Two-way ANOVA Interactions

David Gerard

2018-12-07

Objectives

- Gain intuitive understanding of Two-way ANOVA model.
- Chapter 13 in the book.

Pygmalion Effect Case Study

- Pygmalion Effect: high expectations of a supervisor translate into improved performance of subordinate.
- A company of soldiers contains three platoons each.
- Within each company, one platoon was randomly selected to be the "Pygmalion platoon."
- The platoon leader in the Pygmalion platoon was told by the army psychologist that his platoon was predicted to be superior.
- At end of basic training, all soldiers in each platoon were given a skill test.
- Data consist of average scores for each platoon.

The data

```r
library(Sleuth3)
data("case1302")
head(case1302)
```

```
##   Company     Treat Score
## 1      C1 Pygmalion  80.0
## 2      C1   Control  63.2
## 3      C1   Control  69.2
## 4      C2 Pygmalion  83.9
## 5      C2   Control  63.1
## 6      C2   Control  81.5
```

When to use two-way ANOVA

1. You have a quantitative response variable.
2. You have two categorical explanatory variables.
3. It is called *two-way* ANOVA because each observational unit may be placed into a two-way table according to group status in both categorical variables

Company	1	2	3	4	5	6	7	8	9	10
Pyg										
Non-Pyg										

One-way ANOVA Model

- Model: $Y_{ij} = \mu_i + \epsilon_{ij}$
- Y_{ij}: Value of observational unit j of group i.
- μ_i: Mean value for group i.
- ϵ_{ij}: Individual-specific noise for observational unit j of group i. Assumed to have mean 0 and variance σ^2.
- σ^2 is assumed to be the **same** for all observational units of all groups

Equivalent One-way ANOVA Model

- Model: $Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$
- μ: baseline value.
- α_i: Mean difference from baseline for group i.

Equivalent One-way ANOVA Model

- $\mu_i = \mu + \alpha_i$
- In R, the baseline is the mean of the first group listed when you use the `levels()` command.
- In SAS, it is the mean of the last group listed.
- In some other softwares, baseline is the average of the group means.
- Using this notation makes generalizing to two-way ANOVA easier.

Two-way ANOVA model: The additive model

- Model: $Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$
- Y_{ijk}: Value of observational unit k of group i of the first categorical variable and group j of the second categorical variable.

Two-way ANOVA model: The additive model

- Model: $Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$
- Y_{ijk}: Value of observational unit k of group i of the first categorical variable and group j of the second categorical variable.
- μ: baseline value.
- α_i: Additive effect of being in group i in categorical variable 1.
- β_j: Additive effect of being in group j in categorical variable 2.

Two-way ANOVA model: The additive model

- Model: $Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$
- Y_{ijk}: Value of observational unit k of group i of the first categorical variable and group j of the second categorical variable.
- μ: baseline value.
- α_i: Additive effect of being in group i in categorical variable 1.
- β_j: Additive effect of being in group j in categorical variable 2.
- ϵ_{ijk}: Individual-specific noise for observational unit k of group i of the first categorical variable and group j of the second categorical variable. Assumed to have mean 0 and variance σ^2.
- σ^2 is assumed to be the **same** for all observational units of all groups

Two-way ANOVA model with interaction

- Model: $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$
- $(\alpha\beta)_{ij}$: A single number, represents the *interaction* effect.
- This model says that every group has it's own mean, where a group is defined by the combination of both categorical variables.

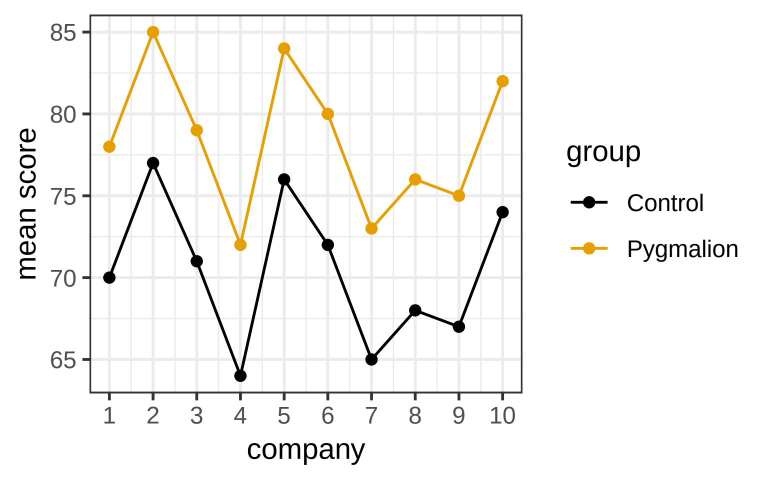
Two-way ANOVA model with interaction

- $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$

is equivalent to

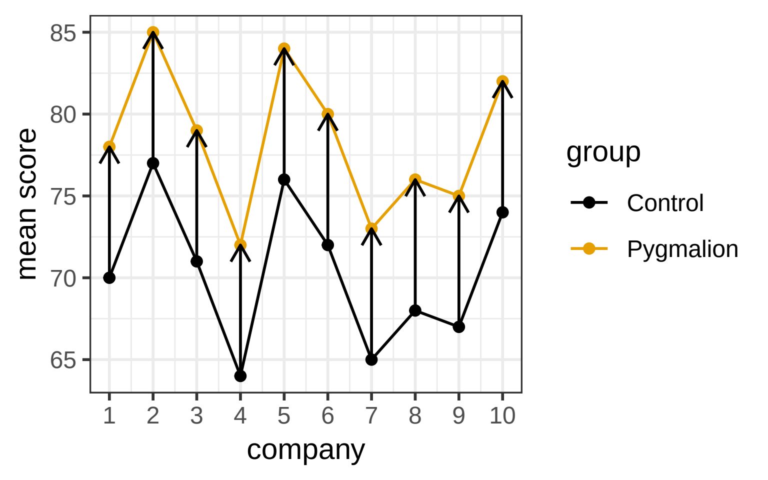
- $Y_{\ell k} = \mu + \tau_{\ell k} + \epsilon_{\ell_k}$, where $\ell = (i, j)$.
- *This is the exact same thing as the one-way ANOVA model.*
- Because each group is allowed to have its own unconstrained mean. In the additive-effect model, there are constraints.
- People often call the two-way ANOVA model with interaction the cell-means model.

The Additive Model



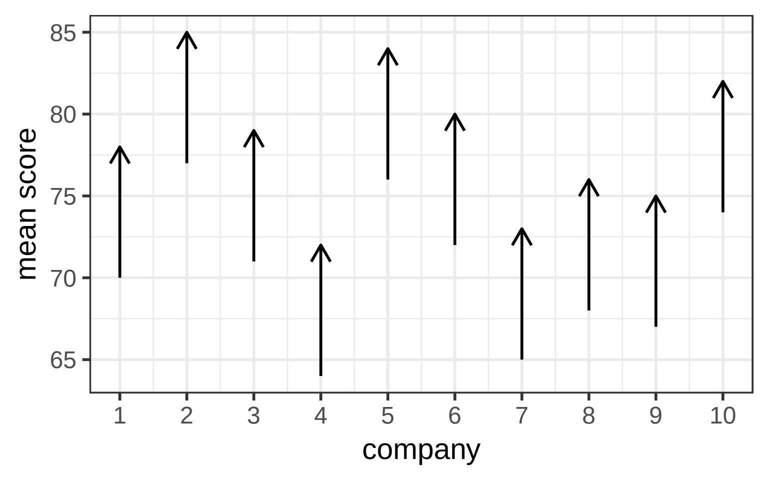
The Additive Model

- The additive effect of treatment is the same for all companies.

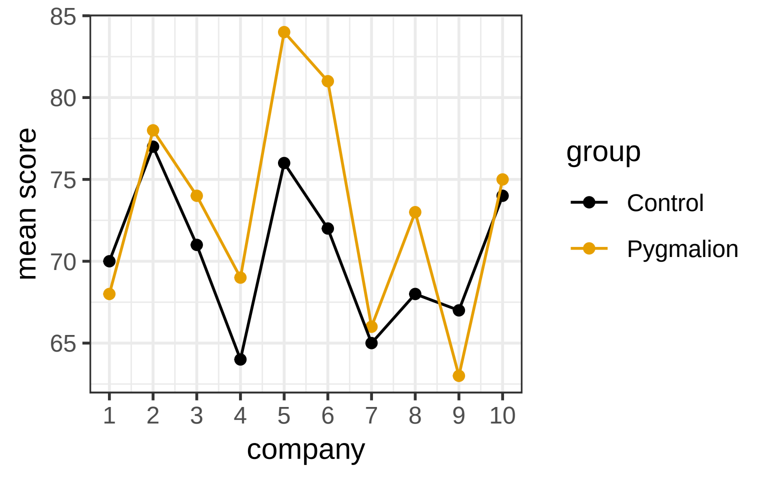


The Additive Model

- The additive effect of treatment is the same for all companies.

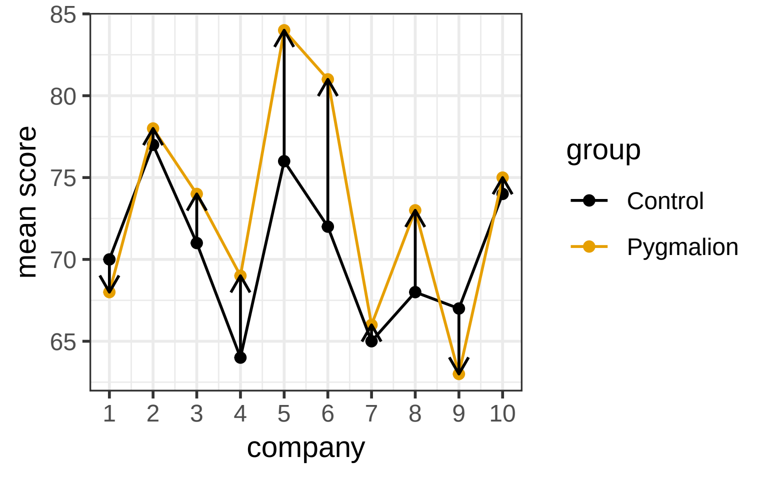


Cell Means Model



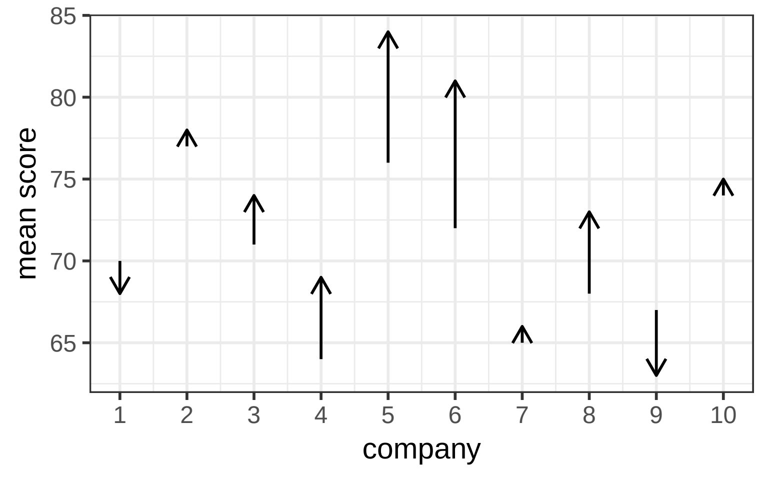
Cell Means Model

- The additive effect differs based on which company you are looking at.
- Not as interpretable if dependent on the company.

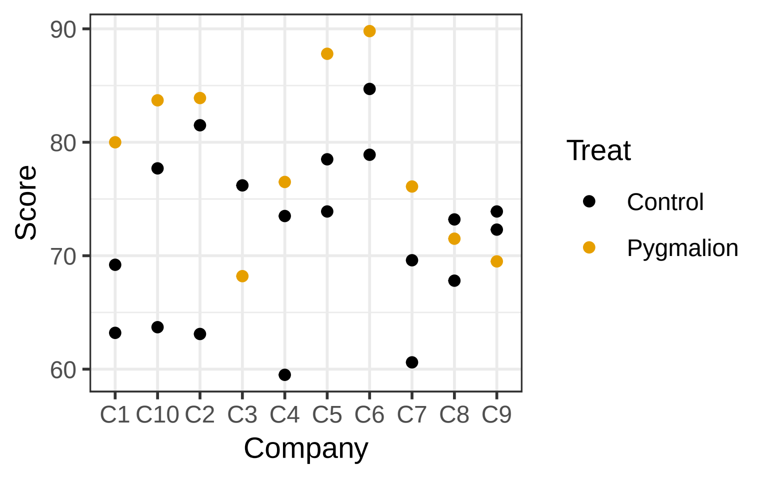


Cell Means Model

- The additive effect differs based on which company you are looking at.
- Not as interpretable if dependent on the company.



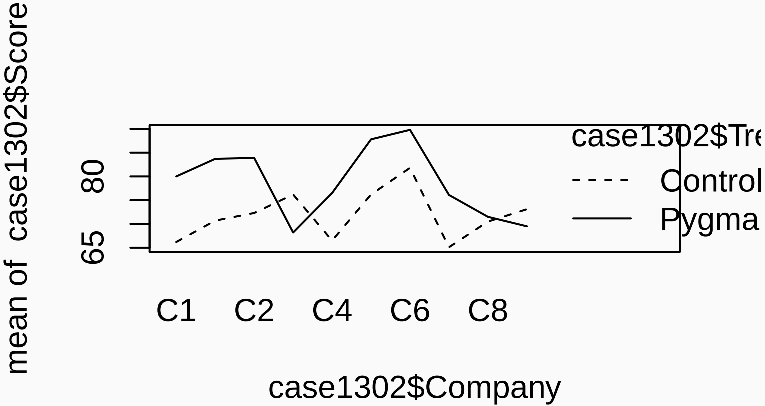
Real Data



90
80
70
60
Score
C1 C10 C2 C3 C4 C5 C6 C7 C8 C9
Company
Treat
Control
Pygmalion

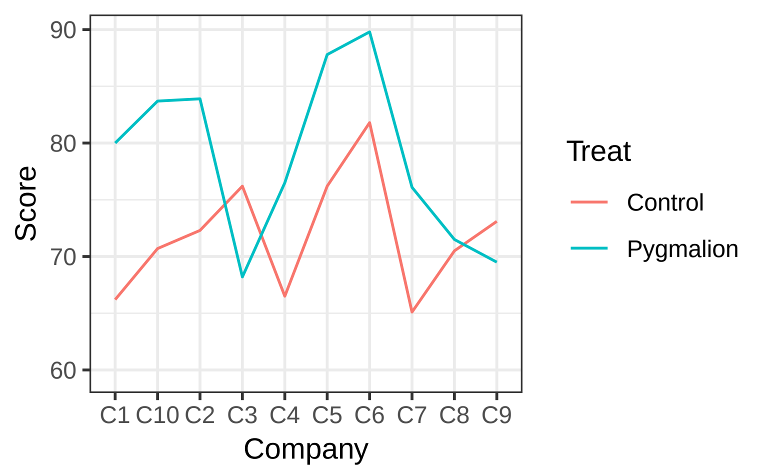
Quick Interaction Plots in R

```
interaction.plot(x.factor = case1302$Company,
                 trace.factor = case1302$Treat,
                 response = case1302$Score)
```



ggplot2 Interaction Plots in R

```
qplot(x = Company, y = Score,
      color = Treat, group = Treat,
      data = case1302, geom = "blank") +
  stat_summary(fun.y = mean, geom = "line")
```



Testing for Interactions

- Often, the first-step of a two-way ANOVA is to test for interactions.
- If we don't see strong evidence for interactions, we often proceed to assume additivity (due to its better interpretability).
- $H_0 : (\alpha\beta)_{ij} = 0$ for all i and all j.
- H_A : At least one $(\alpha\beta)_{ij} \neq 0$.

F-test for Interaction Effects

1. Estimate the group means under the full model (cell-means model; with interactions) and the reduced model (addititve model; without interactions).
2. Calculate residuals under both models: RSS_{full} and $RSS_{reduced}$.
3. Calculate the extra sums of squares: $ESS = RSS_{reduced} - RSS_{full}$.
4. Calculate F-statistic: $\frac{ESS/df_{extra}}{RSS_{full}/df_{full}}$
5. Compare to an $F_{df_{extra},df_{full}}$ distribution.

Degrees of Freedom

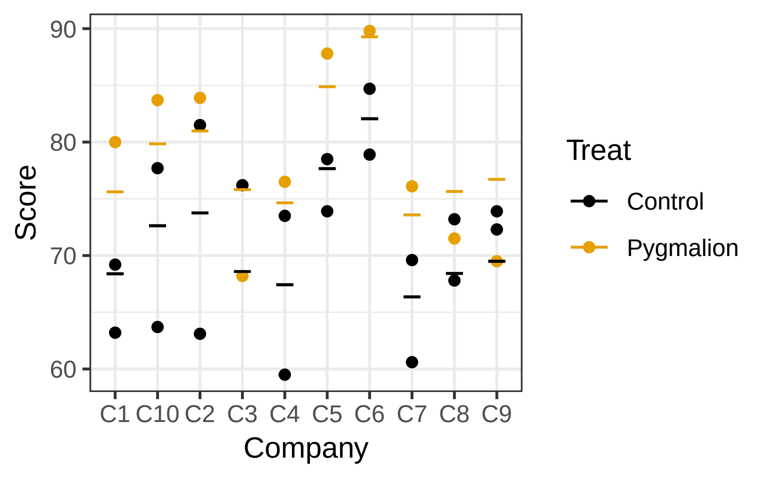
- Let n be the total number of observational units.
- In the full (cell-means) model, there are $I \times J$ parameters (how many groups there are, just like in one-way ANOVA).
- $df_{full} = n - IJ$
- In the reduced (additive) model, there are $I + J - 1$ parameters ($I - 1$ effects for variable 1, $J - 1$ effects for variable 2, and the baseline value).
- $df_{reduced} = n - I - J + 1$
- $df_{extra} = IJ - I - J + 1$

Estimate Under Reduced

- Model: $Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$
- Under **balanced** designs (sample size is the same for all groups), the estimates are provided below.
- $\hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y}_{1..}$
- $\hat{\beta}_j = \bar{Y}_{.j.} - \bar{Y}_{.1.}$
- $\hat{\mu} = \bar{Y}_{.1.} + \bar{Y}_{1..} - \bar{Y}_{...}$
- Mean estimate in cell $(i, j) = \bar{Y}_{.j.} + \bar{Y}_{i..} - \bar{Y}_{...}$
- Estimates are much more complicated in non-balanced designs.

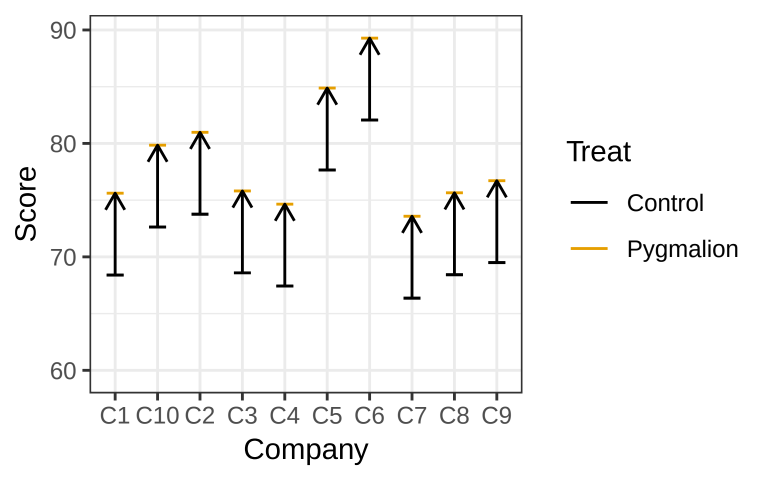
Reduced (Additive) Model

- Mean Estimates



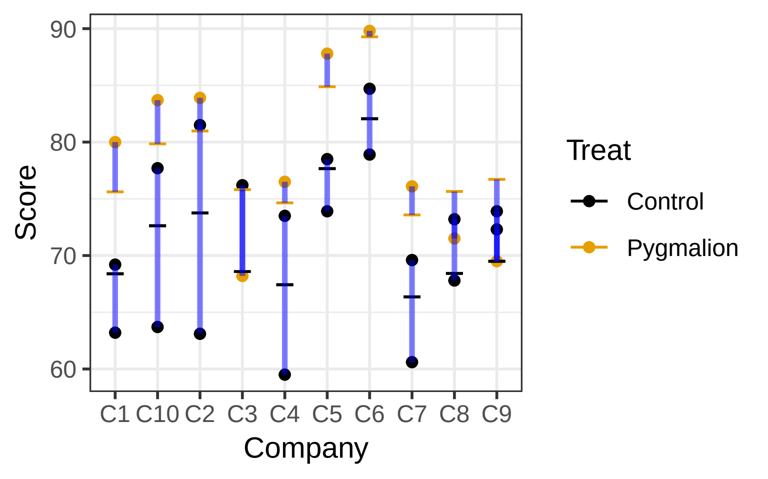
Reduced (Additive) Model

- Effect size estimate same for all groups



Reduced (Additive) Model

- Find the Residuals

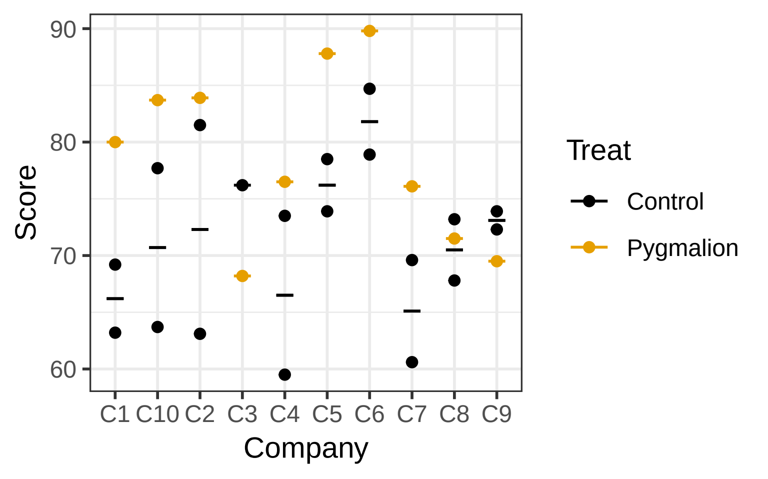


Estimate Under Full

- Model: $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$
- Mean estimate in cell $(i, j) = \bar{Y}_{ij\cdot}$

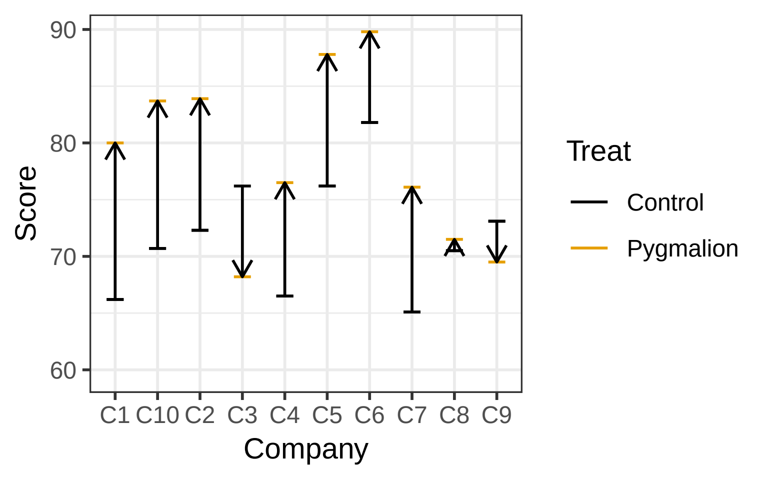
Full (Cell-means) Model

- Mean Estimates



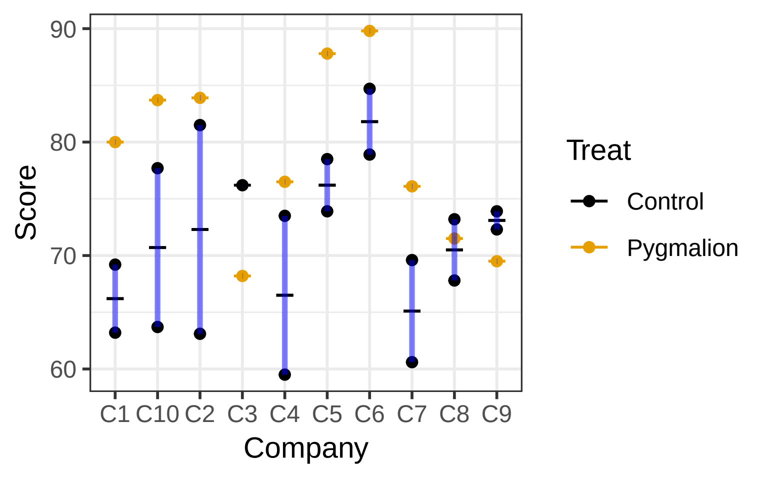
Full (Cell-means) Model

- Effect size estimate *different* for all groups



Full (Cell-means) Model

- Find the Residuals



Sum of Squared Residuals

- $RSS_{full} = 467.04$
- $df_{full} = n - IJ =$ 29 - 10 $\times$ 2 $= 9$

Sum of Squared Residuals

- $RSS_{full} = 467.04$
- $df_{full} = n - IJ =$ 29 - 10 $\times$ 2 $=$ 9
- $RSS_{reduced} = 778.5039$
- $df_{reduced} = n - I - J + 1 =$ 29 - 10 - 2 $+$ 1 $=$ 18

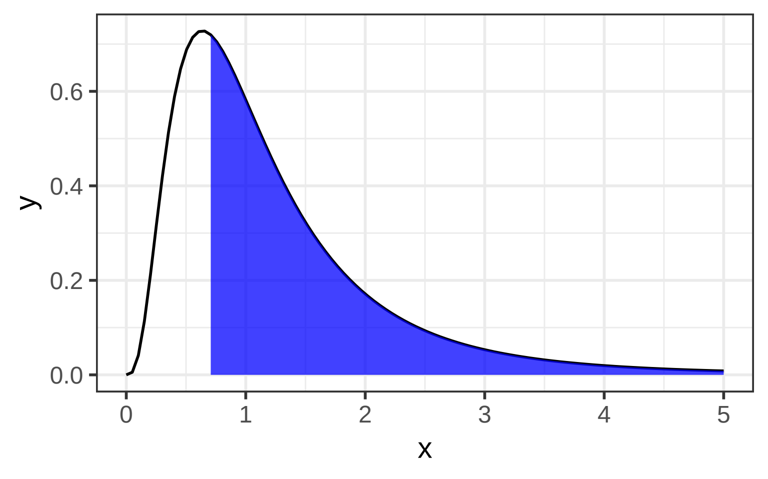
Sum of Squared Residuals

- $RSS_{full} = 467.04$
- $df_{full} = n - IJ = 29 - 10 \times 2 = 9$
- $RSS_{reduced} = 778.5039$
- $df_{reduced} = n - I - J + 1 = 29 - 10 - 2 + 1 = 18$
- $ESS = RSS_{reduced} - RSS_{full} = 311.4639$
- $df_{extra} = df_{reduced} - df_{full} = 9$

Sum of Squared Residuals

- $RSS_{full} = 467.04$
- $df_{full} = n - IJ = 29 \text{ - } 10 \times 2 = 9$
- $RSS_{reduced} = 778.5039$
- $df_{reduced} = n - I - J + 1 = 29 \text{ - } 10 \text{ - } 2 + 1 = 18$
- $ESS = RSS_{reduced} - RSS_{full} = 311.4639$
- $df_{extra} = df_{reduced} - df_{full} = 9$
- F-statistic $= \frac{ESS/df_{extra}}{RSS_{full}/df_{full}} = 0.6669.$

Compare 0.6669 to a $F_{9,9}$ Distribution



Compare 0.6669 to a $F_{9,9}$ Distribution

```r
pf(0.667, df1 = 9, df2 = 9, lower.tail = FALSE)
```

```
## [1] 0.722
```

- There is no evidence that the there is an interaction between Company and Treatment.

In R

```
aout_int <- aov(Score ~ Company * Treat, data = case1302)
anova(aout_int)
```

```
## Analysis of Variance Table
##
## Response: Score
##               Df Sum Sq Mean Sq F value Pr(>F)
## Company        9    671      75    1.44  0.299
## Treat          1    339     339    6.53  0.031
## Company:Treat  9    311      35    0.67  0.722
## Residuals      9    467      52
```

What those numbers mean

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Company	blah	blah	blah	blah	blah
Treat	blah	blah	blah	blah	blah
Company:Treat	df_{extra}	ESS	ESS/df_{extra}	F-stat	p-value
Residuals	df_{full}	RSS_{full}	RSS_{full}/df_{full}		

R syntax

- `*`: include this interaction and all smaller-order terms
- `:`: include this interaction
- `+`: add another term
- `-`: remove a term

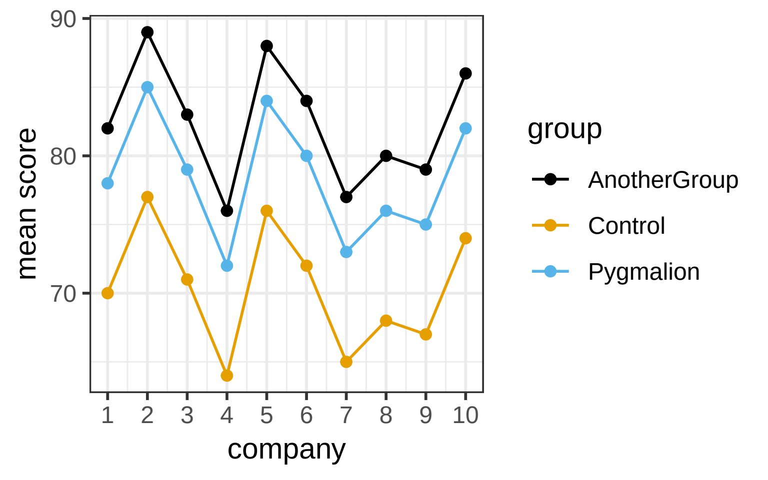
E.g.

- `Company + Treat` fits $\mu + \alpha_i + \beta_j$
- `Company * Treat` fits $\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$
- `Company * Treat - Company:Treat` fits $\mu + \alpha_i + \beta_j$

A Closer Look at Additivity

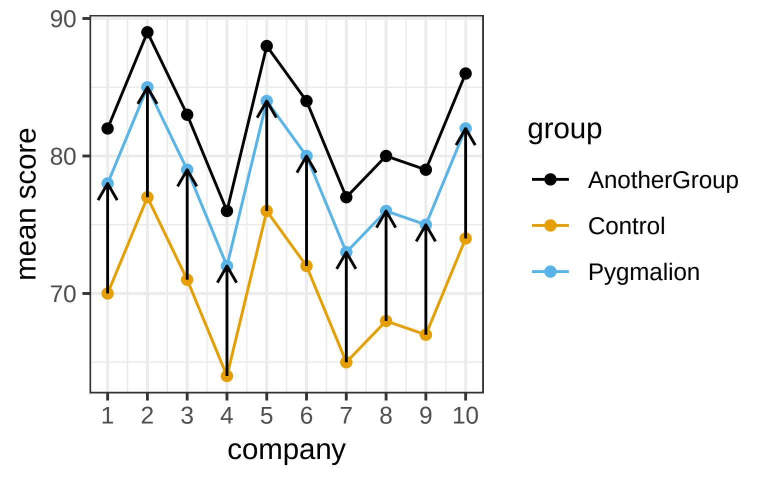
Additive With Three Categories

- More than two categories in each variable.



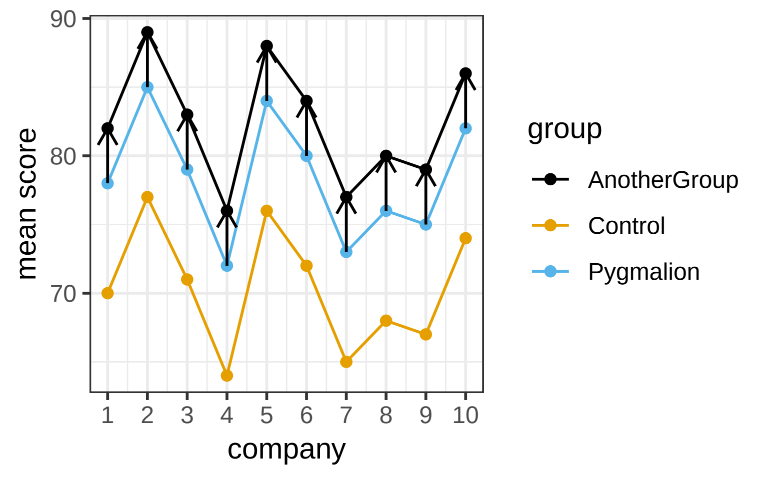
Additive With Three Categories

- The additive effect of treatment is the same for all companies.



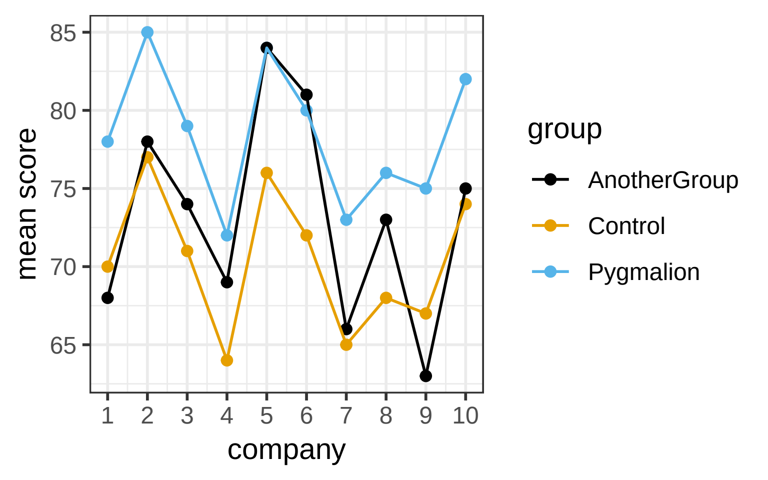
Additive With Three Categories

- The additive effect of treatment is the same for all companies.



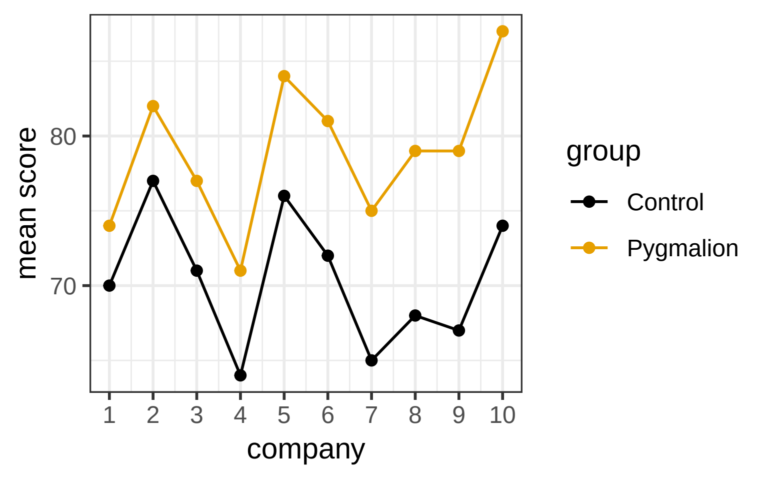
NON-Additive With Three Categories

- More than two categories in each variable.



NON-Additivity

- Same direction of an effect, but non-additive



What to do when there are significant interactions?

- The best course of action is to show an interaction plot (like the one above).

```r
qplot(Company, Score, group = Treat,
      color = Treat, geom = "blank",
      data = case1302) +
  stat_summary(fun.y = mean, geom = "line")
```