index during January and the corresponding changes in the same index for the entire year for the years 1964 to 2001, with the results attached.

```
Regression Analysis - Linear model: Y = a + b*X
------------------------------------------------------------------------------
Dependent variable: YearlyChan
Independent variable: JanuaryCha
------------------------------------------------------------------------------
                              Standard          T
Parameter        Estimate         Error       Statistic        P-Value
------------------------------------------------------------------------------
Intercept       0.0570689      0.022243         2.5657          0.0146
Slope             1.92536      0.413355        4.65789          0.0000
------------------------------------------------------------------------------

Correlation Coefficient = 0.613221
R-squared = 37.604 percent
Standard Error of Est. = 0.131661
```





a) Do you believe that January can serve as a predictor of the stock market's performance? Justify your answer statistically.

b) In January 1999, the S&P500 Index increased by 4%. Predict the performance of the index for the entire year and give an approximate 95% prediction interval.

c) As a fund manager, can you use this phenomenon to design a trading strategy? Do you need to run other statistical tests?

Exercise 2

A motion picture industry analyst wants to estimate the gross earnings generated by a movie. The estimate will be based on different variables involved in the film's production. The independent variables considered are X_1 (COST) = production cost of the movie and X_2 (PROM) = total costs of all promotional activities. A third variable that the analyst wants to consider is the qualitative variable of whether or not the movie is based on a book published before the release of the movie. This third, qualitative variable is handled by the use of an indicator variable: X_3 (BOOK) = 1 if the movie is based on a book, and 0 otherwise. The analyst obtains information on a random sample of 20 Hollywood movies made within the last five years (the inference is to be made only about the population of movies in this particular category). The data are given in Table 1. The variable Y (EARN) is gross earnings, in millions of dollars. The two quantitative independent variables are also in millions of dollars.

TABLE 1. Data on Earnings

EARN	COST	PROM	BOOK
28	4.2	1	0
35	6	3	1
50	5.5	6	1
20	3.3	1	0
75	12.5	11	1
60	9.6	8	1
15	2.5	0.5	0
45	10.8	5	0
50	8.4	3	1
34	6.6	2	0
48	10.7	1	1
82	11	15	1
24	3.5	4	0
50	6.9	10	0
58	7.8	9	1
63	10.1	10	0
30	5	1	1
37	7.5	5	0
45	6.4	8	1
72	10	12	1

The regression model can be presented as:

$$\text{Predicted EARNings} = a + b_1 \times \text{COST} + b_2 \times \text{PROM} + b_3 \times \text{BOOK} + \text{error term}$$

The regression output is shown in Table 2 below.

a) How useful is the model overall?

b) Are all three independent variables relevant?

c) What gross earnings does the model predict for a movie costing nothing to produce or promote, and that is not based on a book? How meaningful is this figure?

d) Explain the meaning of the estimate $b_1 = 2.85$

e) Can you reject the hypothesis that the underlying value of $b_1 = 1$?

f) What would this hypothesis imply?

g) Compare the estimated gross earnings of a movie costing \$6m, with promotion cost of \$3m based on a book, to that of a movie with identical costs, but not based on a book. Explain the meaning of b_3.

h) An authors' association claims that the existence of a book increases gross earnings on average by at least \$7.5m. Can you reject this hypothesis?

TABLE 2. Regression Results: Earnings on COST, PROM & BOOK

Regression Statistics				
Multiple R	0.98322			
R Square	0.96671			
Adjusted R Square	0.96047			
Standard Error	3.689501			
Observations	20			
	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	7.836190	2.333380	3.3582996	0.0039967
COST	2.847692	0.392340	7.2582345	0.0000019
PROM	2.278237	0.253437	8.9893685	0.0000001
BOOK	7.166093	1.817964	3.9418244	0.0011664