ESCUELA DE CIENCIAS MATEMÁTICAS Y TECNOLOGÍA INFORMÁTICA

Cheddar Cheese Taste

Richard Torres *

Andres Riofrío †

Lenin Rueda ‡

July 1, 2017

ABSTRACT

It is presented a study of different linear models and hypothesis taking into account the explanatory variables H2S, lactic and acetic. The data was collected by the cheddar cheese study from the La Trobe Valley of Victoria, Australia, cheese samples were analyzed for their chemical composition and were subjected to taste tests.

KEYWORDS: *Cheese Chaddar, H2S, lactic, acetic, laTrobe Valley of Victoria.*

*richard.torres@yachaytech.edu.ec
†andres.riofrio@yachaytech.edu.ec
‡lenin.rueda@yachaytech.edu.ec

I. INTRODUCTION

Cheese since its beginning has always been considered an important food for humans. Therefore, many years ago, approximately in the XVI century, cheese markets began to settle in the big cities. For that reason the industry has worried about improving the taste of cheeses, via great experimental advances.

This paper intends to make a study of the Cheddar cheese tastes when it is used three different chemical compounds, and how they affect the taste of the testers. Cheddar cheese is popular due to its taste. Spicy and very suitable for the preparation of casseroles and sandwiches. Cheddar cheese obtained its name from a village in the South-West of England. The production of cheddar cheese is distinguished by a special feature, called "cheddarization" (John,2010). This cheese is prepared as in the making of other cheeses. We go to the curd after draining the whey progressive acidification. The curds, or solids, are cut into large blocks, redistributed, and cut again into small and salty pieces. Cheddar has a long maturation period and it is a cheese without lactose. It has very little lactose in itself, in contrast to other cheeses.

The paper is organized with a brief overview of the problem, into the problem explanation, the methodology used and the comparison between the different linear models, calculated graphs and the various hypothesis. It is concluded with a discussion about the factor in the cheddar cheese tasting. Then, it is finished with conclusions, as well as the references biblio-hemerographic.

II. PROBLEM

The main premise of the research is that, the cheese taste in maturation is related to the concentration of several chemicals product. When using one of the chemical products, there is the possibility that the taste of cheese improves, is damaged, or is not affected at all. In the first case, the industry of cheeses that uses these products do not need to make any changes because it helps to increase their profits by making a flavor-enhanced product. In the second case the industry need to remove the product that affects the flavor and leave only the products that enhance the taste, in this way avoiding the loss of consumers who turn to another type of cheese.

On the basis of official figures emanating from the experiments made by G. T. Lloyd and E. H. Ramshaw del CSIRO Division of Food Research,Victoria, Australia, this paper presents the analysis of the chemical composition of the cheese samples. The chemical compounds to be analyzed are acetic acid, hydrogen sulfide and lactic acid. Overall taste scores were obtained by combining the scores from several tasters.

According to Ortega "There are many chemical additives to be added to foods not only to prevent decomposition by microbial growth, but also to maintain the physical-chemical conditions and prevent oxidation, to enhance the taste and this way improve the comfort of the consumer" (2013).

The concentrations of chemicals in the cheese samples, obtained from G. T. Lloyd and E. H. Ramshaw del CSIRO, it has an effect on the taste of the consume cheddar cheese, when using a product or another makes the taste improves or does not has any effect. The industry needs the study of the products used in the preparation of the cheese, for the reason people buy cheese depending its taste. Therefore the industry needs to know which products to enhance the taste of the cheese and thus increase their income. In particular, as stated by Ortega is needed to use chemicals to improve the taste and satisfy the consumer. For this reason, it is important to know that chemicals have effect on the taste of the cheese.

In particular, this research is proposed as objectives the following:

1. Analyze the data and find the relationship between the variables that are involved in the taste of the cheese.

2. Discuss the found results, which show that chemicals affect the taste of cheese, thus satisfying the needs of the consumer.

III. METHODOLOGY

In order to achieve the objectives, it is supposed a model, to compare the relations between the taste and the "H2S", "Lactic" and "Acetic" composition. To develop different linear models that explain which are the relation between the response variable (taste) and the explanatory variables, in order to find in what range the taste depends of its composition. From the results obtained with the used models, and tables that explains a general idea between the chemical composition with the flavor.

IV. DEVELOPMENT

4.1 Simple linear models analysis

It is presented an analysis of different simple linear models, taking in consideration the explanatory variables and response variable:

- Taste (y): Subjective taste test score, obtained by combining the scores of several tasters.
- Acetic (x_1): Natural log of concentration of acetic acid.
- H2S (x_2): Natural log of concentration of hydrogen sulfide.
- Lactic (x_3): Concentration of lactic acid.

Case	taste	Acetic	H2S	Lactic
1	12.3	4.543	3.135	0.86
2	20.9	5.159	5.043	1.53
3	39	5.366	5.438	1.57
4	47.9	5.759	7.496	1.81
5	5.6	4.663	3.807	0.99
6	25.9	5.697	7.601	1.09
7	37.3	5.892	8.726	1.29
8	21.9	6.078	7.966	1.78
9	18.1	4.898	3.85	1.29
10	21	5.242	4.174	1.58
11	34.9	5.74	6.142	1.68
12	57.2	6.446	7.908	1.9
13	0.7	4.477	2.996	1.06
14	25.9	5.236	4.942	1.3
15	54.9	6.151	6.752	1.52
16	40.9	6.365	9.588	1.74
17	15.9	4.787	3.912	1.16
18	6.4	5.412	4.7	1.49
19	18	5.247	6.174	1.63
20	38.9	5.438	9.064	1.99
21	14	4.564	4.949	1.15
22	15.2	5.298	5.22	1.33
23	32	5.455	9.242	1.44
24	56.7	5.855	10.199	2.01
25	16.8	5.366	3.664	1.31
26	11.6	6.043	3.219	1.46
27	26.5	6.458	6.962	1.72
28	0.7	5.328	3.912	1.25
29	13.4	5.802	6.685	1.08
30	5.5	6.176	4.787	1.25

Figure 1: Data of response variable (taste) and explanatory variables (Acetic, H2S and Lactic).

In the Figure 1 it is showed the different compounds data to the cheese elaboration. In this sense, it is postulated a "pure additive" model:

$$\boldsymbol{y = B_0 + B_1 * x_i + \epsilon, i = (1, 2, 3)}$$

4.1.1 Linear regression between response variable taste and explanatory variable acetic acid.

Now we start with the explanatory variable acetic, calculated with the statistical software R. The results were that the acetic compound is significant in the cheddar cheese because we reject the null hypothesis in Table 1. Also we see that the expected value found is that $\widehat{taste} = -61.499 + 15.648 Acetic$ taking into account 30 points. Figure 2 it is showed that the positive slope $\widehat{B_1}$ is the growth rate of the response variable taste, while the intercept $\widehat{B_0}$ is the value of the response variable when acetic is equal to zero. Hence, if it increased in one unit the acetic, allows that the taste increases in 15.648.

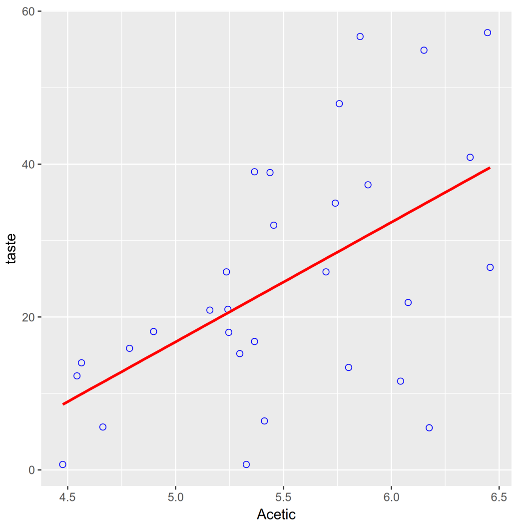

Figure 2: Linear regression between taste and acetic. There is a positive relation between the taste and the acetic acid.

Hypothesis	**p-value**	**Result**
H_0: $\beta_0 = 0$	0.01964	Reject the hypothesis H_0.
H_0: $\beta_1 = 0$	0.00166	Reject the hypothesis H_0.

Table 1: Test the individual nullity hypotheses of each β_j parameter taking only the explanatory variable acetic.

The R^2 value for Table 1 is 0.302 that is near to zero, for this reason is not a good model because the R^2 is near to 0 and not to 1.

4.1.2 Linear regression between response variable taste and explanatory variable hydrogen sulfide.

Then, the other chosen explanatory variable is $H2S$. The results were that the $H2S$ compound is significant in the cheddar cheese because we reject the null hypothesis in Table 2. Also we see that the expected value found is that $\widehat{taste} = -9.7868 + 5.7761 H2S$. Figure 3 informs that the positive slope $\widehat{B_1}$ is the growth rate of the response variable $\widehat{taste}$, while the intercept -9.7868 is the value of the response variable when H2S is equal to zero. Hence, if it increased in one unit the H2S, allows that the taste increases in 5.7761.

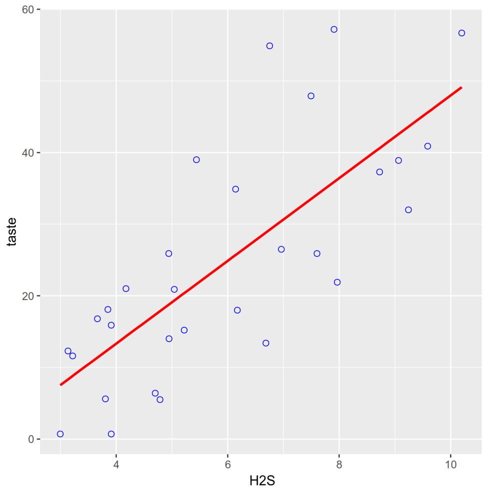

Figure 3: Linear regression between taste and H2S.

Hypothesis	p-value	Result
H_0: $\beta_0 = 0$	0.112	We can not reject the hypothesis H_0.
H_0: $\beta_1 = 0$	1.37e-06	Reject the hypothesis H_0.

Table 2: Test the individual hypotheses of nullity of each β_j parameter, taking only the H2S variable.

The R^2 value for this table is 0.5712, that is not near to 1, for this reason is not a good model.

4.1.3 Linear regression between response variable taste and explanatory variable lactic acid.

Then, the last explanatory variable evaluated is Lactic. The results were that Lactic compound is significant in the cheddar cheese because we reject the null hypothesis in Table 3. Also we see that the expected value found is that $\widehat{taste} = -29.859 + 37.720 Lactic$. Figure 4 shows that if it increased in one unit the lactic, allows that the taste increases in 37.720. In addition there is a positive relation between the taste and the lactic acid.

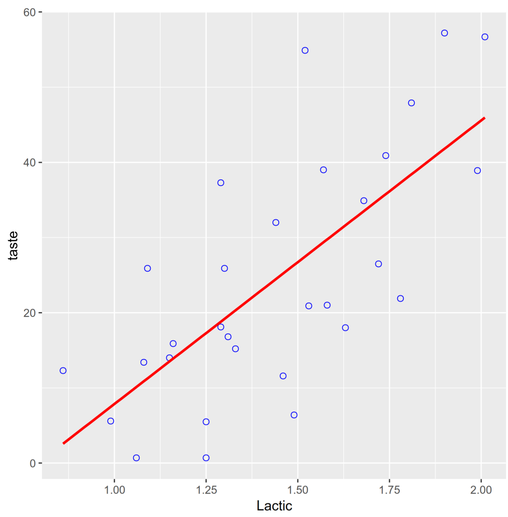

Figure 4: Linear regression between taste and lactic acid.

Hypothesis	p-value	Result
H_0: $\beta_0 = 0$	0.00869	Reject the hypothesis H_0.
H_0: $\beta_1 = 0$	1.41e-05	Reject the hypothesis H_0.

Table 3: Test the individual hypotheses of nullity of each β_j parameter, taking only the Lactic variable.

The R^2 value for this table is 0.4959, that is not near to 1, in this case also is not a good model.

4.2 Multiple linear models analysis

It is necessary to postulate another linear models, but in this case a multiple linear regression model. So, we build a "pure additive" linear model that is the following:

$$y = B_0 + B_1 * x_i + B_2 * x_j + \epsilon, i, j = (1, 2, 3) \; \forall i \neq j$$

4.2.1 Multiple linear regression between response variable taste and explanatory variables H2S and acetic.

We begin with the model multiple linear regression between response variable taste and explanatory variables H2S and acetic. The results were that the $H2S$ compound is significant in the cheddar cheese because we reject the null hypothesis in Table 4. However the acetic compound not is significant in the cheddar cheese because we cannot reject the null hypothesis in Table 4. On the other hand we see that the expected value found is that $\widehat{taste} = -26.940 + 5.146 H2S + 3.801 Acetic$.

Hypothesis	p-value	Result
H_0: $\beta_0 = 0$	0.214536	We can not reject the hypothesis H_0.
H_0: $\beta_1 = 0$	0.000225	Reject the hypothesis H_0.
H_0: $\beta_2 = 0$	0.406245	We can not reject the hypothesis H_0.

Table 4: Multiple linear regression between response variable taste and explanatory variable H2S and acetic.

The R^2 value for this table is 0.5822, that is not near to 1, for this reason the $B's$ are not good estimators.

4.2.2 Multiple linear regression between response variable taste and explanatory variables H2S and lactic.

Next, the other chosen explanatory variables in a multiple linear regression are H2S, and lactic. The results were that both compounds are significant in the cheddar cheese because we reject the null hypothesis in Table 5. In addition we see that the expected value found is that $\widehat{taste} = -27.592 + 3.95H2S + 19.89Lactic$

Hypothesis	p-value	Result
H_0: $\beta_0 = 0$	0.00481	Reject the hypothesis H_0.
H_0: $\beta_1 = 0$	0.00174	Reject the hypothesis H_0.
H_0: $\beta_2 = 0$	0.01885	Reject the hypothesis H_0.

Table 5: Multiple linear regression between response variable taste and explanatory variable H2S and acetic.

The R^2 value for this table is 0.6517, that is not near to 1, for this reason the $B's$ are not good estimators.

4.2.3 Multiple linear regression between response variable taste and explanatory variables acetic and lactic.

Then, the other chosen explanatory variables in a multiple linear regression are acetic, and lactic. The results were that the acetic compound not is significant in the cheddar cheese because we cannot reject the null hypothesis in Table 6. In addition we see that the expected value found is that $\widehat{taste} = -51.366 + 5.571Acetic + 31.392Lactic$

Hypothesis	p-value	Result
H_0: $\beta_0 = 0$	0.02223	Reject the hypothesis H_0.
H_0: $\beta_1 = 0$	0.25217	We can not reject the hypothesis H_0.
H_0: $\beta_2 = 0$	0.00161	Reject the hypothesis H_0.

Table 6: Multiple linear regression between response variable taste and explanatory variable acetic and lactic.

The R^2 value for this table is 0.5203, that is not near to 1, for this reason the

$B's$ are not good estimators.

4.2.4 Multiple linear regression between response variable taste and explanatory variable H2S, acetic and lactic.

Then, the other chosen explanatory variables in a multiple linear regression are H2S, acetic, and lactic. The results were that the H2S, and lactic compounds are significant in the cheddar cheese because we reject the null hypothesis in Table 7, but the acetic compound is not significant in the cheddar cheese because we cannot reject the null hypothesis. In addition we see that the expected value found is that $\widehat{taste} = -28.8768 + 3.91H2S + 0.33Acetic + 19.67Lactic$.

Hypothesis	p-value	Result
H_0: $\beta_0 = 0$	0.15540	We can not reject the hypothesis H_0.
H_0: $\beta_1 = 0$	0.00425	Reject the hypothesis H_0.
H_0: $\beta_2 = 0$	0.94198	We can not reject the hypothesis H_0.
H_0: $\beta_3 = 0$	0.03108	Reject the hypothesis H_0.

Table 7: Multiple linear regression between response variable taste and explanatory variable H2S, acetic and lactic.

The R^2 value for this table is 0.6518, that is not near to 1, for this reason the $B's$ are not good estimators.

4.3 Overall Regression

In order to test the relationship between the response variable and the explanatory variables. It is postulated the test of overall regression, taking the null hypothesis H_0: β_1, knowing our matrix C, the p-value found is 3.81e-06, indicating that the null hypothesis can be rejected with a confidence interval of 0.001. Meaning that the acetic, H2S and lactic are affecting the cheese taste.

$$C = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

However we want to know if in the cheddar cheese tasting, the composition of acetic has the same effect that H2S. So, in this case we postulate that H_0: $\beta_1 = \beta_2$. The p-value found is 0.4852, so there is not enough evidence to reject the null hypothesis. The result show us that the H2S and acetic have same effect in the cheddar cheese tasting. In the same way it is postulated another hypothesis in which we want to know if in the response variable the composition of acetic has the same effect that lactic, in this case the p-value result 0.09, meaning that the lactic and acetic have not a similar effect in the participants tasting.

4.4 Individual hypothesis of nullity

Now we wonder what happened in the individual hypotheses of nullity of each j parameter. The p-value results in the hypothesis are show in the table 8. The result expressed that in the acetic, there is not enough evidence to reject the nullity hypothesis in this context, so maybe the acetic has no influence in the cheese taste. Moreover, the lactic and H2S the null hypothesis is reject individually, so there are significant in our particular model.

Hypothesis	**p-value**	**Result**
H_0: $\beta_0 = 0$	0.1554	
H_0: $\beta_1 = 0$	0.942	No enough evidence to reject H_0.
H_0: $\beta_2 = 0$	0.004247	Reject H_0
H_0: $\beta_3 = 0$	0.03108	Reject H_0

Table 8: Test the individual hypotheses of nullity of each β_j parameter

On the other hand, we propose another hypothesis CB=0, where a null hypothesis is $2\beta_1$=β_2=β_3. The meaning is that we want to know is two times the 2 times the acetic has the same effect in the cheese tasting than the H2S, and that H2S has the same effect that the lactic respectively. The C matrix will be:

$$C = \begin{pmatrix} 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

The p-value is 0.2517, so we cannot reject the null hypothesis because we do not have enough evidence. Also, we think about if there exist a way in which there is a difference of 10 between the lactic and H2S, for this reason we apply the null hypothesis H_0: β_1-β_2=10, following the test for H_0: $C\beta = t$. Then, the obtained p-value was 0.01251, so we can reject the null hypothesis, meaning that there is not such a difference between the previous explanatory variables.

4.5 Best Model

In the Table 9 we made a summary of all the models proposed in this paper and we could observe that our best model is $\widehat{taste} = -27.6 + 3.9H2S + 19.9Lactic$ because has the smallest p-value 6.551e-07. For this reason, we go to work with this two variable model.

Models Summary	**p-value**
$\widehat{taste} = -61.5 + 15.6Acetic$	0.001658
$\widehat{taste} = -9.8 + 5.8H2S$	1.374e-06
$\widehat{taste} = -29.9 + 37.7Lactic$	1.405e-05
$\widehat{taste} = -26.9 + 5.1H2S + 3.8Acetic$	7.645e-06
$\widehat{taste} = -27.6 + 3.9H2S + 19.9Lactic$	6.551e-07
$\widehat{taste} = -51.4 + 5.6Acetic + 31.4Lactic$	4.936e-05
$\widehat{taste} = -28.9 + 3.9H2S + 0.33Acetic + 19.7Lactic$	3.81e-06

Table 9: Models summary obtained with R

Then, in order to understand the best model proposed is necessary analyze the ANOVA Table 10 (analysis of variances) that was calculated with a statistical processor program named R.

Parameters	**Estimate**	**Std. Error**	**t Value**	$Pr(>\|t\|)$
β_0	-27.592	8.982	-3.072	0.004812
β_1	3.946	1.136	3.475	0.00174
β_2	19.887	7.959	2.499	0.01885

Table 10: ANOVA table proposed model.

We have an appropriate model adjustment suggested to the data because we have a 95% of confidence level that permits reject the nullity hypothesis . Then the taste of the cheese is affected by the H2S and Acetic, therefore, our model can be formulated as $\widehat{taste}_i = -27.6 + 3.9H2S_i + 19.9Lactic_i + \epsilon_i, i = 1, 2, ..., 30$. Also is important to observe that we obtained other important values when we calculate the ANOVA table in R, as the residual standard error: 9.942 on 27 degrees of freedom and adjusted R-squared: 0.6259.

4.6 Model validation

Knowing our best model $\widehat{taste} = -27.6 + 3.9H2S + 19.9Lactic$ we can validate it. First, it is used the graph Q-Q plot, in the Figure 5, that tell us if our data is consistent with a normal probability model. This shows the curve normal for the standardized residues. In the normal conditions the points need to coincide with the line that represents the theoretical quartiles of a normal distribution. The points distributions show that could exist normal errors, and there exist suspicious points (8,12,15).

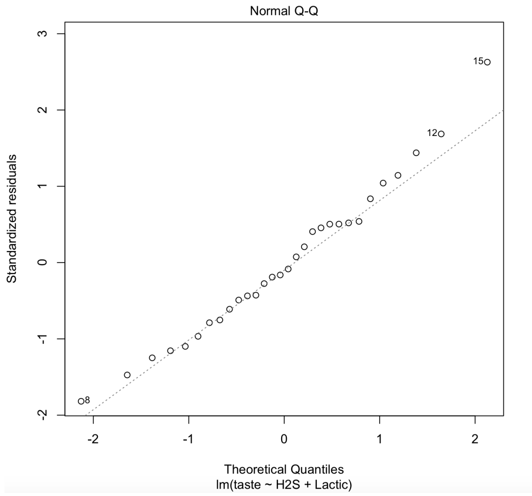

Figure 5: Normal probability plot of residuals.

Now, it is calculated the predicted value with the residuals. The Figure 6 verifies that all the points are disperse, they do not have a tendency, so is consistent with the theory, but still there are the suspicious points that can be outliers.

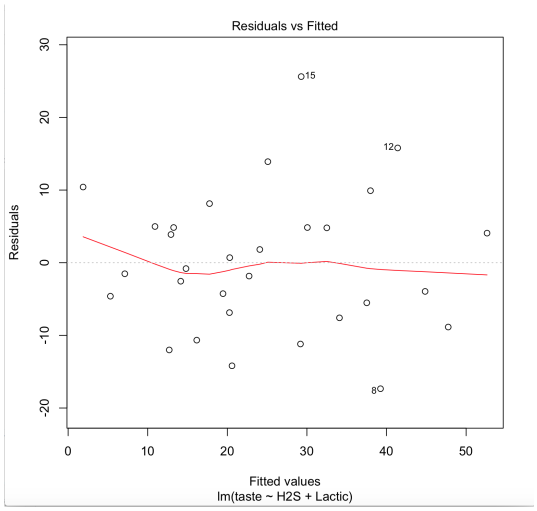

Figure 6: Residuals vs fitted values.

For this reason was necessary to plot the studentized and standardized residuals in order to confirm if that points are outliers. In this sense, the Figure 7 shows the original residuals (blue points), studentized residuals (yellow diamonds) and standardized residuals (triangles). It is demonstrated that the suspicious points are not outliers because the suspicious points are near to the zero cloud in this moment.

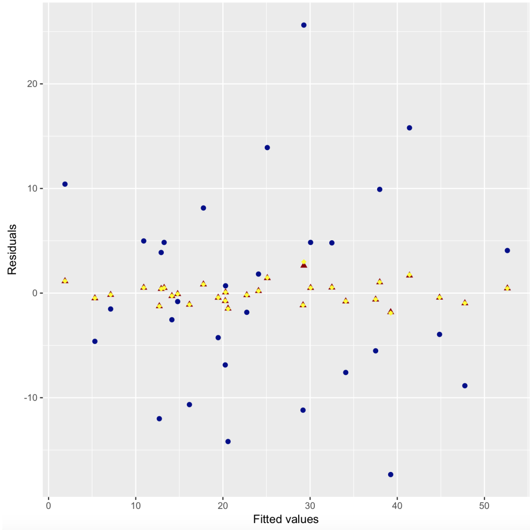

Figure 7: Original, Standardized and Studentized Residuals vs fitted values.

V. CONCLUSION

We conclude, that when was used only one explanatory variable, the square coefficient of correlation was near to 0, so we prove that they are not good models. In the test of overall regression was proved that the acetic, H2S and lactic are affecting the cheese taste because the p-value found is 3.81e-06, indicating that the null hypothesis can be rejected with a confidence interval of 0.001. In the same way, another hypothesis verify that for instance that the H2S and acetic have the same effect in the cheddar cheese tasting, with a p-value of 0.4852, so there is not enough evidence to reject the null hypothesis.

We finally confirm that our best model is $\widehat{taste} = -27.6 + 3.9H2S + 19.9Lactic$, with a p-value 6.551e-07. Our best model adopts the explanatory variables H2S and Lactic in the taste response variable. Also, we validate our model because it has no outliers, and no violation in the normal assumptions of our model.

References

[1] John, M. (2010). *Diversification of the cheddar cheese industry through specialty cheese production, Cornell University.*

[2] Moore, C. (2009). *Introduction to the practice of statistics, W. H. Freeman and Company.*

[3] Moore, C. (1990). *Characteristics of Cheddar Cheese Made with Sodium Chloride, Potassium Chloride, Or Mixtures of Sodium and Potassium Chloride, University of Wisconsin.*

[4] R Development Core Team. (2017). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0.Retrieved from http://www.R-project.org.*

APPENDIX: R CODE

```
library(car)
library(ggplot2)
require(ggplot2)
dat <- read.table("data.txt", header=T, sep=";")
dat

#Graphs taste~Acetic

ggplot(dat, aes(x = Acetic, y = taste)) +
geom_point(colour="blue", shape=21, size=2) +
geom_smooth(method='lm', colour="red",se=FALSE)

# Linear model with Acetic
mod1 <- lm(taste~Acetic, data=dat, x=TRUE)
mod1

tab <- summary(mod1)
tab

#Graphs taste~H2S

ggplot(dat, aes(x = H2S, y = taste)) +
geom_point(colour="blue", shape=21, size=2) +
geom_smooth(method='lm', colour="red",se=FALSE)

# Linear model with H2S
mod2 <- lm(taste~H2S, data=dat, x=TRUE)
mod2

tab <- summary(mod2)
tab

#Graphs taste~Lactic

ggplot(dat, aes(x = Lactic, y = taste)) +
geom_point(colour="blue", shape=21, size=2) +
geom_smooth(method='lm', colour="red",se=FALSE)

# Linear model with Lactic
mod3 <- lm(taste~Lactic, data=dat, x=TRUE)
mod3
tab <- summary(mod3)
tab
```

```
#Multiple linear regression between response variable
#taste and explanatory variable H2S and acetic.

mod4 <- lm(taste~H2S + Acetic, data=dat, x=TRUE)
mod4

tab <- summary(mod4)
tab

#Multiple linear regression between response variable
#taste and explanatory variable H2S, and lactic.
mod5 <- lm(taste~H2S +  Lactic, data=dat, x=TRUE)
mod5

tab <- summary(mod5)
tab
#Multiple linear regression between response variable
#taste and explanatory variables acetic and lactic.

mod6 <- lm(taste~Acetic +  Lactic, data=dat, x=TRUE)
mod6

tab <- summary(mod6)
tab

#Multiple linear regression between response variable
#taste and explanatory variable H2S, acetic and lactic.

mod7 <- lm(taste~H2S + Acetic + Lactic, data=dat, x=TRUE)
mod7

tab <- summary(mod7)
tab

# Test of Overall Regression
X <- mod7$x
X
n <- nrow(X)
n
k <- ncol(X)-1
k
C <- cbind(rep(0,k), diag(1,k))
C

linearHypothesis(mod7,C)
# Test the hypothesis B1=B2 (Acetic has the same
```

```
# composition that the H2S to the tasting)
C <- rbind(c(0,1,-1,0))
C

linearHypothesis(mod7,C)

# Test the hypothesis B1=B3 (Acetic has the same
# composition that the lactic to the tasting)
C <- rbind(c(0,1,0,-1))
C
linearHypothesis(mod7,C)

#Test the individual hypotheses of nullity of each j parameter.
# Individual B0
C <- c(1,0,0,0)
C
linearHypothesis(mod7,C)

# Individual B1
C <- c(0,1,0,0)
C
linearHypothesis(mod7,C)

# Individual B2
C <- c(0,0,1,0)
C
linearHypothesis(mod7,C)

# Individual B3
C <- c(0,0,0,1)
C
linearHypothesis(mod7,C)

#Test the following hypothesis: 2B1=B2=B3
C <- rbind(c(0,2,-1,0),c(0,0,1,-1))
C
linearHypothesis(mod7,C)
# Test B1-B2 = 10; CB = t
C <- rbind(c(0,1,-1,0))
C
linearHypothesis(mod7,C, rhs=10)

# Curve normal for the standardized residues.

res <- residuals(mod5)
res
```

```r
py <- fitted(mod5)
py

pDat <- as.data.frame(cbind(i=seq(1:nrow(dat)), e=res,
yf=py, x3=dat$x3, x4=dat$x4))
pDat

# Scatter plot e vs y
ggplot(pDat, aes(x = yf, y = e)) +
geom_point(colour="darkblue", shape=21, size=2) +
ylab("Residuals") +
xlab("Predicted values")

# Q-Q Plot
# Standarized residuals are:
res
# Now we calculate observed quantiles:
res1_y <- quantile(res, c(0.25, 0.75))
res1_y
# And then the theoretical quantiles
#assuming that the residuals comes
#from a standard normal distribution):
res1_x <- qnorm(c(0.25, 0.75))
res1_x
# Calculating the slope of the line
slope <- diff(res1_y)/diff(res1_x)
# Calculating the intercept of the line
inter <- res1_y[1] - slope * res1_x[1]
# Finally:
ggplot(as.data.frame(res), aes(sample=res))+
stat_qq() +
geom_abline(slope = slope, intercept = inter, color="blue")+
ylab("Residuals") +
xlab("nScores")

# Summary best model
summary(mod5)
plot(mod5)
res <- residuals(mod5)
res
sum(res)
plot(res)

py <- fitted(mod5)
py
```

```
pDat <- as.data.frame(cbind(i=seq(1:nrow(dat)), e=res,
yf=py, H2S=dat$H2S, Lactic=dat$Lactic))
pDat

res1 <- rstandard(mod5)
res1
res2 <- rstudent(mod5)
res2
pDat <- cbind(pDat, rStan=res1, rStud=res2)
pDat

# Fitted values vs residuals
ggplot(pDat) +
geom_point(aes(x = yf, y = e), colour="darkblue",
 shape=16, size=2) +
geom_point(aes(x = yf, y = rStan), colour="darkred",
shape=17, size=2) +
geom_point(aes(x = yf, y = rStud), colour="yellow",
shape=18, size=2) +
ylab("Residuals") +
xlab("Fitted values")
```