_STUDENT SCORE ANALYSIS USING PYTHON (EDA)

● Introduction:

Academic performance is influenced by a range of factors, including demographic traits, parental background, and study habits. Recognizing these influences is essential for educators and policymakers to create effective interventions and support systems. This project examines a dataset of student scores to investigate how factors like gender, parental education, marital status, ethnic group, and weekly study hours impact performance in math, reading, and writing. Through statistical analysis and visualizations, this study seeks to reveal significant patterns and offer practical insights to enhance educational outcomes.

● Objectives:

1. Analyze demographic and parental factors affecting student scores.
2. Assess the impact of weekly study habits on performance.
3. Provide actionable insights for improving academic outcomes.

● Dataset Overview:

1. **Source**: *Expanded_data_with_more_features.csv*.
2. **Content**: Includes demographics, parental background, study habits, and academic scores.
3. **Key Features**: Gender, ethnic group, parental education, marital status, weekly study hours, and math, reading, and writing scores.

4. **Key Features**:

a. **Demographics**: Gender, ethnic group.
b. **Parental Background**: Education level, marital status.
c. **Study Habits**: Weekly study hours, participation in sports.
d. **Performance Metrics**: Math, reading, and writing scores.

● Methodology:

1. **Data Preprocessing**:
    - Load the dataset and remove irrelevant columns.
    - Handle missing values and standardize formats.
2. **Exploratory Data Analysis (EDA)**:
    - Generate descriptive statistics to understand the dataset.
    - Identify trends and potential correlations among variables.
3. **Visualizations**:
    - Use bar plots, pie charts, heatmaps, and box plots to interpret relationships.
4. **Insights Extraction**:
    - Group data by attributes like parental education, ethnic group, and study hours to derive key insights.

● Data Analysis:

1. **Gender Distribution**: Visualized the male-to-female ratio among students.
2. **Parental Factors**: Analyzed the impact of education and marital status on student scores.
3. **Study Hours**: Assessed how weekly study hours correlate with performance.
4. **Ethnic Group Representation**: Explored the distribution and its relationship to performance.
5. **Score Trends**: Examined distributions and patterns in math, reading, and writing scores.

● Visualizations:

1. **Bar Plot**: Gender distribution across the dataset.
2. **Heatmaps**:
    a. Weekly study hours vs. average scores.
    b. Parental education and marital status vs. scores.
3. **Pie Chart**: Representation of ethnic groups.
4. **Box Plots**: Spread of math, reading, and writing scores.

● Key Findings:

1. Students with higher parental education levels consistently perform better in all subjects.
2. Weekly study hours positively impact math, reading, and writing scores.
3. Gender and ethnic groups show disparities in representation and performance.
4. Parental marital status has a measurable effect on academic outcomes.

● Conclusion:

The analysis emphasizes the considerable impact of parental background, study habits, and demographic factors on academic performance. These results highlight the necessity for focused interventions to tackle disparities and promote student success.

● Recommendations:

1. Encourage parental involvement, particularly in educational activities.
2. Promote balanced and consistent study habits for students.
3. Address disparities in resources and opportunities across ethnic groups.
4. Provide additional support for students from less stable family environments.

SUBMITTED BY: Anushka Ajay Lande

COURSE NAME: BCA - Data Science

CLG NAME: Sri Balaji University, Pune (SBUP)

DATE: 16TH DEC, 2024