BILLONES, CRISTEL KAYE P.
ROSALES, FRANCES ANETH C.

APM1111 - Statistical Theory
Formative Assessment 8

Data Analysis for PlantGrowth

Introduction

The PlantGrowth dataset captures the outcomes of an experiment examining plant growth under distinct conditions. Specifically, it measures the dried weight of plants cultivated in three separate conditions: a control group and two distinct treatment conditions. The dataset aims to assess any significant differences in plant yield among these conditions. This dataset serves as the basis for conducting statistical analyses, such as a one-way analysis of variance (ANOVA), to determine if there are statistically significant variations in plant growth between the different treatment groups and the control.

List of Assumptions

Assumptions #1: You have one dependent variable that is measured at the continuous level.

Assumptions #2: The variances of the dependent variable should be roughly equal across all groups.
Assumption #3: Each of the data observations should be independent between each other.

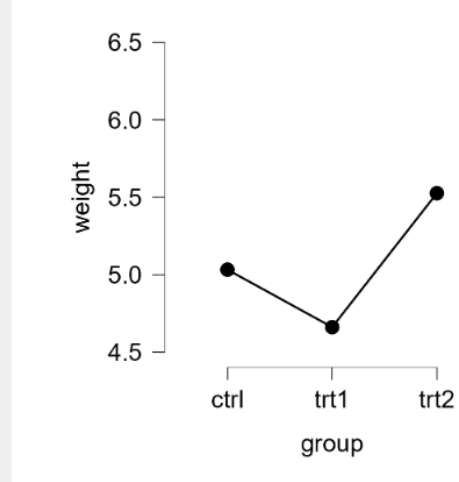
Assumption #4: There should be no significant outliers in the dependent variable within each group.

Assumption #5: Your dependent variable should be approximately normally distributed for each group of the independent variable.

Assumption #6: You have homogeneity of variances (i.e., the variance of the dependent variable is equal in each group of your independent variable).

Checking of Assumptions

Assumption #1: You have one dependent variable that is measured at the continuous level.



Remark: The dependent variable in this dataset, "weight," represents the measured weight of plants. This variable is on a continuous level as it involves quantitative measurements that can take on any value within a range.

Assumption #2: You have one independent variable that consists of two categorical, independent
groups

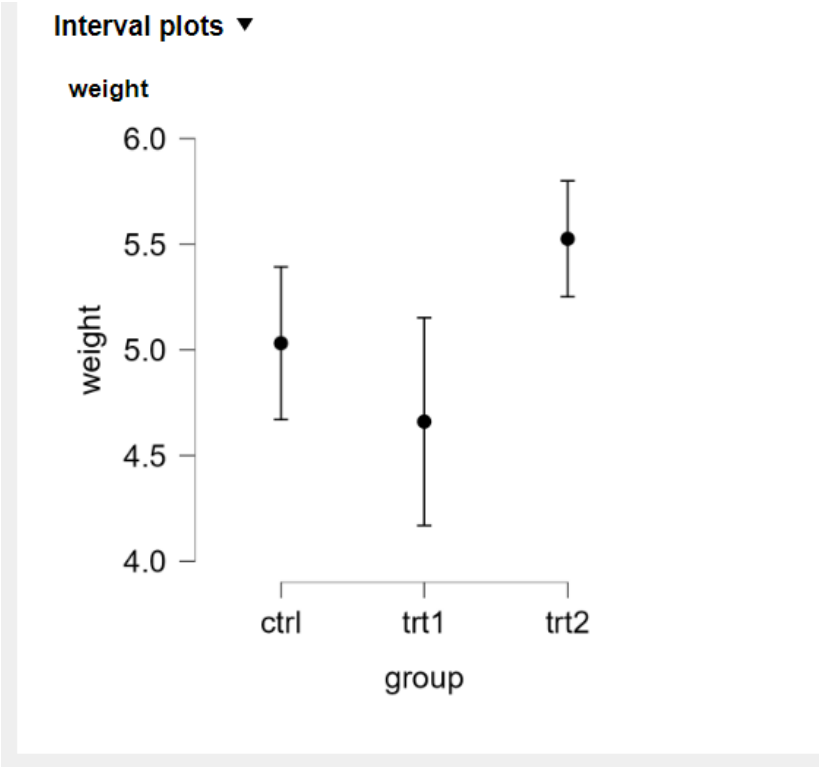
ANOVA - weight ▼

Homogeneity Correction	Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
None	group	3.766	2.000	1.883	4.846	0.016	0.264
	Residuals	10.492	27.000	0.389			
Brown-Forsythe	group	3.766	2.000	1.883	4.846	0.018	0.264
	Residuals	10.492	22.208	0.472			
Welch	group	3.766	2.000	1.883	5.181	0.017	0.264
	Residuals	10.492	17.128	0.613			

Note. Type III Sum of Squares

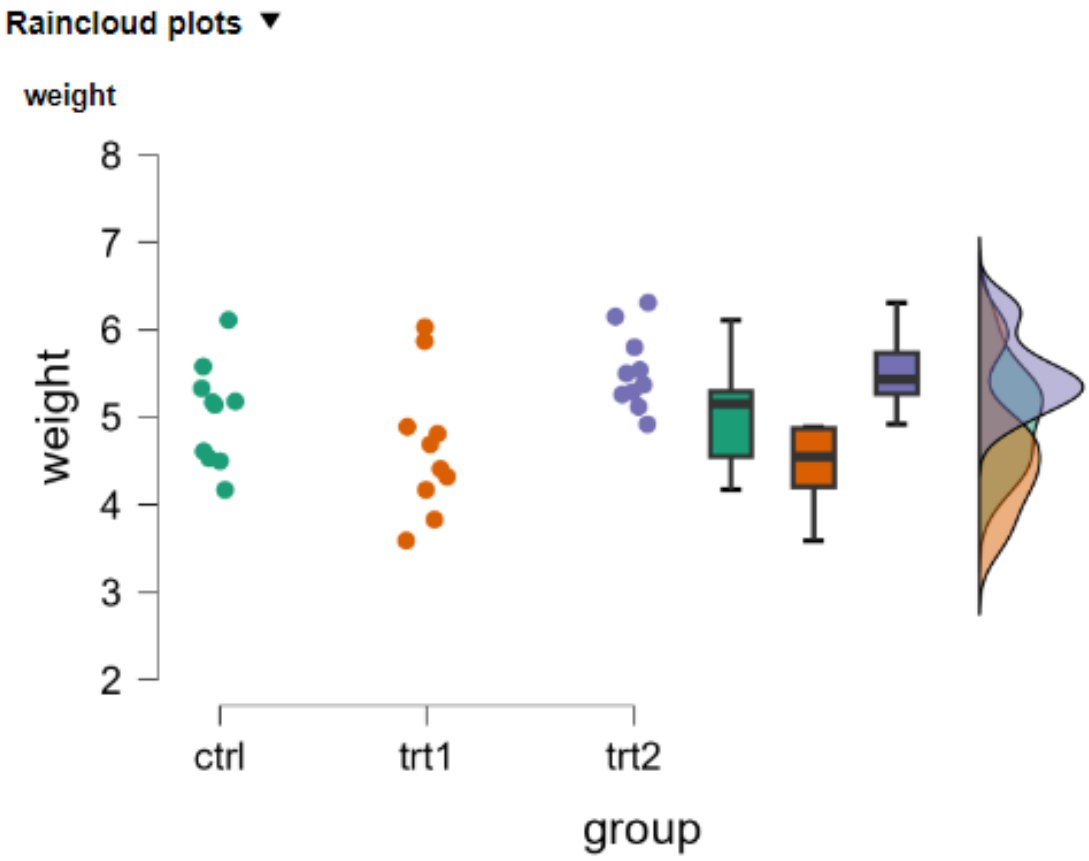
Remarks: As assumed, we have observed a dependent variable of homogeneity of variances, as assessed by the Welch Test accumulated p= 0.017 while on Brown- Forsythe p=0.018.

Assumption #3: You should have independence of observations.



Remark: Observations within each group and between groups appear independent, suggesting no discernible relationship among weight measurements within or across the treatment groups.

Assumption #4: There should be no significant outliers in the dependent variable within each group.



Remarks: Outliers can influence the results of statistical analyses, thus disproportionately can be observed. Therefore as seen in the boxplots, we indeed saw no significant outliers.

Assumption #5: Your dependent variable should be approximately normally distributed for each group of the independent variable.

Descriptive Statistics

Descriptive Statistics

	weight
Valid	30
Missing	0
Mean	5.073
Std. Deviation	0.701
Skewness	−0.162
Std. Error of Skewness	0.427
Kurtosis	−0.553
Std. Error of Kurtosis	0.833
Shapiro-Wilk	0.983
P-value of Shapiro-Wilk	0.892
Minimum	3.590
Maximum	6.310

Remarks: We know that the importance of accurate estimation of parameters and hypotheses in testing statistical methods applications is essential. Using the Shapiro-Wilk test we have obtained a p-value of 0.8, this can only imply that the null hypothesis or normal distribution is not rejected at the 0.05 significance level.

Assumption #6: You have homogeneity of variances (i.e., the variance of the dependent variable is equal in each group of your independent variable).

Test for Equality of Variances (Levene's)

F	df1	df2	p
1.237	2.000	27.000	0.306

Remark: There was homogeneity of variances of the dependent variable for all physical activity groups, as assessed by Levene's test of homogeneity of variances, p = 0.306

Computation:

ANOVA - weight

Homogeneity Correction	Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
None	group	3.766	2.000	1.883	4.846	0.016	0.264
	Residuals	10.492	27.000	0.389			
Brown-Forsythe	group	3.766	2.000	1.883	4.846	0.018	0.264
	Residuals	10.492	22.208	0.472			
Welch	group	3.766	2.000	1.883	5.181	0.017	0.264
	Residuals	10.492	17.128	0.613			

Note. Type III Sum of Squares

Descriptives - weight

group	N	Mean	SD	SE	Coefficient of variation
ctrl	10	5.032	0.583	0.184	0.116
trt1	10	4.661	0.794	0.251	0.170
trt2	10	5.526	0.443	0.140	0.080

Post Hoc Comparisons - group

		Mean Difference	SE	t	p_{tukey}
ctrl	trt1	0.371	0.279	1.331	0.391
	trt2	−0.494	0.279	−1.772	0.198
trt1	trt2	−0.865	0.279	−3.103	0.012

Note. P-value adjusted for comparing a family of 3

Reporting.

As observed using one-way ANOVA, the group is divided and classified into 3 groups namely; ctrl (n = 10), trt1 (n = 10), and trt2 (n = 10) which are divided equally. As also assumed, the database of PlantGrowth was observed to be normally distributed, with the tool of Shapiro-Wilk test obtaining a p-value of 0.8, thus can only mean that normal distribution is not rejected at the 0.05 significance level, and therefore distributed normally. Additionally, using mathematical and statistical tools such as the Welch Test and Brown-Forsyth, with ($p > 0.05$) therefore concluded that there was homogeneity of variances in the data.