STRATEGIES TO CONQUER PROBLEM SOLVING AND DATA ANALYSIS SECTION ON SAT MATH

BY- DEVYANSH BHANSALI
Student of La Martiniere for Boys, Kolkata

SYLLABUS(12-15% Weightage with 5-7 questions asked):

- Ratios, rates, proportional relationships, and unit conversion

- Percentages

- Distributions and measures of center and spread

- Data representations and Scatterplots

- Probability and relative frequency

- Inference from sample statistics and margin of error

- Evaluating statistical claims: observational studies and experiments

Ratio, Rates, Proportional relationships and units:

- Single-Step Unit Conversion:

Conversion factor : $\frac{\text{desired units}}{\text{starting units}}$. We must relate the desired and starting units

Starting units x$\frac{\text{Desired units}}{\text{Starting units}}$ = Desired units , where the conversion factor must be abstractly, not mathematically, equal to 1.

Eg1. 45 mins to hour : 45 ~~mins~~ x $\frac{1 \text{ hour}}{60 \text{ ~~mins~~}}$ = $\frac{3}{4}$ hours=0.75 hours

Eg2. 35 gm to ounces(oz)(note:1gm=0.04 oz) : 35 ~~gm~~ x $\frac{0.04 \text{ oz}}{1 \text{ ~~gm~~}}$ = 1.4 oz

Eg3.14feet(ft)to yards(yds)(note: 3 feet=1 yard): 14 ~~feet~~ x $\frac{1 \text{ yard}}{3 \text{ ~~feet~~}}$ = $\frac{14}{3}$ yd=4.7 yd

Abstractly, 1 hour is the same as 60 mins,0.04 oz is the same as 1 gm and 1 yard is the same as 3 feet. However, mathematically they are 1/60 ,0.04 and 1/3. We use this technique to cancel out the starting units so that we are left with the value corresponding to the desired unit.

- Multiple-step unit conversion:

$$\text{Starting units} \times \frac{\text{Linking units}}{\text{Starting units}} \times \frac{\text{Desired units}}{\text{Linking units}} = \text{Desired units}$$

EG1: A sloth travelled 3.2 meters in 10 minutes. What was the sloth's average speed in kilometer per hour for 10 mins? Round your answer to nearest hundredth

$$\frac{3.2 \text{ m}}{10 \text{ min}} \times \frac{1 \text{ km}}{1000 \text{ m}} \times \frac{60 \text{ min}}{1 \text{ hour}} = 0.02 \text{ km/1 hour.}$$

Here, we needed answer in the form $\frac{km}{hour}$, . so we rightfully took our starting units 3.2 m/ 10 minutes and first changed m to km by multiplying $\frac{1 \text{ km}}{1000 \text{ m}}$ and then changed minute to hour in denominator by multiplying with $\frac{60 \text{ min}}{1 \text{ hour}}$

EG2: 650 mm to feet (note: 1 cm=0.39 inches and 1 cm=10mm)

$$650 \text{ mm} \times \frac{1 \text{ cm}}{10 \text{ mm}} \times \frac{0.39 \text{ inch}}{1 \text{ cm}} \times \frac{1 \text{ feet}}{12 \text{ inch}} = 2.1 \text{ feet}$$

- Rates:

IDENTIFICATION:

Recognize the following words:

-at the rate of or at an average speed of

-"per" or "for every"

-repetition of units

Where,

 Ta= Time taken by a working alone to complete the job

 Tb= Time taken by b working alone to complete the job

 Ta+b= Total time taken by a and b working together to complete the job

FORMULAS:

Rate=$\dfrac{\text{Work}}{\text{Time}}$

Distance=Speed x time and Distance=rate x time

Combined work by two individuals a and b=

$$\frac{1}{Ta} + \frac{1}{Tb} = \frac{1}{Ta+b} \text{ ,}$$

Q1) How far can Vedant go if he drives for 4 hours <u>at a rate of </u>30 miles per hour?

A1)Distance=Rate x time = $30 \frac{miles}{hour}$ x 4 ~~hour~~ = 120 miles

Q2)How long will Vedant take to travel 100 miles in total <u>at the rate of </u>20 mile per hour(mph) ?

A2)time= $\frac{Distance}{Rate}$ =100/20 hour = 5 hours

Q3)One car A is travelling <u>at an average speed </u>of 24 miles <u>per</u> hour and another car B is going <u>at a rate of </u>30 miles per hour. If both cars make a 360 mile trip, how much longer will the slower car take?

A3)Time taken by car A: Time taken by car B:

time= $\frac{Distance}{Rate} = \frac{360}{24} = 15$ hour   time= $\frac{Distance}{Rate} = \frac{360}{30}$ = 12 hour

The extra time taken by slower car(B)= 15-12 hour = 3 hour

Q4)Train X travelled 180 miles in 3 hours and train Y travelled 225 miles in 5 hours in the opposite direction. How fast are the two trains separating?

A4)Rate of travel by X in -> direction = $\frac{180}{3}$mph=60 mph

Rate of travel by Y in <- direction = $\frac{225}{5}$mph=45 mph

 Train Y<---------------|-------------->Train X
45 mph   60 mph
Thus, total separation = 45+60 mph = 105 mph

Q5) A city has an area of 85 square miles and a population of 630000 people. Rounded to the nearest whole number, what is the population density in people <u>per</u> square mile?
A5) Population density is the no. of people living per unit of area

Population density=$\frac{630000 \text{ people}}{85 \text{ square miles}}$

= 7411.76 people per square mile
= 7412 people per square mile

Q6) If Jimmy can paint a room in 3 hours and Corey can paint the same room in 4 hours, How long will it take them to paint the room together?

A6) In 1 hour, Jimmy can paint 1/3 of the room and Corey can paint 1/4 of the room.

Their combined rate $=\dfrac{1}{t} = \dfrac{1}{3} + \dfrac{1}{4}$

$$\frac{1}{t} = \frac{7}{12}$$

$$\Rightarrow t = \frac{12}{7} \text{ hours}$$

Q7)It was observed that a liquid evaporates at a rate of 7.5 x 10^-6 cm^3/s. How much time in <u>days</u> will 72 cm^3 of this liquid will take to evaporate?(*Challenge question*)

Rate=7.5 x 10^-6 $\dfrac{cm\char`^3}{\cancel{s}} \dfrac{60 \cancel{s}}{1 \cancel{min}} \dfrac{60 \cancel{min}}{1 \cancel{hour}} \dfrac{24 \cancel{hours}}{1 \text{ day}}$

= 0.648 cm^3/day

Time= $\dfrac{72 \cancel{cm\char`^3}}{0.648 \cancel{cm^3}/\text{day}} = 111.11 \text{ days}$

- Ratio and proportion:

-A ratio is a comparison of two quantities. The ratio of a to b can be expressed

as a:b or $\frac{a}{b}$. In our final answer, we must simplify the determined ratio to the lowest terms possible.

-Two common types of ratios are part-to-part and part-to-whole.

The ratio of girls to boys=4:3 in a class is a part-to-part ratio but the ratio of girls to no. of students=4:7 is a part-to-whole ratio. It compares the amount of one part to the sum of all parts.

Since all the parts need to add up to the whole, part-to-part and part-to-whole ratios often imply each other and this means we can use the ratios we're provided to find whichever ratios we need to solve a problem!

-A proportion is an equality of two ratios. If we know a ratio and want to apply that ratio to a different scenario or population, we can use proportions to set up equivalent ratios and calculate any unknown quantities.

-Steps To use a proportional relationship to find an unknown quantity:

i)Write an equation using equivalent ratios.

ii)Plug in known values and use a variable to represent the unknown quantity.

iii)Solve for the unknown quantity by isolating the variable.

Ratio and proportion mixed questions:

Q1)If there are 15 apples, 3 bananas and 2 strawberries in a basket, What is the Ratio Of apple to fruit?

A1) Ratio of apple to fruit = $\frac{\text{No.of apples}}{\text{No.of fruits}} = \frac{15}{20}$ = 3/4 or 75% of fruits in the basket are apples

Q2)If 7 eggs cost $5, How many can I buy if I have $10?

A2)Let x = no. of eggs I can buy with $8.

$$\frac{7 \text{ eggs}}{5 \text{ dollars}} = \frac{x \text{ eggs}}{10 \text{ dollars}} \Rightarrow x = 14 \text{ eggs}$$

Q3) There are 340 students at an academy. If the student to teacher ratio is 17 : 2 . How many teachers are there?

A3) We have been given a student is to teacher ratio and the no. of students so we can easily set up a proportional relationship expression.

Let x=No. of teachers at the academy.

$$\frac{17}{2} = \frac{340}{x} => x = \frac{(340)(2)}{17} \Rightarrow x = 40 \text{ teachers}$$

Q4) There are two oxygen atoms and one carbon atom in one carbon dioxide molecule. How many oxygen atoms are in 78 carbon dioxide molecules?

A4) $\dfrac{2 \text{ Oxygen atoms}}{1 \text{ carbon di oxide molecule}} = \dfrac{x \text{ Oxygen atoms}}{78 \text{ carbon di oxide molecules}}$

x=(78)(2)=156 oxygen atoms

Q5) Building A is 140 feet tall, and Building B is 85 feet tall. The ratio of the heights of Building A to Building B is equal to the ratio of the heights of Building C to Building D .If Building C is 90 feet tall, what is the height of Building D to the nearest foot?

A5) $\dfrac{140}{85} = \dfrac{90}{x}$ where x is the height of building D

$x = \dfrac{(90)(85)}{(140)}$ => x=54.6 ≈ 55 foot

Q6) if 7 pounds of plums are needed to make 8 rolls of fruit leather and considering that each roll if fruit leather require the same amounts of plums, how many pounds of plums are required to make 20 rolls of fruit leather?

$\dfrac{7}{8} = \dfrac{x}{20}$ => x=140/8=17.5. Thus,17.5 pounds of plums are required to make 20 rolls of fruit leather.

Percentages:

-A percentage is a ratio out of 100 that represents a part-to-whole relationship. 10%=10/100=0.1

Percent(%) means parts per hundred : p%=$\dfrac{p}{100}$

%=$\dfrac{Part}{Whole}$ x 100

NOTE- The part and the whole must be correctly identified and part is not necessarily the smaller of the two.

The correct way to interpret the following words :
a is what % of b
what=(x)what we are required to find   is=equal to
%= Divide by 100   of=multiply by

$\dfrac{x}{100}=\dfrac{a}{b}$ where x is the required answer in percent.

NOTE- Most likely on the digital SAT, it will be mentioned to submit your answer without the % sign so do not make the mistake of submitting your answer with it as it might lead to an incorrect answer as the questions are corrected using technology and not manually.

-Since all parts of a whole should add up to 100% , we can also use percentages to determine the value of any missing parts. Suppose a coed class has 65% boys. It would indicate that the class has (100-65)%=35% girls. Or if a box containing chocolate, candies and biscuit has 35% chocolates and 30% biscuits. It implies that (100-35-30)% = 35% of the box is filled by candies.

-To calculate percent change, we can use the following formula:

$$\% \text{ Change} = \frac{\text{New value} - \text{Old value}}{\text{Old Value}} \times 100\%$$

If the % change is +ve, it is a % increase
Otherwise if the % change is –ve, it is a % decrease.

Concept of Multiplying factor:
A multiplying factor in percentage is a number you multiply by the original value to obtain a new value that has been increased or decreased by a specified percentage.
We use it when we have been given a % increase or a % decrease

For a given percentage p,

Increasing by percentage=New value=Original value x $(1+\frac{p}{100})$

Decreasing by percentage=New value=Original value x $(1-\frac{p}{100})$

Questions:
Q1) What is 25% of 60?
A1) $\frac{x}{60} = \frac{25}{100}$ => x =(25 x 60)/100 = 15
Q2)7 is 35% of which number?
A2) $\frac{7}{x} = \frac{35}{100}$ => x=(100 x 7)/35 = 20
To verify we can $\frac{35}{100}$ x 20 and this should amount to 7 and it does.
Q3)10 is what % of 50?
$\frac{10}{50} = \frac{x}{100}$ => x=1000/50

The above three questions deal with the word translations discussed above.
They must be dealt with carefully

Q4) A group of 1300 people took a survey in which they had to select their favorite of the four seasons (winter, spring, summer, and fall). The incomplete table below shows the results:

Season	% of respondents
Winter	16
Spring	22
Summer	
Fall	28

i)What percentage of respondents said summer was their favorite season?

ii)How many respondents said summer was their favorite season?

A4) We know that the full group of survey respondents represent 100%

i)   % of respondents saying summer was their favorite season = (100-29-22-16)% = 34%

ii) No. of respondents saying summer was their favorite season: $34\% = \dfrac{part}{1300} \times 100$

Part=334 x 13 = 442(ans)

Q5)A customer's monthly water bill was \$75.74 but due to a rate increase her monthly bill is now \$79.86. To the nearest tenth of a percent, by what percent did the amount of the customer's water bill increase?

A5) % Change=$\dfrac{\text{New value} - \text{Old value}}{\text{Old Value}}$ x 100%

$$= \dfrac{79.86 - 75.74}{75.74} \text{ x 100\%}$$

=5.43%

=5.4%

Q6) The atomic weight of an unknown element in atomic mass units(amu) is approximately is approximately 20% less than that of calcium. The atomic weight of calcium is 40 amu. Which of the following best approximates the atomic weight in amu of the unknown element?

A6)We shall use the concept of multiplying factor as we have been given a %decrease.

New value = old value x (1- $\dfrac{p}{100}$)

$\Rightarrow$ New value = 40 x (1- $\dfrac{20}{100}$)

= 40 x 0.8

=32 (answer)

Q7)The table above shows a summary of 1800 responses to a survey question. Based on the table, how many of those surveyed get most of their news from either newspapers or the radio?

A7)Total % of people getting news from either newspapers or the radio

= 15%+11%=26%

Thus, no. of respondents = $\frac{26}{100}$ x 1800 = 468 (answer)

You can always cross check your answer. Think of this one as:

[468 is what % of 1800 and the value must come 26%]

$$\frac{468}{1800}=\frac{X}{100}$$

X=46800/1800 = 26. Confirmed.

Where Do People Get Most of Their News?

Source	Percent of those surveyed
Internet/social media	43%
Television	23%
Newspapers	15%
Radio	11%
Others/none of the above	8%

Q8)Sahil would owe $15,500 in taxes each year If he were not eligible for any tax deductions. This year he is eligible for tax reductions that reduce the amount of taxes he owes by $2325. If these tax deductions reduce the taxes Sahil owes this year by d%, What is the value of d.

A8)Since the deductions reduce the original amount of taxes by d%, we can use the equation $\frac{2325}{15500}=\frac{d}{100}$ to find the %decrease

Q9) Juhi has read 90% of the books she owns. If Juhi owns 180 books, how many of her books has she <u>not</u> read?

A9)Juhi has read $\frac{90}{100}$ x 180 books =162 books.   OR Juhi has not read (100-90)%=10% books = 10/100 x 180

Thus she has not read 180-162 = 18 books   =18 books

Q10) Katrina is a botanist studying the production of pears by two types of pear trees. She noticed that type A trees produced 20 % more pears than type B trees. Based on Katrina's observation, if the type A trees produced 144 pears how many more pears did the type B trees produce?

A10)A=B x (1+20/100)

   A=1.2B

   B=144/1.2

   B=120

Q11)Today Ebuka opened a cereal box and ate a bowl of cereal which was 8% of cereal in the entire box. Approximately how many bowls of cereal can Ebuka expect to fill up if he continues to eat the same amount of cereal each time?

A11) The no. of bowls that Ebuka can expect to fill up: $\frac{\text{New value} - \text{Old value}}{\text{Old Value}} = \frac{100 - 8}{8} = \frac{92}{8}$ = 11.5 bowls

One-variable data: distributions and measures of center and spread

-In statistics, we describe the overall pattern of distribution of a quantitative variable set using summary statistics-center, spread and shape of distribution. Shape of the distribution is not in our scope.

Center- Describes the typical or central data points.
Measures of central tendency include mean, median and mode.

Spread-Measures how far the data points are from the center
Measures of spread are standard deviation and range.

Outliers – an observation that is surprisingly different from the rest of the data. It can be treated as an extreme value.
EG- Usain bolt's 100 m timings in a data set containing 100 m timings of high school students.
Mean: Mode-The value having highest frequency
*Point at which whole group of data balances
*Affected by outliers(depends on all values)
*order of listing of data not important
Median:
*Divides the measurements into two halves.   *Order of listing of data very important
*Unaffected by outliers(depends on middle values only)

Standard Deviation:

*Basically tells how close the data points are to the mean

*Affected by outliers

*More the difference between mean and data points, more is The S.D

*Flatter the curve of the graph, more the S.D

*Order of listing of data not important

*If more frequencies of data set are closer to mean, S.D is less.

Range:

*Basically tells how close the data points are to the mean

*Affected by outliers

NOTE- Make Desmos your best friend as you approach questions regarding mean, median and standard deviation.

You can access the functions by the following steps : Desmos < keyboard emoticon < functions < statistics

However, you may require the

Formulas for mean and median in

Some problems.

STATISTICS

mean	median	min
max	quartile	quantile
stdev	stdevp	var

Mean = $\dfrac{\text{Total sum of values}}{\text{No.of values}}$

Sometimes the College Board uses the word 'Average' to imply mean

Also note that on making changes to the data sets, the mean changes

Median:

First arrange the values in either ascending or descending order

Count n=no. of values

If n=odd,

Median=$(\dfrac{n+1}{2})$th term

If n=even,

Median=$\dfrac{\frac{\text{nth term}}{2} + \frac{\text{n+1th term}}{2}}{2}$

Mode is simply the value that occurs the most

Range=Highest value – Lowest value

-You don't need to learn the formula for standard deviation. Just use Desmos otherwise you might lose out on time.

-Variance=(S.D)^2. Higher the variance, the more the S.D

Questions:

Q1)This table shows the number of points earned by three students on two Consecutive exams. What is the <u>average</u> increase in the points earned from the first To the second exam for the three students

Points Earned in a Class		
Student	First exam	Second exam
A	78	87
B	84	98
C	91	95

A1) Increase in points of A:87-78=9

Increase in points of B:98-84=14

Increase in points of C:95-91=4

Mean increase=$\dfrac{9+14+4}{3}=\dfrac{27}{3}=9$

On average, the points earned from the first exam increased by 9 in the second exam for the three students.

Q2)Akshay earned 85,95,93 and 80 points on 4 exams each worth 100 points. How much should he score on a fifth exam, worth 100 points, to get an <u>average</u> of 90?

A2) Let x be the marks Akshay has to score in the fifth exam to get an average of 90 over all exams taken

Mean=$\dfrac{\text{Total sum of observation}}{\text{Total no.of observation}}$

$90=\dfrac{85+95+93+80+x}{5}$ => 353+x=450

x=97. Akshay must score at least a 97 if he wants an average of 90.

Q3)The average of 6,19 and x is 19. What is value of x

Mean= $\dfrac{\text{Total sum of observation}}{\text{Total no.of observation}}$

19= $\dfrac{6+19+x}{3}$ => 25 + x = 57

x=57-25=32   x=32(ans)

Q4)If the average of t and t+2 is x and the average of t and t-2 is y, What is the average of x and y?

A4)x= $\dfrac{t+t+2}{2} = \dfrac{2t+2}{2} = \dfrac{\cancel{2}(t+1)}{\cancel{2}}$ y= $\dfrac{t+t-2}{2} = \dfrac{2t-2}{2} = \dfrac{\cancel{2}(t-1)}{\cancel{2}}$

x=t+1   y=t-1

Avg of x and y= = $\dfrac{t+1+t-1}{2} = \dfrac{2t}{2} = t\text{(ans)}$

Q5) The mean score of 8 players in a basketball game was 14.5 points. If the highest individual score is removed, the mean score of the remaining 7 players becomes 12 points. What was the highest score?

14.5= $\dfrac{\text{Total sum of observation}}{8}$ => total=14.5x8 = 116

Let the highest individual score be x

Mean= $\dfrac{\text{Total sum of observation}-x}{\text{Total no.of observation}}$ =>12= $\dfrac{116-x}{7}$

116-x=84 => x=116-84=32   x=32(ans)

Q6) 10, 18, 4, 15, 3, 21, x. If x is the median of the 7 numbers listed above, which of the following could be the value of x?

A) 5 B) 8 C) 9 D) 14

A6)Arranging in ascending order:3,4,10,15,18,21. If x is the median, for n=6, it should be in the middle of 3rd and 4th term that is 10 and 15. The only option that satisfies this id D=14. (Ans)

Q7)The table below lists the yield in grams of a substance X for first four trials of an experiment. After the fifth trial, the mean yield of substance X for the first five trials was 7.386 g. What was the median for all five trials?

Trial	1	2	3	4
Yield	7.398 g	7.309 g	7.421 g	7.379 g

A7) First we need to take out the yield for the fifth trial.

We know, Mean= $\frac{\text{Total sum of observation}}{\text{Total no.of observation}}$ => 7.386= $\frac{\text{Total sum of observation}}{5}$ =>total=7.386x5=36.93

Let the fifth yield be y => 7.398 + 7.309 + 7.421 + 7.379 + y = 36.93

y=7.423

Now, arranging in ascending order,

7.309,7.379,7.398,7.421,7.423

n=5

Median = =$(\frac{n+1}{2})$th term = (6/2)th term = 3rd term = 7.398(ans)

Q8)
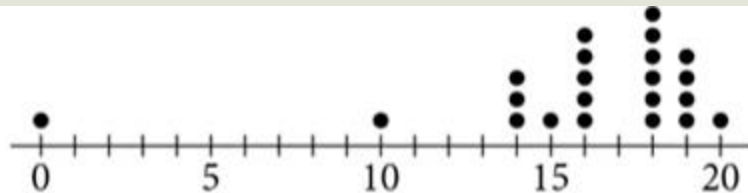
Aric's scores on 22 quizzes, each worth 20 points, are summarized in the dot plot above. The score of 0 was on a quiz given on a day Aric was absent. Aric's teacher will allow him to make up that quiz at a later date and will replace the 0 with his score on the make-up quiz.

i)According to the dot plot, what is Aric's current median score?

ii)If the teacher drops Aric's current highest & lowest score, what will be his median test score?

iii)If Aric scores 18 on the make up quiz, what will be his median score?

A8i)n=22, Median will be the sum of 11th and 12th term divided by 2. Order is already maintained in the dot plot.

We shall count the dots from the beginning until we come across the 11th and 12th dots. Then, we will check the score corresponding to it and do the necessary calculations.

The 11th dot corresponds to a score of 16 and the 12th dot to a score of 18.

Thus median=(16+18)/2 = 17(ans)

A8ii) If we chuck off both the highest and the lowest values in a data set, our median will remain the same. Suppose a set of {2,5,7,10,13). The median is 7. On removal the set is {5,7,10}. The median is unaffected. Thus answer=17

A8iii) We have chucked off 0 and 20. so now our n is 20.

If Aric scores 18 in his make up test, n will become 21 and one dot will be added to the 18 score

Median =22/2 = 11 th term which is also 18.

Ans=18

Q9) What is the mode of the data set : 1,1,2,3,3,3,3,3,8,8,10

A9)1 appears 2 times.   2 appears 1 time.   3 appears 5 times.   8 appears 2 times.   10 appears 1 time.

Thus mode=3

Q10) Ned runs a soybean farm and recorded the yields for 175 different one-acre
      Sections. The results are shown in the graph above. Which of the following could
      be the median yield of Ned's soybean acres?

a)44 bushels   b)48 bushels   c)52 bushels   d)56 bushels

A10)Median=(176)/2 th term = 88th term

In the 40-45 interval, 25 terms are there.

In the 45-50 interval, 25 to (25+70=95) terms are there.

In the 50-55 interval,95 to (95+55=150) terms are there.

In the 55-60 interval,150 to (150+20=170) terms are there.

In the 60-65 interval,170 to (170+5=175) terms are there.

88 th term lies in 45-50 interval. Ans=b(48)
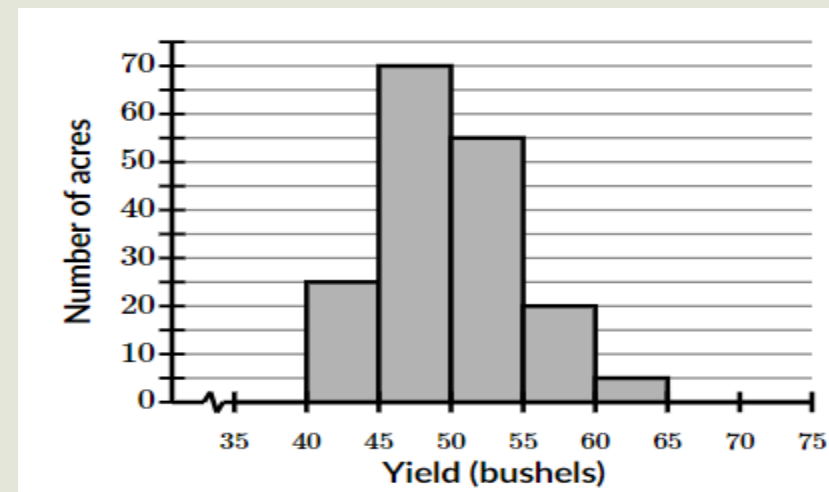
Q11)   Mr. Jadav raised all of his students' scores on a recent exam by 10 points. What effect did this have on the mean and median of the scores?

○ The mean increased by 10 points, but the median remained the same.

○ The median increased by 10 points, but the mean remained the same.

○ The mean increased by 10 points, and the median increased by 10 points.

○ The mean and the median remained the same.

A11) For such a question, it is best to visualize using small data sets. Let the three observations be 20, 30 and 40.
Initial mean = (40+30+20)/3 = 90/3 = 30
New mean=(40+10+30+10+20+10)/3=120/3 = 40
Initial median=30
New median=40

Thus we can say answer is c.
you can use slightly bigger data sets to cross check.

Q12)Which dot plot has a greater standard deviation?
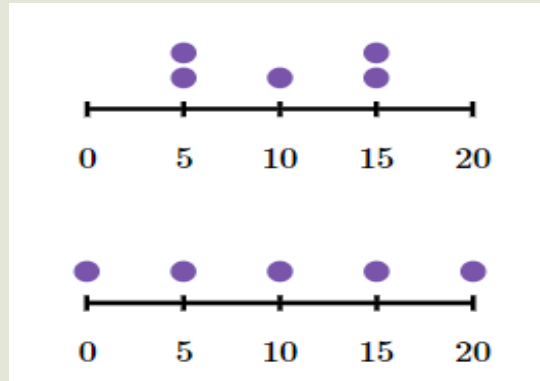
A12)Greater spread = Greater standard deviation
  Thus the second dot plot has higher standard
  deviation.
You can also check using the Desmos:

$\text{stdev}(5,5,10,15,15)$

$= 5$

$\text{stdev}(0,5,10,15,20)$

$= 7.90569415042$
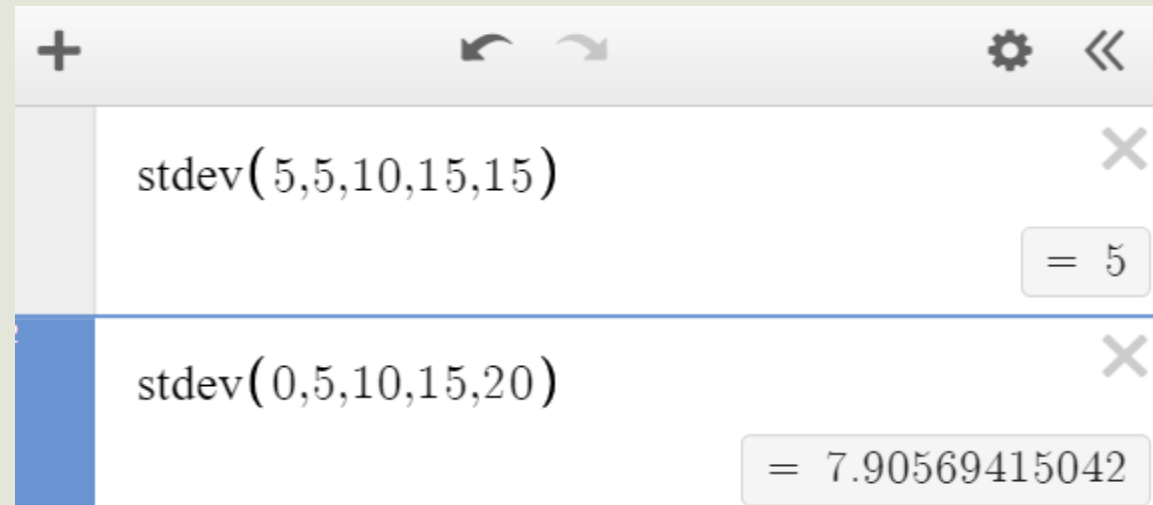
Q13) Which histogram has higher standard deviation?

A13) Histogram 1 because if you try to draw a curve connecting

the top of each bar in each histogram, you will notice that the

curve of the first histogram is much more flatter than the

curve connecting the top sections of each bar in the graph of

histogram 2.
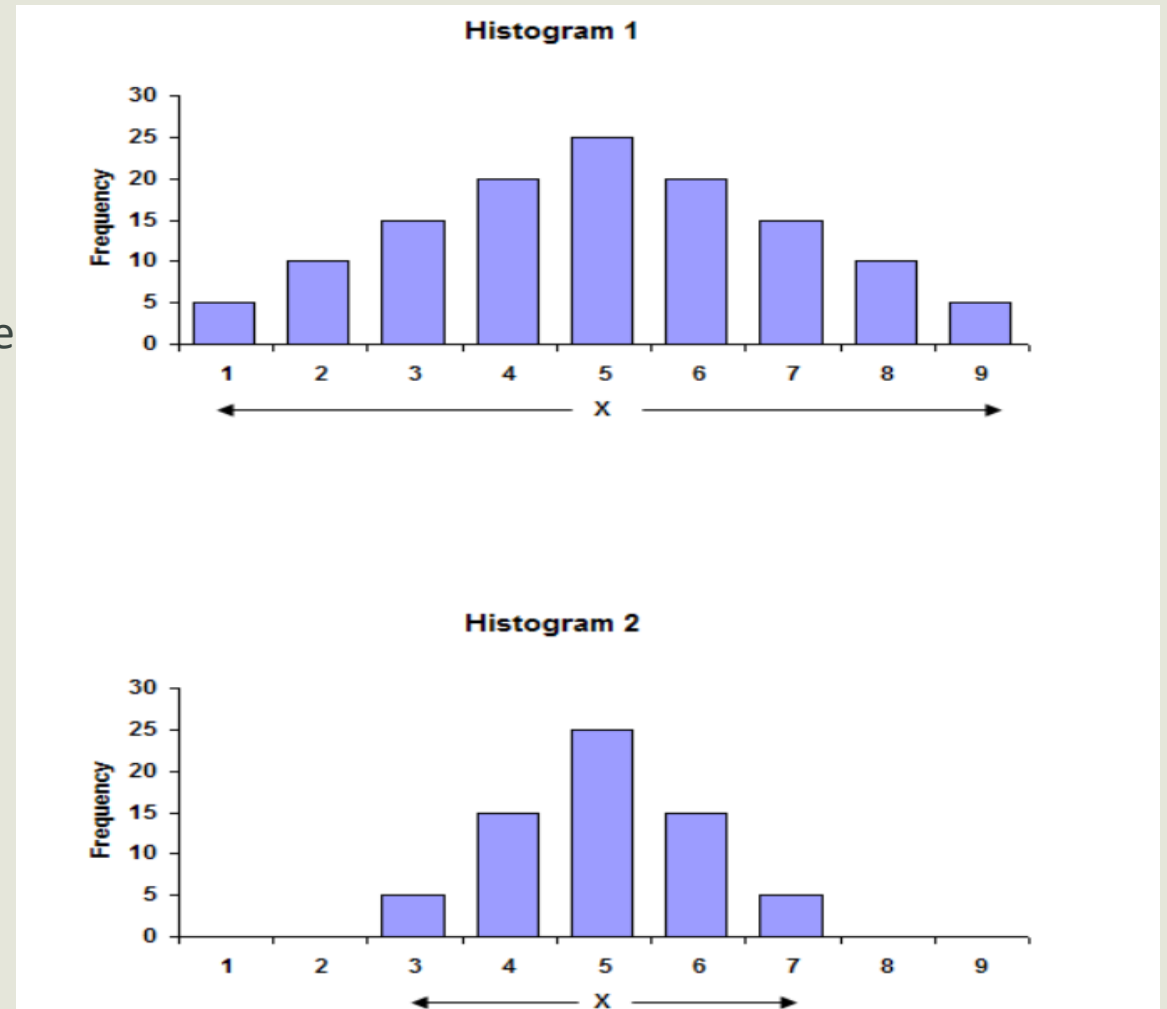
Q14) A farmer records the weight of watermelons from two different fields. Which field likely has a higher standard deviation in watermelon weights?

Field A: Most watermelons weigh between 8 and 12 pounds.

Field B: Watermelon weights range from 5 to 15 pounds.

A14)Field B has a wider range of weights, suggesting a higher standard deviation.

Q15) Two groups of students took the same math quiz. Here are the scores:

Group A: 65, 75, 75, 80, 85

Group B: 50, 70, 80, 90, 100

Which of the following is true about the standard deviations of the two groups?

(A) The standard deviation of Group A is less than the standard deviation of Group B.

(B) The standard deviation of Group A is equal to the standard deviation of Group B.

(C) The standard deviation of Group A is greater than the standard deviation of Group B

(D) The relationship between the standard deviations cannot be determined.

A15) The Standard deviation of B is greater than A.

If you notice, the spread of B is much wider than A.

You can also confirm using Desmos as shown alongside

$$\text{stdev}(65,75,75,80,85)$$
$$= 7.4161984871$$

$$\text{stdev}(50,70,80,90,100)$$
$$= 19.2353840617$$

Q16-18 refer to the table below.

	Class A	Class B
Points Earned	74, 80, 81, 83, 84, 86, 86	64, 74, 78, 83, 86, 86, 103
Mean	82	82
Median	83	83
Mode	86	86

Q16)The table below summarizes the scores of 7 students in two classes A and B on their mid-term exams. Both classes have equivalent measures of tendency. Which class has a higher variation in the distribution of scores?
A16)We are looking at how the observations are spread with respect to the mean.
   If you analyze the data, you'll find that Class B has a very wide set of observations.
   We can safely say Class B has a higher variation in the distribution of scores.
   You can even put the data on Desmos if you are not confident enough
Q17)Calculate the range for both Class A and B.
A17)Range=highest-lowest observational value.
   For class A, Range =86-74=12   For class B, Range=103-64=39.
Q18)Can you generalize that More the range, more is the S.D?
A18)No-This is not always true and is a wrong way of measuring standard deviation.

Data representations and Scatterplot:

1)Bar graphs :
A bar graph can be defined as a graphical representation of data, quantities, or numbers using bars or strips. They are used to compare and contrast different types of data, frequencies, or other measures of distinct categories of data. On a bar graph, the sizes of the bars are related to the size of the quantities: the larger a quantity is, the taller or longer the bar representing it is. The bars are separated horizontally by the same width and are of the same thickness
 Q1)On the bar graph given below, answer the following questions:
   i)Hottest city
   ii)if temperature>25 Degree Celsius, the climate is considered warm. List the warm cities.
   iii)Range for the given set of observations
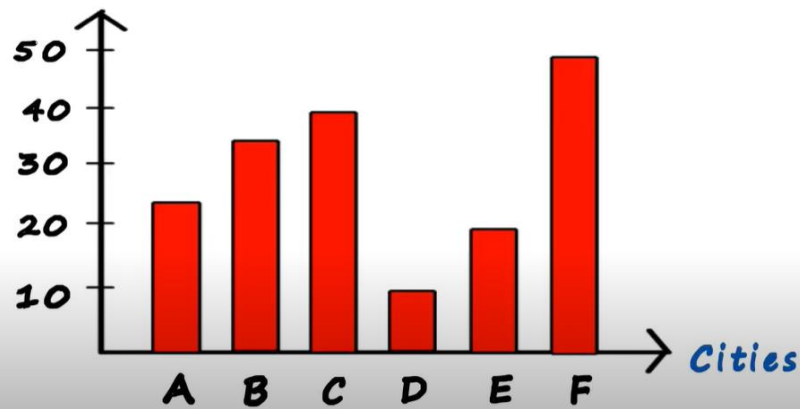
A1) City F

A2)Cities B, C and F

A3)Range= Temp of city F – Temp of city D
     Range =49-9 Degree Celsius.
     Range =40 Degree Celsius.

Q2)
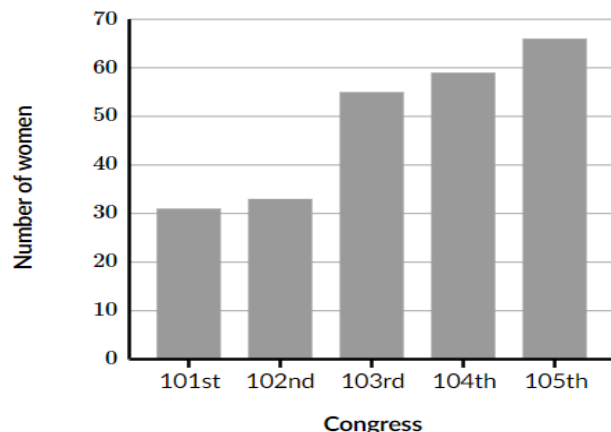
The bar graph here shows the number of women in the 101st to 105th U.S congresses. The number of women in Congress saw the greatest increase between which of the f following Congresses?

A2) From 101-102nd,Increase =33-31=2
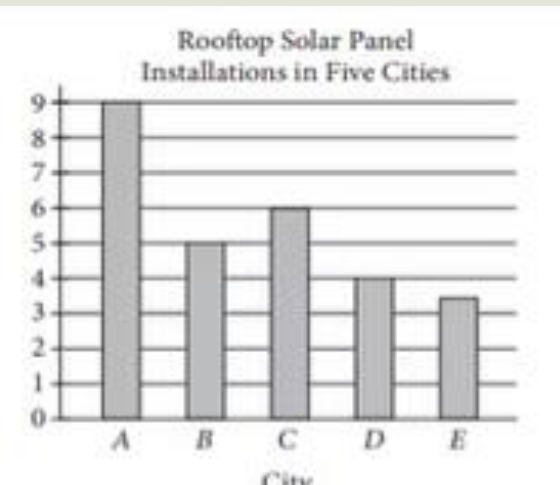
From 102-103^{4d},Increase =54-33=21

From 103-104th,Increase =59-54=5

From 104-105th,Increase=66-59=7

Therefore, highest increase is from 102nd to 103rd U.S congress.

Also note that we use approximate measures in this section when the measures in the axis are not given clearly.

Q3)

The no. of rooftops installations in 5 cities is shown by this bar graph. If the total no. of installations is 27500, what is an appropriate label for the vertical axis of the graph?

a)No. of installations in tens

b)No. of installations in hundreds

c)No. of installations in thousands

d)No. of installations in tens of thousands

A3) A=9, B=5, C=6, D=4, E=3.5. A+B+C+D+E=27.5

27.5 x 1000 will give us total number of installations = 27500.

Thus, ans = c

Q4) For the graph given alongside, What is the <u>average</u> number of seeds per apple?

A4)For this type of question, just multiply each x value corresponding to its y value and add all. Then divide by total no. of the observation that the question has asked, as usual.
2 apples had 3 seeds = 2x3 = 6
4 apples had 5 seeds= 4x5 = 20
1 apple had 6 seeds= 1x6 =6
2 apples had 7 seeds= 2x7 =14
3 apples had 9 seeds= 3x9 =27

Total number of seeds=6+20+6+14+27
=73 seeds
Total number of apples=2+4+1+2+3
=12
Arithmetic mean(average)=73/12 =6.08 (answer)

Number of Seeds in Each of 12 Apples
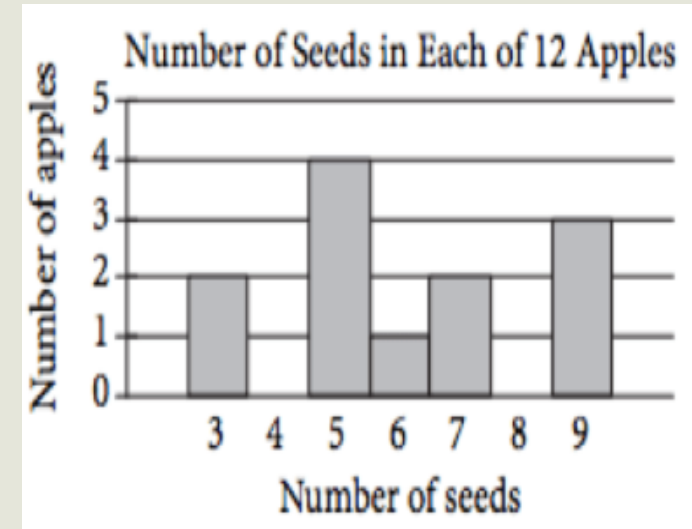
2)Histograms:

A Histogram is a graphical representation of a grouped frequency distribution with continuous classes. Histograms use bars to represent the frequency at which a range of values occurs. In such representations, all the rectangles are adjacent since the base covers the intervals between class boundaries. The heights of rectangles are proportional to corresponding frequencies of similar classes and for different classes, the heights will be proportional to corresponding frequency densities.
We must also learn how to tackle a histogram table as done below.

Q5)The ages of 20 students enrolled in a college class are given in the table shown alongside. What is the correct order of mean, median and mode ?

a)mean<median<mode
b)median<mean<mode
c)mode<mean<median
d)mode<median<mean

Ages of 20 Students Enrolled in a College Class

Age	Frequency
18	6
19	5
20	4
21	2
22	1
23	1
30	1

A5)mode=18(frequency 6)

n=20. Median=(10th+11th term)/2

Median=(19+19)/2

Median=19

Mean=$\frac{18\times6+19\times5+20\times4+21\times2+22\times1+23\times1+30\times1}{20}$ = $\frac{400}{20}$ = 20. Thus mean>median>mode(d)

Q6)Study the histogram given alongside and answer the following questions:
i)How many students altogether are displayed in this histogram?
ii)Students need to score 80-89% to earn a B and above 90% to get A. How many students earned a grade lower than a B?
iii)Which interval range includes the median test score?
A6)For this type of question where we need to find a n, always label the frequency of each bar as shown below.
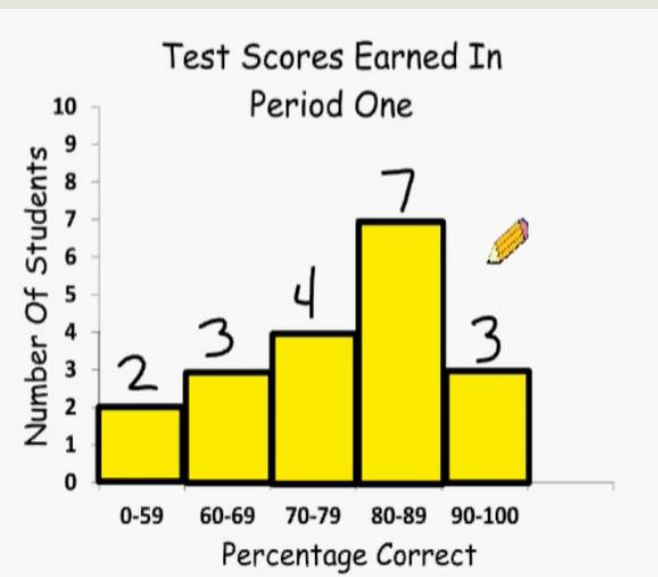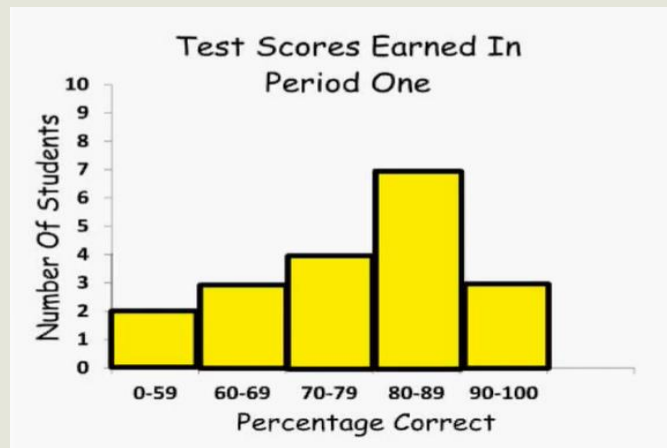
i)Total number. Of students displayed in the histogram=2+3+4+7+3=19

ii)No. of students who earned lower than a B=2+3+4 = 9

iii)n=19 (calculated above)
Median=(n+1)/2 th term
Median=10 th term
0-59 interval contains 2 students.   60-69 interval contains 2+3=5 students
70-79 interval contains 5+4=9 students   80-89 interval contains 9+7=(9 to 16 students)
ANS-80-89 interval

3)Dot plots:

Dot plots use dots to represent the frequency with which particular values occur. Dot plots are usually used for low, easily countable frequencies because it's impractical to draw or count many dots. There is just a horizontal axis on the dot plot and the count of dots for one value is its frequency.

Q7)The 26 students in a computer science class reported the amount of data in gigabytes(GB) stored in their mobile phones. The dot plot below shows the result. Which statement is an accurate analysis of the data?

a)mean>median
b)mean<median
c)mean=median
d)Insufficient information

A7) Mean=$\frac{30x1+40x8+50x9+60x1+80x1+90x4+100x2}{26} = \frac{1500}{26}$ = 57.69

25

Amount of data stored on phone (GB)

Median= ($13^{th}+14^{th}$ term)/2

Median=(50+50)/2

Median=50

Thus, Mean>Median

Answer-(a)
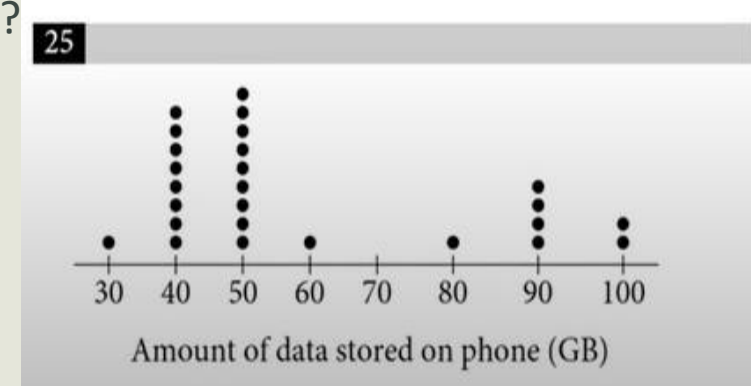
Q8)The dot plot shown each represent a data set. Which of the following statements correctly compares the mean and the standard deviation of the two dot plots A and B?
a)Means are equal; SD of A> SD of B
b)Means are equal; SD of A<SD of B
c)S.Ds are equal; Mean of A<Mean of B
d)S.Ds are equal; Mean of A>Mean of B

A8)Mean of data set A= $\dfrac{10x2+15x5+20x6+25x5+30x2}{2+5+6+5+2}$ = $\dfrac{400}{20}$ =20

Mean of data set B= $\dfrac{10x4+15x4+20x4+25x4+30x4}{4+4+4+4+4}$ = $\dfrac{400}{20}$ =20

For Standard deviation, on analyzing the two data sets you can see that the data points are much more spread in Data set B than Data set A. In Data set A, the dots are more close to the mean(20) while in B there are lots of extreme points .
Thus SD of A<SD of b.
You could also enter values in Desmos and check like this:
Thus answer=b.

Data Set A

10   15   20   25   30

Data Set B

10   15   20   25   30

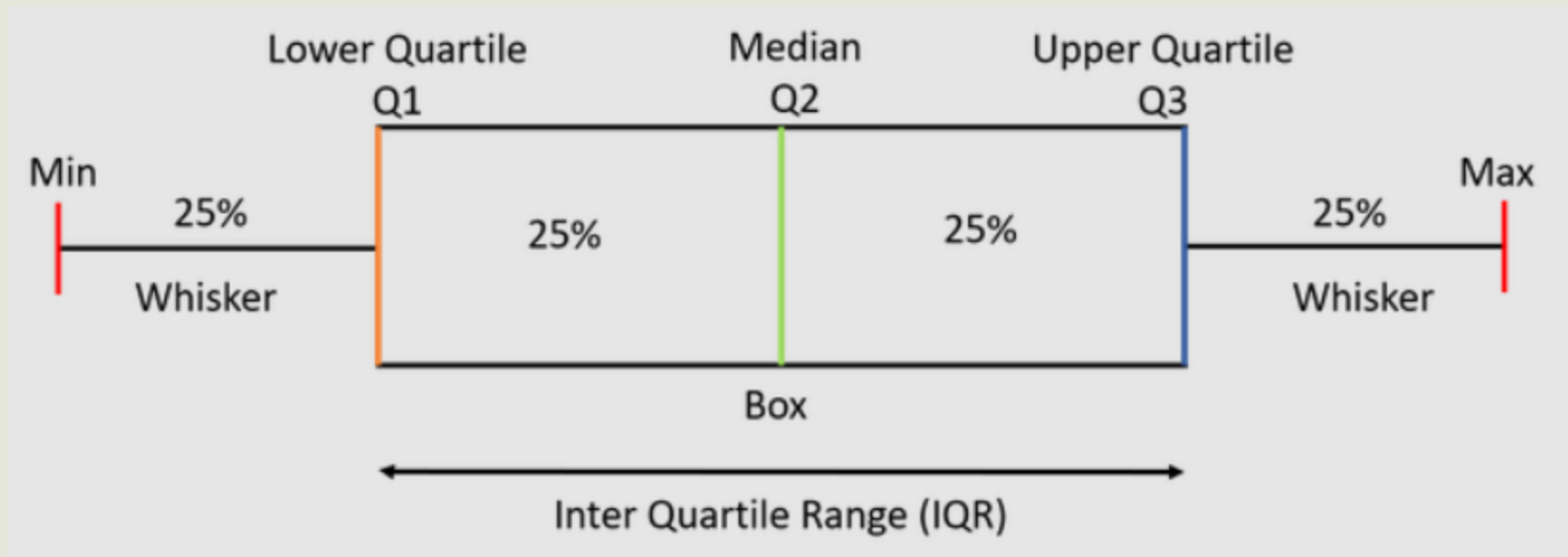stdev(10,10,15,15,15,15,15,20,20,20,20,20,   = 5.84897651866

stdev(10,10,10,10,15,15,15,15,20,20,20,20,   = 7.2547625011

4)Box Plots:

A boxplot is a standardized way of displaying the dataset based on the five-number summary: the minimum, the maximum, the sample median, and the first and third quartiles.

Q9)In the box plot given alongside, Find:
  i)value of the first quartile
  ii)value of the third quartile
  iii)value of the median
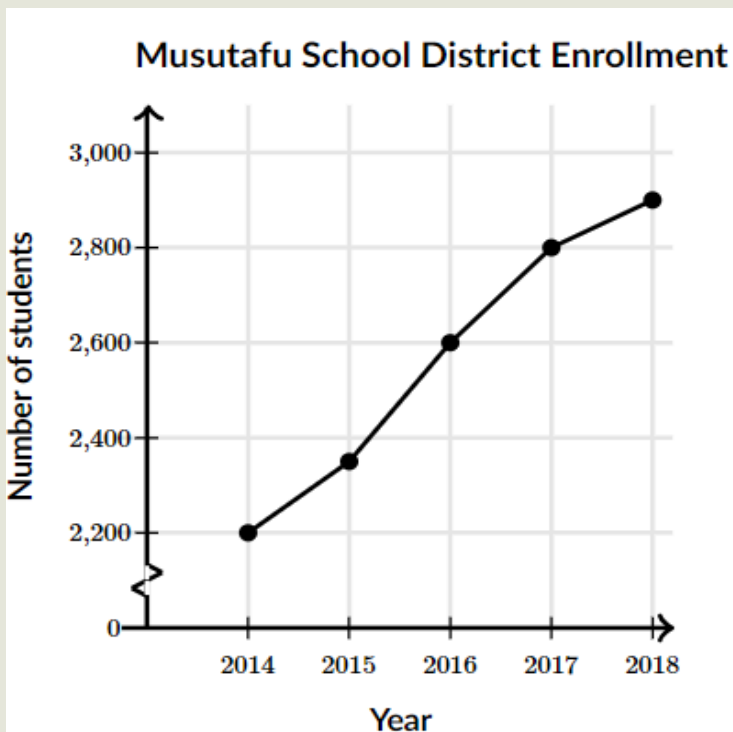 iv)value of the interquartile range
  v)the range
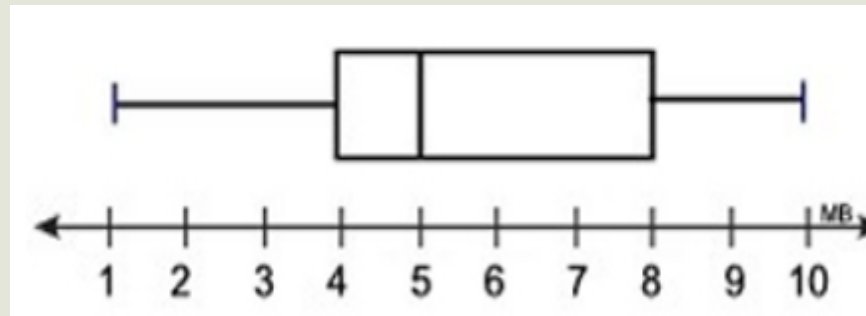
A9)
i)Q1=4

ii)Q3=8

iii)Q2(median)=5

iv)IQR=Q3-Q1=8-4=4

v)Range=highest-lowest=10-1=9

5)Line Graphs:

A line graph is a unique graph which represents the change in a quantity with respect to another quantity

Musutafu School District Enrollment

Q10) With respect to the following line graph, answer the following questions:

i)From 2014 to 2018, student enrollment in Musutafu generally:
a)Increased   ANSWER: (a)
b)Decreased
c)Stayed the same
d)Insufficient answer

ii) The school district saw the smallest change in the number of students enrolled
between: ANSWER: (d)
a)2014-2015   2350-2200=150
b)2015-2016   2600-2350=250
c)2016-2017   2800-2600=200
d)2017-2018   2900-2800=100(ans)

Q11) Paulina bought her first video game in 2014 . The line graph below shows the total number of video games Paulina owned since then. According to the graph, the number of video games she owned in 2015 is what fraction of the number she owned in 2019?

Paulina's Video Game Collection
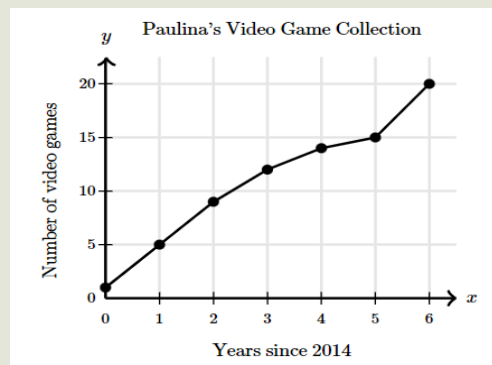
A11)Lets consider the points (1,5) and (5,15).
    They convey that Paulina owned 5 video games in 2015(1 year after 2014) and 15 video game in 2019(5 years after 2014).

Thus, the number of video games she owned in 2015 is $\dfrac{\text{No.of games owned in 2015}}{\text{No.of games owned in 2019}}$ fraction of the number she owned in 2019.

=5/15=1/3(answer)

Q12)Nayeli created a social media account for her company and she tracks the account's no. of followers as shown in the alongside graph.

On what interval did the no. of followers increased the most?

a)Between 1 to 2 months

b) Between 2 to 3 months

c)Between 3 to 4 months

d)Between 4 to 5 months

ANSWER: Since we are not given with measurements, We have to use approximation. We must find the biggest slope here. This indicates the maximum increase in the graph. On closely viewing the graph, we can safely say that the answer is (b) as there is a huge jump from 2 to 3 month
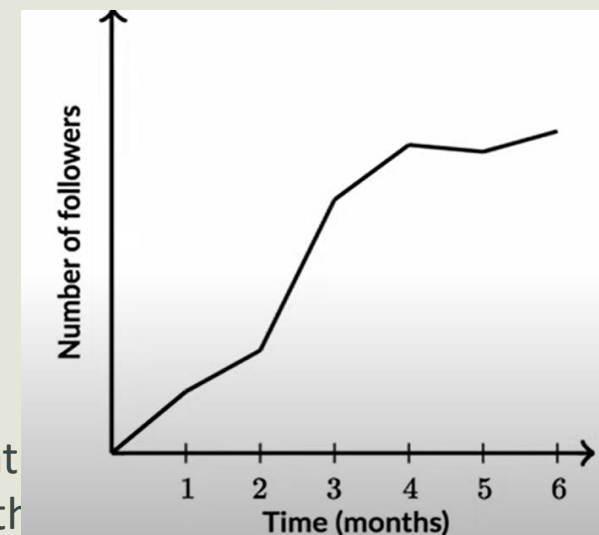
6)Scatterplots:

They are graphs that use a series of points and can be formed around a line of best fit.
The equation of this line helps us display the relationship between the two variables x and y
Consider the following table:

| Description of Best Fit	Relationship between Variables (x and y)
Upward-sloping straight line	Linear and positive
Downward-sloping straight line	Linear and negative
Upward-opening parabola	Quadratic w/ positive coefficient in front of squared term
Downward-opening parabola	Quadratic w/ negative coefficient in front of squared term
Upward-sloping curve w/ increasing slope	Exponential and positive
Downward-sloping curve w/ flattening slope	Exponential and negative

While each point in a scatterplot represents a specific observation, the <u>line of best fit</u> describes the general trend based on all of the points.
For a given data point, we expect to see a difference between its y-value and the y-value predicted by the line of best fit.
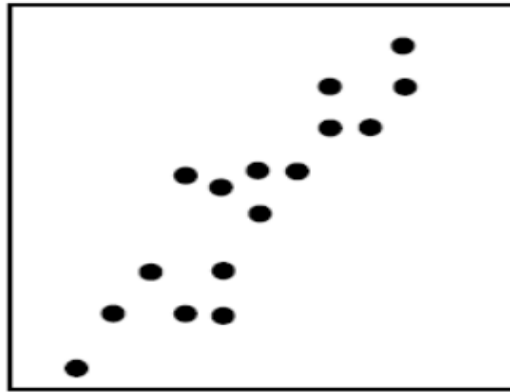These differences are used for more advanced statistical analysis; for the SAT, we only need to calculate the difference.
We can also interpret the slope and y -intercept of the line of best fit the same way we interpret line graphs.
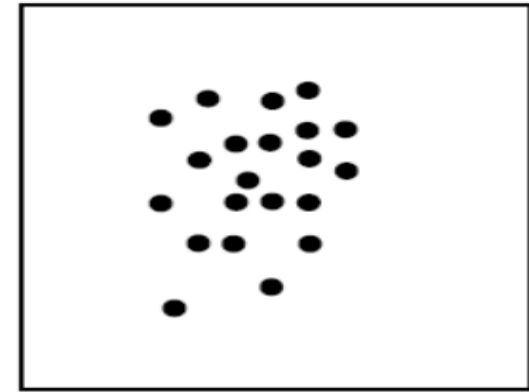
Types of relations:
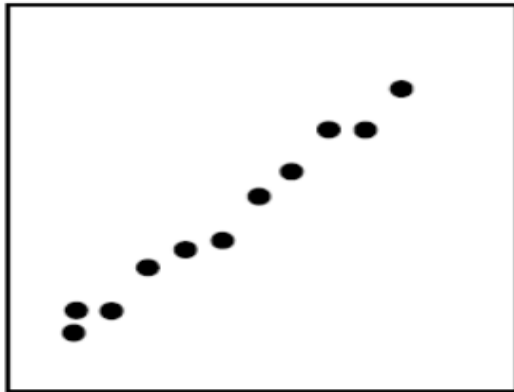
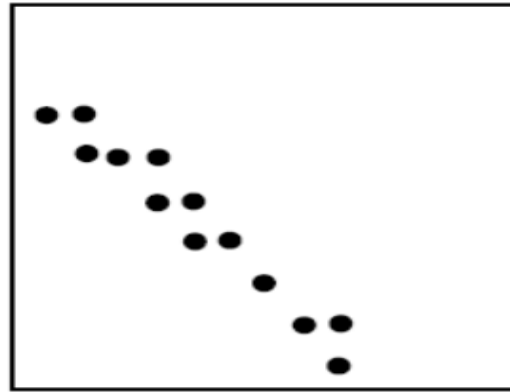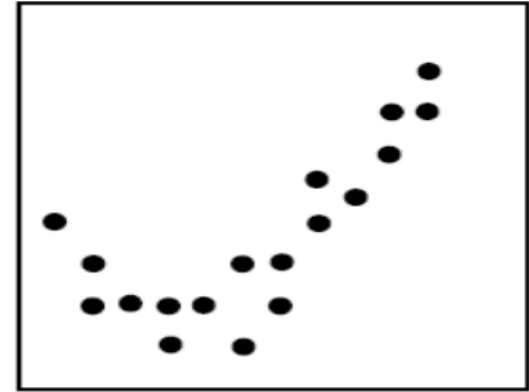Strong positive correlation

Moderate positive correlation

No correlation

Moderate negative correlation

Strong negative correlation

Curvilinear relationship

Q13)In this scatterplot,

   i)The line of best fit passes through the point (15,y). What is y?

   ii)Point m has co ordinates(a,b). What are the values of a and b?

  iii)The positive difference in y-value between the data point and the line of best fit is?

A13i)y=11

  ii)a=15;b=5.

  iii)the value of the data point m is 5 and the y value predicted by the line of best fit is 11

   Thus the positive difference in y-value between the data point and the line of best fit:

   = 11-5 =6
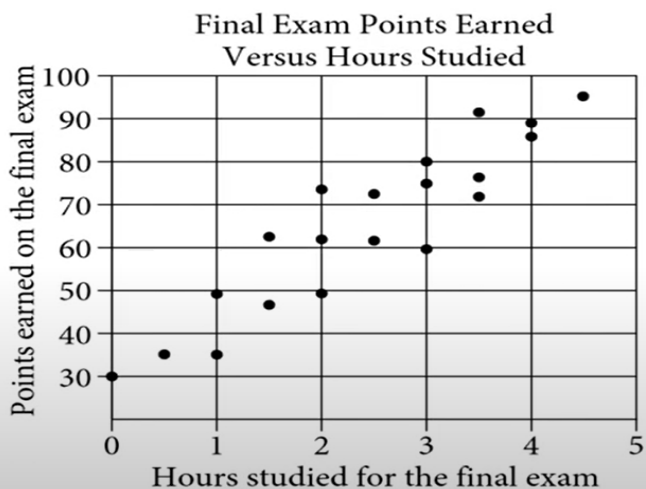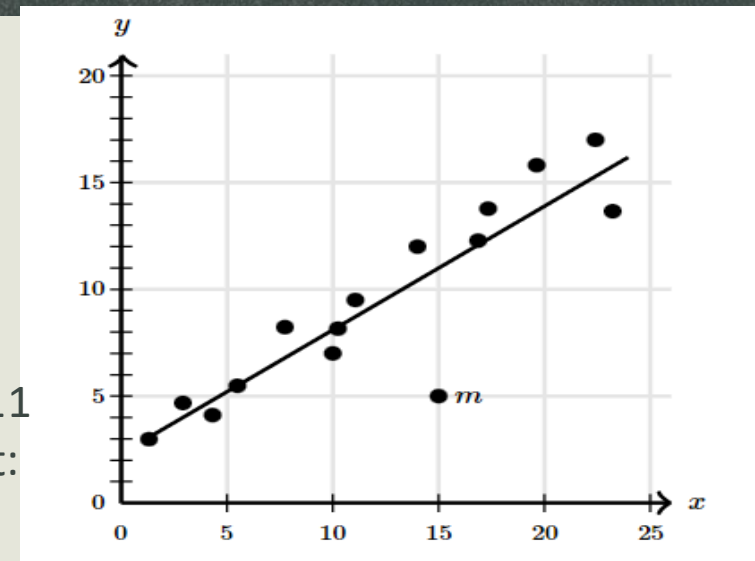
Q14)The histogram on the left shows the number of points earned on the final exam by 20 students and the no. of hours each student spent studying for the exam.

How many students studied no more than two hours and earned less than 60 points?

A14) This is the region we are interested in :

This rightly accounts for hour studied <3 or <=2 and points scored <=60.

Now just count the total number of dots and your answer is 6.

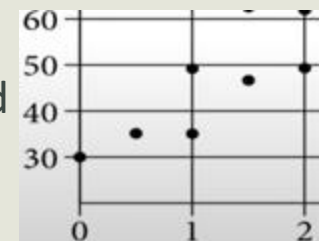Q15)In the scatterplot shown below which of the following best models the relationship between h, the number of hours studied for the final exam and P, points earned?
a)P=14.4h
b)P=14.4h+30
c)P=1.4h + 30
d)P=1.4h

Points Earned Versus
Hours Studied on the Final Exam
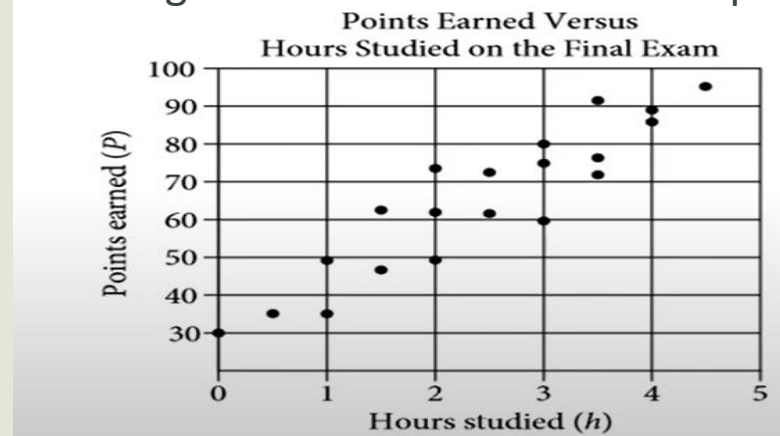
A15)First visualize a line of best fit. You can conclude that it is a linear graph(y=mx+c). Now Take two easy to calculate points (I prefer the left most and the rightmost ones) (0,30)&(4.5,95)

$$m= \frac{y2-y1}{x2-x1} = \frac{95-30}{4.5-0} = \frac{65}{4.5} = 14.4$$

Now calculate the y intercept.

When x=0, y=30. Thus c=30.

We can now form our equation y=14.4x+30

Replacing this with the variables used in this question,

P=14.4h+30 (option-b)

Q16)The scatterplot alongside shows the population of Sumatran Orangutans,in thousands, in 2014 and several years after. According to the line of best fit, to the nearest hundred, what is the yearly decrease in the number of orangutans from 2014 to 2018?

Orangutan Population in Sumatra
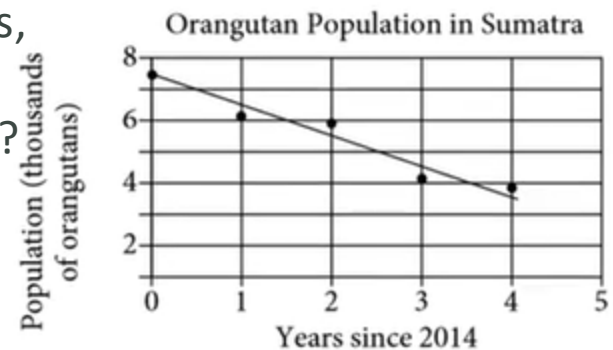
A16)The yearly decrease in the no. of orangutans can be found out using slope.
  Lets take two points: (0,7.5) and (4,3.5).
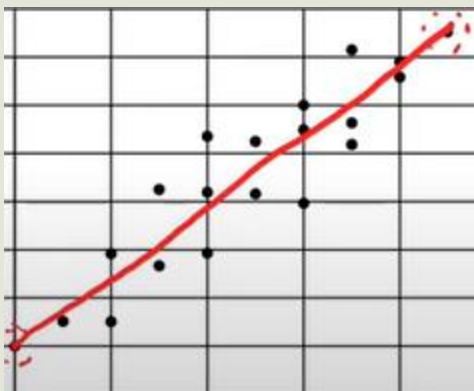 slope= m= $\dfrac{y2-y1}{x2-x1}$ = $\dfrac{7.5-3.5}{0-4}$ = $\dfrac{4 \times 1000 \text{(it was mentioned in thousands)}}{-4}$ = - $\dfrac{4000}{4}$ = -1000

NOTE-If this comes as a grid in question, insert 1000 and not -1000 as in the question it is already mentioned "decrease".
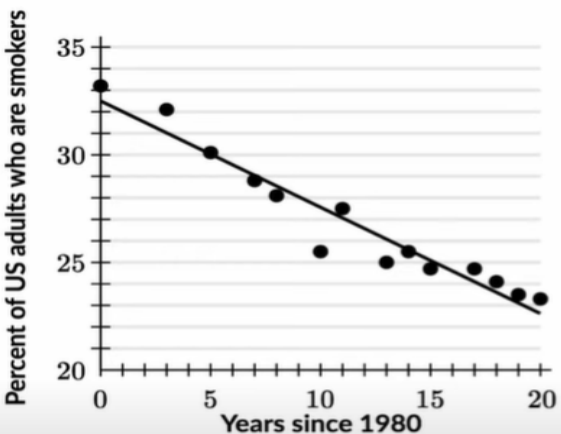
Q17)In the figure alongside, the percent of US adults who are smokers from 1980 to 2000. A line of best fit is also shown. Based on the line of best fit, which of the following is closest to the predicted percent of U.S adults who are smokers in 1981?

  a)26%   b)28%   c)30%   d)32%

A17)1981 would be 1 year after 1980. Find The y value when x=1. It is 32%.
  Thus we can confidently say that the predicted % of U.S adults who are smokers in 1981 is 32%.
  Answer-option (d)

Probability and relative frequency

Probability is the branch of mathematics concerning events and numerical descriptions of how likely they are to occur. The probability of an event is a number between 0 and 1; the larger the probability, the more likely an event is to occur.

$$P = \frac{No.of\ favourable\ outcomes}{Total\ no.of\ outcomes}$$

Q1)The probability a spinner will land on red is 0.7. If the spinner is spun 100 times, How many times will the spinner land on red?

A1)Ans=0.7 x 100 = 70

Q2) Jackie asked his classmates whether they own a bike or a skateboard. The results are shown in the table alongside

i)How many classmates own a skateboard?

ii)How many classmates own both a bike and a skateboard?

iii)How many classmates own either a bike or a skateboard?

iv) If a classmate is selected at random, what is the probability that they do not own a bike?

	Owns a skateboard	Does not own a skateboard	Total
Owns a bike	4	11	15
Does not own a bike	3	7	10
Total	7	18	25

A1i)4+3=7

ii)4

iii)3+11=14

iv) $P = \frac{No.of\ favourable\ outcomes}{Total\ no.of\ outcomes}$

$P = {}^{10}/_{25} = 0.4$

Q3)The table below summarizes the results of 200 law school graduates who took the bar exam. If one of the surveyed graduates who passed the bar exam is chosen at random for an interview, What is the probability that the person chosen did not take the review course?

Results on the Bar Exam of Law School Graduates	Passed bar exam	Did not pass bar exam
Took review course	18	82
Did not take review course	7	93

A3) P=$\frac{\text{No.of favourable outcomes}}{\text{Total no.of outcomes}}$

Here, our no. of favourable outcomes= a person who did not take the review course.

Total no. of outcomes= all people who have passed the bar exam.

P= $\frac{7}{25}$ = 0.28

Q4) Four teams, City, Rovers, Town and United play a competition to win a cup. Only one team can win the cup. The table below shows the probabilities of City or Rovers or Town winning the cup. Find x.

City	Rovers	Town	United
0.38	0.27	0.15	x

A4)0.38+0.27+0.15+x=1   x=0.2(ans_

   x=1-0.15-0.27-0.38=0.2

Q5)For a particular machine that produces beads, 29 out of every 100 beads it produces have a defect. A bead produced by the machine will be selected at random. What is the probability that the bead selected will have a defect?

a)1/2900   b)1/29   c)29/100   d)29/10

A5)100 produced beads will contain -->29 defected beads

1 produced bead will contain--->$\frac{29}{100}$(option-c)(ans)

Q6)The table alongside gives the distribution of votes for a new school mascot and grade level for 80 students. If one of these students is randomly selected, what is the probability of selecting a student whose vote for a new mascot was a lion?

	Grade level			
Mascot	Sixth	Seventh	Eighth	Total
Badger	4	9	9	22
Lion	9	2	9	20
Longhorn	4	6	4	14
Tiger	6	9	9	24
Total	23	26	31	80

a)1/9   b)1/5   c)1/4   d)2/3

A6)First get total no. of lines =20 and total no. of mascots=80

P=$\frac{No.of\ favourable\ outcomes}{Total\ no.of\ outcomes}$ = $\frac{20}{80}$ =1/4 (option c)(ans)

Q7)Each face of a fair 14-sided die is labelled with a number from 1-14, with a different number appearing on each face. If the die is rolled one time. What is the probability of getting a 2?

a)1/14   b)2/14   c)12/14   d)13/14   Also calculate probability of not getting a two.

A7) P=$\frac{No.of\ favourable\ outcomes}{Total\ no\ of\ outcomes}$ = 1/14(option a)(ans).   1-(1/14)=0.92

Q8)The table above represents the medals won at the Summer Olympics. Approximately what percent of gold medals were won by Great Britain and the United States?

	Great Britain	United States	Other nations	Total
Gold	56	23	31	110
Silver	51	12	44	107
Bronze	39	12	56	107
Total	146	47	131	324

A8) P=$\frac{\text{No.of favourable outcomes}}{\text{Total no.of outcomes}}$ = $\frac{(56+23)}{110}$ x 100% = 72% (ans)

NOTE-Students taking 324 in the denominator is a common pitfall.
It is given that we are only interested in gold medals so we will only make a selection
Out of all gold medals and check how much of them have been won by Great Britain & U.S

Q9)Angel's ice cream shop sells ice cream by the scoop and by the pint. The table alongside shows the purchases made by her customers last weekend. Of the customers who bought scoops, what fraction of them bought pints?
a)27/310   b)27/73   c)27/283   d)45/48

	Bought pints	Did not buy pints	Total
Bought scoops	27	283	310
Did not buy scoops	45	3	48
Total	73	286	358

A9) P=$\frac{\text{No.of favourable outcomes}}{\text{Total no.of outcomes}}$ =$\frac{27}{310}$ (option –a)(ans). Again notice that we took 310 in the denominator and not 358 because it
the following is mentioned: "Of the customers who bought scoops, what...."

Inference from sample statistics and margin of error

-When we want to answer questions like "how many voters feel positively about a new law" or "what percentage of Africans exercise regularly", it's often impractical to ask everyone—it would take a lot of time and effort to ask every voter, let alone every African!

So, when we have questions about a large population, we often answer those questions by surveying a **representative sample**: a smaller set of people whose answers can give us a good idea of how the population would answer the same questions.

A random sample drawn from a population is representative of the population. With a representative sample, we can multiply the sample proportion with the population to get an estimate.

While answering these questions, we must look for random samples.

Estimate = sample proportion x population.

Margin of error:

While we can make reasonable estimates using sample proportions, we can never be 100% certain that the population proportion matches the sample proportion exactly.

 Margins of error let us address the uncertainty inherent to sampling.

The margin of error is most commonly given as a percentage. When given a percent estimate and a margin of error, we can establish a range around the estimate by adding and subtracting the margin of error.

Range= Estimate – MOE to Estimate + MOE.

Q1)A group surveys 50 people in the mall about their support or opposition to a politician's policy. What is the main reason that this method of collecting data in order to make a conclusion about the opinions of all the people in the town?

a)Sample size is too small

b)You should not talk about politics in public

c)The sample was not taken randomly

d)Sample size is too big.

A1)Ans-(b).

A sample was not taken randomly. We just took people from the mall. Not everyone in the town visit malls and thus this is not a good method of data collection.

Q2)In a survey of a random sample of 1500 residents aged 25 years or older from a particular county, 399 residents had a bachelors degree or higher. If the entire county had 635000 residents aged 25 years or older, approximately how many county residents could be expected to have a bachelors degree or higher?

A2)$\frac{399}{1500}$ residents had a bachelors or higher degree in the county.

This is our sample proportion

The population given is 635000.

Estimate = sample proportion x population.

$= \frac{399}{1500}$ x 635000 = 168910

Q3) In a recent survey of 600 randomly selected registered voters in the town of Carrington, 482 are in favor of increasing funding for the town's mental health services. Based on the survey results, approximately how many of Carrington's 19310 registered voters are in favor of increasing funding for the town's mental health services?

a)482   b)1082   c)15512   d)19192

A3)Estimate= sample proportion x population

$$=\frac{482}{600} \text{ x } 19310 = 15512 \text{(option c-ans)}$$

Q4)Anya surveyed a random sample of members of a large gym about how often they visit the Gym. Using the sample data, she estimated that 74% of gym members go to the gym at least once a week, with a margin of error of 3% . Which of the following is the most appropriate conclusion about all members of the gym, based on the given estimate and margin of error?

Range= Estimate – MOE to Estimate + MOE

=(74-3)% to (74+3)%

=71 to 77%

a) Anya is 71% to 77% sure that less than half of the gym members go to the gym at least once a week.
b) It is likely that the percentage of gym members who go to the gym at least once a week is between 71% and 77% .
c) At least 71%, but no more than 74% , of the gym members go to the gym at least once a week
d) It is unlikely that more than 74% of the gym members go to the gym at least once a week.

ANSWER-(b)

Evaluating statistical claims: observational studies and experiments

-We routinely conduct research to answer questions such as "how many residents are in favor of a new law" or "is a new medical treatment effective?" While research results can give us powerful insights, we must carefully consider *how* the research is conducted, which in turn affects *what* conclusions can be drawn.

For example:

-If a survey was given to individuals of one ethnicity, then the results of the survey are *not* representative of individuals of other ethnicities.

-If a medical treatment is effective when tested on mice, we *cannot* conclude that the treatment is just as effective on humans without additional testing.

We won't be required to perform any calculations for these problems. Instead, we'll be asked to read fairly lengthy descriptions and then make logical observations or draw valid conclusions.

To make valid conclusions about the population , we need a sample that recreates the characteristics of the population on a smaller scale.

A **good** sample is **representative** and **random**.

Representative means that the sample includes only members of the population being studied.

Random means that every member of the population being studied has an equal chance to be selected for the sample.

Bad sampling methods include those that:

Gather data from outside the population being studied

Gather data that overrepresent or underrepresent a subgroup of the population (not random)

We can draw conclusions about only the population from which the random sample was selected.
Note the following two words:
1)**Correlation** means there is a relationship or pattern between the values of two variables.
2)**Causation** means that one event causes another event to occur.

NOTE-Completely eliminate options which include these two words : "always" and "never" and you can only generalize your conclusion to the population not all people.

Q1)A study was conducted to determine if a new treatment is successful in treating insomnia. 500 <u>participants were selected at random from a large population of people with insomnia.</u> Half of the participants were randomly assigned to receive the treatment, and the other half did not receive the treatment. The resulting data showed that participants who received the treatment slept significantly better than those who did not. Based on the design and results of the study, which of the following in an appropriate conclusion?
a) The new treatment is likely to improve the sleep of anyone who undergoes it.
b) The new treatment is better than all other available treatments in improving the sleep of people with insomnia.
c) The treatment is likely to improve the sleep of people with insomnia
d) None of the conclusions are appropriate.

A1)Eliminate option a because the participants were selected at random from a large population of people with insomnia. Therefore, the results can be generalized only to that population and not to people in general.

Eliminate option b because The study did not include all available treatments, so no conclusion can be made about the relative effectiveness of all available treatments.

Eliminate option d because The study was conducted on a well-defined population with random selection and random assignment, so the results of the study can be used to make conclusions about cause and effect on the population studied (people with insomnia)

Select option c because the study was conducted on a well-defined population with random selection and random assignment, so the results of the study can be used to make conclusions about cause and effect on the population studied.

Since the sleep of participants who received the new treatment improved significantly, we can conclude that the treatment is likely to improve the sleep of all people with insomnia.

Q2) A study was done on the lengths of frogs in a pond. A random sample of frogs were caught and tagged in order to ensure that none were measured more than once. The sample contained 50 American bullfrogs, of which 40% were shorter than 7 inches. Which of the following conclusions is best supported by the sample data?
a)Approximately 40% of all American bullfrogs in the pond are shorter than 7 inches.
b) Approximately 40% of all frogs in the pond are shorter than 7 inches

c) The average length of all frogs in the pond is approximately 7 inches.
d) The majority of all frogs in the pond are longer than 7 inches.

A2) Eliminate option b because: Since the sample contained 50 American bullfrogs, of which 40% are shorter than 7 inches, the largest population to which this result can be generalized is the population of the American bullfrogs in the pond, not all frogs in the pond (some of which may not be bullfrogs).

Eliminate option c because: The results do not give us any information on the *average* length of American bullfrogs or all frogs in the pond.

Eliminate option d because :Since the sample contained 50 American bullfrogs, of which 40% are shorter than 7 inches, the largest population to which this result can be generalized is the population of the American bullfrogs in the pond, not all frogs in the pond (some of which may not be bullfrogs).

Select option a because : The sample of 50 American bullfrogs was selected at random from all the American bullfrogs in the pond, and since 40% of them are shorter than 7 inches, it can be concluded that approximately 40% of all American bullfrogs in the pond are shorter than 7 inches.