11. Twenty high school students took an examination and received the following scores:
70, 60, 75, 68, 85, 86, 78, 72, 82, 88, 88, 73, 74, 79, 86, 82, 90, 92, 93, 73
Determine what percent of the students scored within one standard deviation of the mean. Do the results of the examination approximate a normal distribution? Justify your answer.

12. Mrs. Ramírez is a real estate broker. Last month, the sale prices of homes in her area approximated a normal distribution with a mean of $150,000 and a standard deviation of $25,000.
A house had a sale price of $175,000. What is the percentile rank of its sale price, to the *nearest whole number*? Explain what that percentile means.
Mrs. Ramírez told a customer that most of the houses sold last month had selling prices between $125,000 and $175,000. Explain why she is correct.

MR12 Worksheet (Normal Distribution)

1. The mean of a normally distributed set of data is 56, and the standard deviation is 5. In which interval do approximately 95.4% of all cases lie?
(1) 46–56 (2) 51–61 (3) 46–66 (4) 56–71

2. The amount of juice dispensed from a machine is normally distributed with a mean of 10.50 ounces and a standard deviation of 0.75 ounce. Which interval represents the amount of juice dispensed about 68.2% of the time?
(1) 9.00–12.00 (2) 9.75–11.25 (3) 9.75–10.50 (4) 10.50–11.25

3. In a New York City high school, a survey revealed the mean amount of cola consumed each week was 12 bottles and the standard deviation was 2.8 bottles. Assuming the survey represents a normal distribution, how many bottles of cola per week will approximately 68.2% of the students drink?
(1) 6.4 to 12 (2) 9.2 to 14.8 (3) 6.4 to 17.6 (4) 12 to 20.4

4. On a standardized test, the distribution of scores is normal, the mean of the scores is 75, and the standard deviation is 5.8. If a student scored 83, the student's score ranks
(1) below the 75th percentile (2) between the 75th percentile and the 84th percentile
(3) between the 84th percentile and the 97th percentile (4) above the 97th percentile

5. Scores on PSAT range from 20 to 80. For a certain population of students, the mean is 52 and the standard deviation is 9.
a) How many standard deviations from the mean for a) 49 b) 56 c) 64 d) 65?

6. In a standardized test with a normal distribution of scores, the mean was 75 and the standard deviation was 5.2. Which score could be expected to occur less than 5% of the time?
a) 95 b) 85 c) 75 d) 65

7. In the diagram, the shaded region represents approximately 95% of the scores on a standardized test. If these scores ranged from 78 to 92, what would be the standard deviation?

8. In a standardized test with a normal distribution of scores, the mean is 63 and the standard deviation is 5. Which score could be expected to occur most often?
a) 45 b) 55 c) 65 d) 74

9. Battery lifetime is normally distributed for large samples. The mean lifetime is 500 days and the standard deviation is 61 days. Approximately what percent of batteries have lifetimes *longer than* 561 days?
(1) 16% (2) 68% (3) 34% (4) 84%

10. The national mean for verbal scores on an exam was 428 and the standard deviation was 113. Approximately what percent of those taking this test had verbal scores between 315 and 541?
(1) 68.2% (3) 38.2% (2) 52.8% (4) 26.4%

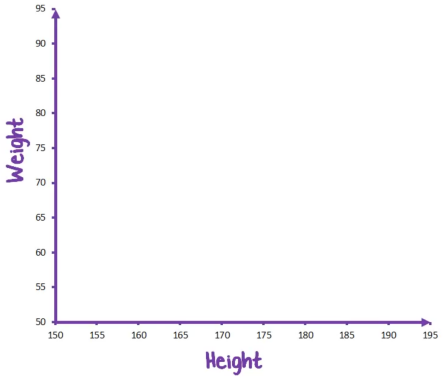
11. Twenty high school students took an examination and received the following scores:
70, 60, 75, 68, 85, 86, 78, 72, 82, 88, 88, 73, 74, 79, 86, 82, 90, 92, 93, 73
Determine what percent of the students scored within one standard deviation of the mean. Do the results of the examination approximate a normal distribution? Justify your answer.

12. Mrs. Ramírez is a real estate broker. Last month, the sale prices of homes in her area approximated a normal distribution with a mean of $150,000 and a standard deviation of $25,000.
A house had a sale price of $175,000. What is the percentile rank of its sale price, to the *nearest whole number*? Explain what that percentile means.
Mrs. Ramírez told a customer that most of the houses sold last month had selling prices between $125,000 and $175,000. Explain why she is correct.

CORE (Data Analysis) Tutorial 17:

Scatterplots

Name	Height (cm)	Weight (kg)
Fred	190	82
Lucy	168	61
Jill	175	68
Li	188	91
Harry	175	85
Gertrude	151	59
Peggy	160	65



4 Characteristics of Scatterplots to look for

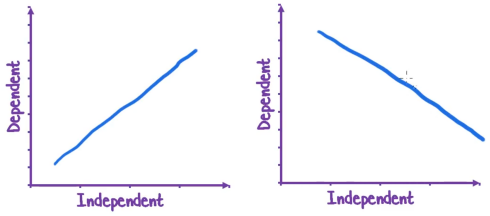
1)

2)

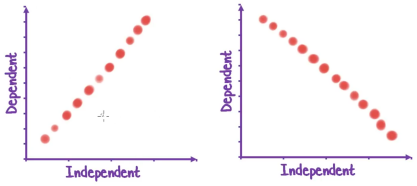
3)

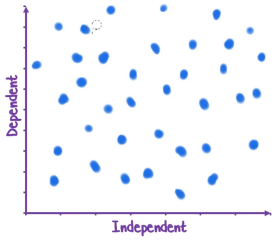
4)

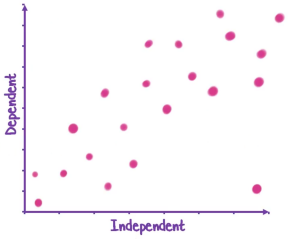
DIRECTION

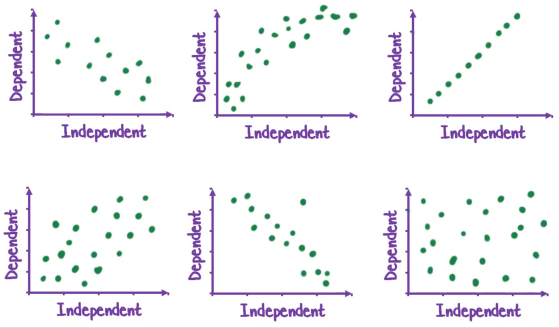


FORM









9.1 Constructing Scatter Plots

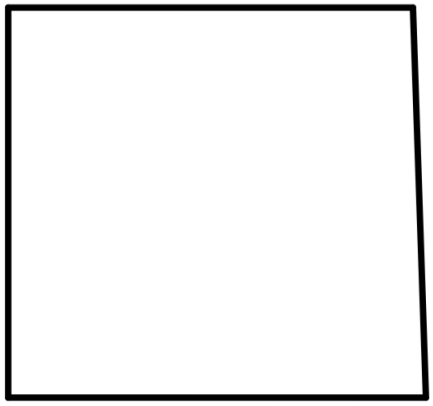
A **scatter plot** is a plot on the coordinate plane used to compare two sets of data and look for a correlation between those data sets. An **association** is a relationship or dependence between data. For example, the price of oil and the price of gasoline have a strong association. The daily price of oil and the number of penguins swimming in the ocean on that day most likely have no association at all. However, to find this association we need to make a scatter plot.

Start with the Data

Before we can make a scatter plot, we need two sets of data that we want to compare. For example, we might compare the number of letters in a student's first name and their math grade. Do people with shorter names tend to score higher in math? Do people with the lowest grades have longer names? These are questions of relationship, or correlation, that we can explore with a scatter plot once have some data. That data set might look like this:

Name	Nichole	Josiah	Kame	Gungar	Roberto	Frank	John	Herman	Sami	Daimon
Letters	7	6	4	6	7	5	4	6	4	6
Grade	58	83	61	70	31	76	81	70	72	57

Name	Yolina	Johanne	Karolinea	Kurt	Addison	Ian	Dennis	Ophelia	Kristina	Bradford
Letters	6	7	9	4	7	3	6	7	8	8
Grade	77	90	87	83	76	78	87	87	80	41



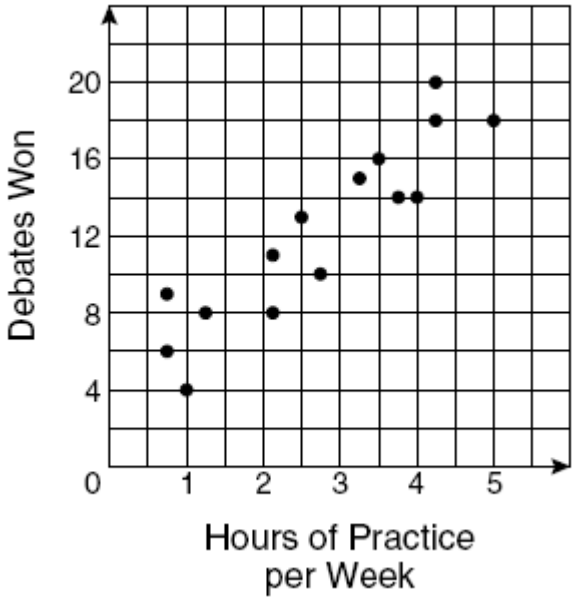
To input data into the STAT list editor:

- Enter STAT edit mode by pressing [STAT] [1].
- Enter the data in the L1 and L2 lists, pressing [ENTER] after each entry.
- Press [2nd] [MODE] to QUIT and return to the home screen.

Graph the scatter plot:

- Press [2nd] [Y=] to access the STAT PLOT editor.
- Press [ENTER] to edit Plot1.
- Press [ENTER] to turn ON Plot1.
- Scroll down and highlight the scatter plot graph type (first option in the first row). Press [ENTER] to select the scatter plot graph type.
- Scroll down and make sure Xlist: is set to L1 and Ylist: is set to L2. To input L1, press [2nd] [1]. To input L2, press [2nd] [2].

The coaches of a group of debate teams answered a survey about hours of debate, team practice and number of team wins. The graph shows the results of this survey.



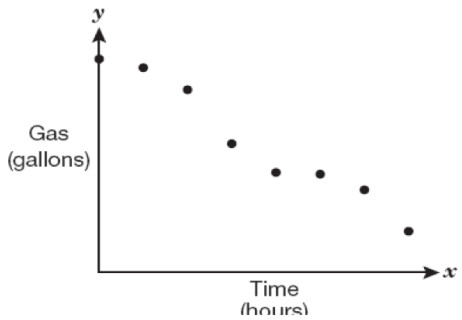
1. **The scatterplot indicates which of the following?**

   A a positive correlation
   B a negative correlation
   C no correlation
   D a parallel correlation

2. **Based on these results, if a team practices 4 hours per week next season, which is the best estimate of the number of debates the team can expect to win?**

   A 20
   B 16
   C 12
   D 1

3. **Josie and some of her friends rode motorcycles all day on Saturday. She made a table that showed the number of gallons of gas remaining at the end of each hour. The scatterplot below shows the gas that remained in terms of the hours that had passed.**



Which of the following describes the correlation between the gas that remained and the hours that had passed?

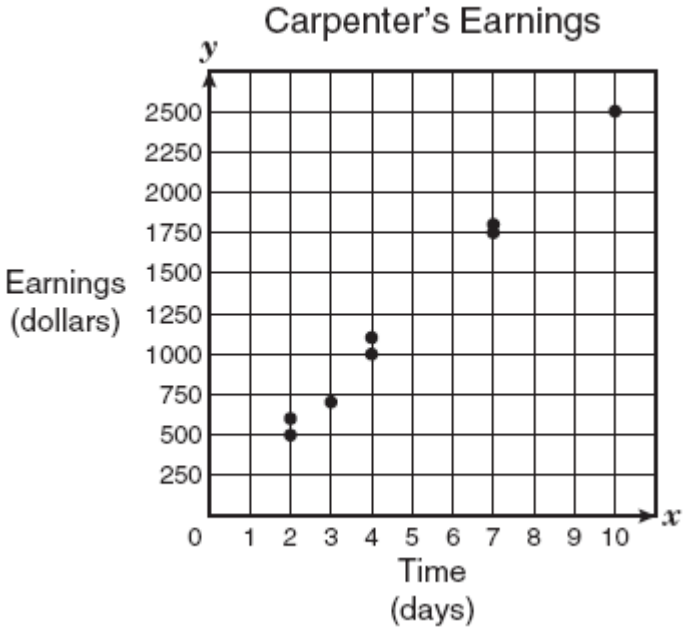
A Positive correlation
B No correlation
C Negative correlation
D Undefined correlation

---

4. **Teresa records the ages and weights of 12 children in her neighborhood. If she records this data in a scatterplot, what type of relationship will she most likely see?**

A Positive correlation
B Negative correlation
C No correlation
D Constant correlation

A carpenter recorded the amount of money he earned for different jobs and the amount of time he spent on each job. The data are shown in the scatterplot below:



5. **Based on the data, which best represents the amount of money the carpenter would earn from a job that took 5 days to complete?**

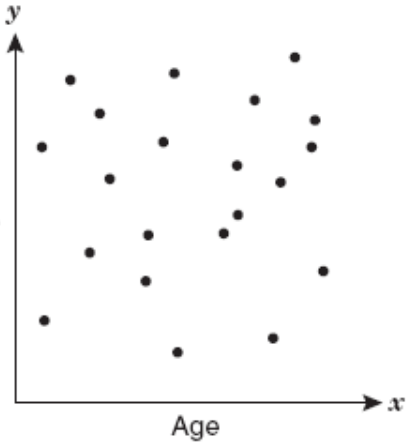
   A $2500
   B $500
   C $1250
   D $200

6. **According to the graph above, the relationship between the carpenter's earnings and the amount of time he spent on jobs is:**

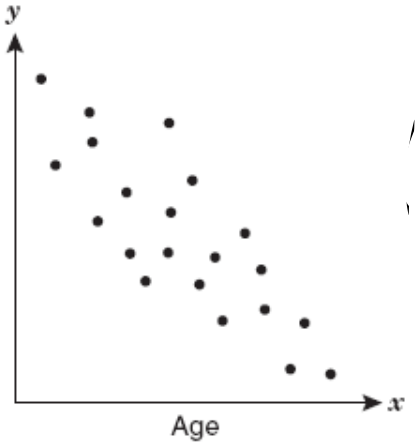
   A a positive correlation
   B a negative correlation
   C no correlation
   D multiple correlations

7. **As the age of a car increases, its value decreases. Which scatterplot best represents this relationship?**

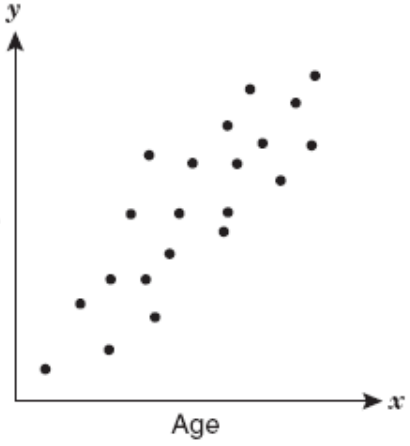
A



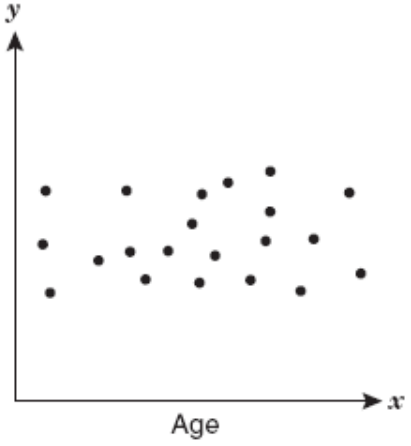
C



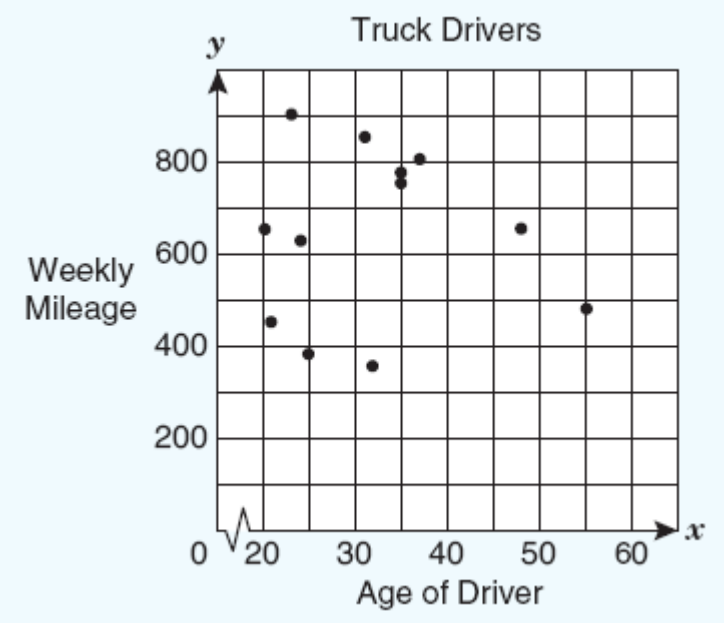
B



D



8. **A trucking company keeps track of the number of miles each of its drivers logs each week. The scatterplot below shows the relationship between a driver's age and the number of miles the driver drove last week.**



Describe the correlation between the number of miles a driver logs and the drivers age.

A There is a positive correlation between the number of miles a driver logs and the drivers age.

B There is a negative correlation between the number of miles a driver logs and the drivers age.

C There is no correlation between the number of miles a driver logs and the drivers age.

D There is a parallel correlation between the number of miles a driver logs and the drivers age.

MULTIPLE CHOICE QUESTIONS

In the following multiple choice questions, circle the correct answer.

1. Which of the following provides a measure of central location for the data?
   a. standard deviation
   b. mean
   c. variance
   d. range

2. A numerıcal value used as a summary measure for a sample, such as sample mean, is known as a
   a. population parameter
   b. sample parameter
   c. sample statistic
   d. population mean

3. Since the population size is always larger than the sample size, then the sample statistic
   a. can never be larger than the population parameter
   b. can never be equal to the population parameter
   c. can be smaller, larger, or equal to the population parameter
   d. can never be smaller than the population parameter

4. μ is an example of a
   a. population parameter
   b. sample statistic
   c. population variance
   d. mode

6. The variance of a sample of 169 observations equals 576. The standard deviation of the sample equals

   a. 13
   b. 24
   c. 576
   d. 28,461

7. The median of a sample will always equal the
   a. mode
   b. mean
   c. 50th percentile
   d. all of the above answers are correct

8. The median is a measure of
   a. relative dispersion
   b. absolute dispersion
   c. central location
   d. relative location

9. The 75th percentile is referred to as the
   a. first quartile
   b. second quartile
   c. third quartile
   d. fourth quartile

10. The *p*th percentile is a value such that at least *p* percent of the observations are
    a. less than or equal to this value
    b. less than this value
    c. more than or equal to this value
    d. more than this value

11. The difference between the largest and the smallest data values is the
    a. variance
    b. interquartile range

c. range
d. coefficient of variation

12. The first quartile
a. contains at least one third of the data elements
b. is the same as the 25th percentile
c. is the same as the 50th percentile
d. is the same as the 75th percentile

13. Which of the following is not a measure of central location?
a. mean
b. median
c. variance
d. mode

14. If a data set has an even number of observations, the median
a. cannot be determined
b. is the average value of the two middle items
c. must be equal to the mean
d. is the average value of the two middle items when all items are arranged in ascending order

15. Which of the following is a measure of dispersion?
a. percentiles
b. quartiles
c. interquartile range
d. all of the above are measures of dispersion

16. The most frequently occurring value of a data set is called the
a. range
b. mode
c. mean
d. median

17. The interquartile range is
a. the 50th percentile
b. another name for the variance
c. the difference between the largest and smallest values
d. the difference between the third quartile and the first quartile

)

20. When data are positively skewed, the mean will usually be
    a. greater than the median
    b. smaller than the median
    c. equal to the median
    d. positive

21. Which of the following is **not** a measure of dispersion?
    a. the range
    b. the 50th percentile
    c. the standard deviation
    d. the interquartile range

22. The interquartile range is used as a measure of variability to overcome what difficulty of the range?
    a. the sum of the range variances is zero
    b. the range is difficult to compute
    c. the range is influenced too much by extreme values
    d. the range is negative

25. The measure of dispersion that is influenced most by extreme values is
    a. the variance
    b. the standard deviation
    c. the range
    d. the interquartile range

27. The descriptive measure of dispersion that is based on the concept of a deviation about the mean is
    a. the range
    b. the interquartile range
    c. the absolute value of the range
    d. the standard deviation

28. The measure of location which is the most likely to be influenced by extreme values in the data set is the
    a. range
    b. median
    c. mode
    d. mean

29. The most important statistical descriptive measure of the location of a data set is the
    a. mean
    b. median

c. mode
d. variance

33. If two groups of numbers have the same mean, then
a. their standard deviations must also be equal
b. their medians must also be equal
c. their modes must also be equal
d. None of these alternatives is correct

35. Which of the following symbols represents the standard deviation of the population?
a. σ^2
b. σ

c. μ
d. $\bar{x}$

36. Which of the following symbols represents the mean of the population?
a. σ^2
b. σ
c. μ
d. $\bar{x}$

37. Which of the following symbols represents the variance of the population?
a. σ^2
b. σ
c. μ
d. $\bar{x}$

38. Which of the following symbols represents the size of the population?
a. σ^2
b. σ
c. μ
d. N

39. Which of the following symbols represents the mean of the sample?
a. σ^2
b. σ
c. μ
d. $\bar{x}$

40. Which of the following symbols represents the size of the sample
a. σ^2
b. σ
c. N
d. n

41. The symbol σ is used to represent
a. the variance of the population
b. the standard deviation of the sample
c. the standard deviation of the population

d. the variance of the sample

42. The symbol σ^2 is used to represent
a. the variance of the population
b. the standard deviation of the sample
c. the standard deviation of the population
d. the variance of the sample

48. When the smallest and largest percentage of items are removed from a data set and the mean is computed, the mean of the remaining data is
    a. the median
    b. the mode
    c. the trimmed mean
    d. any of the above
    Answer: c

49. In a five number summary, which of the following is **not** used for data summarization?
    a. the smallest value
    b. the largest value
    c. the mean
    d. the 25th percentile

51. Which of the following is **not** a measure of dispersion?
    a. mode
    b. standard deviation
    c. range
    d. interqurtile range

52. Since the mode is the most frequently occurring data value, it
    a. can never be larger than the mean
    b. is always larger than the median
    c. is always larger than the mean
    d. None of these alternatives is correct.