Test 13 AP Statistics Name:

Part I: Multiple Choice. *Circle the letter corresponding to the best answer.*

Use the following for questions 1 through 4:

Does listening to music aid short-term memory? A student researcher addressed this question by randomly assigning 90 volunteers to one of three groups. The first group studied a list of 30 words for five minutes while listening to Beethoven, the second group studied the same list of words while listening to Bagpipe Metal, and the third group studied the words in silence. All subjects were then asked to write down as many of the words as they could remember, and One-way Analysis of Variance was performed on the variable Number of words recalled. Here are the results.

```
Source  DF      SS     MS     F      P
Music    2   197.1   98.5  7.58  0.001
Error   87  1130.7   13.0
Total   89  1327.8

S = 3.605   R-Sq = 14.84%   R-Sq(adj) = 12.89%


                                  Individual 95% CIs For Mean Based on Pooled StDev
Level       N    Mean   StDev    -+---------+---------+---------+--------
Beethoven  30  13.933   3.750    (-------*-------)
BP Metal   30  16.500   3.830                    (-------*-------)
None       30  17.433   3.202                          (-------*-------)
                                  -+---------+---------+---------+--------
                                 12.8      14.4      16.0      17.6

Pooled StDev = 3.605
```

1. Which of the following is the appropriate *alternative* hypothesis for the ANOVA F test in this situation?
   A. The variables "Number of words recalled" and "Music" are independent.
   B. The mean number of words recalled is lowest for the 30 subjects who listened to Beethoven.
   C. The mean number of words recalled is the same for all three treatments.
   D. The mean number of words recalled is different for at least one treatment.
   E. The mean number of words recalled is different for each treatment.

2. Which of the following is the correct F-distribution for the ANOVA test comparing these three means?
   A. F(87)
   B. F(89)
   C. F(2, 87)
   D. F(2, 89)
   E. F(87, 89)

3. Which of the following are conditions that must be satisfied in order to carry out the ANOVA *F*-test on these data?

   I. The subjects were randomly assigned to the three treatment groups.
   II. The treatment groups and the individual observations within each treatment group are independent.
   III. The variable Number of words recalled has the same standard deviation for all treatments.

   A. I only
   B. II only
   C. I and II
   D. I and III
   E. I, II, and III

4. Assume the conditions for the ANOVA *F*-test have been satisfied. At the $\alpha = 0.01$ significance level, which of the following is an appropriate conclusion for the ANOVA *F*-test?

   A. Fail to reject H_0. We do not have convincing evidence that the mean number of words recalled is different for at least one of the treatment groups.
   B. Reject H_0. We have convincing evidence that the means of the three treatment groups are not equal.
   C. Reject H_0. We have convincing evidence that the three means are all different.
   D. Fail to reject H_0. We have convincing evidence that the means of the three treatment groups are not equal.
   E. Fail to reject H_0. We have convincing evidence that the three means are all different.

Use the following to answer questions 5 – 8:

At what age do babies learn to crawl? Does it take longer for them to learn in the winter, when babies are often bundled in clothes that restrict their movements? Data were collected from parents who brought their babies into the University of Denver Infant Study Center to participate in one of a number of experiments. Parents reported the birth month and the age at which their child was first able to creep or crawl a distance of four feet within one minute. The resulting data were grouped by month of birth. The data below are for January, May, and September. (Crawling age is given in weeks.)

		Crawling Age		
		Mean	Std. Dev.	n
Birth Month	January	29.84	7.08	32
	May	28.58	8.07	27
	September	33.83	6.93	38

Assume that the data represent three independent SRSs, one from each of the three populations of interest (all babies born in a particular month), and that crawling ages are normally distributed for all three populations. A partial ANOVA table is given below.

Source	df	Sums of Squares	Mean Square	F-ratio
Birth month		505.26		
Error			53.45	
Total				

5. What are the degrees of freedom for birth month (numerator)?

A. 2 B. 3 C. 4 D. 94 E. 97

6. What are the degrees of freedom for error (denominator)?

A. 2 B. 3 C. 4 D. 94 E. 97

7. The null hypothesis for the ANOVA F test is that the population mean crawling ages are equal for all three birth months. Which of the following is an appropriate alternative hypothesis?

A. The population mean crawling age is larger for January than for the other two months.
B. The population mean crawling age is larger for May than for the other two months.
C. The three months all have different population mean crawling ages.
D. The population mean crawling ages for the three months are all within one standard deviation of each other.
E. The population mean crawling age is different for at least one of the three months.

8. Which of the following is the value of the ANOVA F test statistic for equality of the population means of the three birth months?

A. 3.15 B. 3.42 C. 4.73 D. 6.30 E. 9.45

Use the following for questions 9 – 12:

A high school teacher suspects that students of different ages estimate the ages of adults differently. He asks randomly-selected sophomores, juniors, and seniors to guess the age of a person in a photograph and plans to compare the mean age guesses using one-way analysis of variance. Here is the computer output from his analysis:

```
Source        DF      SS     MS     F      P
School Year    2  135.53  67.76  6.85  0.003
Error         43  425.08   9.89
Total         45  560.61

S = 3.144   R-Sq = 24.17%   R-Sq(adj) = 20.65%

                                Individual 95% CIs For Mean Based on
                                Pooled StDev
Level       N    Mean  StDev  ---------+---------+---------+---------+
Senior     15  43.800  2.426  (-------*-------)
Junior     16  47.875  3.948                    (-------*-------)
Sophomore  15  46.733  2.789            (--------*-------)
                                ---------+---------+---------+---------+
                                      44.0      46.0      48.0      50.0
Pooled StDev = 3.144
```

9. Which of the following is the appropriate null hypothesis for the ANOVA *F*-test in this situation?
   A. The population mean age guess for seniors is higher that the mean age guess for juniors and sophomores.
   B. The population mean age guesses for all three age groups are different.
   C. The population mean age guess for at least one age group is different from the others.
   D. The population mean age guesses for all three age groups are equal.
   E. The population mean age guesses for at least two of the three age groups are equal.

10. Which of the following statements about required conditions for the ANOVA *F*-test is correct in this situation?
   A. None of the three distributions of sample guesses should show signs of strong skew.
   B. As long as there are no outliers, the ANOVA test is appropriate.
   C. As long as two of the three distributions of sample guesses are close to Normally distributed, the test is robust with respect to strong skew in the third distribution.
   D. The shapes of the distributions of samples guesses don't matter, because the condition of equal sample standard deviations is violated.
   E. The shapes of the distributions of samples guesses don't matter, because the condition of independence has been violated, since the three grade levels were sampled from the same school.

11. Assuming all necessary conditions have been met, what is the appropriate conclusion for the ANOVA F test?

A. Reject H_o. These data do not provide enough evidence to conclude that there is a difference in the true mean age guesses in the three age groups.

B. Reject H_o. These data provide convincing evidence that there is a difference in the true mean age guesses in these three age groups.

C. Accept H_a. These data provide convincing evidence that there is a difference in the true mean age guesses in these three age groups.

D. Fail to reject H_o. These data do not provide enough evidence to conclude that there is a difference in the true mean age guesses in the three age groups.

E. Fail to reject H_o. These data provide convincing evidence that there is a difference in the true mean age guesses in these three age groups.

12. Based on the numerical summaries in the computer output, which of the follow statement is true?

A. All three samples have about the same range.

B. The mean age guesses by seniors is closest to the person's actual age.

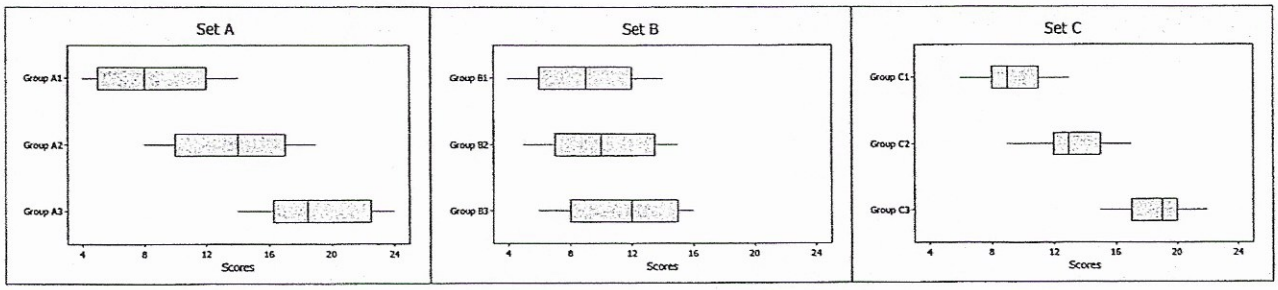
C. There appears to be little difference between the age guesses of the three age groups.

D. Age guesses by juniors are significantly higher than age guesses by sophomores.

E. On average, the age guesses of seniors is much lower than that of the other two age groups.

Use the following for questions 13 – 14 :

Below are three sets of parallel box plots, labeled Set A, Set B, and Set C. Each set of box plots describes the results of random samples of size $n = 30$ from three independent populations.



An ANOVA F test was performed on each set of samples to compare means. Assume conditions for performing the F test were met in each case.

13. Which one of the following statements is supported by these box plots?

A. Set A has much larger within-group variation that either Set B or Set C.

B. Set B has more between-group variation than Set C.

C. Set C has much larger within-group variation than either Set A or Set B.

D. Set B has the lowest between-group variation.

E. Set A has much less between-group variation than Set C.

14. Which of the following describes the relationship between the F-statistics for these three ANOVA tests?

A. $F_{\text{Set A}} > F_{\text{Set B}} > F_{\text{Set C}}$

B. $F_{\text{Set A}} > F_{\text{Set C}} > F_{\text{Set B}}$

C. $F_{\text{Set B}} > F_{\text{Set A}} > F_{\text{Set C}}$

D. $F_{\text{Set C}} > F_{\text{Set B}} > F_{\text{Set A}}$

E. $F_{\text{Set C}} > F_{\text{Set A}} > F_{\text{Set B}}$

Part II: Free Response *Communicate your thinking clearly and completely.*

For exercises 15 – 19: A researcher decided to investigate turkeys in Annadel State Park to see if there is a difference in life satisfaction among turkeys who participate in various athletic pursuits. The researcher considered four groups - those who participate in Track and Field, Curling, Badminton and Water Polo. The *Life Satisfaction Index* questionnaire was given to simple random samples of six turkeys selected from each of the four groups. The researcher wants to know if there is any difference in general life satisfaction among these four groups of turkeys. Past experience has indicated that scores on this index follow a normal distribution. The means and standard deviations for the scores for the four groups are:

	T & F	Curling	Badminton	Water Polo
mean	66.6700	71.1700	37.1700	32.8300
st. dev.	18.6620	15.5618	11.6520	9.6628

15. Verify that the conditions for ANOVA have been met.

16. State the hypotheses H_0 and H_a .

17. Complete the ANOVA table:

Source	D.F.	Sum of Squares	Mean Square	F
Groups				
Error				
Totals		*n/a*	*n/a*	

18. Conduct the appropriate test of significance by recording the important information in the blanks:

F = _____, the distribution is $F($, $)$, and the P-value = _____

19. Write your conclusion.

For exercises 20 – 23: The SRHS Math Club has decided to invest in a Slurpee machine, but wants to conduct some research to help decide which flavors are likely to sell best. After consulting with the AP Statistics class about proper methods for gathering data (assume all conditions for inference have been met), they obtain the following results. Each value represents a score of taste excellence for the corresponding flavor.

Cola	Pepsi & Banana	Root Beer
13	12	7
17	8	19
19	6	15
11	16	14
20	12	10
15	14	16
18	10	18
9	18	11
12	4	14
16	11	11

20. State the hypotheses H_0 and H_a .

21. Complete the ANOVA table:

Source	D.F.	Sum of Squares	Mean Square	F
Groups				
Error				
Totals		*n/a*	*n/a*	

23. Conduct the appropriate test of significance and record the important information in the blanks:

F = _____, the distribution is $F($, $)$, and the P-value = _____

24. Write your conclusion.