Q1) Identify the Data type for the Following:

Activity	Data Type
Number of beatings from Wife	Discrete
Results of rolling a dice	Discrete
Weight of a person	Continuous
Weight of Gold	Continuous
Distance between two places	Continuous
Length of a leaf	Continuous
Dog's weight	Continuous
Blue Color	Discrete
Number of kids	Discrete
Number of tickets in Indian railways	Discrete
Number of times married	Discrete
Gender (Male or Female)	Discrete

Q2) Identify the Data types, which were among the following

Nominal, Ordinal, Interval, Ratio.

Data	Data Type
Gender	Nominal
High School Class Ranking	Ordinal
Celsius Temperature	Ratio
Weight	Interval
Hair Color	Nominal
Socioeconomic Status	Nominal
Fahrenheit Temperature	Ratio
Height	Interval
Type of living accommodation	Ordinal
Level of Agreement	Nominal
IQ(Intelligence Scale)	Interval
Sales Figures	Interval
Blood Group	Nominal
Time Of Day	Ordinal
Time on a Clock with Hands	Nominal

Number of Children	Nominal
Religious Preference	Ordinal
Barometer Pressure	Ratio
SAT Scores	Interval
Years of Education	Interval

Q3) Three Coins are tossed, find the probability that two heads and one tail are obtained?-3/8

Q4) Two Dice are rolled, find the probability that sum is

a) Equal to 1 = 0
b) Less than or equal to 4=1/6
c) Sum is divisible by 2 and 3=1/6

Q5) A bag contains 2 red, 3 green and 2 blue balls. Two balls are drawn at random. What is the probability that none of the balls drawn is blue?

Ans-10/21

6) Calculate the Expected number of candies for a randomly selected child

Below are the probabilities of count of candies for children (ignoring the nature of the child-Generalized view)

CHILD	Candies count	Probability
A	1	0.015
B	4	0.20
C	3	0.65
D	5	0.005
E	6	0.01
F	2	0.120

Child A – probability of having 1 candy = 0.015.

Child B – probability of having 4 candies = 0.20

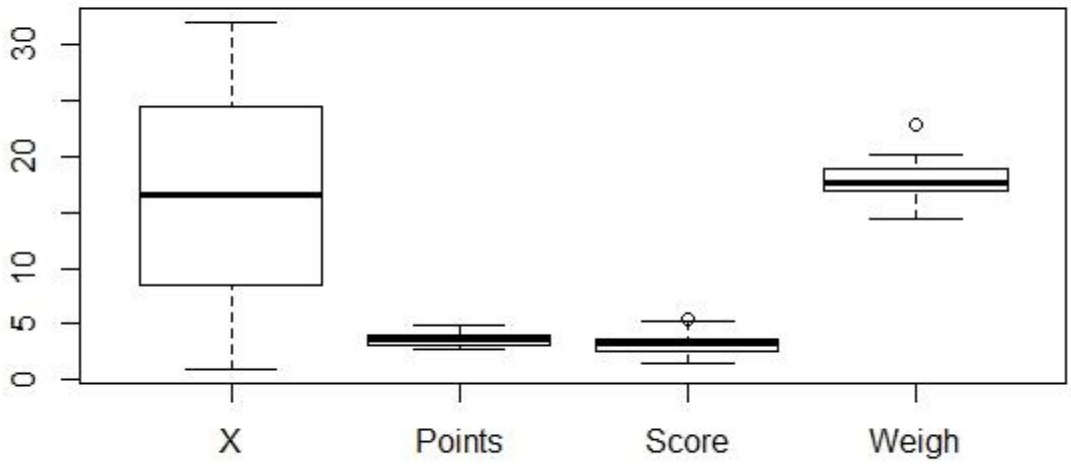
Ans-3.09

Q7) Calculate Mean, Median, Mode, Variance, Standard Deviation, Range & comment about the values / draw inferences, for the given dataset

- For Points,Score,Weigh>
  Find Mean, Median, Mode, Variance, Standard Deviation, and Range and also Comment about the values/ Draw some inferences.

ANS-

	Points	Score	Weight
Mean	3.596	3.217	17.848
Median	3.695	3.325	17.71
Mode	3.891	3.54	17.43
Variance	0.285	0.957	3.19
Standard Deviation	0.534	0.978	1.786
Range	2.76,4.93	1.513,5.424	14.5,22.9


0
5
10
20
30
X
Points
Score
Weigh

Points	Score	Weigh>
3.9	2.62	16.46
3.9	2.875	17.02
3.85	2.32	18.61
3.08	3.215	19.44
3.15	3.44	17.02
2.76	3.46	20.22
3.21	3.57	15.84
3.69	3.19	20
3.92	3.15	22.9
3.92	3.44	18.3
3.92	3.44	18.9
3.07	4.07	17.4
3.07	3.73	17.6
3.07	3.78	18
2.93	5.25	17.98
3	5.424	17.82
3.23	5.345	17.42
4.08	2.2	19.47
4.93	1.615	18.52
4.22	1.835	19.9
3.7	2.465	20.01
2.76	3.52	16.87
3.15	3.435	17.3
3.73	3.84	15.41
3.08	3.845	17.05
4.08	1.935	18.9
4.43	2.14	16.7
3.77	1.513	16.9
4.22	3.17	14.5
3.62	2.77	15.5
3.54	3.57	14.6
4.11	2.78	18.6

Q8) Calculate Expected Value for the problem below

a) The weights (X) of patients at a clinic (in pounds), are
108, 110, 123, 134, 135, 145, 167, 187, 199

Assume one of the patients is chosen at random. What is the Expected Value of the Weight of that patient?

Solution=

```
Q8 <- read.csv("D://data sets//csv//Q8.csv")
> mean(Q8$x)
[1] 145.3333
```

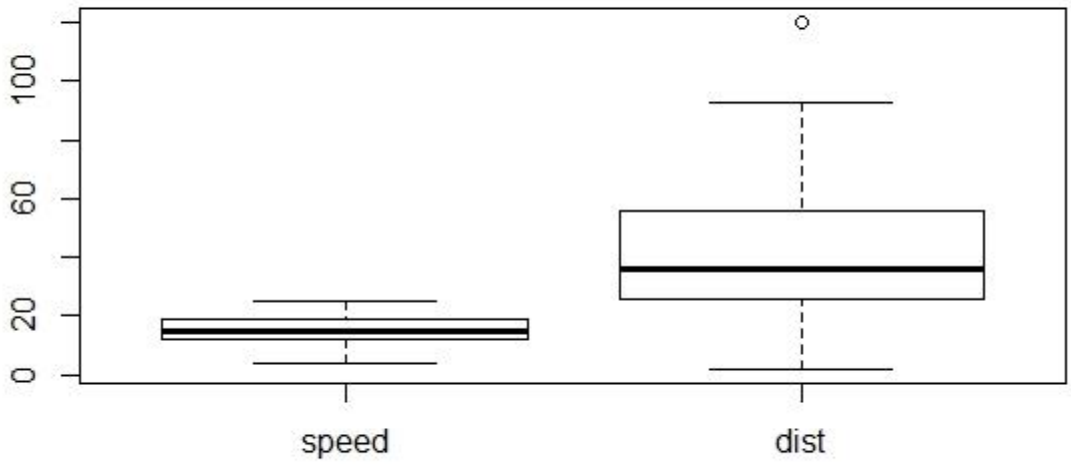
Q9) Calculate Skewness, Kurtosis & draw inferences on the following data

Cars speed and distance

speed	dist
4	2
4	10
7	4
7	22
8	16
9	10
10	18
10	26
10	34
11	17
11	28
12	14
12	20
12	24
12	28
13	26
13	34
13	34
13	46
14	26
14	36
14	60
14	80
15	20
15	26
15	54
16	32

Solution

	Car speed	Distance
Skewness	-0.1139548	0.7824835
Kurtosis	2.422853	3.248019



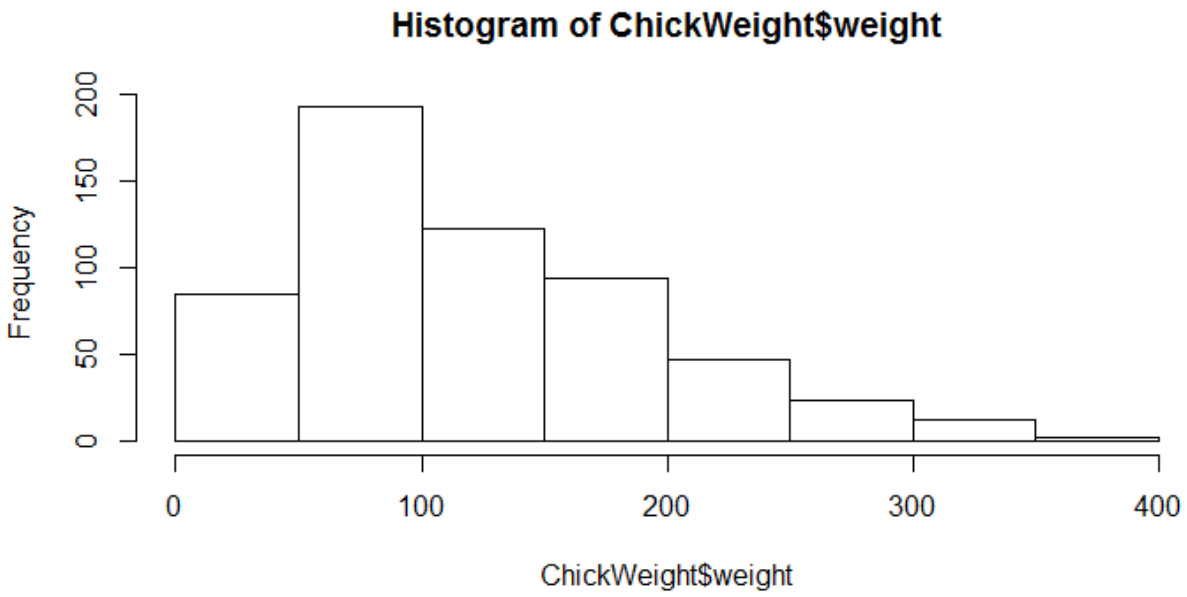
SP and Weight(WT)

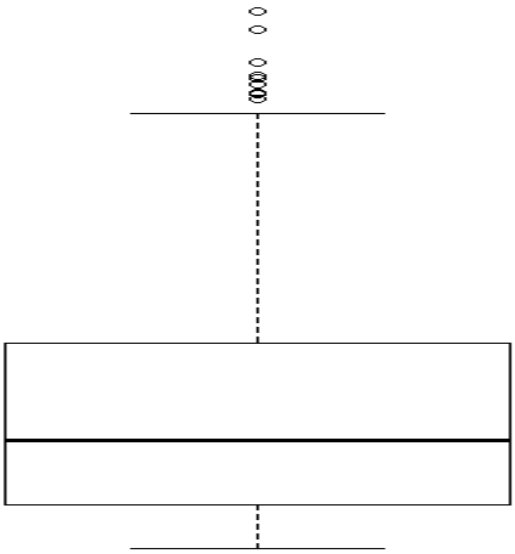
SP	WT
104.1854	28.76206
105.4613	30.46683
105.4613	30.1936
113.4613	30.63211
104.4613	29.88915
113.1854	29.59177
105.4613	30.30848
102.5985	15.84776
102.5985	16.35948
115.6452	30.92015
111.1854	29.36334
117.5985	15.75353
122.1051	32.81359
111.1854	29.37844
108.1854	29.34728
111.1854	29.60453
114.3693	29.53578
117.5985	16.19412
114.3693	29.92939
118.4729	33.51697
119.1051	32.32465
110.8408	34.90821
120.289	32.67583
113.8291	31.83712
119.1854	28.78173
114.5985	16.04317
120.7605	38.06282
119.1051	32.83507
99.56491	34.48321
121.8408	35.54936
113.4846	37.04235
112.289	33.23436
119.9211	31.38004
121.3926	37.57329

SOLUTION

	SP	WT
Skewness	1.581454	-0.6033099
Kurtosis	5.723521	3.819466

Q10) Draw inferences about the following boxplot & histogram





Q11) Suppose we want to estimate the average weight of an adult male in Mexico. We draw a random sample of 2,000 men from a population of 3,000,000 men and weigh them. We find that the average person in our sample weighs 200 pounds, and the standard deviation of the sample is 30 pounds. Calculate 94%,98%,96% confidence interval ?

Solution=

confidence interval	**Z value**	**Range**
confidence interval94%	1.880794	198.74,201.26
confidence interval96%	2.053749	198.62,201.38
confidence interval98%	2.326348	198.43,201.56

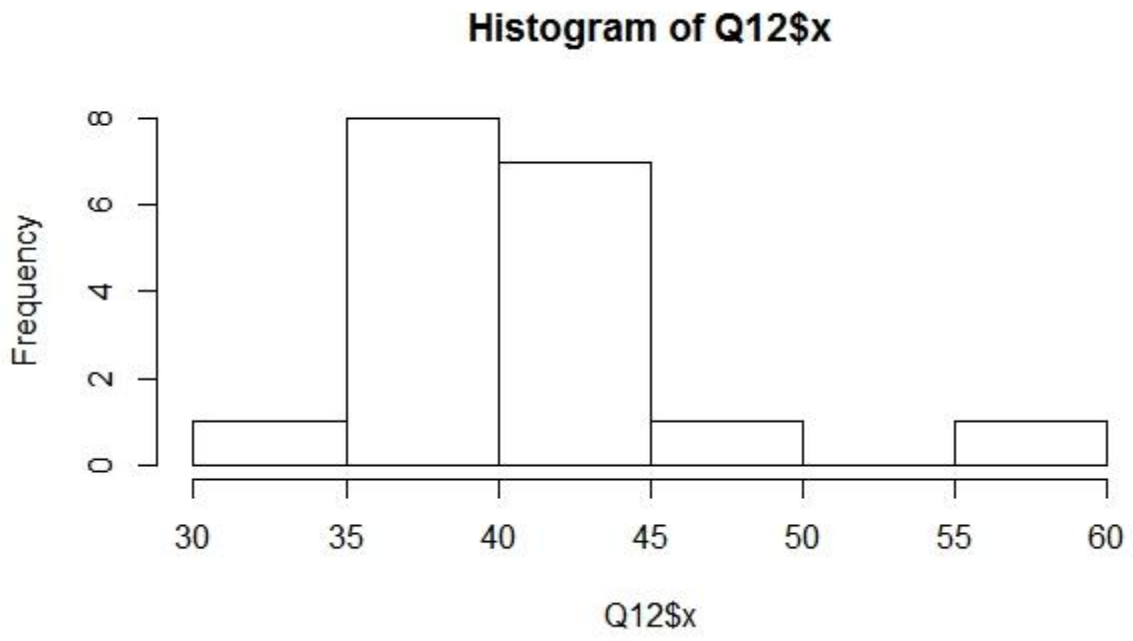
Q12) Below are the scores obtained by a student in tests

34,36,36,38,38,39,39,40,40,41,41,41,41,42,42,45,49,56

1) Find mean, median, variance, standard deviation.
2) What can we say about the student marks?

Solution=

Mean	41
Median	40.5
Variance	25.52
Standard deviation	5.05664



2)Mass of students marks between 38-42.

Skewness(1.52) is positive because mass of marks in left side of plot.

Q13) What is the nature of skewness when mean, median of data are equal?

Solution-Data is normalized and there is no skewness.

Q14) What is the nature of skewness when mean > median ?

Solution-Negative Skewness implies mass of the Distribution concentrated on right side.

Q15) What is the nature of skewness when median > mean?

Solution-Positive Skewness implies mass of the Distribution concentrated on left side.

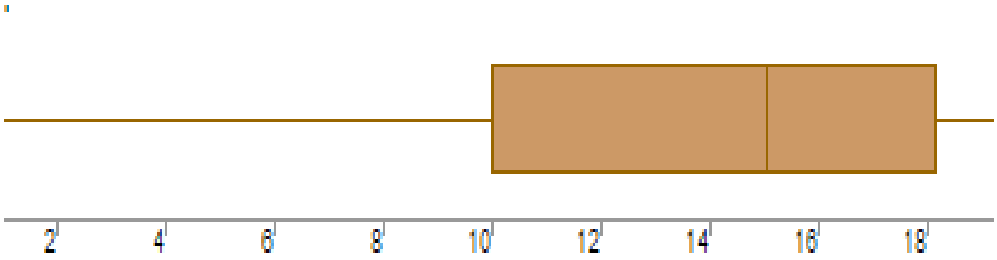
Q16) What does positive kurtosis value indicates for a data ?

Solution- Positive kurtosis value indicates that thinner peak and wider tails.

Q17) What does negative kurtosis value indicates for a data?

Solution-Negative kurtosis value indicates that wider peak and thinner tails.

Q18) Answer the below questions using the below boxplot visualization.



What can we say about the distribution of the data?

-Not normally distributed

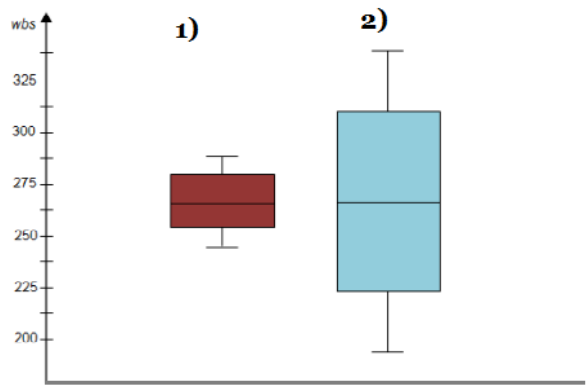
What is nature of skewness of the data?

-Negative skewness

What will be the IQR of the data (approximately)?

-10-18

Q19) Comment on the below Boxplot visualizations?



Draw an Inference from the distribution of data for Boxplot 1 with respect Boxplot 2.

Q 20) Calculate probability from the given dataset for the below cases

Data _set: Cars.csv

Calculate the probability of MPG of Cars for the below cases.

MPG <- Cars$MPG

a. P(MPG>38)-

```
1-pnorm(38,34.422,9.13144)= 0.3475908
```

b. P(MPG<40)-

```
pnorm(40,34.422,9.13144)= 0.7293527
```

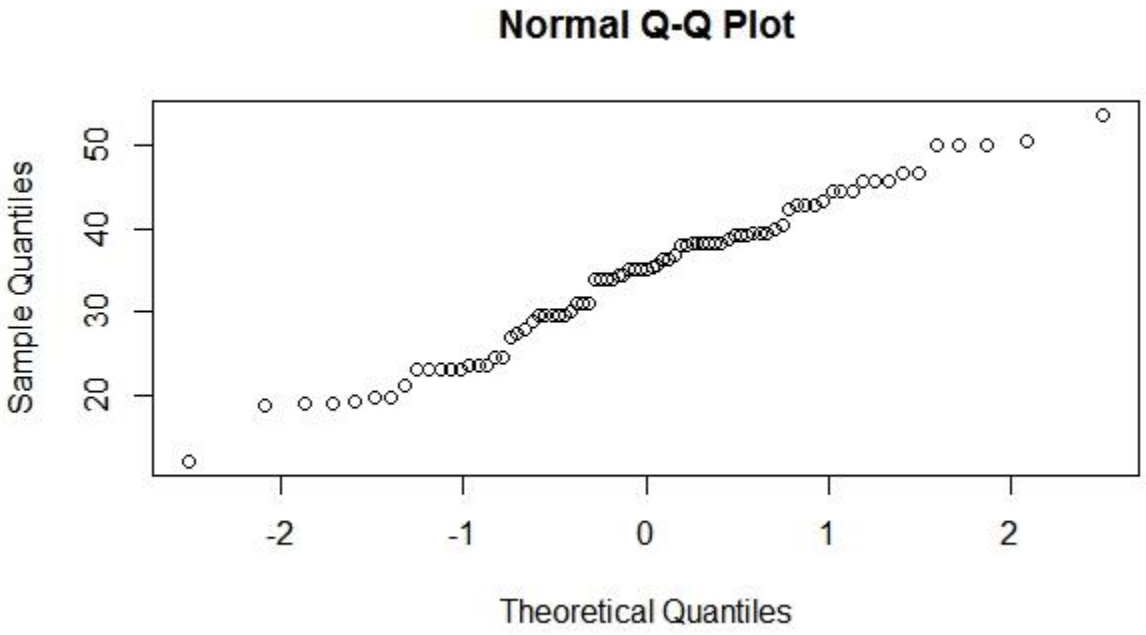
c. P (20<MPG<50)-

```
pnorm(50,34.422,9.13144)-(1-pnorm(20,34.422,9.13144))=
0.01311818
```

Q 21) Check whether the data follows normal distribution

a) Check whether the MPG of Cars follows Normal Distribution
Dataset: Cars.csv

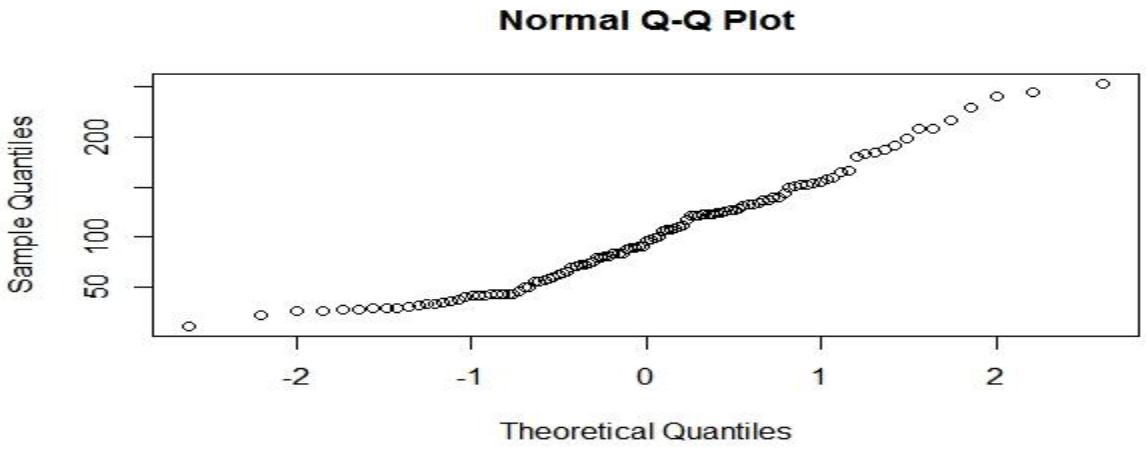
Solution



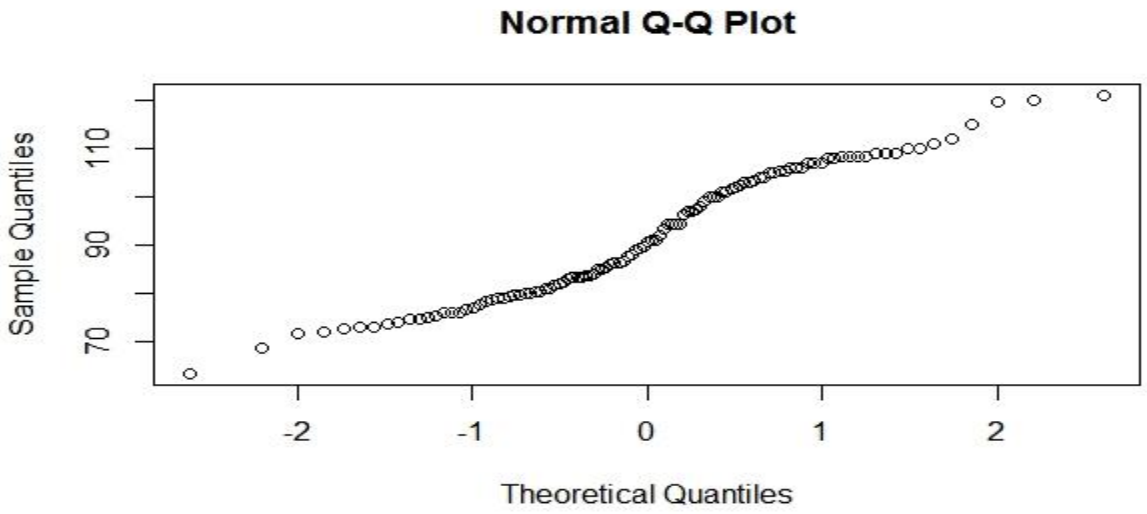
- distributed normally

b) Check Whether the Adipose Tissue (AT) and Waist Circumference(Waist) from wc-at data set follows Normal Distribution

Solution-



Adipose Tissue (AT) normal distributed



-Waist Circumference(Waist) normal distributed

Q 22) Calculate the Z scores of 90% confidence interval,94% confidence interval, 60% confidence interval –

Solution-

Confidence interval	Z scores
60%	0.8416212
90%	1.644854
94%	1.880794

Q 23) Calculate the t scores of 95% confidence interval, 96% confidence interval, 99% confidence interval for sample size of 25.

Solution-

Confidence interval	T scores
95%	2.063899
96%	2.171545
99%	2.79694

Q 24**)** A Government company claims that an average light bulb lasts 270 days. A researcher randomly selects 18 bulbs for testing. The sampled bulbs

last an average of 260 days, with a standard deviation of 90 days. If the CEO's claim were true, what is the probability that 18 randomly selected bulbs would have an average life of no more than 260 days

Solution=52.86%