A study to investigate the best compound in preserving the freshness of a rose flower.

Project: Multivariate and Hierarchical Data (3565)
Project Discovering Associations
2018-2019

1st year Master of Statistics
Hasselt University

Group members:
Kubam Ivo (1850637)
Rondon Bautista Diego Alejandro (1849052)
Kamya Andrew (1849786)
Ferdous Nur A (1848853)
Onyedire Glory Chidima (1850831)

Submission Date: 05/June, 2019

Lecturers:
Prof. Dr. Geert Molenberghs
Prof. Dr. Christel Faes
Prof. Dr. Geert Verbeke
Prof. Dr. Luc Bijnens
Dr. Martin Otava
Dr. Steven Abrams
Dr. Yannick Vandendijck

Abstract

A rose is a universally celebrated flower. This is attributed to its delicate nature, beauty and aroma. It is one of the most marketable cut flowers in the floriculture industry (Elgimabi et al. 2011). However, a limited vase life greatly reduces its marketability. Distilled water is the most commonly used liquid in prolonging the vase life of cut flowers so as to maintain their freshness. Nevertheless, it preserves the freshness up to a limited degree since it lacks essential nutrients to boost its performance. The aim of this study was to determine the best compound in preserving the freshness of a rose by evaluating the efficacy of fifteen different compounds in comparison to distilled water. Three approaches were used to analyse two datsets (Gaussian and non-Gaussian) in which the responses were the number of days for which the rose stayed fresh (Poisson case) which was analyzed using the Poison regression with Generalized Linear Mixed Models (GLMM), then the probability of the flower being fresh (Binary response) which was analyzed using the Generalized Estimating Equations (GEE) and the width of the flower in centimeters (Gaussian response) using Linear Mixed Models. The results in all the three approaches showed three compounds as the best. These were Zest of Zen, Concentrate of Caduceus and Beerse Brew in that order for the non-Gaussian response and then in the order Concentrate of Caduceus, Zest of Zen and Beerse Brew for the Gaussian response. Using contrasts, we concluded Zest of Zen and Concentrate of Caduceus as the best compounds in preserving the freshness of a rose and that they have the similar effects.
Key Words: Vase life; Gaussian; Zest of zen; Concentrate of Caduceus; Beerse Brew

1 Introduction

Tulip flowers are spring-blooming perennial herbaceous bulbiferous geophytes from the family of Liliaceae, genus Tulipa (Christenhusz et al. 2013). As a result of their beauty, they are one of the most loved flowers in the world. Among these flowers, Black tulips are a rare hybrid of tulips and they symbolize power and strength (Scriber and Ording 2005).
Jean Baptiste, who was a botanist, needed a black tulip to offer to his beloved woman for marriage. But this flower was only available on an island in the Atlantic Ocean. He wanted to bring the black tulip from Atlantic, but on average it took 8 days and depending on the different weather conditions it could take up to 21 days. Without any compound, it was difficult to keep the unrooted black tulip fresh. That's why he requested the scientist to make an experiment with 15 compounds to see which compounds would keep the flower fresh for the longer period than water. Since the tulips are very expensive, he agreed to sacrifice rose flowers from his garden, so the experiment could be carried out with rose flowers instead of tulips

1.1 Objective

- The primary objective is to identify how different compounds affect the wilting process of roses and to investigate which compound keeps the flower fresh for a longer period than the voyage.
- The secondary objective is to ascertain how other covariates such as cutting time, rose type, rater affect the freshness of flower.

2 Methods and Materials

2.1 Experimental Design and Data Description

2.1.1 Experimental Design

Two different types of roses with the same growth that is, Floribunda sp. (Rose type 1) and Hybrid tea (Rose type 2) were randomly collected by four assistants from two gardens that is, the northern garden (Garden 1) and the southern garden (Garden 2) under specific process, specific time and with standard and clean tools. Then, they were randomly assigned to fifteen (15) different chemical compounds. These were Distilled water (Compound 1), Apathic Acid (Compound 2), Beerse Brew (Compound 3), Concentrate of Caduceus (Compound 4), Distillate of discovery (Compound 5), Essence of Epiphaneia (Compound 6), Four in December (Compound 7), Granules of Geheref (Compound 8), Kar Hamel Mooh (Compound 9), Lu- cifers Liquid (Compound 10), Noosperol (Compound 11), Oil of Johns son (Compound 12), Powder of Perlimpinpin (Compound 13), Spirit of Scienza (Compound 14) and Zest of Zen (Compound 15). In the laboratory, environmental conditions such as temperature, humidity were kept constant. The width of the roses were measured daily for 24 days. Through this period, the total number of days for which the roses remained fresh were recorded.

2.1.2 Data Description

Two separate dataset was used for the analysis:

- Count data: Had 3900 observations with 8 variables. No missing values recorded. The fresh variable was constructed by assigning the value 1 to if the flower was fresh on a given day and zero not fresh. The total number of 1's till when the flower ceased being fresh were then recorded as days for which the flower stayed fresh. BushID was used as a clustering variable.
- Gaussian Data: Made of 180 observations and 27 variables. Eight (8) flowers had missing measurements. Flowers with some missing measurements were kept because they were less than 5% of the sample size. This dataset was reconstructed and a width variable was created to hold diameter measurement of flowers for each day.

2.2 Sample Size Calculation

For this experiment, the hypotheses are as follows,

H_0: There is no difference for mean number of days between compounds and distilled water
H_1: At least one compound preserve the rose for the longer period than distilled water

A pilot study was conducted to determine the sample size for this study. The mean (10) and variance (5.8) was calculated from the pilot study.We took significance level=0.05/14 , where 14 is the number of pairwise comparisons since every compound is compared with water and the effect size $\delta = 1$ and choose the power $\beta = 0.80$. From this we calculated a sample size by simulation. For each compound we have 260 flowers, so a total sample size of $(15 * 260) = 3900$ flowers.

2.3 Statistcal Analysis

Three approaches were carried out in this study for poisson, binary and gaussian outcome. For poisson outcome, we used normal poisson model no clustering and poisson model for random effect taking into account clustering.Generalized Estimating Equations(GEE) with an Autoregressive working correlation for the Binary response was analyzed. For gaussian outcome, Linear Mixed Model(LMM) with random intercept and random slopes was fitted.

2.3.1 Exploratory analysis

Exploratory data analysis was done and summary were computed to have an overview of the data sets. From table 1, almost all the compounds had the most frequent survival time for the flowers as 24 days with few exceptions like compounds 2,8,10 and 12 with 13,7,6 and 13 days respectively. Compounds 15 and 4 both have the highest average number of flower survival days of 17 and 16 respectively. They were closely followed by compound 3 with 15 days. Compound 1(water)had an average survival days of approximately 13days.

Compound	**Mode**	**Mean**	**Std Dev**	**Compound**	**Mode**	**Mean**	**Std Dev**
compound 1	24	12.72	6.13	compound 9	24	13.39	6.08
compound 2	13	11.02	5.77	compound 10	6	11.37	6.07
compound 3	24	15.42	5.67	compound 11	24	13.85	5.84
compound 4	24	16.33	6.07	compound 12	13	12.87	5.93
compound 5	24	13.51	5.72	compound 13	24	14.79	5.95
compound 6	24	14.42	6.05	compound 14	24	13.05	5.81
compound 7	11	12.85	5.82	compound 15	24	16.89	6.14
compound 8	7	12.37	5.72				

Table 1: Summary Statistics By compound:Count Data

All compound had 260 flowers randomly chosen. Compounds 15 and 4 both had the highest average number of flower survival days of 16.89 and 16.33 respectively. Closely followed by compound 3 with 15 days. Compound 1 (water) had an average survival days of approximately 12.72 days (figure 1).

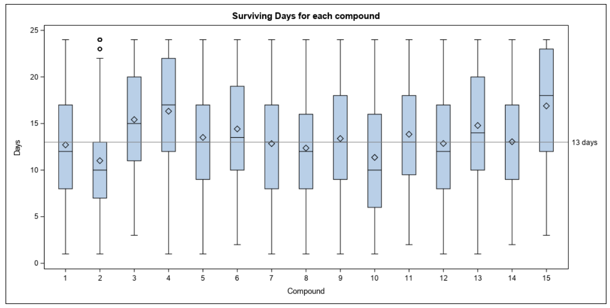

Figure 1: Surviving days of roses for Compounds

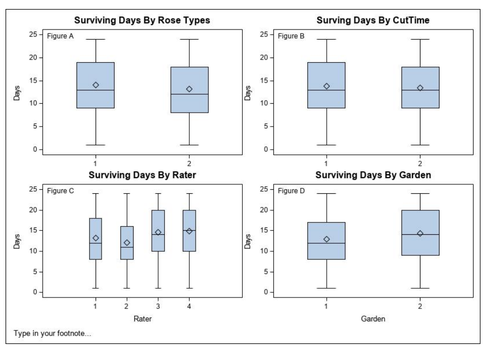

Figure 2: Surviving days of roses for covariates

From figure 2, the first sub-figure top right shows a slightly higher average number of surviving days for the flowers of rose type 1 than rose type 2. The second sub-figure top left also shows a slight difference in the average number of surviving days between the different cut times. Flowers from Rater 3 and 4 both show a higher average number of surviving days than those from Rater 1 and 2 (first sub-figure on the bottom). Garden 2 produced flowers with a higher average number of surviving days than garden 1 as shown in the second sub-figure on the bottom.

Figures 3 and 4 show the frequency of dead flowers (0) verses fresh flowers (1) amongst the different compounds and other covariates respectively. Generally, the figures show that there were more dead flowers than fresh flowers for all the covariates, compound 15 and 4 had more fresh flowers than the other compounds (figure 3). Also, there was really little difference in the number of fresh flowers for both cut times, both gardens and both rose types. Comparing fresh flowers for the raters, rater 3 and 4 had more fresh flowers than rater 2 and 1 (figure 4)

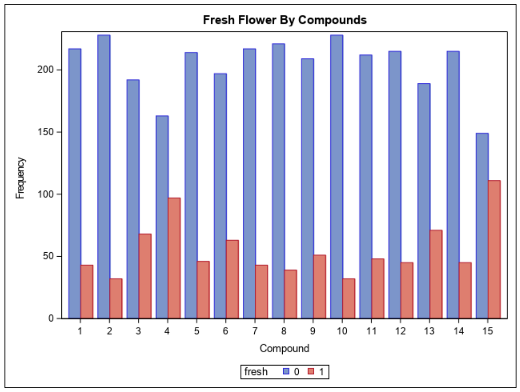

Figure 3: Dead vs Fresh roses for different compounds after 20 days

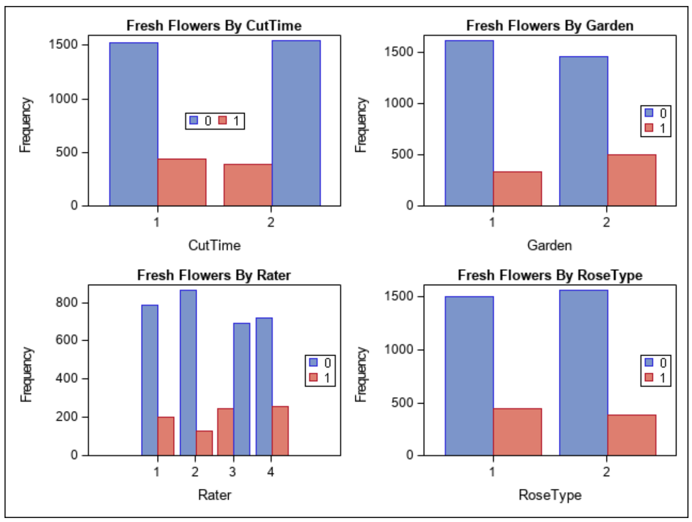

Figure 4: Dead vs Fresh roses for other covariates after 20 days

The figure 5 indicates that there was just slight difference in the starting widths of the roses (nested in subplots) implying different intercepts and differing rate of increase in the width over time (different evolution over time). A linear mixed model with a random intercept and random slope was thus used to analyze the data.

Compounds 4,3 and 15 all showed flowers with a small diameter range of 6.10, 6.9, 8.4 respectively.This implies small change of diameter over time. Compound 1 had a width range of 11.4 while compounds 8 and 12 both had wide range over time. As shown in the 2 below

Compound	Max	Min	Range	Compound	Max	Min	Range
Compound 1	14.60	3.20	11.40	Compound 9	15.40	3.40	12.00
Compound 2	17.90	3.90	14.00	Compound 10	16.10	3.50	12.60
Compound 3	11.60	3.20	8.40	Compound 11	12.80	3.70	9.10
Compound 4	9.60	3.50	6.10	Compound 12	16.40	3.50	12.90
Compound 5	14.50	3.20	11.30	Compound 13	12.00	4.00	8.00
Compound 6	12.70	3.80	8.90	Compound 14	14.10	3.70	10.40
Compound 7	15.00	3.60	11.40	Compound 15	10.20	3.30	6.90
Compound 8	16.60	3.60	13.00				

Table 2: Summary Statistics By compound:Gaussian Data

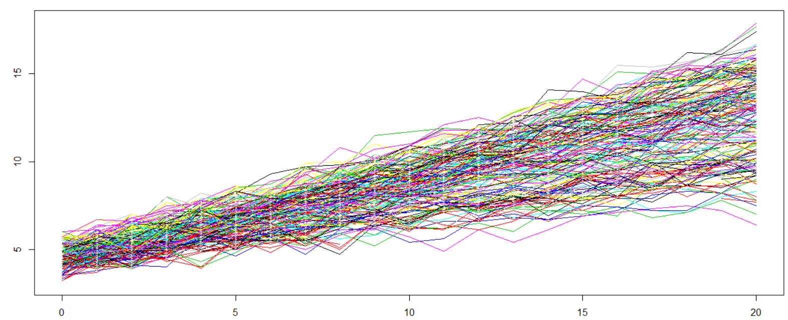

Figure 5: Individual profiles by subplot

2.3.2 Poisson Regression Model

Poisson models were build counting the mean number of days that the flower survived. The Poisson model can be described as (Agresti 2012):

$$log(\mu_i) = \alpha + \sum_j \beta_j x_{ij}, \qquad i = 1, \cdots, N$$

where

$$f(y; \mu) = \frac{e^{-\mu}\mu^y}{y!}, \qquad i = 0, 1, 2, \cdots$$

In this case models with and without cluster were take into account. The cluster approach was handeled using mix models creating specific model for the cluster

2.3.3 Binary Model

Binary data was analyzed using marginal models, in this case the estimates are solutions of Generalized Estimating Equations (GEE). With this model we are looking for a model that takes into account the interaction compound and time and as a result the probability of being fresh by using the compound c at time t(Agresti 2012). The model of interest is:

$$logit[P(Y_{ct} = 1)] = log\left[\frac{P(Y_{ct} = 1)}{1 - P(Y_{ct} = 1)}\right] = \alpha + \beta * CutTime + \sum_{c=1}^{15} Compound_c * t$$

This model assume that all flower did start at the same point, and each compound have a different slope.

2.3.4 Linear Mixed Model

Linear mixed models are an extension of simple linear models to allow both fixed and random effects(Agresti 2012). In this case, we are looking for a model whose response is the width of the flower and main effect of rose type and interaction between compound and time as covariates. The response (diameter of rose) can be modeled by a linear regression model, but with subject-specific regression coefficients. The model define by (Verbeke and Molenberghs 2009) has the following form,

$$Y_i = X_i\beta + Z_ib_i + \epsilon_i$$

Where, β is fixed effect (rose type and Compound*Time) and b_i is the random effect (random slope for each flower) and b_i $N(0, D)$, where D is the covariance matrix.

$$D = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}$$

d_{11} equals the variance of the intercepts b_{1i}
d_{22} equals the variance of the slopes b_{2i}
d_{12} equals the convariance between the intercepts b_{1i} and the slopes b_{2i}
The correlation between the intercepts and slopes is

$$corr(b_{1i}, b_{2i}) = \frac{d_{11}}{\sqrt{d_{11}}\sqrt{d_{22}}}$$

3 Result

3.1 Poisson Response

3.1.1 Model without clustering

The model for the main objective, average number of days as a function of the compounds. The values with * on the estimate show compound that are not significant.
In table 3,compound 1 (water) the value of $log(\mu) = 2.542$ which show that the effect of water is $\mu = exp2.542 = 12.705$ days. Negative values of the parameter estimates depict compounds that performed worse than water whereas positive values of the parameter

Coefficients	**Estimate**	**Coefficients**	**Estimate**
Intercept	2.542(0.017)	Compound9	0.051(0.024)
Compound2	-0.143(0.025)	Compound10	-0.112(0.025)
Compound3	0.19(30.023)	Compound11	0.085(0.024)
Compound4	0.250(0.023)	Compound12	0.011(0.024)*
Compound5	0.060(0.024)	Compound13	0.151(0.023)
Compound6	0.125(0.023)	Compound14	0.025(0.024)*
Compound7	0.010(0.024)*	Compound15	0.284(0.023)
Compound8	-0.027(0.024)*		

Table 3: Estimate(Std. Error) for Model without clustering

estimates reflect compounds that performed better than water. The compound that had the greatest effect on the response variable is 15 ($log(\mu) = 2.542 + 0.284 = 2.826 \implies \mu = exp2.542 = 16.887$), which means that the effect of compound 15 will improve the life of the rose from 12.7 to 16.877 mean number of days. The p-values are sign that some compounds no different effect from water. The null deviance of the model is 11263 with 3899 degrees of freedom, (11263/3899 = 2.88, which is sign of over dispersion). The AIC for the model is 27497. However this model does not take cluster into account.

3.1.2 Model with clustering

Taking clustering for the BushID into account lead to the parameter estimates in table 4. Between the two models the coefficient have not change that much but in this case is good to remember that we have to add the effect of the random part because it is not an aggregate model.

Coefficients	**Estimate**	**Coefficients**	**Estimate**
Intercept	2.523(0.064)	Compound9	0.051(0.024)
Compound2	-0.143(0.025)	Compound10	-0.112(0.025)
Compound3	0.193(0.023)	Compound11	0.085(0.024)
Compound4	0.250(0.023)	Compound12	0.011(0.024)*
Compound5	0.060(0.024)	Compound13	0.151(0.023)
Compound6	0.125(0.023)	Compound14	0.025(0.024)
Compound7	0.010(0.024)*	Compound15	0.284(0.023)
Compound8	-0.027(0.024)	**Random effects**	Std.Dev=0.1956

Table 4: Estimate(Std. Error) Poisson model with random effects

Several mix models were use to find the best approach, The selection probability was made with the lowest AIC value (Agresti 2012). Different summaries statistics can be seen on table 5. Model 2 present the best approach and the estimates coefficients (std. error) can be seen on table 4

Model	Predictor	Std. Dev	Deviance	Df	AIC	BIC
1	$(1\|Bu) + Ra + Cu + Co + Ga + Ro$	0,190	25022.1	3878	25066.1	25204.3
2	$(1\|Bu) + Ra + Cu + Co + Ro$	0,197	25022.8	3879	<u>25064.8</u>	<u>25196.4</u>
3	$(1\|Bu) + Cu + Co + Ro$	0,196	25416	3882	25452	25564.8
4	$(1\|Bu) + Co + Ro$	0,195	25427.2	3883	25461.2	25567.7
5	$(1\|Bu) + Cu + Co$	0,196	25474.4	3883	25508.4	25615
6	$(1\|Bu) + Co$	0,195	25484.6	3884	25516.6	25616.9
7	$(1\|Bu)$	0,195	26221.7	3898	26225.7	26128.2

Rater(Ra), Compound(Co), Cutting Time(Cu), Garden(Ga), Rose type(Ro), random intercept for bush$(1|Bu)$

Table 5: Result of Fitting Several Poisson Models

Fixed effects	Estimate	Fixed effects	Estimate
Intercept	2.527(0.065)	Compound7	0.014(0.024)*
Rater2	-0.087(0.012)	Compound8	-0.015(0.024)*
Rater3	0.115(0.012)	Compound9	0.042(0.024)
Rater4	0.122(0.012)	Compound10	-0.118(0.025)
CutTime2	-0.0283(0.008)	Compound11	0.091(0.024)
Compound2	-0.145(0.025)	Compound12	0.012(0.024)*
Compound3	0.191(0.023)	Compound13	0.154(0.023)
Compound4	0.248(0.023)	Compound14	0.025(0.024)*
Compound5	0.061(0.024)	Compound15	0.280(0.023)
Compound6	0.130(0.023)	RoseType2	-0.061(0.008)

Table 6: Paramet(Std. Error) for model with the lowest AIC value

The intercept of this model can be interpreted as the average effect of Rater1, at a cutting time between 10:00-14:00 using water with a rose type Floribunda which is $exp(2.52) = 12.42\,days$. From the model we can see that positive coefficients will estimate better results.

Label	Den DF	F Value	p-value
C3-C4	9	6.736	0.029
C3-C15	9	16.701	0.003
C4-C15	9	2.232	0.169

Table 7: Contrast statement Poisson model

Table 7 show all the comparison between compounds 3, 4 and 15, low p values will reject the null hypothesis that both contrast are equal. For this case we can see that the compound

3 and 4 are different and also 3-15.
Taking the information of tables 6 and 7 we are able to conclude that compound 15 will have a better performance over the others. Calculating the highest average possible number of days is given by: Rater 4 and compound 15 2.527+0.122+0.280 = 2.929, that is $exp(2.89) = 18.708\, days$ on average for the flower to live. Based on these results we recommend cutting time 10:00-14:00 and compound 15 to improve the rose live.

3.2 Binary Response

Generalized Estimated Equation was used to study binary outcome in a marginal way. The correlation matrix was define as autoregresive because it treats observations close in time as more correlated (Agresti 2012).

Parameter	**Estimate**	**95%LCI**	**95%HCI**
Intercept	3.638(0.055)	3.530	3.746
CutTime 1	0.087(0.055)*	-0.022	0.196
C15*time	-0.205(0.006)	-0.216	-0.193
C4*time	-0.215(0.006)	-0.227	-0.203
C3*time	-0.233(0.007)	-0.247	-0.219
C13*time	-0.242(0.007)	-0.256	-0.227
C6*time	-0.247(0.007)	-0.262	-0.232
C11*time	-0.260(0.007)	-0.275	-0.244
C9*time	-0.265(0.008)	-0.281	-0.249
C5*time	-0.269(0.008)	-0.286	-0.253
C14*time	-0.275(0.008)	-0.292	-0.259
C1*time	-0.276(0.008)	-0.293	-0.259
C7*time	-0.279(0.008)	-0.295	-0.263
C8*time	-0.293(0.009)	-0.310	-0.275
C10*time	-0.308(0.010)	-0.328	-0.289
C2*time	-0.332(0.011)	-0.354	-0.310

Table 8: Estimates (std Error) and confidence interval for Binary response

The interpretation of the GEE estimates is based on $logit[P(Y_t = 1)]$, the common intercept indicate that all the flowers start with the same odd of being fresh. The interpretation of the exponent of the slope of compound i (exp^{β_c}) is how much the odd of being fresh decrease for an increase in one day
Each probability can be plot as a function of time as show in figure 6. Compound 1 is represented with the thick black line. We can conclude that all the lines below water will be worse than water. Parameters with an estimate close to zero have better response (lines above).
The contrast statements were used to make pairwise comparison between the three compounds with the highest estimates, in this case are 3, 4 and 15. For each comparison we are

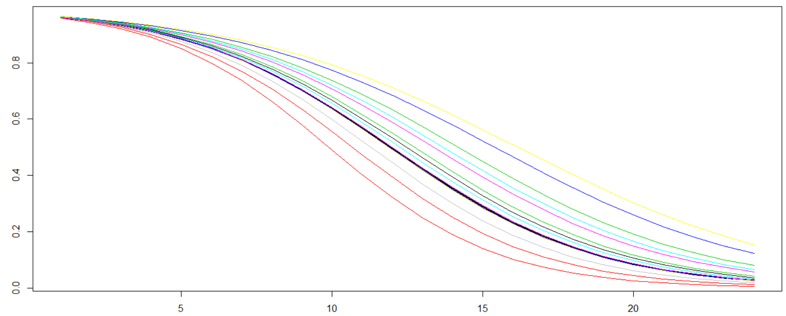

Figure 6: Probability of been fresh in function of time for each compound

testing if the compounds are equal or not. Table 9 show the p values for each pair, small p-values allow to reject the hypothesis that the compounds are equal, therefore we can only say that compound 4 and 15 are equal. With this information we select compound 15 as the one that offer the best condition for the travel.

Constrast	**DF**	**Chi-Square**	**p-value**
Compound 3-4	1	27.53	<.0001
Compound 3-15	1	33.07	<.0001
Compoud 4-15	1	0.30	0.5863

Table 9: Contrast Results for GEE Analysis

At day 21 the maximum probability that we can ensure is calculated by the 95% lowest confidence interval of component 15 and CutTime1 $3.530 - 0.022 + (-0.216 * 21) = -1.028$ taking the expit function will give us the probability of success $\frac{e^{-1.028}}{1+e^{-1.028}} = 0.263$

3.3 Gaussian Data

For the Gaussian response (width of a rose), a linear mixed model with a random intercept and random slope for rose nested in subplot was fitted. The model takes into account different widths at day 1 and the variability between slopes.

Table 11 is the test for fixed effects of the model, with this table we are testing if all the parameters related to a covariate are equal to zero, in this case, it shows that the interaction between compound and time was significant. This means that there was an effect of compounds on the width of the flowers and that the rate of change in this width over time was not the same for all the compounds.

The petals of a freshly cut rose are always closed and they start to open when it starts tilting. This means that the rate of opening is directly proportional to the rate of losing freshness. Since the interaction of compound 4 with time gives the least estimate, followed by compound 15 and then compound 3, the preference for the compound that preserves

the freshness of a flower is that order. The contrast estimates show that compound 4 and 15 have the same effect and are the best in preserving the freshness of a rose followed by compound 3.
The correlation between the random intercept and the random slope for the rose nested in the subplots (-0.0013) indicates that a larger rose in width at time 0 increases at a lower rate than a rose with a smaller initial width.

$$covariance\ matrix = \begin{bmatrix} 0.1967 & -0.00130 \\ -0.0013 & 0.000618 \end{bmatrix}$$

Parameter	**Estimate**	**Parameter**	**Estimate**
Intercept	4.587(0.049)	C8*time	0.530(0.008)
Type2	0.067(0.069)	C9*time	0.437(0.008)
C2*time	0.572(0.008)	C10*time	0.545(0.008)
C3*time	0.286(0.008)	C11*time	0.327(0.008)
C4*time	0.193(0.008)	C12*time	0.512(0.008)
C5*time	0.419(0.008)	C13*time	0.298(0.008)
C6*time	0.299(0.008)	C14*time	0.432(0.008)
C7*time	0.465(0.008)	C15*time	0.239(0.008)
C1*time	0.429(0.008)		

Table 10: Estimates (std. Erro) for gaussian data

Effect	**Den DF**	**F Value**	**p-value**
Type	3340	0.93	0.3343
Compound*time	3340	2328.35	<.0001

Table 11: Type 3 Tests of Fixed Effects

Label	**Num DF**	**Den DF**	**F Value**	**p-value**
Compound 3-4	1	3340	361.89	<.0001
Compound 3-15	1	3340	399.46	<.0001
Compound 4-15	1	3340	0.81	<0.3676

Table 12: Contras statement gaussian data

The random effect for the intercept and slope can be plot to check the behavior of independent measures, outlier values and correlation between random slope and intercept. Figure 7 have in the x-axis the slope and y-axis the intercept. Observation close to the vertical red line have had start the experiment with an average value, observation close to the horizontal line had a grow with in the average, this plot does not show outlier or any systematic patron that inform us about (Verbeke and Molenberghs 2009).

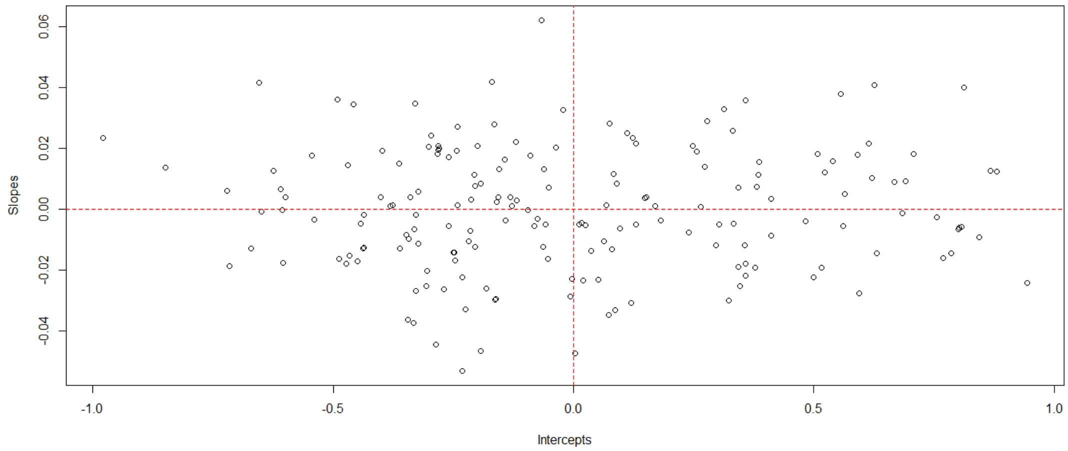

Figure 7: Random Intercept vs Slope

4 Discussion

This studied took into consideration three approaches in other to find the best compound to preserve two types of roses. These approaches were based on the outcome of interest. Considering the mean number of surviving days as the dependent variable (poisson model with clustering), the top three compounds are Beerse Brew (compound 3), Concentrate of Caduceus (compound 4) and Zest of Zen (compound 15); these were identified as the best based on their small parameter estimates, after performing pairwise comparison, Zest of Zen (Compound 15) was regarded as the best compound to keep the rose flowers for the longest mean surviving days. Also cutting the flowers between 10h- 14h will positively affect the mean surviving days. Modelling the probability of a flower to remain fresh for each day (GEE approach), Beerse Brew, Concentrate of Caduceus and Zest of Zen were found as the best three compounds based on smallest absolut parameter estimates. Pairwise comparison led to conclusion to accept Zest of Zen and Concentrate of Caduceus and not significant difference was found between the two. Eventhough cut time was not significant, cutting flowers between 10h-14h will reduce the probability of flowers been fresh. The third approach modeled the width of flowers using linear mixed model, same three compounds were found as the best as in the other two approaches and same conclusion was drawn as in the binary approach. Concentrate of Caduceus and Zest of Zen were found to be the best.

References

Agresti, A. (2012). *Categorical Data Analysis*. Wiley Series in Probability and Statistics. Wiley. ISBN: 9780470463635. URL: `https://books.google.be/books?id=UOrr47-2oisC`.

Christenhusz, Maarten JM et al. (2013). "Tiptoe through the tulips–cultural history, molecular phylogenetics and classification of Tulipa (Liliaceae)". In: *Botanical journal of the Linnean Society* 172.3, pp. 280–328.

Elgimabi, MENE et al. (2011). "Vase life extension of Rose cut flowers (Rosa Hybrida) as influenced by silver nitrate and sucrose pulsing". In: *American Journal of Agricultural and Biological Sciences* 6.1, pp. 128–133.

Scriber, J Mark and Gabriel J Ording (2005). "Ecological speciation without host plant specialization; possible origins of a recently described cryptic Papilio species". In: *Entomologia experimentalis et applicata* 115.1, pp. 247–263.

Verbeke, G. and G. Molenberghs (2009). *Linear Mixed Models for Longitudinal Data*. Springer Series in Statistics. Springer New York. ISBN: 9781441902993. URL: `https://books.google.be/books?id=jmPkX4VU7h0C`.

Appendix - R/SAS code

Simulation

```
# 1. Choose one fixed N
N <-500

# 2. Select parameters
control = 10   # e.g. pilot experiment
treated = 11   # i.e. increase of control mean by 11 or relative increase of 110% (= 11/10)
alpha = 0.05/14    # significance level

# 3. Simulate huge number of experiments and test
numberSimulation <- 1000
pval <- numeric(numberSimulation)

set.seed(1234) # set the seed for reproducible resuts
for (i in 1:numberSimulation){
  # we simulate from Poisson distribution
  controlGroup <- rpois(N, lambda = control)
  treatedGroup <- rpois(N, lambda = treated)
  simData <- data.frame(response = c(controlGroup, treatedGroup),
        treatment = rep(c(0,1), each = N))
  # we use a GLM model for Poisson regression to test effect of treatment
  pval[i] <- summary(glm(response ~ treatment, data = simData,
        family=poisson()))$coeff["treatment", "Pr(>|z|)"]
}
# 4. Estimate power
sum(pval < alpha)/numberSimulation
```

Poisson model

```
model1<-glm(Days~Compound, data=data, family = poisson()) #model with out cluster
#model with cluster
model2<-glmer(Days~Compound+(1|BushID), data=data,
family=poisson(link = "log"))

model2.1<-glmer(Days~Compound+Rater+CutTime+Garden+RoseType+(1|BushID),
            data=data, family=poisson(link = "log"))
model2.2<-glmer(Days~Rater+CutTime+Compound+RoseType+(1|BushID), data=data,
```

```
            family=poisson(link = "log"))
model2.3<-glmer(Days~CutTime+Compound+RoseType+(1|BushID), data=data,
            family=poisson(link = "log"))
model2.4<-glmer(Days~Compound+RoseType+(1|BushID), data=data, family=poisson(link = "log"))
model2.5<-glmer(Days~CutTime+Compound+(1|BushID), data=data, family=poisson(link = "log"))
model2.5<-glmer(Days~Compound+(1|BushID), data=data, family=poisson(link = "log"))
model2.6<-glmer(Days~(1|BushID), data=data, family=poisson(link = "log"))
```

GEE

```
proc genmod data=flower2 desc;
    class flowerid bushid compound CutTime;
model response= CutTime compound*time/ dist=b;
    repeated subject=flowerid(bushid)/modelse type=ar;
    contrast '3-4' compound*time 0 0 1 -1 0 0 0 0 0 0 0 0 0 0 0;
    contrast '3-15' compound*time 0 0 1 0 0 0 0 0 0 0 0 0 0 0 -1;
    contrast '4-15' compound*time 0 0 0 1 0 0 0 0 0 0 0 0 0 0 -1;
    estimate '3' compound*time 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0/ alpha=0.0167;
    estimate '4' compound*time 0 0 0 1 0 0 0 0 0 0 0 0 0 0/ alpha=0.0167;
    estimate '15' compound*time 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1/ alpha=0.0167;
run;
```

Gaussian Data

```
proc mixed data = flower_rev;
    Title 'Linear Mixed Model considering Random Intercept';
    class compound (ref=first) garden(ref=first) type(ref=first) subplot index;
    model width = type time2*compound/solution;
    random  intercept time2/ type = un subject =index(subplot) g gcorr solution ;
    ods listing excludesolutionr;
    ods output solutionr=out;
    contrast 'compound3 vs compound 4' time2*compound 0 0 1 -1 0 0 0 0 0 0 0 0 0 0 0;
    contrast 'compound3 vs compound 15' time2*compound 0 0 1 0 0 0 0 0 0 0 0 0 0 0 -1;
    contrast 'compound4 vs compound 15' time2*compound 0 0 0 1 0 0 0 0 0 0 0 0 0 0 -1;
run;
```