

## CAREER SUMMARY

### Data Engineer • Business Intelligence • Data Scientist • Data Analyst

- Data Scientist with extensive experience in leveraging data engineering and modeling abilities and machine learning knowledge to build algorithms that drive operational excellence.
  - Over 18 years of IT experience, adept at collaborating with cross-functional teams to deliver actionable insights while upholding data accuracy and integrity.
  - Successfully led teams through time-critical projects, leveraging advanced tools such as Python, R, and SQL, while effectively managing multiple assignments concurrently.
  - Demonstrated sound judgment, emotional intelligence, and decorum in navigating sensitive, confidential, and unpredictable scenarios, ensuring optimal outcomes.
- 

## SKILLS

**Expert:** SQL (SQL SERVER, Tera Data Studio, DB2) • R (Shiny, Tidyverse, ggplot2) • R-Markdown • VBA • Excel

**Proficient:** Alteryx Designer • pySpark • Python (PyTorch, NumPy, Pandas, TensorFlow, Scikit-learn, Keras, Flask, Django) • SAS • ACCESS • GIT • Data Visualisation (Tableau, Power BI, Matplotlib, Seaborn) • JSON/XML • Machine learning (Supervised and unsupervised learning models - Random Forest, Linear Regression, Logistic Regression, Clustering) • Statistical analysis • Time Series Analysis

**Novice:** Linux • Social Network Analysis • Natural Language Processing (SciPy, NLTK) • Convolutional Neural Networks

---

## PROFESSIONAL EXPERIENCE

**Health Protection Surveillance Center**, Dublin, Ireland

**June 2023 – Present**

### Data Scientist

- Leading the SQL Server database design for epidemiology, overseeing all aspects of architecture, design, development and implementation of HPSC's data infrastructure.
- Panel expert responsible for ascertaining the HPSC requirements in acquiring the cloud version of R.
- Spearheaded the integration and utilisation of visualisation tools as part of the data strategy at HPSC, resulting in improved data interpretation, BI and analytics needs and decision-making processes.
- Spearheaded the visualisation of wastewater data products, incorporating maps and plots such as Sankey diagrams. This was accomplished in record time, showcasing our efficiency and expertise.
- Pioneered the application of LSTM models for time series data analysis, specifically in forecasting influenza trends, resulting in more accurate predictions and proactive public health measures.
- Conceptualised and implemented a comprehensive data pipeline capable of analysing multiple diseases, facilitating epidemiological analyses and enhancing public health surveillance capabilities.
- Developed GAM smoothing (thin plate) techniques in R to visually depict trends in notifiable disease data, enabling stakeholders to identify patterns and make informed decisions.
- Acted as a dynamic leader in advocating for the adoption of R for data analysis, visualisation, and machine learning within HPSC, driving efficiency and innovation in data-driven approaches. Mentored and coached

team members to foster a culture of collaboration, knowledge sharing, and continuous learning, contributing to professional development and team cohesion.

**Dublin Business School**, Dublin, Ireland

**October 2023 – Present**

#### **Dissertation Supervisor (part-time)**

- Provided comprehensive guidance and mentorship to Master's program students throughout their dissertation process, facilitating successful completion and fostering academic excellence. Notably, mentored eight students over 6 months, leveraging strong communication skills and subject matter expertise to support their research endeavors and ensure achievement of academic goals.

**Bank of Ireland**, Dublin, Ireland

**April 2019 – June 2023**

#### **Fraud Analytics, Rules and Strategy (Manager)**

- Conceptualised a robust SQL server design to ensure data integrity, enhance data governance, and mitigate unauthorised access to sensitive information.
- Skillfully engineered a robust database for ETL pipelines, employing a star schema design to facilitate transaction analysis and uncover insights into fraudulent activities and patterns.
- Created a data extraction and cleaning pipeline to reduce human effort in debit fraud refund instances by 90%, saving roughly five hours effort each day. I led a two-person team that designed, developed, tested, and deployed numerous data pipelines using Alteryx Designer and SQL Server in less than two months.

#### **Senior Data Analyst and Senior Analyst – IRB Model Development**

- Headed a team from conceptualising, analysing, building and implementing a data pipeline to remedy the 700,000-customer data and fed to Lexis Nexis in the specific format within very aggressive regulatory timelines and presented the project at the monthly departmental tech meet.
- Actively identified the potential in the data to drive the roadmap and enhance the data quality by utilising analytical and technical capabilities. Built a workflow using Tableau and Teradata SQL.
- Implemented word macros to standardise the Model development document; formatted the document automatically with predefined formatting attributes.
- **Proactively identified an opportunity to automate the monthly ECL (expected credit losses) runs**, reducing manual interventions and eliminating the risk of being non-compliant with publishing the monthly ECLs.

**NTT DATA Services Ireland**, Ireland

**April 2015 – April 2019**

#### **Solution Architecture Advisor**

- Implemented a system monitoring framework that cut down on customer complaints by 50% while increasing the stability of the system. Led a team to design, create, test, and implement the framework.
- Headed a team of developers and testers to design, develop, and implement EFT (Electronic Fund Transfer) solutions with aggressive timelines and within budget, allowing the bank to do away with cheque-based mortgage funds transfers. This improved the customer experience, reduced paper use, and sped up turnaround time.

**Tata Consultancy Services**, Noida, India

**Oct 2010 – Feb 2015**

#### **Assistant Consultant**

- Introduced MS Word macros to standardise the testing findings in documents, which reduced manual errors by 90% and removed document-related concerns for the client. I trained staff in the use of the word macros. Along with the project activity, the solution was managed and put into practice.
- Managed and led teams for project deliveries to clients such as Kuehne + Nagel, AXA insurance and Caremark, delivering on the onshore-offshore model.

**Programmer**

- Led a technical five-person team that designed, developed, and put the automation of deal confirms for third-party mutual funds into place. Improved the turnaround time and reduced the manual efforts by 95%, allowing the system to onboard more third parties to the system.
- Devised automated system health checks, strengthening overall stability of the system of records and improving turnaround time for support teams.

---

**PERSONAL PROJECT – Ireland University Admissions Analysis Project (Python)**

- Wrote Python program to scrape university admissions per secondary schools in Ireland by extracting data from schools' websites using BeautifulSoup, requests.
- Utilised data available on <https://www.gov.ie/en/collection/63363b-data-on-individual-schools/> to enrich the dataset.
- Visualised the data in graph format using NetworkX, Matplotlib and Seaborn.
- Currently working on clustering these data and formulating the hypothesis that more than 50% of university admissions are to university near the secondary schools.

**PERSONAL PROJECT – Stanford Car Computer Visual Project (Python)**

- Wrote Pytorch code to model TinyVGG structure layers of CNN to classify the Model of the car from images.
- Code - <https://www.kaggle.com/code/ankusharya/stanford-cars>
- Blog - <https://www.linkedin.com/pulse/computer-vision-pytorch-model-trained-stanford-cars-dataset-arya-bdcoe/?trackingId=07nALv91MWNzU8swhlwhlw%3D%3D>

---

**EDUCATION**

**Master of Sciences in Big Data Analytics and Artificial Intelligence**

Atlantic Technological University (ATU) Donegal, Ireland

Dissertation: Road Network traffic analysis using Graph Neural Networks (using PyTorch framework)

**Postgraduate Certificate in Applied Statistics** from Technological University Dublin, Ireland

Applied linear and logistic regression models in R

**Master of Business Administrator (Business Management)** from Dayalbagh Educational Institute, Agra, India

Dissertation submitted - Online Grocery Shopping Preferences, used SPSS for analytical inferences.

**Bachelor of Technology in Information Technology**

Kurukshetra University, Kurukshetra, Haryana, India

**Certified ScrumMaster® Jan 2017**

---

**VOLUNTEER EXPERIENCE**

- Organised learning of MS Office and basic computer skills for 20 students over two years for Dayalbagh Educational Institute.
  - Planned, organised and formulated the IT infrastructure for virtual seminars and workshops for ESG (Environmental, social, and corporate governance) for SPHEEHA (Society for Preservation of Healthy Environment and Ecology and Heritage of Agra).
-