Foundations of Data Science 2023/2024

Activity 4: data visualization and analytics

This activity is dedicated to data visualization and analytics. Using the research questions and data of activity 2, and Python's matplotlib, do the following and

1. Create one visualization of each type according to Abela 2006 (see class slides):
    a. Comparison
    b. Distribution
    c. Composition
    d. Relationship
2. Create two visualizations using the same data:
    a. The first one should follow the best principles outlined in class
    b. The second one should break several of the best principles
    c. Critically compare the two figures.
3. Write a report detailing each step
    a. Present each visualization with a proper caption
    b. Discuss choices and decisions
    c. Include the original data in table format as annex.
    d. Describe the open source tools and libraries employed in your project, describing briefly how they were used.
    e. Include an estimation of hours each student contributed to the project as an annex.
    f. Use the ACM template (2 page limit excl. references and annexes): https://www.acm.org/publications/proceedings-template

References:

https://matplotlib.org/