NAME: _____
SECTION: _____

SAMIE's FINAL EXAM REVIEW- PACKAGE # 1

PLEASE PUT DOWN YOUR FINAL ANSWERS ☺. Good Luck and Study Well!

SFR101-Based on past experience, 25% of the contacts made by a firm's sales representatives result in a sale being made. Jane has contacted 100 potential customers, but has made only 15 sales. Assume that Jane's contacts represent a simple random sample of those who could have been called upon. Given this information
a. What is the sample proportion p = proportion of contacts that resulted in a sale being made? **FINAL ANSWER:**
b. For simple random samples of this size, what is the probability of finding a sample this small or smaller? **FINAL ANSWER:**

SFR102-Collecting data on traffic accident fatalities, the National Highway Traffic Safety Administration has found that 47.9% of the victims have 0.0% blood alcohol content (BAC); 11.1% of the victims have from 0.01 to 0.09 percent BAC; and 41.0% of the victims have at least 0.10 percent BAC.
For a randomly selected victim,
a. What is the probability that the victim's BAC was at least 0.01%? **FINAL ANSWER:**
b. Given that the victim had been drinking prior to the accident, what is the probability that this victim's BAC was at least 0.1%? **FINAL ANSWER:**

SFR103-The reasons given by workers for quitting their jobs generally fall into one of two categories: (1) worker quits seeking or taking a different job, or (2) worker quits withdrawing from the labor force. Economic theory suggests that wages and the quit rates are related. The table below lists quit rates (**y:** quits per 100 employees) and the average hourly wage (**x** in $) in a sample of fifteen manufacturing industries.

Industry	1	2	3	…	14	15
Average Wage (X)	8.20	10.35	6.18	…	10.93	8.80
Quit Rate (Y)	1.40	0.70	2.60	…	1.80	2.00

Assuming that a simple linear regression model is appropriate for analyzing the above data, the following results are obtained from a least squares fit of the model:

$$b_0 = 4.8615,\ \ S_{b_0} = 0.5201,\ \ b_1 = -0.3466,\ \ S_{b_1} = 0.0587$$

$$\sum x = 129.05,\ \ \sum y = 28.2,\ \ SS_{xx} = 68.6999,\ \ SS_{yy} = 11.324,\ \ SSE = 3.0733$$

a. At 5% significance level, is there any evidence to conclude that average hourly rate contributes useful information for prediction of quit rates? What does the model suggest about the relationship between quit rates and wages?
FINAL ANSWER:

b. What proportion of the variation in quit rate is accounted for by average hourly wage?
FINAL ANSWER:
c. Estimate with a 95% confidence interval the mean quit rate in an industry with an average hourly wage of $9.00. Interpret the result.
FINAL ANSWER:
d. Find a 95% prediction interval for the quit rate in an industry with an average hourly wage of $9.00. Interpret the result.
FINAL ANSWER:

SFR104-The manager at a real-estate office has provided the following probabilities for the number of houses sold per week by all three agents working at his office:

Number of houses sold	0	1	2	3	4	5
Probability	.15	.24	.31	.14	.10	.06

a. What is the mean of the distribution? **FINAL ANSWER:**
b. What is the standard deviation of the distribution? **FINAL ANSWER:**
c. If the average profit for the office from a sale of a house is $2,000, what is the expected weekly profit of the office? **FINAL ANSWER:**
d. What is the probability that the total number of houses sold in two weeks (assuming independence of results in the two weeks) is less than 3? **FINAL ANSWER:**

SFR105-In an effort to ensure the quality of incoming goods, a buyer will often do acceptance sampling. Samples of 10 items are chosen for inspection and if 2 or more are defective, the entire shipment may be returned as being below standard; otherwise the shipment is approved as being up to standard. Given the hypotheses:
Ho: shipment is below standard
Ha: shipment is up to standard, what are the consequences of a Type I and Type II error?

FINAL ANSWER:

SFR106-The length of time it takes Revenue Canada to send refunds to taxpayers after the return is received is normally distributed with a mean of 12 weeks and a standard deviation of 3 weeks.
a. What proportion of taxpayers should get sent a refund within 6 weeks.
FINAL ANSWER:
b. What proportion of refunds will be sent more than 15 weeks after the tax return has been received? **FINAL ANSWER:**
c. How long will it take before 90% of taxpayers have been sent their refunds?
FINAL ANSWER:

SFR107-A supplier of inexpensive floppy disks claims that less than 4% of the disks are defective. In a random sample of 80 disks, it is found that 10% are defective, but the supplier claims that this is only a sample fluctuation. At the 1% level of significance, test the supplier's claim. State the p-value. **FINAL ANSWER:**

SFR108-A researcher has obtained the number of hours worked per week during the summer for a sample of twelve students. The results are shown below:

40 25 35 30 22 43 35 24 38 10 8 15

a. Construct a stem-and-leaf display for the number of hours worked per week.
b. Provide the five numbers required in order to draw the boxplot.
c. Compute the mean and standard deviation of the distribution.
   MEAN:
   STANDARD DEVIATION:
d. Determine the 40th percentile of the data set. **FINAL ANSWER:**

SFR109-According to an estimate, more than four hundred thousand Canadians are fired from their jobs every year. A study by ABC Inc. identified the following reason for firing (shown in the first column):

	ABC Inc.	**Recent Sample**
Incompetence	38%	400
Negative attitude	32%	350
Dishonest	17%	150
Other	13%	100

a. Is the current distribution of reasons why employees were fired different from the results obtained by ABC Inc.? Test the hypothesis using a 5% level of significance.
REJECT Ho? YES or NO?
b. Of the 400, 350, 150 and 100 employees interviewed, respectively, 350, 320, 140 and 90 had received a warning from their employer prior to the firing. At a 10% level of significance, can it be concluded that the reason for firing is related to previous warnings given by the employer? **REJECT Ho? YES or NO?**

SFR110-The length of time required to complete a small-business income tax form by a professional tax accountant is approximately normally distributed with a mean of 4.5 hours and standard deviation of 0.6 hours. The accountant has 80 small-business tax forms to complete.
a. About how many of them will take more than 4 hours to complete? **FINAL ANSWER:**
b. At least how long would it take to complete a form that is one of the 10 most time-consuming forms to complete? **FINAL ANSWER:**
c. If a form takes longer than average to complete, what is the probability that it will take longer than 5.1 hours? **FINAL ANSWER:**

SFR111-The pharmaceutical industry is a very fast-growing segment of the U.S. and international economies. Recently, there has been controversy over how studies are done to show that drugs are both safe and effective. One drug product, Cymbalta, whish is an antidepressant, was purported in a published abstract of an article in a medical journal to be superior to other competing products. Yet the article itself stated that no studies had actually been done to show such comparisons between Cymbalta and other competing products. In an August 2005 report in an article titled " Reading Fine Print, Insurers Question Drug Studies" in the Wall Street Journal, the following data were presented showing the U.S. sales of antidepressant drugs by major brand. The sales data for the first half of 2005 are shown in the following table:

Antidepressant Drug	**Sales (First Half 2005 in Billions)**
Effexor XR	$ 1.29
Lexapro	$ 1.03
Zoloft	$ 1.55
Cymbalta	$ 0.27
Other	$ 0.97

SFR112-A multiple regression is used in cost analysis to shed light on the factors that cause costs to be incurred and the magnitudes of their effects. The sample data from a firm's accounting and production records for 20 weeks provide cost information about the firm's shipping department. LABOR (Y, in hours) is the dependent variable, while the independent variables are POUNDS shipped (X_1, in 1000s), percentage of UNITS shipped by truck (X_2, %) and average shipping WEIGHT (X_3, in pounds). The table below shows partial Excel results obtained from fitting a multiple regression model (Labor versus Pounds, Units, Weight):

	Coef	**SE Coef**	**t**	**p**
Intercept	131.92	25.69		
Pounds	2.726	2.275		
Units	0.04722	0.09335		
Weight	-2.5874	0.6428		

Analysis of Variance					
Source	**DF**	**SS**	**MS**	**F**	**P**
Regression					
Residual Error		1539.9			
Total		6698.2			

a. Is there sufficient evidence at the 5% level of significance to conclude that the model is useful in predicting labor cost? State your conclusion in the context of the problem. **REJECT Ho? YES or NO?**
b. Is there sufficient evidence at the 5% level of significance to conclude that firm's hours of labor are related to average shipping weight? What is the approximate *p*-value? **P-VALUE:**
c. Estimate the coefficient of determination and explain its meaning in the context of the problem. **FINAL ANSWER:**
d. Estimate the average hours of labor based on a shipment of 6,000 pounds, 40% of which are trucked with average shipment weight of 20 pounds. **FINAL ANSWER:**

SFR113-A scatter diagram can be used to do which of the following?
A) Determine the trend in a variable
B) Analyze the relationship between two variables
C) Describe the basic distribution for a quantitative variable
D) Show the percentage of a variable that is associated with each category into which that variable has been divided

SFR114-At Joe's Restaurant, 80 percent of the diners are new customers (*N*), while 20 percent are returning customers (*R*). Fifty percent of the new customers pay by credit card, compared with 70 percent of the regular customers. If a customer pays by credit card, what is the probability that the customer is a new customer? **FINAL ANSWER:**

SFR115-The manager for State Bank and Trust has recently examined the credit card account balances for the customers of her bank and found that 20% have an outstanding balance at the credit card limit. Suppose the manager randomly selects 15 customers and finds 4 that have balances at the limit. Assume that the properties of the binomial distribution apply. What is the probability that 4 or fewer customers in the sample will have balances at the limit of the credit card?
FINAL ANSWER:

SFR116-If a binomial distribution applies with a sample size of $n = 20$, find the expected value, $n = 20$, $p = 0.20$. **FINAL ANSWER:**

SFR117-Magic Valley Memorial Hospital administrators have recently received an internal audit report that indicates that 15% of all patient bills contain an error of one form or another. After spending considerable effort to improve the hospital's billing process, the administrators are convinced that things have improved. They believe that the new error rate is somewhere closer to 0.05.Suppose that recently the hospital randomly sampled 10 patient bills and conducted a thorough study to determine whether an error exists. It found 3 bills with errors. Assuming that managers are correct that they have improved the error rate to 0.05, what is the probability that they would find 3 or more errors in a sample of 10 bills? **FINAL ANSWER:**

SFR118-Dell Computers receives large shipments of microprocessors from Intel Corp. It must try to ensure the proportion of microprocessors that are defective is small. Suppose Dell decides to test five microprocessors out of a shipment of thousands of these microprocessors. Suppose that if at least one of the microprocessors is defective, the shipment is returned. If Intel and Dell agree that Intel will not provide more than 5% defective chips, calculate the probability that the entire shipment will be returned even though only 5% are defective. **FINAL ANSWER:**

SFR119-Last week, 108 cars received parking violations in the main university parking lot. Of these, 27 had unpaid parking tickets from a previous violation. Assuming that last week was a random sample of all parking violators, find the 95 percent confidence interval for the percentage of parking violators that have prior unpaid parking tickets.
FINAL ANSWER:

SFR120-A manufacturer of TV sets wants to find the average selling price of a particular model. A random sample of 28 different stores gives the mean sale price as $334 with a standard deviation of $16. Give a 90% confidence interval for the mean selling price of the TV model. **FINAL ANSWER:**

SFR121-A particular camera is sold with a standard deviation of $12. In a sample of 55 randomly selected stores, what is the probability that the sample mean falls within $5 of the population mean? **FINAL ANSWER:**

SFR122-A walk-in medical clinic believes that arrivals are uniformly distributed over weekdays (Monday through Friday). It has collected the following data based on a random sample of 100 days.

	Frequency
Mon	25
Tue	22
Wed	19
Thu	18
Fri	16
Total	100

Assuming that a goodness-of-fit test is to be conducted using a 0.10 level of significance, the critical value is: **FINAL ANSWER:**

SFR123-A study published in the *American Journal of Public Health* was conducted to determine whether the use of seat belts in motor vehicles depends on ethnic status in San Diego County. A sample of 792 children treated for injuries sustained from motor vehicle accidents was obtained, and each child was classified according to (1) ethnic status (Hispanic or non-Hispanic) and (2) seat belt usage (worn or not worn) during the accident. The number of children in each category is given in the table below.

	Hispanic	Non-Hispanic
Seat belts worn	31	148
Seat belts not worn	283	330

Referring to these data, which of the following conclusions should be reached if the appropriate hypothesis is conducted using an alpha = .05 level?
A) The mean value for Hispanics is the same as for Non-Hispanics.
B) There is no relationship between whether someone is Hispanic and whether they wear a seat belt.
C) The use of seat belts and whether a person is Hispanic or not is statistically related.
D) None of the above

SFR124-We want to test whether type of car owned (domestic or foreign) is independent of gender. A contingence table is obtained from a sample of 990 people as

	Gender	
Car	Yes	No
Domestic	215	400
Foreign	110	265

At alpha = 0.05 level, we conclude that:
x^2=
FINAL CONCLUSION:

SFR125-Which of the following statements is true with respect to a simple linear regression model?
A) The percent of variation in the dependent variable that is explained by the regression model is equal to the square of the correlation coefficient between the *x* and y variables.
B) If the correlation coefficient between the *x* and *y* variables is negative, the sign on the regression slope coefficient will also be negative.
C) If the correlation between the dependent and the independent variable is determined to be significant, the regression model for *y* given *x* will also be significant.
D) All of the above are true.

SFR126-Customers of the Sky Mountain Grocery chain are routinely asked at the checkout whether they prefer paper or plastic bags for their purchases. In a recent study, researchers observed the type of bag specified and surveyed the customer for other information, including his or her level of education. For the 175 persons in the sample, bag selection and education levels were shown below. At the 0.01 level, is bag selection independent of education level? Based on the chi-square table, what is the most accurate statement that can be made about the p-value for the test?

Education Level

		High School	**Some College**	**College Grad**	**Graduate Study**	
Bag Selection	**Paper**	14	13	34	2	63
	Plastic	17	19	19	3	58
No Preference		8	28	13	5	54
		39	60	66		175

FINAL CONCLUSION:

SFR127-Use the following regression results to answer the question below.

Regression Statistics	
Multiple R	0.8851
R Square	0.7835
Adjusted R Square	0.7474
Standard Error	5.4006
Observations	8

ANOVA

	df	*SS*	*MS*	*F*
Regression	1	633.242	633.242	21.711
Residual	6	175.000	29.167	
Total	7	808.242		

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	5.93118	4.17721	1.41989	0.20545
Total Bill	−2.71551	0.58279	−4.65952	0.00347

Which of the following is true?
A) *x* explains about 88.5 percent of the variation in *y*.
B) *y* explains about 88.5 percent of the variation in *x*.
C) *x* explains about 78.4 percent of the variation in *y*.

D) *y* explains about 78.4 percent of the variation in *x*.

SFR128-A recent study by a major financial investment company was interested in determining whether the annual percentage change in stock price for companies is linearly related to the annual percent change in profits for the company. The following data was determined for 7 randomly selected companies:

% Change Stock Price	% Change in Profit
8.4	4.2
9.5	5.6
13.6	11.2
-3.2	4.5
7	12.2
18.4	12
-2.1	-13.4

Based upon this sample information, which of the following is the regression equation?
A) $\hat{y}$ = 4.19 + .61*x*
B) $\hat{y}$ = 15.04 + 4.25*x*
C) $\hat{y}$ = 1.19 - 3.00*x*
D) $\hat{y}$ = 20.19 + .005*x*

SFR129-Consider the following partially completed computer printout for a regression analysis where the dependent variable is the price of a personal computer and the independent variable is the size of the hard drive.

SUMMARY OUTPUT					
Regression Statistics					
Multiple R	0.819361805				
R Square					
Adjusted R Square	0.661687702				
Standard Error					
Observations	36				
ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	33116034.84	33116034.84		
Residual		16211214.72			
Total	35	49327249.56			
	Coefficients	*Standard Error*	*t Stat*	*P-value*	
Intercept	50.84102383	246.9869514	0.205844979	0.838139607	
Hard Drive Capacity	217.7539792	26.12854674		9.95844E-10	

Based on the information provided, what is the F statistic? **FINAL ANSWER:**

SFR130-The editors of a national automotive magazine recently studied 30 different automobiles sold in the United States with the intent of seeing whether they could develop a multiple regression model to explain the variation in highway miles per gallon. A number of different independent variables were collected. The following regression output (with some values missing) was recently presented to the editors by the magazine's analysts:

Regression Statistics						
Multiple R	0.906876					
R Square						
Adjusted R Square						
Standard Error						
Observations	30					
ANOVA						
	df	*SS*	*MS*	*F*		
Regression						
Residual		89.88341785				
Total		506.1666667				
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	41.77425	5.612843566				
Curb Weight	-0.00608	0.00145286				
cylinders	0.959134	0.887531206				
Horse Power	0.052399	0.023938405				
Torque	-0.01747	0.024737052				
0 to 60 mph	0.492998	0.486960357				
Price as Tested	-0.00015	4.31238E-05				
Displacement	-2.11223	0.88675014				

Based on this output and your understanding of multiple regression analysis, what is the critical value for testing the significance of the overall regression model at a 0.05 level of statistical significance? **FINAL ANSWER:**

SFR131-Which of the following is not an indication of potential multicollinearity problems?
A) The sign on the standard error of the estimate is positive.
B) A sign on a regression slope coefficient is negative when the sign on the correlation coefficient was positive.
C) The standard error of the estimate increases when a variable enters the model in the presence of other independent variables.
D) An independent variable goes from being statistically significant to being insignificant when a new variable is added to the model.

SFR132-The Fitness Center manager has collected data on the number of visits to the club each week for the past 8 weeks. These data are shown as follows. Which of the following statements is most correct?

Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
1415	1623	1934	1879	2102	2156	2511	2499

A) The proper graph for displaying these data is a pie chart.
B) There has been a gradual downward trend in these data.
C) A frequency histogram should be developed to help identify the trend in these data.
D) The data lend themselves to a line chart.

COMM 215 Business Statistics
List of formula provided in the Final Examination

Chapter 3 Describing Data

Sample mean: $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$

Sample variance:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{1}{n-1}\left[\sum x^2 - \frac{(\sum x)^2}{n}\right]$$

Sample standard deviation: $s = \sqrt{s^2}$

Z score: $z = \dfrac{x - \text{mean}}{\text{standard deviation}}$

Coefficient of variation: $\dfrac{\text{standard deviation}}{\text{mean}} \times 100\ \%$

Chapter 4 Probability

The rule of complement: $P(\bar{E}) = 1 - P(E)$

The addition rule for two events:

$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) - P(E_1 \text{ and } E_2)$$

Conditional probability: $P(E_1 \mid E_2) = \dfrac{P(E_1 \text{ and } E_2)}{P(E_2)}$

The general multiplication rule:

$$P(E_1 \text{ and } E_2) = P(E_2)\,P(E_1 \mid E_2)$$

Chapter 5 Discrete Probability Distributions

Mean (expected value) of a discrete random variable

$$E(x) = \sum xP(x)$$

Variance and standard deviation of a discrete random variable

$$\sigma_x^2 = \sum [x - E(x)]^2 P(x) \qquad \sigma_x = \sqrt{\sigma_x^2}$$

Binomial probability formula

$$p(x) = \frac{n!}{x!(n-x)!}\, p^x q^{n-x}$$

Mean, variance, and standard deviation of a binomial random variable

$$\mu_x = np,\ \sigma_x^2 = npq, \text{ and } \sigma_x = \sqrt{npq}$$

Chapter 6 Continuous Probability Distribution

Standard normal random variable: $z = \dfrac{x - \mu}{\sigma}$

Chapter 7 Sampling distribution

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Sample proportion: $\bar{p} = \dfrac{x}{n}$

$$\mu_{\bar{p}} = p$$

Standard error of $\bar{p}$:

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$$

Chapter 8 Estimating Single Population Parameters

A *z*-based confidence interval for a population mean μ with σ known:

Margin of error: $e = z\dfrac{\sigma}{\sqrt{n}}$

Confidence interval = $\bar{x} \pm z\dfrac{\sigma}{\sqrt{n}}$

A *t*-based confidence interval for a population mean μ with σ unknown: $\bar{x} \pm t\dfrac{s}{\sqrt{n}}$

Confidence interval for the proportion: $\bar{p} \pm z\sqrt{\dfrac{\bar{p}(1-\bar{p})}{n}}$

Chapter 9 Hypothesis Testing

$$z - \textit{Test for mean}\quad z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$t - \textit{Test for mean}\quad t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

$$z - \textit{Test for proportion}\quad z = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

Chapter 13 Goodness-of-Fit Tests

Chi-square goodness-of-fit test statistic

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

Chi-square contingency test statistic

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \text{ with } df = (r-1)(c-1)$$

Chapter 14 Simple Linear Regression and Correlation Analysis

Sample correlation coefficient

$$r = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\left[\sum (x-\bar{x})^2\right]\left[\sum (y-\bar{y})^2\right]}}$$
$$= \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n(\sum x^2) - (\sum x)^2\right]\left[n(\sum y^2) - (\sum y)^2\right]}}$$

Simple linear regression model: $y = \beta_0 + \beta_1 x + \varepsilon$

Least squares point estimates

$$b_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

Sum of squared residuals (Sum of squares error)

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum y^2 - b_0 \sum y - b_1 \sum xy$$

Total sum of squares: $SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$

Sum of squares regression: $SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$

Simple regression estimator for the standard error of the estimate: $s_\varepsilon = \sqrt{\frac{SSE}{n-2}}$

Coefficient of determination: $R^2 = r^2 = \frac{SSR}{SST}$

F test for the simple linear regression model:

$$F = \frac{\frac{SSR}{1}}{\frac{SSE}{(n-2)}} \qquad df = (D_1 = 1, D_2 = n-2)$$

Simple regression estimator for the standard error of the slope: $s_{b_1} = \frac{s_\varepsilon}{\sqrt{\sum (x-\bar{x})^2}} = \frac{s_\varepsilon}{\sqrt{\sum x^2 - \frac{(\sum x)^2}{n}}}$

Test of hypothesis for slope: $t = \frac{b_1 - \beta_1}{s_{b_1}} \qquad df = n-2$

Confidence interval for $E(y) \mid x_p$:

$$\hat{y} \pm t s_\varepsilon \sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum (x-\bar{x})^2}}$$

Prediction interval for $y \mid x_p$: $\hat{y} \pm t s_\varepsilon \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum (x-\bar{x})^2}}$

Chapter 15 Multiple regression

The multiple regression model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon$$

Standard error: $s_\varepsilon = \sqrt{\frac{SSE}{n-k-1}} = \sqrt{MSE}$

Multiple coefficient of determination: $R^2 = \frac{SSR}{SST}$

An F test for the linear regression model:

$$F = \frac{\frac{SSR}{k}}{\frac{SSE}{n-k-1}} \qquad df = (D_1 = k, D_2 = n-k-1)$$

Testing the significance of each regression coefficient:

$$t = \frac{b_j - 0}{s_{b_j}} \qquad df = n-k-1$$