DATA-51000
INTRODUCTION TO DATA MINING AND ANALYTICS

GENERAL TIPS FOR WRITING DATA SCIENCE REPORTS:

- Write as if you are the expert data scientist and the instructor is your client for whom you need to analyze the data.
- Remember that the purpose of data science is to find new knowledge in data. The whole report needs to be written around this purpose. The conclusions should be about new insights that come from the analysis of the data and how they could be applied.
- When choosing data to work on, think about the problem the analysis will solve in this data. Also, make sure you focus on finding up to date, real data sets. For example, choose a dataset on current crime data from city portals or recently gathered data from social networks. Using old, well-used datasets that are now primarily used for teaching purposes is not interesting. Find data about something that interests you.d
- Make the title specific. Instead of using "Clustering on Data". Write something that relates to the data and the problem around it: "Identifying Groups of Customers for Good Market Segmentation".
- The introduction section should do several things:
    1. Begin by stating the problem, which in the case of data science will be based on the data.
    2. Motivate why your analysis work was useful for this data.
    3. Provide a short overview of what was done in the process and the general outcomes.
    4. Outline the rest of the paper. For example: "In section II, I provide an overview of the data. Then in section III, the analysis methodology is presented. Section IV, describes the results and discusses the analysis. Lastly, section V provides conclusions of the analysis."
- DO NOT PASTE SCREENSHOTS! (that is, unless you are actually writing about what is going on in the computer's screen, e.g. talking about graphical user interfaces – but that's an exception). Only show what is needed to help the user understand your methods or results. Make sure everything is clearly legible (sufficiently large fonts, easily distinguished features on the graphs, etc.).
- Figures and tables should be used to help the reader understand the writing and you should refer to them in the text. They should be numbered and labeled per IEEE specifications.
- Make sure to provide references to things you mention in the text: data sources, software, algorithms, theorems, facts about the problem or data, etc. You need to attribute the source, otherwise, it's plagiarism.
- Make sure to adhere to IEEE formatting guidelines – use the template for IEEE Transactions articles.
- Proofread your paper to make sure you avoid spelling and grammar mistakes and that the paper flows well. Get help on writing if necessary.
- Look up papers in IEEE Transactions journals for examples on how papers should be written. Look for journals with high impact scores that are also relevant to the field.
- Make sure to justify your methodology: why did you pick these particular algorithms? How did you go about finding the optimal parameters for the algorithms? Why did you preprocess the data in a particular way (e.g. normalized to mean of zero)?