

May 20, 2024

The Effects of Sunlight and Water on Geranium Growth

Report prepared by:

Addison Andrix, Ruth Frahm, Minh Tran

Department of Mathematics, Computer Science, and Statistics

Gustavus Adolphus College

For: MCS - 243, Course Project

Executive Summary

Levels of sunlight and water may affect the growth of geraniums. A developer of a new variety of geraniums is interested in the impact of these variables on growth. To investigate these effects, 16 identical pots were selected, each planted with the same number of seeds. Random assignment allocated each pot to one of two sunlight levels (high or low) and one of two water levels (high or low). All pots were placed in the same greenhouse environment. After 12 weeks, plant height (in cm) was measured.

Results indicated that there is an interaction between sunlight and water to affect geranium growth. The pots that had the most growth were the pots with low levels of both sunlight and water. The pots that had the least growth were the pots with low levels of sunlight and high levels of water. Pots with high levels of sunlight and water resulted in growth that fell between pots that had low levels of sunlight and water and pots that had low levels of sunlight and high levels of water. There was a lack of evidence to prove a significant difference for pots that had high levels of sunlight and low levels of water.

Study Design and Data Description

The purpose of this study was to understand the effect of sunlight and water on plant height. Sixteen identical pots were obtained and filled with soil. Each pot then had the same number of seeds planted. The pots were brought into one greenhouse and each pot was randomly assigned one of two levels of sunlight: high or low, as well as one of two levels of water: high or low. This results in four unique combinations of treatments. This study uses a balanced design as all treatment groups have 4 pots which leads to a total of 16 pots. After 12 weeks, the plants' heights were measured in centimeters.

Figure 1 displays an interaction plot for sunlight and water with the mean yield on the y-axis. Since the two lines cross each other and have almost opposite slopes there appears to be an interaction between sunlight and water. Since an interaction appears to be present, we will analyze the combined effect. Table 1 displays summary statistics for the combined factors (treatments).

The results indicate that the medians are similar for high sun (18.5 cm). The spread for high sun and low sun with low water levels is similar (High.High IQR = 5 cm, High.Low IQR = 7.75 cm, and Low.Low IQR = 8.25 cm). Low sun and low water had the smallest IQR and the lowest median (IQR = 1.5 cm, med = 6.5 cm). Low sun and low water had an outlier at 40 cm (max height in the data set). There were no missing values in the dataset.

Figure 1: Interaction plot for sunlight and water on geranium growth

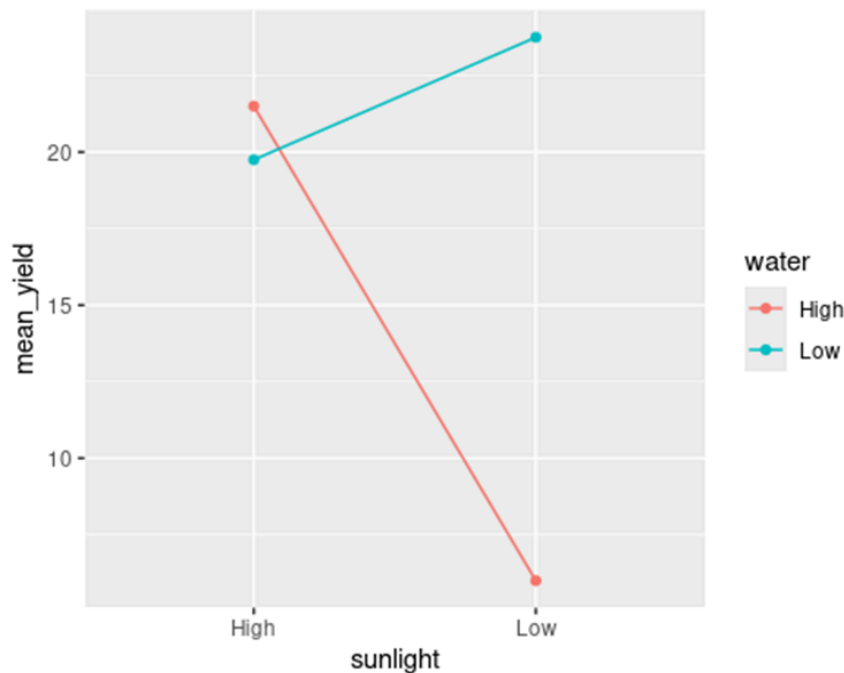


Table 1: Summary Statistics of Sunlight by Water

sunlight.water	min	Q1	median	Q3	max	mean	sd	n
High.High	16	17.50	18.5	22.5	33	21.5	7.77	4
Low.High	4	5.50	6.5	7.0	7	6.0	1.41	4
High.Low	14	14.75	17.5	22.5	30	19.75	7.32	4
Low.Low	13	18.25	21.0	26.5	40	23.75	11.50	4

Data Analysis and Results

The study analyzes the means of two levels of categorical variables (sunlight and water). For this reason, a two-way ANOVA was used to analyze the data. The assumptions were not originally met, so a transformation of the response was required. After transformation, the assumptions were met as can be found in the appendix. The model is:

$$\ln(\text{Yield}_{ijk}) = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ij}$$

where μ is the natural log average plant height (cm) of plants, α_i is the effect of sunlight where i is the level of sunlight, β_j is the effect of water where j is the level of water. $(\alpha\beta)_{ij}$ is the interaction between sunlight level i and water level j . ε_{ij} is the random error that is assumed to be normally distributed with a mean of 0 and a variance of σ^2 . k represents the k^{th} pot with combination i and j . A two-way ANOVA was used to analyze the data. The results of this ANOVA are shown in Table 2. All further results are in terms of the natural log.

Table 2: ANOVA Results

Variable	F-statistic	df	p-value
sunlight	9.694	1 and 12	0.0089
water	11.873	1 and 12	0.0048
sunlight:water	15.608	1 and 12	0.0019

There is strong evidence that there is an interaction between sunlight and water present ($F = 15.608$, $df = 1$ and 12 , $p = 0.0019$). Because of the interaction, we cannot analyze sunlight and water separately. The parameter estimates for these treatments can be found in Table 3.

Table 3: Parameter Estimates

Term	Estimate	SE	95% CI
High.High (Intercept)	3.026	0.178	2.638, 3.414
Low.High	-1.259	0.252	-1.807, -0.710
High.Low	-0.090	0.252	-0.639, 0.459
Low.Low	1.408	0.356	0.631, 2.184

Note: These parameters were obtained after taking the natural log of Yield (centimeters). Estimates and CIs of Low.High, High.Low and Low.Low represent the differences from High.High and are in terms of the natural log of centimeters.

With no specific comparisons being requested by the client, conditional pairwise comparisons were then calculated. These results are found in Table 4.

Table 4: Pairwise Point Estimates and Confidence Intervals

	Estimate	SE	95% CI
For High sunlight on water			
High - Low	0.09	0.252	-0.459, 0.639
For Low sunlight on water			
High - Low	-1.32	0.252	-1.867, -0.769
For High water on sunlight			
High - Low	1.259	0.252	0.710, 1.810
For Low water on sunlight			
High - Low	-0.149	0.252	-0.698, 0.400

Note: A Bonferroni adjustment was used. Results were obtained after a natural log transformation was conducted on yield.

These results indicate a significant difference between Low.High and Low.Low, and between High.High and High.Low. For low sunlight, low levels of water result in a taller yield (cm). For high water, high levels of sunlight result in a taller yield (cm).

Concluding Remarks

Differences exist among the geranium height based on the sunlight and water levels they received. The combination of sunlight and water that resulted in the least growth was low sunlight with high water. There was a medium yield for high sunlight and high water levels as it fell between the worst-growing and best-growing geraniums. The sunlight and water combination resulting in the most growth was low levels of both sunlight and water. Geraniums that grew under high sunlight and low water levels lacked evidence of being different from both high sunlight and high water as well as low sunlight and low water. Lacking strong evidence, we are unable to conclude the precise ordering of Geraniums grown under high sunlight and low water between those that grew with low sunlight, low water and high sunlight, high water.

There are other variables that can affect the growth of geraniums such as nutrients in the soil, temperature, fertilizer, etc. We assume that in this study all of these other variables are controlled from pot to pot and as such the only differences are the amount of sunlight and water received. Any additional experiment done should expand the levels of sunlight and water used. Additional experiments could also look at different soil types and how water levels are affected by that or if natural or artificial lights result in different growths.

Appendix

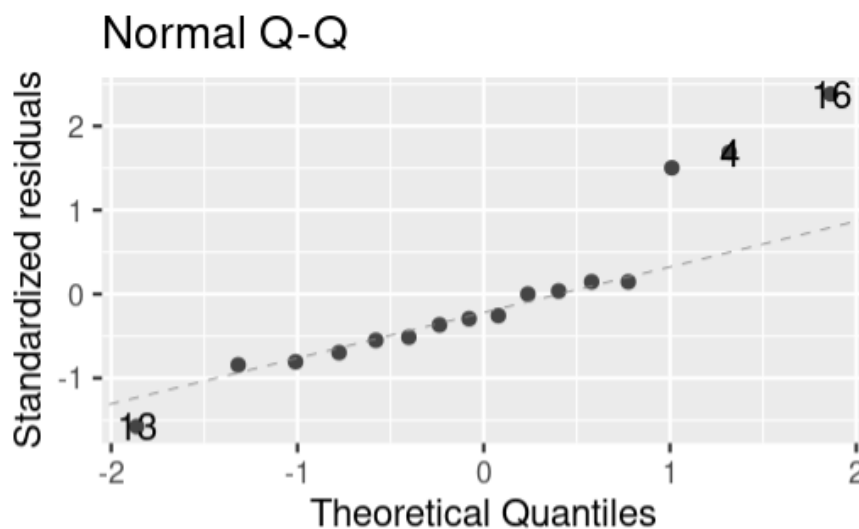
Three assumptions, conducted on the original dataset must be met for the analysis to be valid. Those are:

- Independence: Independence is met because our ou=eu=geraniums plants, there are no obvious clusters like where the seeds came from, and there were no repeated measures since the plants were measured once after 12 weeks.
- Equal Variance: Equal variance is violated because the largest and lowest standard deviations had a ratio of around 8 ($11.5/1.41 = 8.156 > 2$) so a natural log transformation was performed to normalize the yield.
- Normality: Normality is not met because both tails in Figure A1 stray from the reference line and we have a small sample size.

Table A1: Summary Statistics of Sunlight by Water for original dataset

sunlight.water	min	Q1	median	Q3	max	mean	sd	n
High.High	16	17.50	18.5	22.5	33	21.5	7.77	4
Low.High	4	5.50	6.5	7.0	7	6.0	1.41	4
High.Low	14	14.75	17.5	22.5	30	19.75	7.32	4
Low.Low	13	18.25	21.0	26.5	40	23.75	11.50	4

Figure A1: Normal Q-Q plot for sunlight and water



Three assumptions, conducted on the natural log dataset must be met for the analysis to be valid. Those are:

- Independence: Independence is met because our ou=eu=geraniums plants, there are no obvious clusters like where the seeds came from, and there were no repeated measures since the plants were measured once after 12 weeks.
- Equal Variance: After the transformation was performed, equal variance was met because largest and lowest standard deviations had a ratio of less than 2 ($0.46/0.26 = 1.769 < 2$).
- Normality: After the natural log transformation, normality is met because all points are close to the reference line (Figure A2).

Table A2: Summary Statistics of Sunlight by Water after Natural Log Transformation

sunlight.water	min	Q1	median	Q3	max	mean	sd	n
High.High	2.773	2.861	2.917	3.082	3.497	3.026	0.322	4
Low.High	1.386	1.690	1.869	1.946	1.946	1.767	0.264	4
High.Low	2.639	2.691	2.852	3.097	3.097	2.936	0.346	4
Low.Low	2.565	2.888	3.043	3.241	3.241	3.085	0.463	4

Figure A2: Normal Q-Q plot for sunlight and water after natural log transformation on “yield”

