You may answer the problems and their subparts in any order. Just give us a clue which answer is which. Write your answers on separate paper. Each part of each problem is worth 3 points.

PROBLEM 1

This question is based on

Lauenroth, W.K., Burke, I.C., Paruelo, J.M. (2000). Patterns of production and precepitation-use efficiency of winter wheat and native grasslands in the central Great Plains of the United States. *Ecosystems, 3*, 344-351.

The authors drew a transect line across the grass and wheat growing areas of Colorado and Kansas. They then sampled tracts along that transect. The following variables are available for each tract.

ANPP (g/m^2) Aboveground Net Primary Production (the higher the number, the greater the biological "output" from the ecosystem, ranging from about 50 to 500)

MAP Mean annual precipitation (mm, ranging from about 300 to 800)

CROP grassland or wheat

Use the above variables to specify MODELs C and A that would answer the following questions.

A. On average, is there a difference in Aboveground Net Primary Production (ANPP) between grassland and wheat tracts?

Later (in Chapter 10) we will learn more powerful and sophisticated methods for asking the following questions, but for now we will just be using the wheat tracts. Just assume that we no longer have the data for the grassland tracts so they will not be included in any of the model questions below. This isn't supposed to be a tricky part of this question. If this doesn't make sense, ask!

B. To be making efficient use of this type of ecosystem, the ANPP should be at least 325. Is average ANPP different from 325 for the wheat tracts?

C. Does Aboveground Net Primary Production increase with precipitation? [Note: the slope of this relationship is referred to as "precipitation-use efficiency." The higher the slope the better the ecosystem is at converting rain to biological output.]

D. For reasons I won't explain here, there was particular interest in ANPP when there is 575 mm of annual precipitation. Assuming that ANPP does increase with precipitation for the wheat tracts, does ANPP differ from 325 for the wheat tracts when mean annual precipitation is 575 mm? [Use the most powerful test.]

E. The precipitation-use efficiency (PUE, the rate at which rain water is converted into ANPP) for grasslands is known to be .82. Does the PUE for the wheat tracts differ from the known PUE for the grasslands?

F. Just using the grassland tracts, the authors regressed ANPP on mean annual precipitation (MAP) and obtained the following prediction equation:

$$\widehat{ANPP} = -156 + 0.82\ MAP$$

Does the equation for grasslands fir the data from wheat tracts (include only the wheat tracts in the models to answer this question)?

<u>The following are NOT Model A/C questions.</u>

G. After presenting this wheat model

$$\widehat{ANPP}\ 201 + 0.2\ MAP$$

the authors gave this explanation: "The wheat model has a positive *y*-intercept, implying a positive ANPP as a result of water stored during the fallow year even in the absence of annual precipitation [p. 347]." What would make you very cautious about such a statement?

H. After presenting the wheat model

$$\widehat{ANPP} = 201 + 0.2\ MAP$$

the authors write: "…the slope of the wheat model suggests that on average just 0.2 g/m^2 of ANPP is added for every millimeter of mean annual precipitation above 201 mm [p. 347.]" Briefly explain the very serious error the authors made in interpreting the parameters of their model. [Fortunately, the authors are at that other state university and not this one.]

I. For the wheat model they report $r^2 = .67$, $n = 19$, $p < .01$. Suppose they were planning to conduct a study along a transect further north for which there would only be 10 wheat tracts available. What would be their *approximate* chances of finding a significant relationship between ANPP and MAP for those 10 tracts if the relationship were the same?

J. Given that the authors observed consecutive tracts along a transect (a straight line drawn on the map), what assumption about error was likely violated?

PROBLEM 2

The attached output comes from data collected by U.S. News and World Report of all colleges and universities in the United States. In this output, data are used from four different variables:

GRADRATE	The percentage of enrolled students who graduate.(0 to 100)
SFRATIO	Student to faculty ratio
NFULL	Number of full time students
PUBPRI	Whether the school is public (+1) or private (-1)
SFRATIOMD	SFRATIO minus its mean (14.7567).

The output includes descriptive statistics on the four variables and five **proc reg** outputs from the following models:

1: model gradrate=/clb;
2: model gradrate=sfratio/clb;
3: model gradrate=sfratiomd/clb;
4: model gradrate=nfull/clb;
5: model gradrate=pubpri/clb;

Using the results from the SAS output, answer the questions below. In many instances the numbers you need are directly available on the outputs. In other instances you will need to calculate additional statistics based on the numbers available in the outputs.

A. Is student to faculty ratio a useful predictor of the graduation rate at these colleges? (Provide Pre, F*, pa-pc, N-pa, and a one sentence substantive description of the relationship.)

B. In the model where SFRATIO is used to predict GRADRATE, provide short interpretations of the slope and intercept.

C. At what level of SFRATIO do we expect the graduation rate to be only about 50%?

D. Using the simplest test, test whether the mean graduation rate in these colleges is different from .60. (Provide Pre, F*, pa-pc, N-pa, and a one sentence substantive conclusion.)

E. Given that SFRATIO is a useful predictor of GRADRATE, test once again whether the mean graduation rate in these colleges is different from .60. (Provide Pre, F*, pa-pc, N-pa, and a one sentence substantive conclusion.)

F. Is there a significant difference between public and private universities in their graduation rates? (Provide Pre, F*, pa-pc, N-pa, and a one sentence substantive conclusion.)

G. What is the mean graduation rate at public universities? What is the mean graduation rate at private universities?

H. In Model 5, where GRADRATE is regressed on PUBPRI, provide a one sentence interpretation of the intercept.

J. If we tested the null hypothesis that the mean graduation rate at private universities was 20 percentage points higher than the mean graduation at public universities, would we reject that null hypothesis? (Don't try to calculate PRE or F* but do provide a justification for your answer.)

The MEANS Procedure

Variable	Label	N	Mean	Std Dev
GRADRATE	Graduation Rate	1201	60.4704413	18.8612743
SFRATIO	Student/Faculty Ratio	1201	14.7567027	4.7723691
NFULL	Number of Fulltime Students	1201	3819.85	4666.06
PUBPRI		1201	-0.2822648	0.9597362

Variable	Label	Minimum	Maximum
GRADRATE	Graduation Rate	8.0000000	118.0000000
SFRATIO	Student/Faculty Ratio	2.5000000	91.8000000
NFULL	Number of Fulltime Students	59.0000000	31643.00
PUBPRI		-1.0000000	1.0000000

The REG Procedure
Model: MODEL1
Dependent Variable: GRADRATE Graduation Rate

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	0	0	.	.	.
Error	1200	426897	355.74767		
Corrected Total	1200	426897			

Root MSE	18.86127	R-Square	0.0000
Dependent Mean	60.47044	Adj R-Sq	0.0000
Coeff Var	31.19090		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value
Intercept	Intercept	1	60.47044	0.54425	111.11

Parameter Estimates

Variable	Label	DF	Pr > |t|	95% Confidence Limits	
Intercept	Intercept	1	<.0001	59.40265	61.53823

The REG Procedure
Model: MODEL2
Dependent Variable: GRADRATE Graduation Rate

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	49718	49718	158.05	<.0001
Error	1199	377179	314.57833		
Corrected Total	1200	426897			

Root MSE	17.73636	R-Square	0.1165
Dependent Mean	60.47044	Adj R-Sq	0.1157
Coeff Var	29.33062		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value
Intercept	Intercept	1	80.37353	1.66384	48.31
SFRATIO	Student/Faculty Ratio	1	-1.34875	0.10729	-12.57

Parameter Estimates

Variable	Label	DF	Pr > |t|	95% Confidence Limits	
Intercept	Intercept	1	<.0001	77.10916	83.63790
SFRATIO	Student/Faculty Ratio	1	<.0001	-1.55924	-1.13826

The REG Procedure
Model: MODEL3
Dependent Variable: GRADRATE Graduation Rate

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	49718	49718	158.05	<.0001
Error	1199	377179	314.57833		
Corrected Total	1200	426897			

Root MSE	17.73636	R-Square	0.1165
Dependent Mean	60.47044	Adj R-Sq	0.1157
Coeff Var	29.33062		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value
Intercept	Intercept	1	60.47045	0.51179	118.15
sfratiomd		1	-1.34875	0.10729	-12.57

Parameter Estimates

Variable	Label	DF	Pr > |t|	95% Confidence Limits	
Intercept	Intercept	1	<.0001	59.46634	61.47455
sfratiomd		1	<.0001	-1.55924	-1.13826

The REG Procedure
Model: MODEL4
Dependent Variable: GRADRATE Graduation Rate

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	3671.86635	3671.86635	10.40	0.0013
Error	1199	423225	352.98193		
Corrected Total	1200	426897			

Root MSE	18.78781	R-Square	0.0086
Dependent Mean	60.47044	Adj R-Sq	0.0078
Coeff Var	31.06942		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value
Intercept	Intercept	1	61.90246	0.70074	88.34
NFULL	Num of Fulltime Students	1	-0.00037489	0.00011623	-3.23

Parameter Estimates

Variable	Label	DF	Pr > |t|	95% Confidence Limits	
Intercept	Intercept	1	<.0001	60.52764	63.27728
NFULL	Num of Fulltime Students	1	0.0013	-0.00060293	-0.00014684

The REG Procedure
Model: MODEL5
Dependent Variable: GRADRATE Graduation Rate

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	69641	69641	233.72	<.0001
Error	1199	357257	297.96210		
Corrected Total	1200	426897			

Root MSE	17.26158	R-Square	0.1631
Dependent Mean	60.47044	Adj R-Sq	0.1624
Coeff Var	28.54548		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value
Intercept	Intercept	1	58.22994	0.51920	112.15
PUBPRI		1	-7.93759	0.51920	-15.29

Parameter Estimates

Variable	Label	DF	Pr > |t|	95% Confidence Limits	
Intercept	Intercept	1	<.0001	57.21129	59.24859
PUBPRI		1	<.0001	-8.95624	-6.91895