Analyzing student Academic performance using Python. Examining factors influencing student performance in Uganda Primary schools.

**Problem Statement**

Student performance is of deep concern in primary schools of Uganda with various challenges which affect their performance. The government of Uganda through the Ministry of Education and Sports implemented initiatives like the Universal Primary Education with the aim of ensuring the completion of the primary cycle of education and reducing poverty by equipping every individual with basic skills (National Planning Authority, 2018) Despite the formation of such initiatives, primary schools in Uganda still face factors that affect the performance of their students. The factors include lack of scholastic materials like textbooks, exercise books, pens, pencils especially in rural areas. Others include lack of skilled teachers, distance from schools, and attendance rate of students, age, overcrowding, and inadequate funding from stakeholders like grants, inadequate study time and high teacher to child ratios, low enrollment and dropout rates especially among districts in rural areas.

This project aims to analyze existing primary data using python to identify patterns and trends in performance of students to provide recommendations and interventions which are evidence based.

**Relevance to a Real-World Context**

Analyzing student performance is relevant to National development goals hence enabling the project to generate evidence.

It enables effective evidence based decision making. The use of python provides evidence based data that shows the impact of specific inputs like funding levels and teacher student ratios on student performance (National Planning Authority, 2020). The allows the stakeholders and policy makers to prioritize budget allocations and allocate funds and resources to schools with low performance rates hence making informed and practical decisions. This will enable schools to also use python as a system to turn raw data into actionable knowledge that supports informed and accountable decisions in order to meet their targets (Han & Kamber, 2022).

**Improves effective planning.** This is through identifying gaps that lead to a decrease in student performance and implementing strategies that reduce dropout rates and increase enrollment rates (Manideep & Kumar, 2022). Use of python will provide data that will enable schools set up programs and interventions that enables students to stay in school for example provision of scholastic materials, counselling and mentoring programs. The government can use the data to plan for the education sector by allocating more funds to them which will used to recruit and train more teachers, introducing the use of Information, communication and Technology in schools which will ease the learning and research purposes. This also increases the number of students competing at the job market with equitable skills.

Enhancement of Accountability. This demonstrates accountability to funders, beneficiaries, stakeholders and policy makers on how resources are utilized to increase student performance (Hallack, 2000). Since the data is evidence based it shows currently how resources are used by showing the initial situation on the school used to function and what changes the resources have brought onto the school for example the funds provided for infrastructure development used and increasing on the enrollment rate of students, provision of scholastic materials like textbooks, computers. This will encourage stakeholder involvement which increases more funding through grants. This also enable policy makers to form policies that are transparent and evidence based.

Use of python will enable schools set targeted interventions and programs to increase on the performance of students like recruitment and training of teachers, capacity building through organizing workshops and seminars to set up policies and initiatives to increase student performance.

**Scope of the Project**

The project will focus on analyzing primary school student performance by examining the factors that affect the performance. The analysis will use secondary data collection method through using data from the Ministry of education and sports Annual school primary school census data and find the relationship between factors.  
**Proposed Python Approach and Tools**

This proposed approach will involve the use of python libraries and tools which will perform various activities and they include the following;

Pandas will used for importing, cleaning and organizing the dataset by removing errors and missing values. It will transform variables to be suitable for analysis.

Matplotlib will be used for generating descriptive statistics and visualizations to identify patterns and trends by checking for correlations and detect outliers.

Python will also build statistical models to quantify the impact of each factor and predict student performance. This involves the use of scikit-learn for regressions models and Numpy for providing efficient array operations and numerical computation.

Python will extract feature importance from the models to identify the strongest predicators, validating and structuring the findings and making actionable recommendations

**Expected Outcomes and Deliverables**

The expected outputs of this project include:

Factors that show the statistical significance and magnitude of impact of impact on student performance in Ugandan primary schools.

A report containing visualizations that illustrate the patterns and trends identified during the exploratory data analysis like performance differences across groups for gender or study time.

A report containing evidence based recommendations for stakeholders on where to utilize resources.

**REFERENCES**

McKinney, W. (2018). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython (2nd ed.).

Ministry of Education and Sports [Uganda]. (2021). Education Management Information System (EMIS) Report 2021. Government of Uganda.

UNESCO. (2023). Leveraging data for education policy and practice in Sub-Saharan Africa. United Nations Educational, Scientific and Cultural Organization.

Han, J., Kamber, M., & Pei, J. (2012). Data mining: Concepts and. Techniques, Waltham: Morgan Kaufmann Publishers.

National Planning Authority. (2018). Comprehensive evaluation of the Universal Primary Education (UPE) policy. Thematic Report 3: Primary teacher training for competent teachers to deliver UPE. Government of Uganda.

Hallack, J. (2000). "Accountability in Education: Crisis and Reform." UNESCO/International Institute for Educational Planning. (This reference highlights the global shift towards data-driven accountability as a core element of education reform and good governance.)

Manideep, R., & Kumar, P. V. A. (2022). Predicting student’s performance based on machine learning. Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology.