Statistical Inference

Multiway Factorial Analysis of Variance

Contents

Multiway Factorial Analysis of Variance

- Two Way Anova can be extended to assess the effects of simultaneous applications of three or more factors.

- **Example:** If there are three factors, say A,B and C then we can study

  - The Main effects of A,B and C
  - Two way interactions A*B, A*C and B*C
  - Three way interaction A*B*C

- The researcher may decide to exclude higher order interactions as they are difficult to interpret.

Case Study

Background

Two new marketing campaigns are tested along with traditional campaign(Control).
The campaigns are tested in mid size and large size stores of 3 regions (North ,west, east).

Objective

To test whether there is significant difference in growth among three campaigns, three regions & two sizes.

Sample Size

Sample size: 72
Variables: campaign, region, size, growth

Data Snapshot

Three Way Anova

Variables

Observations

campaign	region	size	growth
Test1	north	mid	11.9
Test1	north	mid	11.8
Test1	north	mid	11.6
Test1	north	mid	11.4
Test1	north	large	11.8
Test1	north	large	11.7
Test1	north	large	11.4
Test1	north	large	11.5
Test1	west	mid	12.3
Test1	west	mid	12.1
Test1	west	mid	12
Test1	west	mid	12.6

Columns	Description	Type	Measurement	Possible values
campaign	Campaign	Character	Control, Test1, Test2	3
region	Region	Character	east, north, west	3
size	Size of the store	Character	large, mid	2
growth	Growth in sales	Numeric	Percentage	+/- values

Three-way ANOVA in R

```
# Import data
data<-read.csv("Three Way Anova.csv", header=TRUE)

# Anova table
anovatable<-aov(formula=growth~campaign*region*size,data=data)
summary(anovatable)
```

- *'anovatable' is user defined object name created to store output.*
- *'aov' is the R function for ANOVA .*
- *formula specifies 'growth' as analysis (dependent) variable and 'campaign','region',' size' as factor (independent) variable.*
- *summary function displays the ANOVA table output.*

Three-way ANOVA in R

```
# Output:
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
campaign	2	1.818	0.909	24.475	2.71e-08	***
region	2	24.656	12.328	332.024	< 2e-16	***
size	1	0.009	0.009	0.239	0.6266	
campaign:region	4	1.102	0.275	7.418	7.75e-05	***
campaign:size	2	0.370	0.185	4.986	0.0103	*
region:size	2	0.175	0.088	2.360	0.1041	
campaign:region:size	4	0.221	0.055	1.485	0.2196	
Residuals	54	2.005	0.037			

Interpretation :

- *Since p-value is <0.05 for campaign, region, reject H0. there is significant difference in growth among three campaigns and three regions. Also, campaign*region and Campaign*size interaction is significant.*

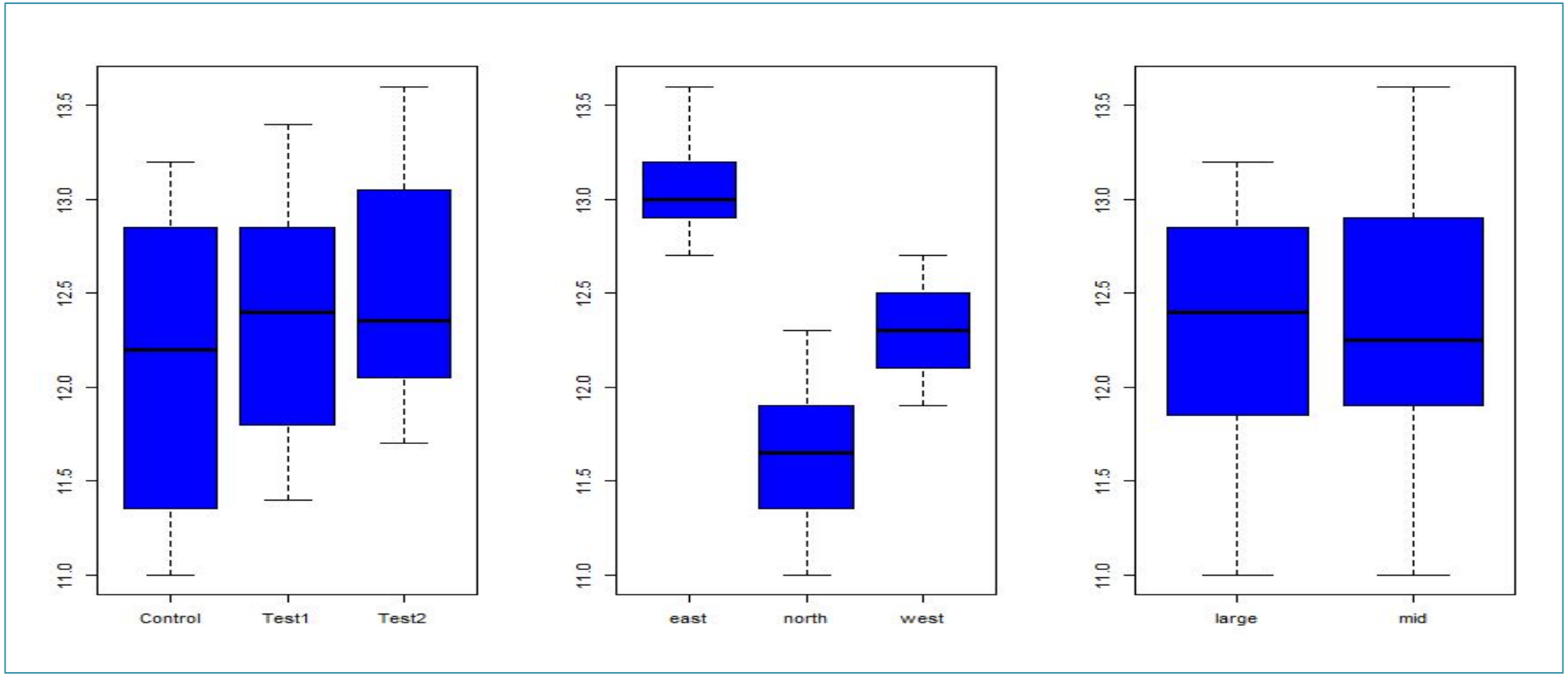
Visualize main effects

```
# Box plots for main effects

par(mfrow=c(1,3))

boxplot(growth~campaign,data=data,col="blue")
boxplot(growth~region,data=data,col="blue")
boxplot(growth~size,data=data,col="blue")
```

- *'par' function creates partition for graph output in 1 row and 3 columns.*
- *Boxplots are plotted for each factor with growth variable.*

Visualize main effects

```
# Output:
```



Interpretation :

- *There is significant difference in growth among three campaigns and three regions.*

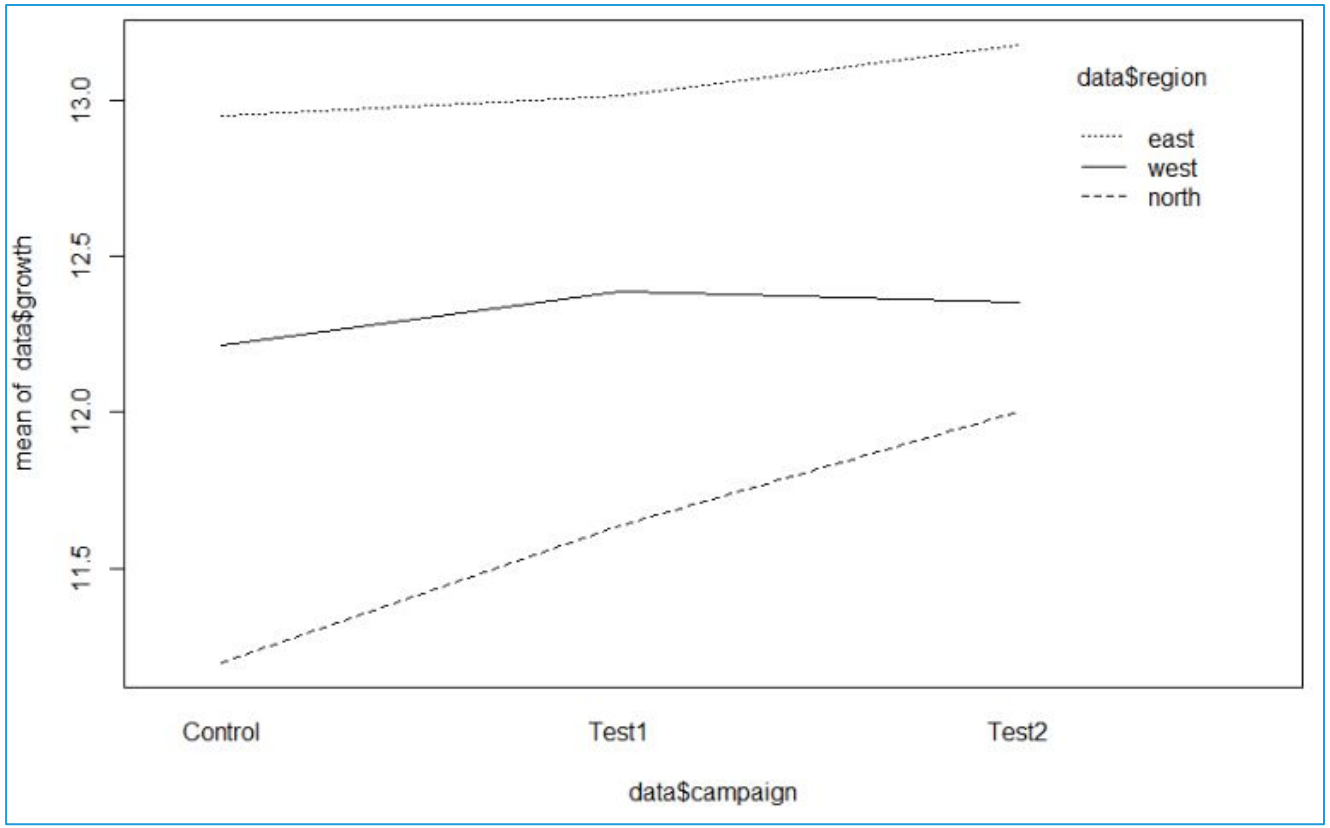
Visualize interaction effects

```
# Box plots for interaction effects

par(mfrow=c(1,1))

interaction.plot(data$campaign, data$region, data$growth)
```

- *interaction.plot function plots the mean (or other summary) of the response for two-way combination of factors (campaign, region), thereby illustrating possible interactions.*

Visualize interaction effects

```
# Output:
```



Interpretation :

- *Campaign and region interaction is having impact on growth of sales.*

Quick Recap

Three Way ANOVA	Two Way Anova can be extended to assess the effects of simultaneous applications of 3 or more factors.
Main Effects and Interactions	There are 3 main effects- One for each factor Interaction effects are 2-way and 3 –way.
Data Visualization	Box-Plot for main effect Interaction effect plot