Design and Analysis of Experiments

Dr. Crunk

Kaylyn Vo

May 22nd, 2020

On the Rise of Bread: Which Variables Contribute to the Height of Bread?

Executive Summary

Motivation

The experiment was partly motivated by a statement by Gemma Stafford of Bigger Bolder Baking on baking bread that “Mixing salt and yeast directly inhibits the rise of bread because salt kills yeast.” Stafford and several other bread’s making recipes[4] recommend adding salt only after the initial mix of yeast with other ingredients to maximize the livelihood of the active yeast. We built a statistical model and conducted experiments to examine whether the average rise of a bread is affected by different methods of combining of salt and active yeast in flour.

Objective

The objective of the experiment is to identify and analyze the effects of three different ingredients and their combinations on rise of a bread. The variables that we are interested in are the amount of active yeast used, fermentation time and placement of salt and active yeast in flour.

Result

The result of the experimental design shows that amount of active yeast used in the flour mixture, combination of salt and active yeast, and the interaction between time and combination have significant influences on the rise a bread. Additional result shows that not including salt in the flour mixture contribute to the highest rise of a bread. To save time and resources while maximizing profits, it is recommended that the flour mixture only including yeast as an active ingredient.

Table of Contents

Analysis of the Problem

Problem Statement

A good bread is defined by the flavorful golden crispy crust on the outside and the soft and fluffiness on the inside. The soft and airy component is partly due to the reaction of active yeast in combination with other ingredients in the dough during the fermentation process. During the process, yeast begins feeding on the sugars of the flour and releases carbon dioxide that makes the bread rise and increase in volume[2]. Therefore, to ensure the maximum height of a bread, the yeast has to be viable and reactive with other ingredients.

Another component that influences the rise of a bread is gluten. It helps the bread rise by trapping gas bubbles during fermentation process. Gluten is formed when two protein, glutenin and gliadin, came into contact with water[3]. Gluten development occurs when water is added flour and continues to develop during the mixing process. In the experiment, the mixing time is not of our interest, therefore it is held constant for all batches of dough. We are interested in the fermentation time and how varying the time could affect the rise of a bread.

The last variable is the combination of salt and yeast. A good bread is not only defined by the golden crust and the airy component but also the flavor. We add flavor by adding salt to the flour mixture; however, combining of salt directly with yeast may adversely affect the rise of a bread. Many recipes and bakers recommend to separate salt and yeast during the mixing stage to reduce direct contact of yeast with salt. We will examine different combinations of salt and yeast and test whether the recommendation is true.

The response variable

Fluffiness and airy are characteristics of a properly proofing bread; however, they can't be directly measured. We will measure the height of the bread because it is directly related to the amount of active used, fermentation time, and the combination of salt and active yeast.

Factors

The height of a bread is dependent on the amount of active yeast used, fermentation time, mixing time, temperature during fermentation process, type of flour, temperature of the oven, and other factors. In this experiment, we chose to study the influences on the rise of a bread of three variables: fermentation time, amount of active yeast, combination of active and salt, and their interactions.

Factors Levels

Factor	Level 1	Level 2	Level 3
Active Yeast	2.625 g	5.25 g	10.5 g
Fermentation Time	3 Hours	5 Hours	
Combination of Salt and Active Yeast	No salt	Mixed	Side

Description of Factor Levels

Active Yeast: The amount of yeast is divided into three levels. The variable was chosen because it is directly related to the rise of bread.

Fermentation Time: The factor is divided into two levels and in the experiment, we want to examine whether the proofing time has an influenced on the rise of bread.

Combination of Salt and Active Yeast: The factor is divided into 3 levels: No salt, Mixed, and Side. The level No salt means that the flour mixture doesn't include salt. The level Mixed means that salt and yeast were mixed thoroughly with water before they were added to the flour. The level Side means that salt was mixed with flour and water thoroughly before the yeast was added to the flour. The factor was chosen because we want to determine how salt interacts with yeast and whether it could suppress the rise of a bread.

Choice of Experimental Design

Original, the full 3^k factorial design was considered; however, due to limited resources, the number of levels of Fermentation Time was reduced to two. In addition, we are more interested in the factors Active Yeast and Combination of Salt and Yeast, therefore we want to have more levels and observations of the two factors. In the model, we assumed that all factors are fixed. The model was fit with the main factors and all the possible interactions. There were two replications for each level combination.

Performing the Experiment

Ingredients for Each Bread:

- 64 grams of Gold Medal All Purpose Flour
- X grams of Morton Iodized Salt
- 32 grams of water with temperature around 72°F
- X grams of active yeast

Constant Factors:

- Mixing the dough for each bread for 3 minutes
- Containers to store the dough
- Muffin baking tray
- Baking oven was preheated before each experiment
- Temperature of the oven: 450°F

- Baking time: 30 minutes
- Each bread was cool down for 10 minutes before measurement
- 10 pounds of Gold Medal All Purpose Flour
- 1 pound of Morton Iodized Salt
- Fleischmann’s Active Dry Yeast
- Plastic ruler
- Bread knife
- Mixing glass bowl

Prepare the experiment:

The order of each batch of dough was prepared randomly. For each batch, 64 grams of flour, 32 grams of water, x amount of active yeast, and different combination of salt and active yeast were mixed for three minutes. The order for which ingredient was placed in the mixing bowl was the same for all batches. After 3 minutes of mixing, the dough was transferred to a container for fermentation. After x amount of fermentation time, each batch was divided equally into two smaller batches and rolled into two uniform and round batches. They were placed in a muffin tray and baked for 30 minutes at 450°F. After 30 minutes, the breads were cool down for 10 minutes before measurement. The two breads were cut right in the middle using a bread knife and were measured by a ruler. There were two breads for each batch, each bread was measure twice and the average of the measurements was recorded as an observation. The choice of cutting each batch into two and were placed in a muffin try was to prevent the bread expanding rather than rising without a confined space which made it difficult to determine the correct height.

Statistical Analysis

Data:

Time	**3 Hours**	**3 Hours**	**3 Hours**		**5 Hours**	**5 Hours**	**5 Hours**
Amount of Active Yeast	**No Salt**	**Mixed**	**Side**		**No Salt**	**Mixed**	**Side**
2.625 g	6.15 6.575	6.05 5.825	5.15 5.25		6.375 6.6375	5.7 5.7	6 5.8
5.25 g	6.2 7	6.425 5.625	5.675 5.5		6.175 6.525	5.95 5.825	5.65 5.725
10.5	6.275 7.1	6.25 6.575	5.675 6.025		6.4 6.7	6.3 6.225	5.725 5.8

Model of Original Response

The complete interaction model for a three-factor completely randomized design is:

$$y_{ijkl} = \mu + T_i + \beta_j + \gamma_k + (T\beta)_{ij} + (T\gamma)_{ik} + (\beta\gamma)_{jk} + (T\beta\gamma)_{ijk} + \varepsilon_{ijkl} \quad \textbf{(1.1)}$$

T_i: Factor A, Amount of active yeast used at level i^{th} and i = {1,2,3} = {2.625, 5.25, 10.5}.

β_j: Factor B, the combination of active and salt at level j^{th} and j = {1,2,3} = {No Salt, Mixed, Side}.

γ_k: Factor C, the number of hours for flour fermentation at level k^{th} and k = {1,2} = {3 Hours, 5 Hours}.

ε_{ijkl}: the random error associates with every batch of bread from $(i,j,k)^{th}$ treatment and replication l where l = {1,2}. We assume $\varepsilon_{ijkl} \sim N(0, \sigma^2)$.

y_{ijkl} : the height of a bread from $(i,j,k)^{th}$ treatment and replication l where l = {1,2}.

μ: the overall mean.

Models Diagnostic 1

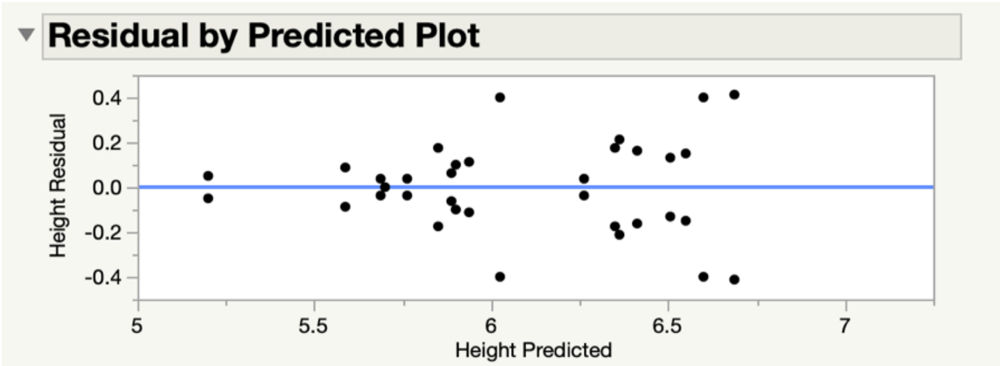

Figure 1.A

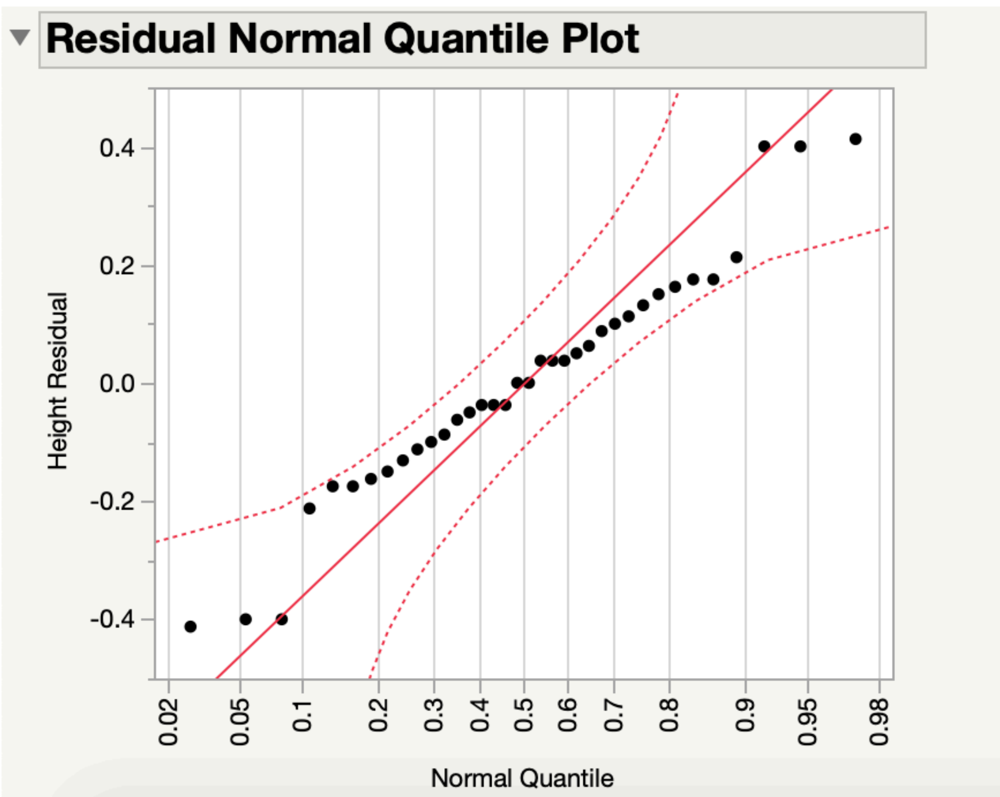

Figure 1.B

The model was built under the assumptions that errors of the model must have a constant variance and independently distributed as Normal(0, σ^2). There is generally no ordering to the levels of the factor variables, therefore we won't be looking for out-ward opening funnel rather, we will look for large differences in the vertical spread in the residual versus the response plot.

In figure 1.A, the assumption of constant variance is violated as the residuals exhibit an increase trend as the predicted response increases and in figure 1.B, the quantiles of the residuals diverge from the theoretical Normal quantiles. Therefore, the response variable must be transformed to stabilizing the variance of the response and making the response distribution to approximately distributed as Normal.

Model with Transformed Response

The Box-Cox method was used to find the best power that minimizes the model's Sum of Error. The best transformation is $\lambda = -2$ and the corresponding model is y^{-2} which is a monotonically decreasing function for y > 0, therefore we can transform back to the original response variable. The model was refitted with the transformed response variable.

The transformed fitted model:

$\frac{1}{y_{ijkl}^2} = \mu + T_i + \beta_j + \gamma_k + (T\beta)_{ij} + (T\gamma)_{ik} + (\beta\gamma)_{jk} + (T\beta\gamma)_{ijk} + \varepsilon_{ijkl}$, where all the parameters have the same interpretations in 1.1.

Model Diagnostic 2

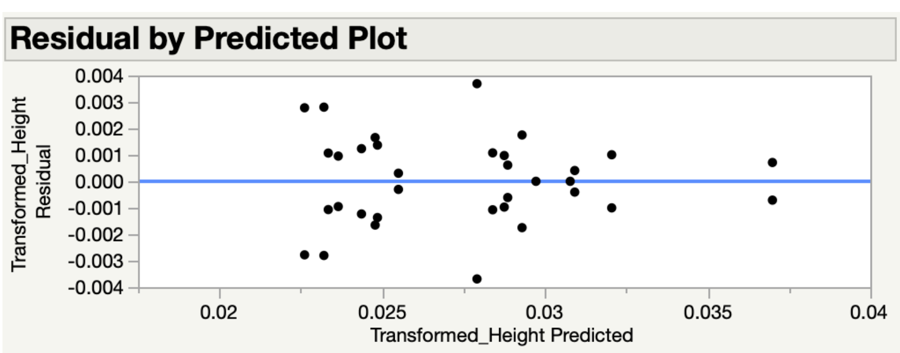

Figure 2.A: Residual by Transformed Predicted Variable

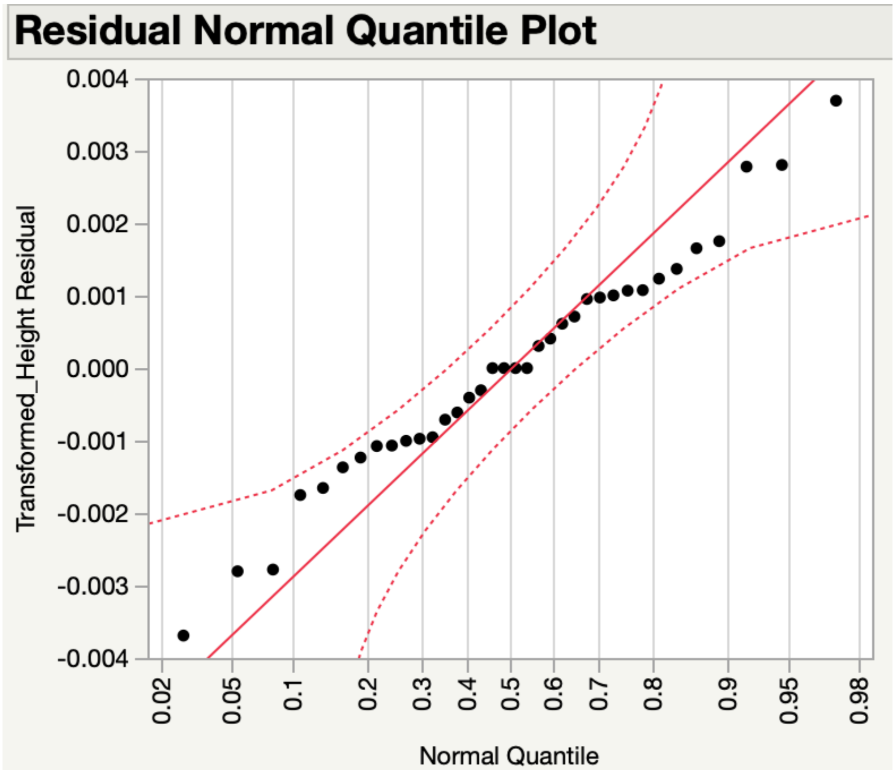

Figure 2.B: Sample Quantiles of Transformed Response by Normal Quantiles

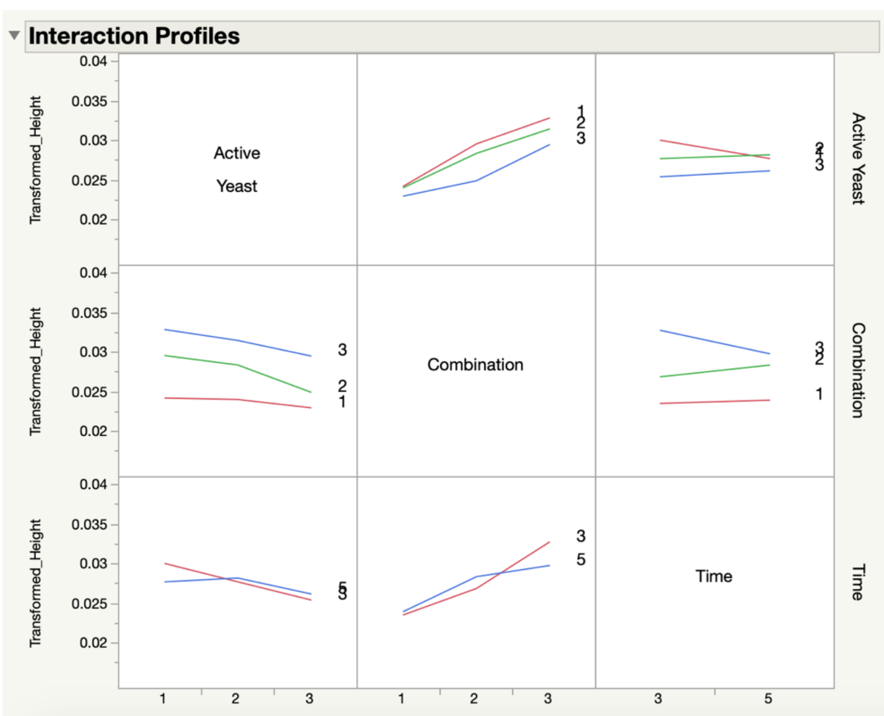

Figure 2.C: Two- Way Interactions of All Three Factors

Summary Table 1

Table 2.A: Summary of Fit

RSquare	0.847
RSquare Adj	0.703
Root Mean Square Error	0.002
Mean of Response	0.028
Observation (or Sum Wghts)	36

Table 2.B: Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	17	0.000497	0.000029	5.88
Error	18	0.0000894	$4.986 * 10^{-6}$	Prob > F
C.Total	35	0.000586		0.0003*

Table 2.C: Effect Tests

Source	**Nparm**	**DF**	**Sum of Squares**	**F Ratio**	**Prob > F**
Active Yeast	2	2	0.00006	6.05	0.0098*
Combination	2	2	0.000342	34.39	<.0001*
Time	1	1	0.0000012	0.236	0.6331
Time * Active Yeast	2	2	0.0000176	1.769	0.198
Time * Combination	2	2	0.0000327	3.289	0.061
Combination * Active Yeast	4	4	0.0000130	0.652	0.633

Combination * Time*Active Yeast	4	4	0.0000303	1.527	0.237

Analysis of Diagnostics Plots and Model's Factors 1

The diagnostic plots of the predicted transformed response and residuals in Figure 2.A and Sample Quantiles in Figure 2.B showed improvements from the original response diagnostic plots. Although they are not perfect and we need more observations to conclude that residuals deviate from normality, we will consider that the assumptions constant variance and Normal($0,\sigma^2$) distributed of errors are approximately satisfied. We will use the transformed response to fit the model. Beside the main factors, we were also interested in the interactions between factors because they provide additional insights into our response.

The figure 2.C shows all two-way interactions between the factors. From the plot, the interactions between Active Yeast*Time and Combination*Time are significant at $\alpha = 0.1$ and the interaction between Time*Combination is not significant. The JMP outputs in Table 2.A summarizes the fit of the model where it contains the Root Mean Square Error which is the estimate standard deviation of the Error terms and Sample Average of Transformed Response. We will used RSquare adjusted to analyze the model. The percentage of the response variation can be explained by the factors and their interactions are approximately 70% which is a satisfactory fit.

ANOVA test for the main factors was used to compare the sample means across three factors and the results can be seen in Table 2.B. The p-value is less than the significant level, therefore at least one of the true means of the main factors is different from the other. The Table 2.C shows the significant results in red. In this table, the true mean of the transformed response of the factors Active Yeast, Combination, and Time*Combination are significant difference from each other. We will refit the model with only the significant terms. In addition, because the Time factor was not significant, yet the interaction Time*Combination is significant, we will include the Time factor in the refitted model.

Refitting Model with Only Significant Factors

$\frac{1}{y_{ijkl}^2} = \mu + T_i + \beta_j + (\beta\gamma)_{jk} + + \varepsilon_{ijkl}$, where all the parameters have the same interpretations in 1.1.

Model Diagnostic 3

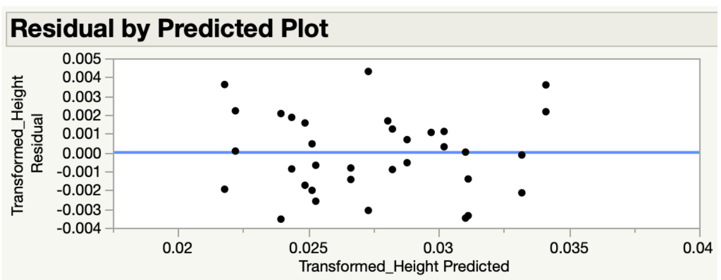

Figure 3.A: Transformed Response Residuals with Predicted Transformed Response

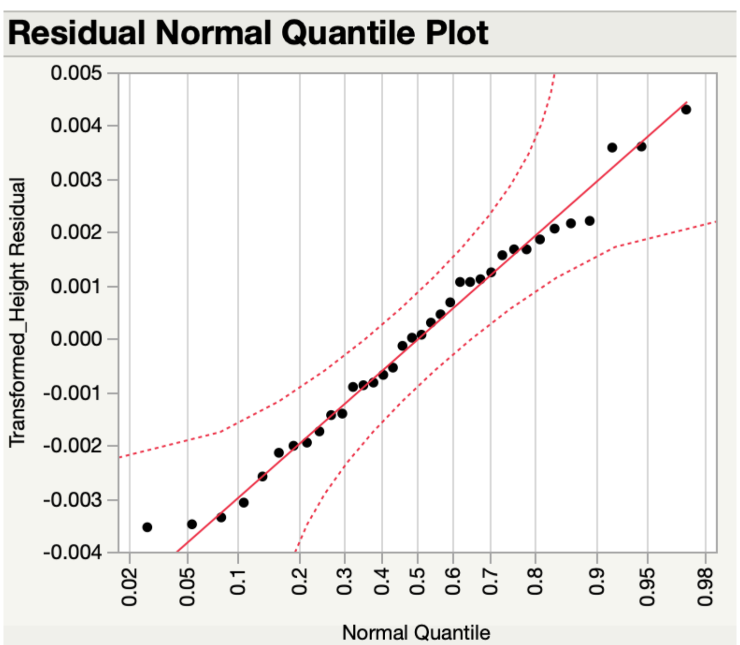

Figure 3.B Sample Quantile

Summary Tables 2

Table 3.A: Summary of Fit

RSquare	0.744
RSquare Adj	0.679
Root Mean Square Error	0.0023
Mean of Response	0.0275
Observation (or Sum Wghts)	36

Table 3.B: Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	7	0.000436	0.000062	11.596
Error	28	0.00015	$5.367 * 10^{-6}$	Prob > F
C.Total	35	0.000586		0.0001*

Table 3C: Effect Tests

Source	**Nparm**	**DF**	**Sum of Squares**	**F Ratio**	**Prob > F**
Active Yeast	2	2	0.0000601	5.60	0.009*
Combination	2	2	0.000342	31.84	<.0001*
Time	1	1	$1.2*10^{-6}$	0.218	0.6440
Time * Combination	2	2	$3.27*10^{-5}$	3.044	0.064

Table 3D: Parameter Estimates:

Term	**Estimate**	**Std Error**	**T Ratio**	**Prob > |t|**
Intercept	0.0276	0.000386	71.37	<.0001*
Active Yeast[1]	0.00134	0.000546	2.45	0.0210*
Active Yeast[2]	0.000412	0.000546	0.75	0.4568
Combination[1]	-0.0038	0.000546	-6.98	<.0001*
Combination[2]	$7.89*10^{-5}$	0.000546	0.14	0.8861
Time[3]	0.00018	0.000386	0.47	0.6440
Time[3]*Combination[1]	$-3.87*10^{-4}$	0.000546	-0.71	0.4843
Time[3]*Combination[2]	$9.24*10^{-4}$	0.000546	-1.69	0.1017

Analysis of Diagnostics Plots and Refitted Model's Factors 2

The model was refitted with only the significant terms. The diagnostic plots of Residual by Transformed Response in Figure 3.A and the Sample Quantiles in Figure 3.B suggest that the variance of the error is constant and distributed as N$(0,\sigma^2)$ therefore we can proceed with ANOVA testing and analyze the results.

Table 3.A shows the RSquare Adj of the model whose value has decreased about 2% from the previous model. The table also includes the estimate standard deviation of error whose value increase from the previous model's estimate. Both decreasing in RSquare Adj and increasing in sample standard deviation are the consequences of reducing the number of factors in the model.

The results from ANOVA in Table 3.B show an F value of 11.596 which is a statistically significant result that at least one of the true means is different from the other. The Sum of

Squares in the Effect Test in Table 3.C can be interpreted as given other factors already in the model, the Sum of Squares contributed by a factor is reported in its column along with the F ratios. The factors Active Yeast, Combination, and Time * Combination contribute significantly to the Sum of Square of the model. The results from Table 3.D list the parameter estimates and their contribution to the model. From the table, intercept, Active Yeast[1], and Combination[1] are significant parameters.

Conclusion

In the experiment we attempt to identify which factor Time, Combination of Active Yeast and Salt, and Amount of Active and their interactions contribute to the rise of bread. The experimental design was a factorial model with two factors of three levels and one factor of two levels. We transformed the response so that assumptions of the model are approximately satisfied. The model was refitted with the significant factors Active Yeast, Combination, and Time*Combination. From the final result tables, the amount of active yeast has a significant influence on the rise of a bread. This is not a surprising result because the reaction between yeast with other ingredients during the fermentation process helps the flour to release carbon dioxide which help the bread to rise. The height of bread is similar for the fermentation time = {3,5}. This is a surprising result because the gluten in the dough needs time to develop and gluten helps a bread to rise by trapping gas bubbles during the fermentation process. Combination of Active Yeast and Salt is also significant in contributing to the height of a bread. From the Parameter Estimates Table 3.D, not including salt in the mixture contributes to the highest rise of a bread and the difference between average height either with mixing or on the side is not significant.

JMP Experimental Design

Active Yeast	Combination	Time	Height
1	1	3	6.15
1	1	3	6.575
2	1	3	6.2
2	1	3	7
3	1	3	6.275
3	1	3	7.1
1	2	3	6.05
1	2	3	5.825
2	2	3	6.425
2	2	3	5.625
3	2	3	6.25
3	2	3	6.575
1	3	3	5.15
1	3	3	5.25
2	3	3	5.675
2	3	3	5.5
3	3	3	5.675
3	3	3	6.025
1	1	5	6.375
1	1	5	6.6375
2	1	5	6.175
2	1	5	6.525
3	1	5	6.4
3	1	5	6.7
1	2	5	5.7
1	2	5	5.7
2	2	5	5.95
2	2	5	5.825
3	2	5	6.3
3	2	5	6.225
1	3	5	6
1	3	5	5.8
2	3	5	5.65
2	3	5	5.725
3	3	5	5.725
3	3	5	5.8

References:

1. Math 261B Text Book: Design and Analysis of Experiments by Douglas C. Montgomery
2. *"Bread Dough and Baker's Yeast: An uplifting Synergy"* by Nore Struyf, Eva Van der Maelen, Sami Hemdane, Joran Verspreet, Kevin J. Verstrepen, Christophe M. Courtin.
3. The Science of Cooking: Baking 101 https://www.exploratorium.edu/cooking/bread/bread_science.html
4. https://thebreadguide.com/bread-dense-as-a-brick/