Please circle your:

<table>
<thead>
<tr>
<th>Instructor Name</th>
<th>Section (time)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anwar Joarder</td>
<td>6 (8:00 – 8:50)</td>
</tr>
<tr>
<td>Mohammad H. Omar</td>
<td>1 (9:00 – 9:50)</td>
</tr>
<tr>
<td>Mohammad Saleh</td>
<td>3 (7:00 – 7:50)</td>
</tr>
<tr>
<td></td>
<td>2 (10:00 – 10:50)</td>
</tr>
<tr>
<td></td>
<td>5 (11:00 – 11:50)</td>
</tr>
</tbody>
</table>

Directions:
1) You must show all work to obtain full credit for questions on this exam.
2) DO NOT round your answers at each step. Round answers only if necessary at your final step to 4 decimal places.

<table>
<thead>
<tr>
<th>Question No</th>
<th>Full Marks</th>
<th>Marks Obtained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Q2</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Q4</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Q5</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Q6</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Q7</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
Solution for Question Six

<table>
<thead>
<tr>
<th>Question number</th>
<th>Correct answer</th>
<th>Question number</th>
<th>Correct answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Solution for Question Seven

<table>
<thead>
<tr>
<th>Question number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
</tbody>
</table>
Question One (5 points)
Before leaving a particular restaurant, customers are asked to respond to the questions listed below. For each question, determine whether the possible responses are interval, nominal, ordinal or ratio.

1. What is the approximate distance of the restaurant from your residence?

2. Have you eaten at the restaurant previously?

3. If your answer to part (b) was yes, on how many occasions?

4. Which of the following attributes of the restaurant do you find most attractive: service, prices, quality of the food, or varied menu?

5. Would your overall rating of the restaurant be excellent, good, fair, or poor?

Question Two (4+4+2+10= 20 points)
The ordered array below resulted from taking a sample of 25 batches of 500 computer chips and determining how many in each batch were defective.

<table>
<thead>
<tr>
<th>Defects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 4 4 5 5 6 7 9 9 12 12 15 17 20 21 23 23 25 26 27 27 28 29 29</td>
</tr>
</tbody>
</table>

1. If a frequency distribution for the defects data is constructed, using "0 but less than 5" as the first class, the frequency of the "20 but less than 25" class would be what?

2. Construct a Stem-and-Leaf plot of this data, comment on the shape

3. What type of variable is defects?

4. Construct a box plot of defects, Identify the outliers (if any!), and Determine whether the data is symmetric.
Question Three (6+6+3 = 15 points)

A water pump manufacturing company sells model I at $100 per unit, model II at $125 per unit, model III at $175 per unit, model IV at $250 per unit and model V at $600 per unit. Based on a sample of size 1000 units sold, the relative frequency distribution of sales of various models is 0.3, 0.25, 0.22 and 0.13 for models I, II, III, IV respectively.

1. Find the cumulative frequency distribution of models sold.

2. Compute the average selling price per unit, the median selling price per unit and the corresponding mode.

3. What can you say about the skewness of this distribution?
Question Four (3+9+4+4=20 points)
You are designing a study of students’ preference for food at the cafeteria of the College of Industrial Management lounge. Design your study by doing the following:

1. List **three** important variables about preference for food you must collect.  
   *(HINT: don’t list demographic questions)*

2. For each of these variables, provide description for its correct data type and measurement level

3. What data collection method is the most appropriate? Provide your rationale

4. What is the most appropriate sampling method if you need to sample 30 students to represent the whole KFUPM student population? Provide your rationale
Question Five (6+9+5=20 points)
In a comparative study of traffic congestion in KSA cities, the time needed to drive from the Central Business District (CBD) to the airport in two different cities has been observed for 50 working days during the afternoon rush hour. The respective frequency distributions of the CBD – airport travel time for the two cities are shown hereafter:

<table>
<thead>
<tr>
<th>Travel time (in minutes)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 to under 15</td>
<td>City I: 5</td>
</tr>
<tr>
<td></td>
<td>City II: 10</td>
</tr>
<tr>
<td>15 to under 20</td>
<td>City I: 10</td>
</tr>
<tr>
<td></td>
<td>City II: 15</td>
</tr>
<tr>
<td>20 to under 25</td>
<td>City I: 30</td>
</tr>
<tr>
<td></td>
<td>City II: 15</td>
</tr>
<tr>
<td>25 to under 30</td>
<td>City I: 5</td>
</tr>
<tr>
<td></td>
<td>City II: 10</td>
</tr>
</tbody>
</table>

1. Calculate the mean travel times in each city.

2. Calculate the corresponding coefficients of variation. Which city exhibits greater relative variability in the travel time from the CBD to the airport?

3. Find the cumulative frequency distribution for the sample obtained in city I, and sketch it graphically. What is the median travel time interval?
Question Six (10 points)
Circle the correct answer
1. You asked five of your classmates about their height. On the basis of this information, you stated that the average height of all students in your university or college is 67 inches. This is an example of:
   a. Descriptive statistics
   b. Statistical inference
   c. Parameter
   d. Population
   e. Statistics

2. When the population is divided into mutually exclusive sets, and then a simple random sample is drawn from each set, this is called
   a. Stratified random sampling
   b. Simple random sampling
   c. Cluster sampling
   d. Systematic random sampling
   e. Selection bias

3. Since the population is always larger than the sample, the population mean:
   a. is always larger than the sample mean
   b. is always smaller than the sample mean
   c. is always larger than or equal to the sample mean
   d. is always smaller than or equal to the sample mean
   e. can be smaller than, or larger than, or equal to the sample mean

4. Which of the following is a continuous quantitative variable?
   a. the amount of milk produced by a cow in one 24-hour period
   b. the color of a student's eyes
   c. the number of gallons of milk sold at the local grocery store yesterday
   d. the number of employees of an insurance company
   e. All of the above.

5. When a questionnaire is poorly worded, the kind of survey error that could possibly occur is
   a. Hawthorne effect
   b. Selection bias
   c. Measurement error
   d. Nonresponse error
   e. Sampling error.

6. The **best type** of chart for comparing two sets of categorical data is:
   a. a line chart
   b. a pie chart
   c. a histogram
   d. a bar chart
   e. a stem and leaf plot.

7. In a right-skewed distribution, which of the following is correct?
   a. The median is less than the arithmetic mean.
   b. The median is larger than the arithmetic mean.
   c. The median equals the arithmetic mean
   d. The geometric mean equals the arithmetic mean
   e. The median is less than the geometric mean.
8. A marketing research firm, in conducting a comparative taste test, provided three types of peanut butter to a sample of households randomly selected within the state. Which of the following methods of data collection is involved when people are asked to compare the three types of peanut butter?
   a. Surveying
   b. Observation
   c. Published sources
   d. Online sources
   e. Experimentation

9. The following random numbers were obtained from a random number table to draw a sample from a population of employees at Al-Marai company:
   82 30 55 14 32 23 45 09 55 94

Which of the following method is the most likely sampling method used to obtain this sample?
   a. Simple random sampling with replacement
   b. Systematic random sampling
   c. Simple random sampling without replacement
   d. Cluster sampling
   e. Convenience sampling

10. Which of the following statements about the arithmetic mean is not always correct
   a. The sum of the deviations from the mean is zero
   b. Half of the observations are on either side of the mean
   c. The mean is a measure of the center of a distribution
   d. The value of the mean times the number of observations equals the sum of all of the observations
   e. The mean is the most common measure of central tendency.

---

**Question Seven (10 points)**

Which of the following statements true and which false

1. AlMarai inventory cost from January to December of 2007 is considered a time series data.__________

2. Data collected using internet online survey is an example of a primary data.__________

3. “Wouldn’t you agree that foreign cars are better than American cars?” is an example of leading questions.__________

4. The largest value in a set of data is 140, and the lowest value is 70. If the resulting frequency distribution is to have five classes of equal width, the class width will be 14.__________

5. If we draw a straight line through the points in a scatter diagram and most of the points fall close to the line, we say that there is a positive linear relationship between the two variables.__________

6. A data sample has a mean of 107, a median of 122, and a mode of 134. The distribution of the data is positively skewed.__________

7. The coefficient of variation allows us to compare two sets of data based on different measurement units.__________

8. Since the sample is always smaller than the population, the sample mean is always smaller than the population mean.__________

9. The median of the values 3.4, 4.7, 1.9, 7.6, and 6.5 is 1.9.__________

10. As a general rule, an observation is considered an extreme value if it’s Z score is less than 3.__________