BNCA Project : Student Progress Analysis

Project Overview

The "BNCA" project focuses on analyzing student progress using data extracted from a PDF file containing student marks. This document details the process from data extraction to the creation of an interactive Power BI dashboard.

Data Sources

- **Primary Source**: PDF file containing student marks-sheet.

Tools

- **Microsoft Power BI**: Used for data manipulation, visualization, and dashboard creation.
- **Python Script (within Power BI)**: Employed for data extraction from the PDF.

Data Cleaning/Preparation

1. **PDF Extraction**:
    - A Python script embedded within Power BI was used to automate the process of extracting relevant information from the PDF. This script involved techniques like pattern matching to identify and isolate student data points (e.g., names, marks, semester information) and also created dataframes for subjects.
2. **Data Transformation**:
    - **Handling Missing Values**: Identifying and addressing missing data points .
    - **Data Type Conversion**: Ensuring data types are consistent (e.g., converting text marks to numerical values).
    - **Creating Calculated Fields**: Deriving new metrics from existing data (e.g., calculating semester GPA based on individual class marks).

Exploratory Data Analysis (EDA)

- Initial exploration of the extracted data involved:
    - **Data Distribution**: Visualizing the distribution of marks and other metrics (e.g., histograms, boxplots) to identify patterns.
    - **Identifying Relationships**: Exploring potential relationships between variables like semester, class, and student performance.

Data Analysis

The core analysis focused on understanding student progress across semesters using techniques such as:

- **Trend Analysis**: Identifying trends in student performance over time (e.g., marks across semesters, year-wise).
- **Grouping and Aggregation**: Grouping students based on specific criteria (e.g., year of enrolment) and analyzing their performance as a group.

Results/Findings

The primary outcome of this project is a user-friendly Power BI dashboard with the following key features:

- **Semester-wise Performance**: Visualizations (e.g., bar charts, line graphs) displaying student performance metrics (e.g., average marks, pass rates) for each semester.
- **Drill-Through Functionality**: Users can click on specific data points within the visualizations to access detailed information for individual students. This includes:
    - Student name and identification number
    - Class-wise marks and semester GPA/CGPA
    - Performance trends over time

Recommendations

Based on the analysis, the following recommendations are proposed:

- **Student Support**: Identifying students who might require additional support based on their performance trends.
- **Performance Improvement**: Highlighting specific classes or semesters where overall performance is lower, prompting further investigation.
- **Curriculum Adjustments**: Suggesting curriculum adjustments based on identified student struggles.