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

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 | Interval, ratio |
| Celsius Temperature | Ratio |
| Weight | Ratio |
| Hair Color | Ordinal |
| Socioeconomic Status | Nominal |
| Fahrenheit Temperature | Ratio |
| Height | Ratio |
| Type of living accommodation | ordinal |
| Level of Agreement | Ratio |
| IQ(Intelligence Scale) | Ratio |
| Sales Figures | Ratio |
| Blood Group | nominal |
| Time Of Day | Ratio |
| Time on a Clock with Hands | Ratio, interval |
| Number of Children | Ratio |
| 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?

1/2 New answer - **0.375**

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

1. Equal to 1 = 0
2. Less than or equal to 4 = 66.666 3/36 = 0.08333
3. Sum is divisible by 2 and 3 = - HELP

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?

- HELP

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.

- HELP- did not understand what has to be done

Child B – probability of having 4 candies = 0.20

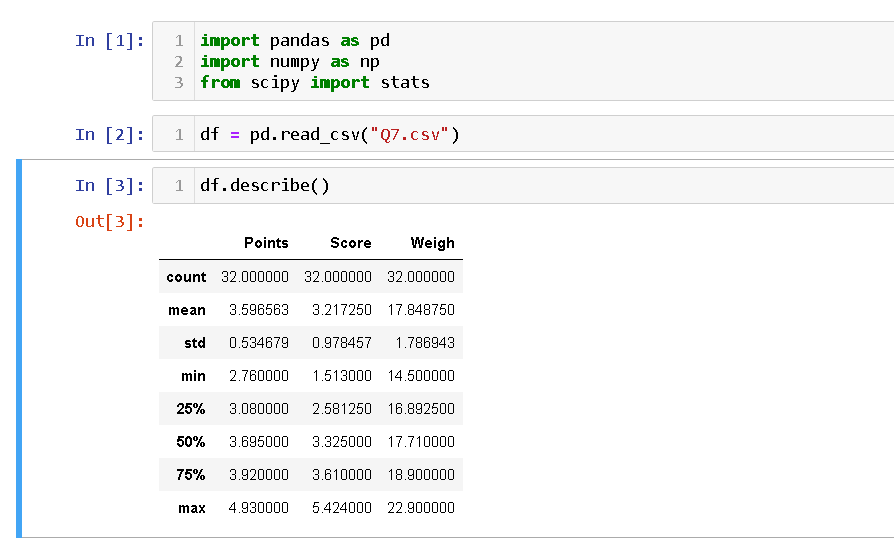
- HELP- did not understand what has to be done

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

1. 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?

- HELP

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

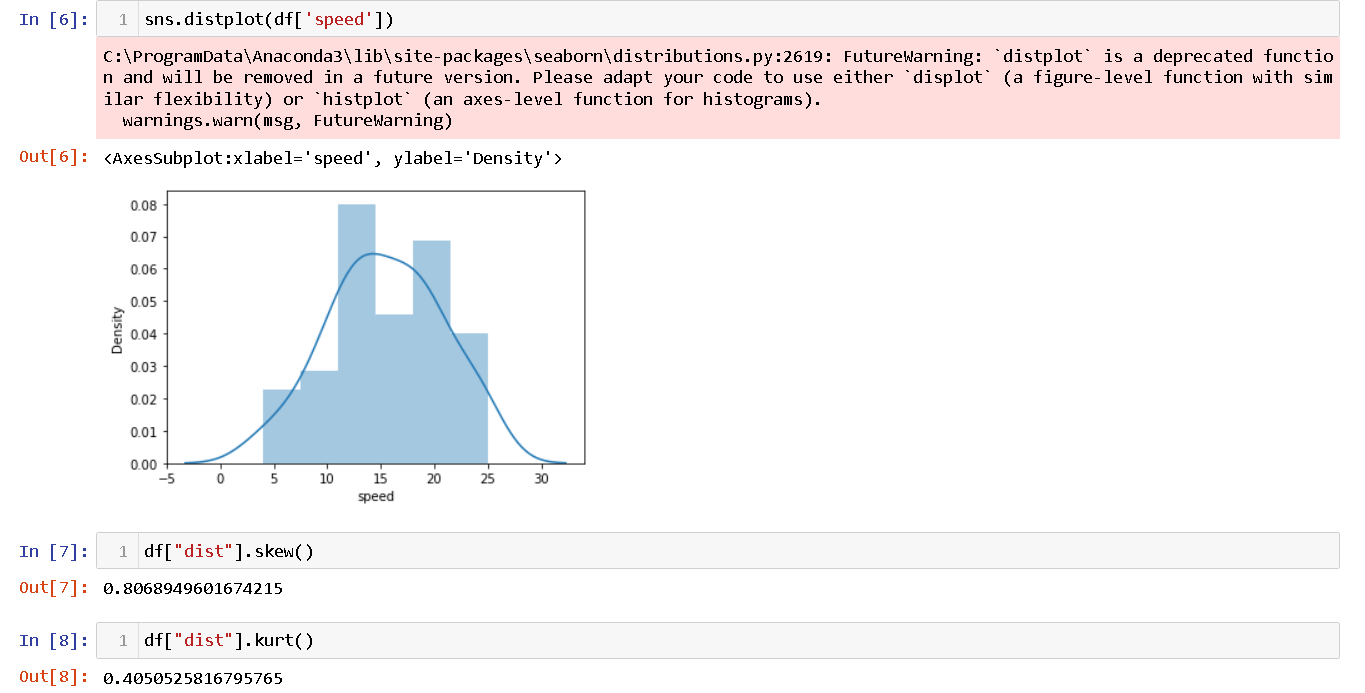
**Cars speed and distance**

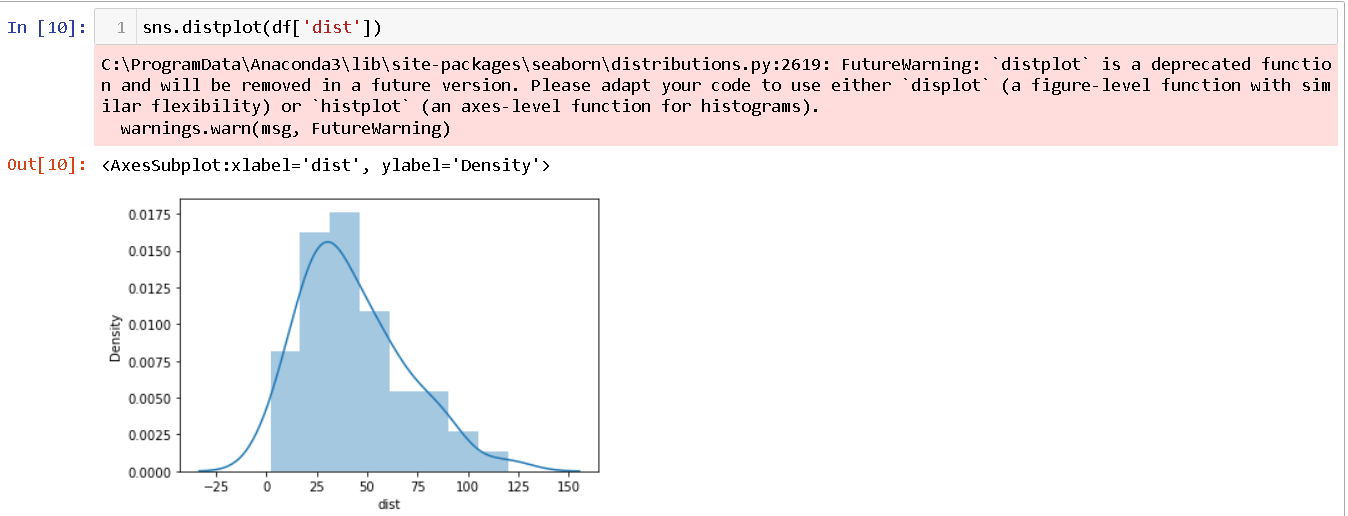
**Use Q9\_a.csv**

speed = -ve skewness and -ve kurtosis

Distance = +ve skewness and +ve kurtosis





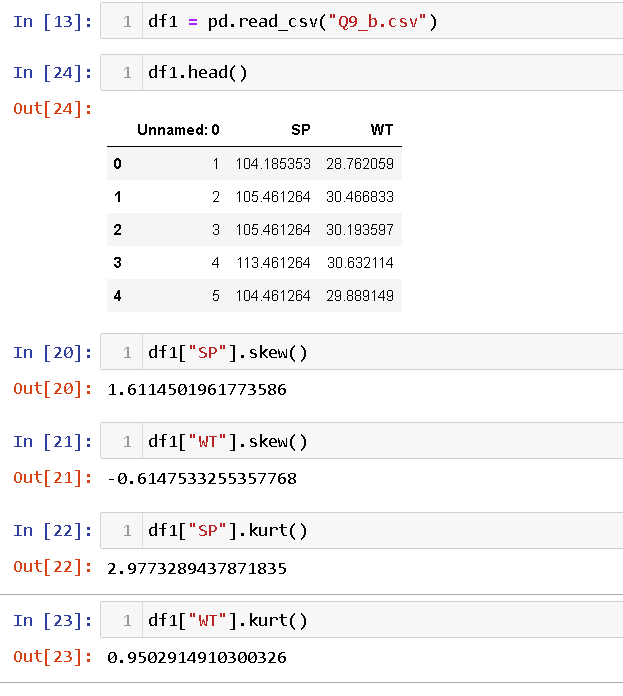


**SP and Weight(WT)**

**Use Q9\_b.csv**

SP = +ve skewness and +ve kurtosis

WT = -ve skewness and +ve kurtosis



**Q10) Draw inferences about the following boxplot & histogram**



IQR

median

lower quartile

upper quartile

outlier

upper extreme

Lower extreme



**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?

- according to class we need population standard deviation and z-score for 94%,98%,96% which is missing. HELP

**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.

Mean -41

Median- 40.5

Variance - 25.529

Standard deviation - 5.052

1. What can we say about the student marks?

- ASSUMING THE TOTAL MARKS IS 100 THE STUDENT MARKS CAN BE SAID AS PASS AS THE MEAN IS 41 MORE THAN 35 PASSING MARKS / HELP

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

Non skew

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

- HELP

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

- HELP

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

The data set has high value

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

The data set have balanced value

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



What can we say about the distribution of the data?

25% of data is in 0-10 scale (lower extreme to lower quartile)

25%-50% of data is in 10-15 scale (lower quartile to median)

50%-75% of data is in 15-18 scale (median to upper quartile)

75% - 100% of data is in 18+ scale (upper quartile to upper extreme)

What is nature of skewness of the data?

-ve skewness most of the data is before median point

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.

- WHAT IS INFERENCE?

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

* 1. P(MPG>38)
  2. P(MPG<40)

c. P (20<MPG<50)

A. 65%

B. 72%

C. 89%



Q 21) Check whether the data follows normal distribution

1. Check whether the MPG of Cars follows Normal Distribution

Dataset: Cars.csv

- help

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

Dataset: wc-at.csv

- help

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

- data set?

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

- data set?

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

- help