Assignment NO. 1 week1-week4

Student Full Name:_____ .

Student ID:_____ .

CRN No:_____ .

Branch: _____.

STATISTICS
(STAT-101)

Total Points

True/False   ____/6

MCQ   ____/6

Short Answer   ____/18

Total   ____/30

Good Luck

STATISTICS (STAT-101)

Marks- 30

Answer all the Questions on the same question paper.

Section-I

State whether the following statements are True or False. (6 marks*, 1 Mark Each)*

1. The ages (in years) of survey respondents are quantitative data.	T
2. "In a given class, 14 students had the academic major of Business Administration, 12 students had E-commerce, and 7 students had Accounting". This is an example based on quantitative data.	F
3. A parameter is a numerical measurement describing some characteristic of a sample.	F
4. A histogram and a relative frequency histogram constructed from the same data always have the same basic shape, but the vertical scales are different.	T
5. A distribution of data is skewed to the right if mean and median are to the right of the mode	T
6. S = {Head, Tail} is the Sample Space for the experiment of flipping two coins.	F

(Multiple Choice Questions)   *(6 marks, 1 Mark Each)*

1. If you classified the files in a folder as pdf, word or in excel files, this would be an example of which level of measurement
    a. Nominal
    b. Ordinal
    c. Interval
    d. Ratio

2. _____ are the numbers used to separate classes, but without the gaps created by class limits
    a. Lower Class Limits
    b. Upper Class Limits
    c. Class Boundaries
    d. Class Midpoints

3. Which graph is best for paired data consisting of the advertising expenditures and sales of 30 randomly selected companies?
    a. Histogram
    b. Dot plot
    c. Scatter plot
    d. Pie chart

4. Which of the followings is not a measure of center:
    a. Median
    b. Midrange
    c. Range
    d. None of the above

5. Consider the following sample data values: 170, 430, 470, 300, 600. Find the standard deviation:
    a. 394
    b. 147.3
    c. 164.7
    d. None of the above

6. P(A or B) =
    a. P(event A occurs in a first trial and event B occurs in a second trial)
    b. P(event A and event B occur in the same trial)

c. P(event A occurring after it is assumed that event B has already occurred)
d. P(either event A occurs or event B occurs or they both occur as the single outcome of the procedure)

Filling the table below is mandatory!

| MCQ	1	2	3	4	5	6
Answers	a	c	c	c	C	d

Section –III

Answer the following Essay Type Questions   (18 marks, *3 Mark Each*)

1. Listed below are blood groups of O, A, B, and AB of 50 randomly selected blood donors.
   O, A, B, O, O, O, O, O, AB, O, O, O, O, B, O, B, O, A, A, A, B, O, A, A, AB, A, B, A, A, A, O, B, A, O, A, O, O, A, A ,B, AB, B, O, O, O, O, A, A, AB and A:
      a. Construct a table summarizing the frequency distribution of these blood groups.
      b. Draw a bar graph to represent this data set.

Blood group	Frequency
O	21
A	17
B	8
AB	4
Total	50
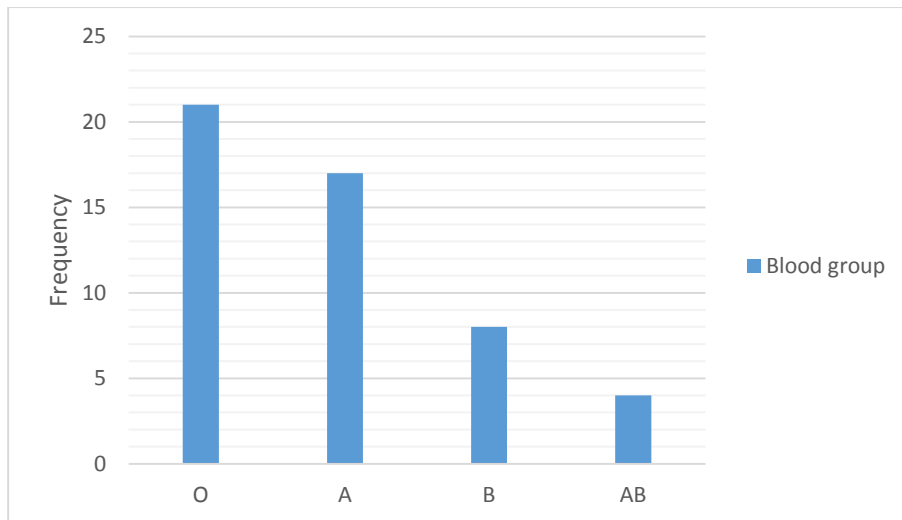
2. The table below represents frequency distribution of the measured cotinine levels of a simple random sample of 40 smokers.

Cotinine Level (in ng/ ml)	Frequency
0 - 99	11
100 - 199	12
200 - 299	14
300 - 399	1
400 - 499	2
Total	40

    a. What is the class width in this table?
    b. Compute the relative frequencies.
    c. Compute the cumulative frequencies.
    d. Compute the <u>cumulative relative</u> frequencies.

Class width = 100 – 0 = 100.

Cotinine Level (in ng/ ml)	Frequency	Relative frequency (%)
0 - 99	11	27.5
100 - 199	12	30
200 - 299	14	35
300 - 399	1	2.5
400 - 499	2	5
Total	40	100

Cotinine Level (in ng/ ml)	Cumulative frequency	Cumulative relative

		frequency (%)
Less than 100	11	27.5
Less than 200	23	57.5
Less than 300	37	92.5
Less than 400	38	95
Less than 500	40	100

3. The company XYZ collects the service times of 6 randomly selected customers from its sales outlet A. The service times in the sales outlet A (in minute) are the following:

6.3	7.2	6.7	7	6.6	7

Find the mean, Mode and Median of the data set of the sales outlet A.

Mean = (6.3 + 7.2 + 6.7 + 7 + 6.6 + 7)/6 =6.8
Mode = 7 (highest frequency)
Median? Data set values in an increasing order:
6.3   6.6   6.7   7   7   7.2
Median = (6.7 + 7)/2 = 6.85

4. Consider the sample data values in sales outlet A in question 3. The company XYZ collects sample service times of its sales outlet B, where the mean = 2.8 minutes and the standard deviation = 0.28 minutes. Compare the coefficient of variations in sample data of both stores A and B.

Means of the samples are very different, so we have to use the coefficient of variation to compare the variations in both samples.
$S_A=\sqrt{(6.3 - 6.8)^2 + (7.2 - 6.8)^2 + \cdots + (7 - 6.8)^2}$=0.33
$CV_A=S_A/\bar{x}_A$ .100= 0.33/6.8 = 4.83%
$CV_B=S_B/\bar{x}_B$ .100= 0.28/2.8 = 10%
$CV_B > CV_A$ so the sales outlet B has a greater variability in customers service times.

5. 45% of the children in a school have food allergies, 30% have seasonal allergies, and 18% have food and seasonal allergies.
What is the probability of those who have food allergies given that they have seasonal allergies?

Event A: Have a food allergy,

Event B: Have a seasonal allergy,

$P(A|B) = \frac{P(A \ and \ B)}{P(B)} = 0.18/0.30 = 60\%$

6. Refer to the question 5 statement. What is the probability of a randomly selected child is a food or a seasonal allergic?
What is the probability that a randomly selected child has no allergy?

$P(A \ or \ B) = P(A) + P(B) - P(A \ and \ B)$
$\qquad = 0.45 + 0.3 - 0.18 = 0.57$
$P(\overline{A \ or \ B}) = 1 - P(A \ or \ B) = 1 - 0.57 = 0.43$