




Dennis Mwendwa

 github.com/deninjo  0700362004  linkedin.com/in/dennis-mwendwa  imwendwadennis@gmail.com

PROFESSIONAL SUMMARY

Innovative and zealous programmer seeking to leverage extensive background in data analysis and trends. Proficient in Python programming, I have successfully developed and maintained scalable and efficient software applications. Demonstrated strong problem-solving skills by implementing optimized algorithms and data structures in Python, significantly improving system performance. All in all, I am a one that is conversant with new and emerging technologies that tend to benefit me in the various endeavours that I undertake within the tech space with a current peak interest in back-end development.

EDUCATION

Primary School Education

January 2006 - June 2016

Kenya Certificate of Primary Education(K.C.P.E)

GPA: 404/500

Whitestar Academy, Nairobi, Kenya

Secondary School Education

January 2017 - April 2021

Kenya Certificate of Secondary Education(K.C.S.E)

GPA: A-

Nairobi School, Nairobi, Kenya

Undergraduate Education

September 2021 - Current

B.Sc. Computer Science

Jomo Kenyatta University of Agriculture and Technology

SKILLS

Languages: C, Python, SQL, PHP, HTML/CSS, \LaTeX

Tools: Git/GitHub, Unix Shell, VS Code, MySQL Workbench, PyCharm, IBM Cognos, Excel, Power BI, Ms Office, Microsoft Dynamics 365

Operating Systems: Linux, Windows

PROJECTS

Analyzed the Relationship Between HIV Prevalence and Multidimensional Poverty | *Python, Pandas, GeoPandas, Seaborn, Matplotlib, Jupyter*

- Conducted a cross-national analysis of the relationship between HIV prevalence and multidimensional poverty using datasets from the WHO Global Health Observatory and World Bank
- Cleaned and preprocessed multi-source datasets then integrated HIV and poverty datasets by country and year, forming a comprehensive panel dataset to support temporal and spatial analysis.
- Built and interpreted correlation matrices and a linear model to explore statistical relationships between HIV prevalence and poverty sub-indicators.
- Uncovered monetary poverty and lack of access to drinking water as factors with the strongest positive correlations with HIV prevalence. Higher education enrollment was significantly associated with lower HIV prevalence.
- Visualized key trends using time-series plots and scatter plots to support pattern discovery and hypothesis testing using interactive visualization libraries (e.g., Plotly) to enhance data exploration and stakeholder presentation quality.
- Isolated countries accounting for 75% of the global HIV burden with Sub-Saharan Africa dominating where South Africa alone accounts for a significant share followed by Nigeria, Mozambique, India, and Kenya
- Analyzed under-five and neonatal mortality in East African Community (EAC) countries using UN IGME data.
- Filtered the mortality dataset for 8 EAC countries then used GADM shapefiles to visualize latest estimates of under-five and neonatal mortality for these countries via choropleth maps, reinforcing geographic storytelling

- Demonstrated ability to work with real-world, imperfect data and draw actionable insights from public health and socio-economic datasets.

Analyzed Kenya Data | *Jupyter Notebook ,HTML, Git,VS Code*

- Extracted Kenya's 2019 census data from KNBS
- Performed comparative and descriptive insights and visualized the Kenya 2019 census data
- Cleaned and transformed Kenya's weather data from 1991-2016, providing quality data
- Identified correlations and applied statistical methods to find the best fit model for average data across the years
- Analyzed Stock Prices from Historical Data (NSE from 2007 - 2012) to understand patterns, predict future movements, and make data-driven investment decisions.
- Predicted the closing price of stocks based on key predictors using a linear regressor with a MAE of 0.17 and RMSE of 0.60
- Analyzed Road accidents data to predict how risky a road is prone to traffic accidents. The Random Forest classifier model managed to achieve an accuracy of 99%
- Exported the Jupyter Notebook file to HTML to allow interactive visualizations to be seamlessly integrated with the code and analysis

Analyzed Personal Spotify and Apple Music Data | *Power BI ,Excel*

- Extracted and integrated personal Spotify and Apple Music streaming data, handling multi-source data consolidation across different platforms and time periods
- Implemented advanced data filtering and slicing using Power BI's interactive features including date range sliders (spanning 2021-2024), time-of-day analysis (24-hour breakdown), and cross-platform comparison filters
- Developed comprehensive KPI metrics including total tracks played (22K+ and 4.9K+ respectively), total listening minutes (83.77K+ and 14.8K+), unique artists discovered (1,218+ and 541+), and average session duration patterns
- Created temporal analysis frameworks to identify listening patterns by year, month, day, and hourly intervals, revealing peak activity periods and seasonal trends in music consumption
- Built multi-dimensional artist performance analytics tracking top artists by both total minutes listened and song count frequency, enabling identification of preferred content types
- Engineered behavioral insights dashboards analyzing listening completion rates through "End Reason Type" segmentation (natural completion vs. manual skips vs. track forwarding) to understand user engagement patterns
- Applied DAX queries to create calculated measures and columns, enabling insights such as most-streamed genres, artists, and time-based listening patterns.
- Created dynamic time-series visualizations with trend analysis showing listening evolution over multi-year periods, identifying peak listening months and correlation with external factors

LeafLens – Maize Leaf Disease Detection System | *Python, TensorFlow, CNN, MobileNetV2, Android Integration*

- Developed a deep learning-based maize disease prediction system using Convolutional Neural Networks (CNNs), trained on 4,000+ images sourced from Kaggle across four disease classes.
- Conducted comparative analysis of CNN architectures (DenseNet, Xception, and MobileNetV2); selected MobileNetV2 for its superior performance and efficiency
- Applied transfer learning with hyperparameter tuning to improve generalization and reduce training time.
- Preprocessed images via normalization (0–1 pixel scaling) and resizing to 224x224 pixels; augmented training data with random flipping, 90° rotation, and brightness adjustment
- Implemented model checkpointing, saving best-performing models based on validation accuracy; achieved 99.3% testing accuracy with a confusion matrix showing only 5 misclassifications across 2 classes
- Achieved > 97% precision, recall, and F1-score across all four classes, ensuring robust multi-class performance.
- Designed a binary classifier (maize vs. non-maize leaf) using 8,000 images (balanced across both classes); achieved 99% accuracy, serving as an effective preliminary stage for classification.
- Collaborated with Android developer colleague to convert models to tensorflowlite and integrate into a mobile app, enabling real-time disease detection via camera input.

Movie Ticket Booking System | *Python, MySQL, Linux Shell, Power BI, PHP*

- Built a robust MySQL database for a ticket booking system, incorporating essential tables for customers, movies, showtimes, theatres, seats, and bookings.
- Advanced logic was implemented using triggers to handle tasks like: Dynamic price calculations based on seat category, showtime, and customer type (e.g., applying a 25% discount for students).
- The backend system was developed in Python, leveraging libraries like mysql-connector to connect to the MySQL database. Python scripts acted as the core interface between users and the database.
- Python classes (Customer, Movie, Theatre, Seat, etc.) handled operations like: Customer data entry and validation, Displaying movies and their available showtimes, Viewing and managing seating charts, Adding, confirming, and canceling bookings.
- Distributed the Database via Master-Slave Replication:
 - * A master-slave replication model was implemented to improve database performance and scalability.
 - * The master database managed all write operations, such as adding bookings, updating customer information, and modifying seat statuses.
 - * The slave databases were optimized for read-heavy operations like retrieving seating charts or displaying available showtimes.
 - * Automated the backup process using cron jobs in combination with database-specific tools (e.g., mysqldump, mysqlbackup).
 - * This architecture enhanced system performance and reduced latency for users by reducing the load on the master node. It also provided redundancy to safeguard against failures.
- A data warehouse was designed and set up to aggregate and store data for analytics.
 - * ETL pipelines extracted transactional data from the MySQL database, transformed it into meaningful metrics (e.g., total revenue by showtime), and loaded it into the data warehouse.
 - * A star schema was used to optimize the warehouse for analytical queries, with fact tables (e.g., booking data) linked to dimension tables (e.g., customers, movies, showtimes).
 - * The warehouse enabled deeper insights into system operations, such as:
 - Revenue trends and peak booking times.
 - Customer behavior analysis (e.g., preferences for certain movies or showtimes).
 - Utilization rates of seat categories.
- Power BI Integration:
 - * Connected Power BI to MySQL: Installed the MySQL ODBC driver (9.1). Successfully connected Power BI to my MySQL database using the ODBC connector.
 - * Power BI was used for reporting and visualization, connected either directly to the MySQL database and the data warehouse:
 - Dashboards were created to display real-time insights like popular movies, peak booking periods, seat occupancy rates, and revenue breakdowns.
 - Dynamic, user-friendly reports empowered stakeholders to make data-driven decisions.
- Currently building the web application implementation

Chama App | *Python, C, XAML, .NET Framework, APIs (Google), Git, VS Code*

- Collaboration project for a sacco management system
- Developed python SMTP scripts that enable email automation increasing efficiency and time saved
- Learned how to connect to gmail servers and use databases in conjunction with Google APIs
- Generated reports - individual member and sacco reports

CERTIFICATIONS

Microsoft Certified: Azure Fundamentals: Microsoft

Introduction to Data Analytics: Coursera

Data Analysis with Python: Freecodecamp

SQL Intermediate: HackerRank

Excel for Data Analysis: IBM

Data Visualization and Dashboards with Excel and Cognos: IBM

Alteryx Foundational Micro-Credential: Alteryx

International Computer Drivers License - Level One (ICDL): ICDL Africa