David Moraes

moraes.david02@gmail.com

+44 7470258187

davidanand02.github.io/Portfolio/

I have a passion for using AI and statistics to turn complex data into actionable insights. My experience spans multiple technologies from Python and SQL to R and Excel on a variety of topics.

EDUCATION

BSC (HONS) ACTUARIAL SCIENCE - LONDON SCHOOL OF ECONOMICS & POLITICAL SCIENCE

London School of Economics and Political Science

09/2021 - 06/2024 *London, UK*

Courses

- Relevant Courses: Inferential & Descriptive Statistics, Calculus & Linear Algebra, Corporate Finance, Financial & Management Accounting, Survival Modelling, Actuarial Financial Investigations, Machine Learning, AI, Ethics for Data Science & Databases.

INTERNATIONAL A-LEVELS

St Christophers School

05/2019 - 05/2021

Courses

- Pearson Edexcel: Mathematics - A*, Further Mathematics - A, Economics - A | Oxford AQA: Physics - A*

EXPERIENCE AND AWARD

Student Consultant

Grove Finance

02/2024 - 03/2024

Financial start-up focused on using traditional equity release to finance childcare costs.

Tasks

- **Reported** on how direct lending and equity release can be leveraged in the UK's childcare industry.
- **Researched** Grove's industry, market gaps, competitors, target demographic, trends and opportunities.
- **Collaborated** with an international team to present report to founder.

1ST PLACE IN THE INTER-AMERICAN DEVELOPMENT BANK CHALLENGE

Inter American Development Bank

01/2024 - 02/2024

Tasks

- Accounted for 20 economic, socio-political and financial indicators using **AIC**, **Ridge** and **Lasso regression**.
- **Predicted** future **correlations** across several industries using a **neural net** & presented **statistically significant** findings.

SKILLS

- Python Programming Libraries: seaborn, matplotlib, Ml models
- R Programming Packages: dplyr, ggplot, tidy models
- Excel Tools: Pivot Table, Pivot Chart, XLOOKUP, learning VBA
- Tableau Tools: Charts, dynamic maps, joins
- SQL Software: Microsoft SQL Server Express

PERSONAL PROJECTS

Python: Real and Synthetic Image Detection with CNN's (05/2024)

- **Trained** CNN models to classify human-made and AI-made images with a training set of 100,000 images from the CIFAKE dataset.
- Evaluated performance of models using **ROC-AUC** and **confusion matrices**.
- **Reported** on the comparative performance of the models as well as presented the model architecture to a lay audience.

R Programming: Text and Cluster Analysis for Drug Reviews and Medical data (05/2024)

- Conducted **Latent Dirichlet Allocation (LDA)** to uncover key themes in migraine prevention drug reviews, revealing insights into drug efficacy, side effects, and treatment duration.
- Utilized **quanteda** and **tm** packages to **clean, tokenize, and preprocess** a large dataset of 161,297 drug reviews, including removal of stopwords, punctuation, and stemming.
- Applied the AFINN sentiment dictionary to analyze **sentiment** in drug reviews and created visualizations using **ggplot2** to highlight areas of high and low patient satisfaction.

PYTHON/MONGODB: HOSPITAL AND PHYSICIAN MANAGEMENT DATABASE SYSTEM (01/2024)

- Executed **MongoDB queries** to retrieve and display essential information about doctors, patients, finances and more.
- Developed an **aggregation pipeline** to calculate total revenue by patient, considering different procedure types.
- Visualised **temporal queries** e.g health evolution of patients, using Matplotlib.
- **Performed queries** involving update/delete structures, calculations, views, retrieve, filter and aggregate.

SQL/TABLEAU: COVID-19 EX PLORATORY DATA ANALYSIS & DASHBOARD (05/2023)

- Utilized **joins**, **CTEs**, **temp tables**, and data type conversions for efficient data retrieval and aggregation.
- Applied **window functions and aggregate functions** to calculate death & infection % and rolling counts of vaccinations.
- Developed a **view** to simplify data access and integrated with **Tableau**, creating a dashboard.

EXCEL: BIKE SALES DYNAMIC DASHBOARD & VISUALISATION (05/2023)

- Ensured **data quality** of 1000 rows by removing duplicates & null values.
- Utilized **pivot tables and charts** to visually represent trends in data, enabling clear and insightful analysis.
- Integrated **slicers** to customize the map view, enhancing user interaction and exploration.