|  |  |
| --- | --- |
| **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** | Continuous |
| **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** | Nominal |
| **Level of Agreement** | Ordinal |
| **IQ(Intelligence Scale)** | Interval |
| **Sales Figures** | Ratio |
| **Blood Group** | Nominal |
| **Time Of Day** | Ordinal |
| **Time on a Clock with Hands** | Interval |
| **Number of Children** | Ratio |
| **Religious 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?**

**A)** Formula : Probability of an event (E) = Number of favorable outcomes / Total number of outcomes.

Possible outcomes: (H,H,H), (H,H,T), (H,T,H), (H,T,T), (T,H,H), (T,H,T), (T,T,H), (T,T,T)

Total number of outcomes = 8 and Number of outcomes that gives two heads and one tail = 3

Thus, 3/8 is answer

The probability of getting two heads and one tail on tossing three coins is 3/8.

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

1. **Equal to 1**
2. **Less than or equal to 4**
3. **Sum is divisible by 2 and 3**

**A) a)** Two dice thrown is =36

a) answer is =0 (the sum is equal to 1 is zero because they starts with (1,1) .... likewise . other than in the dice we are not having zero.)

b)answer is = (3/36)=(1/12) (the sum is equal to 4 the possible outcomes are (1,3),(2,2),(3,1) therefore n( b) = 3/36 = 1/12)

c) answer is =(36/36)= 1 (their sum is less than 13. here the total sample space will come . n(c)= 36/36 = 1)

**b)** The set of possible outcomes when we roll a die are {1, 2, 3, 4, 5, 6}

So, when we roll two dice there are 6 × 6 = 36 possibilities.

When we roll two dice, the possibility of getting number 4 is (1, 3), (2, 2), and (3, 1).So, The number of favorable outcomes = 3 ;Total number of possibilities = 36

Probability = The number of favorable outcomes / Total number of possibilities = 3 / 36 = 1/12. Thus, 1/12 is the probability of rolling two dice and getting a sum of 4.

**C)** When two dice are rolled, sample space is given as:

(1,1) (1,2) (1,3) (1,4) (1,5) (1,6)

(2,1) (2,2) (2,3) (2,4) (2,5) (2,6)

(3,1) (3,2) (3,3) (3,4) (3,5) (3,6)

(4,1) (4,2) (4,3) (4,4) (4,5) (4,6)

(5,1) (5,2) (5,3) (5,4) (5,5) (5,6)

(6,1) (6,2) (6,3) (6,4) (6,5) (6,6)

Therefore, Total number of possible outcomes = 36, Favorable outcomes = sum is divisible by 2 and 3, Sum should be divisible by both 2 and 3

Favorable outcomes = (1 , 5) , (3 , 3) , (4 , 2) , (5 , 1) , (6 , 6)

Therefore,

Number of favorable outcomes = 5

The probability of an event is given as: Thus the probability that sum is divisible by 2 and 3 is 5/36

**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?**

**A)** Total number of balls = (2 + 3 + 2) = 7

Let S be the sample space.

Then, n(S) = Number of ways of drawing 2 balls out of 7

= 7C 2

​= (7×6)/(2×1) = 21

Let E = Event of drawing 2 balls, none of which is blue.

∴n(E)= Number of ways of drawing 2 balls out of (2 + 3) balls.

= 5C2

= (5×4)/(2×1) = 10

∴P(E)= n(S) n(E) = 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**

**A)** Expected number of candies for a randomly selected child = 3.09

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 \* 0.12

= 0.015 + 0.8 + 1.95 + 0.025 + 0.06 + 0.24

= 3.090

= 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**

**A)** <https://github.com/Aashish7979/Basic-statistics-level-1-Assignmentss/blob/ac0028b3fe45e8b889928f50a242f194c4ab2de3/Q7_Basic_Stats_L1.ipynb>

**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?**

**A)** Expected Value = ∑ ( probability \* Value )

∑ P(x).E(x)

there are 9 patients

Probability of selecting each patient = 1/9

Ex = 108, 110, 123, 134, 135, 145, 167, 187, 199

P(x) = 1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9

Expected Value = (1/9)(108) + (1/9)110 + (1/9)123 + (1/9)134 + (1/9)135 + (1/9)145 + (1/9(167) + (1/9)187 + (1/9)199

= (1/9) ( 108 + 110 + 123 + 134 + 135 + 145 + 167 + 187 + 199)

= (1/9) ( 1308)

= 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**

**A)** <https://github.com/Aashish7979/Basic-statistics-level-1-Assignmentss/blob/ac0028b3fe45e8b889928f50a242f194c4ab2de3/Q9a_Basic_Stats_L1.ipynb>

**SP and Weight(WT)**

**Use Q9\_b.csv**

**A)** <https://github.com/Aashish7979/Basic-statistics-level-1-Assignmentss/blob/ac0028b3fe45e8b889928f50a242f194c4ab2de3/Q9b_Basic_Stats_L1.ipynb>

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



**A)** The Histogram is right side skewed or positively skewed and the interface for this box plot is positively 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?**

**A)** <https://github.com/Aashish7979/Basic-statistics-level-1-Assignmentss/blob/ac0028b3fe45e8b889928f50a242f194c4ab2de3/Q11_Basic_Stats_L1.ipynb>

**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?**

**A)** <https://github.com/Aashish7979/Basic-statistics-level-1-Assignmentss/blob/b0d53c19db730cbefe5f7cfbd7f40774410f6622/Q12_Basic_Stats_L1.ipynb>

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

1. This means skewness = 0, perfectly symmetrical or Normalized skewness.

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

1. Right skewed

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

1. Left skewed

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

1. Sharp peak in the plot and less gap between tails to X-axis.

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

1. Flat peak and border peak under the curve and more gap between tails and x-axis

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



**What can we say about the distribution of the data?**

**A)** The data is distributed in D-assigned format.

**What is nature of skewness of the data?**

**A)** Left side skewed

**What will be the IQR of the data (approximately)?   
A) =** Q3-Q1; = 18-10; = 8 is IQR

**Q19) Comment on the below Boxplot visualizations?**



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

1. The box plot 1 designed with range = 3; The second one range is = 1.5

**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)** <https://github.com/Aashish7979/Basic-statistics-level-1-Assignmentss/blob/cb42969801bbb8d53d9a588803a529e6932243ab/Q20_Basic_Stats_L1.R>

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

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

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

**Q 21) Check whether the data follows normal distribution**

1. **Check whether the MPG of Cars follows Normal Distribution**

**Dataset: Cars.csv**

1. <https://github.com/Aashish7979/Basic-statistics-level-1-Assignmentss/blob/d2bb3731ce8509a70e1e5c6af310da18742f94d7/Q21a_Basic_Stats_L1.ipynb>
2. **Check Whether the Adipose Tissue (AT) and Waist Circumference(Waist) from wc-at data set follows Normal Distribution**

**Dataset: wc-at.csv**

1. <https://github.com/Aashish7979/Basic-statistics-level-1-Assignmentss/blob/413c858af1aef4b4cce935e8601199664886b9da/Q21b__Basic_Stats_L1.ipynb>

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

**A)** <https://github.com/Aashish7979/Basic-statistics-level-1-Assignmentss/blob/2e5b53f12abf060fcdebe59f5dc0e6060371f94a/Q22_Basic_Stats_L1.ipynb>

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

**A)** <https://github.com/Aashish7979/Basic-statistics-level-1-Assignmentss/blob/a320961dafd7f0184dd98c596c06c13befd8d6dd/Q23_Basic_Stats_L1.ipynb>

**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**

1. <https://github.com/Aashish7979/Basic-statistics-level-1-Assignmentss/blob/7dcbceec270081039fa258a6c51114dbfc077706/Q24_Basic_Stats_L1.ipynb>