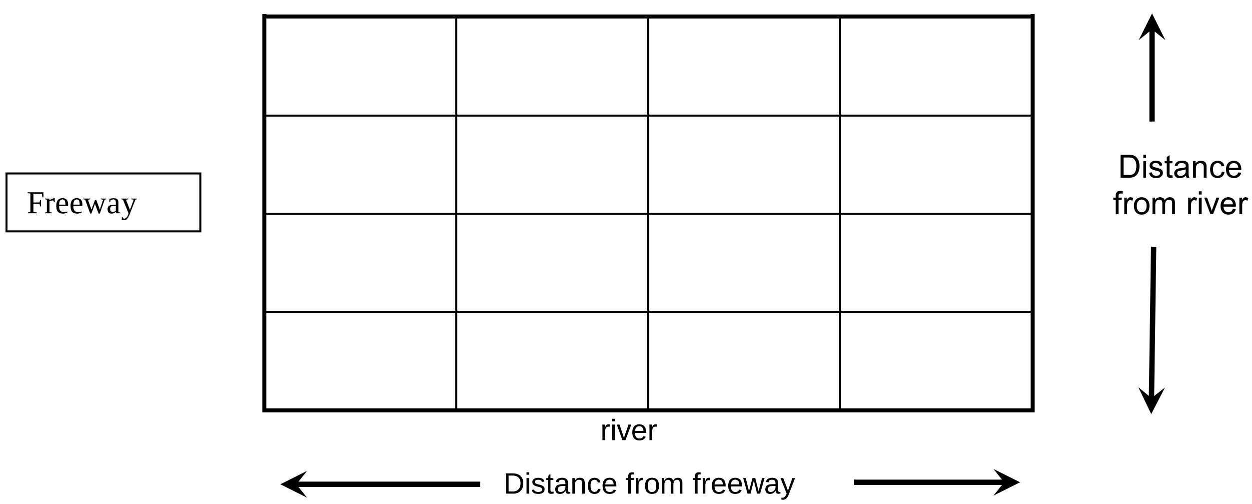
Stat 500 **Latin Square Example**

This is a study of four management treatments for increasing rare grass in a meadow (data file: brome.txt and sas file: brome.sas). Two blocking factors are considered in the design. One is the distance from a stream (river) and another is the distance from the freeway. The yield of rare grass is the response variable.

The data files for SAS are often organized in a way such that:
Each row contains one observation.
Data values are separated by one or more spaces.
There may be header line giving the variable names.

```
brome.txt

freeway stream trt yield
1 1 uncut 34.45
1 2 harvest 64.16
1 3 windrow 53.13
1 4 insitu 49.02
2 1 harvest 66.42
2 2 uncut 31.31
2 3 insitu 49.6
2 4 windrow 51.94
3 1 insitu 52.54
3 2 windrow 54.94
3 3 harvest 62.73
3 4 uncut 25.84
4 1 windrow 58.05
4 2 insitu 53.73
4 3 uncut 29.33
4 4 harvest 56.79
```



ANOVA table for brome data from a LS design:

Source	df	SS	MS	F	p-value
Row	3	108.96	36.32		
Column	3	3.02	1.005		
Treatment	3	2275.77	758.6	312	<0.0001
Error	6	14.60	2.43		

Efficiency: did Latin Square help? (compare to RCBD)

1) Did including columns as blocks help? Compare with RCBD with only the row blocks.

   $MSE_{LS} = 2.43$
   $MSE_{RCBD} = (SS_{col} + SS_{Error}) / (df_{col} + df_{Error}) = (3.02 + 14.60) / (6 + 3) = 1.96 < 2.43$
   (RCBD with only the row blocks is more precise!)

Relative Efficiency = $MSE_{RCBD} / MSE_{LS} = 1.96 / 2.43 = 0.81$
   RCBD requires 0.81 as many observations.

2) Did including rows as blocks help? Compare with RCBD with only the column blocks.

   $MSE_{LS} = 2.43$
   $MSE_{RCBD} = (SS_{Row} + SS_{Error}) / (df_{Row} + df_{Error}) = (108.96 + 14.60) / (6 + 3) = 13.73$
   (LS is more precise than the RCBD with only column blocks!)

   Relative Efficiency = $MSE_{RCBD} / MSE_{LS} = 13.73 / 2.43 = 5.65$
   RCBD requires 5.65 times as many observations to get the same precision as the Latin square design.

Note that we haven't talked about testing rows and columns here. Row blocks and column blocks are a way to reduce the variability.

In a Latin Square design, the experimental unit is the intersection of a row and column in a field (one tire of an auto), that is what is randomly assigned to a treatment. Neither rows nor columns were assigned to any treatment factor.

What do you give up with Latin Square designs?

1) Cannot look at the look at the interactions, e.g., row and treatment interaction
2) When t is small, very few df in error → multiple squares

If you have multiple Latin squares, the analysis depends on how the squares are arranged.

Example 1, Latin rectangle. 3 treatments (t = 3). 3 x 6 rectangle, in s=2 squares.

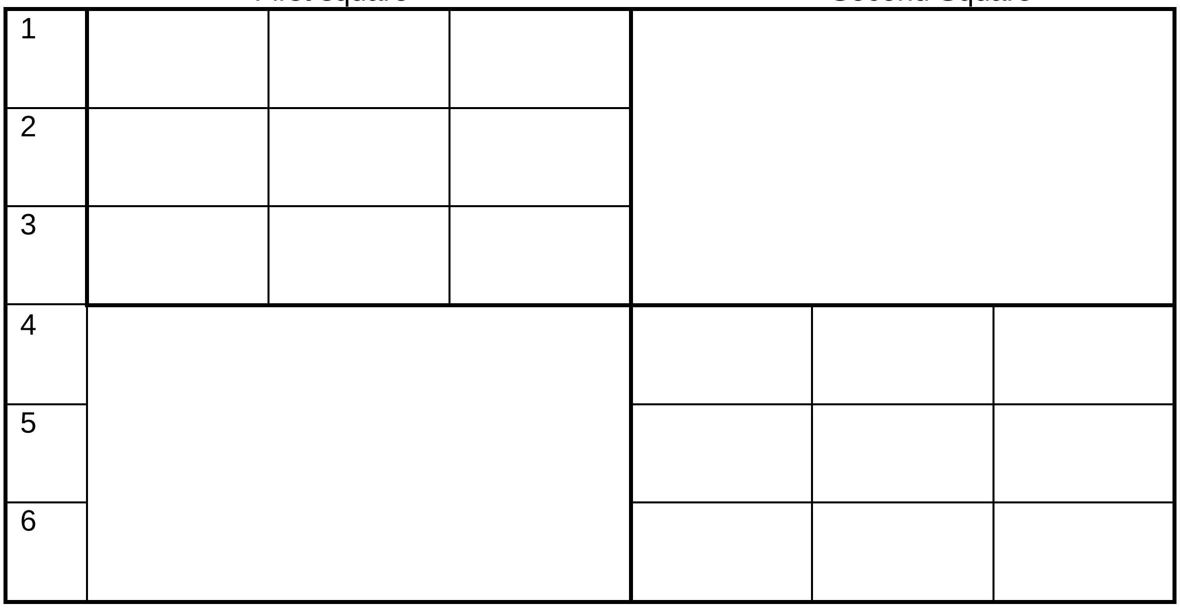
	First square			Second Square		
1						
2						
3						

ANOVA table for t rows, st columns and t treatments (s squares)

Source	df	df (example)
Row	t-1	2
Column	***st-1***	5
Treatment	t-1	2
Error	(t-1)(***st-2***)	8
Total	st^2-1	17

Example 2, multiple squares

First square | Second Square



Need to lay out the table, not just t rows and t columns. The analysis is trickier.

ANOVA table for the example 2 (with s=2 squares)

Source	df	df (example, s= 2, t = 3)
Square	s-1	1
Row (square)	s(t-1)	4
Column (square)	s(t-1)	4
Treatment	t-1	2
Error	(st-s-1)(t-1)	6
Total	st^2-1	17

Example: brome2.sas (separate squares)