

QUESTIONS FOR REVIEW AND CRITICAL THINKING

- What is covariance?
- How are covariance and correlation different?
- How does a researcher determine if a correlation coefficient is significant?
- The management of a regional bus line thought the company's cost of gas might be correlated with its passenger/mile ratio. The data and a correlation matrix follow. Comment.

Total Gas		
Year	Expenditures	Passengers/Miles
1	56.5	8.37
2	59.4	8.93
3	63.0	9.15
4	65.6	9.79
5	89.0	11.20

	Year	Expenditures	Miles
Year (r)	1.00000	0.87016	0.95127
p-value	0.00000	0.05510	0.01280
Price (r)	0.87016	1.00000	0.97309
p-value	0.05510	0.00000	0.00530
Mile (r)	0.95127	0.97309	1.00000
p-value	0.01280	0.00530	0.00000

Interpret the following data:

- $\hat{Y} = 5.0 + .30X_1$
Where the dependent variable equals turnover intentions for line managers and the independent variable equals number of employees supervised.
 - $\hat{Y} = 250 - 4.0X_1$
Where the dependent variable is the number of hits on a new banner ad and the independent variable is the number of weeks the ad has run.
- What are some different terms used to refer to the slope coefficient estimated in regression analysis?
 - The following ANOVA summary table is the result of a regression of sales on year of sales. Is the relationship statistically significant at the 0.05 significance level? Fill in the value for Sum of Squares in the SST row. Comment.

Source of Variation	Sum of Squares	d.f.	Mean Square	F-Value	p-Value
SSR	605,370,750	1	605,370,750	3.12	0.115
SSE	1,551,381,712	8	193,922,714		
SST		9			

- Address the following questions about regression analysis:
 - Define simple linear regression.
 - When is it most appropriate to rely on raw parameter coefficients and when is it most appropriate to rely on standardized parameter coefficients?

- Why is the Y-intercept estimate equal to 0 for estimates?

- What are the steps in interpreting a regression model?
- The following table gives a basketball team's season percentage of games won, and number of active players years 2002–2011.

Year	Season-Ticket Sales	Percentage of Games Won	Number of Active Players
2002	4,995	40	30
2003	8,599	54	34
2004	8,479	55	32
2005	8,419	58	31
2006	10,253	63	35
2007	12,457	75	37
2008	13,285	36	38
2009	14,177	27	39
2010	15,730	63	40
2011	15,805	70	41

- Compute a correlation matrix for the variables. A statistical package is recommended. Interpret the correlation between each pair of variables.
 - Estimate a regression model for sales = percentage of games won.
 - Estimate a regression model for sales = number of active alumni.
 - If sales is the dependent variable, which of the two independent variables do you think explains sales better? Explain.
- Are the different forms of consumer installment credit in the following table highly correlated?

Debt Outstanding (millions of dollars)						
Year	Gas Cards	Travel and Entertainment Cards	Bank Credit Cards	Retail Cards	Total Credit Cards	Total Installment Credit
1	\$39	\$61	\$828	\$9,400	\$11,228	\$79,428
2	1,119	76	1,312	10,200	12,707	87,745
3	1,298	110	2,639	10,900	14,947	98,105
4	1,650	122	3,792	11,500	17,064	102,064
5	1,804	132	4,490	13,925	20,351	111,225
6	1,762	164	5,408	14,763	22,097	127,332
7	1,832	191	6,838	16,395	25,256	140,458
8	1,823	238	8,281	17,933	28,275	156,088
9	1,893	273	9,501	18,002	29,569	164,955
10	1,981	238	11,351	19,052	32,622	185,489
11	2,074	284	14,262	21,082	37,702	216,572

- b. Using the same website, record how many days in January are typically sunny. Test whether or not the number of sunny days meets your standard.
- c. For each location, record whether or not there was measurable precipitation yesterday. Test the following hypothesis:

H_1 : Among places you would like to live, there is less than a 33.3 percent chance of rain/snow on a given day (five days out of fifteen).

2. **ETHICS** Examine the statistical choices under Analyze in SPSS. Click on Compare Means. To compare an observed mean to some benchmark or hypothesized population mean, the available choice is a one-sample t -test. A researcher is preparing a report and finds the following result testing a hypothesis that suggests the sample mean did not equal 14.

- a. What is the p -value? Is the hypothesis supported?
- b. Write the 95% confidence interval which corresponds to an α of 0.05.
- c. Technically, since the sample size is greater than 30, a Z -test might be more appropriate. However, since the t -test result is readily available with SPSS, the research presents this result. Is there an ethical problem in using the one-sample t -test?

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
1997-2000	67	14.5337	16.02663	1.95796

Test Value = 14						
				95% Confidence Interval of the Difference		
	t	df	Sig. (two-tailed)	Mean Difference	Lower	Upper
1997-2000	0.273	66	0.786	0.53373	-3.3755	4.4429

Premier Motorcars

CASE 21.1

Premier Motorcars is the new Fiat dealer in Delavan, Illinois. Premier Motorcars has been regularly advertising in its local market area that the new Fiat 500 averages 30 miles to a gallon of gas and mentions that this figure may vary with driving conditions. A local consumer group wishes to verify the advertising claim. To do so, it selects a sample of recent purchasers of the Fiat 500. It asks them to drive their cars until two tanks of gasoline have been used up and to record the mileage. The group then calculates and records the miles per gallon for each purchaser. The data in Case Exhibit 21.1-1 portray the results of the tests.



CASE EXHIBIT 21.1-1

Miles per Gallon Information

Purchaser	Miles per Gallon	Purchaser	Miles per Gallon
1	30.9	13	27.0
2	24.5	14	26.7
3	31.2	15	31.0
4	28.7	16	23.5
5	35.1	17	29.4
6	29.0	18	26.3
7	28.8	19	27.5
8	23.1	20	28.2
9	31.0	21	28.4
10	30.2	22	29.1
11	28.4	23	21.9
12	29.3	24	30.9

Questions

1. Formulate a statistical hypothesis appropriate for the consumer group's purpose.
2. Calculate the mean average miles per gallon. Compute the sample variance and sample standard deviation.
3. Construct the appropriate statistical test for your hypothesis, using a 0.05 significance level.

a sample of employees are as follows (use your computer and statistical software to solve this problem):

Name	Skill Before	Skill After	Name	Skill Before	Skill After
Ed	4.84	5.43	Kathy	4.00	5.00
Mark	5.24	5.51	Susie	4.67	4.50
Jason	5.37	5.42	Ron	4.95	4.40
Raj	3.69	4.50	Jen	4.00	5.95
Heidi	5.95	5.90	Matt	3.75	3.50
Donna	4.75	5.25	Doug	3.85	4.00
Rob	3.90	4.50	Bob	5.00	4.10

9. Conduct a Z-test to determine whether the following two samples indicate that the population proportions are significantly different at the 0.05 level:

	Sample 1	Sample 2
Sample Proportion	0.77	0.68
Sample Size	55	46

10. In an experiment with wholesalers, a researcher manipulated perception of task difficulty and measured level of aspiration for performing the task a second time. Group 1 was told the task was very difficult, group 2 was told the task was somewhat difficult but attainable, and group 3 was told the task was easy. Perform an ANOVA on the resulting data:

Level of Aspiration (10-Point Scale)			
Subjects	Group 1	Group 2	Group 3
1	6	5	5
2	7	4	6
3	5	7	5
4	8	6	4
5	8	7	2
6	6	7	3
Cases	6	6	6

11. Interpret the following output examining group differences for purchase intentions. The three groups refer to consumers from three states: Illinois, Louisiana, and Texas.

Tests of Between-Subjects Effects					
Dependent Variable: int2					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	6681.746 ^a	2	3340.873	3.227	0.043
Intercept	308897.012	1	308897.012	298.323	0.000
State	6681.746	2	3340.873	3.227	0.043
Error	148068.543	143	1035.444		
Total	459697.250	146			
Corrected Total	154750.289	145			

^aR Squared = 0.043 (Adjusted R Squared = 0.030)

Law				
Dependent Variable: int2				
State	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
IL	37.018	4.339	28.441	45.595
LA	50.357	4.965	40.542	60.172
TX	51.459	4.597	42.373	60.546