Name: _____

Student ID: _____

- Please avoid all unethical behaviors (e.g., looking at others' solutions, using unethical means and asking others). I will follow very strict policy concerning this. First of all the exam will be cancelled, and any student found to be cheating will be reported to the dean of concerned faculty by the relevant faculty member for disciplinary action.

- 🗄 means there is a datafile with the question.

- 🅇 means only for this questions you can keep your complete answer in the Excel file, no need to write the answer in the exam paper. But except these questions, all the other ones must be answered in the answer script.

- A good practice is do not work with the original data file, keep it intact, copy the file and work with the copied file.

- Please save your Excel files with this naming convention -

  `id-questionnumber`

and after the exam, go to the google classroom and upload the file in the specified place.

- The exam is worth 50 points, and total duration of the exam is 90 minutes.

Question:	1	2	3	4	5	6	Total
Points:	8	10	7	13	8	4	50
Score:							

1. (8 points) Suppose prior to an election Institute for Public Policy Dialogue conducted a survey with **sample of** 512 **Bangladeshi adults**. In the survey they asked different questions.

    (a) (1 point) What do you think is the population for this study?

    (b) (1 point) Suppose the **total sum of the income** of the respondents is $7,680,000$ taka. What is the sample average of the income?

    (c) (1 point) Suppose the respondents were asked whether they **filed income tax** in the last financial year, and 294 respondents said "yes", what proportion of the people filed income tax in the last financial year?

    (d) (5 points) Suppose later you learned that the sample was **restricted only to the family members of the Institute for Public Policy Dialogue employees**. With this sample can we make conclusions for the entire Bangladeshi adults, why or why not?

1   Descriptive Stats Part

2. (10 points) Suppose we collected a dataset of Bachelors and Masters students from a University. The dataset has information from which departments the students have finished their Bachelors and Masters. The departments are

    - Electrical and Electronic Engineering (EEE),
    - Computer Science and Engineering (CSE),
    - Mathematics (Math),
    - Bachelors in Business Administration (BBA),
    - Economics (ECO),
    - Sociology (SOC) and
    - Pharmacy (Pharm).

    The dataset is available in the file 🗄 `Graduates.xlsx`. Using this dataset answer following questions,

    (a) (2 points) 🅧 - Provide a frequency, relative frequency and percent frequency distribution of different departments for the Bachelor students .

    (b) (2 points) 🅧 - Provide a frequency, relative frequency and percent frequency distribution of different departments for the Masters students .

    (c) (2 points) 🅧 - Construct a bar chart using frequency of different departments for the Bachelor students .

    (d) (2 points) According to the data, from which department majority Bachelors students have finished their Bachelors?

    (e) (2 points) Which department has the largest increase in percentage from Bachelors to Masters?

3. (7 points) For this question assume we have a data set of 50 BPL players. The data set contains runs from their best innings in the last t-20 season. The dataset is available in the file 🗄 `BPL.xlsx`. Using this dataset answer following questions

    (a) (2 points) ☒ Develop a frequency distribution using classes starting at 40 and ending at 60 in increments of 2 (this means classes will start from 40, and will end at 60, and each class size will be size of 2).

    (b) (2 points) ☒ Develop a histogram of the runs using the frequency distribution.

    (c) (2 points) Do the data appear to be skewed or symmetric?

    (d) (1 point) What percentage of the players made at least 50 runs?

2 Probability Theory Part

4. (13 points) Suppose that you are a real estate agent who sells apartments. Now you have a new task of contacting new customers. When you call any customer there are two possible options, either the customer "receives (R)" or "ignores (I)" your call. Now answer following questions,

    (a) (2 points) Is contacting a customer a random experiment? What is random here? What are the possible outcomes of this experiment and what is the sample space?

    (b) (2 points) If you apply the classical definition, what is the probability of the event that the customer will receive your call?

    (c) (2 points) If you know that the probability of the event that the customer will ignore your call is higher than the probability of the event that the customer will receive your call, then is it possible to apply classical definition in this case? Why or why not?

    (d) (2 points) Now suppose again you want to calculate the probability of the event that the customer will receive your call, how can we calculate this probability using frequency definition?

    (e) (3 points) Finally suppose now you will contact not one but three customers. This is a different random experiment now. What is the sample space now?

    (f) (2 points) Applying classical definition, calculate what is the probability of the event that all three customers will receive your call?

5. (8 points) In these days with random Hartals/Oborodhs it's becoming increasingly difficult for students to attend universities. Suppose that **if** there is an Oborodh on Thursday (this means <u>conditional on Oborodh on Thursday</u>) then there is a 0.7 probability that a randomly selected student will NOT go to the university. But if there is NO Oborodh (this means <u>conditional on NO Oborodh</u> on Thursday) then there is a 0.8 probability that a randomly selected student will go to the university. Suppose that there is a 0.8 probability that there will be an Oborodh on Thursday, now answer the following questions,

Hint: To help you, the structure of the joint probability table is given below. Here "O" means Oborodh and "NO" means NO Oborodh, "G" means go to Uni, and "NG" means NOT go to Uni.

	O	NO	
G	?	?	?
NG	?	?	?
	.8	.2	1

    (a) (2 points) If there is an Oborodh on Thursday (this means <u>conditional on Oborodh</u> on Thursday) what is the probability that a randomly selected student will go to the university?

(b) (2 points) What is the probability that a randomly selected student will go to the university?

(c) (2 points) What is the probability that a randomly selected student won't go to the university?

(d) (2 points) If a randomly selected student goes to the university on Thursday, what is the probability that there was an Oborodh on Thursday?

Finally we can calculate the all the joint probabilities

- $\mathbb{P}(G \cap O) = \mathbb{P}(G|O)\mathbb{P}(O) = 0.3 * 0.8 = 0.24$
- $\mathbb{P}(G \cap NO) = \mathbb{P}(G|NO)\mathbb{P}(NO) = 0.8 * 0.2 = 0.16$
- $\mathbb{P}(NG \cap O) = \mathbb{P}(NG|O)\mathbb{P}(O) = 0.7 * 0.8 = 0.56$
- $\mathbb{P}(NG \cap NO) = \mathbb{P}(NG|NO)\mathbb{P}(NO) = 0.2 * 0.2 = 0.04$

fill the joint probability table (although it was not asked, but you can get some bonus points..),

	O	NO	
G	.24	.16	.4
NG	.56	.04	.6
	.8	.2	1

6. (4 points) During an epidemic in a town, 40% of its inhabitants became sick. Given that a person is sick there is .10 probability that she will need to be admitted to an emergency ward. What is the probability that a randomly chosen person from this town will be admitted to an emergency ward?