# EDA Report

**Project Overview**

The project involved conducting an exploratory data analysis (EDA) for "cetrisured," who organized a workshop in a college spanning 9 days. During the workshop, students took 5 quizzes, and feedback forms were collected on instructor knowledge, overall ratings, and individual help provided. The client's objective was to identify the top 6 students for an internship based on quiz scores and attendance and to analyze feedback ratings.

**Data Preparation**

1. **Merging Feedback Data**

* Data from two Excel files, "FDF1.xlsx" and "FDF2.xlsx," was merged into a single dataset.
* Duplicate entries were removed based on email addresses.
* Unnecessary columns were dropped.
* Column names were renamed for clarity.
* Missing values were replaced with 0.
* Categorical feedback ratings were mapped to numerical values (Excellent: 5, Very good: 4, Good: 3, Fair: 2, Bad: 1).
* An "Average" column was created to calculate the average rating across different columns.

1. **Attendance Data Analysis**

* The attendance data was read from the "Score\_Table.xlsx" file.
* Rows with missing values (NaN) were removed.
* The data was grouped by email addresses to count the number of timestamps for each student.
* A new DataFrame, "student\_with\_4\_att," was created to store students who attended all 5 quizzes.
* An inner join was performed to combine the attendance data with students who attended all 5 quizzes.
* The result was saved to a new Excel file, "result.xlsx."

1. **Selecting the Top 6 Students for Internship**

* The "result.xlsx" file was read into a DataFrame.
* The data was grouped by email addresses to calculate the total score for each student, sorted in descending order.
* The top 6 students with the highest total scores were selected for the internship.

**Conclusion**

The EDA process successfully prepared the data for further analysis, making it structured and ready for insights and visualization. The top 6 students were selected for the internship based on quiz scores and attendance. Additional analysis can be performed on feedback ratings to provide further insights and recommendations.

Please note that the report includes code snippets for reference, but actual data and results are not provided to maintain data privacy and confidentiality.

**BY: AmithRaj k**

**Sandesh N Christy**