**Project Assignment Guideline: ANOVA Analysis**

**Objective:** This project aims to provide practical experience with conducting and interpreting ANOVA, a statistical method used to determine if there are any statistically significant differences between the means of three or more independent groups.

**Instructions:**

1. **Topic**:
   * Conduct an ANOVA analysis. The topic should involve comparing group means in a scenario that is either of academic interest or practical relevance.
2. **Data Collection**:
   * **Find Existing Data**: You may use data from online repositories, academic publications, or other credible sources.
   * **Collect Your Own Data**: Alternatively, you can conduct a survey or experiment to gather data specifically for this project.
   * Ensure that the data involves at least three groups or categories for comparison.
3. **Data Description and Objectives**:
   * Describe your dataset, including the source, nature, and any relevant background information.
   * Clearly state the aim of the data analysis. What are the hypotheses or questions you are addressing with ANOVA?
4. **Analysis Requirements**:
   * **Data Summary and Visualization**: Provide a statistical summary of the data. Include appropriate graphs (e.g., box plots, histograms) to illustrate the data distributions and relationships among groups.
   * **Model Diagnostics**: Conduct diagnostics to check ANOVA assumptions regarding normality and homogeneity of variances.
5. **Conclusion**:
   * Summarize the findings of the data analysis. Indicate whether the results support your initial hypotheses.
   * Discuss any potential implications or recommendations based on your findings.
6. **Presentation**:
   * Prepare a slide presentation summarizing your project.
   * The presentation is scheduled for May 16th. Each presentation is limited to 5-6 minutes.
   * Include the following slides:
     + **Title Slide**: Project title, team members, and specify the team leader.
     + **Introduction**: Background information and objectives.
     + **Methodology**: Data collection methods and analysis techniques.
     + **Results**: Summary statistics, visualizations, model diagnostics.
     + **Conclusion**: Main findings and implications.
     + **Reference**: The resources in your project.
7. **Submission**:
   * Upload your presentation slides to the class online platform before **May 16th (Deadline: May 15th, 12 PM)**.
8. **Presentation Order**:
   * Teams will present in a random order, determined on the day of presentation.

**Evaluation Criteria:**

* **Data Appropriateness**: How well the chosen data suits the ANOVA analysis.
* **Analysis Accuracy**: Correct application of ANOVA and diagnostics.
* **Clarity of Presentation**: Effectiveness in communicating findings.
* **Visual and Statistical Thoroughness**: Quality and relevance of statistical summaries and visualizations.
* **Insight and Depth of Conclusion**: Depth of analysis in supporting or refuting initial hypotheses.

**Data Links for Reference:**

1. **UCI Machine Learning Repository**:
   1. **Description**: A popular repository for machine learning datasets but also extensively used for statistical analysis in various fields.
   2. **Link** <https://archive.ics.uci.edu/>
2. **Kaggle Datasets**:
   1. **Description**: Kaggle offers a wide range of datasets uploaded by users around the world. It includes data from competitions as well as datasets uploaded for sharing.
   2. **Link**: <https://www.kaggle.com/datasets?fileType=csv>
3. **National Development Council's Open Data Portal (Taiwan)**
   1. **Description**: This portal is aimed at promoting transparency and civic engagement by making government data available to the public. The data can be used for research, analysis, and educational purposes.
   2. **Link**: <https://data.gov.tw/>