

1 Instructions

- Only one submission per group, deadline = 2355 hrs, 7 May 2023
- Mention the names of all the group members along with the roll numbers in your submission
- One group can consist of 3-5 members
- Submit the report as a PDF file (maximum 8 pages) on MS Teams. You should also upload the supporting files (Python notebook/Excel file, raw data etc., so that the results presented in the report can be verified). If there is a dataset that is too big to be uploaded, you should provide the link to the dataset.
- Name the file as G(Groupnumber).pdf. Eg, Group 11's submission should be named as G11.pdf
 - I will send you the group numbers once I have the list of all the groups
- Fill the details of the group members in this link by 21st April 2023 - <https://forms.office.com/r/8nmCAKcfRi>
- No late submissions will be accepted
- Assignment should reflect your own group's work. Any case of copying will lead to strict disciplinary action
- Only one member of the group should submit the assignment

2 Assignment

1. Take a topic of your choice, and identify a problem that can be solved, or any interesting insight that can be obtained, or any common notion about something that you would like to prove/disprove. Collect data either by creating a survey or take data from sources on internet depending on the problem you have formulated.
2. Our institute has subscription of statista.com . You may be able to get some good dataset to work with on this link.
3. The analysis should use one of the methods learnt in the course. Eg. - Logistic regression, multiple regression, SVM, K-means clustering, GMM, hypothesis testing, constructing confidence interval, ANN, text analytics, etc.
4. The main report should clearly contain the motivation of the problem, a snapshot of data, methodology, analysis and results along with interesting findings/insights/observations. The length of the report should not be more than 8 pages.
5. If any dataset is taken from internet sources, the source of data must be clearly mentioned
6. Along with the report, you are also required to submit the entire raw data file, the results of the survey(if survey is taken), and other related files.
7. The report and analysis should be original. Any report copied directly from internet or elsewhere shall attract a score of zero.
8. Some of the groups at random may be called to present the work via online meeting. If the group fails to prove that the work is original, it will lead to a score of zero.