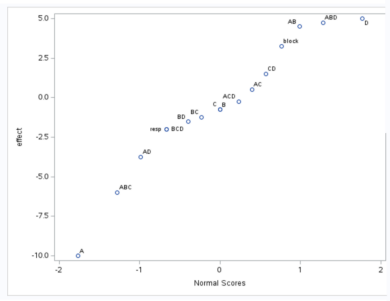
```
/* Montgomery 7.5 */
proc import datafile="/home/u63048916/STAT571B/Homework/Homework 6/Q7-5.xlsx"
    dbms=xlsx
    out=mont7_5
    replace;
    getnames=yes;
run;

data inter;
    set mont7_5;
        A=A;
        B=B;
        C=C;
        D=D;
        AB=A*B;
        AC=A*C;
        AD=A*D;
        BC=B*C;
        BD=B*D;
        CD=C*D;
        ABC=AB*C;
        ABD=AB*D;
        ACD=AC*D;
        BCD=BC*D;
        block=ABC*D;
        resp=Yield;

proc glm data=inter;  /* GLM Proc to Obtain Effects for unreplicated case */
 class A B C D AB AC AD BC BD CD ABC ABD ACD BCD block;
 model resp=block A B C D AB AC AD BC BD CD ABC ABD ACD BCD;
 run;


proc reg outest=effects data=inter;    /* REG Proc to Obtain Effects */
 model resp=A B C D AB AC AD BC BD CD ABC ABD ACD BCD block;
 run;

 proc print data=effects;
 run;

data effect2; set effects;
 drop y intercept _RMSE_;
 run;

proc transpose data=effect2 out=effect3;
run;

data effect4; set effect3; effect=col1*2;
run;

proc sort data=effect4; by effect;
run;

proc print data=effect4;
run;

proc rank data=effect4 out=effect5 normal=blom;
 var effect;
ranks neff;
 run;

proc sgplot data=effect5;
scatter x=neff y=effect/datalabel=_NAME_;
xaxis label='Normal Scores';
run;

/* rerun glm with specific significant effects */
proc glm data=inter;                  /* GLM Proc to Obtain Effects */
 class A B AB block;
 model resp= A B C D AB AC AD BC BD ABC ABD block;
 output out=two r=res p=pred;
 run;

proc univariate data=two normal;
var res;
qqplot;
run;
/* check constant variance using graph*/
title 'residual plot: res vs predicted value ';
proc sgplot data=two;
scatter x=pred y=res;
refline 0;
run;
```

7.5. Consider the data from the first replicate of Problem 6.7. Construct a design with two blocks of eight observations each with *ABCD* confounded. Analyze the data.

6.7. An experiment was performed to improve the yield of a chemical process. Four factors were selected, and two replicates of a completely randomized experiment were run. The results are shown in the following table:

Treatment Combination	Replicate I	Replicate II	Treatment Combination	Replicate I	Replicate II
(1)	90	93	*d*	98	95
a	74	78	*ad*	72	76
b	81	85	*bd*	87	83
ab	83	80	*abd*	85	86
c	77	78	*cd*	99	90
ac	81	80	*acd*	79	75
bc	88	82	*bcd*	87	84
abc	73	70	*abcd*	80	80

Factor Analysis



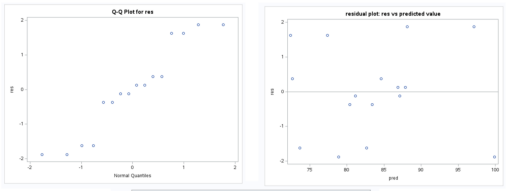
Dependent Variable: resp

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	12	934.5000000	77.8750000	9.25	0.0463
Error	3	25.2500000	8.4166667		
Corrected Total	15	959.7500000			

R-Square	Coeff Var	Root MSE	resp Mean
0.973691	3.479639	2.901149	83.37500

Source	DF	Type III SS	Mean Square	F Value	Pr > F	
A	1	400.0000000	400.0000000	47.52	0.0063	←
B	1	2.2500000	2.2500000	0.27	0.6408	
C	1	2.2500000	2.2500000	0.27	0.6408	
D	1	100.0000000	100.0000000	11.88	0.0410	←
AB	1	81.0000000	81.0000000	9.62	0.0532	←
AC	1	1.0000000	1.0000000	0.12	0.7531	
AD	1	56.2500000	56.2500000	6.68	0.0814	←
BC	1	6.2500000	6.2500000	0.74	0.4522	
BD	1	9.0000000	9.0000000	1.07	0.3772	
ABC	1	144.0000000	144.0000000	17.11	0.0256	←
ABD	1	90.2500000	90.2500000	10.72	0.0466	←
block	1	42.2500000	42.2500000	5.02	0.1109	

Figure 7.5.1
Residual Normality Diagnostics



Tests for Normality

Test	Statistic		p Value	
Shapiro-Wilk	W	0.905865	Pr < W	0.0998
Kolmogorov-Smirnov	D	0.144801	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.072576	Pr > W-Sq	0.2450
Anderson-Darling	A-Sq	0.534162	Pr > A-Sq	0.1479

-2: Supposed to make model based on selected factors and do analysis on that model