Practice Exam 3
(Answers are found on the last page.)

Name:_____

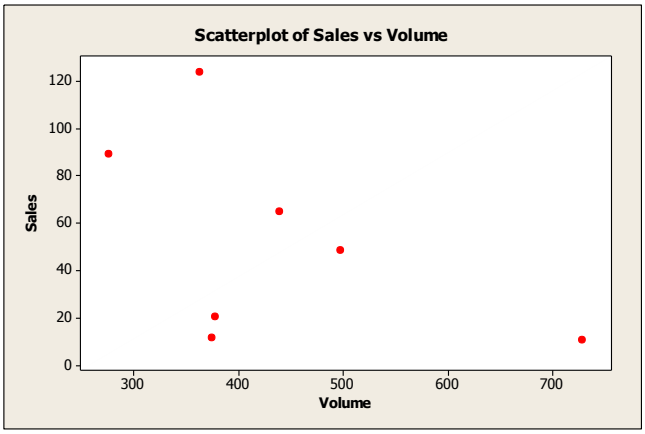
Instructions:

1. Do not start until instructed to do so.
2. You may use a scientific calculator (no graphing calculators allowed).
3. No other aids are allowed.
4. The work you turn in must be your own.
5. Use $\alpha = .05$, unless otherwise specified.
6. SHOW ALL WORK to receive full credit.

Questions 1 - 5: Shown below are data and analyses on annual sales ($ billions) and annual stock volume (millions of shares) traded for a random sample of large companies.

Consider the model: $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

Company	Sales	Volume
Merck	10.5	728.6
Altria	48.1	497.9
IBM	64.8	439.1
Eastman Kodak	20.1	377.9
Bristol-Myers Squibb	11.4	375.5
General Motors	123.8	363.8
Ford Motors	89.0	276.3



$\bar{x} = 437.0142857$
$\bar{y} = 52.52857143$
$SS_{yy} = 11{,}089.5543$
$SS_{xy} = -19{,}131.55286$
$SS_{xx} = 127{,}201.5686$

Predictor	Coef	SE Coef	T	P
Constant		51.97	2.28	0.072
Volume		0.1136	-1.32	0.243

S = 40.5268 R-Sq = 25.9% R-Sq(adj) = 11.1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	2877.5	2877.45	1.75	0.243
Error	5	8212.1	1642.42		
Total	6	11089.6			

1. **5 points** Find the equation of the least-squares regression line.

2. **2 points** Carefully draw this line on the scatterplot.

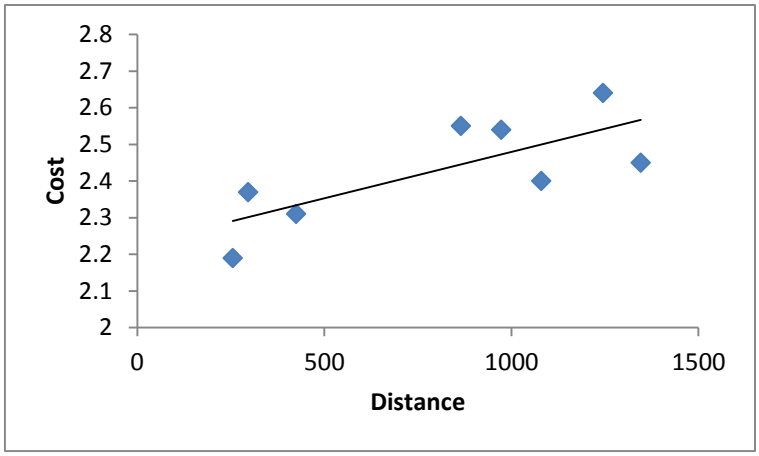
3. **6 points** Is there enough evidence of a linear relationship between sales and volume for all large companies? Show the hypotheses, the value of the test statistic, the rejection region, the p-value, and say either "yes" or "no."

4. **5 points** Construct a 95% confidence interval for β_1. Does this result agree with your result of the test above? Explain.

5. **3 points** Find the value of r, the correlation coefficient.

Questions 6 – 8: Wisconsin is an important milk-producing state. Some people might argue that because of transportation costs, the cost of milk increases with the distance (in miles) of markets from Wisconsin. A random sample of milk prices (in dollars per gallon) from 8 cities is shown below along with some regression summaries.

Distance	Cost
1245	2.64
425	2.31
1346	2.45
973	2.54
255	2.19
865	2.55
1080	2.4
296	2.37



$\bar{x} = 810.625$
$SS_{xx} = 1{,}299{,}457.875$

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.083162279	0.083162	7.757046	0.031780344
Residual	6	0.064325221	0.010721		
Total	7	0.1474875			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	2.226180023	0.082228118	27.07322	1.68E-07	2.024975067	2.427385
Distance	0.000252978	9.0831E-05	2.785147	0.03178	3.07223E-05	0.000475

6. **3 points** Interpret the value of $\hat{\beta}_1$ in the context of the problem.

7. **3 points** Find the value of s, the standard deviation of the regression.

8. **5 points** Find a 95% interval estimate for the mean price of a gallon of milk in all cities that are 1000 miles from Wisconsin.

9. Suppose female entrepreneurs, when asked to state their personal definition of success in terms of several categories, select them in the proportions shown below (Female %). In order to compare the population of males to this distribution, a random sample of male entrepreneurs were given a similar survey and the numbers of males who responded in each category are also shown below (Male Counts).

Definition of Success	Female %	Male Counts
Happiness	39	42
Sales/profit	12	95
Helping others	18	27
Achievements/challenge	31	63

10 points Conduct a test to see whether the distribution of how male entrepreneurs define success differs from that for females. Show your hypotheses, compute the value of the test statistic, give either the rejection region or the p-value, and state your conclusion in terms of the problem.

Questions 10 – 12: A random sample of 30-year-olds is interviewed to see if there is evidence that music preference is related to geographic region. The data are shown below.

	Classical	Country	R&B	Rock	Total
Northeast	18	5	32	140	195
South	8	52	41	134	235
West	13	8	27	154	202
Total	39	65	100	428	632

10. **2 points** What percentage of southerners in the sample listen to R&B? What percentage of those in the sample living the northeast listen to R&B?

11. **2 points** If music preference is not related to geographic region, how many westerners in the sample would we have expected to listen to classical?

12. **4 points** In the test of

    H_0: Music preference and geographic region are not related for all 30-year-olds
    H_a: Music preference is related to geographic region for all 30-year-olds

    $x^2 = 64.920$

    Give the rejection region for this test and make a conclusion in terms of the problem.

Answers

1. $\hat{y} = 118.2570 - .1504x$

2. Line goes through $(300, 73.1)$ and $(700, 13.0)$.

3. H_0: $\beta_1 = 0$ H_a: $\beta_1 \neq 0$

   $t = -1.32$ RR: $|t| > 2.571$ $p - value = .243$ No.

4. $(-.4425, .1417)$
   This agrees with the result of the hypothesis test because the interval does not lie entirely in the space of the alternative hypothesis; 0 is included in the interval.

5. $-.5094$

6. A 1 mile increase in distance from Wisconsin is associated with an estimated \$0.00025 increase in price per gallon of milk.

7. $.1035$

8. (\$2.386 , \$2.572)

9. H_0: $p_{Happiness} = .39,\ p_{Sales} = .12, p_{Help} = .18, p_{Achieve} = .31$
   H_a: $not\ H_0; at\ least\ one\ proportion\ differs\ from\ the\ specified\ values$

   $x^2 = 198.483$ RR: $x^2 > 7.81473$ $p - value < .005$

   We have enough evidence that the distribution of how male entrepreneurs define success differs from that for females.

10. 17.45%; 16.41%

11. 12.465

12. $x^2 > 12.5916$; We have enough evidence that music preference is related to geographic region for all 30-year-olds.