Lab #15: Final Exam Review Key

1) When the F ratio is computed and there are more than two groups, it is called an omnibus F or overall F ratio. If this is significant, then additional analyses are needed to localize or pinpoint the effect. One way of doing these additional analyses are called post hoc comparisons.

2) **a.** 4 **b.** 3 **c.** 1 **d.** 16

3) The factorial design has three advantages:
1. It gives you more information for the same amount of work
2. Has increased control and generality.
3. Allows us to get at interactions.

4) **a.** Between Groups, Within Groups, & Total.
b. The F ratio is formed by dividing the between groups variance by the within groups variance. If the ratio is larger than 1, it suggests a treatment effect. If the ratio is less than or equal to 1, there are no differences between groups.

5) There are 5: Main effect A, Main of B, Interaction of AxB, Within Groups, & Total.

6) Main Effect – a difference in population means for a factor collapsed over the levels of all other factors in the design.
Interaction - occurs when the effect on one factor is not the same at the levels of another.

7) B

8) D

9) **1. Research Question**
Does music influence worker productivity?

2. Hypotheses

	Symbols	Words
H_O	$\mu_1=\mu_2=\mu_3$	Music does not affect worker productivity.
H_A	Not H_o	Music does affect worker productivity.

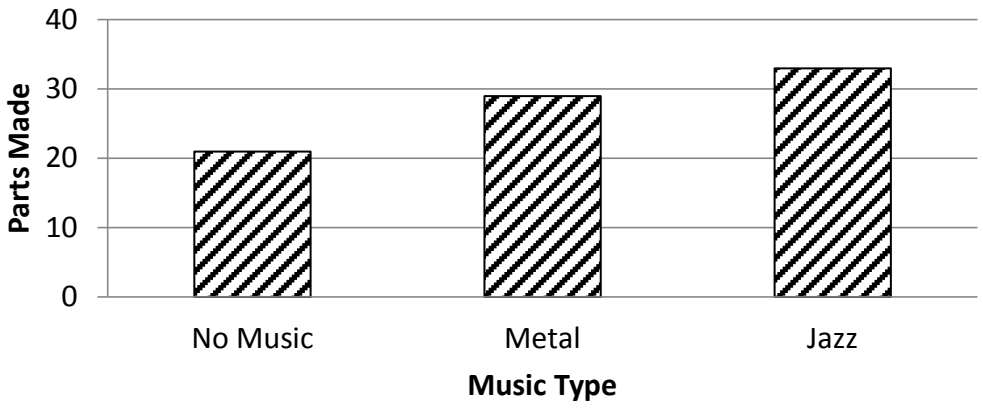
3. Assumptions
1) H_o
2) Subjects are sampled randomly.
3) Population distribution of the DV is normal in shape.
4) Groups are independent.
5) Population variances are homogenous.

4. Decision rules

Given 3 groups with 3, 3, and 3 subjects, respectively, we have (3-1) 2 df for the between groups variance estimate and (2+2+2) 6 df for the within groups variance estimate. Alpha level of .05, F_{crit}=5.14. If $F_{obs} \geq F_{crit}$, reject H_o, otherwise do not reject H_o.

5. Computation

	None	X^2	Metal	X^2	Jazz	X^2	
	21	441	27	729	31	961	
	25	625	30	900	30	900	
	16	256	29	841	38	1444	
T_j	62		86		99		247 = **T**
n_j	3		3		3		9 = **N**
$\bar{x}_j$	20.67		28.67		33.00		
$\sum_{i=1}^{n_j} X_{ij}^2$		1322		2470		3305	7097 = **II**
$\frac{T_j^2}{n_j}$	1281.33		2465.33		3267.00		7013.67 = **III**



The relevant quantities:

I. $\frac{T^2}{N} = \frac{247^2}{9} = \frac{61009}{9} = 6778.781$

II. $\sum\left(\sum_{i=1}^{n_j} X_{ij}^2\right) = 7097.00$

III. $\sum\left(\frac{T_j^2}{n_j}\right) = 7013.67$

$$\mathbf{SS_B} = \text{III} - \text{I} = 7013.67 - 6778.78 = 234.89$$

$$\mathbf{SS_w} = \text{II} - \text{III} = 7097.00 - 7013.67 = 83.33$$

$$\mathbf{SS_T} = \text{II} - \text{I} = 7097.00 - 6778.78 = 318.22$$

$$\mathbf{MS_B} = \frac{SS_B}{df_B} = \frac{234.89}{2} = 117.44$$

$$\mathbf{MS_W} = \frac{SS_W}{df_W} = \frac{83.33}{6} = 13.88$$

$$\mathbf{F} = \frac{MS_B}{MS_W} = \frac{117.44}{13.88} = 8.461$$

Source	SS	df	MS	F	p
Between	234.89	2	117.44	8.461	<.05
Within	83.33	6	13.88		
Total	318.22	8			

6. Decision

Since F_{obs} (8.461) is > F_{crit} (5.14), reject H_o and conclude that Music effects productivity. In particular, Jazz improves worker productivity compared to no music. Because the omnibus F is significant, we need post hoc tests to further localize the effect.

$$\mathbf{F_{Comp}} = \frac{\left(\bar{x}_1 - \bar{x}_2\right)^2}{MS_W\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$\mathbf{F_{1x2}} = \frac{\left(20.67 - 28.67\right)^2}{13.88\left(\frac{1}{3} + \frac{1}{3}\right)} = \frac{\left(-8\right)^2}{13.88\left(0.67\right)} = \frac{64}{9.2996} = 6.882$$

$$\mathbf{F_{1x3}} = \frac{\left(20.67 - 33.00\right)^2}{13.88\left(\frac{1}{3} + \frac{1}{3}\right)} = \frac{\left(-12.33\right)^2}{13.88\left(0.67\right)} = \frac{152.0289}{9.2996} = 16.348$$

$$\mathbf{F_{2x3}} = \frac{\left(28.67 - 33.00\right)^2}{13.88\left(\frac{1}{3} + \frac{1}{3}\right)} = \frac{\left(-4.33\right)^2}{13.88\left(0.67\right)} = \frac{18.7489}{9.2996} = 2.016$$

The F_{crit} (with 1 and 6 df) = 5.99, thus, the first 2 differences are significant. In other words, both forms of music increased productivity. Furthermore, there was no difference between the two music types.

10) 1. Research Questions

1. Is past drug experience related to current drug interest?
2. Is age related to current drug interest?
3. Does the relationship of past drug experience to current drug interest depend on age?

2. Hypotheses

Factor A:

	Symbols	Words
H_O	$\alpha_1 = \alpha_2$	Past drug experience is not related to current drug interest.
H_A	Not H_o	Past drug experience is related to current drug interest.

Factor B:

	Symbols	Words
H_O	$\beta_1 = \beta_2$	Age is unrelated to current drug interest.
H_A	Not H_o	Age is related to current drug interest

A x B Interaction:

	Symbols	Words
H_O	$\alpha\beta_{11}=\alpha\beta_{12}=\alpha\beta_{21}=\alpha\beta_{22}$	Relationship of past drug experience to current drug interest does not depend on age.
H_A	Not H_o	Relationship of past drug experience to current drug interest depends on age.

3. Assumptions

1. H_o
2. Subjects are sampled randomly.
3. DV is normally distributed.
4. Groups are independent.
5. Population variances are homogenous.
6. n's are equal and greater than 1.
7. Factors are fixed.

4. Decision rules

We have a 2x2 factorial design with 6 subjects per group:

df_A	=p-1	=2-1	=1
df_B	=q-1	=2-1	=1
df_{AxB}	=(p-1)(q-1)	=(2-1)(2-1)	=1
df_{Within}	=pq(n-1)	=(2*2)(6-1)	=20
Df_T	=npq-1	=(6*2*2)-1	=23

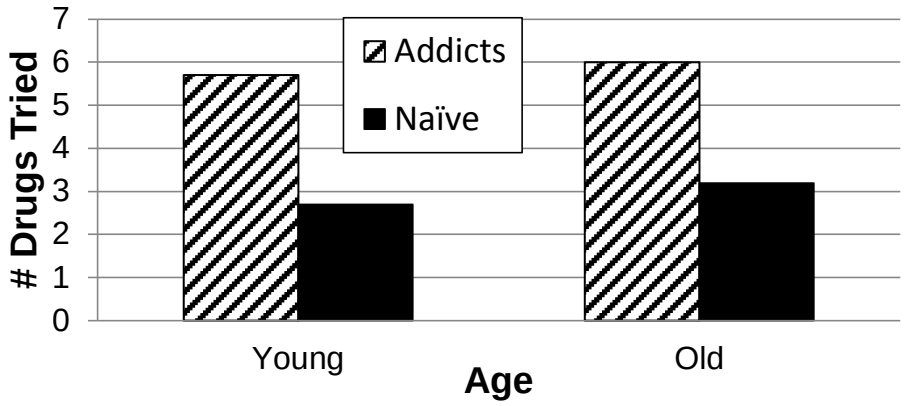
Thus, the critical values are:

Source	A	B	AxB
df	1/20	1/20	1/20
F	4.35	4.35	4.35

For the 2x2 case, all three critical values are the same.
If $F_{obs} \geq F_{crit}$, reject H_o. Otherwise do not reject H_o.

5. Computation

Age	**b_1**				**b_2**				
Past drug experience	**a_1**		**a_2**		**a_1**		**a_2**		
	X	**X^2**	**X**	**X^2**	**X**	**X^2**	**X**	**X^2**	
	8	64	3	9	9	81	5	25	
	2	4	4	16	4	16	4	16	
	7	49	6	36	7	49	5	25	
	6	36	1	1	3	9	0	0	
	6	36	2	4	8	64	2	4	
	5	25	0	0	5	25	3	9	
$\sum T_{jk}$	34		16		36		19		105 = **T**
n_{jk}	6		6		6		6		24 = **N**
$\bar{x}_{jk}$	5.7		2.7		6.0		3.2		
$\sum_{i=1}^{n} X_{ijk}^2$		214		66		244		79	603 = **II**



And the marginal sum table:

	b_1	**b_2**		
a_1	34	36	70	**$T_{j.}$'s**
a_2	16	19	35	
	50	55	105	**$T_{..}$**
	$T_{.k}$'s			

The relevant quantities:

I. $\frac{T_{..}^2}{N} = \frac{105^2}{24} = \frac{11025}{24} = 459.375$

II. $\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{q}\left(X_{ijk}\right)^2 = 603$

III. $\frac{\sum_{j=1}^{p} T_{j.}^2}{nq} = \frac{\left(70^2 + 35^2\right)}{\left(6*2\right)} = \frac{\left(4900+1225\right)}{12} = \frac{6125}{12} = 510.417$

IV. $\frac{\sum_{k=1}^{q} T_{.k}^2}{np} = \frac{\left(50^2 + 55^2\right)}{\left(6*2\right)} = \frac{\left(2500+3025\right)}{12} = \frac{5525}{12} = 460.417$

V. $\frac{\sum_{j=1}^{p}\sum_{k=1}^{q} T_{jk}^2}{n} = \frac{\left(34^2 + 36^2 + 16^2 + 19^2\right)}{6}$

$= \frac{\left(1156+1296+256+361\right)}{6} = \frac{3069}{6} = 511.5$

Thus:

$\mathbf{SS_A}$	= III-I	= 510.417-459.375	= 51.042
$\mathbf{SS_B}$	= IV-I	= 460.417-459.375	= 1.042
$\mathbf{SS_{AxB}}$	= V+I-III-IV	= 511.5+459.375-510.417-460.417	= 0.41
$\mathbf{SS_W}$	= II-V	= 603-511.5	= 91.5
$\mathbf{SS_T}$	= II-I	= 603-459.375	= 143.625

$\mathbf{MS} = \frac{SS}{df}$ and $\mathbf{F} = \frac{MS_{source}}{MS_{error}}$

$\mathbf{MS_A} = \frac{51.042}{1} = 51.042$ and $\mathbf{F_A} = \frac{51.042}{4.575} = 11.157$

$\mathbf{MS_B} = \frac{1.042}{1} = 1.042$ and $\mathbf{F_B} = \frac{1.042}{4.575} = 0.228$

$\mathbf{MS_{AxB}} = \frac{0.41}{1} = 0.41$ and $\mathbf{F_{AxB}} = \frac{0.41}{4.575} = 0.090$

Source	SS	df	MS	F	P
A	51.042	1	51.042	11.157	< .05
B	1.042	1	1.042	0.228	> .05
AxB	0.41	1	0.41	0.079	> .05
Within	91.5	20	4.575		
Total	143.625	23			

6. Decision

1. There is a main effect of past drug experience which says that individual's current drug use is increased with past drug experience compared to naïve users.
2. There is no main effect of age which says that older subjects' current drug use was similar to the younger subjects' current drug use.
3. The lack of an interaction indicates that age and past drug experience do not depend on one another when analyzing current drug use.

11) There are numerous questions which could be asked. Some examples include:

1. Is there a correlation between the students' GPA and their ACT Scores? (Pearson's r)

2. Is there a correlation between class rank and income? (Spearman's Rho)

3. Do the students who make under $10,000 have significantly higher GPAs than the students who make $10,000 or more? (2 sample t-test)

4. Are there gender differences in ACT scores? (2 sample t-test)

5. Is class rank related to GPA? (Spearman's Rho or a one way ANOVA)

6. Two-Way ANOVA:
1) Does GPA change with class rank?
2) Does GPA change with gender?
3) Does the relationship between class rank and GPA depend on gender?