1. Specify the data type our variable of interest is:
   - What is his education level?(primary, high school, UG, PG)
2. What are the quartiles for the following set of numbers?
   8, 11, 20, 10, 2, 17, 15, 5, 16, 15, 25, 6
3. Create a histogram for the following test scores: 99, 97, 94, 88, 84, 81, 80, 77, 71, 25.
4. Create required plot to explain the data given below.

Sport	Boys	Girls	Total
Basketball	15	20	35
Volley ball	25	10	35
Badminton	15	15	30

Given the data set Pizza sales per day

Place	No of orders		Online Ordering		Age group of 15-30		Average time to sale(in min)		
	Females	Males	Females	males	Females	males	Small	Medium	large
A	400	350	300	300	200	200	15	20	15
B	100	250	70	100	80	200	17	22	24
C	70	140	50	100	30	100	20	21	17
D	387	150	300	70	250	110	18	15	26
E	500	600	250	250	300	400	20	25	30

5. What is the percentage of females among the age group 15-30 ordering pizza at Place D? Show graphically.
6. What is the best place to get the maximum sales?
7. Which gender is purchasing maximum number of pizzas?
8. What is the average time taken to sell a Pizza?

9. An ice cream shop keeps track of how much ice cream they sell versus the temperature on that day. The two Variables are Ice cream sales and temperature.

   Here are their figures for the last 12 days: Find any relationship?

Ice cream sales vs temperature	
Temperature	Ice cream sales
14.2°	$215
16.4°	$325
11.9°	$185
15.2°	$332
18.5°	$406
22.1°	$522
19.4°	$412
25.1°	$614
23.4°	$544

18.1°	$421
22.6°	$445
17.2°	$408

10. A town has 15 neighborhoods. If you interviewed everyone living in one particular neighborhood, would you be interviewing a population or a sample from the town? Would this be a random sample?

11. For the following stock price one-year percentage changes, plot the data and identify any outliers. Find the mean, median and variance.

Intel	-6.9%
AT&T	46.5
General Electric	12.1
ExxonMobil	20.7
Microsoft	16.9
Pfizer	17.2
Citigroup	16.5

12. Refer to the box plot below to answer the questions.
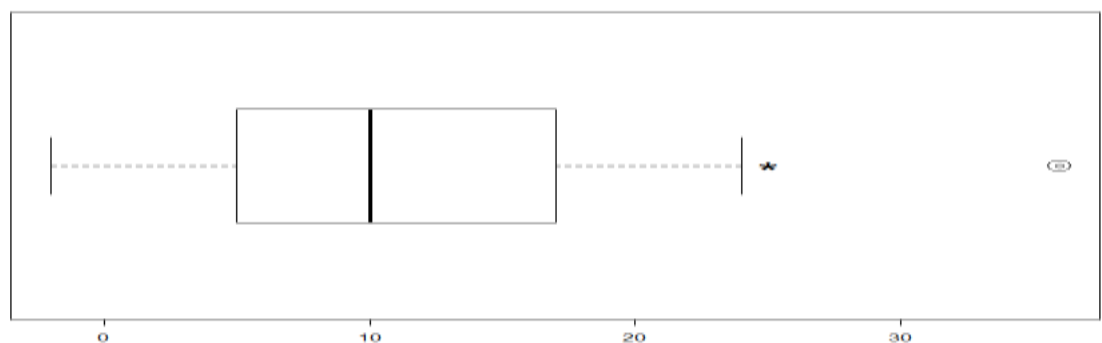
1. What is the interquartile range for this data set?
2. What can you say about the skewness of this data set?
3. For this data set, the value of 9.5 is more likely to be (choose one)
   a. The first quartile rather than the median.
   b. The median rather than the first quartile.

        c.   The mean rather than the mode.

        d.   The mode rather than the mean.

4.   If a data point that was originally 26 is changed to 0, how would the boxplot be effected?

Solved Examples:

1.   Specify the data type our variable of interest is:
   - Is he present in the class? (yes or no) --- It is binary nominal categorical variable.

2.   π = 3.14159265358979323846264343383...
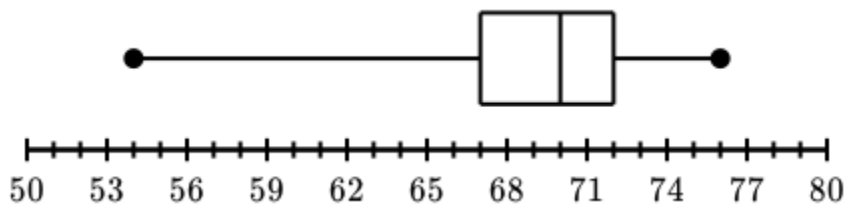   What is the median of the first 12 digits of π?

Sol: Median of 12 numbers is mean of 6th and 7th number. Therefore mean of 9 and 2 which is 5.5.

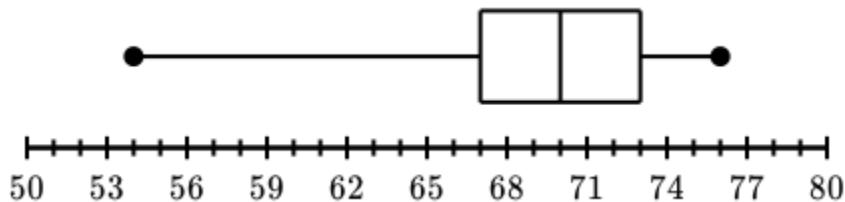3.   The data below represents the number of desks on each floor of Texter Corporate.

    54,60,65,66,67,69,70,72,73,75,76

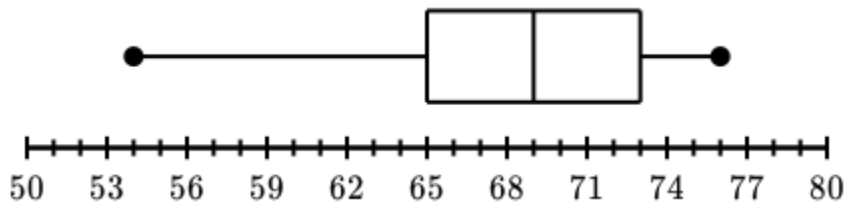   Which box plot correctly summarizes the data?

    a)

    b)

    c)

The media of the data given is 69 and 1st quartile is 65.5 and 3rd quartile 72.5. So the boxplot c is rightly explaining the data.

The media of the data given is 69 and 1st quartile is 65.5 and 3rd quartile 72.5. So the boxplot c is rightly explaining the data.

Study the following line graph and answer the questions.

Exports from Three Companies Over the Years (in Rs. crore)
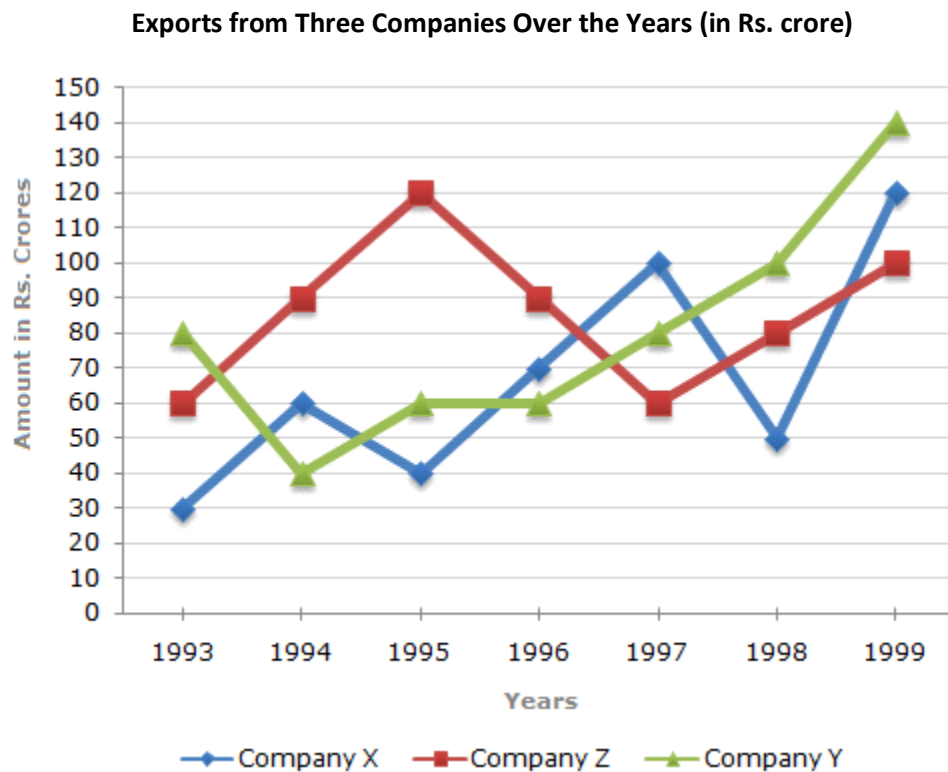
4. For which of the following pairs of years the total exports from the three Companies together are equal?

A. 1995 and 1998

B. 1996 and 1998

C. 1997 and 1998

D. 1995 and 1996

Correct Answer: D

Explanation:

Total exports of the three Companies X, Y and Z together, during various years are:

In 1993 = Rs. (30 + 80 + 60) crores = Rs. 170 crores.

In 1994 = Rs. (60 + 40 + 90) crores = Rs. 190 crores.

In 1995 = Rs. (40 + 60 + 120) crores = Rs. 220 crores.

In 1996 = Rs. (70 + 60 + 90) crores = Rs. 220 crores.

In 1997 = Rs. (100 + 80 + 60) crores = Rs. 240 crores.

In 1998 = Rs. (50 + 100 + 80) crores = Rs. 230 crores.

In 1999 = Rs. (120 + 140 + 100) crores = Rs. 360 crores.

Clearly, the total exports of the three Companies X, Y and Z together are same during the years 1995 and 1996.

5. Average annual exports during the given period for Company Y is approximately what percent of the average annual exports for Company Z?

A. 87.12%

B. 89.64%

C. 91.21%

D. 93.33%

Correct Answer: D

Explanation:

Analysis of the graph: From the graph it is clear that

The amount of exports of Company X (in crore Rs.) in the years 1993, 1994, 1995, 1996, 1997, 1998 and 1999 are 30, 60, 40, 70, 100, 50 and 120 respectively.

The amount of exports of Company Y (in crore Rs.) in the years 1993, 1994, 1995, 1996, 1997, 1998 and 1999 are 80, 40, 60, 60, 80, 100 and 140 respectively.

The amount of exports of Company Z (in crore Rs.) in the years 1993, 1994, 1995, 1996, 1997, 1998 and 1999 are 60, 90,, 120, 90, 60, 80 and 100 respectively.

Average annual exports (in Rs. crore) of Company Y during the given period

= (80 + 40 + 60 + 60 + 80 + 100 + 140)/7 = 560/7= 80.

Average annual exports (in Rs. crore) of Company Z during the given period

= (60 + 90 + 120 + 90 + 60 + 80 + 100)/7 = 600/7.

Therefore Required percentage = [(80/600)/7x 100]% ~= 93.33%.

6. In which year was the difference between the exports from Companies X and Y the minimum?

      A.1994

      B.1995

      C.1996

      D.1997

Correct Solution: C

The difference between the exports from the Companies X and Y during the various years are:

In 1993 = Rs. (80 - 30) crores = Rs. 50 crores.

In 1994 = Rs. (60 - 40) crores = Rs. 20 crores.

In 1995 = Rs. (60 - 40) crores = Rs. 20 crores.

In 1996 = Rs. (70 - 60) crores = Rs. 10 crores.

In 1997 = Rs. (100 - 80) crores = Rs. 20 crores.

In 1998 = Rs. (100 - 50) crores = Rs. 50 crores.

In 1999 = Rs. (140 - 120) crores = Rs. 20 crores.

Clearly, the difference is minimum in the year 1996.

7. What was the difference between the average exports of the three Companies in 1993 and the average exports in 1998?

      A. Rs. 15.33 crores

      B. Rs. 18.67 crores

      C. Rs. 20 crores

      D. Rs. 22.17 crores

Correct Solution: C

Explanation:

Average exports of the three Companies X, Y and Z in 1993

  = Rs. [(30 + 80 + 60)/3]crores = Rs.(170/3)crores.

Average exports of the three Companies X, Y and Z in 1998

= Rs. [(50 + 100 + 80)/3]crores = Rs.(230/3)crores.

Difference = Rs.[(230/3)-(170/3)]crores = Rs.(60/3)crores

= Rs. 20 crores.

8. In how many of the given years, were the exports from Company Z more than the average annual exports over the given years?

     A.2

     B.3

     C.4

     D.5

Correct Solution: C

Explanation:

Average annual exports of Company Z during the given period

  = (60 + 90 + 120 + 90 + 60 + 80 + 100)/7

  = Rs.(600/7)crores = Rs. 85.71 crores.

From the analysis of graph the exports of Company Z are more than the average annual exports of Company Z (i.e., Rs. 85.71 crores) during the years 1994, 1995, 1996 and 1999, i.e., during 4 of the given years.

The table below shows the dividend yields of six companies in the New

York Stock Exchange energy sector.

Company	Dividend Yield July 2012 (%)
BP	4.8
Chevron	3.41
Exxon Mobil	2.66
PetroChina	3.5
Petroleo Brasiliero	1.2
Royal Dutch Shell	4.3

9. Find the sample mean.

Solution:

The data in this example are the dividend yields for each company. The sample mean is

$$\bar{X} = \frac{\sum X}{n}$$

= (4.80+ 3.41+ 2.66 +3.50+ 1.20+ 4.30)/6

=3.312

10. Find the variance.

Solution:

Use a table to compute the differences from the mean and the squared differences from the mean.

X	(X-mean)	(X-mean)^2
4.8	1.488333	2.215136111
3.41	0.098333	0.009669444
2.66	-0.65167	0.424669444
3.5	0.188333	0.035469444
1.2	-2.11167	4.459136111
4.3	0.988333	0.976802778

Variance =$\sigma^2 = \frac{\sum(X-\bar{X})^2}{n}$

=8.121/(6)

=1.3535

11. Find the standard deviation.

Solution:

The standard deviation is the square root of the variance,

$\sigma = \sqrt{\sigma^2}$ = $\sqrt{1.3535}$

=1.16