Report:

1-Introduction: we are looking at effects of the percentage of hardwood concentration in raw pulp, vat pressure, and pulp cooking duration on paper strength.

There are three levels of hardwood concentration (2,4,8), three pressure levels (400,500,650), and two cooking times (3 hours, 4 hours) to choose from.

After conducting a 3-factorial way anova design using R studio, we get the following results:

```
                                 Df Sum Sq Mean Sq F value   Pr(>F)
concentration                     2  7.764   3.882  10.619   0.0009 ***
pressure                          2 19.374   9.687  26.499 4.33e-06 ***
cookingtime                       1 20.250  20.250  55.395 6.75e-07 ***
concentration:pressure            4  6.091   1.523   4.166   0.0146 *
concentration:cookingtime         2  2.082   1.041   2.847   0.0843 .
pressure:cookingtime              2  2.195   1.097   3.002   0.0750 .
concentration:pressure:cookingtime 4  1.973   0.493   1.350   0.2903
Residuals                        18  6.580   0.366
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Looking at the p-value of the result we can see that concentration level, pressure, cooking time, the interaction between concentration level and pressure, the interaction term between concentration levels and cooking time are all significant in a 5% level. Whereas The interaction between concentration level and pressure, and the interactions between pressure and cooking time are significant in the 10% level. However, the interaction between all the factor concentrations, pressure, and cooking time is not significant both levels 5% and 10% since the p-value is bigger than 0.1.

We are interested in a 5% significance level; therefore, we are going to take out non-significant term from our model those with p-value bigger than 0.05.

```
                    Df Sum Sq Mean Sq F value   Pr(>F)
concentration        2  7.764   3.882   7.867  0.00213 **
pressure             2 19.374   9.687  19.631 6.37e-06 ***
cookingtime          1 20.250  20.250  41.037 8.71e-07 ***
concentration:pressure  4  6.091   1.523   3.086  0.03322 *
Residuals           26 12.830   0.493
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

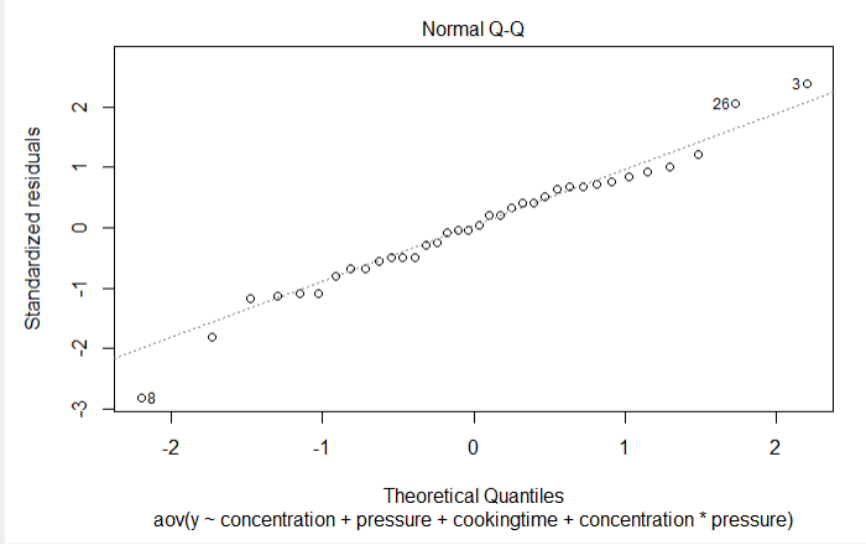
Model selected:

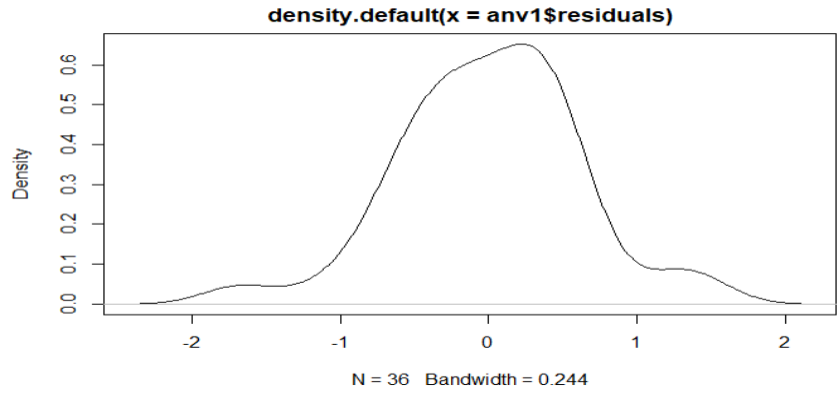
Hence, our model will be y= concentration + pressure +cookingtime + concentration*pressure.

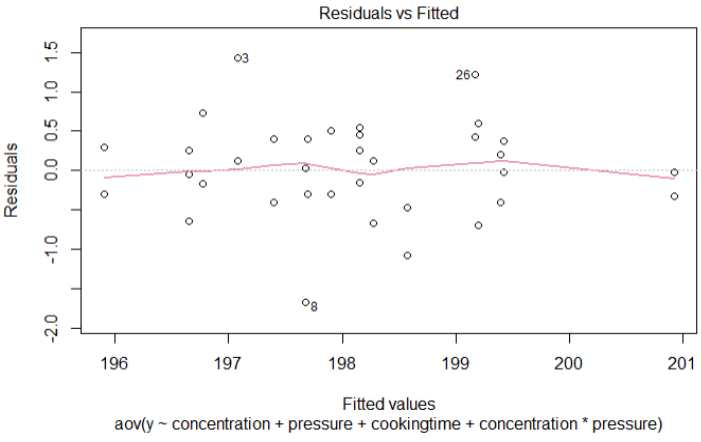
Residual analysis:

Then we Check the model's adequacy from residual plots.

the QQ-plot shows that the data is normally distributed, moreover the density plot confirms the normality of the data. also, the residuals in the fitted plot show that there is no pattern meaning that the variance is constant, and the data is independent.







2-introduction: we are looking at the vibration level at the surface of circuit board generated by a router to cut locating notches. There are two factor that influence vibration: (A) bit size and (B) cutting speed. We selected Two-bit sizes (1/16 and 1/8 inch) and two speeds (40 and 90 rpm). Then we repeated the same experiment on 4 boards.

After conducting a 2^2 factorial design we get the following results:

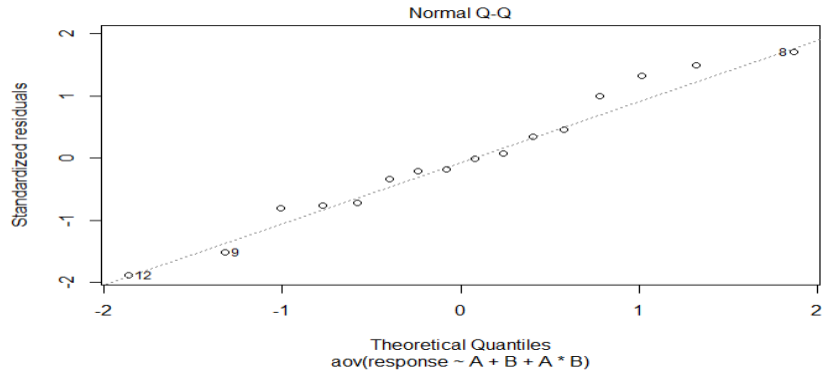
```
            Df Sum Sq Mean Sq F value   Pr(>F)
A            1 1107.2  1107.2  185.25 1.17e-08 ***
B            1  227.3   227.3   38.02 4.83e-05 ***
A:B          1  303.6   303.6   50.80 1.20e-05 ***
Residuals   12   71.7     6.0
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

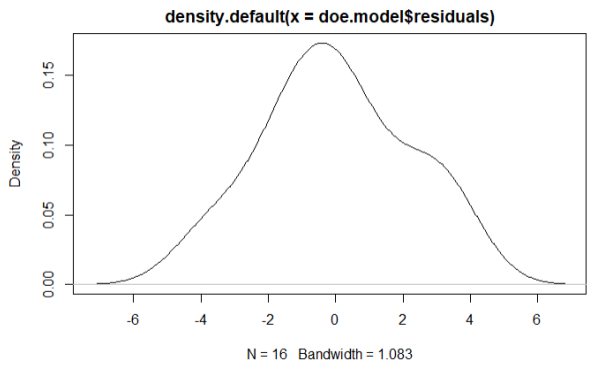
Form the p-value we see that all the factor A, B are significant including the interaction term A*B.

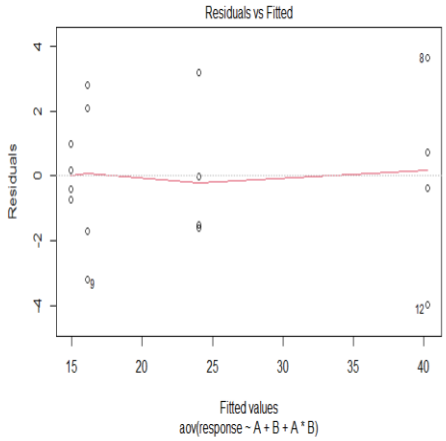
Model selected: Therefore, our model will be y= A + B+ A * B

Residual analysis:

Then we Construct a normal probability plot of the residuals and plot the residuals versus the predicted vibration level. the QQ-plot shows that the data is normally distributed, moreover the density plot confirms the normality of the data. also, the residuals in the fitted plot show that there is no pattern meaning that the variance is constant, and the data is independent







3-Introduction: The quality control department are looking at the effect of 3 factors on the dyeing of cotton-synthetic cloth used to manufacture men's shirts. First factor they have is the three operators (1,2,3), then they selected three cycle times (40,50,60).the last factor they selected is tow temperatures (300, 350).

After conducting a 3-factorial way anova design using R studio, we get the following results:

```
                          Df Sum Sq Mean Sq F value   Pr(>F)
operator                   2  261.3  130.67  39.864 7.44e-10 ***
cycle                      2  436.0  218.00  66.508 8.14e-13 ***
temperature                1   50.1   50.07  15.277 0.000393 ***
operator:cycle             4  355.7   88.92  27.127 1.98e-10 ***
operator:temperature       2   11.3    5.63   1.718 0.193895
cycle:temperature          2   78.8   39.41  12.023 0.000100 ***
operator:cycle:temperature 4   46.2   11.55   3.523 0.015870 *
Residuals                 36  118.0    3.28
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

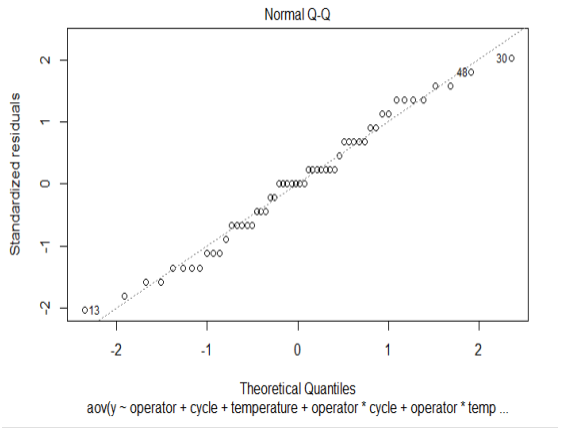
Based on the p-value we can see that all the factors are significant at 5% level, besides all the interaction terms except for the interaction between operator and temperature.

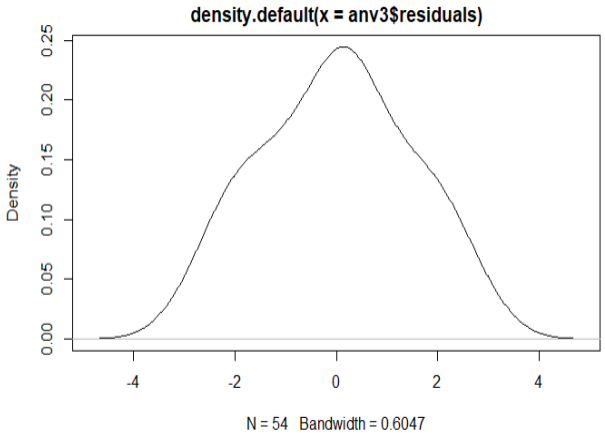
Thus, our model will be as the following:

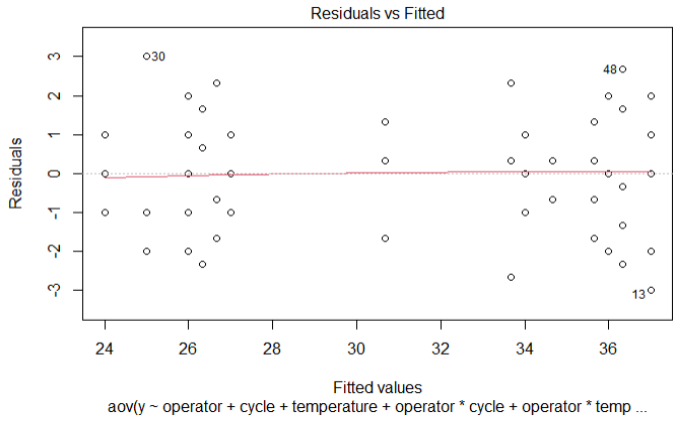
Model selected: y= operator + cycle +temperature + operator*cycle +operator*cycle*temperature.

Residual analysis:

There is nothing unusual about the residual plots. the QQ-plot shows that the data is normally distributed, moreover the density plot confirms the normality of the data. also, the residuals in the fitted plot show that there is no pattern meaning that the variance is constant, and the data is independent







4-introduction: Four factors were selected (A, B, C,D), to see the effect of an experiment that was performed to improve the yield of a chemical process. This experiment was replicated twice.

After conducting a 2^4 factorial design we get the following results:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
A	1	657.0	657.0	85.816	7.87e-08	***
B	1	13.8	13.8	1.800	0.198445	
C	1	57.8	57.8	7.547	0.014317	*
D	1	124.0	124.0	16.200	0.000979	***
A:B	1	132.0	132.0	17.245	0.000749	***
A:C	1	3.8	3.8	0.494	0.492302	
B:C	1	2.5	2.5	0.331	0.573296	
A:D	1	38.3	38.3	5.000	0.039945	*
B:D	1	0.3	0.3	0.037	0.850417	
C:D	1	22.8	22.8	2.976	0.103793	
A:B:C	1	215.3	215.3	28.118	7.15e-05	***
A:B:D	1	175.8	175.8	22.959	0.000200	***
A:C:D	1	7.0	7.0	0.918	0.352162	
B:C:D	1	7.0	7.0	0.918	0.352162	
A:B:C:D	1	47.5	47.5	6.208	0.024077	*
Residuals	16	122.5	7.7			

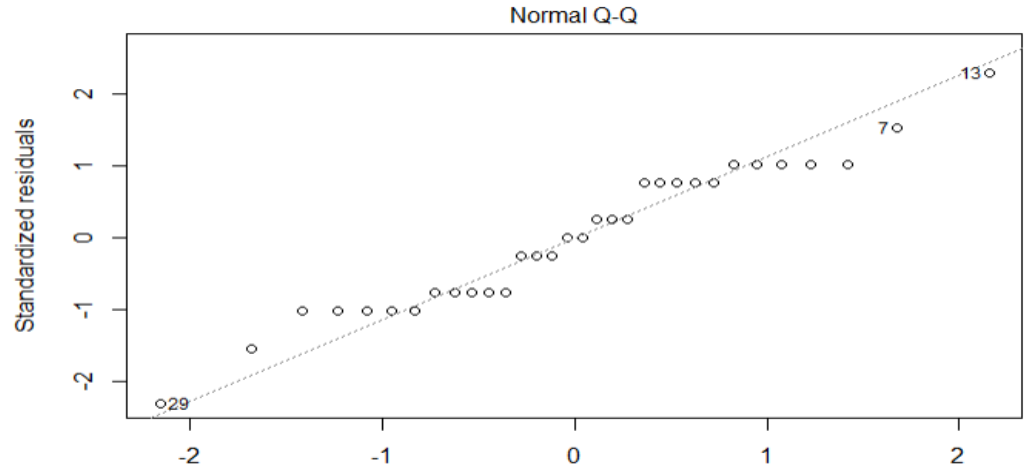
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

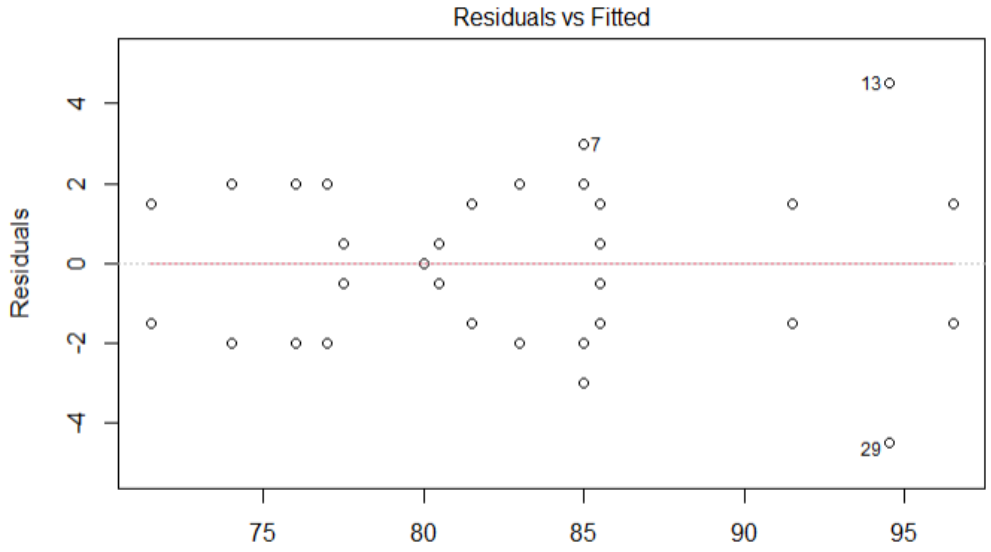
From R outcome we can see that some factors and some interaction terms are not significant at 5 % level, therefore we are going to take them out from our prediction model. And thus, our model will be as following:

Model selected: Y = A + C + D + A*B + A*D + A*B*C + A*B*D + A*B*C*D

Residual analysis:

There is nothing unusual about the residual plots. the QQ-plot shows that the data is normally distributed. also, the residuals in the fitted plot show that there is no pattern meaning that the variance is constant, and the data is independent


Normal Q-Q
Standardized residuals
-2
-1
0
1
2
13
7
29


Residuals vs Fitted
Residuals
-4
-2
0
2
4
75
80
85
90
95
13
7
29