Some Sample Problems for the Final Exam

These are only a *few* suggestions to *help* you prepare for the exam.

1. You should completely understand the questions and ideas in Exam 1 & 2 and the Sample Problems for the exams.

2. You should also be certain that you completely understand the Problems Sets, Reading Assignments, in-class work, and your class notes since Exam 2.

3. See onCourse for a description of the 1 page essay that is due at the beginning of the exam.

4. The National Oceanic and Atmospheric Administration publishes temperature information from cities around the world. The scatterplot in #1 shows the average high and average low temperatures in January for 50 randomly selected cities. #2 shows some additional Minitab output for this data.

   (a) Use Graph #1 to describe the association between the average low and high temperatures.

   (b) How much of the variation in average high temperatures in January is explained by the linear model?

   (c) For the city with average low temperature 29, is the residual positive or negative? How about for the city with average low temperature of 65?

   (d) What average high temperature does this model predict for a city with average low in January of 60 degrees?

5. A household-appliance manufacturer wants to analyze the relationship between total sales and the company's three primary means of advertising: television, internet, and radio. A random sample of 10 sales periods is selected, where all data are in millions of dollars. #3 gives the output for a multiple regression analysis on the data using the variables television, internet, and radio advertising expenditures as predictor variables for sales.

   (a) Use the regression equation to predict total sales if the amounts spent on television, internet, and radio advertising are \$9.5 million, \$4.3 million, and \$5.2 million, respectively.

   (b) How could you improve this model? Explain. What is the practical significance of your proposal?

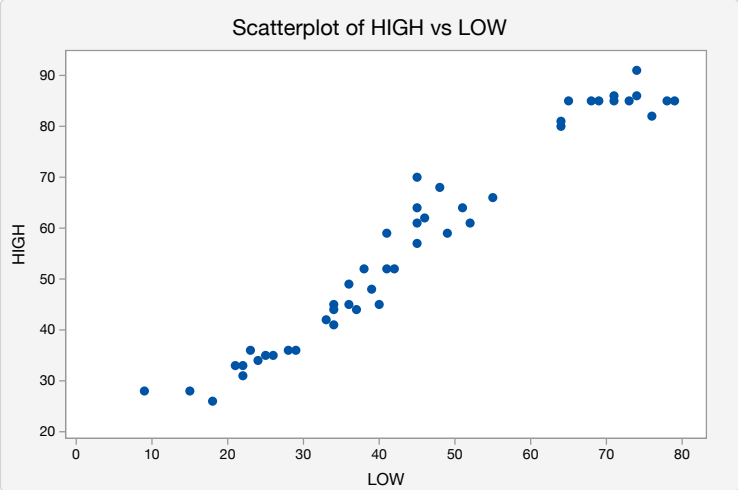
6. The U.S. Census Bureau compiles information on income of people by type of residence. The following data came from Independent simple random samples of people residing inside principal cities (IPC), outside principal cities but within metropolitan areas (OPC), and outside of metropolitan areas (OMA).

Income Level	IPC	OPC	OMA
Under \$5000	30	49	18
\$5000 – \$9999	36	45	20
\$10,000 – \$14,999	41	57	27
\$15,000 – \$24,999	82	122	46
\$25,000 – \$34,999	69	108	41
\$35,000 – \$49,999	73	126	40
\$50,000 – \$74,999	67	135	34
\$75,000 & over	68	146	20

   Does the data support that the location of residence varies based on income level? Test an appropriate hypothesis and clearly state your conclusion.

   If you determine that location does vary, which cells contribute the most to your conclusion? What does this mean in the context of this data?

#1



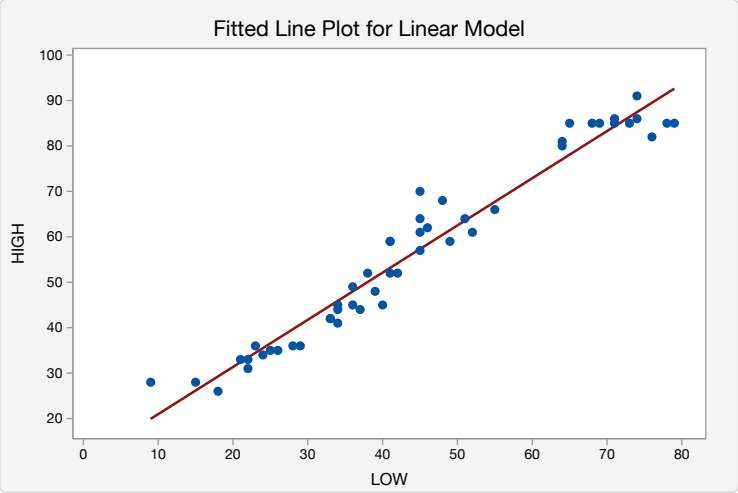
Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
4.41794	95.22%	95.12%	94.77%

Regression Equation

HIGH = 10.584 + 1.03850 LOW

#2



#3

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
4.41791	91.07%	86.61%	73.51%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	266.23	16.34	16.29	<0.0001	
TV	6.727	1.344	5.01	0.0024	1.37
INTERNET	3.257	1.642	1.98	0.0946	1.38
RADIO	4.507	3.703	1.22	0.2692	1.80

Regression Equation

SALES = 266.23 + 6.727 TV + 3.257 INTERNET + 4.507 RADIO

#4

Tabulated Statistics: Income Level, Worksheet columns

Rows: Income Level Columns: Worksheet columns

	IPC	OPC	OMA	All
Under $5000	30	49	18	97
	30.13	50.96	15.91	
	0.00	0.08	0.28	
$5000 - $9999	36	45	20	101
	31.38	53.06	16.56	
	0.68	1.22	0.71	
$10,000 - $14,999	41	57	27	125
	38.83	65.67	20.50	
	0.12	1.14	2.06	
$15,000 - $24,999	82	122	46	250
	77.67	131.33	41.00	
	0.24	0.66	0.61	
$25,000 - $34,999	69	108	41	218
	67.73	114.52	35.75	
	0.02	0.37	0.77	
$35,000 - $49,999	73	126	40	239
	74.25	125.55	39.20	
	0.02	0.00	0.02	
$50,000 - $74,999	67	135	34	236
	73.32	123.98	38.70	
	0.54	0.98	0.57	
$75,000 & over	68	146	20	234
	72.70	122.93	38.38	
	0.30	4.33	8.80	
All	466	788	246	1500

Cell Contents: *Count*
Expected count
Contribution to Chi-square

Chi-Square Test

	Chi-Square	DF	P-Value
Pearson	24.54	14	0.0394