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	Discrete
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

Q1) Identify the Data type for the Following:

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

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

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

ANS

(P,P,T)+(P,T,P)+(T,P,P)

NO . Of . Possible = 8

1/8+1/8+1/8 = 3/8

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

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

ANS

  NO . Of . possible = 36
  A) Equal to 1
      There is no **possibility** because if you two roll dices it will start from 2 .
  B) Less than or equal to 4
      Less than 4 = (1,1)(1,2)(2,1) = 3 possible 3/36
      Equal to 4 = (2,2)(1,3)(3,1) = 3 possible 3/36
  C) Sum is divisible by 2 and 3
      (6 and 12 are the number divided by both 2 and 3)
      (5,1)(6,6)(1,5)(3,3)(4,2)(2,4)

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

Total number of balls is (2+3+2) = 7

Total number of possible is 7^2 = 21

Probability that get none of the bals is drawn is blue = (2+3) = 5^2 = 10

So probability is 10/21

Q6) 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

The expected number of candies for a randomly selected child

(1*0.015) + (4*0.20) + (3*0.65) + (5*0.005) + (6*0.01) + (2*0.120) = **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.

Use Q7.csv file

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?

ANS

Mean = (180+110+123+134+135+145+167+187+199) = 1308

1308/9 = 145.33

Expected Value of the Weight of that patient = 145.33

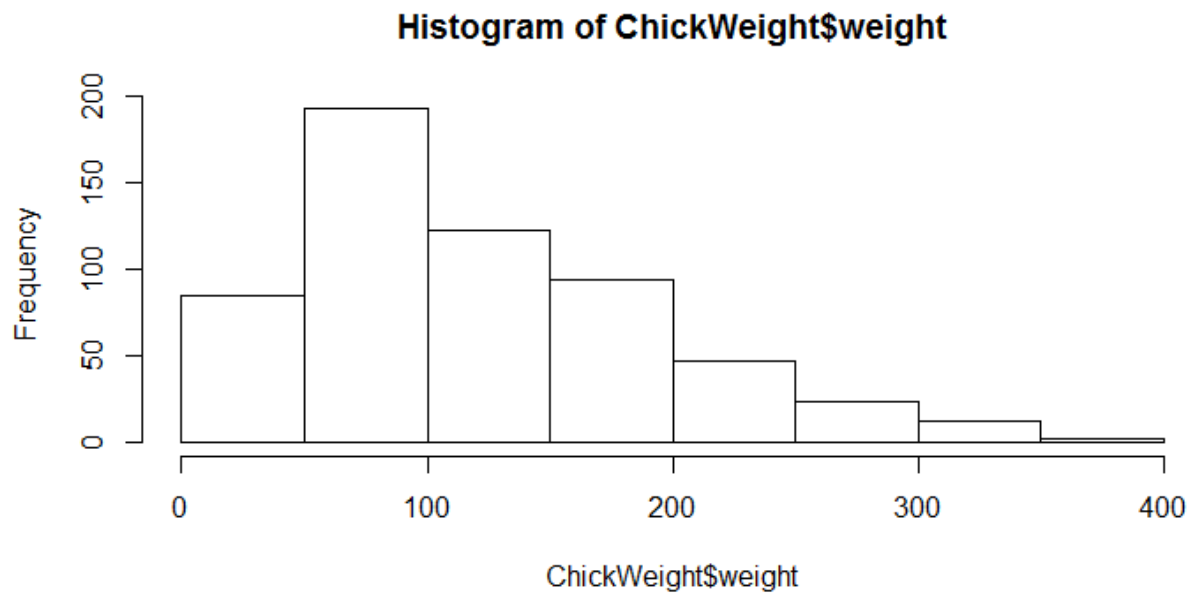**Q9) Calculate Skewness, Kurtosis & draw inferences on the following data**

Cars speed and distance

Use Q9_a.csv

SP and Weight(WT)

Use Q9_b.csv

Q10) Draw inferences about the following boxplot & histogram
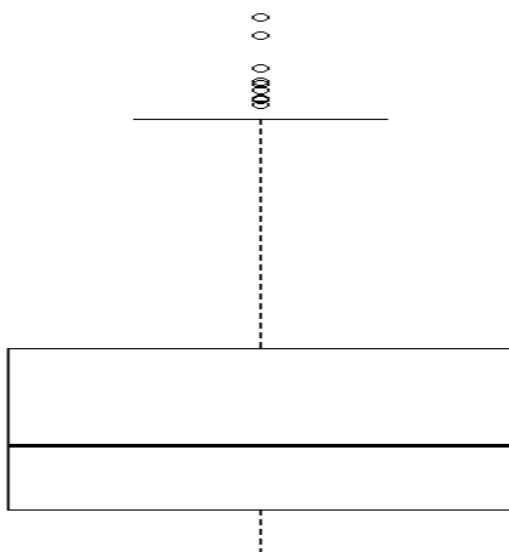
Histogram of ChickWeight$weight

ANS

The higher frequency lies between 50 to 100 chick weight

Range of chick weight is between 0 to 400

Maximum value lies between 50 to 200

It is a Right skewed

ANS

The outliner is at upper side so this is a right skewed

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?

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?

ANS

The mean and median of student marks are 41.0 and 40.5

Mean > median so this is a Right skewed

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

ANS

Null skew, because when mean and median are equal then skewness distribution will be zero

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

ANS

Right skew, because distribution will be towards right

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

ANS

Left skew, because distribution will be towards left

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

ANS

A distribution with a positive kurtosis value indicates that the distribution has heavier tails than the normal distribution

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

ANS

A distribution with a negative kurtosis value indicates that the distribution has lighter tails than the normal distribution

Q18) Answer the below questions using the below boxplot visualization.
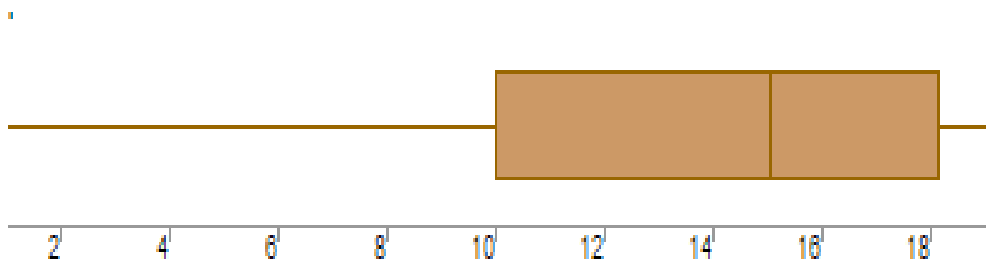
What can we say about the distribution of the data ?

ANS

Normal distribution because mean and median will be more or less equal to each other

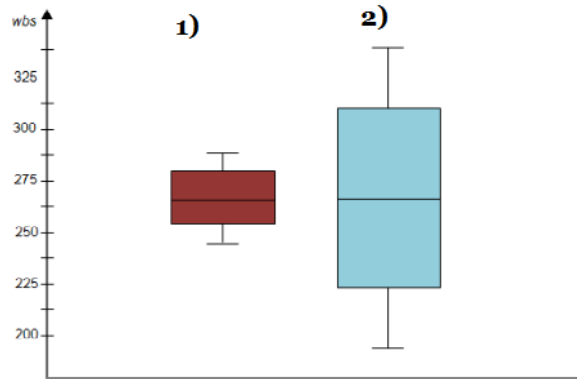What is nature of skewness of the data?

ANS

Negative skewness

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

ANS

IQR of the data will be 15

Q19) Comment on the below Boxplot visualizations?

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

ANS

Both boxplot medians and means are equal

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)
b. P(MPG<40)
c. P (20<MPG<50)

Q 21) Check whether the data follows normal distribution
a) Check whether the MPG of Cars follows Normal Distribution
Dataset: Cars.csv

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

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

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

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

Hint:

rcode → pt(tscore,df)

df → degrees of freedom