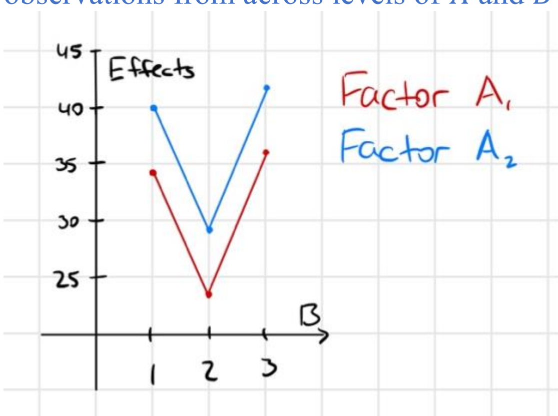
*19.4. In a two-factor study, the treatment means μ_{ij} are as follows:

	Factor B		
Factor A	B_1	B_2	B_3
A_1	34	23	36
A_2	40	29	42

a. Obtain the factor A level means.
b. Obtain the main effects of factor A.
c. Does the fact that $\mu_{12} - \mu_{11} = -11$ while $\mu_{13} - \mu_{12} = 13$ imply that factors A and B interact? Explain.
d. Prepare a treatment means plot and determine whether the two factors interact. What do you find?

a. $\mu_1 = \frac{34+23+36}{3} = 31$; $\mu_2 = \frac{40+29+42}{3} = 37$.
b. $\mu_{12} = \frac{31+37}{2} = 34$. $\alpha_1 = 31 - 37 = -6$. $\alpha_2 = 34 - 37 = -3$.
c. No because both differences come from within the level A_1. You would need to compare observations from across levels of A and B to determine an interaction effect.
d.



As evidenced by the parallel treatment effects, the factors do not interact.

19.5. In a two-factor study, the treatment means μ_{ij} are as follows:

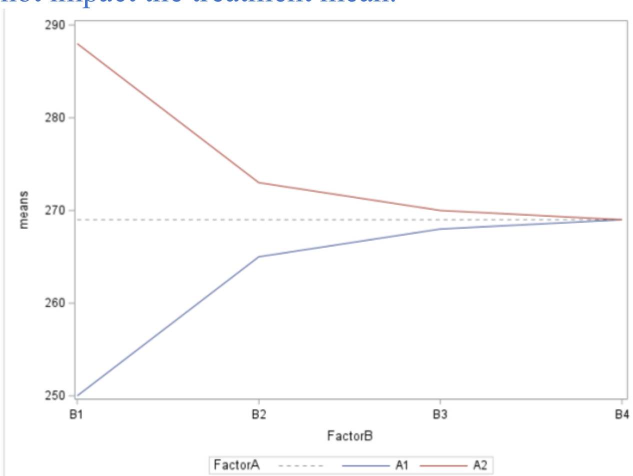
	Factor B			
Factor A	B_1	B_2	B_3	B_4
A_1	250	265	268	269
A_2	288	273	270	269

a. Obtain the factor B main effects. What do your results imply about factor B?

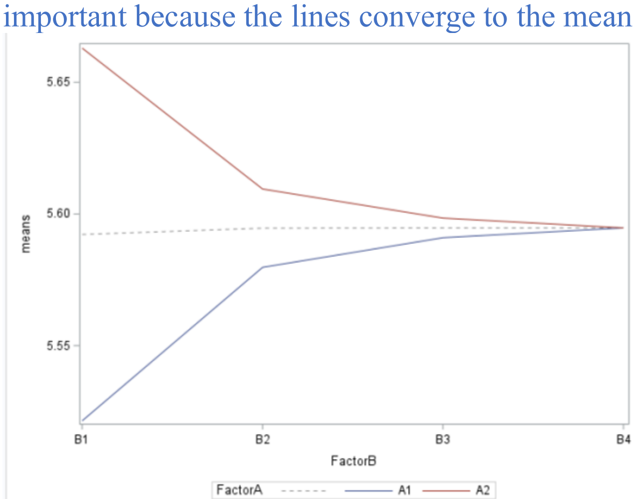
b. Prepare a treatment means plot and determine whether the two factors interact. How can you tell that interactions are present? Are the interactions important or unimportant?

c. Make a logarithmic transformation of the μ_{ij} and plot the transformed values to explore whether this transformation is helpful in reducing the interactions. What are your findings?

a. B_1 through B_4 have the same mean, 269, so the implication is that Factor B *by itself* does not impact the treatment mean.

b. 

Level A_1 is represented by the blue line and A_2 is represented by the red line. There does appear to be interactions between the two factors, as evidenced by the non-parallel nature of the lines. These interactions are important because the lines converge to the mean 269.

c. 

Applying the natural log transformation to the response variable does not appear to reduce the interaction effect.

*19.14. **Hay fever relief.** A research laboratory was developing a new compound for the relief of severe cases of hay fever. In an experiment with 36 volunteers, the amounts of the two active ingredients (factors A and B) in the compound were varied at three levels each. Randomization was used in assigning four volunteers to each of the nine treatments. The data on hours of relief follow.

Factor A (ingredient 1)		Factor B (ingredient 2) $j=1$ Low	$j=2$ Medium	$j=3$ High
$i=1$	Low	2.4	4.6	4.8
		...	...	...
		2.5	4.7	4.6
$i=2$	Medium	5.8	8.9	9.1
		...	...	.
		5.3	9.0	9.4
$i=3$	High	6.1	9.9	13.5
		...	...	...
		6.2	10.1	13.2

a. Obtain the fitted values for ANOVA model (19.23).

b. Obtain the residuals.

c. Plot the residuals against the fitted values. What departures from ANOVA model (19.23) can be studied from this plot? What are your findings?

d. Prepare a normal probability plot of the residuals. Also obtain the coefficient of correlation between the ordered residuals and their expected values under normality. Does the normality assumption appear to be reasonable here?

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	373.1050000	46.6381250	774.91	<.0001
Error	27	1.6250000	0.0601852		
Corrected Total	35	374.7300000			

R-Square	Coeff Var	Root MSE	TreatmentMean Mean
0.995664	3.415221	0.245327	7.183333

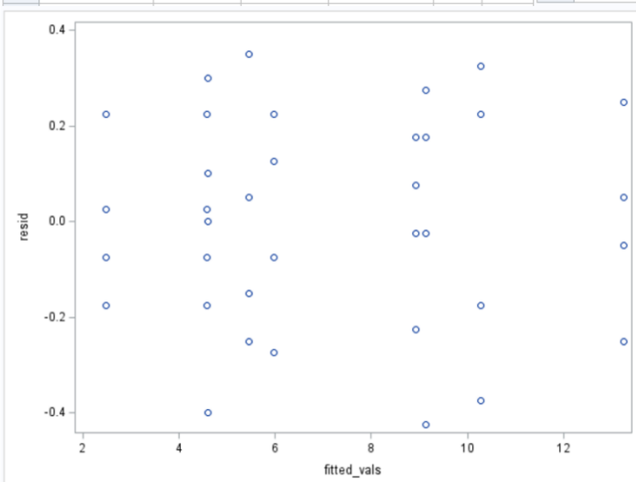
Source	DF	Type I SS	Mean Square	F Value	Pr > F
IngredientA	2	220.0200000	110.0100000	1827.86	<.0001
IngredientB	2	123.6600000	61.8300000	1027.33	<.0001
Ingredien*Ingredient	4	29.4250000	7.3562500	122.23	<.0001

a.

Source	DF	Type III SS	Mean Square	F Value	Pr > F
IngredientA	2	220.0200000	110.0100000	1827.86	<.0001
IngredientB	2	123.6600000	61.8300000	1027.33	<.0001
Ingredien*Ingredient	4	29.4250000	7.3562500	122.23	<.0001

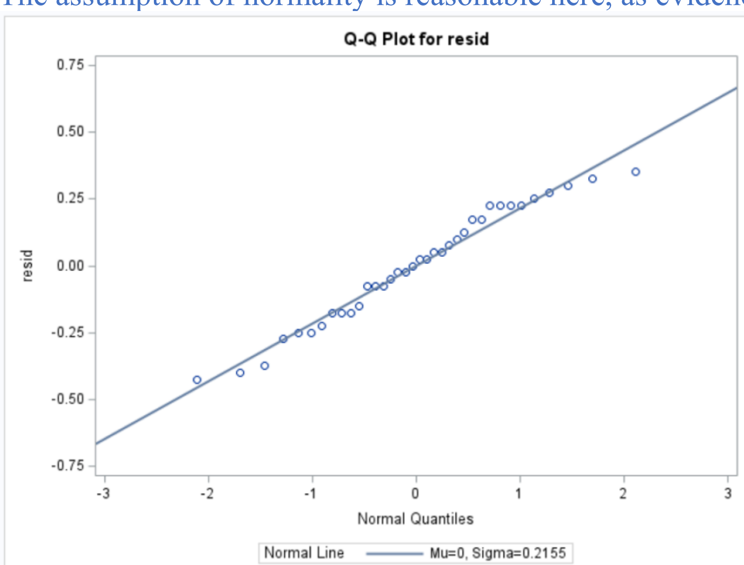
b.

Obs	TreatmentMean	IngredientA	IngredientB	VolunteerNum	resid	pred
1	2.4	1	1	1	-0.075	2.475
2	2.7	1	1	2	0.225	2.475
3	2.3	1	1	3	-0.175	2.475
4	2.5	1	1	4	0.025	2.475
5	4.6	1	2	1	-0.000	4.600
6	4.2	1	2	2	-0.400	4.600
7	4.9	1	2	3	0.300	4.600
8	4.7	1	2	4	0.100	4.600
9	4.8	1	3	1	0.225	4.575
10	4.5	1	3	2	-0.075	4.575
11	4.4	1	3	3	-0.175	4.575
12	4.6	1	3	4	0.025	4.575
13	5.8	2	1	1	0.350	5.450
14	5.2	2	1	2	-0.250	5.450
15	5.5	2	1	3	0.050	5.450
16	5.3	2	1	4	-0.150	5.450
17	8.9	2	2	1	-0.025	8.925
18	9.1	2	2	2	0.175	8.925
19	8.7	2	2	3	-0.225	8.925
20	9.0	2	2	4	0.075	8.925
21	9.1	2	3	1	-0.025	9.125
22	9.3	2	3	2	0.175	9.125
23	8.7	2	3	3	-0.425	9.125
24	9.4	2	3	4	0.275	9.125
25	6.1	3	1	1	0.125	5.975
26	5.7	3	1	2	-0.275	5.975
27	5.9	3	1	3	-0.075	5.975
28	6.2	3	1	4	0.225	5.975
29	9.9	3	2	1	-0.375	10.275
30	10.5	3	2	2	0.225	10.275
31	10.6	3	2	3	0.325	10.275
32	10.1	3	2	4	-0.175	10.275
33	13.5	3	3	1	0.250	13.250
34	13.0	3	3	2	-0.250	13.250
35	13.3	3	3	3	0.050	13.250
36	13.2	3	3	4	-0.050	13.250

c. 

As evidenced by the residual plot, there doesn't seem to be any outliers or nonconstant variance.

d. The assumption of normality is reasonable here, as evidenced by the qqplot.



*19.15. Refer to **Hay fever relief** Problem 19.14. Assume that ANOVA model (19.23) is applicable.

a. Prepare an estimated treatment means plot. Does your graph suggest that any factor effects are present? Explain.

b. Obtain the analysis of variance table. Does any one source account for most of the total variability in hours of relief in the study? Explain.

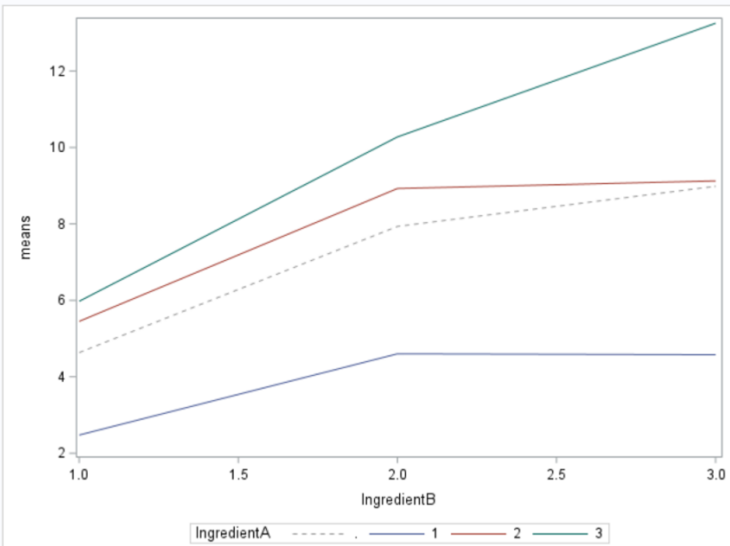
c. Test whether or not the two factors interact; use $\alpha = .05$. State the alternatives, decision rule, and conclusion. What is the P-value of the test?

d. Test whether or not main effects for the two ingredients are present. Use $\alpha = .05$ in each case and state the alternatives, decision rule, and conclusion. What is the P-value of each test? Is it meaningful here to test for main factor effects? Explain.

e. Obtain an upper bound on the family level of significance for the tests in parts (c) and (d); use the Kimball inequality (19.53).

f. Do the results in parts (c) and (d) confirm your graphic analysis in part (a)?

a. The different color lines denote the different levels of Ingredient A. Because they are non-parallel there seems to an interaction effect seems to exist between A and B.



b. Ingredient A seems to account for the most variability, as evidenced by its greatest Type III SS value:

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	373.1050000	46.6381250	774.91	<.0001
Error	27	1.6250000	0.0601852		
Corrected Total	35	374.7300000			

R-Square	Coeff Var	Root MSE	TreatmentMean Mean
0.995664	3.415221	0.245327	7.183333

Source	DF	Type I SS	Mean Square	F Value	Pr > F
IngredientA	2	220.0200000	110.0100000	1827.86	<.0001
IngredientB	2	123.6600000	61.8300000	1027.33	<.0001
Ingredien*Ingredient	4	29.4250000	7.3562500	122.23	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
IngredientA	2	220.0200000	110.0100000	1827.86	<.0001
IngredientB	2	123.6600000	61.8300000	1027.33	<.0001
Ingredien*Ingredient	4	29.4250000	7.3562500	122.23	<.0001

c. The F^* of testing whether interaction occurs is 122.23, which is greater than $F(0.95, 4, 27) = 0.8606$. The corresponding p-value is < 0.0001, which is much less than 0.05. Hence there exists significant statistical evidence to conclude that Factors A and B interact to affect the response.
d. For Factor A, $F^* = 1827.86$ and for Factor B, $F^* = 1027.33$. These values are both greater than $F(0.95, 2, 27) = 0.7113$. Both p-values are less than 0.0001, significantly less than 0.05. Thus there exists significant statistical evidence to conclude that Factors A and B both independently exert a mean effect on the response.
e. The upper bound is given by $\alpha \leq 1 - (1 - 0.5)(1 - 0.5)(1 - 0.5) \leq 0.143$.
f. Yes because they both confirm the existence and significance of an interaction effect.

19.28. A two-factor study was conducted with $a = 6$, $b = 6$, and $n = 10$. No interactions between factors A and B were found, and it is now desired to estimate five contrasts of factor A level means and four contrasts of factor B level means. The family confidence coefficient for the joint set of estimates is to be 95 percent. Which of the three procedures at the bottom of page 852 and the top of page 853 will be most efficient here?

The most efficient procedure will have the smallest coefficient. To that end, let's examine:; $S = \sqrt{(6-1)F(1-0.5, 6-1, (10-1)(6)(6))} = 2.088$; $B = t\left(1 - \frac{\alpha}{2g}, (n-1)ab\right) = t\left(1 - \frac{0.05}{2(9)}, (10-1)(6)(6)\right) = t(0.9722, 324) = 1.921$; $S_{multiple} = \sqrt{(6+6-2)F(1-0.5, 6+6-2, (10-1)(6)(6))} = 3.060$. We conclude that the Bonferroni procedure is most efficient.

6. A clay tile company was interested in studying the effects of oven and cooling temperature on the strength of their tiles. The company's five ovens, used to bake the tiles, and four cooling temperatures (°C) were considered. The data are shown below.

Cooling Temp	Oven 1	2	3	4	5	Mean
5	5	10	7	4	3	5.80
10	3	8	12	2	8	6.6
15	9	13	15	4	10	10.20
20	7	12	9	6	13	9.40
Mean	6.00	10.75	10.75	4.00	8.50	8.00

a. Here MSE=5.375, compute the F-statistic to determine if there is a difference among the four cooling temperatures and the five ovens (use $\alpha = .05$). If significant, perform pairwise comparisons using Tukey's procedure (HINT: Use Table B.9)

b. Suppose the company believes there is a jump in the tile strength at 12.5 °C but otherwise cooling temperature has no effect (i.e., step function ___‾‾‾). Find a set of three contrasts that would allow you to test this (HINT: Contrasts in this set need to test the jump but also the relationship among the means before and after 12.5 °).

c. Test these contrasts using SAS (or by hand). State your conclusions.

a. The F statistic is 5.57 with a p value of 0.0048, so we proceed with the Tukey tests:

The SAS System

The GLM Procedure

Dependent Variable: Strength

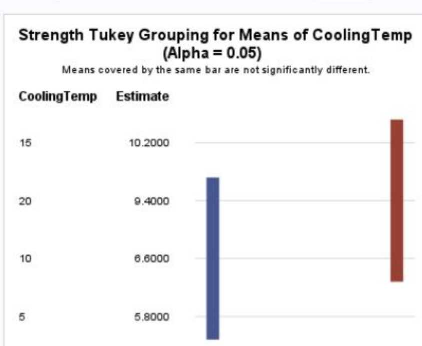
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	209.5000000	29.9285714	5.57	0.0048
Error	12	64.5000000	5.3750000		
Corrected Total	19	274.0000000			

R-Square	Coeff Var	Root MSE	Strength Mean
0.764599	28.98006	2.318405	8.000000

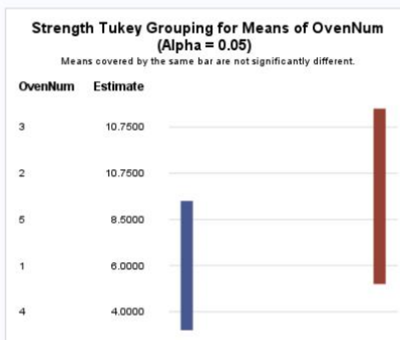
Source	DF	Type I SS	Mean Square	F Value	Pr > F
CoolingTemp	3	68.0000000	22.6666667	4.22	0.0298
OvenNum	4	141.5000000	35.3750000	6.58	0.0048

Source	DF	Type III SS	Mean Square	F Value	Pr > F
CoolingTemp	3	68.0000000	22.6666667	4.22	0.0298
OvenNum	4	141.5000000	35.3750000	6.58	0.0048

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	5.375
Critical Value of Studentized Range	4.19851
Minimum Significant Difference	4.3531



Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	5.375
Critical Value of Studentized Range	4.50766
Minimum Significant Difference	5.2253



b. Let $\mu_1 = 5°C$, $\mu_2 = 10°C$, $\mu_3 = 15°C$, and $\mu_4 = 20°C$. Three possible contrasts are,
   1. $\frac{\mu_1+\mu_2}{2} - \frac{\mu_3+\mu_4}{2}$
   2. $\mu_3 - \mu_2$
   3. $\mu_4 - \mu_1$

c. As observed by the contrasts, at a $\alpha = 0.05$ significance level there is no difference strength between mean temperatures 15 and 10, and 20 and 5. However, there is a significant difference between the jump from 5-10 to 15-20. Hence we can conclude there is a significant jump in tile strength at 12.5 degrees Celsius.

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
5&10 v 15&20	1	64.80000000	64.80000000	5.03	0.0394
15 v 10	1	32.40000000	32.40000000	2.52	0.1322
20 v 5	1	32.40000000	32.40000000	2.52	0.1322

*20.4. Refer to **Coin-operated terminals** Problem 20.2. Conduct the Tukey test for additivity; use $\alpha = .025$. State the alternatives, decision rule, and conclusion. If the additive model is not appropriate, what might you do?

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	84.12855763	16.82571153	126.30	0.0079
Error	2	0.26644237	0.13322119		
Corrected Total	7	84.39500000			

R-Square	Coeff Var	Root MSE	hours_not_used Mean
0.996843	2.182330	0.364995	16.72500

Source	DF	Type I SS	Mean Square	F Value	Pr > F
location	3	37.00500000	12.33500000	92.59	0.0107
week	1	47.04500000	47.04500000	353.13	0.0028
yhat*yhat	1	0.07855763	0.07855763	0.59	0.5228

Source	DF	Type III SS	Mean Square	F Value	Pr > F
location	3	1.32330343	0.44110114	3.31	0.2406
week	1	1.32895606	1.32895606	9.98	0.0873
yhat*yhat	1	0.07855763	0.07855763	0.59	0.5228

We have $H_0: D = 0$ and $H_A: D \neq 0$. We have $F^* = 126.30$ for the additive model. The corresponding p value is 0.00079, which is less than 0.025. We then reject H_0 and conclude there exists significant statistical evidence to suggest that interaction effects are present between terminal location and week number.