

Question I. [8.0 marks]

For the following survey situations, and in the context of the situation described state the specified terminologies in the accompanying tables.

- a) A school is interested in estimating the average score that may be obtained on a reading comprehensive exam for the 200 students in sixth year. In order to achieve this, a sample of size n is drawn.

Appropriate Survey design	
Target Population	
Frame	
Sampling unit	
Variable of interest	
Population size	

- b) A researcher wishes to do a nutritional survey in order to estimate the average amount of fiber consumed by individuals in a certain city. Suppose there is no list of families available for the city but there is a map of the city showing each block in detail.

Appropriate Survey design	
Target Population	
Sampling unit	
Variable of interest	

Question I. [cont.]

- c) A wholesale food distributor in a large city wants to know whether demand is great enough to justify adding a new product to his stock. To aid in making his decision, he plans to add this product to a sample of the stores he services in order to estimate average monthly sales. He only services four chains in the city. There are 24 stores in chain 1, 36 in chain 2, 30 in chain 3, and 30 in chain 4. The distributor has enough time and money to obtain data on monthly sales in n stores.

Appropriate Survey design	
Target Population	
Sampling unit	
Variable of interest	
What do chain 1, chain 2, chain 3 and Chain 4 indicate?	
Population size	

Question II. [4.0 marks]

Select from the list of words to complete the following statements, (Note: Not all list of words are used). Write only the number of the selected words.

List of words

- | | | |
|---------------------|--------------------------|--------------|
| 1. cluster | 7. data collections | 13. sampling |
| 2. simple random | 8. sampling plan | |
| 3. stratified | 9. estimation procedures | |
| 4. nonresponse | 10. parameters | |
| 5. nonobservational | 11. statistics | |
| 6. observational | 12. coverage | |

- a) In a sample survey, the major statistical components are referred to as the sample design and it involves two items. The ----- is the methodology used for selecting the sample from the population. The ----- are algorithms or formulas used for obtaining estimates of population values known as ----- from the sample data.
- b) Inability of the interviewed person to answer the question of interest is an example of ----- . It is considered as the most serious of all the ----- errors in surveys.
- c) If the population elements fall into groups, the variance of the estimator of the population parameter is usually reduced by using ----- sampling rather than ---- sampling.
- d) A sociologist wants to estimate the pre-capita income in a certain small city. No list of resident adults is available. ----- sampling seems to be the logical choice for the survey design.

Question III. [8.0 marks]

A simple random sample of 5 factories was taken from a group of 100 factories. The total number of accidents in each sampled factory during the current year was recorded and presented in the accompanying table.

<i>Factory #</i>	17	64	33	72	25
current year	3	5	3	3	1

- a) A researcher is interested in estimating the average number of accidents in all factories during the current year. Determine the number of factories to be sampled in the study in order to estimate the population mean with a bound on the error of the estimation equal to 0.50. Many similar studies have been run in the past. Using data from these studies, the investigator found that the population variance was approximately equal 4 accidents². Determine the required sample size.

[1.50 marks]

- b) Estimate the total number of accidents in all factories and place a bound on the error of the estimation.

[2.50 marks]

Question III. [cont.]

Suppose the total number of accidents in each sampled factory during the previous year was provided and presented in the accompanying table. The total number of accidents during the previous year over all factories is known to be 420.

<i>Factory #</i>	17	64	33	72	25
previous year	4	7	3	4	2

- c) Use the ratio estimation to estimate the total number of accidents in all factories and place a bound on the error of the estimation.

[3.0 marks]

- d) Which estimator is more appropriate to use as an estimate of the total number of accidents in the 100 factories? Explain your answer.

[1.0 mark]

Question IV. [12.0 marks]

Wage earners in a large firm are classified into management and clerical classes, the first having 300 and the second having 500 employees. To assess attitude on sick-leave policy, independent random samples of 100 employees each were selected, one from each of the classes. Each selected employee was asked "if he/she likes the policy" and his/her response was recorded. The obtained data are given below.

Number of employees who like the policy	
Management	70
Clerical	68

- a) Estimate the proportion of managers who like the policy and place a bound on the error of the estimation.

[1.5 marks]

Question IV. [cont.]

- b) Estimate the proportion of wage earners who like the policy and place a bound on the error of the estimation.

[3.0 marks]

Question IV. [cont.]

- c) Estimate the total number of wage earners who like the policy and place a bound on the error of the estimation.

[2.5 marks]

Question IV. [cont.]

- d) The data in the previous table are out of date. A new study will be conducted in the same firm for the purpose of estimating the proportion of wage earners who like the policy with a bound on the error of estimation equal to ~~2~~. The researcher decided to collect the data by sending questionnaires to employees. The cost of obtaining an observation is \$8 for the management class and \$4 for the clerical class. A prior survey suggests that the proportion of employees who like the policy in the two classes are approximately 0.75 and 0.55, respectively. Find the sample size n and the classes sample sizes n_1 and n_2 , that will give the desired bound at minimum cost.

0/0
3/2/17
Base 10

[5.0 marks]

Question V. [8.0 marks]

A researcher wishes to estimate the average number of college students per household in a given governorate, travel costs from household are substantial. Therefore, the 5000 households in the community are listed in 500 geographical clusters of 10 households each, and a simple random sample of 4 clusters is selected. The results are given in the following table:

Cluster	Number of College Students									
1	2	4	0	2	3	2	5	2	3	2
2	1	1	4	2	1	4	0	3	2	2
3	2	1	2	3	4	2	0	2	1	3
4	4	0	3	1	1	0	1	2	1	2

- a) Estimate the average number of college students per household for the governorate and place a bound on the error of the estimation.

[3.0 marks]

Question IV. [cont.]

- b) Assuming that the above data represent a simple random sample from the 5000 households, estimate the average number of College students per household for the governorate and place a bound on the error of the estimation.

[3.0 marks]

Question IV. [cont.]

- c) Find the relative efficiency of the estimator obtained in part (a) to the estimator obtained in part (b). Comment on your result.

[2.0 marks]
