

Name Key
 (Please ALSO write your name on the back of the last page.)

Ag 400 - Quiz 1
 Makeup - Fall 2006

- A. Professor I. M. Offensive takes attendance at each meeting of History 101 which he teaches. For the first eight class meetings, he compiles the following information:

| <u>Class Meeting</u> | <u>Number of Students Absent</u> | <u>X</u> | <u>X²</u> |
|----------------------|----------------------------------|-----------|----------------------|
| 1 | 5 | 12 | 144 |
| 2 | 7 | 10 | 100 |
| 3 | 10 | 7 | 49 |
| 4 | 12 | 6 | 36 |
| 5 | 5 | 5 | 25 |
| 6 | 3 | 5 | 25 |
| 7 | 6 | 3 | 9 |
| 8 | 0 | 0 | 0 |
| | | <u>48</u> | <u>388</u> |

- (5) 1. Calculate the mean number of absences per class or indicate why it cannot be calculated?

$$\bar{X} = \frac{48}{8} = 6.0$$

- (5) 2. What is the median number of absences per class?

$$\text{Md position} = \frac{9}{2} = 4.5$$

$$\text{Md} = 5.5$$

- (6) 3. Calculate the standard deviation for the above distribution.

$$S = \sqrt{\frac{388 - \frac{(48)^2}{8}}{8}} = \sqrt{\frac{100}{8}} = \sqrt{12.5} = 3.5$$

(16) 4. Indicate whether each of the following statements is true (T) or false (F) in terms of the above data. If any part of a statement is untrue, it should be marked false (F). Add comments if you wish, to clarify your answers.

T a. The statistical unit in the above data is a class meeting.

T b. The variable, number of students absent, is measured by an interval scale.

F c. The range in absences per class is 7.

T d. The above data are what we have called ungrouped data.

F e. The mode for the above distribution is 12.

F f. The figures calculated in 1, 2, and 3 above are descriptive statistics that measure the central tendency of the distribution.

F g. Since Class Meeting #8 had no absences, that class meeting can be deleted from the analysis with no loss of data.

F h. Suppose Professor Offensive erroneously recorded zero for the number of absences for Class Meeting #8, when, in reality there were 2 students absent that period. The mean and median for the "corrected data" would both increase over the figures calculated above.

- (20) B. To ascertain whether students who live on campus differ from those who live off campus in regard to the number of times they miss class, Professor Offensive selected a random sample of 36 students from the 100 students in the class and compiles the following table.

| # Times missed class in first two months | Residence | |
|---|------------------------------|------------|
| | On Campus | Off Campus |
| | -----number of students----- | |
| 5 or more | 11 | 9 |
| < 5 | 11 | 5 |

Indicate whether each of the following statements is true (T) or false (F) in terms of the above data. If any part of a statement is untrue, it should be marked false (F). Add comments, if you wish, to clarify your answers.

- F 1. The statistical unit here is an absence during the first two months of class.
- T 2. The independent variable, student's residence is measured by a two category nominal scale.
- T 3. The dependent variable is treated as a nominal scale in this table.
- F 4. For on-campus students, there is no relationship between residence location (on vs. off campus) and number of classes missed.
- F 5. On-campus students were more likely than off-campus students to miss class 5 or more times in the first 2 months.
- F 6. The ratio of off-campus to on-campus students in the above table is 22:14.
- F 7. Calculations (percentages, proportions, and ratios) based on the above table are called parameters.

- F 8. The relationship between student's residence and number of times he/she missed class in the first 2 months can be described as follows:

There are more on-campus than off-campus students in the sample and more students missed 5 or more classes than missed fewer than 5 classes in the first two months.

- T 9. Among students who missed fewer than 5 classes in the first two months, the ratio of on-campus to off-campus students is 2.2.

- I 10. 64% of the off-campus students missed 5 or more classes; 45% of those who missed 5 or more classes lived off campus.

(20) C. Indicate whether each of the following statements is true (T) or false (F). If any part of a statement is untrue, it should be marked false (F). Add comments, if you wish, to clarify your answer.

T 1. In a unimodal distribution that is skewed to the right, the mean is greater than the median.

F 2. The sum of the deviations of X about the mean of X is large when the standard deviation is large.

T 3. $\sum X^2$ is never less than $(\sum X)^2 \div N$

F 4. The variance equals the square root of the standard deviation.

T 5. The mean and standard deviation cannot be calculated from grouped data if the distribution is open-ended.

F 6. Descriptive statistics are calculated from sample data; inferential statistics are obtained from populations.

F 7. The SEX RATIO measures the number of females for every 100 males.

T 8. The mean, median, and mode are all measures of central tendency that can be specified/calculated for intervally scaled variables that are presented as ungrouped data.

T 9. If the standard deviation for a distribution is zero, the range is also zero.

T 10. The variance and standard deviation are never less than zero.

- (10) D. Professor Offensive has used a final exam in History 101 over the last three years in which Section A of the exam is a series of multiple choice questions and Section B is an essay test. Data from a random sample of students who took the test were compiled and saved as an SPSS data file. The file contains the following variables.

ID = Identification number
SEX = Student's gender
1 Male
2 Female
0 No data on gender

SCOREA = Student's Score on Section A of the Test

SCOREB = Student's Score on Section B of the Test

Indicate whether each of the following questions is true (T) or false (F). If any part of a statement is untrue, it should be marked false (F). Add comments, if you wish, to clarify your answers.

- T 1. The statistical unit here is a student and there is a line in the data file for each student in the sample.
- T 2. There are at least four columns in the data file.
- F 3. You could calculate each student's total score for the two sections of the test combined by running FREQUENCIES and specifying that you want STATISTICS=SUM.
- F 4. You could calculate the mean score of Section A per student for the sample by using:

Transform → Compute

And specifying the calculation of the mean.

- T 5. To calculate the mean score of SCOREB for the sample of Male students, you could use SELECT IF. . . and specify SEX = 1, then run FREQUENCIES, and ask for STATISTICS = MEAN

(18) E. The attached output was obtained using the data set described in D above.

Page 1 contains analysis dealing only with males in the sample

Page 2 contains analysis dealing only with females in the sample

TEST is a computed variable obtained by combining information on the test performance of each student from Sections A and B of the exam. The higher the TEST value, the higher the student's grade.

Indicate whether each of the following questions is true (T) or false (F). If any part of a statement is untrue, it should be marked false (F). Add comments, if you wish, to clarify your answers.

T 1. Overall, females had higher test scores than did males.

T 2. The dispersion in TEST scores is greater for males than for females.

T 3. The distribution of TEST scores is skewed slightly to the right for males, but skewed slightly to the left for females.

T 4. The modal score for males is 80, while the modal score for females is 85.

F 5. The distribution of TEST scores for males is more platykurtic than a normal distribution.

T 6. Two males in the sample have a blank space in the data file where their test scores should be recorded.

I 7. The range in TEST scores for males is 10; for females the range is 7.

(T) 8. More than one third of the females in the sample have TEST scores above 85, but only 2% of the males have such high scores.

F 9. The "statistics" reported on this output are inferential or inductive statistics.

```

USE ALL.
COMPUTE filter_$=(SEX = 1).
VARIABLE LABEL filter_$ 'SEX = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .
FREQUENCIES
  VARIABLES=Test
  /STATISTICS=STDDEV MEAN MEDIAN SKEWNESS SESKEW KURTOSIS SEKURT
  /ORDER= ANALYSIS .

```

Frequencies

[DataSet1] y:\willits\AAATeaching\AAATeaching\Ag400\Quizzes400\Q1m_06.sav

SEX = 1

Statistics^a

Test

| | | |
|------------------------|---------|---------|
| N | Valid | 48 |
| | Missing | 2 |
| Mean | | 80.1250 |
| Median | | 80.0000 |
| Std. Deviation | | 2.15984 |
| Skewness | | .347 |
| Std. Error of Skewness | | .343 |
| Kurtosis | | .412 |
| Std. Error of Kurtosis | | .674 |

a. SEX = 1 MALE

Test^a

| | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|-----------|---------|---------------|--------------------|
| Valid | | | | |
| 76.00 | 2 | 4.0 | 4.2 | 4.2 |
| 77.00 | 4 | 8.0 | 8.3 | 12.5 |
| 78.00 | 4 | 8.0 | 8.3 | 20.8 |
| 79.00 | 8 | 16.0 | 16.7 | 37.5 |
| 80.00 | 10 | 20.0 | 20.8 | 58.3 |
| 81.00 | 9 | 18.0 | 18.8 | 77.1 |
| 82.00 | 6 | 12.0 | 12.5 | 89.6 |
| 83.00 | 2 | 4.0 | 4.2 | 93.8 |
| 84.00 | 1 | 2.0 | 2.1 | 95.8 |
| 85.00 | 1 | 2.0 | 2.1 | 97.9 |
| 86.00 | 1 | 2.0 | 2.1 | 100.0 |
| Total | 48 | 96.0 | 100.0 | |
| Missing | System | 2 | 4.0 | |
| Total | | 50 | 100.0 | |

a. SEX = 1 MALE

```

USE ALL.
COMPUTE filter_$(SEX = 2).
VARIABLE LABEL filter_$ 'SEX = 2 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .
FREQUENCIES
  VARIABLES=Test
  /STATISTICS=STDDEV MEAN MEDIAN SKEWNESS SESKEW KURTOSIS SEKURT
  /ORDER= ANALYSIS .

```

Frequencies

[DataSet1] y:\willits\AAATeaching\AAATeaching\Ag400\Quizzes400\Q1m_06.sav

SEX = 2

Statistics^a

Test

| | | |
|------------------------|---------|---------|
| N | Valid | 46 |
| | Missing | 4 |
| Mean | | 84.6522 |
| Median | | 85.0000 |
| Std. Deviation | | 1.89991 |
| Skewness | | -.185 |
| Std. Error of Skewness | | .350 |
| Kurtosis | | -.900 |
| Std. Error of Kurtosis | | .688 |

a. SEX = 2 FEMALE

Test^a

| | Frequency | Percent | Valid Percent | Cumulative Percent |
|----------------|-----------|---------|---------------|--------------------|
| Valid | | | | |
| 81.00 | 2 | 4.0 | 4.3 | 4.3 |
| 82.00 | 6 | 12.0 | 13.0 | 17.4 |
| 83.00 | 6 | 12.0 | 13.0 | 30.4 |
| 84.00 | 5 | 10.0 | 10.9 | 41.3 |
| 85.00 | 11 | 22.0 | 23.9 | 65.2 |
| 86.00 | 7 | 14.0 | 15.2 | 80.4 |
| 87.00 | 7 | 14.0 | 15.2 | 95.7 |
| 88.00 | 2 | 4.0 | 4.3 | 100.0 |
| Total | 46 | 92.0 | 100.0 | |
| Missing System | 4 | 8.0 | | |
| Total | 50 | 100.0 | | |

a. SEX = 2 FEMALE