

Assignment 1 – Part 2 - Solution

1. **Due Date & Time: Tuesday, May 21th, 2013 by 23:59 hrs.**
2. **You MUST Upload the Assignment on MyStatLab in “pdf” format.**
3. **You MUST also attach the Printed/ Signed “Integrity Statement”.**
4. **Use MiniTab if or Excel if you have to. However you should cut and paste your result and sufficient explanation.**
5. **You should provide detailed explanation for each of your solution. You must show your complete manual calculation if the question involves using formula or any kind of calculations.**

Question 1. (10 marks)

The following table displays some of the data collected by the above research firm in their study on online shopping. List the variables in the data set. Indicate whether each variable is categorical or quantitative. If the variable is quantitative, give the units.

<i>Age</i>	<i>Gender</i>	<i>Household Income</i>	<i>Participated in online auction?</i>
25	Male	\$55,000	Yes
47	Female	\$60,000	No
38	Female	\$110,000	No
30	Male	\$62,000	No

Answer:

Categorical: Gender, Participated in online auction? (2 marks each)

Quantitative: Age (years), Household Income (\$). (2 marks for each variable, 2 marks for each unit)

Question 2. (10 marks)

For each of the following, indicate whether the data are nominal, ordinal, interval or ratio data.

a) A respondent of a survey indicates that he is a resident in the state of Ohio.

Answer : nominal data. (2 marks)

b) The Graduate Management Admission Test (GMAT) is a standardized test used by schools to determine the aptitude of individuals who are applying for MBA programs. The range of the GMAT score is 200-800. Brian has recently taken the exam and scored 720.

Answer : interval data. (2 marks)

c) A respondent of a survey indicates that she is currently in her junior year at the University of Texas.

Answer : ordinal data. (2 marks)

d) A respondent of a survey indicates that he owns the home that he currently resides.

Answer : nominal data. (2 marks)

e) The number of iPhones sold today at an Apple store

Answer : ratio data. (2 marks)

Question 3. (4 marks)

For each of the following, indicate best data type that represents them among: cross-sectional, nominal, time series or ordinal.

a) The number of televisions sold at the local Best Buy store over the past seven days.

Day	Number Sold
Monday	5
Tuesday	9
Wednesday	5
Thursday	7
Friday	12
Saturday	19
Sunday	15

Answer : time series data (2 marks)

b) the number of people, in thousands, in the United States at various levels of household income in 2010.

Household Income	Number of People
Less than \$25,000	60,140
\$25,000 to \$49,999	70,680
\$50,000 to \$74,999	57,359
\$75,000 or more	117,931

Answer : cross-sectional (2 marks)

Question 4. (10 marks)

The following contingency table shows the number of customers who bought various brands of inkjet printers and their gender.

Printer Brand	Men	Women
HP	30	21
Canon	18	6
Epson	9	6
Brother	3	3
Other	24	30

a) What is the probability that a randomly selected customer from this sample bought a Canon printer or a Brother printer?

Answer: $P(\text{Canon or Brother}) = \frac{24}{150} + \frac{6}{150} = \frac{30}{150} = 0.20$ (2 marks)

b) What is the probability that a randomly selected customer from this sample bought an HP printer or is a female?

Answer: $P(\text{HP or Female}) = \frac{51}{150} + \frac{66}{150} - \frac{21}{150} = \frac{96}{150} = 0.64$ (2 marks)

c) The probability that a randomly selected customer from this sample bought an Epson printer and is a male ?

Answer: $P(\text{Epson and Male}) = \frac{9}{150} = 0.06$ (2 marks)

d) Given that a randomly selected printer from this sample was a Canon, determine the probability that it was purchased by a man.

Answer: $P(\text{Man} | \text{Canon}) = \frac{P(\text{Man and Canon})}{P(\text{Canon})}$ (4 marks)

$$P(\text{Man and Canon}) = \frac{18}{150} = 0.12$$

$$P(\text{Canon}) = \frac{24}{150} = 0.16$$

$$P(\text{Man} | \text{Canon}) = \frac{0.12}{0.16} = 0.75$$

Question 5. (10 marks)

A college student population of freshman and seniors was surveyed to determine if they prefer traditional textbooks or digital textbooks. According to this sample, 60% of freshman students preferred digital textbooks. Forty percent of the sample consisted of freshman students. Also, 20% of senior students preferred digital textbooks. Given that a student prefers digital textbooks, use Bayes' Theorem to determine the probability that the student is a freshman.

Answer: Define A_1 = student is a freshman, A_2 = student is a senior, B = student prefers a digital textbook.

$$P(A_1) = 0.40 \text{ (2 marks)}$$

$$P(A_2) = 1 - 0.40 = 0.60 \text{ (2 marks)}$$

$$P(B | A_1) = 0.60 \text{ (2 marks)}$$

$$P(B | A_2) = 0.20 \text{ (2 marks)}$$

$$P(A_1 | B) = \frac{P(A_1) P(B|A_1)}{P(A_1) P(B|A_1) + P(A_2) P(B|A_2)} \text{ (2 marks)}$$

$$P(A_1 | B) = \frac{(0.40)(0.60)}{(0.40)(0.60) + (0.60)(0.20)}$$

$$P(A_1 | B) = \frac{0.24}{0.24 + 0.12} = \frac{0.24}{0.36} = 0.67$$