Middle East Technical University
Department of Statistics
Fall 2022-2023

STAT 112– INTRODUCTION TO DATA PROCESSING ANA VISUALIZATIONS
TABLEAU PROJECT

Here is the outline for your term project:

- Open Tableau Public or Desktop Student version.
- Load the project datasets and union/join the two datasets to work on one dataset.
- Examine the variables and try to understand them. Which ones are categorical, numerical, date, or location items? Are they correctly classified in Tableau? If not correct them.
- Examine the existence of duplicates or missing values. Remove the duplicates and logically treat the missing ones.
- Look at the descriptive statistics and interpret them. Have you seen any strange data points? Out-of-range observations?
- Do necessary cleaning if it is needed.
- Ask at least 5 research questions based on a story in your mind related to this dataset.
- For each research question, use summary statistics or any appropriate visualization to answer your research question. Be creative in the selection of visualization methods. Do not use univariate plots too much. Most of the plots must be multivariate.
- Use proper techniques to improve your visualizations.
- Interpret the results.
- Upload your report to ODTUCLASS. (Please be sure that the report is written in your wording. Turnitin will check for similarities. If the similarity is more than % 40, you will get 0 from this project). Turnitin checks the AI usage of your paper too.
- Based on your story or the message you want to give your audience, choose 4-5 visualizations or descriptive/tables create your dashboard, and publish it under Tableau Public. Link this address with the GitHub page of the class. Give the link to your dashboard at the end of your report.

REPORT:

- Start with the title of your project and your name.
- Write a FORMAL report, including all the sections that should be in a report (such as Introduction, etc.). Include only the most important findings.
- Your report should be around 5 to 7 pages in the IEEE conference template format, in Times New Roman font 12, including graphs & tables.
- Start with an introduction by introducing the data.
- Data Preprocessing: Explain the data cleaning tools that you used.
- Exploratory Data Analysis: Give the descriptive statistics and interpretation of crucial variables. Write your research questions and the output that you applied to answer the question. Interpret the results.
- Conclusion and discussion: Give an overall conclusion that you observe after your EDA. Explain what kind of important findings you get. Are they known facts or have you seen any interesting findings that you were surprised when you saw the results?
- References (optional): If you used any.
- Link to your dashboard.