For the purpose of this project, we want to find the best combination of Line Speed and Loading of Additives that maximize the production capacity for the manufacturer. The main method used for this project is 2-way ANOVA analysis

```
Analysis of Variance Table

Response: rate
                          Df Sum Sq Mean Sq F value  Pr(>F)
loading_reform             2  28022 14011.1  4.1230 0.03357 *
speed_reform               2  23945 11972.3  3.5230 0.05114 .
loading_reform:speed_reform 4   7844  1961.1  0.5771 0.68293
Residuals                 18  61169  3398.3
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

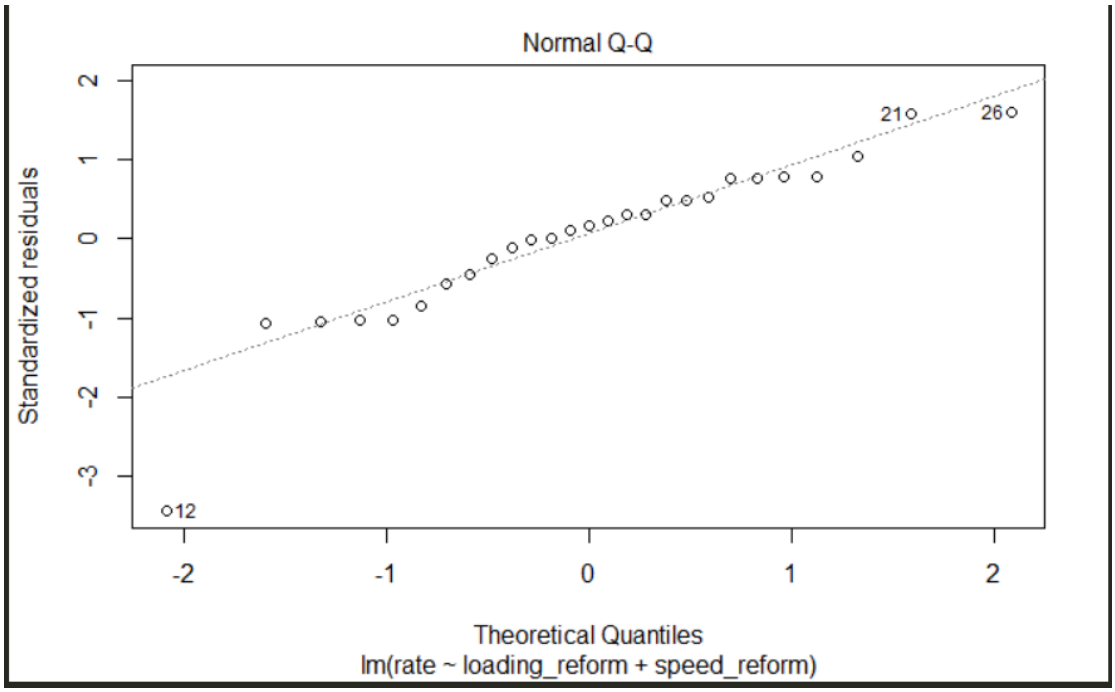
I start with the full model include factors Line Speed and Loading of Additive and also include the interaction term. However, by looking at the anova table for the full model, the p value for interaction term is about 0.683, therefore we can conclude that the interaction term is insignificant

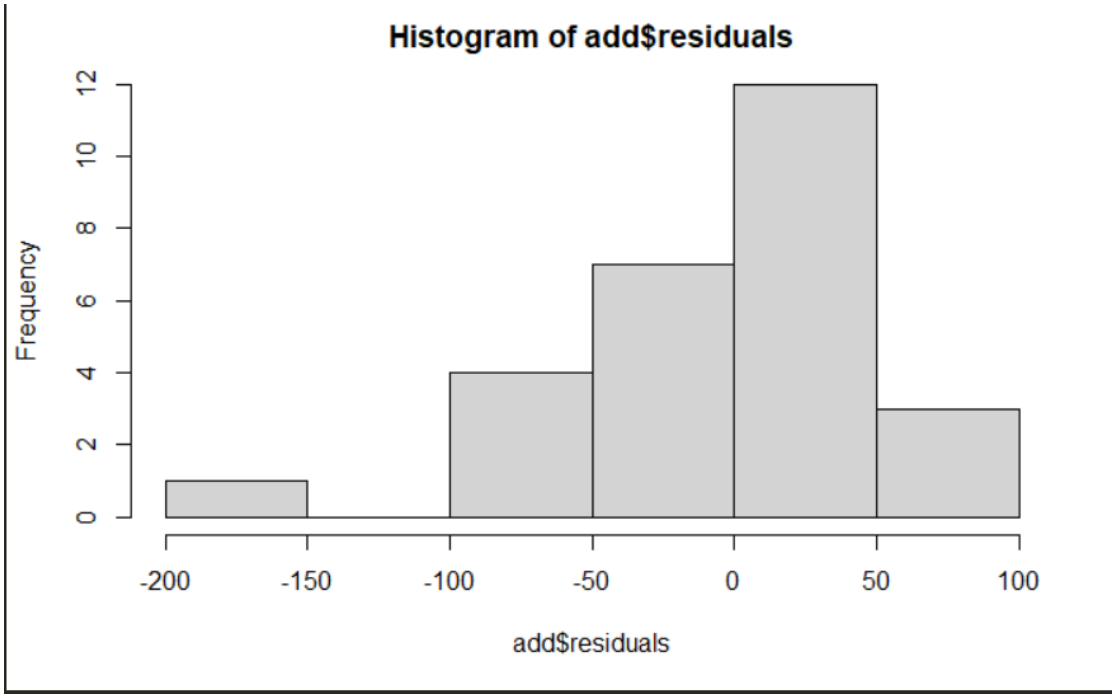
```
Analysis of Variance Table

Model 1: rate ~ speed_reform
Model 2: rate ~ loading_reform + speed_reform
  Res.Df   RSS Df Sum of Sq      F  Pr(>F)
1     24 97036
2     22 69014  2     28022 4.4664 0.02355 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Analysis of Variance Table

Model 1: rate ~ loading_reform
Model 2: rate ~ loading_reform + speed_reform
  Res.Df   RSS Df Sum of Sq      F  Pr(>F)
1     24 92958
2     22 69014  2     23945 3.8165 0.03777 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Since the interaction term is insignificant, I refit an additive model with factors Line Speed and Loading of Additives without interaction. Also, to test the significance of each factor, I fit two reduced model each one has only 1 corresponding factor. Then by looking at the anova table for reduced model with Line Speed only vs additive model and reduced model with Loading of Additives only vs additive model. The P value is 0.02355 and 0.03777 for each. Thus we conclude that the full additive model is needed and both factors are statistically significant.

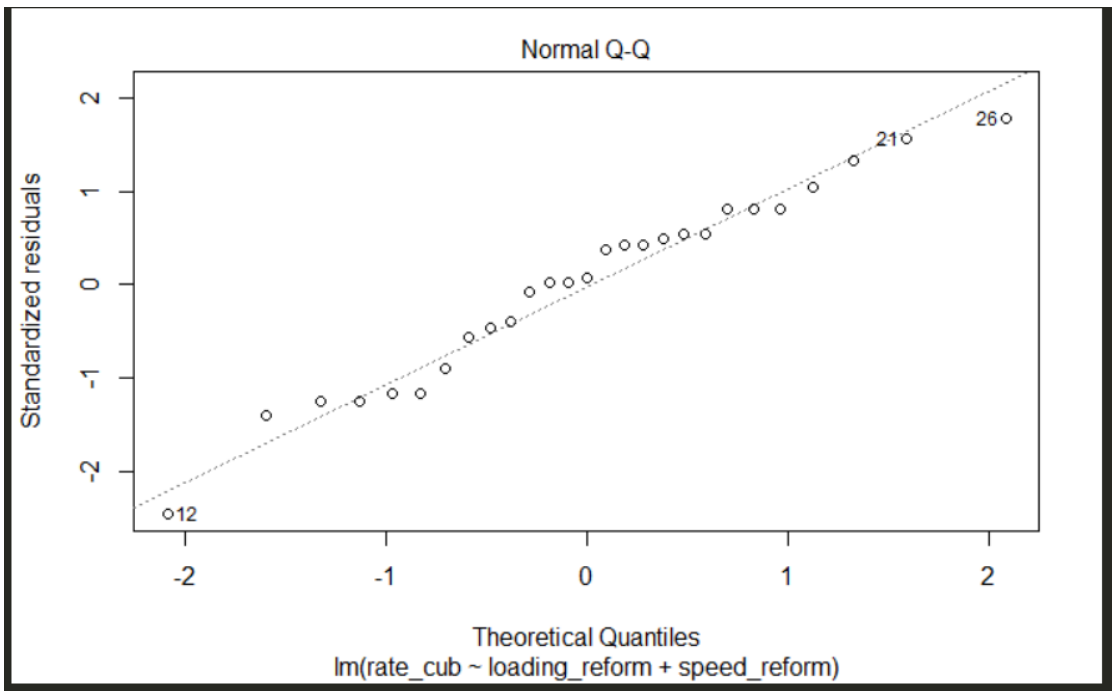
```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = rate ~ loading_reform + speed_reform, data = bubblewrap)

$loading_reform
         diff        lwr       upr     p adj
2-0 -1.111111 -67.436661  65.21444 0.9990241
4-0 67.777778   1.452228 134.10333 0.0445852
4-2 68.888889   2.563339 135.21444 0.0408082

$speed_reform
           diff        lwr        upr     p adj
37-36  44.33333  -21.99222 110.658883 0.2353973
38-36 -28.00000  -94.32555  38.325550 0.5478326
38-37 -72.33333 -138.65888  -6.007784 0.0308834
```

Then to choose the best combination that results in highest production rate, we use the TukeyHSD function to find the pairwise difference for levels in each factor. By looking at the result we find that the Loading of Additives type 2 and 0's difference is insignificant and type 4 is greater than type 2 and type 0. The Line Speed 37 and 36's difference is not significant. Also, the difference between 38 and 36 is not significant. However, the difference between 38 and 37 is significant and type 37 is greater than type 38. Thus, we conclude that the best combination is Loading of Additive 2 and Line Speed 37.





Finally, I did some diagonal check. First by looking at the half normal plot and histogram of residuals, the normality assumption seems fail. To verify this hypothesis, I did Shapiro test since the data size is small. The p value for Shapiro test is 0.01321 thus we can reject the normality assumption. Then I did the bp test for the constant variance assumption, the p value is about 0.117. Thus, the constant variance assumption holds.



To recover the normality, I observed the histogram of residuals and find it has a left tail. Thus, I tried that transform the production rate to its cube and fit a new additive model. The data in graph above is approximately in a straight line and it seems that we recover the normality by this transformation. To verify this, I did another Shapiro

test for the new model and the p value is about 0.598 thus we can conclude that normality assumption holds for the new model

In conclusion, to maximize the production rate the best combination is Loading of Additive 2 and Line Speed 37.