

**FINAL EXAM
STAT 302**

Fall 1996

Instructions: Show work for partial credit. If you run out of time (I hope not), delineate the procedure you would use to solve the problem.

1. Data were recorded to study a new method of measuring body composition. In this study, the focus was on % body fat. The main objective was to investigate the relationship between % fat and age, and whether this relationship was different for males and females. The data appear below (source: SDS, p. 13).

4 Cols	<input checked="" type="checkbox"/> C <input type="checkbox"/>	<input checked="" type="checkbox"/> C <input type="checkbox"/>	<input checked="" type="checkbox"/> N <input type="checkbox"/>
	Age	% Fat	Gender
1	23	9.5	M
2	23	27.9	F
3	27	7.8	M
4	27	17.8	M
5	39	31.4	F
6	41	25.9	F
7	45	27.4	M
8	49	25.2	F
9	50	31.1	F
10	53	34.7	F
11	53	42	F
12	54	29.1	M
13	56	32.5	F
14	57	30.3	M
15	58	33	F
16	58	33.8	F
17	60	41.1	F
18	61	34.5	F

A model containing the interaction was fitted to the data. Recall that JMP creates the dummy variable for gender as: 1 if F and -1 if M.

- a) Is this model useful? Interpret the value of R^2 . Are all the variables contributing to the model?

Response: % Fat

Summary of Fit

RSquare	0.81582
RSquare Adj	0.776353
Root Mean Square Error	4.324506
Mean of Response	28.61111
Observations (or Sum Wgts)	18

Lack of Fit

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	7.6117098	4.024894	1.89	0.0795
Age	0.4482701	0.08743	5.13	0.0002
Gender[F-M]	11.631096	4.024894	2.89	0.0119
Age*Gender[F-M]	-0.178409	0.08743	-2.04	0.0606

Effect Test

Source	Nparm	DF	Sum of Squares	F Ratio	Prob>F
Age	1	1	491.61704	26.2878	0.0002
Gender	1	1	156.17280	8.3509	0.0119
Age*Gender	1	1	77.87221	4.1640	0.0606

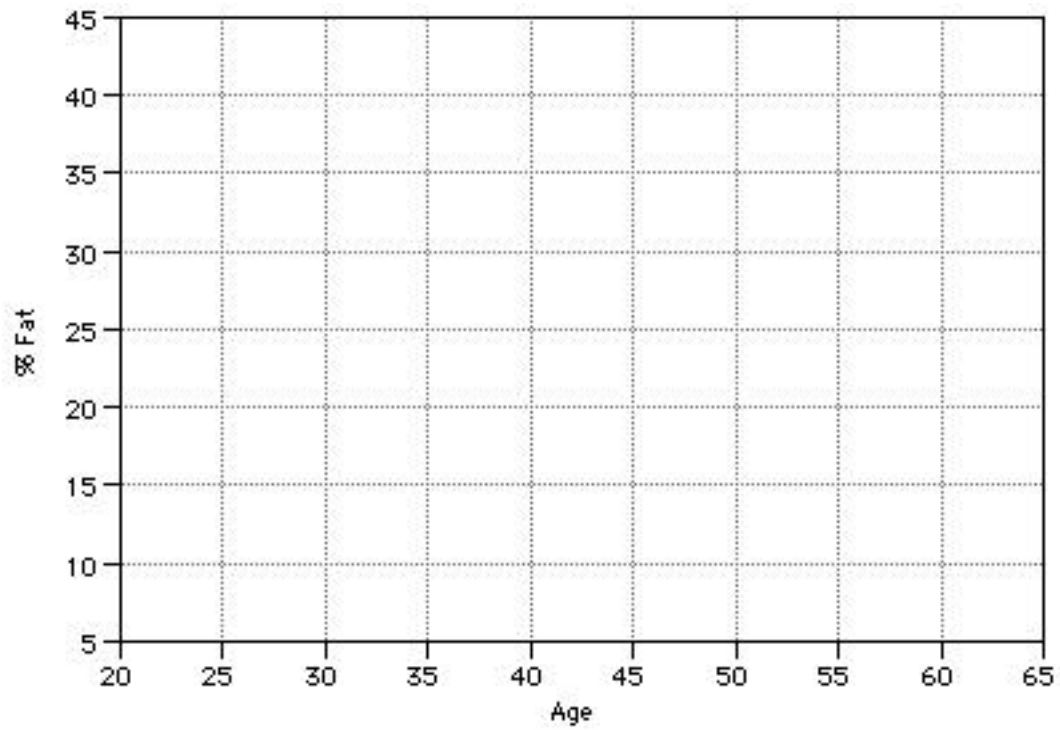
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	1159.7189	386.573	20.6709
Error	14	261.8189	18.701	Prob>F
C Total	17	1421.5378		<.0001

b) Quantify the effect (slope) of age for females. [Hint: Write the model]

c) Quantify the effect (slope) of age for males. [Hint: Write the model]

d) Draw a picture of the model and explain its meaning.



e) Obtain the residual for the subject number 10.

2. A survey was conducted to investigate whether snoring was related to, among other things, heart disease. Those surveyed were classified according to the amount they snored, on the basis of report from their spouses. The data appear in the following table (source: SDS, p. 19).

Heart disease	Non-snorers	Occasional snorers	Snore nearly every night	Snore every night	Total
Yes	24	35	21	30	110
No	1355	603	192	224	2374
Total	1379	638	213	254	2484

a) Explain in plain words what the hypothesis of independence means in this case.

b) Obtain the number of occasional snorers with heart disease that are expected under the assumption of independence.

The results of the analysis appear below.

		Snore				
		Every night	Nearly everu night	Occasionally	No	
Heart Disease	Count	224	192	603	1355	2374
	Col %	88.19	90.14	94.51	98.26	
	N					
	Y	30	21	35	24	110
		11.81	9.86	5.49	1.74	
		254	213	638	1379	2484

Tests

Source	DF	-LogLikelihood	RSquare (U)
Model	3	32.95224	0.0732
Error	2480	417.46136	
C Total	2483	450.41360	
Total Count	2484		

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	65.904	<.0001
Pearson	72.782	<.0001

c) According to these results, what do you conclude?

d) Is there any trend in the data is worth mentioning?

3. Suppose that you are told that the p-value in an ANOVA F-test to compare the average of each of four populations is $p = 0.002$. What would you conclude?