Dr. Justin B. Keeler
BSAN 710

Term Project – Python Programming

Activity Title: Open Data Exploration and Assignment Creation (No ChatGPT)

Objective: Students will explore open data sources, perform data analysis using Python and Jupyter Notebook, and design an assignment for future students based on their findings. Students will also write a 5-8 page report (double spaced) detailing their process and findings in Microsoft Word.

Instructions:

1. Find a suitable open data source that you find interesting and relevant. You may use Kaggle, the World Bank, or any other open data repositories.

2. Download the dataset and import it into a Jupyter Notebook.

3. Perform an exploratory data analysis using Python libraries such as pandas and NumPy. Make sure you cover the following tasks:
    - Data type identification
    - Basic function creation
    - Repetition statements (loops)
    - Lists, strings, and dictionaries manipulation
    - Data preparation (handling missing values, encoding categorical variables, etc.)
    - Sorting, filtering, and splicing data using pandas DataFrames and Series

4. Based on your data analysis, identify an interesting problem or question that future students can solve or answer.

5. Design an assignment for future students that includes:
    - A brief introduction to the dataset and its context
    - Clear objectives and requirements
    - Step-by-step instructions for completing the assignment
    - Sample code snippets and explanations (where applicable)
    - Evaluation criteria and rubric

6. Write a 5-8 page report in Microsoft Word detailing your data source, exploratory data analysis, and assignment design. The report should include the following sections:
    - Introduction
    - Data Source Description
    - Exploratory Data Analysis
    - Assignment Design
    - Conclusion

7. You may NOT use ChatGPT for guidance or assistance during this activity.

Term Project – Python Programming

Requirements:

- Students must choose an open data source that is relevant and engaging.
- Students must use Python and Jupyter Notebook to perform exploratory data analysis and demonstrate the use of key concepts like data types, basic functions, loops, lists, strings, dictionaries, data preparation, and pandas DataFrames and Series.
- Students must design an assignment for future students based on their findings, which includes a clear problem or question, objectives, requirements, and step-by-step instructions.
- Students must write a 5-8 page report in Microsoft Word detailing their data source, exploratory data analysis, and assignment design.

Disclaimer: Usage of ChatGPT is Strictly Forbidden

Attention students: Please note that the use of ChatGPT or any similar AI language model is strictly forbidden and disallowed for Version 1 of the "Designing and Solving Real-World Optimization Problems using Linear and Integer Programming" assignment. Any team found to be in violation of this rule will receive a score of zero for the entire assignment, and all team members will be referred to the Office of Student Conduct and Community Standards (https://www.wichita.edu/about/student_conduct/) for further investigation and potential disciplinary action in accordance with the Academic Integrity Policy.

By choosing to work on this assignment, you and your team members are committing to complete the assignment without the use of ChatGPT or any similar AI assistance. It is the responsibility of each team member to ensure that no such tools are used during the completion of the assignment.

Failure to comply with this disclaimer may have serious consequences for your academic standing and future prospects within the institution. Please make sure to abide by the guidelines and maintain the highest standards of academic integrity.

Term Project – Python Programming

Assessment Criteria (Rubric):

1. Data Source Selection (20 points)
    - Level 1: Inappropriate or irrelevant data source chosen (0-5 points)
    - Level 2: Data source chosen is somewhat relevant but lacks complexity or depth (6-10 points)
    - Level 3: Data source chosen is relevant and has moderate complexity (11-15 points)
    - Level 4: Data source chosen is highly relevant, complex, and engaging (16-20 points)

2. Exploratory Data Analysis (40 points)
    - Level 1: Limited or no understanding of data types, functions, loops, data structures, and data preparation techniques (0-10 points)
    - Level 2: Some understanding of the aforementioned concepts, but with significant errors or omissions (11-20 points)
    - Level 3: Good understanding of the concepts with minor errors or omissions (21-30 points)
    - Level 4: Excellent understanding of the concepts, well-executed analysis with clear and concise code (31-40 points)

3. Assignment Design (50 points)
    - Level 1: Poorly designed assignment with unclear objectives, lack of context, and confusing instructions (0-12 points)
    - Level 2: Assignment is partially well-designed, but objectives and instructions are not fully clear or complete (13-25 points)
    - Level 3: Assignment is well-designed with clear objectives, context, and instructions, but could be further refined (26-37 points)
    - Level 4: Assignment is excellently designed with a clear and engaging problem, comprehensive instructions, and a well-thought-out evaluation rubric (38-50 points)

4. Report Quality (40 points)
    - Level 1: Poor organization, lack of clear structure, and numerous grammatical errors or typos (0-17 points)
    - Level 2: Some organization, but with unclear structure and a few grammatical errors or typos (18-25 points)
    - Level 3: Good organization with a clear structure, and few grammatical errors or typos (26-33 points)
    - Level 4: Excellent organization with a clear and concise structure, and minimal to no grammatical errors or typos (34-40 points)