Programming for Data Science

1. **Data Collection (10 points):**
    a. Completeness and quality of data sources identified
    b. Appropriateness of data collection methods used
    c. Adequacy of data sampling techniques
    d. Adherence to ethical and legal considerations
2. **Data Pre-processing (40 points):**
    a. Completeness and correctness of data cleaning steps
    b. Appropriate handling of missing and inconsistent data
    c. Effectiveness of data normalization and standardization
    d. Suitability of data transformation and feature selection techniques
    e. Appropriateness of data integration and fusion techniques
3. **Exploratory Data Analysis (30 points):**
    a. Adequacy and correctness of data visualization techniques
    b. Suitability and effectiveness of descriptive statistics methods
    c. Clarity and accuracy of statistical inference and hypothesis testing
    d. Appropriateness of feature engineering and selection methods
    e. Insightfulness and novelty of exploratory data analysis results
4. **Modeling (20 points):**
    a. Appropriateness of machine learning algorithms and techniques used
    b. Adequacy of model training and evaluation methods
    c. Suitability of hyperparameter tuning and regularization techniques
    d. Accuracy and robustness of model predictions and performance metrics
    e. Novelty and significance of model results and insights