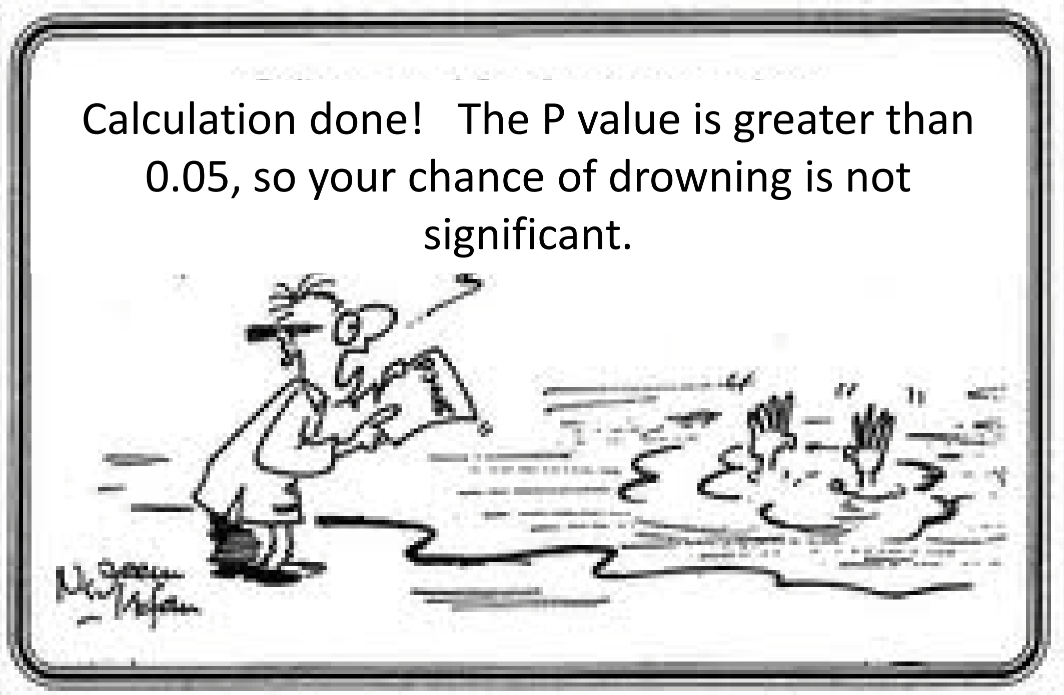
DESIGN OF EXPERIMENTS

- Where do DOEs fit in the Quality tool box?
- What goes into doing a simple DOE?



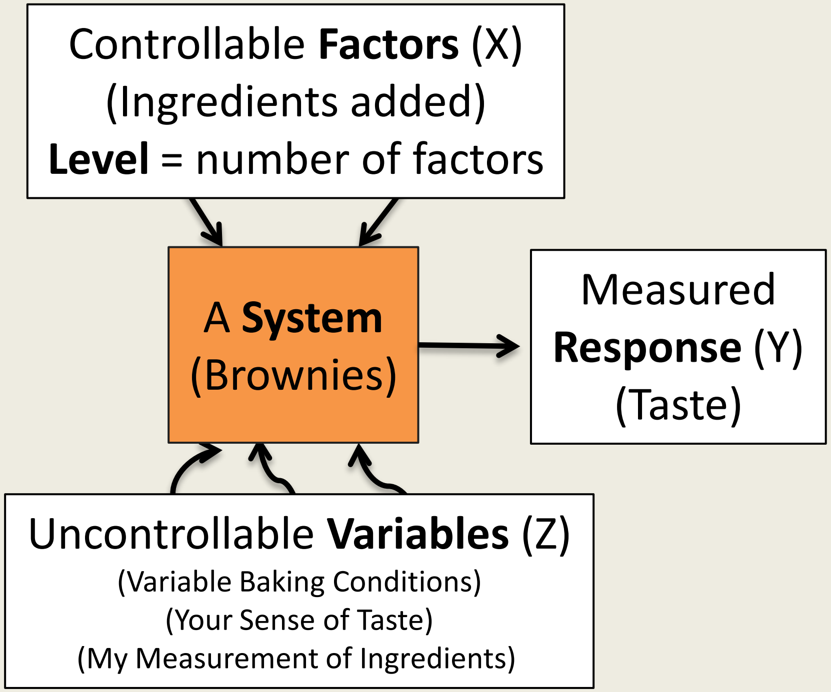
When to use a DOE?

- As a complement statistical process control.
  - SPC separates special causes from common causes.
  - Used to eliminate upsets.
- Need experiments to understand common causes and improve processes.
  - Can change one factor at a time (OFAT).
    - Not very efficient.
  - A DOE can probe several factors at once.
    - For 2 factors, a complete DOE requires just 4 runs.

Do you need a background it Stats?

- No.   But it sure would help!
- Modern statistics packages have greatly improved access to DOE methods.
- However, these must be used with:
  - An understanding of some basic statistical principles.
  - An appreciation for the limitations and pitfalls of DOE design and execution.

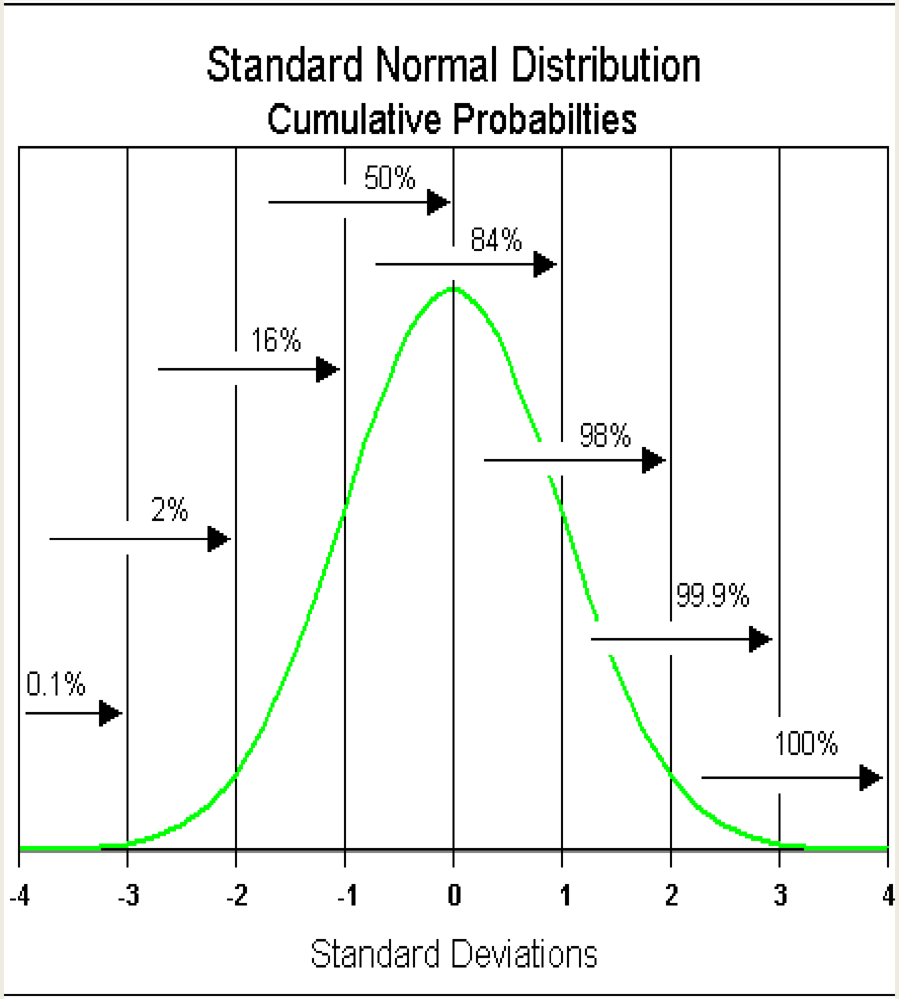
Some DOE Terms:
1. The System



In planning a DOE, the design will depend on:

- The number of factors and the number of levels in each.
- Whether the factors are categorical or numerical.
- How many experiments you can afford to do. (Partial Factorials)
- What you do to minimize uncontrolled variables. (Blocking)

DOE Terms: 2. This <u>is</u> Statistics



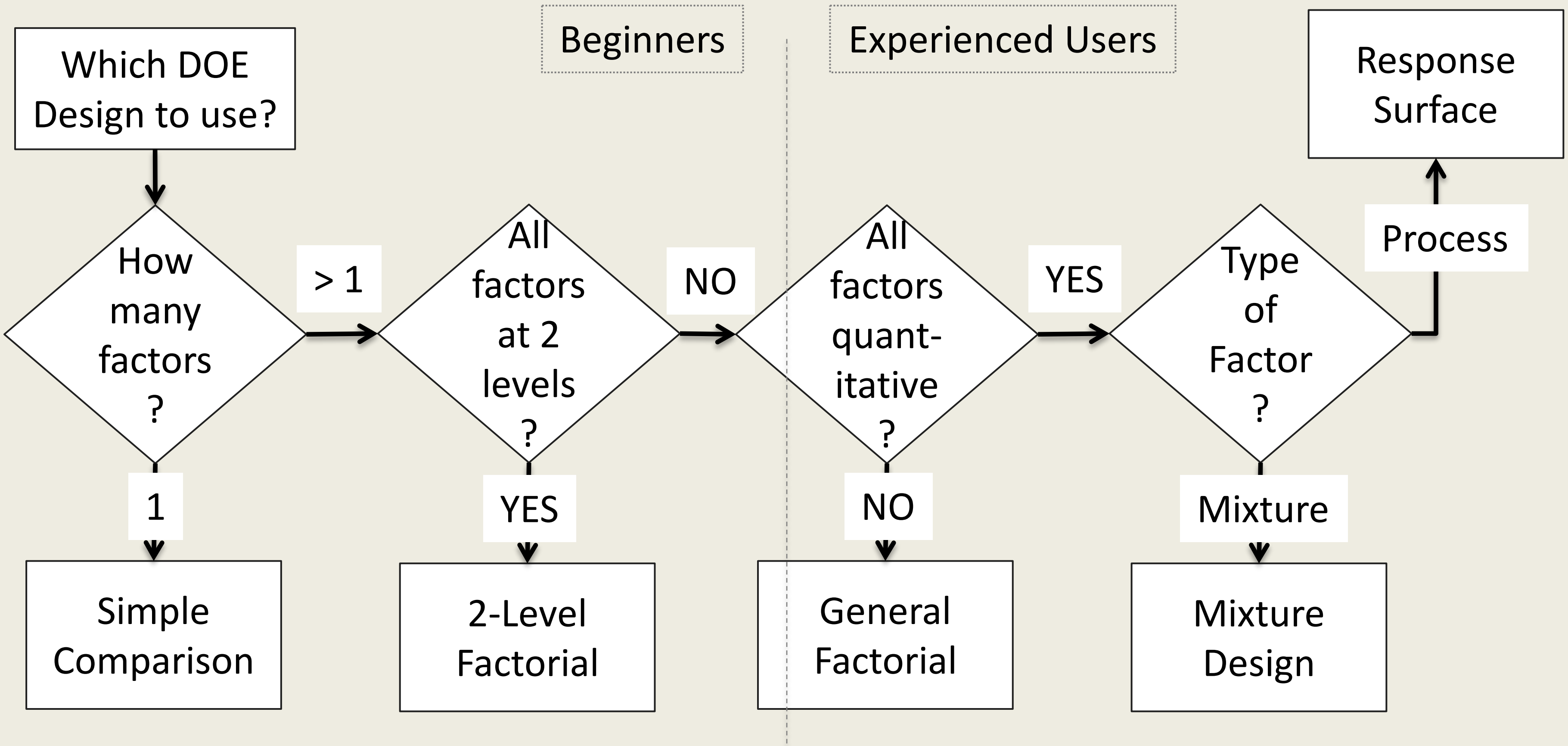
- Results always have a **confidence interval**.
  - Usually set at 95%.
  - Measured with the "F-test", ratio of V_{ind}/V_{pooled}.
  - Result: Prob > F, less than .05
- The calculations assume a **normal distribution** of data.
  - Must check that results show normal behavior.
- Must do enough runs to have **degrees of freedom** left over for the calculation of error.

Remember: V (the variance) equals σ^2 (or better, $\hat{s}^2$), the standard deviation of the sample

The Three Steps of a Simple DOE

- Design a Set of Experiments:
  - Decide on Factors, Levels.   (Tradeoff)
  - Use “blocking” to reduce the effect of random variations.
- Execution:
  - Experiments must follow the design’s random order.
  - Repeats must be of the entire experiment.
- Analysis:
  - ANOVA (analysis of variance)
  - Check results for normal behavior.

Designing a DOE: 1st start your stats program



Taken from "DOE Simplified: Practical Tools for Effective Experimentation", by Mark J. Anderson and Patrick J. Whitcomb, Productivity, Inc., Portland, OR (2000)

Simple Comparison: A One Factorial DOE

- The bowling team has one open slot; Pat, Mark and Shari have each tried out, and now the team captain must decide who is the best player.
- The experiment includes:
  - One factor (bowler)
  - Three levels (the players)
  - Six replicates (games)
  - One response (the score)

GAME	PAT	MARK	SHARI
1	160	165	166
2	150	180	158
3	140	170	145
4	167	185	161
5	157	195	151
6	148	175	156
MEAN	153.7	178.3	156.2
ST. DEV.	9.6	10.8	7.4

A One Factor DOE – 1. The Design

(Generated by the StatEase program)

STD	RUN	FACTOR 1 (BOWLER)	RESPONSE 1 SCORE(PINS)
1	9	Pat	
2	7	Pat	
3	2	Pat	
4	16	Pat	
5	8	Pat	
6	5	Pat	
7	6	Mark	
8	15	Mark	
9	4	Mark	
10	11	Mark	
11	14	Mark	
12	3	Mark	
13	10	Shari	
14	1	Shari	
15	17	Shari	
16	12	Shari	
17	18	Shari	
18	13	Shari	

→

STD	RUN	FACTOR 1 (BOWLER)	RESPONSE 1 SCORE(PINS)
14	1	Shari	
3	2	Pat	
12	3	Mark	
9	4	Mark	
6	5	Pat	
7	6	Mark	
2	7	Pat	
5	8	Pat	
1	9	Pat	
13	10	Shari	
10	11	Mark	
16	12	Shari	
18	13	Shari	
11	14	Mark	
8	15	Mark	
4	16	Pat	
15	17	Shari	
17	18	Shari	

One Factor DOE: 2. The Experiment

STD	RUN	FACTOR 1 (BOWLER)	RESPONSE 1 SCORE(PINS)
14	1	Shari	158
3	2	Pat	140
12	3	Mark	175
9	4	Mark	170
6	5	Pat	148
7	6	Mark	165
2	7	Pat	150
5	8	Pat	157
1	9	Pat	160
13	10	Shari	166
10	11	Mark	185
16	12	Shari	161
18	13	Shari	156
11	14	Mark	195
8	15	Mark	180
4	16	Pat	167
15	17	Shari	145
17	18	Shari	151

One Factor DOE: 3. The Analysis

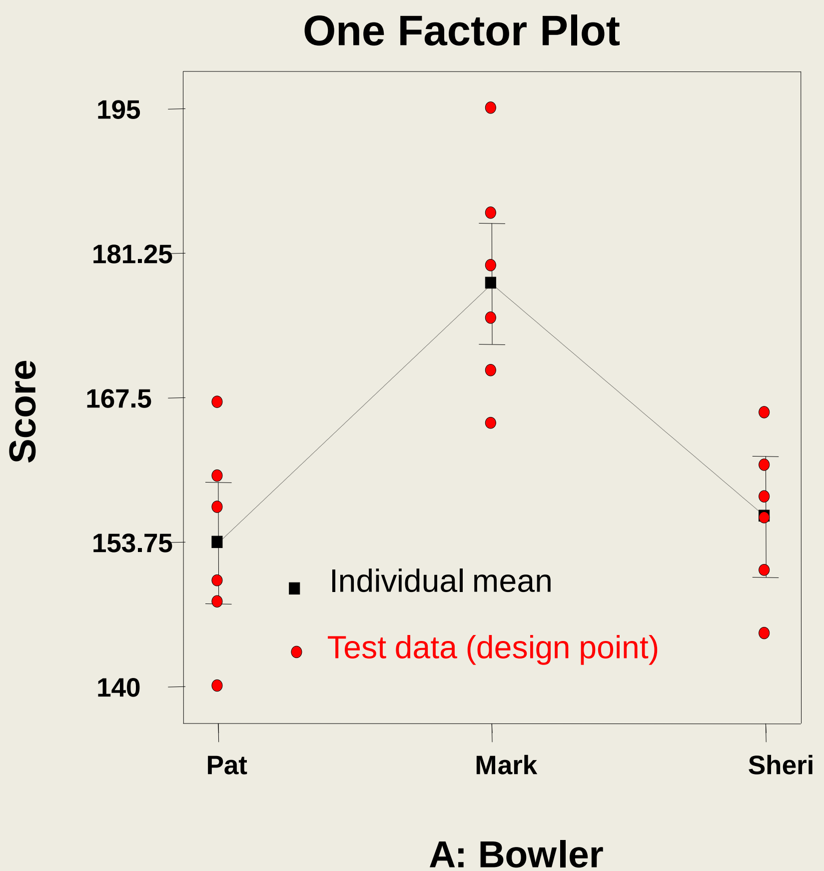
Source	Sum of Squares	Degrees of Freedom	Mean Square	F Value	Prob > F	
Model	2212	2	1106	12.6	0.0006	significant
A*	2212	2	1106	12.6	0.0006	significant
Pure Error	1320	15	88			
Corrected Total	3532	17				
* Factor (= bowler)						

- Looks good – The mean of at least one player’s score is significantly different from the rest.
  - The F value is >> 1 and the probability the result is due to random variation is very small: 0.06%
  - There are ample degrees of freedom used for calculating the error.
- But:
  - Who is the better player?
  - And, is the data valid – are the variations distributed normally?

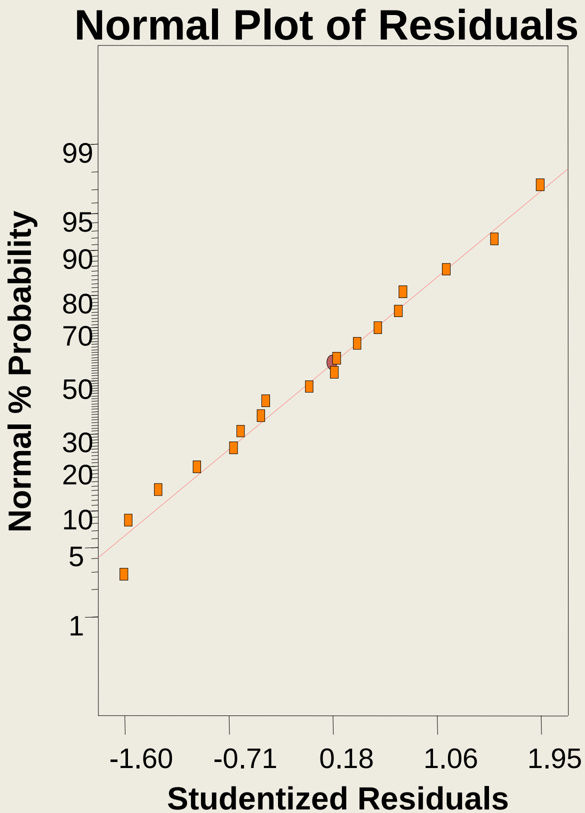
Further down the ANOVA output…..

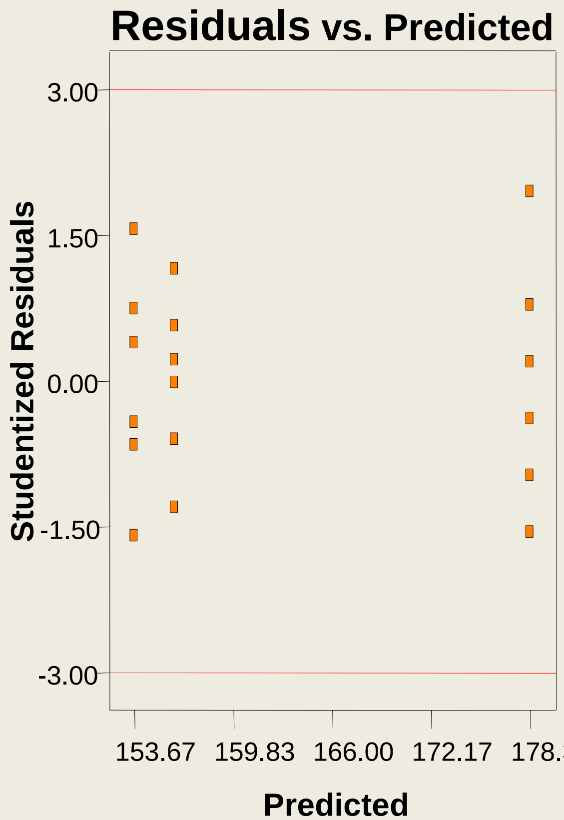
Treatment	Mean Difference	Degrees of Freedom	Standard Error	Coeff=0	Prob > |t|
Pat vs Mark	-24.7	1	5.4	-4.6	0.0004
Pat vs Sheri	-2.5	1	5.4	-0.5	0.6509
Mark vs Sheri	22.2	1	5.4	4.1	0.0010

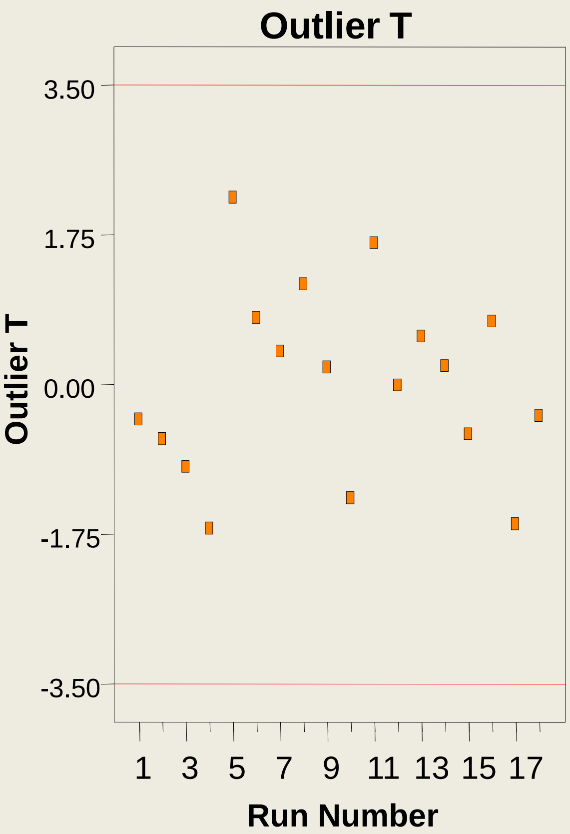
- Still need to know if any one of the bowlers is statistically better than the others.
  - Bars: 95% Confidence Intervals
  - Overlap: means are not significantly different.
  - Red points are test data.
- Still need to check that the data and residuals are normally distributed.



Checks for Normal Behavior



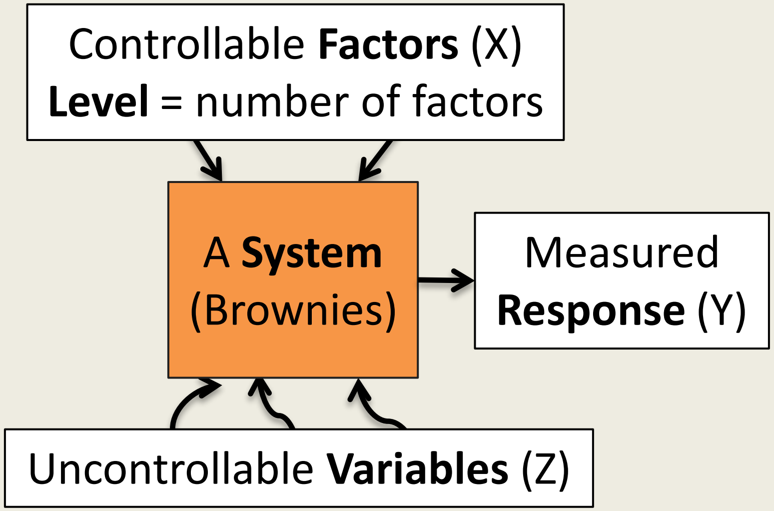




- Residuals: The difference between the data and the model-generated values at each data point.
  - They should follow a normal distribution
  - Their scatter should be independent of the model-generated data.
- There should not be outliers .

Brownies: The Ultimate Chocolate Experience
1. Planning the Experiment

- Design the experiment:
  - Factors: 1 : Flavoring
  - Levels: 4 : V, C, V+C & 0
  - Responses: 1 : Taste
  - Repeats: 9
  - What variables need control?
- Let's eat!
  - Draw up the Experimental Plan
  - Rules for the eaters



2. Doing the Experiment

- Testing Scale:
  - 1: Ordinary
  - 2: Not Bad
  - 3: Tasty
  - 4: Delicious!
- Tester Bias
- Other Testing Variables

Std	Run	Block	Factor 1 A:Flavor Enhancers	Response 1 Flavor 1(ordinary) to 4(delicious)
10	1	Block 1	B	
19	2	Block 1	C	
1	3	Block 1	A	
28	4	Block 1	D	
20	5	Block 2	C	
2	6	Block 2	A	
29	7	Block 2	D	
11	8	Block 2	B	
3	9	Block 3	A	
30	10	Block 3	D	
21	11	Block 3	C	
12	12	Block 3	B	
22	13	Block 4	C	
31	14	Block 4	D	

[redacted]

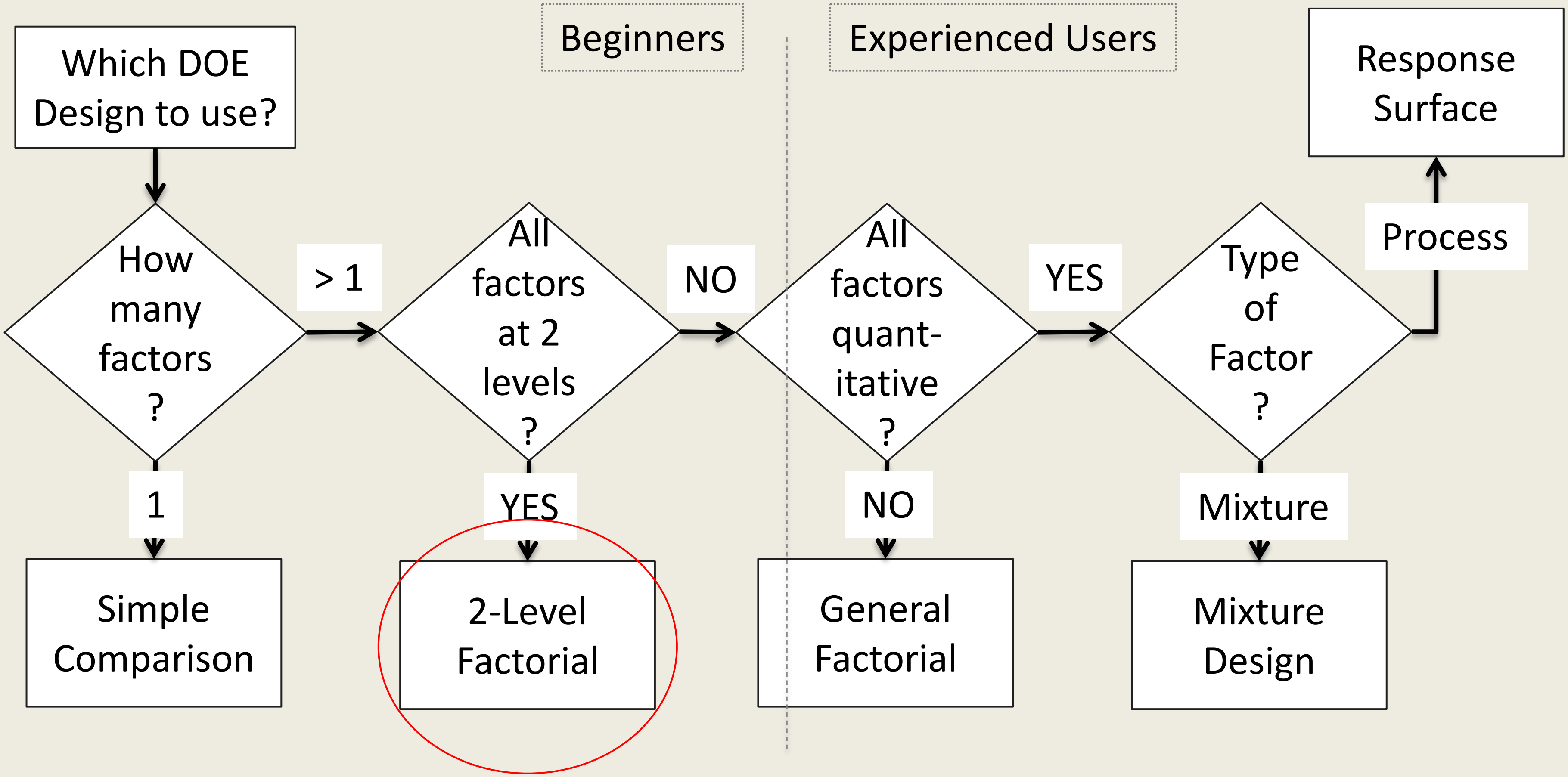
Std	Run	Block	Factor 1 A:Flavor Enhancers	Response 1 Flavor 1(ordinary) to 4(delicious)
4	16	Block 4	A	
32	17	Block 5	D	
23	18	Block 5	C	
5	19	Block 5	A	
14	20	Block 5	B	
24	21	Block 6	C	
33	22	Block 6	D	
15	23	Block 6	B	
6	24	Block 6	A	
25	25	Block 7	C	
16	26	Block 7	B	
7	27	Block 7	A	
34	28	Block 7	D	
8	29	Block 8	A	
35	30	Block 8	D	
17	31	Block 8	B	

[redacted]

The Ultimate Chocolate Experience: 3. Evaluation

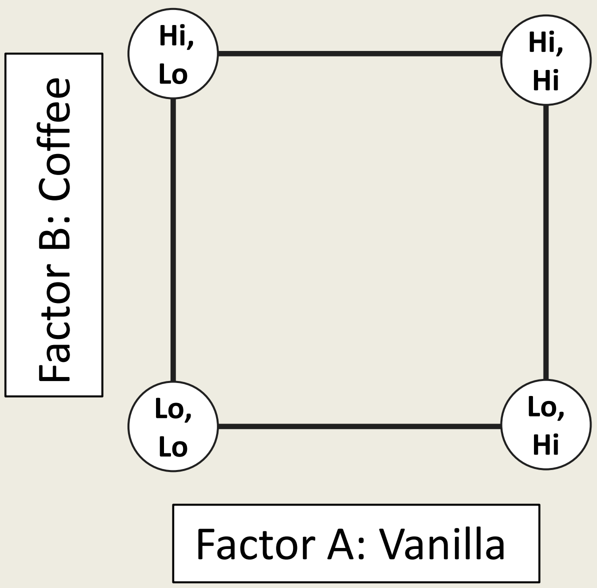
- Using JMP “jump”, a statistics program for business applications from SAS inc.
- Once data entered and numbers crunched:
  - View plot of Responses
  - Run ANOVA to find Means
  - Check normal behavior (Distrib. – Norm Q plot)
  - Compare Means: Student’s t: RESULTS!

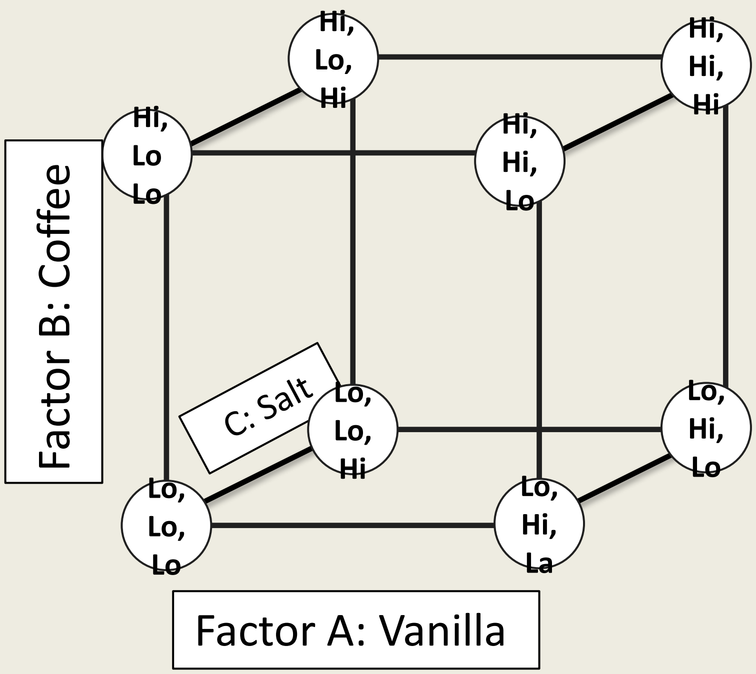
On Beyond One Factor:
What else might a beginner undertake?



2-Level Factorial Design (vs. OFAT)

- Reveals interactions between factors.
- Explores the full volume of experimental space.
- 3 Factors = 8 Expts, 4 = 16, 5 = 32.   (Fractional Factorial)



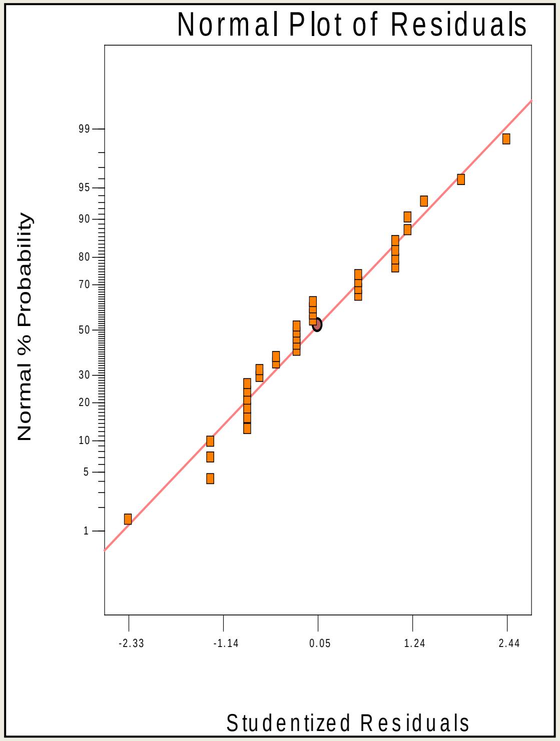


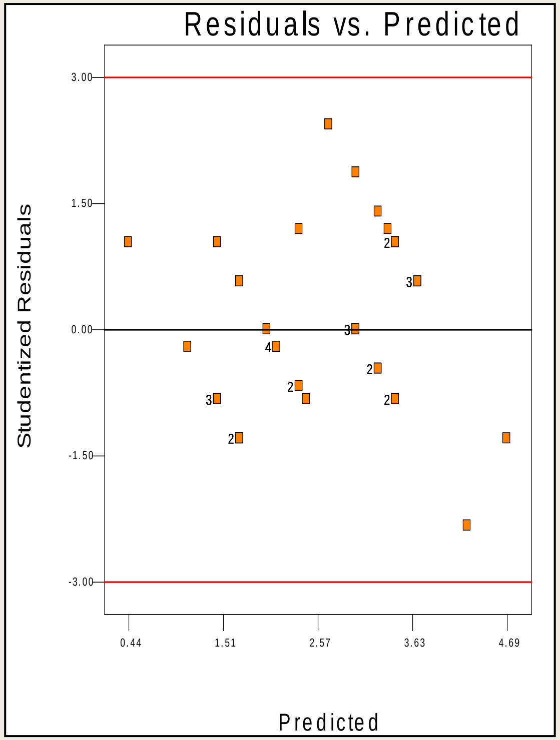
2-Level Factorial: Ultimate Brownie

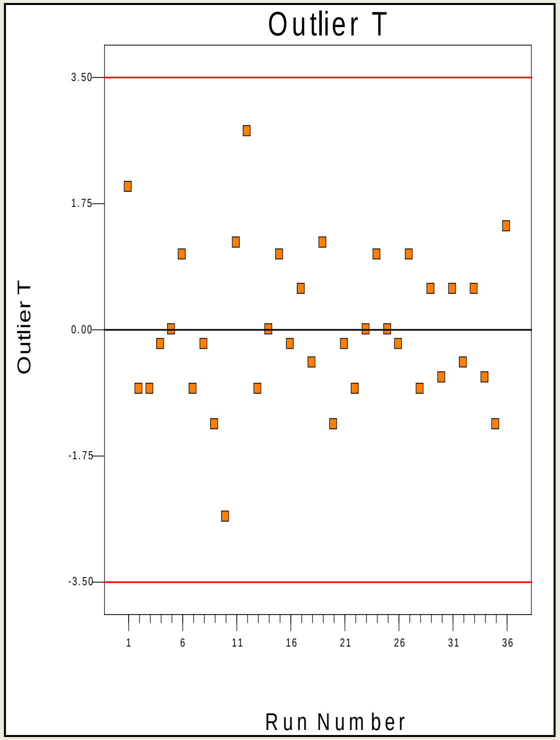
- This same data can be evaluated as a two-level factorial experiment:
  - Factors: Coffee and Vanilla (two factors)
  - Levels:
    - Coffee: none (=low) and 1 tsp instant powder (=high).
    - Vanilla: none (=low) and 1 tsp extract (=high).
  - An interaction between coffee and vanilla shows in the results.

2-Level Factorial ANOVA Results

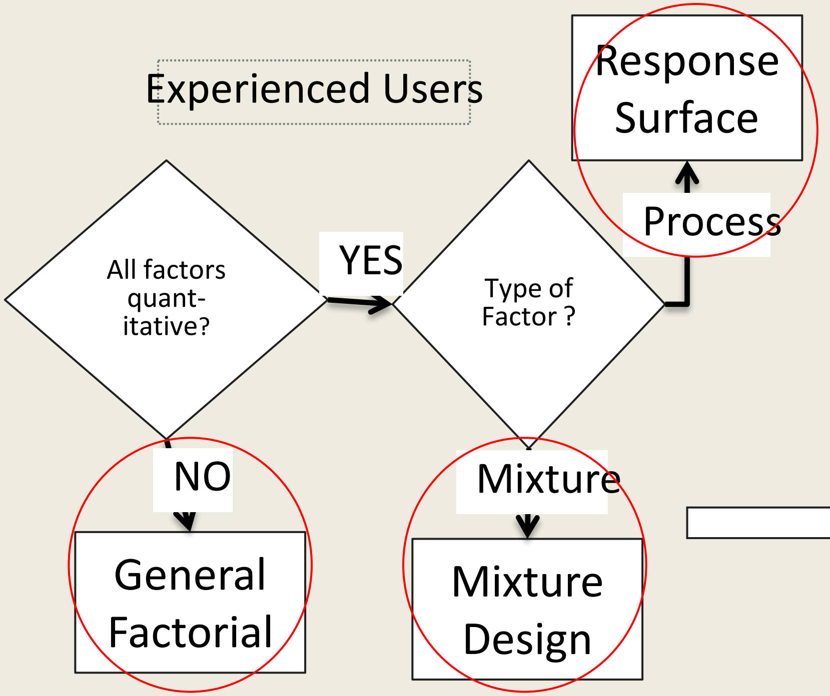
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	Comment
Block	10.56	8	1.32			
Model	21.67	3	7.22	16.77	<0.0001	Significant
Vanilla	18.78	1	18.78	43.61	<0.0001	Significant
Coffee	0.11	1	0.11	0.26	0.6161	Not Signifcant
Both	2.78	1	2.78	6.45	0.0180	Significant
Residual	10.33	24	0.43			
Corrected Total	42.56	35				

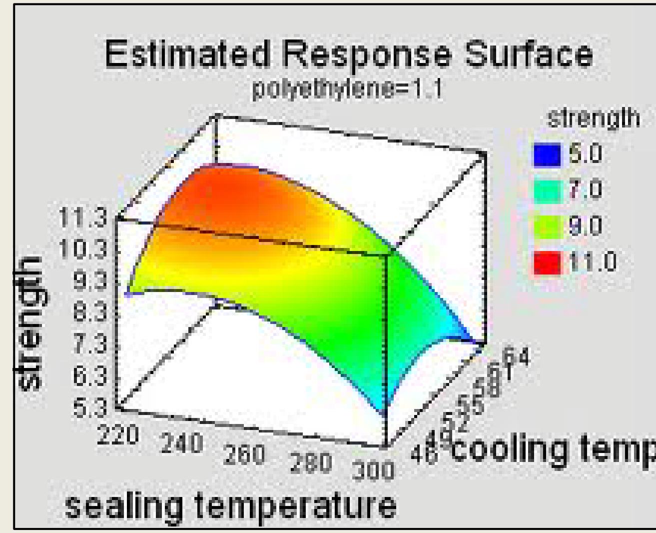


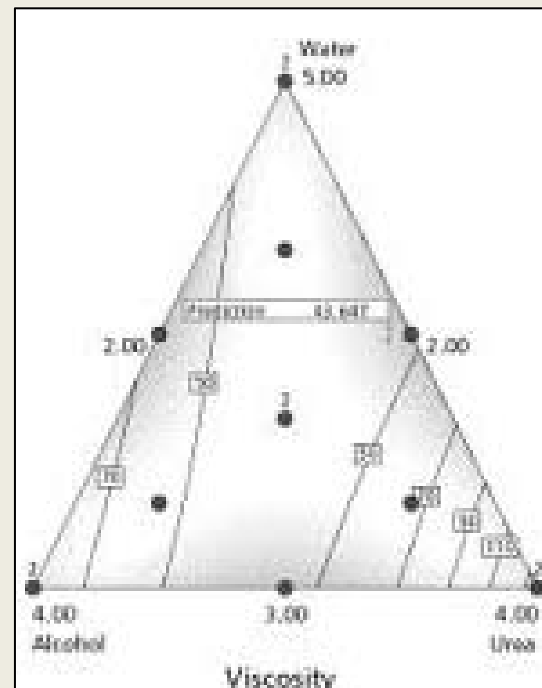




Further Possibilities







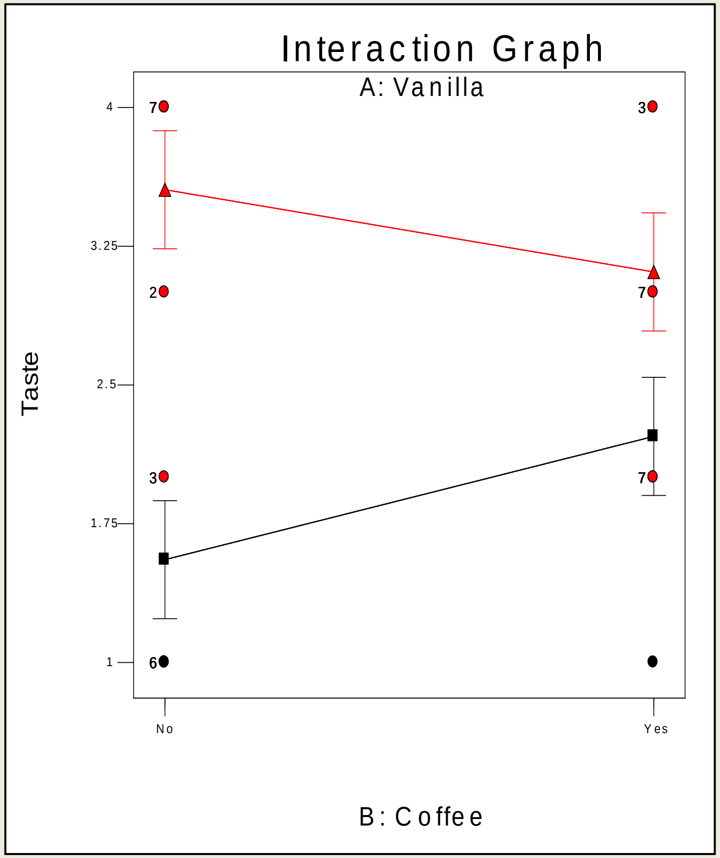
Questions?

Thank You, One and All!

Graphic Results

Shows how adding coffee effects the taste at both levels of vanilla.

- ? Does coffee increase vanilla's effects at low vanilla levels?



Shows how adding vanilla effects the taste at both levels of coffee.

- ? What is significant here?