# JACQUES DIAMBRA ODI

DATA SCIENTIST STONECREST, GA, UNITED STATES 6784090623

## DETAILS

Stonecrest, Ga, United States
6784090623
jodiambra@gmail.com

## LINKS

Github

Portfolio

LinkedIn

## SKILLS

Python

SQL

Machine Learning

Exploratory Data Analysis

Tableau

Scikit-Learn

Statistics

Deep Learning

Fast API

KERAS/ TensorFlow

Regression

Classification

Cat Boost

XG Boost

Light GBM

NLP

## PROFILE

**Data Scientist**

As a seasoned scientist, I have spent over 8 years working in a variety of research and development fields. During this time, I have honed my skills in critical thinking, problem-solving, data analysis, and communicating complex findings to non-technical stakeholders. My passion for using data to uncover insights and drive decision-making led me to pursue a career in data science, where I can combine my scientific knowledge with advanced analytics techniques. I am eager to continue building my skills and applying them to real-world challenges and am excited about the opportunities that the data science field has to offer.

## INTERNSHIPS

**Data Science Intern at Confidential Company**

April 2023 — June 2023

- Contributed research to company's infrastructure, with the goal of training a MENET deep learning model to predict geolocation based on individual text.
- Median and mean differences between predicted and actual distances were 1,146 km and 1,784 km, respectively, demonstrating the model's accuracy.
- Implemented Hugging Face NLP pipelines to extract text features: embeddings, sentiment, topic, language, location.
- Acknowledged the workflow's potential to provide valuable geolocation prediction capabilities, with the possibility of scaling and integrating it into the existing infrastructure for real-time application.

## PROJECTS

**Spotify EDA and Web Application**

December 2022 | Project

- Collaborated with software engineers to build a website for our project
- Led a team of 4 to conduct exploratory data analysis and statistical analysis on Spotify playlists, with interactive data visualizations
- Discovered popularity trends was correlated with music events and pop culture: music awards, fashion shows, movies, super bowl, festivals
- Demonstrated presentation skills with YouTube livestream

**Ice Retail**

November 2022 | Web Notebook

- EDA on video game retail store sales
- Created interactive visualizations, identified patterns that lead to successful sales and other data driven insights
- Formulated business strategy for advertising campaign to maximize profit in regional markets

**Megaline Plus**

November 2022 | Web Notebook

- Used hypothesis testing to determine that the Ultimate plan brings in more revenue than the Surf plan
- Data suggested a marketing push towards the Surf plan would yield a higher cash on cash return
- Illustrated data usage to be the largest contributor to revenue
- Created a classification model to predict upgrade options for legacy plan holders, with 80% accuracy on an imbalanced dataset

**Sure Tomorrow Insurance**
December 2022 | Web Notebook

- Applied machine learning and model tuning to identify similar customers
- Predicted whether a customer was likely to receive an insurance benefit
- Predicted the number of insurance benefits a customer was likely to receive

**Sweet Lift Taxi**
January 2023 | Web Notebook

- Used ML on time series data to predict the number of orders during peak hours
- Tested different regression models, tuned hyperparameters, and chose the best model for the task

**Rusty Bargain Used Car Sales**
January 2023 | Web Notebook

- Conducted EDA on used car sales and cleaned data of missing values and outliers
- Illustrated the difference between simple and boosted models, with model training times and RMSE scores
- Concluded the Light GBM model had a good balance of speed and prediction

## EMPLOYMENT HISTORY

**QC Analyst | Arxada**
December 2019

- Analyzed chemical, raw materials, in-process materials, or release test samples in support of compliance with company standards
- Interpreted and evaluated the analyses in terms of accuracy and precision compared against established specifications; recommended, and implemented corrective action where necessary
- Successfully took on the challenge of analyzing a 120% increase in lab samples
- Reduced the analysis time for HPLC azole method by 130%
- Team lead of 3 chemists

**Research Associate | RCUOG**
September 2016 — April 2019

- Teamed with 2 researchers in successfully identifying pro-inflammatory molecules involved in Areca nut carcinogenesis, using HPLC fractions (doi: 10.1038/s41598-017-18996-2)
- Contributed to leading-edge Areca nut carcinogenicity data; compiled, analyzed, and interpreted results leading to 1 publication, and Areca nut intervention studies with Guam Cancer Research Center
- Presented data to community at conferences and developed further studies and experimental designs to address similar research concerns

**Research Lab Assistant | PITT Dept. Biological Sciences**
September 2011 — April 2013

- Enhanced quantitative and qualitative analyses, such as sampling, testing, and measuring using specialized equipment, shaving 10% off analysis times
- Conducted DNA miniprep (purification) experiments with drosophila DNA with minimal supervision, followed experimental protocols

## COURSES

**Data Science Bootcamp, Practicum US**
2022 — 2023

## EDUCATION

**M.S. Biology, University of Guam**
August 2016 — April 2019

NIH/U54 Scholar

**B.S. Biology, University of Pittsburgh, Pittsburgh**
August 2010 — April 2015