

Answer Key

BUS 230: Business and Economics Communication and Research
In-class Exercise: Interpreting SPSS output for hypothesis testing
Instructor: Dr. James Murray

Directions: Work in groups of up to four people and answer the questions on the following pages. All papers will be collected, but only one member's paper will be randomly selected and graded and all members of the group will receive the same grade. Failure to complete the worksheet with a group will result in a 20 percentage point penalty.

In the following pages are six research questions and the SPSS output for the appropriate statistical analysis to answer the research question. The purpose of this assignment is to demonstrate you can conduct all the steps these hypotheses tests, and that you can read SPSS output and report summary statistics and hypothesis tests conclusions in plain English.

For all hypothesis tests, you should do the following steps:

1. Report the null and alternative hypothesis.
2. Report the p-value.
3. Report whether you reject the null hypothesis or fail to reject the null hypothesis.
4. Report the conclusion in plain English, in the context of the research problem.

Note, many of the following scenarios also ask you to report additional information. Include this additional information in your answers.

By signing below, you agree that the following work represents the efforts of everyone in the group, and you are willing you accept as your own grade for the group assignment the grade earned from this representation of the group's work.

_____ Signature Group Member 1	_____ Print Name	_____ Date
_____ Signature Group Member 2	_____ Print Name	_____ Date
_____ Signature Group Member 3	_____ Print Name	_____ Date
_____ Signature Group Member 4	_____ Print Name	_____ Date

1. A researcher is interested in determining whether there is a relationship between grade school children's interests and the class standing. The researcher asked 478 grade school children whether being good at sports, being popular, or getting good grades was their most important goal. Class standing is 1st grade, 2nd grade, 3rd grade, 4th grade, and 5th grade. The researcher runs a Chi-squared test of Independence to determine if Goals and Class Standing is related. Use the SPSS output below to answer this research question. *In your answer, also describe what percentage of 3rd grade children care most about getting good grades.*

Note: Goals = 1 if primary goal is Sports, Goals = 2 if primary goal is being Popular, Goals = 3 if primary goal is getting good grades.

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Goals * Grades	478	100.0%	0	.0%	478	100.0%

Goals * Grades Crosstabulation

Count		Grades				Total
		1	2	3	4	
Goals	1	10	24	33	23	90
	2	14	33	45	49	141
	3	70	66	55	56	247
Total		94	123	133	128	478

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	31.090 ^a	6	.000
Likelihood Ratio	31.836	6	.000
N of Valid Cases	478		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 17.70.

H_0 : Class Standing & Goals are independent (i.e. not related)

H_A : Class Standing & Goals are dependent (i.e. they are related)

p-value = 0.000

Reject H_0 .

We found statistical evidence that class standing & goals are related,

133 total third graders, 55 care about grades

$$= \frac{55}{133} = 0.414 = 41.4\%$$

2. A researcher is interested in how important getting good grades are for 4th grade school children versus 5th grade school children. She gives the children a survey and asks them to rate how important getting good grades are to them on the scale, "Very Important," "Somewhat Important," and "Not important." The researcher conducts a Mann-Whitney U-Test to determine whether the median response for 4th grade students is different than the median response for 5th grade students. Use the SPSS output below to answer this research question.

NPar Tests

Mann-Whitney Test

Ranks				
	Grade	N	Mean Rank	Sum of Ranks
Grades	4	119	145.19	17277.50
	5	176	149.90	26382.50
	Total	295		

Test Statistics ^a	
	Grades
Mann-Whitney U	10137.500
Wilcoxon W	17277.500
Z	-.482
Asymp. Sig. (2-tailed)	.630

a. Grouping Variable: Grade

$H_0: \text{median}_1 = \text{median}_2$
 $H_A: \text{median}_1 \neq \text{median}_2$
 p-value = 0.630
 Fail to reject,

Failed to find evidence that median response of 4th graders is different than 5th graders.

3. A researcher is interested in determining whether there is a relationship between income inequality and crime rate. In 1960, researchers collected data on crime rate and income inequality for 47 U.S. states. The researcher estimates a Pearson Correlation Coefficient on crime rate and income inequality. Use the output from SPSS given below to answer this research question. *In your answer, also report what is the sample estimate for the correlation between crime rate and income inequality.*

R: Crime rate (number of offenses reported to authorities per 1,000,000 population)
 X: Income inequality (number of families living below 1/2 median income per 1,000 population)

Correlations

		R	X
R	Pearson Correlation	1	-.179
	Sig. (2-tailed)		.229
	N	47	47
X	Pearson Correlation	-.179	1
	Sig. (2-tailed)	.229	
	N	47	47

H_0 : correlation = 0,

H_A = correlation \neq 0.

p-value = 0.229

Fail to reject H_0 .

Failed to find statistical evidence that income inequality & crime rate are related,

Correlation in sample = -0.179

4. A researcher is interested in whether there is a difference in the unemployment rate of men ages 14-24 (Variable U1 is number of unemployed men aged 14-24 per 1000 men in the age group) and men ages 34-39 (Variable U2 is the number of unemployed men aged 35-39 per 1000 men in the age group). In 1960, the researcher collected data on the unemployment rate of men in these age groups for 47 U.S. states. The researcher estimates a paired-sample t-test to determine if there is a difference in the unemployment rates. Use the SPSS output below to answer this research question. *In your answer, also report the average difference in the unemployment rate. Which age group experiences a higher unemployment rate?*

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	U1	95.47	47	18.029	2.630
	U2	33.98	47	8.445	1.232

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	U1 & U2	47	.748	.000

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	U1 - U2	61.489	13.008	1.897	57.670	65.309	32.407	46	.000

$H_0: \mu_1 = \mu_2$

$H_A: \mu_1 \neq \mu_2$

p-value = 0.000

Reject H_0

Found statistical evidence that there is a difference in average unemployment rate between younger & older age group.

Mean of young age group = 95.47

Mean of old age group = 33.98,

Younger people have higher unemployment-

5. Health officials for the State of New Mexico are interested in estimating the average number of patient-days in the state's 52 licensed nursing homes. In particular, they are interested in whether the average number of patient days is greater than 250. The health officials conduct a one-sample T-test to determine whether the average number of patient days is statistically significantly different from 250. Use the SPSS output below to answer the research question. *In your answer, also report the sample estimate for the average number of patient-days in New Mexico nursing homes.*

T-Test

	N	Mean	Std. Deviation	Std. Error Mean
TDAYS	52	280.17	120.847	16.758

	Test Value = 250				95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
TDAYS	1.800	51	.078	30.173	-3.47	63.82

$H_0: \mu = 250$

$H_A: \mu \neq 250$

p-value = 0.078

Reject H_0 at 10% level.

or
Fail to Reject H_0 at 5% level.

We found statistical evidence (at 10% level) that the average number of patient days is different from 250.

Sample mean = 280.17,

6. Health officials for the State of New Mexico are interested in whether the average number of patient days in rural nursing homes (RURAL=1) is different than in urban nursing homes (RURAL=0). To answer this question, the health officials conduct an independent samples T-test for difference in means. Use the SPSS output to answer the researcher's question. In your answer, also report the average of number of patient days in New Mexico nursing homes in rural areas, urban areas, and the difference between the two.

T-Test

Group Statistics					
	RURAL	N	Mean	Std. Deviation	Std. Error Mean
TDAYS	0	18	322.11	95.035	22.400
	1	34	257.97	128.257	21.996

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval Difference	
									Lower	Upper
TDAYS	Equal variances assumed	.446	.507	1.865	50	.068	64.141	34.401	-4.955	
	Equal variances not assumed			2.043	44.349	.047	64.141	31.394		.884

use top row.

$$H_0: \mu_1 = \mu_2$$

$$H_A: \mu_1 \neq \mu_2$$

$$p\text{-value} = 0.068$$

Reject H_0 at 10% level

or
Fail to reject H_0 at 5% level

We found statistical evidence (at 10% level) that there is a difference in average number of patient days between rural & urban nursing homes.

$$\text{Difference} = 322.11 - 257.97 = 64.14$$

7. Health officials in the State of New Mexico are interested in estimating what percentage of state licensed nursing homes is struggling financially. Historically, when a nursing home's profit margins fell below 1% the nursing home would typically close within 5 years. Researchers collect data on 52 nursing homes and create a dummy variable for low profit margin that is equal to 1 when the profit margin is below 1% and is equal to 0 otherwise. The researchers conduct a one-sample T-test to determine whether the percentage of nursing homes struggling financially is different than 20%. Use the SPSS output below to answer the researcher's question. Also in your answer, report the sample estimate for the percentage of nursing homes struggling financially.

T-Test

	N	Mean	Std. Deviation	Std. Error Mean
lowmargin	52	.3462	.48038	.06662

	Test Value = 0.2					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
lowmargin	2.194	51	.033	.14615	.0124	.2799

H_0 : proportion = 0.2 (i.e. 20%)

H_A : proportion \neq 0.2

p-value = 0.033

Reject H_0 .

We found statistical evidence that the percentage of nursing homes struggling financially is different than 20%.

Sample proportion = 0.3462 = 34.62%