Statistics_Assignment1

September 19, 2023

Q1. What is Statistics?

[ ]:
```
Ans: Statistics is the science of collecting,organizing,and analyzing the data.
```

Q2. Define the different types of statistics and give an example of when each type might be used.

[ ]:
```
Ans:
Types of Statistics :
    1)Descriptive Statistics
    Def: It consists of organizing and summarising the data.
    example:Average height of students in a class room (sum(x1,x2,x3,...xn)/n)␣
 ↪where x->heights of students,n->number of students.
        i)Measure of Entral(Mean,Median,Mode)
        ii)Measure of Dispersion (Variance,standard deviation)
        iii) Different types of distribution of data
        eg: Histogram,pdf.
    2)Inferential Statistics
    Def:It consists of using data you have measured to form a conclusion.
    Example: Uisng the sample data (heights and number of students of a class␣
 ↪room) predict population data(Average height of entire college students)
```

Q3. What are the different types of data and how do they differ from each other? Provide an example of each type of data.

[ ]:
```
Ans:
Data:
    1) Quantitative Data (Numerical +,-,/,*)
        i) Discrete (whole number)(Number of bank account)
        ii) Continuous (Weight,height,Temperature)
    2) Qualitative Data (Categorical)
        i)Nominal (Gender-M/F,Blood gp)
        ii) Ordinal (Rank)(Customer feedback Good/Bad/Better)
```

Q4. Categorise the following datasets with respect to quantitative and qualitative data types: (i) Grading in exam: A+, A, B+, B, C+, C, D, E (ii) Colour of mangoes: yellow, green, orange, red (iii) Height data of a class: [178.9, 179, 179.5, 176, 177.2, 178.3, 175.8,...] (iv) Number of mangoes exported by a farm: [500, 600, 478, 672, ...]

[ ]:
```
Ans:
(i) Grading in exam: A+, A, B+, B, C+, C, D, E --> qualitative (Ordinal)
(ii) Colour of mangoes: yellow, green, orange, red --> qualitative (Ordinal)
(iii) Height data of a class: [178.9, 179, 179.5, 176, 177.2, 178.3, 175.8,...]␣
 ↪--> quantitative (Continuous)
(iv) Number of mangoes exported by a farm: [500, 600, 478, 672, ...]-->␣
 ↪quantitative (discrete)
```

Q5. Explain the concept of levels of measurement and give an example of a variable for each level.

[ ]:
```
Ans:
Concept : We need to do measurement of data once we get the business data.It␣
 ↪will be helpful for data analysis and data science, basically it will help␣
 ↪us to find the patterns to solve business problems.
1) Nominal Scale Data
2) Ordinal Scale Data
3) Interval Scale Data
4) Ratio Scale Data

1) Nominal Scale Data:
    i) Qualitative / Categorical
       eg: Gender,Colors
    ii) Order does not matters
Let say we have 10 students in a class.We conduct an Survey which color they␣
 ↪like most between Red,Green,Blue.
5 students selected Red,3 Green,2 Organge.Here Red !> Blue (Not depends on Rank)

2) Ordinal Scale Data:
    i) Ranking is imprtant
    ii) Order is matters
    iii) Difference can be measured
eg: App Reveiw.
1->Best ,2-> Good ,3->Bad

3) Interval Scale Data :
    i) Order is matters
    ii) Difference can be measured
    iii)Ratio can be measured
    iv) No True Zero setting point

eg: Temperature Variable : 30F,60F,90F,120F

4) Ratio Scale Data  :
    i) Order is matters
    ii) Difference can be measured
    iii)Ratio can be measured
    iv) True Zero setting point present
```

```
eg: Students marks in a class in ascending order
30,40,50,60

Ratio = 60/30=2:1
```

Q6. Why is it important to understand the level of measurement when analyzing data? Provide an example to illustrate your answer.

[ ]:
```
Ans:
   To understand, change, and improve the data and found business ideas and␣
 ↪implement to solve big problems.
   eg:
   1)Let say we have 10 students in a class.We conduct an Survey which color␣
 ↪they like most between Red,Green,Blue.
5 students selected Red,3 Green,2 Organge.Here Red !> Blue (Not depends on␣
 ↪Rank). (Nominal)
2)eg: App Reveiw.
1->Best ,2-> Good ,3->Bad (ordinal)
```

Q7. How nominal data type is different from ordinal data type.

[ ]:
```
Ans:
    1) Nominal Scale Data:
    i) Qualitative / Categorical
       eg: Gender,Colors
    ii) Order does not matters
Let say we have 10 students in a class.We conduct an Survey which color they␣
 ↪like most between Red,Green,Blue.
5 students selected Red,3 Green,2 Organge.Here Red !> Blue (Not depends on Rank)

2) Ordinal Scale Data:
    i) Ranking is imprtant
    ii) Order is matters
    iii) Difference can be measured
eg: App Reveiw.
1->Best ,2-> Good ,3->Bad
```

Q8. Which type of plot can be used to display data in terms of range?

[ ]:
```
Ans:
    Histogram
```

Q9. Describe the difference between descriptive and inferential statistics. Give an example of each type of statistics and explain how they are used.

[ ]:
```
Ans:
     1)Descriptive Statistics
    Def: It consists of organizing and summarising the data.
```

```
   example:Average height of students in a class room (sum(x1,x2,x3,...xn)/n)␣
↪where x->heights of students,n->number of students.
    2)Inferential Statistics
   Def:It consists of using data you have measured to form a conclusion.
   Example: Uisng the sample data (heights and number of students of a class␣
↪room) predict population data(Average height of entire college students)
```

Q10. What are some common measures of central tendency and variability used in statistics? Explain how each measure can be used to describe a dataset.

[ ]:
```
Ans:
    Measure of Central Tendency :
    1) Mean or Average
    2) Median
    3) Mode

    1) Mean:
    Let, Population ->N
    Sasmple ->n
    x={1,1,2,2,3,3,4,4,4,5}
    Population Mean(miu(u)) = Sum i=1 to N (xi/N)
    Sample Mean (x bar)= Sum i=1 to n (xi/n)

    2) Medican :

    x={1,2,2,3,4,5}
    No of elements = 6

    If count == even
    2+3/2=2.5 will be the Median

    If count == odd
    x={1,2,2,3,4,5,6}
    Median= 3

    3) Mode:
    Frequecy (Maximum occurance of an element)
    {2,1,1,1,4,6,7}
    Mode= 1
```

[ ]:
```

```