Determining Statistically Significant Effects

CASE 2: There are no independent replications of any of the treatment combinations.

CASE 2 (A): n is large

CASE 2 (B): n is small





n is large

Many researchers have observed that four-way, five-way, and higher order interaction are rarely significant, and hence, they most likely correspond to contrasts that are measuring experimental error rather than effects of the sampled treatment combinations.





n is large

When n is large, say n > 5, the sums of squares for 4-way and higher order interactions are often pooled together and used to estimate the experimental error variance.





Example: Consider a 2⁶ experiment.

Effect Type	Number
Main	6
2-Way Interactions	15
3-Way Interactions	20
4-Way Interaction	15
5-Way Interactions	6
6-Way Interaction	1
Total	63





By pooling the four-way and higher order interaction sums of squares into error in a 2^6 experiment, we can get an approximate estimate of the experimental error variance that has

$$15 + 6 + 1 = 22$$
 degrees of freedom.





n is small

When n is small, say n < 6, then half-normal plots can be used to identify the statistically significant effects. Then after the statistically significant effects are identified, another half-normal plot can be created to help get an estimate of the experimental error variance.





Remark: Many authors suggest pooling 3-way and higher order interactions into error. While I have only occasionally observed statistically significant 4-way and higher order interactions during my career as a statistician, I have seen lots of statistically significant 3-way interactions.

Consequently, I would rather use half-normal plots to help estimate the experimental error than depend on being able to use 3-way interaction effects to estimate experimental error.





Table 7.4 Definition of the Treatment Structure in a 2⁵ Factorial Experiment to Study Cake Quality

Level

Variable	Low	High _
W: amount of water	40	120
M: mixing time	3	6
T: temperature	300	500
C: cooking oil	6	12
P: mixer type	\mathbf{P}_1	P_2
_		

Obs	W	M	T	C	P	QUALITY
1	0	0	0	0	0	4.8
2	1	0	0	0	0	3.9
3	0	1	0	0	0	5.0
4	1	1	0	0	0	2.2
5	0	0	1	0	0	3.9
6	1	0	1	0	0	4.2
7	0	1	1	0	0	3.0
8	1	1	1	0	0	2.2
9	0	0	0	1	0	5.7
10	1	0	0	1	0	2.2
11	0	1	0	1	0	8.4
12	1	1	0	1	0	8.3
13	0	0	1	1	0	5.3
14	1	0	1	1	0	2.3
15	0	1	1	1	0	8.6
16	1	1	1	1	0	8.9
17	0	0	0	0	1	4.2
18	1	0	0	0	1	5.0
19	0	1	0	0	1	5.8
20	1	1	0	0	1	5.2
21	0	0	1	0	1	4.6
22	1	0	1	0	1	4.1
23	0	1	1	0	1	5.4
24	1	1	1	0	1	5.2
25	0	0	0	1	1	2.9
26	1	0	0	1	1	3.0
27	0	1	0	1	1	6.7
28	1	1	0	1	1	6.6
29	0	0	1	1	1	5.0
30	1	0	1	1	1	2.7
31	0	1	1	1	1	7.0
32	1	1	1	1	1	7.1





Obs	W	M	T	C	P	QUALITY
1	0	0	0	0	0	4.8
2	1	0	0	0	0	3.9
3	0	1	0	0	0	5.0
4	1	1	0	0	0	2.2
5	0	0	1	0	0	3.9
6	1	0	1	0	0	4.2
7	0	1	1	0	0	3.0
8	1	1	1	0	0	2.2
9	0	0	0	1	0	5.7
10	1	0	0	1	0	2.2
11	0	1	0	1	0	8.4
12	1	1	0	1	0	8.3
13	0	0	1	1	0	5.3
14	1	0	1	1	0	2.3
15	0	1	1	1	0	8.6
16	1	1	1	1	0	8.9





	ı	i i	ì	ì		i .
17	0	0	0	0	1	4.2
18	1	0	0	0	1	5.0
19	0	1	0	0	1	5.8
20	1	1	0	0	1	5.2
21	0	0	1	0	1	4.6
22	1	0	1	0	1	4.1
23	0	1	1	0	1	5.4
24	1	1	1	0	1	5.2
25	0	0	0	1	1	2.9
26	1	0	0	1	1	3.0
27	0	1	0	1	1	6.7
28	1	1	0	1	1	6.6
29	0	0	1	1	1	5.0
30	1	0	1	1	1	2.7
31	0	1	1	1	1	7.0
32	1	1	1	1	1	7.1





PROC ANOVA;

CLASSES W M T C P;

MODEL QUALITY = W M W*M T W*T
M*T W*M*T C W*C M*C W*M*C T*C
W*T*C M*T*C P W*P M*P W*M*P T*P
W*T*P M*T*P C*P W*C*P M*C*P
T*C*P;

RUN;





Class Level Information							
Cl Lev Val							
ass	els	ues					
W	2	0 1					
M	2	0 1					
T	2	0 1					
C	2	0 1					
P	2	0 1					

Number of observations





Source	D F	~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~			Pr > F
Model	25	116.9262	4.6770500	12.63	0.0023
Error	6	2.22250	0.3704167		
Corrected Total	31	119.1487			

R-Square	Coeff Var	Root MSE	QUALITY Mean
0.981347	12.21819	0.608619	4.981250





Source	D F	Sum of Squares		F Value	Pr > F
Model	25	116.9262	4.6770500		
Error	6	2.22250	0.3704167		
Corrected Total	31	119.1487			

R-Square	Coeff Var	Root MSE	QUALITY Mean
0.981347	12.21819	0.608619	4.981250







				F	
Source	DF	Anova SS	Mean Square	Value	Pr > F
W	1	5.44500000	5.44500000	14.70	0.0086
M	1	31.60125000	31.60125000	85.31	<.0001
W*M	1	0.72000000	0.72000000	1.94	0.2127
T	1	0.00500000	0.00500000	0.01	0.9113
W * T	1	0.03125000	0.03125000	0.08	0.7812
M*T	1	0.04500000	0.04500000	0.12	0.7393
W * M * T	1	0.78125000	0.78125000	2.11	0.1966
С	1	15.12500000	15.12500000	40.83	0.0007
W * C	1	0.45125000	0.45125000	1.22	0.3120
M * C	1	34.44500000	34.44500000	92.99	<.0001
W * M * C	1	5.28125000	5.28125000	14.26	0.0092
T*C	1	1.36125000	1.36125000	3.67	0.1037
W * T * C	1	0.40500000	0.40500000	1.09	0.3360
M * T * C	1	0.06125000	0.06125000	0.17	0.6984
P	1	0.08000000	0.08000000	0.22	0.6585
W * P	1	1.90125000	1.90125000	5.13	0.0641
M * P	1	0.32000000	0.32000000	0.86	0.3885
W * M * P	1	0.21125000	0.21125000	0.57	0.4787
T*P	1	0.45125000	0.45125000	1.22	0.3120
W * T * P	1	1.62000000	1.62000000	4.37	0.0815
M * T * P	1	0.01125000	0.01125000	0.03	0.8674
C * P	1	11.28125000	11.28125000	30.46	0.0015
W * C * P	1	0.00500000	0.00500000	0.01	0.9113
M * C * P	1	5.28125000	5.28125000	14.26	0.0092
T * C * P	1	0.00500000	0.00500000	0.01	0.9113





				F	
Source	DF	Anova SS	Mean Square	Value	Pr > F
W	1	5.44500000	5.44500000	14.70	0.0086
M	1	31.60125000	31.60125000	85.31	<.0001
W*M	1	0.72000000	0.72000000	1.94	0.2127
T	1	0.00500000	0.00500000	0.01	0.9113
W*T	1	0.03125000	0.03125000	0.08	0.7812
M*T	1	0.04500000	0.04500000	0.12	0.7393
W*M*T	1	0.78125000	0.78125000	2.11	0.1966
C	1	15.12500000	15.12500000	40.83	0.0007
W*C	1	0.45125000	0.45125000	1.22	0.3120
M*C	1	34.44500000	34.44500000	92.99	<.0001
W*M*C	1	5.28125000	5.28125000	14.26	0.0092
T*C	1	1.36125000	1.36125000	3.67	0.1037
W*T*C	1	0.40500000	0.40500000	1.09	0.3360
M*T*C	1	0.06125000	0.06125000	0.17	0.6984





P	1	0.08000000	0.08000000	0.22	0.6585
W*P	1	1.90125000	1.90125000	5.13	0.0641
M*P	1	0.32000000	0.32000000	0.86	0.3885
W*M*P	1	0.21125000	0.21125000	0.57	0.4787
T*P	1	0.45125000	0.45125000	1.22	0.3120
W*T*P	1	1.62000000	1.62000000	4.37	0.0815
M*T*P	1	0.01125000	0.01125000	0.03	0.8674
C*P	1	11.28125000	11.28125000	30.46	0.0015
W*C*P	1	0.00500000	0.00500000	0.01	0.9113
M*C*P	1	5.28125000	5.28125000	14.26	0.0092
T*C*P	1	0.00500000	0.00500000	0.01	0.9113





Significant Effects

The statistically significant effects are:

W, M, C, M*C, W*M*C, C*P, M*C*P





Half-normal Plot Method

To obtain the standardized effects for utilizing a half-normal plot to identify the statistically significant effects, we fit a full 5-way factorial model to the data.





```
PROC ANOVA;

TITLE2 'ANALYSIS USING A HALF NORMAL PLOT';

CLASSES W M T C P;

MODEL QUALITY = W|M|T|C|P;

ODS OUTPUT MODELANOVA=EFFECTS;

RUN;
```





Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	31	119.1487500	3.8435081	•	•
Error	0	0.0000000			
Corrected Total	31	119.1487500			





				F	
Source	DF	Anova SS	Mean Square	Value	Pr > F
W	1	5.44500000	5.44500000		
M	1	31.60125000	31.60125000		
W*M	1	0.72000000	0.72000000		
Т	1	0.00500000	0.00500000		
W*T	1	0.03125000	0.03125000		
M*T	1	0.04500000	0.04500000		
W*M*T	1	0.78125000	0.78125000		
C	1	15.12500000	15.12500000		
W*C	1	0.45125000	0.45125000		
M*C	1	34.44500000	34.44500000		
W*M*C	1	5.28125000	5.28125000		
T*C	1	1.36125000	1.36125000		
W*T*C	1	0.40500000	0.40500000		
M*T*C	1	0.06125000	0.06125000		
W*M*T*C	1	0.00000000	0.00000000		
P	1	0.08000000	0.08000000		
W*P	1	1.90125000	1.90125000		
M*P	1	0.32000000	0.32000000		
W*M*P	1	0.21125000	0.21125000		
T*P	1	0.45125000	0.45125000		
W*T*P	1	1.62000000	1.62000000		
M*T*P	1	0.01125000	0.01125000		
W*M*T*P	1	0.40500000	0.40500000		
C*P	1	11.28125000	11.28125000	•	
W*C*P	1	0.00500000	0.00500000	-	_





```
PROC PRINT DATA=EFFECTS;

VAR SOURCE MS;
RUN;
```





Obs	Source	MS
1	W	5.44500000
2	M	31.60125000
3	W*M	0.72000000
4	T	0.00500000
5	W*T	0.03125000
6	M*T	0.04500000
7	W*M*T	0.78125000
8	С	15.12500000
9	W*C	0.45125000
10	M*C	34.44500000
11	W*M*C	5.28125000
12	T*C	1.36125000
13	W*T*C	0.40500000
14	M*T*C	0.06125000
15	W*M*T*C	0.00000000





16	P	0.08000000
17	W*P	1.90125000
18	M*P	0.32000000
19	W*M*P	0.21125000
20	T*P	0.45125000
21	W*T*P	1.62000000
22	M*T*P	0.01125000
23	W*M*T*P	0.40500000
24	C*P	11.28125000
25	W*C*P	0.00500000
26	M*C*P	5.28125000
27	W*M*C*P	1.28000000
28	T*C*P	0.00500000
29	W*T*C*P	0.03125000
30	M*T*C*P	0.40500000
31	W*M*T*C*P	0.10125000





```
DATA EFFECTS2;

SET EFFECTS;

STDEFF = SQRT(MS);

KEEP SOURCE MS STDEFF;

RUN;
```





```
PROC RANK OUT=RANKS;

RANKS R;

VAR STDEFF;

RUN;
```





```
DATA PLOTDATA;
 SET RANKS;
 RSTAR = (R - .5)/31;
 P = (RSTAR + 1)/2;
 V = PROBIT(P);
RUN;
```





```
PROC SORT; BY R;

PROC PRINT;

TITLE3 'PRINTOUT OF RANKED
ABSOLUTE VALUES IN RANK
ORDER';
RUN;
```





10	11 112 2 2	0.1020000	0.000.0	10.0	0.10,,,	0.,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	0.0 = 100
16	M*T*C*P	0.40500000	0.63640	16.0	0.50000	0.75000	0.67449
17	W*C	0.45125000	0.67175	17.0	0.53226	0.76613	0.72616
18	T*P	0.45125000	0.67175	18.0	0.56452	0.78226	0.77984
19	W*M	0.72000000	0.84853	19.0	0.59677	0.79839	0.83587
20	W*M*T	0.78125000	0.88388	20.0	0.62903	0.81452	0.89466
21	W*M*C*P	1.28000000	1.13137	21.0	0.66129	0.83065	0.95672
22	T*C	1.36125000	1.16673	22.0	0.69355	0.84677	1.02270
23	W*T*P	1.62000000	1.27279	23.0	0.72581	0.86290	1.09346
24	W*P	1.90125000	1.37886	24.0	0.75806	0.87903	1.17016
25	M*C*P	5.28125000	2.29810	25.0	0.79032	0.89516	1.25445
26	W*M*C	5.28125000	2.29810	26.0	0.82258	0.91129	1.34874
27	W	5.44500000	2.33345	27.0	0.85484	0.92742	1.45684
28	C*P	11.28125000	3.35876	28.0	0.88710	0.94355	1.58528
29	С	15.12500000	3.88909	29.0	0.91935	0.95968	1.74695
30	M	31.60125000	5.62150	30.0	0.95161	0.97581	1.97395
31	M*C	34.44500000	5.86899	31.0	0.98387	0.99194	2.40598



PROC GPLOT;

```
TITLE 'A HALF-NORMAL PLOT OF THE EFFECTS IN CAKE QUALITY STUDY';

SYMBOL V=DOT I=NONE;

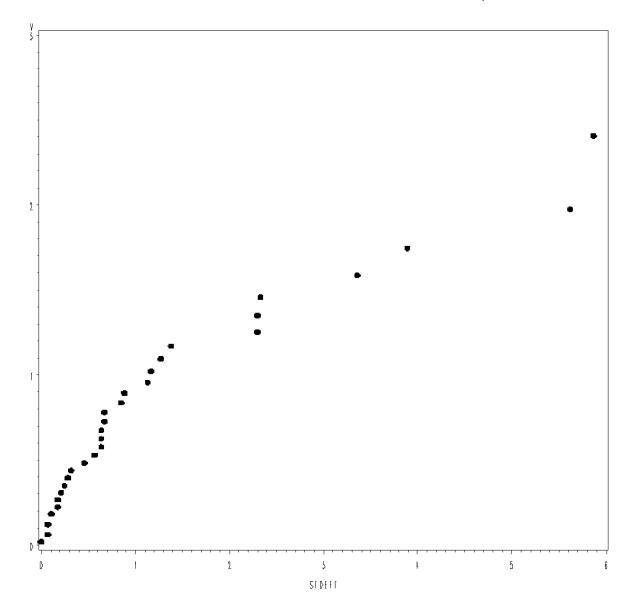
PLOT V*STDEFF;
```

RUN;





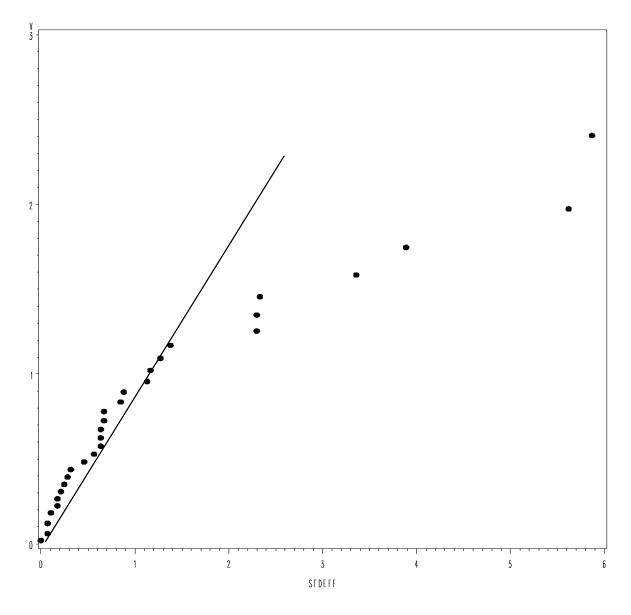
A HALF-NORMAL PLOT OF THE EFFECTS IN CAKE QUALITY STUDY







A HALF-NORMAL PLOT OF THE EFFECTS IN CAKE QUALITY STUDY



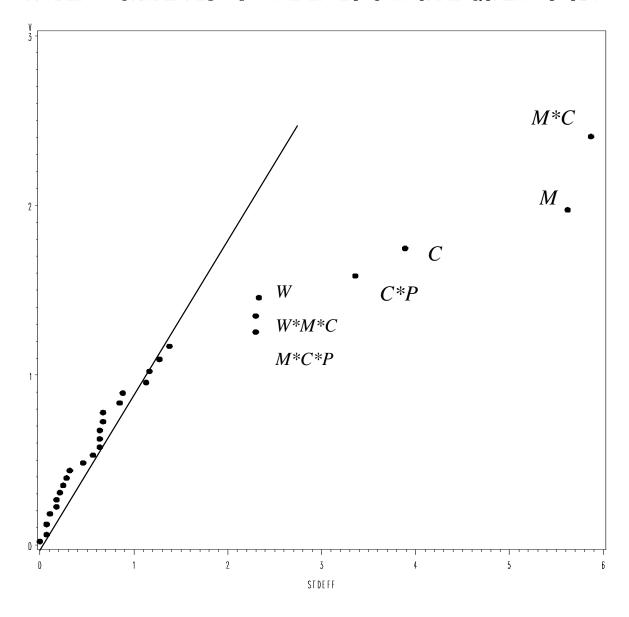




15	W*M*T*P	0.40500000	0.63640	15.0	0.46774	0.73387	0.62456
16	M*T*C*P	0.40500000	0.63640	16.0	0.50000	0.75000	0.67449
17	W*C	0.45125000	0.67175	17.0	0.53226	0.76613	0.72616
18	T*P	0.45125000	0.67175	18.0	0.56452	0.78226	0.77984
19	W*M	0.72000000	0.84853	19.0	0.59677	0.79839	0.83587
20	W*M*T	0.78125000	0.88388	20.0	0.62903	0.81452	0.89466
21	W*M*C*P	1.28000000	1.13137	21.0	0.66129	0.83065	0.95672
22	T*C	1.36125000	1.16673	22.0	0.69355	0.84677	1.02270
23	W*T*P	1.62000000	1.27279	23.0	0.72581	0.86290	1.09346
24	W*P	1.90125000	1.37886	24.0	0.75806	0.87903	1.17016
25	M*C*P	5.28125000	2.29810	25.0	0.79032	0.89516	1.25445
26	W*M*C	5.28125000	2.29810	26.0	0.82258	0.91129	1.34874
27	W	5.44500000	2.33345	27.0	0.85484	0.92742	1.45684
28	C*P	11.28125000	3.35876	28.0	0.88710	0.94355	1.58528
29	С	15.12500000	3.88909	29.0	0.91935	0.95968	1.74695
30	M	31.60125000	5.62150	30.0	0.95161	0.97581	1.97395
31	M*C	34.44500000	5.86899	31.0	0.98387	0.99194	2.40598



A HALF-NORMAL PLOT OF THE EFFECTS IN CAKE QUALITY STUDY







```
DATA ERROR;
 SET EFFECTS2;
  IF MS > 2 THEN DELETE;
RUN;
PROC RANK OUT=RANKS2;
RANKS R;
VAR STDEFF;
RUN;
```





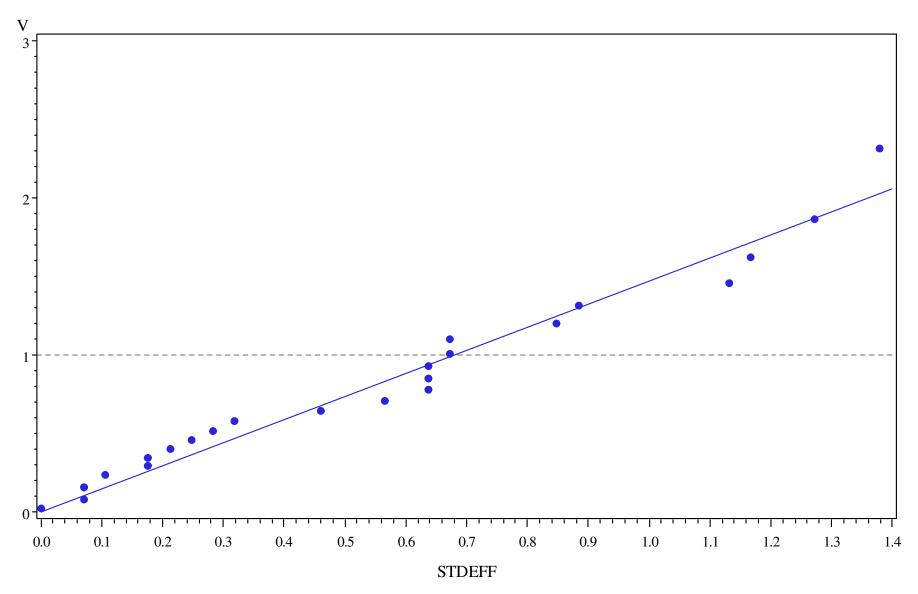
```
DATA PLOTDATA2;
 SET RANKS2;
 RSTAR = (R - .5)/24;
 P = (RSTAR + 1)/2;
 V = PROBIT(P);
RUN;
```





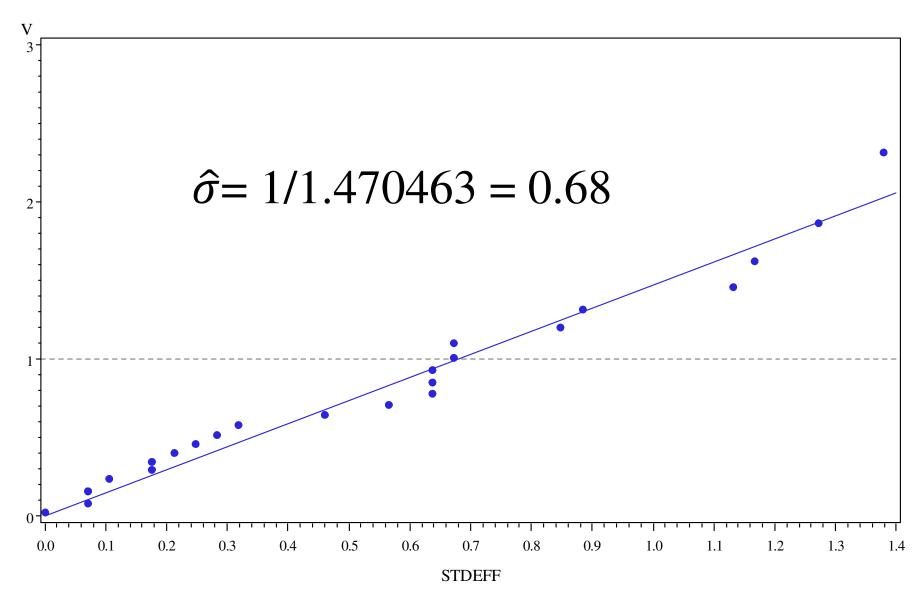
```
PROC GPLOT;
 TITLE2 'FITTED LINE THROUGH THE
ORIGIN';
 SYMBOL V=DOT I=RL0;
 PLOT V*STDEFF/VREF=1 LVREF=2
    REGEON;
 RUN;
ODS RTF CLOSE;
```

A HALF-NORMAL PLOT OF THE EFFECTS IN CAKE QUALITY STUDY FITTED LINE THROUGH THE ORIGIN



Regression Equation: V = 0 + 1 470463*STDFFF

A HALF-NORMAL PLOT OF THE EFFECTS IN CAKE QUALITY STUDY FITTED LINE THROUGH THE ORIGIN



Regression Equation: V = 0 + 1470463*STDFFF **NOTE:** In the Cake Quality study, temperature did not occur in any of the significant effects. So it should not matter whether temperature is set at 300 or 500.

NOTE: The only qualitative factor, mixer type, was involved in the significant interactions, C*P and M*C*P. Therefore, the optimum choices for the other factors, W, M, and C, will likely be different for each mixer type. Consequently, we should look at the data for each mixer separately.





Mixer Type: P₁

Mixer P1			
	Low Temp	Hi Temp	Total
(1)	4.8	3.9	8.7
W	3.9	4.2	8.1
m	5	3	8
wm	2.2	2.2	4.4
С	5.7	5.3	11
WC	2.2	2.3	4.5
mc	8.4	8.6	17
wmc	9.3	8.9	17.2





Mixer Type: P₁

	Total	Step 1	Step 2	Step 3	Std Eff
(1)	8.7	16.8	29.2	78.9	
W	8.1	12.4	49.7	-10.5	-2.625
m	8	15.5	-4.2	14.3	3.575
wm	4.4	34.2	-6.3	3.7	0.925
С	11	-0.6	-4.4	20.5	5.125
WC	4.5	-3.6	18.7	-2.1	-0.525
mc	17	-6.5	-3	23.1	5.775
wmc	17.2	0.2	6.7	9.7	2.425

$$LSD_{0.05} \doteq 2 \cdot \hat{\sigma} = 2 \cdot (0.68) = 1.36$$





The W^*M^*C interaction effect is significant for Mixer Type P_1 , therefore we will look at the W^*M^*C treatment combination means.





The W^*M^*C interaction effect is significant for Mixer Type P_1 , therefore we will look at the W^*M^*C treatment combination means.

$$Var(diff in two means) = \sigma^2 \left(\frac{1}{2} + \frac{1}{2}\right) = \sigma^2$$

$$LSD_{0.05} \doteq 2 \cdot (0.68) = 1.36$$





	Means
(1)	4.35
W	4.05
m	4
wm	2.2
C	5.5
WC	2.25
mc	8.5
wmc	8.6



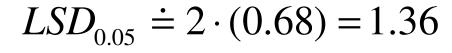




Table 7.4 Definition of the Treatment Structure in a 2⁵ Factorial Experiment to Study Cake Quality

Level

Variable	Low	High _
W: amount of water	40	120
M: mixing time	3	6
T: temperature	300	500
C: cooking oil	6	12
P: mixer type	\mathbf{P}_1	P_2
_		

Note that the treatment combinations wm and wmc are not significantly different from each other, but are significantly higher than each of the other means.

Therefore, for high cake quality with mixer P_1 , mixing time should be near 6 and cooking oil should be near 12. When mixing time is near 6 and cooking oil is near 12, it does not matter whether water is near 40 or 120.



Mixer P2					
	Total	Step 1	Step 2	Step 3	Std Eff
(1)	8.8	17.9	39.5	80.5	
W	9.1	21.6	41	-2.7	-0.675
m	11.2	13.6	-0.5	17.5	4.375
wm	10.4	27.4	-2.2	1.1	0.275
С	7.9	0.3	3.7	1.5	0.375
WC	5.7	-0.8	13.8	-1.7	-0.425
mc	13.7	-2.2	-1.1	10.1	2.525
wmc	13.7	0	2.2	3.3	0.825

$$LSD_{0.05} \doteq 2 \cdot \hat{\sigma} = 2 \cdot (0.68) = 1.36$$





The significant effects for mixer type P_2 are M and M*C. So we look at the M*C treatment combination means for this mixer type.





Mixer P2					
	(1)	t	W	tw	Total
(1)	4.2	4.6	5	4.1	17.9
m	5.8	5.4	5.2	5.2	21.6
С	2.9	5	3	2.7	13.6
mc	6.7	7	6.6	7.1	27.4





Mixer P2				
	Total	Step 1	Step 2	Std Eff
(1)	17.9	39.5	80.5	
m	21.6	41	17.5	4.375
С	13.6	3.7	1.5	0.375
mc	27.4	13.8	10.1	2.525

$$LSD_{0.05} \doteq 2 \cdot \hat{\sigma} = 2 \cdot (0.68) = 1.36$$





trt comb	means
(1)	4.475
m	5.4
C	3.4
mc	6.85





$$Var(diff in two means) = \sigma^2 \left(\frac{1}{4} + \frac{1}{4}\right) = \frac{\sigma^2}{2}$$

$$LSD_{0.05} \doteq \frac{2 \cdot (0.68)}{\sqrt{2}} = 0.962$$





trt comb	means
(1)	4.475
m	5.4
C	3.4
mc	6.85

$$LSD_{0.05} \doteq \frac{2 \cdot (0.68)}{\sqrt{2}} = 0.962$$





We see that the treatment combination mc is significantly higher than each of the other treatment combination means. Thus when using Mixer Type P_2 , mix time should be near 6 and cooking oil should be near 12. The amount of water used and the temperature used does not matter.





trt comb	Means
(1)	4.35
W	4.05
m	4
wm	2.2
C	5.5
WC	2.25
mc	8.5
wmc	8.6

Mixer P2

trt comb	means
(1)	4.475
m	5.4
C	3.4
mc	6.85

trt comb	Means
(1)	4.35
W	4.05
m	4
wm	2.2
C	5.5
WC	2.25
mc	8.5
wmc	8.6

Mixer P2

trt comb	means
(1)	4.475
m	5.4
С	3.4
mc	6.85

How do mc and wmc for P1 compare to mc for P2?

trt comb	Means
(1)	4.35
W	4.05
m	4
wm	2.2
C	5.5
WC	2.25
mc	8.5
wmc	8.6

Mixer P2

trt comb	means
(1)	4.475
m	5.4
С	3.4
mc	6.85

How do mc and wmc for P1 compare to mc for P2?

Recall that each of the mixer P1 means are the mean of two observations, and that each of the P2 means are the average of four observations. Therefore, the difference between a P1 mean and and P2 mean will have a standard error equal to

$$\hat{\sigma}\sqrt{\frac{1}{2} + \frac{1}{4}} = \hat{\sigma}\sqrt{\frac{3}{4}}$$

and the 5% LSD value would be approximately

LSD =
$$2\hat{\sigma}\sqrt{\frac{3}{4}} = 2 \cdot (0.68)(0.866) = 1.178$$

trt comb	Means
(1)	4.35
W	4.05
m	4
wm	2.2
C	5.5
WC	2.25
mc	8.5
wmc	8.6

Mixer P2

trt comb	means
(1)	4.475
m	5.4
С	3.4
mc	6.85

$$LSD = 1.178$$

Mixer P1

trt comb	Means
(1)	4.35
W	4.05
m	4
wm	2.2
C	5.5
WC	2.25
mc	8.5
wmc	8.6

Mixer P2

trt comb	means
(1)	4.475
m	5.4
С	3.4
mc	6.85

LSD = 1.178

Thus, we want to use Mixer P1 with mix time near 6 and cooking oil near 12. The amount of water and temperature don't matter.

You can now do Assignment 2.