

Test 2 M349R Name: _____

UTeid: _____

Problem 1 (house price data) (38 points)

Construct a 90% confidence interval for adding a 300 sqft bedroom ($t^*=1.66$)

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	-19315	31047	-0.62	0.5355
sqft	sqft	1	128.43621	13.82446	9.29	<.0001
bdrms	bdrms	1	15198	9483.51703	1.60	0.1127

Covariance of Estimates				
Variable	Label	Intercept	sqft	bdrms
Intercept	Intercept	963892569.6	-136222.4363	-180600630.8
sqft	sqft	-136222.4363	191.11564506	-69678.578
bdrms	bdrms	-180600630.8	-69678.578	89937095.296

Write down the point estimator for calculating the interval (6 pts)

Calculate the point estimate (6 pts)

Calculate the variance of the point estimator (10 pts)

Put together the interval and write a conclusion (6 pts)

Do you predict that adding a 300 square feet bedroom will increase the average price of a house in the neighborhood by \$40,000? (8 pts for work and 2 pts for writing the correct hypothesis)

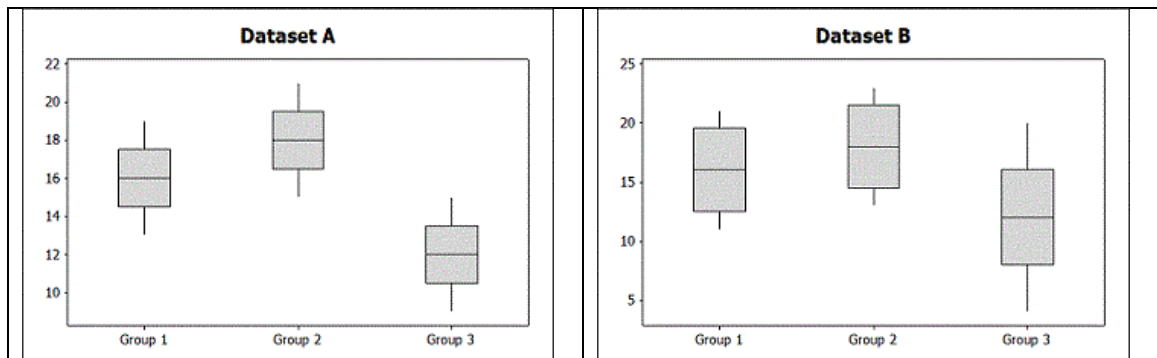
Problem 2

Which of these datasets provides stronger evidence of a difference between the means of the groups?
Circle the dataset with the stronger evidence. (8 pts)

Dataset A			Dataset B		
Group 1	Group 2	Group 3	Group 1	Group 2	Group3
12	25	8	12	19	12
8	15	10	11	18	13
15	12	16	12	18	14
9	9	17	12	18	13
17	28	20	12	17	13
11	19	7	13	18	13
$\bar{x}_1 = 12.0$	$\bar{x}_2 = 18.0$	$\bar{x}_3 = 13.0$	$\bar{x}_1 = 12.0$	$\bar{x}_2 = 18.0$	$\bar{x}_3 = 13.0$

Explain:

[b] Which of these datasets provides stronger evidence of a difference between the means of the groups? Circle the dataset with the stronger evidence. (8 pts)



Explain:

Problem 3 (Overlays)

A study of two surgical methods compare recovery times, in days, for two treatments, the standard and the new method. Three randomly chosen patients got the new treatment; the remaining three patients got the standard. Here are the results:

New procedure 16, 20, 24

Standard 28, 33, 35

[a] Fit (with R) a one-way additive model “days = treatment + error” and write a conclusion (14 points)

[b] For the data above decompose the response value as a sum of grand mean + treatment effects + residuals. (14 points)

16	28
20	33
24	35

=

+

+

Make sure that the sum of square residuals and the sum of square treatment effects is the same as the Anova table from part [a]

Problem 4 (Randomization Matched Pairs Test)

"To exploit the data flood, America will need many more [data analysts]... The story is similar in fields as varied as science and sports, advertising and public health-a drift toward data-driven discovery and decision-making."

Steve Lohr*

The 2008 Olympics were full of controversy about new swimsuits possibly providing unfair advantages to swimmers, leading to new international rules that came into effect January 1, 2010, regarding swimsuit coverage and material. Can a certain swimsuit really make a swimmer faster? A study tested whether wearing wetsuits influences swimming velocity. Twelve competitive swimmers and triathletes swam 1500 m at maximum speed twice each, once wearing a wetsuit and once wearing a regular bathing suit. The order of the trials was randomized. Each time, the maximum velocity in meters/sec of the swimmer was recorded. These data are shown below.



Maximum velocity swimming with and without a wetsuit												
Swimmer	1	2	3	4	5	6	7	8	9	10	11	12
Wetsuit	1.57	1.47	1.42	1.35	1.22	1.75	1.64	1.57	1.56	1.53	1.49	1.51
No Wetsuit	1.49	1.37	1.35	1.27	1.12	1.64	1.59	1.52	1.50	1.45	1.44	1.41

[a] What is the parameter in this problem? (2 points)

[b] What is the best estimator and the best estimate for the parameter above part [a]? (2 points)

[c] What is the null hypothesis and alternative hypothesis for this problem? (2 points)

[d] Explain the algorithm in order to construct a randomization distribution. (6 points)

[e] Write R code with a replicate function or a loop in order to complete the problem and report a randomization p-value and conclusion. (6 points)

