Split plots

STAT 6160 • Experimental Design • Fall 2026

Siddharth Vishwanath

University of Virginia

§1. Chocolate cakes

Chocolate cakes

Do the recipe and the baking temperature change how tender a chocolate cake is?

Factor	Levels
R: recipe	A, B, C
T: baking temperature	175, 185, 195, 205, 215, 225 °C

- Batter is mixed one **batch** at a time, from one recipe. A batch can make up to six cakes.
- Each cake is baked at one temperature.
- Response: **breakage angle** of the cake, in degrees, a measure of tenderness.

18 cakes: every recipe at every temperature. How would you make them?

Cook (1938); Cochran and Cox (1957), p. 300. Real data.

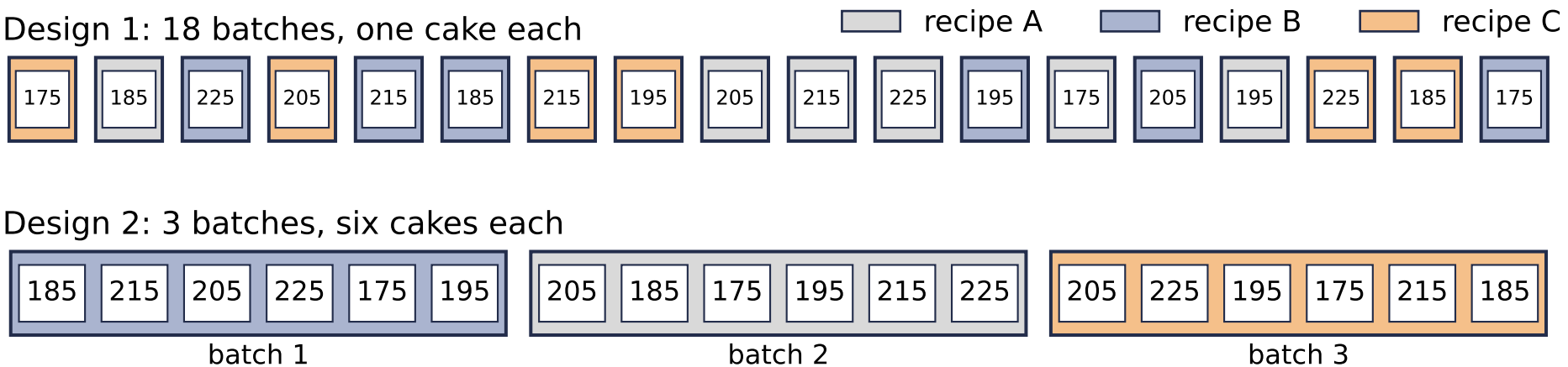


The **breakage angle** of a cake refers to **an objective measurement of cake tenderness derived from classic food science experiments** (such as Frances E. Cook's 1938 thesis and Cochran & Cox's *Experimental Designs*, featured in the **R `cake` dataset**). www.sumsar.net +2

How It Is Measured

- **Method:** One half of a slab of cake is held fixed, while the other half is pivoted around the middle until it breaks.
- **Scale:** The angle through which the moving half revolves is read on a circular scale.
- **Key Finding:** The breakage angle increases as the baking temperature increases. For example, in the classic dataset comparing three recipes (A, B, C) across six temperatures (175–225°C), **Recipe C baked at 225°C** yielded the highest breakage angle. r-packages.io +2

Two ways to run it



- **Design 1.** Every cake gets its own batch of batter. 18 batches are mixed, in random order.
- **Design 2.** One batch per recipe. Each batch is split into six cakes, one per temperature, in random order.
- The experiment: design 2, run 15 times. 45 batches, 270 cakes.

What does design 2 save? What does it give up?

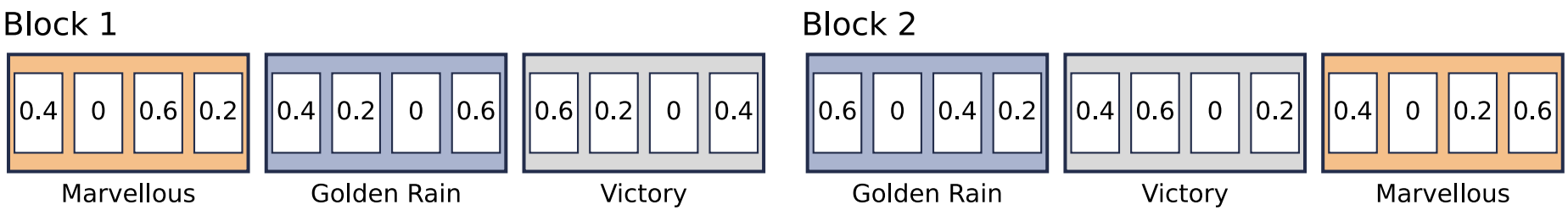
Whole plots and subplots

A **whole plot** is the experimental unit for one factor, the **whole-plot factor**. Each whole plot is split into **subplots**, the experimental units for a second factor, the **subplot factor**.

	Unit	How many	Factor	Randomized
Block	one run of design 2	15		
Whole plot	a batch of batter	45	recipe	within each block
Subplot	a cake	270	temperature	within each batch

- Every run holds one batch of every recipe: the 15 runs are **blocks**.
- Every level of the subplot factor appears in every whole plot.

Where split plots come from



The name is from field experiments. Each plot of land is sown with one variety of oats, then split into four subplots that get different amounts of nitrogen.

Examples:

- A process in stages: a batch is made, then divided and finished in several ways.
- A treatment applied at a large scale: irrigation, a curriculum for a classroom.
- A factor that is hard to change: the temperature of a furnace.

Yates (1935), nitrogen in cwt per acre.

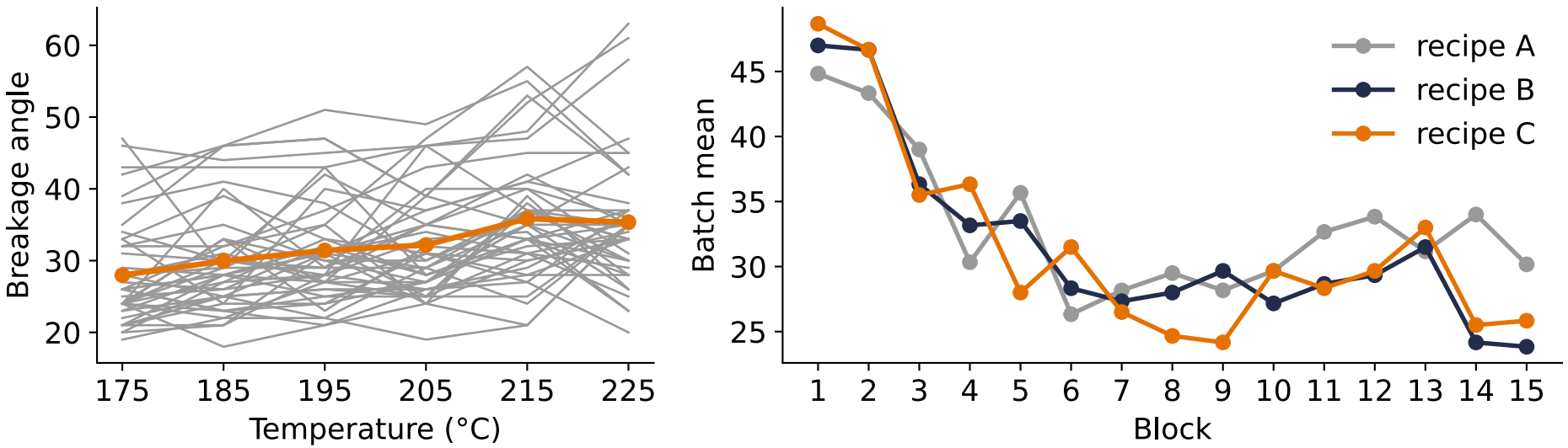
The data

```
cake = pd.read_csv("data/cake.csv").rename(columns={"replicate": "block"})
cake["batch"] = cake.recipe + cake.block.astype(str)
wide = cake.pivot(index=["block", "recipe"], columns="temp", values="angle")
wide["batch mean"] = wide.mean(axis=1)
print(wide.head(9).round(1))
```

```
temp          175  185  195  205  215  225  batch mean
block recipe
1     A        42   46   47   39   53   42        44.8
      B        39   46   51   49   55   42        47.0
      C        46   44   45   46   48   63        48.7
2     A        47   29   35   47   57   45        43.3
      B        35   46   47   39   52   61        46.7
      C        43   43   43   46   47   58        46.7
3     A        32   32   37   43   45   45        39.0
      B        34   30   42   35   42   35        36.3
      C        33   24   40   37   41   38        35.5
```

Real data from Cochran and Cox (1957)

Batch by batch



- Left: one gray line per batch. Right: the 45 batch means, by block.

What moves a whole line up or down? What separates the six points on one line?

Two experiments in one

	Recipe	Temperature
Assigned to	a batch	a cake within a batch
Replicates of each level	15 batches	45 cakes
Compared	between batches	within batches
Its blocks	the 15 runs of design 2	the 45 batches

- Keep only the 45 batch means: a **randomized block design** for recipe, in 15 blocks.
- Look inside the batches: a **randomized block design** for temperature. Each batch is a block.

Which effects cancel from a recipe difference? From a temperature difference?

Your turn: two districts

Two school districts, four classrooms each. Every classroom teaches one term with the new curriculum and one with the current one, in a random order, and students are tested after each term.

- What is the whole plot? What is the subplot?
- Which factor is the whole-plot factor, and which is the subplot factor?
- Which comparison is made between classrooms, and which within?

§2. The model

The split-plot model

y_{ijk} = cake-angle at temperature k from the batch of recipe i in block j

$$i = 1, \ldots a; \qquad j = 1, \ldots b; \qquad k = 1, \ldots c$$

$$y_{ijk} = \underbrace{\mu + \tau_i + \beta_j + \eta_{ij}}_{\text{the batch}} + \underbrace{\gamma_k + (\tau\gamma)_{ik} + \epsilon_{ijk}}_{\text{the cake within the batch}}$$

$$\eta_{ij} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2_W), \qquad \epsilon_{ijk} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2), \qquad \text{independent of each other}$$

- τ_i :
- β_j :
- η_{ij} :
- γ_k :
- $(\tau\gamma)_{ik}$:

Here: $a = 3, b = 15, c = 6, N = abc = 270$. The fixed effects sum to zero.

Two error terms

A batch mean, and a cake's deviation from its batch mean:

$$\bar{y}_{ij\bullet} = \mu + \tau_i + \beta_j + \eta_{ij} + \bar{\epsilon}_{ij\bullet}, \qquad y_{ijk} - \bar{y}_{ij\bullet} = \gamma_k + (\tau\gamma)_{ik} + \epsilon_{ijk} - \bar{\epsilon}_{ij\bullet}$$

Recipe:

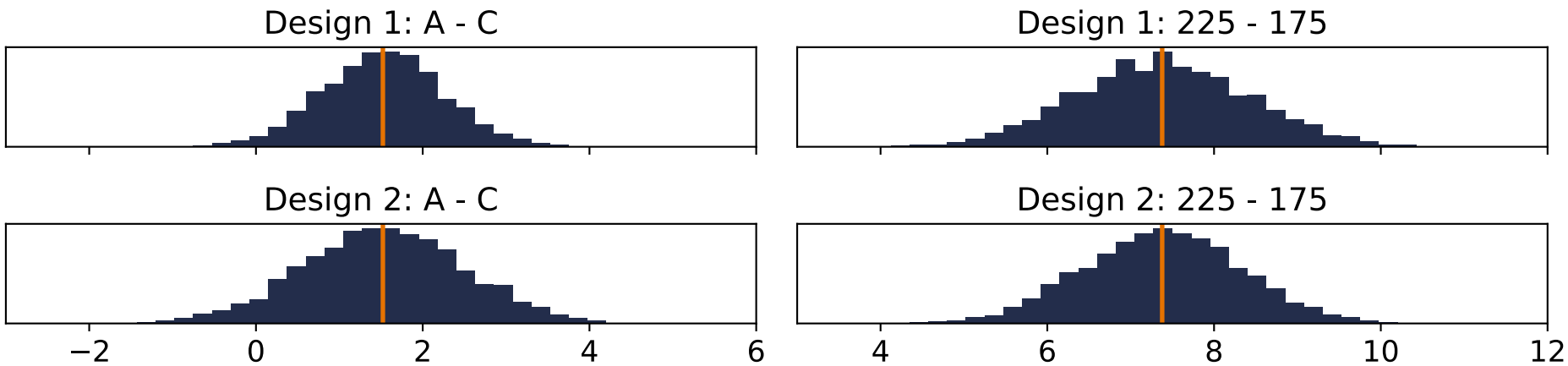
$$\text{Var}\,(\bar{y}_{2\bullet\bullet} - \bar{y}_{1\bullet\bullet}) =$$

Temperature:

$$\text{Var}\,(\bar{y}_{\bullet\bullet 2} - \bar{y}_{\bullet\bullet 1}) =$$

Repeat the experiment

Simulated: 15 blocks, $\sigma_W = 1.93$, $\sigma = 4.52$, 4000 experiments per design.



SD	A - C: formula	simulated	225 - 175: formula	simulated
Design 1	0.73	0.73	1.04	1.04
Design 2	0.98	0.97	0.95	0.95

Which factor would you put on the subplots?

Orange: the true differences, 1.52 and 7.38. Code in the appendix.

§3. The analysis

Between batches and within batches

$$y_{ijk} - \bar{y}_{\bullet\bullet\bullet} = \underbrace{(\bar{y}_{ij\bullet} - \bar{y}_{\bullet\bullet\bullet})}_{\text{between batches}} + \underbrace{(y_{ijk} - \bar{y}_{ij\bullet})}_{\text{within batches}}, \qquad I = H_{\text{batches}} + (I - H_{\text{batches}})$$

With Z the $N \times ab$ matrix of batch indicators, as for the classrooms:

$$y = X\beta + Z\eta + \epsilon, \qquad H_{\text{batches}} = Z(Z^\top Z)^{-1} Z^\top, \qquad \text{Var}\,(y) = \sigma^2 I + c\,\sigma_W^2\, H_{\text{batches}}$$

```
Z = pd.get_dummies(cake.batch).to_numpy(float)
H_batches = Z @ np.linalg.inv(Z.T @ Z) @ Z.T
d = (cake.angle - cake.angle.mean()).to_numpy()
between = d @ H_batches @ d
within = d @ (np.eye(270) - H_batches) @ d
print(np.trace(H_batches).round())
print(between.round(), within.round())
```

```
45.0
11538.0 6605.0
```

- Between batches: 45 batch means, 44 df. Within batches: $270 - 45 = 225$ df.

The ANOVA table

	Source	df	$\mathbb{E}(\text{MS})$
Between batches	Blocks	$b-1=14$	$\sigma^2+c\sigma_W^2+\frac{ac}{b-1}\sum_j \beta_j^2$
	Recipe	$a-1=2$	$\sigma^2+c\sigma_W^2+\frac{bc}{a-1}\sum_i \tau_i^2$
	Whole-plot error	$(a-1)(b-1)=28$	$\sigma^2+c\sigma_W^2$
Within batches	Temperature	$c-1=5$	$\sigma^2+\frac{ab}{c-1}\sum_k \gamma_k^2$
	$R\times T$	$(a-1)(c-1)=10$	$\sigma^2+\frac{b}{(a-1)(c-1)}\sum_{i,k}(\tau\gamma)_{ik}^2$
	Subplot error	$a(b-1)(c-1)=210$	σ^2

- The three between-batch lines carry $c\sigma_W^2$. The three within-batch lines do not.
- Whole-plot error is the block by recipe interaction of the 45 batch means.

Why is the subplot error $a(b-1)\times(c-1)$?

Three F tests

$$F_R = \frac{\text{MS}_R}{\text{MS}_{\text{WP}}}, \qquad F_T = \frac{\text{MS}_T}{\text{MS}_{\text{SP}}}, \qquad F_{RT} = \frac{\text{MS}_{RT}}{\text{MS}_{\text{SP}}}$$

- Recipe was applied to batches. F_R is the randomized-block F of the ab batch means.
- Temperature was applied to cakes. F_T and F_{RT} use the subplot error.
- Cakes: F_R on $(2, 28)$ df, F_T on $(5, 210)$, F_{RT} on $(10, 210)$.

With one block, what is left to test recipe against?

§4. The cakes in Python

The ANOVA table*

```
fit = smf.ols("angle ~ C(block) * C(recipe) + C(recipe) * C(temp)", cake).fit()
table = anova_lm(fit).loc[["C(block)", "C(recipe)", "C(block):C(recipe)",
                           "C(temp)", "C(recipe):C(temp)", "Residual"]]
table.index = ["blocks", "recipe", "whole-plot error",
               "temperature", "R x T", "subplot error"]
print(table.round(3))
```

```
                    df     sum_sq  mean_sq       F  PR(>F)
blocks            14.0  10204.244  728.875  35.605   0.000
recipe             2.0    135.089   67.544   3.300   0.039
whole-plot error  28.0   1198.467   42.802   2.091   0.002
temperature        5.0   2100.300  420.060  20.520   0.000
R x T             10.0    205.978   20.598   1.006   0.439
subplot error    210.0   4298.889   20.471     NaN     NaN
```

- `C(block):C(recipe)` is the batch. Its line is the whole-plot error.
- `anova_lm` divides every line by the subplot error: $F = 3.30$, $p = 0.039$ for recipe.

Recipe against batches

```
ms = table["mean_sq"]
F_R = ms["recipe"] / ms["whole-plot error"]
print(f"F_R = {F_R:.2f} on (2, 28) df,  p-value = {stats.f.sf(F_R, 2, 28):.3f}")
means = cake.groupby(["block", "recipe"]).angle.mean().reset_index()
rcbd = smf.ols("angle ~ C(block) + C(recipe)", data=means).fit()
print(anova_lm(rcbd).round(3))
```

```
F_R = 1.58 on (2, 28) df,  p-value = 0.224
             df    sum_sq  mean_sq       F  PR(>F)
C(block)   14.0  1700.707  121.479  17.029   0.000
C(recipe)   2.0    22.515   11.257   1.578   0.224
Residual   28.0   199.744    7.134     NaN     NaN
```

- $F_R = 67.5/42.8 = 1.58$, the randomized-block F of the 45 batch means.
- Dividing by the subplot error is right for temperature and for $R \times T$.

Each recipe has 90 cakes. How many replicates does it have?

Variance components

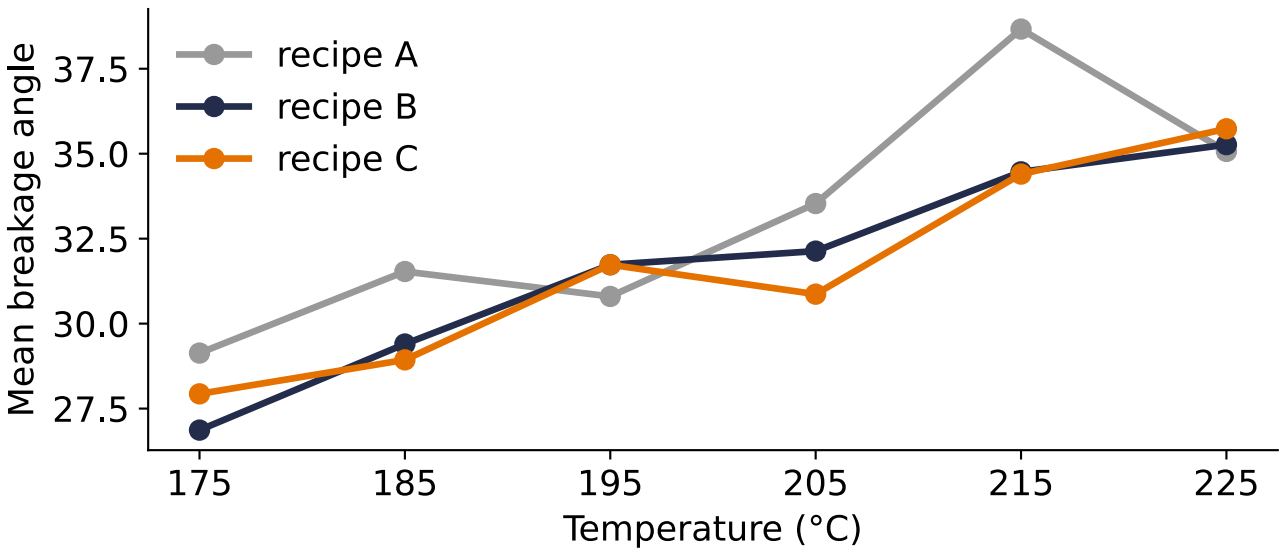
$$\hat{\sigma}^2 = \text{MS}_{\text{SP}}, \qquad \hat{\sigma}^2_W = \frac{\text{MS}_{\text{WP}} - \text{MS}_{\text{SP}}}{c}$$

```
msWP, msSP = ms["whole-plot error"], ms["subplot error"]
mixed = smf.mixedlm("angle ~ C(block) + C(recipe) * C(temp)", cake, groups="batch").fit()
print(f"by hand: {(msWP - msSP) / 6:.2f}  {msSP:.2f}")
print(f"mixedlm: {mixed.cov_re.iloc[0, 0]:.2f}  {mixed.scale:.2f}")
```

```
by hand: 3.72  20.47
mixedlm: 3.72  20.47
```

- $\hat{\sigma}_W = 1.93$ and $\hat{\sigma} = 4.52$.
- The block mean square is 728.9. Batches differ far more between blocks than within a block.

Recipe by temperature



- Each point is the mean of 15 cakes, one from each batch of the recipe.
- Two temperatures for the same recipe share all 15 batches. The standard error of their difference is $\sqrt{2\,\mathrm{MS}_{\mathrm{SP}}/b} = \sqrt{2(20.47)/15} = 1.65$.

What do these lines say about $R \times T$?

Plot code in the appendix.

Summary

	Recipe	Temperature
Experimental unit	a batch	a cake within a batch
Compared within	a block	a batch
Error line	whole-plot error, 28 df	subplot error, 210 df
Error mean square	42.8	20.5
F	1.58, $p = 0.22$	20.5, $p < 0.001$
SE of a difference	$\sqrt{2\,\mathrm{MS}_{\mathrm{WP}}/(bc)} = 0.98$	$\sqrt{2\,\mathrm{MS}_{\mathrm{SP}}/(ab)} = 0.95$

- The whole-plot factor is tested against the whole-plot error.
- The subplot factor and the interaction are tested against the subplot error.

Reading

- Owen, *A First Course in Experimental Design*: §§9.1 and 9.2.
- Meier, *ANOVA and Mixed Models*, Chapter 7 (online).
- Box, Hunter, and Hunter, *Statistics for Experimenters*, 2nd ed.: §9.2.
- Dean and Voss, *Design and Analysis of Experiments*, 2nd ed.: §§19.2 and 19.3.

Appendix

The data plots

```
angles = wide.drop(columns="batch mean")                          # batch by batch
angles.columns = angles.columns.astype(int)
bm = wide["batch mean"].unstack()
fig, axes = plt.subplots(1, 2, figsize=(10, 3.1), width_ratios=[1, 1.4])
axes[0].plot(angles.columns, angles.values.T, "-", color="0.6", lw=1)
axes[0].plot(angles.columns, angles.mean(), "o-", color=orange, lw=2.5, ms=6)
axes[0].set(xticks=angles.columns, xlabel="Temperature (°C)",
            ylabel="Breakage angle")
for (r, col), color in zip(bm.items(), ["0.6", blue, orange]):
    axes[1].plot(bm.index, col, "o-", lw=2, color=color, label=f"recipe {r}")
axes[1].set(xticks=bm.index, xlabel="Block", ylabel="Batch mean")
axes[1].legend(frameon=False, loc="upper right")
```

```
cells = cake.groupby(["temp", "recipe"]).angle.mean().unstack()   # recipe by temperature
fig, ax = plt.subplots(figsize=(7, 3.2))
for (r, col), color in zip(cells.items(), ["0.6", blue, orange]):
    ax.plot(cells.index, col, "o-", lw=2.5, ms=7, color=color, label=f"recipe {r}")
ax.set(xticks=cells.index, xlabel="Temperature (°C)", ylabel="Mean breakage angle")
ax.legend(frameon=False, loc="upper left")
```

The simulation

```
rng = np.random.default_rng(6160)
mu = cake.groupby(["recipe", "temp"]).angle.mean().unstack().values   # 3 x 6
beta = cake.groupby("block").angle.mean().values - cake.angle.mean()  # 15 blocks
sigma_W, sigma, R = 1.93, 4.52, 4000
def experiments(cakes_per_batch):              # R experiments, 15 blocks each
    batches = (R, 3, 15, 1) if cakes_per_batch == 6 else (R, 3, 15, 6)
    eta = rng.normal(0, sigma_W, batches)                # one per batch
    eps = rng.normal(0, sigma, (R, 3, 15, 6))            # one per cake
    return mu[None, :, None, :] + beta[None, None, :, None] + eta + eps
def differences(y):                            # A minus C, and 225 minus 175
    rec, temp = y.mean(axis=(2, 3)), y.mean(axis=(1, 2))
    return rec[:, 0] - rec[:, 2], temp[:, 5] - temp[:, 0]
true = (mu[0].mean() - mu[2].mean(), mu[:, 5].mean() - mu[:, 0].mean())
sims = [("Design 1", differences(experiments(1))), ("Design 2", differences(experiments(6)))]
lims, names = [(-3, 6), (3, 12)], ["A - C", "225 - 175"]
fig, axes = plt.subplots(2, 2, figsize=(10, 2.6), sharex="col")
for row, (label, diffs) in zip(axes, sims):
    for ax, d, t, lim, name in zip(row, diffs, true, lims, names):
        ax.hist(d, bins=40, range=lim, color=blue)
        ax.axvline(t, color=orange, linewidth=2)
        ax.set(yticks=[], xlim=lim, title=f"{label}: {name}")
```