AxB ANOVA Exercises

1. To study the effects of a cow's age and breed on the percentage of butterfat in milk, ten cows were sampled from each of the six age-breed groups. First, calculate the main effects and the interaction effects. Then, plot the means and predict if you will find significant interaction. Create an ANOVA summary table and run a complete analysis of this data. State your conclusions.

	Guernsey	**Holstein-Fresian**	**Jersey**	
Mature	Mean = 4.85 SD = 0.503 n = 10	Mean = 3.72 SD = 0.329 n = 10	Mean = 5.24 SD = 0.547 n = 10	Mean = 4.603 n = 30
Young	Mean = 5.05 SD = 0.466 n = 10	Mean = 3.62 SD = 0.166 n = 10	Mean = 5.34 SD = 0.674 n = 10	Mean = 4.670 n = 30
	Mean = 4.95 n = 20	Mean = 3.67 n = 20	Mean = 5.29 n = 20	Mean = 4.637 n = 60

2. Do women pay more than men for haircuts? Does the price for a haircut depend on the region in which you live? The following table summarizes statistics about the price paid per haircut for 60 individuals. Run an appropriate analysis and state your conclusion.

	Female	**Male**	
Rural	Mean = 18.8 SD = 6.089 n = 10	Mean = 8.45 SD = 3.467 n = 10	Mean = 13.625 SD = 7.172 n = 20
Urban	Mean = 31.0 SD = 7.348 n = 10	Mean = 12.095 SD = 4.980 n = 10	Mean = 21.5475 SD = 11.462 n = 20
Suburban	Mean = 20.6 SD = 10.167 n = 10	Mean = 9.8 SD = 2.616 n = 10	Mean = 15.2 SD = 9.105 n = 20
	Mean = 23.467 SD = 9.5 n = 30	Mean = 10.115 SD = 3.987 n = 30	Mean = 16.7908 SD = 9.874 n = 60

3. A history professor decides to give an essay final to his class. He randomly gives blue-books to half the class and computers to the other half. In addition, the students were partitioned into three groups according to their typing ability. Answers written in blue-books were later typed and scoring was done blindly. The essays were then graded. The following table summarizes the results of this study. Run an appropriate analysis and state your conclusion.

	Blue-book	**Computer**	
No Typing Ability	Mean = 33.67 SD = 1.5275 n = 3	Mean = 34.0 SD = 2.6458 n = 3	Mean = 33.83 SD = 1.9408 n = 6
Some Typing Ability	Mean = 46.0 SD = 3.0 n = 3	Mean = 32.67 SD = 2.3094 n = 3	Mean = 39.33 SD = 7.6855 n = 6
High Typing Ability	Mean = 22.67 SD = 2.0817 n = 3	Mean = 33.0 SD = 3.0 n = 3	Mean = 27.83 SD = 6.1128 n = 6
	Mean = 34.11 SD = 10.301 n = 9	Mean = 33.22 SD = 2.3863 n = 9	Mean = 33.67 SD = 2.3863 n = 18

4. A group of students were presented with a one-week lecture on an unfamiliar topic. At the end of the week, the students were administered a test on the material from the lecture. University officials were interested in the effects of classroom size on student achievement, so they randomly assigned students to one of four groups:

   LL Group: These students received the lecture in a large classroom and were tested in a large classroom
   LS Group: These students received the lecture in a large classroom but were tested in a small classroom
   SL Group: These students received the lecture in a small classroom but were tested in a large classroom
   SS Group: These students received the lecture in a small classroom but were tested in a small classroom

   The following table displays the data from this study:

		Lecture Room	
		Small	**Large**
Test Room	**Small**	22 15 20 17 16	5 8 1 1 5
	Large	1 4 2 5 8	15 20 11 18 16

   Run an AxB ANOVA on this data and state appropriate conclusions. If a significant interaction effect exists, make sure you test for simple effects. If no significant interaction effect exists, then test for main effects.

Solutions:

	Guernsey	Holstein-Fresian	Jersey	
Mature	Mean = 4.85 SD = 0.503 n = 10	Mean = 3.72 SD = 0.329 n = 10	Mean = 5.24 SD = 0.547 n = 10	Mean = 4.603 n = 30
Young	Mean = 5.05 SD = 0.466 n = 10	Mean = 3.62 SD = 0.166 n = 10	Mean = 5.34 SD = 0.674 n = 10	Mean = 4.670 n = 30
	Mean = 4.95 n = 20	Mean = 3.67 n = 20	Mean = 5.29 n = 20	Mean = 4.637 n = 60

$$SS_{Breed} = 20\left[(4.95-4.637)^2+(3.67-4.637)^2+(5.29-4.637)^2\right] = 29.18934$$

$$SS_{Age} = 30\left[(4.603-4.637)^2+(4.67-4.637)^2\right] = 0.06735$$

$$SS_{BreedxAge} = 10\begin{bmatrix}(4.85-4.603-4.95+4.637)^2+(3.72-3.67-4.603+4.637)^2+\\(5.24-4.603-5.29+4.637)^2+(5.05-4.67-4.95+4.637)^2+\\(3.62-4.67-3.67+4.637)^2+(5.34-5.29-4.67+4.637)^2\end{bmatrix} = 0.23335$$

$$SS_{Error} = 9(.503^2+.329^2+.547^2+.466^2+166^2+.674^2) = 12.235$$

Source	Sums of Squares	Degrees of freedom	Mean Square	Mean Square Ratio
Breed (A)	29.189	2	14.595	64.295
Age (B)	0.067	1	0.067	0.295
Interaction (AB)	0.233	2	0.117	0.515
Error	12.235	54	0.227	
Total	41.724	59		

Conclusions:

1. **No significant interaction (we can look at main effects)**
2. **No significant age effect**
3. **Significant breed effect. The breed of cow impacts the percentage of butterfat in milk. We would have to conduct follow-up tests to find where the differences exist. A Tukey procedure finds that Jersey cows have a significantly higher percentage of butterfat than Guernsey cows.**
4. **Breed and Age account for 71% of the total variance**

Solutions:

	Female	Male	
Rural	Mean = 18.8 SD = 6.089 n = 10	Mean = 8.45 SD = 3.467 n = 10	Mean = 13.625 SD = 7.172 n = 20
Urban	Mean = 31.0 SD = 7.348 n = 10	Mean = 12.095 SD = 4.980 n = 10	Mean = 21.5475 SD = 11.462 n = 20
Suburban	Mean = 20.6 SD = 10.167 n = 10	Mean = 9.8 SD = 2.616 n = 10	Mean = 15.2 SD = 9.105 n = 20
	Mean = 23.467 SD = 9.5 n = 30	Mean = 10.115 SD = 3.987 n = 30	Mean = 16.7908 SD = 9.874 n = 60

$$SS_{Sex} = 30\left[(23.467-16.7908)^2 + (10.115-16.7908)^2\right] = 2674$$

$$SS_{\mathrm{Re}\,gion} = 20\left[(13.625-16.7908)^2 + (21.5475-16.7908)^2 + (15.2-16.7908)^2\right] = 703.58$$

$$SS_{SxR} = 10\begin{bmatrix}(18.8-13.625-23.467+16.79)^2 + (8.45-13.625-10.115+16.79)^2 + \\ (31-21.5475-23.467+16.79)^2 + (12.095-21.5475-10.115+16.79)^2 + \\ (20.6-15.2-23.467+16.79)^2 + (9.8-15.2-10.115+16.79)^2\end{bmatrix} = 231.8026$$

$$SS_{Error} = 9(6.089^2 + 3.467^2 + 7.348^2 + 4.98^2 + 10.167^2 + 2.616^2) = 2142.908$$

Source	Sums of Squares	Degrees of freedom	Mean Square	Mean Square Ratio
Sex	2674	1	2674	90.08
Region	703.58	2	351.79	8.865
Interaction	231.8026	2	115.9013	2.921 (CV = 3.15)
Error	2142.908	54	39.6835	
Total	Sum of the above values	59		

Conclusions:

1. **No significant interaction (we can look at main effects). Note the interaction is significant at a .10 level**
2. **Significant Sex effect. Females pay more for haircuts.**
3. **Significant Region effect. We would have to run follow-up tests.**

	Blue-book	Computer	
No Typing Ability	Mean = 33.67 SD = 1.5275 n = 3	Mean = 34.0 SD = 2.6458 n = 3	Mean = 33.83 SD = 1.9408 n = 6
Some Typing Ability	Mean = 46.0 SD = 3.0 n = 3	Mean = 32.67 SD = 2.3094 n = 3	Mean = 39.33 SD = 7.6855 n = 6
High Typing Ability	Mean = 22.67 SD = 2.0817 n = 3	Mean = 33.0 SD = 3.0 n = 3	Mean = 27.83 SD = 6.1128 n = 6
	Mean = 34.11 SD = 10.301 n = 9	Mean = 33.22 SD = 2.3863 n = 9	Mean = 33.67 SD = 2.3863 n = 18

$$SS_{Format} = 9\left[(34.11-33.67)^2 + (33.22-33.67)^2\right] = 3.5649$$

$$SS_{Typing} = 6\left[(33.83-33.67)^2 + (39.33-33.67)^2 + (27.83-33.67)^2\right] = 397$$

$$SS_{FxT} = 3\begin{bmatrix}(33.67-33.83-34.11+33.67)^2 + (34-33.83-33.22+33.67)^2 + \\ (46-39.33-34.11+33.67)^2 + (32.67-39.33-33.22+33.67)^2 + \\ (22.67-27.83-34.11+33.67)^2 + (33-27.83-33.22+33.67)^2\end{bmatrix} = 423.1974$$

$$SS_{Error} = 2(1.5275^2 + 2.6458^2 + 3^2 + 2.3094^2 + 2.0817^2 + 3^2) = 74$$

Source	Sums of Squares	Degrees of freedom	Mean Square	Mean Square Ratio
Format	3.5649	1	3.5649	0.578
Typing	397	2	198.5	32.187
Interaction	423.1974	2	211.5987	34.311*
Error	74	12	6.167	
Total	897.7623	17		

1. **Significant interaction. We need to examine the simple effects. Since we're interested in comparing bluebooks to computers, we'll split by typing ability.**

	Blue-book	Computer	
No Typing Ability	Mean = 33.67 SD = 1.5275 n = 3	Mean = 34.0 SD = 2.6458 n = 3	Mean = 33.83 SD = 1.9408 n = 6

	Blue-book	Computer	
Some Typing Ability	Mean = 46.0 SD = 3.0 n = 3	Mean = 32.67 SD = 2.3094 n = 3	Mean = 39.33 SD = 7.6855 n = 6

	Blue-book	Computer	
High Typing Ability	Mean = 22.67 SD = 2.0817 n = 3	Mean = 33.0 SD = 3.0 n = 3	Mean = 27.83 SD = 6.1128 n = 6

I need to run an ANOVA for each of these.

	Blue-book	Computer	
No Typing Ability	Mean = 33.67 SD = 1.5275 n = 3	Mean = 34.0 SD = 2.6458 n = 3	Mean = 33.83 SD = 1.9408 n = 6

$$SS_{Format} = 3\left[(33.67-33.83)^2 + (34-33.83)^2\right] = 0.1635$$

I'm going to work under the assumption that these variances are equal, so I can use MSE from the AxB ANOVA.

Source	Sums of Squares	Degrees of freedom	Mean Square	Mean Square Ratio
Format	0.1635	1	0.1635	0.0265
Error	74 (from AxB analysis)	12 (from AxB analysis)	6.167	

	Blue-book	Computer	
Some Typing Ability	Mean = 46.0 SD = 3.0 n = 3	Mean = 32.67 SD = 2.3094 n = 3	Mean = 39.33 SD = 7.6855 n = 6

$$SS_{Format} = 3\left[(46-39.33)^2 + (32.67-39.33)^2\right] = 266.5335$$

These variances can be assumed to be equal, so I can use MSE from the AxB ANOVA.

Source	Sums of Squares	Degrees of freedom	Mean Square	Mean Square Ratio
Format	266.5335	1	266.5335	43.2*
Error	74	12	6.167	

	Blue-book	Computer	
High Typing Ability	Mean = 22.67 SD = 2.0817 n = 3	Mean = 33.0 SD = 3.0 n = 3	Mean = 27.83 SD = 6.1128 n = 6

$$SS_{Format} = 3\left[(22.67-27.83)^2 + (33-27.83)^2\right] = 160.0635$$

Even though the variances are similar (and would pass an Fmax test), I'm going to calculate SSE for this data separately.

$$SS_E = 2(2.0817^2) + 2(3.0^2) = 26.67$$

Degrees of freedom for the error term would be N – a = N – 2 = 6 – 2 = 4

Source	Sums of Squares	Degrees of freedom	Mean Square	Mean Square Ratio
Format	160.0635	1	160.0635	24.0*
Error	26.67	4	6.67	

CONCLUSIONS

Students with high typing ability earn higher essay scores when they use a computer.

Students with some typing ability earn lower scores via computer.

		Lecture Room: Small	Lecture Room: Large	
Test Room	**Small**	N = 5 Mean = 18 SD = 2.915	N = 5 Mean = 4 SD = 3.000	N = 10 Mean = 11 SD = 7.888
	Large	N = 5 Mean = 4 SD = 2.739	N = 5 Mean = 16 SD = 3.391	N = 10 Mean = 10 SD = 6.960
		N = 10 Mean = 11 SD = 7.846	N = 10 Mean = 10 SD = 7.008	**N = 20** **Mean = 10.5** **SD = 7.258**

$$SS_{Lecture} = 10\left[(10-10.5)^2 + (11-10.5)^2\right] = 5.00$$

$$SS_{Test} = 10\left[(10-10.5)^2 + (11-11.5)^2\right] = 5.00$$

$$SS_{AxB} = 5\left[\begin{array}{l}(18-11-11+10.5)^2 + (4-10-11+10.5)^2 + \\ (4-11-10+10.5)^2 + (16-10-10+10.5)^2\end{array}\right] = 845$$

$$SS_{Error} = 4(2.915^2 + 3.000^2 + 2.739^2 + 3.391^2) = 146.00$$

Source	Sums of Squares	Degrees of freedom	Mean Square	Mean Square Ratio
Lecture	5.00	1	5.00	0.548
Test	5.00	1	5.00	0.548
Interaction	845.00	1	845.00	92.603
Error	146.00	16	9.125	
Total	1001.00	19		

Significant interaction. With 4 means, we could run Tukey's test to compare each pair of means. We'll examine the simple effects of the lecture room (split by test).

	Small Lecture	**Big Lecture**	
Small Test	N = 5 Mean = 18 SD = 2.915	N = 5 Mean = 4 SD = 3.000	N = 10 Mean = 11 SD = 7.888

	Small Lecture	**Big Lecture**	
Large Test	N = 5 Mean = 4 SD = 2.739	N = 5 Mean = 16 SD = 3.391	N = 10 Mean = 10 SD = 6.960

I need to run an ANOVA for each of these.

	Small Lecture	Big Lecture	
Small Test	N = 5 Mean = 18 SD = 2.915	N = 5 Mean = 4 SD = 3.000	N = 10 Mean = 11 SD = 7.888

$$SS_{Lecture} = 5\left[(18-11)^2 + (4-11)^2\right] = 490$$

Source	Sums of Squares	Degrees of freedom	Mean Square	Mean Square Ratio
Format	490	1	490	53.7*
Error	146 (from AxB analysis)	16	9.125	

	Small Lecture	Big Lecture	
Large Test	N = 5 Mean = 4 SD = 2.739	N = 5 Mean = 16 SD = 3.391	N = 10 Mean = 10 SD = 6.960

$$SS_{Format} = 5\left[(4-10)^2 + (16-10)^2\right] = 360$$

Source	Sums of Squares	Degrees of freedom	Mean Square	Mean Square Ratio
Format	360	1	360	39.45*
Error	146 (from AxB analysis)	16	9.125	

The small lecture room is preferable if the testing room will be small.
The large lecture room is preferable if the testing room will be large.