### Big Idea:

• Designing out the variability in the response from a particular source

### **Example 3: Strawberry Storage**

A study (Smith & Skog, 1992) was conducted in order to investigate the effects of storage method on the firmness of strawberries. As strawberries harvested, they are placed into storage containers called clamshells. Because clamshells have holes in them, air can circulate from outside the clamshell into the clamshell and vice versa, while the clamshell protects the fruit.

V J Matthew/Shutterstock.com

are

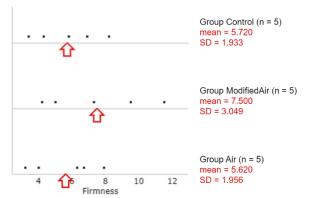
To test the potential effect of changing the composition of the air in which strawberries are stored, researchers obtained 15 clamshells filled with

strawberries. The strawberries came from 5 different varieties: Allstar, Bounty, Kent, Selva and Vesper. (Some varieties are naturally very hardy, others are more fragile.) (The berries in a clamshell were all from the same variety.)

The treatments of interest were: regular air (approximately 21% oxygen and 0.04% carbon dioxide), modified air (18% oxygen and 15% carbon dioxide) and a control condition. Air and modified air clamshells were measured after 42 hours. The control conditional clamshells were measured immediately after packing. The response of interest was the firmness (Newtons) of each clamshell. (The firmness of each berry was measured and then averaged to obtain 'clamshell firmness'.)

#### **1.** Sources of variation diagram:

Observed Variation in:	Sources of explained variation	Sources of unexplained variation
Clamshell firmness (Newtons)	<ul> <li>Storage method (Air, Control, Modified Air)</li> </ul>	<ul> <li>Variety of the berries (5 different varieties used)</li> </ul>
<ul> <li>Inclusion criteria</li> <li>Strawberries in same area of California; these 5 varieties</li> <li>Design</li> <li>Time at which the berries were picked, size and color all controlled</li> </ul>	TVIOUTIEU ATT	<ul> <li>Slight variations due to size, color, etc.</li> <li>Slight variations due to packing, weather, etc.</li> <li>Measurement process</li> <li>unknown</li> </ul>



 Source
 df Sum Sq Mean
 Sq F p-value

 Treatments
 2 11.19 5.59 1.00 0.3981

 Error
 12 67.44 5.62

Total 14 78.62

- **2.** Why is the difference in means not statistically significant?
- **3.** How could the researchers design their study to 'take into account' the variety-to-variety variation

in clamshell firmness?

# Analysis needs to match the design:

**2**Account for the variety-to-variety variation in clamshell firmness within the analysis

**Method 1:** <u>Quantify how much variety-to-variety variation there is in firmness</u>, then <u>move this variation</u> from the unexplained column into the explained sources of variation column.

Observed Variation in:	Sources of explained variation	Sources of unexplained variation
Avg firmness of clamshell	Storage method (Air, Control, Modified Air)	Variety of the berries (5 different varieties used)
<ul> <li>Inclusion criteria</li> <li>Strawberries in same area of California; these 5 varieties</li> <li>Design</li> <li>Time at which the berries were picked, size and color all controlled</li> </ul>	Variety (Allstar, Bounty, Kent, Selva, Vesper)	<ul> <li>Slight variations due to size, color, etc.</li> <li>Slight variations due to packing, weather, etc.</li> <li>The mechanism used to measure firmness</li> <li>unknown</li> </ul>

Word Model:



**4.** Determine the effect of each variety:

effect of Allstar = group mean - overall mean = 7.7 - 6.28 = 1.42 N effect of Bounty = -1.85 effect of Kent = 0.19

effect of Selva = 2.92

effect of Vesper = -2.68

Determine the SSvariety:

SSVariety = 
$$\sum n_j \left(effect_j^2\right) = 3(1.42)^2 + ... + 3(-2.68)^2$$

	Variety	Storage	Firmness
1	Allstar	Air	6.7
2	Allstar	ModifiedAir	9.5
3	Allstar	Control	6.9
4	Bounty	Air	4
5	Bounty	ModifiedAir	5
6	Bounty	Control	4.3
7	Kent	Air	6.3
8	Kent	ModifiedAir	7.3
9	Kent	Control	5.8
10	Selva	Air	7.9
11	Selva	ModifiedAir	11.5
12	Selva	Control	8.2
13	Vesper	Air	3.2
14	Vesper	ModifiedAir	4.2
15	Vesper	Control	3.4

SSVariety = 63.51

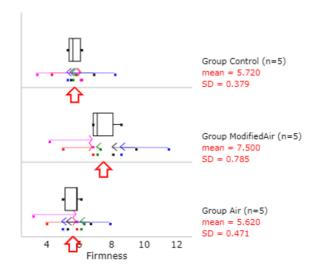
**5.** Fill in the ANOVA table below accounting for variety. The SS for storage method and total are already filled in to help you out. The following relationships need to hold: SSTotal = SSvariety + SSstorage method + SSError and df Total = df variety + df storage method + df Error.

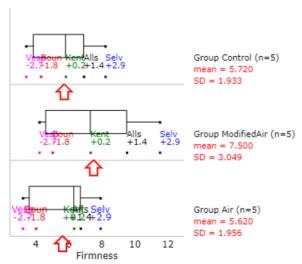
Source	DF	Sum of Squares	Mean Squares	Adjusted F-stat
Variety (= Block)				
Storage Method		11.188		
Error				
Total		78.624		•

**6.** When we remove the effect of each variety from each of the firmness values, what happens to the storage means? What happens to the variation within the storage groups?

Figure 1: Clamshell Firmness

Figure 2: **Block-adjusted Firmness** 





 Source
 df Sum Sq Mean
 Sq F p-value

 Treatments
 2 11.19 5.59 1.00 0.3981

 Error
 12 67.44 5.62

 Total
 14 78.62

 Source
 df
 Sum
 Sq
 Mean
 Sq
 F
 p-value

 Treatments
 2
 11.19
 5.59
 11.40
 0.0046

 Blocks
 4
 63.51
 15.88
 32.36
 0.0000

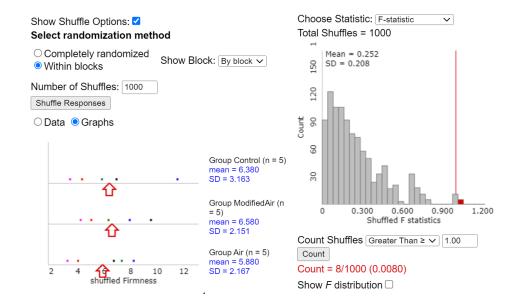
 Error
 8
 3.93
 0.49

 Total
 14
 78.62

## Method 2: Simulation-based Inference

### Adjust the simulation to take into account the variety-to-variety variation

Use the original F-statistic = 1.00 and shuffle within blocks. (Don't want to account for blocks twice.)



## Null and alternative hypotheses:

 $H_o$ :  $\mu_{\text{Air}}$  =  $\mu_{\text{MA}}$  =  $\mu_{\text{Control}}$  , after accounting for strawberry variety

H<sub>a</sub>: At least one of the μi differs, after accounting for strawberry variety

OR

 $\rm H_{o}$ : There is not a true or real effect of storage method on strawberry firmness after accounting for strawberry variety

H<sub>a</sub>: There is a true or real effect of storage method on strawberry firmness after accounting for strawberry variety