Abe Mirza **Sample Test #1** **Statistics**

This is only a sample test!!!

A. Given the data set of ***ages of people*** with diabetes

12, 36, 28, 32, 41, 19, 29, 16, 38, 42, 17, 27, 37, 28 ,**Compute the following**

1. **Mean** 2. **Median** 3. **Mode** 4. **Variance** 5. **St. Dev** 6. **Q1, Q2, Q3**

7. Draw the Box- Plot and **comment** on it 8. Apply all three empirical rules.

B.

A Marketing firm wished to determine whether or not the number of television commercials broadcast was linearly correlated to the sales of its product. The data, obtained from each of several cities, are shown in the following table**,**

X= # of TV Commercials	12	6	9	15	11	15	8	16	12	6
Y= Sales Unit(Y)	7	5	10	14	12	9	6	11	11	8

9. Use the data and plot the data as a scattered diagram and **comment** on the pattern of the points.
10. Compute the correlation coefficient and **comment** on that
11. Compute the slope and y-intercept and write the equation of regression line
12. Compute average and standard deviation for both x and y variables
13. If no. of TV Commercials is 10, then use regression Eq. and estimate sales.
14. If sales is 13 units, then use regression Eq. and estimate no. of TV commercials.
15. Explain the slope.
16. Compute the coefficient of determination and comment on it.

C.

Time(sec)	f	?	?	?	?
6 - 12	4				
12 - 18	8				
18 - 24	10				
24 - 30	16				
30 - 36	20				
36 - 42	14				
42 - 48	6				
48 - 54	2				

17. Draw the Histogram (write your observation) **18**. Draw the Frequency polygon

Compute.
19. Mean **20**. Variance **21**. Standard deviation **22**. Apply all three empirical rules.

D. If out of 12 men and 14 women we select two people at random, find the probability that

23. Both are men **24**. both are women **25**. One of each

E. Give an example of each sampling methods. **F**. What the outlier means in a regression problem?

G. What z-score value is considered to be unusual and why? **H**. What ordinary and unusual z-values mean?
I. What standard deviation is trying to measure? Give an example, and explain how it can be useful.

J. Give two examples of positively and two examples of negatively correlated variables.

K. Discuss the Principles of Causation on page 15 of topic review handout (including type of association and establishing cause-and effect).
L. What significant information histogram, box-plot intends to provide?

M. There were six different stat classes that were offered last semester; one student was randomly selected from each class with his/her final score and the class average and standard deviation.

Joe got score of 32 when the class average was 15 with standard deviation of 9.
Moe got score of 54 when the class average was 24 with standard deviation of 12.
Nielo got score of 66 when the class average was 22 with standard deviation of 16.
April got score of 77 when the class average was 18 with standard deviation of 25.
Max got score of 38 when the class average was 17 with standard deviation of 11.

a)Who did relatively better in the class than the rest?
b)Who did relatively worse in the class than the rest?
c)Who got scores that can be considered unusual?

M1. Three students take equivalent tests of statistics and after the test their papers were graded. Which is the highest relative score?

a. A score of 144 on a test with a mean of 128 and standard deviation of 34.
b. A score of 90 on a test with a mean of 86 and standard deviation of 18.
c. A score of 18 on a test with a mean of 15 and standard deviation of 5.

N.
1. The researcher was an observer at a police sobriety checkpoint at which every 5^{th} driver was stopped and interviewed. He was stunned to witness the arrest of a former student.police5

2. On the day of the presidential election, the news media organize an exit poll in which specific stations are randomly selected and all voters are surveyed as they leave the premises.

3. A researcher interviews 19 work colleagues who work in his building about the new work policies.

4. An IRS researcher investigates cheating on income tax reports by surveying all waiters and waitresses at 20 randomly selected restaurants.

5. A marketing researcher selects 500 drivers under 30 years of age and over 30 years of age.

6. The author surveyed all of his students to obtain sample data consisting of the number of credit cards students possesses.

7. A sample consists of every 49^{th} students from a group of 496 students.

8. The name of each contestant is written on a separate card, the cards are placed in a bag and three names are picked from the bag.

9. At Honda, a researcher has partitioned all registered cars into categories of subcompact, compact, midsize, intermediate, and full size. She is surveying 200 car owner from each category.

10. A John Hopkins University marketing researcher surveys all cardiac patients in each of 30 randomly selected hospitals.

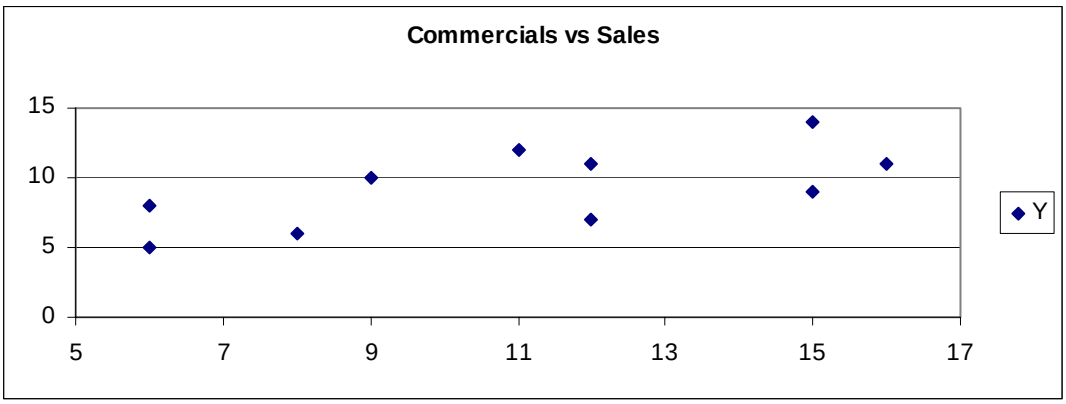
Answers

A.

1. Mean ***28.71*** 2. Median ***28.5***. 3. Mode ***28*** 4. Variance ***94.07***

5. **St.** Dev ***9.699*** 6. Q1, Q2, Q3 ***19, 28.5, 37***

8. All three empirical rules. *-.39 <* ***99%*** *< 57.81* *9.11 <* ***95%*** *< 48.11* 19.01 < **68%** < 38.41



9 .*Positive Correlation* **10**. ***0.661, Positive Correlation*** **11.** Slope and y-intercept ***0.508, 3.710***

The equation of regression line $y = .508\ x + 3.71$

12. Average and standard deviation for both x and y variables ***11, 9.3, 3.68, 2.83***

13. If no. of TV Commercials is 10, then use regression Eq. and estimate sales. ***y' = 8.79***

14. If sales is 13 units, then use regression Eq. and estimate no. of TV commercials. ***x' = 18.29***

15. For every additional hour of TV commercial sales goes up by .51 units.

16. 43.69% of variations of sales is due to hours of TV commercial and 56.31% on other factors.

C.

19. Mean? $\overline{x}$ = **29.70** **20**. Variance? s^2 = **101.98**

21. Standard deviation? s = **10.10** **22**. Apply all three empirical rules.

0 < **99.7 % of data** < 59.7 9.7 < **95 % of data** < 49.7 19.7 < **68 % of data** < 39.7

D. **23**. $\frac{12}{26} \bullet \frac{11}{25} = 20.31\%$ **24**. $\frac{14}{26} \bullet \frac{13}{25} = 28\%$ **25.** One of each $\frac{12}{26} \bullet \frac{14}{25} + \frac{14}{26} \bullet \frac{12}{25} = 51.59\ \%$

M

Name	x	$\overline{x}$	**s**	$z = \frac{x - \overline{x}}{s}$
Joe	**32**	**15**	**9**	**1.889**
Moe	**54**	**24**	**12**	**2.500**
Nielo	**66**	**22**	**16**	**2.750**
April	**77**	**18**	**25**	**2.360**
Max	**38**	**17**	**11**	**1.909**

a) Nielo
b) Joe
c) Moe, Nielo, April

M1

Name	x	$\overline{x}$	**s**	$z = \frac{x - \overline{x}}{s}$
a	**144**	**128**	**34**	**0.471**
b	**90**	**86**	**18**	**0.222**
c	**18**	**15**	**5**	**0.600**

Student C has the highest relative score.

N.

1. Systematic **2.** Cluster. **3.** Convenience. **4.** Cluster. **5**. Stratified

6. Convenience. **7.** Systematic. **8**. Random. **9**. Stratified **10**. Cluster