Data Analyst/ Data Scientist

OBINNA OGBUAGU

(+234) 8118879872 | rnoogbuagu@gmail.com
LinkedIn: https://www.linkedin.com/in/noel-ogbuagu-61813414a/
Portfolio: https://noelogbuagu.github.io/

Professional Summary

Highly motivated and detail-oriented professional with a strong data science and analytics background. Skilled in leveraging data to drive business insights and support decision-making processes. Experienced in developing and implementing predictive models, conducting data visualization, and performing statistical analysis. Proficient in programming languages such as Python and R, with a solid understanding of machine learning algorithms and techniques. Excellent problem-solving and communication skills, with a demonstrated ability to translate complex data into actionable recommendations. Committed to delivering high-quality results and continuously expanding knowledge in data science.

Skills

- SQL (SQL Server, SQLite3, PostgreSQL, SQLAlchemy)
- Python (Pandas, NumPy, SciPy, MatPlotLib)
- Excel (VLookup, Conditional Formatting, Pivot Tables)
- Tableau, MongoDB
- R (tidyverse, dplyr, ggplot2)
- Web Scraping (BeautifulSoup, Selenium, Scrapy)

Project Experience

SMALL BUSINESS OWNNER SEGMENTATION

- Developed a **K-means clustering** model that groups small business owners in the United States into subgroups for partner proximity and customer convenience.
- Demographic characteristics like education, age and race were examined using **Pandas**, **Seaborn** and **MatplotLib**.
- Conducted **feature selection** for clustering using variance and dimensionality reduction using PCA.
- Designed, built, and deployed a dynamic web application using **Dash**, **Plotly express**, **Scikit learn** and **SciPy**.

BANKRUPTCY CLASSIFICATION

- Built a classification model using **Python** and leveraging financial data from 9500+ companies that predict companies at risk of bankruptcy—improving financial management, risk mitigation, and the overall health and resilience of the small business ecosystem.
- Addressed imbalanced data with resampling and compared the performance of multiple classification algorithms.
- Improved validation test accuracy by six percentage points after hyperparameter tuning to achieve 96% test accuracy, a precision score of 57% and a recall score of 59%.
- Created a Python module that wrangles and utilizes the model to predict bankruptcy on new data.

SALES MANAGEMENT DASHBOARD

- Created a sales management dashboard using **Google Sheets** to visualize an electronics store's product sales, offering real-time visibility into key sales metrics and enabling data-driven decision-making, goal tracking, and performance evaluation.
- Provided 3 growth-driving **recommendations** based on trends and insights from analysis.

TORONTO HOUSE PRICE PREDICTION MODEL

- Implemented a Regression model that predicts housing prices in Toronto using house features for an investment syndicate to inform decision-making and provide market insight. I extracted, processed, and analyzed 2000 listings using **Selenium**, **Regex** and Python.
- Model achieved a mean absolute error of $876,000. A $65,000 reduction in the baseline proves its efficacy.
- Developed and deployed an interactive dashboard that predicts house prices based on user input.

Work Experience

DATA ANALYST – Freelance **–** Lagos, NG   May 2023 - June 2023

- Developed a classification model that predicts whether a building will suffer severe damage in an earthquake to inform policies around disaster management and construction practices in earthquake-prone areas.
- Leveraged **SQLite3** and **SQLAlchemy** to create, connect, and query the database. Utilized Python for processing and Analysis of over 39,000 observations from queries.
- Achieved 70% accuracy and deployed an interactive dashboard using **IPyWidgets** that reports whether a building is at risk of severe damage based on user input.

DATA ANALYST – Freelance **–** Lagos, NG   March 2023 - April 2023

- Created an Autoregressive time series model that predicts particulate matter readings throughout the day in Victoria Island, Lagos, for an environmental research group—utilizing **MongoDB** to query over 1.25 million air quality sensor measurements.
- Evaluated the model using **walk-forward** validation and achieved a 66.5% mean absolute error reduction on the test set compared to the baseline.

FINANCE ANALYST – Quintessence Properties Limited **–** Lagos, NG   September 2021 - February 2023

- Analyzed procurement and sales data using Google Sheets, which led to a ₦15,000,000 reduction in expenditure.
- Implemented a new reporting using Google Sheets Pivot, which helps reduced processing time by 50%.
- Generated three weekly process reports for continuous financial and developmental analysis accounts.

DATA ANALYST – Freelance **–** Lagos, NG   July 2021 - August 2021

- Utilized **R** to analyze 1000 Airbnb listings and gain insights into local and diasporan consumer preference for mid-short-term housing over long-term housing investment for a property technology company: leveraged Gauteng, South Africa, as a proxy for Lagos State.
- Aggregated data for visualization using **tidyverse and dplyr.** Created an interactive dashboard using **Tableau.**
- Proposed adding short-term housing options on the marketplace and partnering with industry stakeholders to meet demand. The former led to an increase of 50 users over 3 months.

BUSINESS TRANSFORMATION, INTERN – Ernst & Young – Lagos NG   February 2021 - June 2021

- Collaborated with a three-person team to facilitate process improvement and logistics analysis for a national cash-in-transit firm. Developed maturity models using Excel and outlined efficiency optimization and business process re-engineering phases.
- Conducted market research for a commercial bank implementing a company-wide rebrand to drive growth. Documented current state and future state process maps for stakeholders. Implementation of the rebrand led to a 10% in user growth within six months.

Education

BACHELOR OF SCIENCE – Ashesi University – Eastern Region, Ghana   May 2020
Majors: Electrical and Electronics Engineering

GOOGLE DATA ANALYST PROFESSIONAL CERTIFICATE   October 2021 - Present
Key Coursework: Data Analysis with R Programming, Introduction to SQL

BRAINNEST BUSINESS ANALYSIS INDUSTRY TRAINING   October 2021 - Present
Key Coursework: PESTL Analysis, SWOT Analysis

DATA SCIENTIST WITH PYTHON – Datacamp   March 2023 - Present
Key Coursework: Statistics in Python, Supervised Learning, Unsupervised Learning