A study was conducted to predict the company stock price using several variables that were taken into consideration. The list of variables is as follows;

Y: Stock price as of 9/6/91.
X1: Annual 3-S-year growth rate in sales as a percentage
X2: Total sales in millions of dollars for last four quarters.
X3: Earnings per share annual growth rate as a percentage.
X4: Profits for last four quarters.
X5: Stock price 1 year earlier.
X6: Price earnings (P/E) ratio over last four quarters
X7: Stock Market for which the company is traded,(1: if Over the Counter (OTC), 2: if New York Stock Exchange (NYSE), 3: if NASDAQ)

Minitab analysis was conducted and gave the following results;

Correlations: Y, X1, X2, X3, X4, X5, X6, X8, X9

	Y	X1	X2	X3	X4	X5	X6	X8
X1	-0.068							
	0.503							
X2	0.160	-0.090						
	0.112	0.375						
X3	0.014	0.057	-0.107					
	0.890	0.572	0.289					
X4	0.371	-0.079	0.609	0.053				
	0.000	0.437	0.000	0.601				
X5	0.330	-0.114	0.224	-0.099	0.239			
	0.001	0.257	0.025	0.326	0.017			
X6	0.294	0.025	-0.128	-0.220	-0.082	0.077		
	0.003	0.805	0.204	0.028	0.416	0.449		
X8	0.032	-0.019	-0.256	0.017	-0.143	-0.174	-0.054	
	0.753	0.855	0.010	0.868	0.156	0.084	0.591	
X9	0.000	0.025	0.268	0.150	0.187	0.114	-0.125	-0.804
	0.997	0.806	0.007	0.135	0.062	0.260	0.215	0.000

Cell Contents: Pearson correlation
P-Value

Best Subsets Regression: Y versus X1, X2, X3, X4, X5, X6, X8, X9

Response is Y

			Mallows		X	X	X	X	X	X	X	X
Vars	R-Sq	R-Sq(adj)	Cp	S	1	2	3	4	5	6	8	9
1	13.8	12.9	20.7	18.714				X				
1	10.9	10.0	24.7	19.028					X			
2	24.4	22.8	8.4	17.614				X		X		
2	19.9	18.3	14.4	18.125				X	X			
3	29.1	26.9	4.0	17.146				X	X	X		
3	25.6	23.2	8.8	17.569				X			X	X
4	31.1	28.1	3.4	16.997				X	X	X	X	
4	29.9	26.9	4.9	17.141			X	X	X	X		
5	32.1	28.5	3.9	16.952				X	X	X	X	X
5	31.8	28.2	4.3	16.990			X	X	X	X	X	
6	32.6	28.2	5.3	16.986			X	X	X	X	X	X
6	32.4	28.0	5.5	17.011		X		X	X	X	X	X
7	32.7	27.6	7.1	17.061		X	X	X	X	X	X	X
7	32.6	27.5	7.2	17.071	X		X	X	X	X	X	X
8	32.8	26.9	9.0	17.147	X	X	X	X	X	X	X	X

Stepwise Regression: Y versus X1, X2, X3, X4, X5, X6, X8, X9

Forward selection. Alpha-to-Enter: 0.05

Response is Y on 8 predictors, with N = 100

Step	1	2	3
Constant	19.94	14.74	15.39
X4	0.115	0.123	0.106
T-Value	3.96	4.50	3.85
P-Value	0.000	0.000	0.000
X6		0.164	0.153
T-Value		3.69	3.52
P-Value		0.000	0.001
X5			0.117
T-Value			2.52
P-Value			0.013
S	18.7	17.6	17.1
R-Sq	13.79	24.40	29.10
R-Sq(adj)	12.91	22.84	26.89

Regression Analysis: Y versus X1, X2, X3, X4, X5, X6, X8, X9

Predictor	Coef	SE Coef	T	P	VIF
Constant	6.488	7.500	0.86	0.389	
X1	-0.01078	0.03712	-0.29	0.772	1.024
X2	-0.001249	0.002846	-0.44	0.662	1.781
X3	0.01241	0.01815	0.68	0.496	1.170
X4	0.11505	0.03439	3.35	0.001	1.668
X5	0.13332	0.04723	2.82	0.006	1.119
X6	0.17457	0.04625	3.77	0.000	1.151
X8	11.162	6.489	1.72	0.089	3.213
X9	7.502	7.194	1.04	0.300	3.301

S = 17.1469 R-Sq = 32.8% R-Sq(adj) = 26.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	___	13051.6	1631.5	5.55	0.000
Residual Error	___	26755.4	294.0		
Total	___	39807.0			

Use the above MINITAB output to solve the following questions:

1. From the correlation analysis, the predictor that is LEAST significant to the response Y is:

a. X9
b. X4
c. X5
d. X6
e. All the variables

2. In testing the validity of the overall regression model, the numerator and denominator degrees of freedom (respectively) for the critical value of F will be:

a. 8 & 91.
b. 9 & 99.
c. 9 & 91.
d. 8 & 99.
e. None of the above.

3. From the best subsets regression analysis, According to the C-p value, number of the acceptable models

a. 10
b. 9
c. 6
d. 1
e. 7

4. From best subsets regression analysis, the number of predictors in the best regression model is:

a. 5
b. 6
c. 7
d. 3
e. 4

5. From the regression analysis of the FULL model, to test H_0 $\beta_8 \leq 0$, the P – value equal to

a. 0.0445
b. 0.089
c. 0.179
d. 0.9555
e. None of the above

6. For the standard stepwise regression analysis and using the full model analysis, calculate the C-p

a. 4.01041667
b. - 11.58805
c. 5.06547619
d. - 10.70440
e. We cannot find the c- p

$$C_p = \frac{(1-R_k^2)(n-T)}{1-R_T^2} - (n-2(k+1))$$

Where k = number of independent variables included in a particular regression model

T = total number of parameters to be estimated in the full regression model

= coefficient of multiple determination for model with k independent variables

= coefficient of multiple determination for full model with all T estimated parameters