SECTION 6: TWO-FACTOR ANOVA

- Factorial design involves two or more factors affecting the response variable and each factor has two or more levels
- When there are two factors affecting the response variable, it is known as two-factor ANOVA
- Just like one-factor ANOVA, there is only one variable being measured, that is, the response variable
- There are two factors being tested at the same time to determine whether they have an effect on the variable being measured
- For each of the two factors, there are two or more levels or treatments being applied to the individuals or experimental units being test
- Often the two factors are categorical
- Sometimes one or both of the two factors are quantitative, having several measureable levels or treatments
    - In this case, two-factor ANOVA is very similar to multiple linear regression
- Some of the advantages setting up one experiment to test the effects of two-factors on the response variable at the same time and analyzing the data with two-factor ANOVA, as opposed to doing two separate experiments for the effects of these two factors and analyzing them separately using one-factor ANOVA are as follows:
    1. Can test the effect of two factors at the same time, thus saving time and expenses
    2. Can compare the significance of the effects of the two factors
    3. Allows testing for interaction between the two factors
- **Main effects** = the effect of each factor considered separately
- **Interaction effect** = interaction between the two factors, that is, the effect of one factor on the response variable depends on the level of the other factor
    - The interaction effect may or may not occur

Fixed Effect Factors and Random Effect Factors

- **Fixed Effect Factors** are factors whereby the researcher deliberately fixes the levels of the factor because those are the levels of interest to him/her
- **Random Effect Factors** are factors for which the levels are selected or occur at random from a collection of possible levels

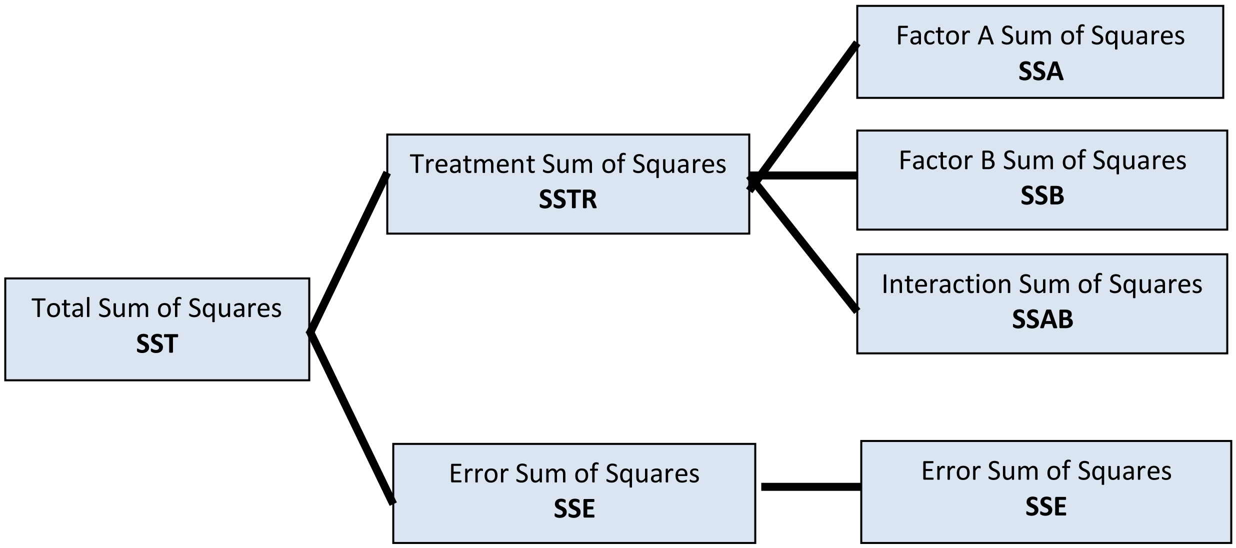
Assumptions of Two-Factor ANOVA

1. Random sampling
2. Independent observations: The observations of the response variable are independent of one another, though the levels of the factors do not need to be independent
3. Normal Distributions: For each combination of treatments, the response variable is normally distributed
4. Equal Standard deviations: The standard deviations of the response variable are the same for all combinations of treatments

Example

Factor A

		Level 1	level 2	level 3
Factor B	level 1	5 8 10 11	— — — —	— — — —
	level 2	— — — —	— — — —	— — — —



Response = Overall mean + A Main Effect + B Main Effect + AB Interaction Effect + Error

6.1 Two-Factor ANOVA with Replication, Balanced Data

- Sample size (i.e., number of observations) at least 2 for each combination of treatments
- **Balanced data** means that the sample size is the same for all combinations of treatments

Two-Factor ANOVA Identity for Sums of Squares (balanced data):

Total Sum of Squares = Factor A Sum of Squares + Factor B Sum of Squares
+ AB Interaction Sum of Squares + Error Sum of Squares

SST = SSA + SSB + SSAB + SSE

Two-Factor ANOVA Identity for Degrees of Freedom (balanced data):

df(SST) = df(SSA) + df(SSB) + df(SSAB) + df(SSE)

Or

$n - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1) + (n - ab)$

Two-Factor ANOVA Hypothesis Test (With Interaction) (Non-Additive Model)

Purpose: To perform hypothesis tests for the main effects and interaction effects of two factors
Assumptions: Given above

Null and Alternative Hypotheses

Overall Model: H_0: The overall model is not useful for making predictions
H_a: The overall model is useful for making predictions
Factor A main effect: H_0: There is no main effect due to Factor A
Ha: There is a main effect due to Factor A
Factor B main effect: H_0: There is no main effect due to Factor B
Ha: There is a main effect due to Factor B
AB interaction effect: H_0: The two factors do not interact
Ha: The two factors interact

ANOVA table for Two-Factor Analysis of Variance

Source of variation	SS	df	MS = SS/df	F-statistic
Overall model Corrected model	SSA +SSB +SSAB	df(A) +df(B) +dfAB) = ab – 1	$Corr\ MS = \frac{Corr\ SS}{Corr\ df}$	$F_{Overall} = \frac{Corr\ MS}{MSE}$
Factor *A*	*SSA*	*a* – 1	$MSA = \frac{SSA}{a-1}$	$F_A = \frac{MSA}{MSE}$
Factor *B*	*SSB*	*b* – 1	$MSB = \frac{SSB}{b-1}$	$F_B = \frac{MSB}{MSE}$
AB Interaction	*SSAB*	(*a* – 1)(*b* – 1)	$MSAB = \frac{SSAB}{(a-1)(b-1)}$	$F_{AB} = \frac{MSAB}{MSE}$
Error (within)	*SSE*	*n* – *ab*	$MSE = \frac{SSE}{n-ab}$	
Total	***SST***	***n* – 1**		

F (Overall model) = F (Corrected model) $= \frac{\text{Corrected } SS\,/\,(ab-1)}{\text{Error } SS\,/\,(n-ab)} = \frac{\text{Corrected } MS}{\text{MSE}}$

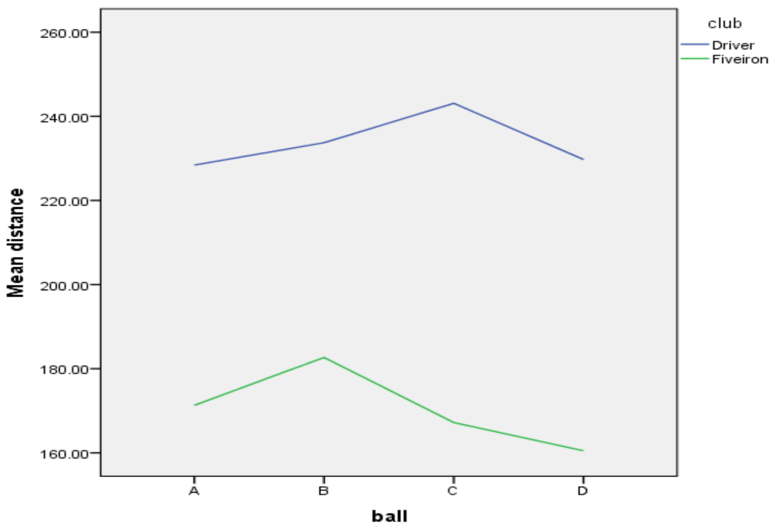
$$F_A = \frac{SSA\,/\,(a-1)}{SSE\,/\,(n-ab)} = \frac{MSA}{MSE} \qquad F_B = \frac{SSB\,/\,(b-1)}{SSE\,/\,(n-ab)} = \frac{MSB}{MSE} \qquad F_{AB} = \frac{SSAB\,/\,(a-1)(b-1)}{SSE\,/\,(n-ab)} = \frac{MSAB}{MSE}$$

Where: a = number of levels of Factor A
b = number of levels of Factor B
n = total number of observations
= a x b x (no. of replicates per combination of treatments)

Research Problem Involving Two-Factor ANOVA, Followed by Multiple Comparisons Tests
Example on Golf Balls

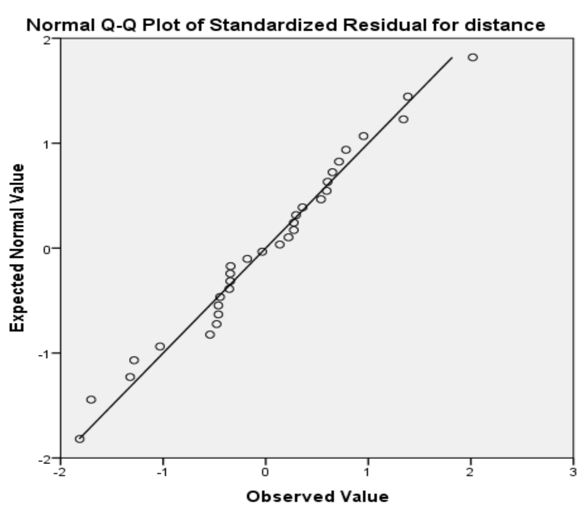
A golfer investigated the distances that golf balls are hit and how this is affected by the brand of ball used and the club used. This is a balanced design, with 4 replicates per combination of the two factors. Thus he was testing the following effects:

1. Effect of the four brands of balls (A, B, C and D)
2. Effect of two types of clubs (driver and five iron)
3. Effect of interaction between club and ball



Observe the following in this Line Chart:

1. The huge separation of the two lines indicates a virtually certain effect of the club factor, that is, a difference between the average of the means for the driver and the average of the means for the five iron.
2. The fact that the two lines go considerably up and down from A to B to C to D indicates a likely effect of the brand of the balls.
3. The non-parallel lines indicate an interaction effect between club and ball.
    - The segments from B to C run in completely different directions
    - They don't have to cross; if they are not parallel that indicates interaction



Checking for Normality
The above Q-Q Plot shows that the points are reasonably close to a straight line. Thus the distance variable is approximately normally distributed.

Tests of Between-Subjects Effects

Dependent Variable: distance

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	33652.830[a]	7	4807.547	140.689	.000
Intercept	1306819.028	1	1306819.028	38243.115	.000
club	32086.778	1	32086.778	938.996	.000
ball	801.348	3	267.116	7.817	.001
club * ball	764.703	3	254.901	7.459	.001
Error	820.113	24	34.171		
Total	1341291.970	32			
Corrected Total	34472.942	31			

a. R Squared = .976 (Adjusted R Squared = .969)

- Suppose the numbers highlighted in yellow are not given

>>>>>>>>>>

(a) At the 5% significance level, perform the most appropriate test to determine whether the overall model is significant.

H_0: All treatment combinations have equal means

H_A: At least 2 treatments have different means

Overall SS = SSA + SSB + SSAB

32086.778 + 801.348 + 764.703

= 33652.829

or Correct total SS − Error SS

df = 7   overall model df = df(A) + df(B) + df(AB) = ab − 1 = 3+3+[illegible] = 7

$F(\text{overall}) = F(\text{corrected}) = \dfrac{\text{corrected SS} / \text{corrected df}}{\text{Error SS} / \text{Error df}}$

$\dfrac{3362.830/7}{820.113/24} = 140.689$

df = (7, 24) $p < 0.001$ Extremely strong evidence $p < \alpha$ of 0.05 ∴ we reject H_0.

Conclusion: you get the idea.

(b) At the 5% significance level, perform the most appropriate test to determine whether there is a main effect of club type on mean distance.

H_0: There is no main effect on club type

H_A: there is a main effect on club type

F (main effect of club type)

$$F_A = \frac{SSA/(a-1)}{SSE/(n-ab)} = \frac{MSA}{MSE} = \frac{32086.778/(2-1)}{820.113 \quad /(32-(2)(4))} = 938.996$$

df = (1, 24) $p < 0.001$ since $p < \alpha$ of 0.001

At the 5% sig level, the data provide sufficient evidence that there is a significant main effect of club type, that is, the mean distance are not all the same for the different club types, averaging over all brands.

(c) At the 5% significance level, perform the most appropriate test to determine whether there is a main effect of ball brand on mean distance.

H_0: there is no effect on ball brand

H_A: There is an effect on ball brand

$$F_B = \frac{SSB/(b-1)}{SSE/(n-ab)} = \frac{MSB}{MSE} = \frac{801.348 \quad /(4-1)}{820.113 \quad /(32-(2)(4))} = 2.817$$

df = (3, 24) $p < 0.001$ Extremely strong evidence

Since $p < \alpha$ of 0.05 we reject H_0.

you get the idea at this point for the conclusion.

(d) At the 5% significance level, perform the most appropriate test to determine whether the effect of club type on mean distance depends on ball brand. (In other words, test whether there is an interaction effect between club type and ball brand.)

H_0: There is no interaction between club type and ball brand

H_A There is an interaction between club type and ball brand

F (interactions) $F_{AB} = \frac{SSAB/(a-1)(b-1)}{SSE/(n-ab)} = \frac{MSAB}{MSE} = \frac{769.703 \ /(2-1)(4-1)}{820.113 \quad /(32-(2)(4))}$

df (3, 24) $0.001 < p < 0.005$ very strong evidence $= 7.459$

Since $p < 0.05$ we reject H_0. against H_0

you get the idea at this point for the conclusion.

\>>>>>>>>>

Multiple Comparisons

Point Estimates + Confidence Intervals for each Combination of Club Type and Ball Brand

Estimates					
Dependent Variable: distance					
club	Ball	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Driver	A	228.425	2.923	222.393	234.457
	B	233.725	2.923	227.693	239.757
	C	243.100	2.923	237.068	249.132
	D	229.750	2.923	223.718	235.782
Fiveiron	A	171.300	2.923	165.268	177.332
	B	182.675	2.923	176.643	188.707
	C	167.200	2.923	161.168	173.232
	D	160.500	2.923	154.468	166.532

Pairwise Comparisons: For each ball brand, comparisons between club types

Pairwise Comparisons							
Dependent Variable: distance							
ball	(I) club	(J) club	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
A	Driver	Fiveiron	57.125*	4.133	.000	48.594	65.656
	Fiveiron	Driver	-57.125*	4.133	.000	-65.656	-48.594
B	Driver	Fiveiron	51.050*	4.133	.000	42.519	59.581
	Fiveiron	Driver	-51.050*	4.133	.000	-59.581	-42.519
C	Driver	Fiveiron	75.900*	4.133	.000	67.369	84.431
	Fiveiron	Driver	-75.900*	4.133	.000	-84.431	-67.369
D	Driver	Fiveiron	69.250*	4.133	.000	60.719	77.781
	Fiveiron	Driver	-69.250*	4.133	.000	-77.781	-60.719
Based on estimated marginal means							
*. The mean difference is significant at the .05 level.							
b. Adjustment for multiple comparisons: Bonferroni.							

- The point estimate for the multiple comparisons
  = Difference between each pair of means (SPSS calls this Mean Difference)
  = $(x_i - x_j)$
- For example, difference between the means for Ball A, Driver and Ball A, Fiveiron
  $=(x_i - x_j) = (228.425 - 171.300) = 57.125$ (shown in the table above)

Pairwise Comparisons for Interactions (Club*Ball)
(Shown in this table for each club type, comparisons between ball brands)

- For example, the first comparison is: (Driver*Ball A compared with Driver*Ball B)

Pairwise Comparisons

Dependent Variable: distance

club	(I) ball	(J) ball	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
Driver	A	B	-5.300	4.133	1.000	-17.184	6.584
		C	-14.675*	4.133	.010	-26.559	-2.791
		D	-1.325	4.133	1.000	-13.209	10.559
	B	A	5.300	4.133	1.000	-6.584	17.184
		C	-9.375	4.133	.196	-21.259	2.509
		D	3.975	4.133	1.000	-7.909	15.859
	C	A	14.675*	4.133	.010	2.791	26.559
		B	9.375	4.133	.196	-2.509	21.259
		D	13.350*	4.133	.021	1.466	25.234
	D	A	1.325	4.133	1.000	-10.559	13.209
		B	-3.975	4.133	1.000	-15.859	7.909
		C	-13.350*	4.133	.021	-25.234	-1.466
Fiveiron	A	B	-11.375	4.133	.067	-23.259	.509
		C	4.100	4.133	1.000	-7.784	15.984
		D	10.800	4.133	.092	-1.084	22.684
	B	A	11.375	4.133	.067	-.509	23.259
		C	15.475*	4.133	.006	3.591	27.359
		D	22.175*	4.133	.000	10.291	34.059
	C	A	-4.100	4.133	1.000	-15.984	7.784
		B	-15.475*	4.133	.006	-27.359	-3.591
		D	6.700	4.133	.709	-5.184	18.584
	D	A	-10.800	4.133	.092	-22.684	1.084
		B	-22.175*	4.133	.000	-34.059	-10.291
		C	-6.700	4.133	.709	-18.584	5.184

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

\>>>>>>>>>> → for Driver

Means Comparisons Diagrams (based on the Bonferroni Method)

A	D	B	C
228.425	229.250	233.725	243.100

you get the idea for conclusion

We can be 95% confident that, when using the Driver, the mean distance for ball C is different from A and D but no other difference between the rest.

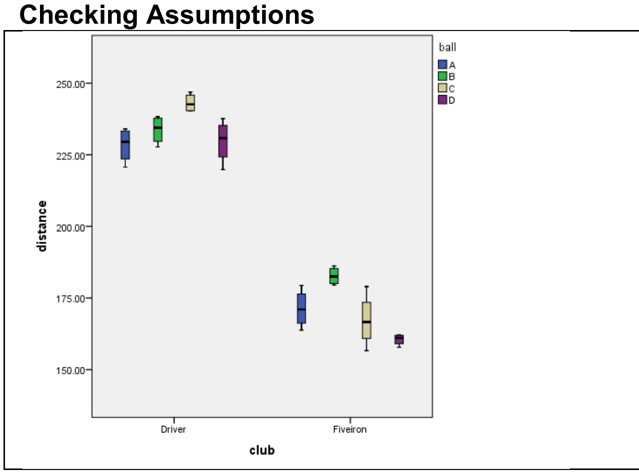
\>>>>>>>>>>

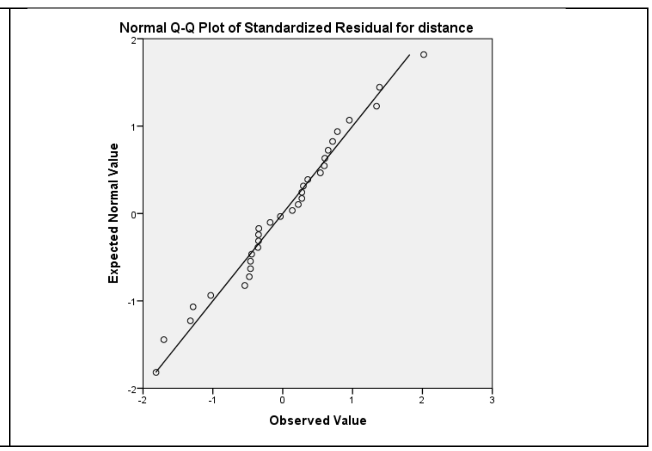
- The very different results of the multiple comparisons for Driver as opposed to Five Iron (since the ascending order of the means are different) also indicates interaction between club type and ball type
- These could be joined together to compare all 8 means in one diagram, but there would be complete separation between the means for Driver and means for Five Iron.

Means Comparisons for Interactions

- The multiple comparisons for interactions on the previous page does not show all of them because it just gives them separately for the two club types.
- Combinations = 4 x 2 = 8, so there are actually [8(8-1)]/2 = 28 multiple comparisons for interactions

Checking Assumptions





Levene's Test of Equality of Error Variances[a]

Dependent Variable: distance

F	df1	df2	Sig.
1.269	7	24	.307

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + club + ball + club * ball

Non-Additive and Additive Models in Two-Factor ANOVA

- Non-additive Model (includes interaction) (Full Model)
- Additive Model (does not include interaction) (Reduced Model)
- Often compared using the Extra Sum-of-Squares F-test

Previous Example on Golf Balls (Comparing Non-Additive and Additive Model)

A golfer investigated the distances that golf balls are hit and how this is affected by the brand of ball used and the club used and the interaction between ball brand and club type.

Consider the following two models

$$\text{Model 1}: \mu(\text{Distance} \mid \text{Club, Ball}) = \beta_0 + \text{Club} + \text{Ball}$$

[No interaction –Additive model >>> Reduced model]

$$\text{Model 2}: \mu(\text{Distance} \mid \text{Club, Ball}) = \beta_0 + \text{Club} + \text{Ball} + (\text{Club*Ball})$$

[With Interaction – Non-additive model >>> Full model]

Additive Model (No Interaction – Reduced Model)

Tests of Between-Subjects Effects

Dependent Variable: distance

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	32888.127	4	8222.03175	140.076	.000
club	32086.778	1	32086.778	546.652	.000
ball	801.348	3	267.116	4.551	.010
Error	1584.816	27	58.6969		
Corrected Total	34472.942	31			

a. R Squared = .976 (Adjusted R Squared = .969)

Non-Additive Model (With Interaction – Full Model)

Tests of Between-Subjects Effects

Dependent Variable: distance

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	33652.830[a]	7	4807.547	140.689	.000
Intercept	1306819.028	1	1306819.028	38243.115	.000
club	32086.778	1	32086.778	938.996	.000
ball	801.348	3	267.116	7.817	.001
club * ball	764.703	3	254.901	7.459	.001
Error	820.113	24	34.171		
Total	1341291.970	32			
Corrected Total	34472.942	31			

a. R Squared = .976 (Adjusted R Squared = .969)

Testing for An Interaction Effect by comparing an Additive and Non-additive Model

(a) At the 5% significance level, perform an Extra Sum-of-Squares F-test, using the additive and non-additive models presented in the tables above, in order to determine whether the effect of club type depends on ball brand (in other words, whether there is an interaction effect between club type and ball brand), after accounting for club type and ball brand.

>>>>>>>>>>

H_0: μ(Distance | club, ball) = β_0 + Club + Ball

H_A: μ(Distance | club, ball) = β_0 + club + Ball + (club * ball)

$$F = \frac{(SS_E(\text{reduced}) - SS_E(\text{full}))/(df_E(\text{reduced}) - df_E(\text{full}))}{SS_E(\text{full})/df_E(\text{full})}$$

$$= \frac{(1584.816 - 820.113)/(27-24)}{820.113/24} = \frac{254.901}{34.1714} = 7.459$$

df = (3, 24) $0.001 < p < 0.005$ very strong evidence against H_0

Since $p < 0.05$ we reject H_0.

Conclusion: you get the idea.

Compared with the approach on page 6.

(b) Now suppose the Interaction is ignored, use the additive model to determine whether either club type or ball brand have an effect on mean distance. Perform the test at the 5% significance level.

H_0: Neither factor has an effect on mean distance

H_A: At least one factor has an effect on mean distance

$$F = \frac{[32086.778 + 801.348]/[(2-1)+(4-1)]}{1584.816/27}$$

$$= \frac{32888.126/4}{1584.816/27} = 140.076$$

df = (4, 27) $p < 0.001$ Since $p < \alpha$ of 0.05 we reject H_0.

Extremely strong evidence against H_0.

At the 5% sig level, the data provide sufficient evidence to conclude that either club type or ball brand or both have an effect on mean distance.

>>>>>>>>>>

Compare the Results of the non-additive and additive models

	Full Model, OR Non-additive Model (Interaction)	**Reduced Model, OR Additive Model (No interaction)**
Overall Model	F = 140.689 Df = (7, 24) P-value = 0.00000000 $F_{0.001}$ = 5.23 (may be more significant than additive)	F = 140.076 Df = (4, 27) P-value = 0.00000000 $F_{0.001}$ = 6.33
Effect of Club type	F = 938.996 Df = (1, 24) P-value = 0.00000000	F = 546.652 Df = (1, 27) P-value = 0.00000000
Effect of Ball brand	F = 7.817 Df = (3, 24) P-value = 0.000824	F = 4.551 Df = (3, 27) P-value = 0.010478
Effect of interaction	F = 7.459 Df = (3, 24) P-value = 0.001074	**None**

Note: The overall model and the effect of club type and ball brand were all more significant with the non-additive model (with interaction) than the additive model. This is because, when the interaction term is significant (such as in this example), the interaction model is more accurate and more effective in showing significant effects.

6.2 Randomized Block Design

- Analyzed with Randomized Block Analysis of Variance (ANOVA)
- Considered as an extension of the paired design
- A special type of Two-Factor ANOVA where the “block” factor is not of interest to the researcher.
- One advantage is that it allows the researcher to eliminate the effect of the block factor so that it does not affect the real factor of interest.
- Thus it is much more powerful in testing the effect of the factor of interest than ordinary ANOVA

Blocks in space

- Suppose an experiment was conducted to determine whether there is a difference in the effectiveness of four new types of fertilizers (A, B, C, and D)
- However, there is a gradient of conditions in the test area due to a gradual slope towards a river
- Experimental area is divided into blocks such that it can be assumed that the conditions are homogenous within each block, even though conditions vary among blocks.
  - Thus it eliminates the effect of extraneous variables in space.
- In each block, each treatment is represented once.

Gradient in moisture &nutrients

Blocks (in space)				
1	C	A	D	B
2	B	D	A	C
3	B	C	D	A
4	D	A	B	C
5	A	C	D	B
	RIVER			

Blocks in time

- Eliminate the effect of time on the observations
  - E.g., If a researcher wants to compare 4 sites and he cannot take several measurements in all sites at the same time, he can take 1 measurement in each site every month
- This will eliminate the effect of temporal variation (e.g., seasonal variation or day-to-day variation) on the results.

Gradient in Time

	Abundance of birds			
Blocks (in time)	**Site A**	**Site B**	**Site C**	**Site D**
Oct	11	8	9	15
Nov	13	10	12	16
Jan	4	2	3	7
Feb	9	6	7	14
March	20	16	17	25

Applications in Medical Sciences, Education, Psychology, Etc.

- The before-and-after “treatment” can be extended to monitoring patients every few hours, once a month, etc., or subjects during some intervention program in education or psychology
- In this case, the blocks are the subjects or patients
- Greatly increases the power of the test in detecting responses of subjects/patients to treatments or programs

Randomized Block ANOVA

Extra Assumption (in addition to the other assumptions of two-way ANOVA:
There is no significant interaction between the main factor of interest (treatment) and the blocks factor (which is not of interest). This can be checked graphically with a line graph.

Null and alternative hypotheses for the factor of interest:
H_0: There is no significant difference between population means of the main factor of interest (treatment)
H_a: There is a significant difference between population means of the main factor of interest (treatment)

Null and alternative hypotheses for blocks (not always tested because not main interest):
H_0: There is no effect of the block factor.
Ha: There is an effect of the block factor

Calculations for Randomized Block ANOVA:

Sum of Squares	Defining formula
Treatment (SSTR)	$SSTR = \sum_{i=1}^{k} b(\overline{x}_{Ti} - \overline{x})^2$
Block (SSBL)	$SSBL = \sum_{i=1}^{b} k(\overline{x}_{Bj} - \overline{x})^2$
Total (SST)	$SST = \sum_{i,j} (x_{ij} - \overline{x})^2$
Error (SSE)	$SSE = SST - SSTR - SSBL$

Source of variation	SS	df	MS = SS/df	F-statistic
Treatment	SSTR	k – 1	MSTR = SSTR / (k – 1)	F = MSTR/MSE
Block	SSBL	b - 1	MSBL = SSBL / (b – 1)	F = MSBL/MSE
Error	SSE	(k – 1)(b – 1)	MSE = SSE / (k – 1)(b – 1)	
Total	SST	n - 1		

Note: Error df = (k – 1)(b – 1) = n – k – b + 1 = Total df – (Treatment df + Blocks df)

$$F_{Treatment} = \frac{SSTR/(k-1)}{SSE/(k-1)(b-1)} = \frac{MSTR}{MSE}$$

$$df = [(k-1),(k-1)(b-1)]$$

$$F_{Blocks} = \frac{SSBL/(b-1)}{SSE/(k-1)(b-1)} = \frac{MSBL}{MSE}$$

$$df = [(b-1),(k-1)(b-1)]$$

Where k = number of treatments, b = number of blocks, n = kb = total number of observations

Research Problem: An experiment was conducted to test the effectiveness of four types of fertilizers on eggplants (*Solanum melongena*). The test area was a low-lying area near a river. Eggplants are particularly sensitive to soil fertility and structure as well as soil moisture. Therefore the gradients in these factors towards the river would have affected the experiment. Thus, a randomized block design was used to eliminate the effect of these gradients.

Table of raw data, including treatment means, block means and grand mean (Units = kg/m²):

Block	Fertilizer Treatment				Block Means
	A	B	C	D	
1	2.7	2.8	2.9	2.8	$\bar{x}_{B1} = 2.8$
2	2.9	3.0	3.2	2.9	$\bar{x}_{B2} = 3.0$
3	3.0	3.1	3.3	3.0	$\bar{x}_{B3} = 3.1$
4	3.3	3.3	3.6	3.4	$\bar{x}_{B4} = 3.4$
5	3.3	3.4	3.5	3.4	$\bar{x}_{B5} = 3.4$
Treatment Means	$\bar{x}_{T1} = 3.04$	$\bar{x}_{T2} = 3.12$	$\bar{x}_{T3} = 3.30$	$\bar{x}_{T4} = 3.10$	$\bar{x} = 3.14$

Calculate Four Sums of Squares

Treatment SS between treatments

$$SSTR = \sum_{i=1}^{k} b(\bar{x}_{Ti} - \bar{x})^2 = 5(3.04-3.14)^2 + 5(3.12-3.14)^2 + 5(3.30-3.14)^2 + 5(3.10-3.14)^2 = 0.188$$

Block SS between blocks

$$SSBL = \sum_{i=1}^{b} k(\bar{x}_{Bj} - \bar{x})^2$$
$$= 4(2.8-3.14)^2 + 4(3.0-3.14)^2 + 4(3.1-3.14)^2 + 4(3.4-3.14)^2 + 4(3.4-3.14)^2 = 1.088$$

Total SS

$$SST = \sum_{i,j} (x_{ij} - \bar{x})^2 = (2.7-3.14)^2 + (2.8-3.14)^2 + (2.9-3.14)^2 + \ldots + (3.4-3.14)^2 = 1.308$$

Error SS

$$SSE = SST - SSTR - SSBL = 1.308 - 0.188 - 1.088 = 0.032$$

>>>>>>>>>>

Source of variation	SS	df	MS	F
Treatment	0.188	$k-1$ $4-1=3$	$0.188/3$ $=0.06267$	23.50
block	1.088	$b-1$ $5-1=4$	$1.088/4$ $=0.27200$	101.87
Error	0.032	$(k-1)(b-1)$ $3\cdot4=12$	$0.032/12$ $=0.00267$	
Total	1.308	$n=kb-1$ $(4\cdot5)-1=19$		

(a) At the 5% significance level, test whether there is a difference in mean eggplant yield between the four fertilizer treatments (the factor of interest).

H_0: There is no difference in mean eggplant yield between the 4 fertilizer
H_A: There is a difference in mean eggplant yield between the 4 fertilizer

$$F_{Treatment} = \frac{SSTR / k-1}{SSE / ((k-1)(b-1))} = \frac{0.188/(4-1)}{0.032/3 \cdot 4} = \frac{0.06267}{0.00267} \approx 23.50$$

df(3,12) $p < 0.001$ extremely strong evidence against H_0

Since $p < \alpha$ of 0.05 we reject H_0.

Conclusion you get the idea. only compared between fertilizer

(b) At the 5% significance level, test whether there is a difference in mean eggplant yield between blocks (in space), which is not the factor of interest.

H_0: There is no effect on the block factor on mean eggplant yield
H_A: There is a effect on the block factor on mean eggplant yield

$$F_{block} = \frac{SSBL / b-1}{SSE / ((k-1)(b-1))} = \frac{1.088/5-1}{0.032/3 \cdot 4} = \frac{0.27200}{0.00267} \approx 101.87$$

df(4,12) $p < 0.001$ extremely strong evidence against H_0

Since $p < \alpha$ of 0.05 we reject H_0. P-value $= 3.71 \times 10^{-9}$

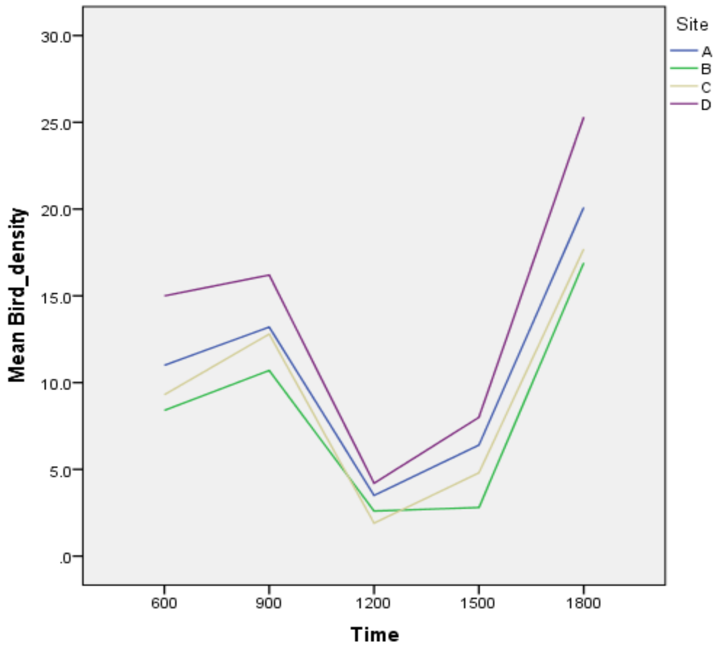
Conclusion you get the idea. only compared between blocks

>>>>>>>>>>

Demonstrate what would happen if we IGNORE the Blocks Factor and analyze this data set with Single-Factor ANOVA (Use Excel)

Result: F = 0.845 **, df =** 3, 16 **, P =** 0.46501

Example of Applying Randomized Block ANOVA to Blocks in Time
As part of planning wildlife conservation strategies, a group of ecologists wanted to determine whether there was a significant difference in bird density between four sites (A, B, C, and D). They chose a randomized block design where blocks were times of the day and they recorded bird density simultaneously at the four sites at 5 times of the day (600 hrs, 900 hrs, 1200 hrs, 1500 hrs, and 1800 hrs). Make use of the line graph and ANOVA table with missing values to answer the questions below.



Note:

1. If a block design was not used, the great variation in time of day would hide any differences between sites.
2. The lines are almost parallel, indicating little or no interaction between subject and test.

Tests of Between-Subjects Effects

Dependent Variable: Bird_density

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	798.279[a]	7	114.040	80.267	.000
Intercept	2221.832	1	2221.832	1563.844	.000
Site	84.876	3	28.292	19.913	.000
Time	713.403	4	178.351	125.533	.000
Error	17.049	12	1.421		
Total	3037.160	20			
Corrected Total	815.328	19			

a. R Squared = .979 (Adjusted R Squared = .967)

(a) Does the line graph above indicate interaction? Explain your answer.

The lines are more or less parallel, which means that there is no interaction between the factor of interest (site) and blocks (time of the day), which is not of interest, in their effect on bird density.

(b) At the 5% significance level, test whether there is a difference in mean bird density between the four sites (the factor of interest).

H_0: There is no difference in mean bird density between the four sites
Ha: There is a difference in mean bird density between sites (means of at least two sites are different).

$$F_{Treatment} = \frac{SSTR/(k-1)}{SSE/(k-1)(b-1)} = \frac{84.876/(4-1)}{17.049/(4-1)(5-1)} = \frac{28.292}{1.421} = 19.913$$

$$df = [(k-1),(k-1)(b-1)] = [(4-1),(4-1)(5-1)] = (3,12)$$

P < 0.001 Since P < α (0.05), reject Ho since there is very strong evidence.

Conclusion: At the 5% significance level, the data provide sufficient evidence to conclude that there is a difference in mean bird density between sites (the means of at least two sites are different).

(c) At the 5% significance level, test whether there is a difference in mean bird density between blocks (time of day), which is not the factor of interest.

H_0: There is no effect of the block factor (time of day).
Ha: There is an effect of the block factor

$$F_{Blocks} = \frac{SSBL/(b-1)}{SSE/(k-1)(b-1)} = \frac{MSBL}{MSE} = \frac{713.403/(5-1)}{17.049/(4-1)(5-1)} = \frac{178.351}{1.421} = 125.533$$

$$df = [(b-1),(k-1)(b-1)] = [(5-1),(4-1)(5-1)] = (4,12)$$

P < 0.001, Since P < α (0.05), reject Ho.

Conclusion: At the 5% significance level, the data provide sufficient evidence to conclude that there is a significant effect of blocks (time of day) on bird density.

Note: The block effect is probably much greater than the effect of site which can be seen in the Line Graph.

(d) Judging by the line graph and the ANOVA output do you think the same conclusion would have been reached if a completely randomized design (analyzed with one-way ANOVA) had been applied by the researchers instead of the randomized block design. Explain the logic of your answer.

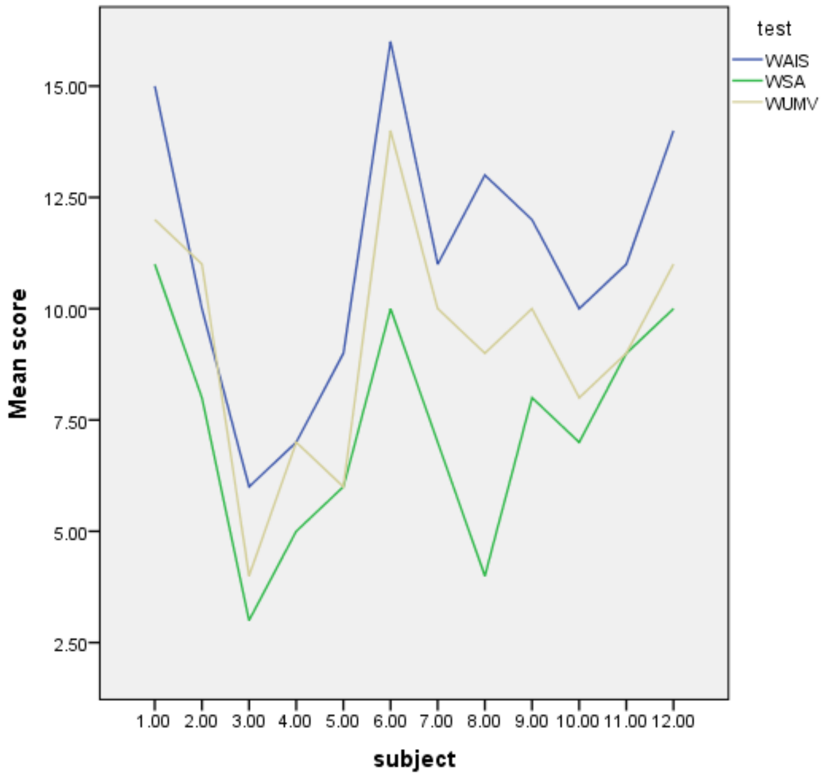
The line graph shows a huge variation in bird density from one time of the day to another. In fact, this variation is much greater than the variation in bird density between sites. Therefore, if the data were analyzed with completely randomized one-way ANOVA it would likely have shown no difference between sites because this difference would have been overshadowed by the difference over time. With such a research design, randomized block ANOVA is much more powerful than One-Way ANOVA.

Example on Comparing Language Test Scores
A random sample of 12 subjects were given three different types of language tests as follows:

1. WAIS vocabulary (linguistic)
2. Willner Unusual Meanings Vocabulary (WUMV) pragmatic
3. Willner-Sheerer Analyogy Test (WSA) pragmatic

Since people have very different linguistic ability, in order to account for, or eliminate, the subjects' abilities, the randomized block design was selected. Therefore the same 12 subjects took all three language tests. At the 5% significance level, test whether there is any difference in mean scores attained on the three language tests.

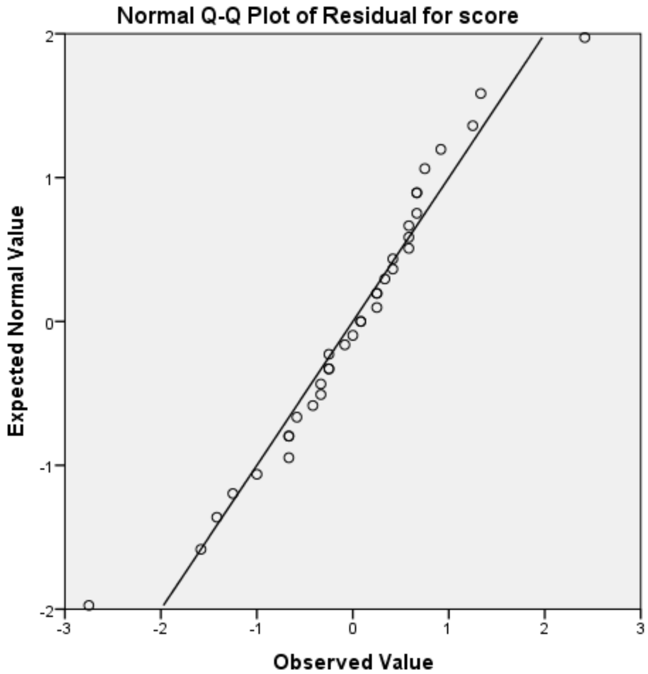
Line Graph



Note:

1. If a block design was not used, the great variation from one subject to the other would hide any differences between language tests.
2. The lines are almost parallel, indicating little or no interaction between subject and test.

Q-Q Plot to examine the assumption of normality



Note: the data points fall roughly along a straight line, indicating that the data are approximately normally distributed.

Randomized Block ANOVA (SPSS Output)

Tests of Between-Subjects Effects

Dependent Variable: score

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	310.250[a]	13	23.865	17.214	.000
Intercept	3080.250	1	3080.250	2221.820	.000
test	88.167	2	44.083	31.798	.000
subject	222.083	11	20.189	14.563	.000
Error	30.500	22	1.386		
Total	3421.000	36			
Corrected Total	340.750	35			

a. R Squared = .910 (Adjusted R Squared = .858)

Test for a Difference in Mean Score Between the Three Language Tests
H_0: There is no difference in mean score between the three language tests
Ha: There is a difference in mean score between the three language tests

F (Treatment)

$$F_{Treatment} = \frac{SSTR/(k-1)}{SSE/(k-1)(b-1)} = \frac{MSTR}{MSE} = \frac{88.167/(3-1)}{30.500/(3-1)(12-1)} = \frac{44.0835}{1.3864} = 31.798$$

$$df = [(k-1),(k-1)(b-1)] = [(3-1),(3-1)(12-1)] = (2,22)$$

$P < 0.005$ Since $P < \alpha$ (0.05), reject Ho with very strong evidence

Conclusion: The data provide sufficient evidence that there is a significant difference in mean score between the three language tests

Test for Effect of Blocks
H_0: There is no effect of the block factor.
Ha: There is an effect of the block factor

F (Blocks)

$$F_{Blocks} = \frac{SSBL/(b-1)}{SSE/(k-1)(b-1)} = \frac{MSBL}{MSE} = \frac{222.083/(12-1)}{30.500/(3-1)(12-1)} = \frac{20.189}{1.3864} = 14.563$$

$$df = [(b-1),(k-1)(b-1)] = [(12-1),(3-1)(12-1)] = (11,22)$$

$P < 0.005$ Since $P < \alpha$ (0.05), reject Ho with very strong evidence

Conclusion: The data provides sufficient evidence that there is a significant effect of blocks.

Note: The block effect is probably much greater than the effect of language score (even though the F-statistic is lower) due to the higher numerator df. Also that can be seen in the Line Graph.

Multiple Comparisons

Dependent Variable: score
Tukey HSD

(I) test	(J) test	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
WAIS	WSA	3.8333*	.48069	.000	2.6258	5.0409
	WUMV	1.9167*	.48069	.002	.7091	3.1242
WSA	WAIS	-3.8333*	.48069	.000	-5.0409	-2.6258
	WUMV	-1.9167*	.48069	.002	-3.1242	-.7091
WUMV	WAIS	-1.9167*	.48069	.002	-3.1242	-.7091
	WSA	1.9167*	.48069	.002	.7091	3.1242

Based on observed means.
The error term is Mean Square(Error) = 1.386.
*. The mean difference is significant at the 0.05 level.

Conclusion: All pairwise comparisons show significant differences.