Data Science for Smart Cities Essay (Group Work)

Objective: The purpose is to evaluate groups of two students (or more, but with additional conditions) on their ability to analyze, discuss, and propose solutions to issues emerging in urban environments by employing Data Science and Sociology. The essay must be composed in English and organized into two distinct parts. The first should describe the chosen problem, its societal significance, and relevant indicators, referencing appropriate scientific sources. The second should present the data analytics component, including dataset selection, data wrangling, spatio-temporal correlation analysis (if pertinent to the problem), visualization, predictive or classification modeling, and policy recommendations aimed at decision makers.

Instructions:

Part 1: Problem Description, literature analysis and Indicators

- Students must select an urban problem—examples include loneliness in urban settings, drinking water management, crime and safety, homelessness, the societal impact of infrastructure development, traffic management, energy consumption, waste management, housing market and Airbnb regulations... A comprehensive explanation of the problem is required, emphasizing its societal significance and impact on the city's inhabitants.
- A thorough review of the scientific literature should be provided to contextualize the issue.
- One or more indicators must be identified or designed to measure the problem and assess the potential impact of any proposed solution. The rationale behind the selection and relevance of these indicators must be discussed. Indicators can include, for example, traffic volume over time, air quality index, energy consumption per capita, waste generation per household, social capital, income...
- Ethical and social implications of both the problem and the selected indicators must be considered and analyzed.

Part 2: Data Analytics and Policy Suggestions

- Relevant datasets must be identified or created (including any necessary scraping activities) to support the analysis of the selected problem.
- Implement the analysis to identify any correlations or patterns in the data according with the objectives of your research. Visualize the data using appropriate tools and techniques.

- Where applicable, predictive models, classification/clustering techniques, or optimization models should be developed to further understand and address the problem.
- The results of the analysis should be discussed in detail. The results should be compared with the literature, whether they confirm the theses in the literature or identify critical points.
- Based on the findings, policy recommendations aimed at addressing the urban problem must be formulated. These recommendations should be critically discussed with consideration given to their ethical and social implications.
- Implemented code must be shared, preferably through an online repository, for potential inspection in the phases of project evaluation.

Essay Guidelines:

- The essay should be written in English and extend to at least ten pages (excluding the cover page and references). An Overleaf compatible LaTeX template is provided; do not alter it.
- Groups should consist of two students. We accept groups including up to 4 students. However, since the class evaluation basically consists in the project, the essay describing it, and the presentation / discussion, works produced by more than 2 students should provide a sensibly wider or deeper discussion of the considered topic. Members of the group are asked to describe their individual contribution to the overall work, adopting a CRediT (Contributor Roles Taxonomy) author statement.
- The cover page must include the names and student IDs of both participants, along with the essay title.
- The final submission must be in PDF format.
- Appropriate citations and references to published research must be provided throughout the essay.
- The language should be scientific (precise, clear, and concise), maintaining an objective tone throughout.
- Submission deadlines will be announced via the eLearning site. Late submissions will be not considered for the evaluation.
- The essay must be presented and discussed in person after the submission.
- Presentations can be held in English or Italian. One presentation date for each exam session is scheduled (check “segreterieonline” for actual dates).
- Questions on the course topics may be asked during/after the presentation to verify the achievement of the course objectives.

Grading Criteria:

- Clarity and coherence in describing the problem and the choice of indicators.
- Quality and relevance of the datasets selected and the data wrangling procedures
- performed.
- Since the course itself is structured around these two complementary perspectives,
- the essay must devote equal depth and balance to both parts:
    - Problem description, societal analysis, indicators, and policy recommendations
    - Data analytics
- Relevance and coherence of the reported scientific literature. A thorough and critical review of the scientific literature must be provided to contextualize the issue. The literature should serve as the foundation for both the problem framing and the interpretation of subsequent analytical results
- Accuracy and validity of the analytical and visualization methods used.
- Correctness of any predictive, classification, or optimization models developed.
- Effectiveness and practicality of the proposed policy suggestions.
- Depth of discussion regarding the ethical and social implications of both the problem and the proposed solutions.
- Overall quality of the essay, including its structure, language, and adherence to academic standards, as well as the effectiveness of the in-person presentation.

Note: Plagiarism is strictly prohibited and will result in disciplinary action. Please ensure that all sources, including pictures, are properly cited and referenced.

Generative AI can be used for grammar and text refinement, but not to generate content of the essay. GenAI can be used as a support in the implementation of code supporting the development of part 2 of the project, but the students must be able to describe it, comment it, respond to questions about it. It goes without saying that the usage of GenAI should be explicitly mentioned and described in the essay. Failing to comply to the above rules about the usage of GenAI will lead to penalizations (loss of points in the final grades), but it might also lead to the rejection of the essay and disciplinary action, based on the nature of the violations.

If you have any questions or concerns about the assignment, please contact the course instructors.