

NISHANT RAI

DATA SCIENCE | ANALYTICS

nishant_rai123@hotmail.com 
www.linkedin.com/in/nishant-rai-
81a993329/ 
+44-747-542-9317 
https://github.com/NishantRai567 
Portfolio: https://nishantrai567.github.io/

MOTIVATION *I am passionate about **solving business problems** using Data Science & Analytics. I systematically & creatively use my skillset to **add tangible value** to the team, the business, and the end-user. I am constantly learning, and always looking to improve.*

EDUCATION **MSci in Chemistry**
2020 - 2024 - University of Birmingham, UK
First Class Honours
Head of School Prize for Outstanding Academic achievement

A-Levels
2016 - 2018 - Lawrence Sheriff School, UK
Grades: A*,A*,A*(Chemistry, Biology, Maths), A(Further Maths, AS level)

SKILLS & TOOLS
Programming: SQL, Python (Base, Pandas, Numpy, Matplotlib, Scikit-Learn, Keras), R
Tools: Excel, Tableau, Github, AWS (S3, Lambda, IAM, EC2, SageMaker, RDS, DynamoDB, Glue)
Math: Linear Algebra, Statistics (Hypothesis Testing, AB Testing, Central Limit Theorem, Distributions)
Machine Learning: Linear Regression, Logistic Regression, Decision Trees, Random Forest, KNN, k-means, PCA, Association Rule Learning, Causal Impact Analysis, Neural Networks

PROJECTS **"You Are What You Eat" Customer Segmentation**

- Used **k-means clustering** on grocery transaction data to split out customers into distinct "shopper types" that could be used to better understand customers over time, and to more accurately target customers with relevant content & promotions

Predicting Customer Loyalty Scores Using Machine Learning

- Designed and implemented a **machine learning** solution to predict customer loyalty for a grocery retailer, addressing data gaps. Tested and compared three regression models (Linear Regression, Decision Tree, and Random Forest), with **Random Forest** achieving an Adjusted R-Squared of 0.955 and R-Squared of 0.925 on cross-validation. The model improved customer segmentation and marketing strategies, enabling better decision-making.

A/B Testing for Delivery Club Mailer Campaign

- Used **Python** (Pandas, SciPy) to conduct a **Chi-Square Test** for Independence comparing signup rates between low-cost and high-cost mailer groups. Cleaned and aggregated campaign data into a 2x2 contingency table, and provided actionable insights to the client, showing that the difference in signup rates was not statistically significant, helping optimize future marketing strategies.

NISHANT RAI

DATA SCIENCE | ANALYTICS

nishant_rai123@hotmail.com 
www.linkedin.com/in/nishant-rai-
81a993329/ 
+44-747-542-9317 
<https://github.com/NishantRai567> 
Portfolio: <https://nishantrai567.github.io/>

PROJECTS (Con't)

E-Commerce Power BI Report

- Performed [ETL](#) on a 120,000-record e-commerce dataset, extracting and transforming data from [Azure Storage](#) and [SQL](#) Database into [Power BI](#). Created a star schema model, implemented 25 [DAX](#) measures, and developed a 4-page report with 40 visualizations to generate actionable insights. Additionally, constructed 5 [SQL](#) queries for deeper insights outside Power BI.

Exploratory Data Analysis: Online Shopping in Retail

- Built a custom [Python](#) class using [SQLAlchemy](#) to extract and clean 54,000-record retail dataset. Conducted data preparation with [Pandas](#), including missing value imputation, outlier removal, and transformation. Analysed 2,000+ records, answering 10 business questions and identifying key trends to inform marketing strategies and improve customer targeting.

COURSES & CERTS

Data Science Professional Certification (Data Science Infinity)

Actionable Learnings: Developed skills in data manipulation with [SQL](#), statistical analysis (A/B testing), and [Python](#) for analysis and visualization. Applied machine learning algorithms (regression, classification, clustering) and built ML pipelines for preprocessing and deployment using [Streamlit](#). Gained experience in data preparation (missing values, encoding, scaling) and model validation. Utilized [Tableau](#) for data visualization and [GitHub](#) for version control, while translating business problems into data-driven solutions.

Data Analytics (AiCore)

Actionable Learnings: Specialized in data analytics, mastering [Python](#) for data cleaning, exploration (EDA), and manipulation using [Pandas](#). Gained expertise in statistical analysis and visualizing data insights with [Power BI](#). Developed skills in version control with [GitHub](#) and cloud computing with [AWS](#). Learned best practices for handling various data file types and building data-driven solutions.

EXPERIENCE AI Prompt Engineer - Outlier

DECEMBER 2024 - PRESENT

- Designed complex chemistry prompts to evaluate and challenge LLM reasoning capabilities.
- Identified and addressed domain-specific knowledge gaps, improving model accuracy across 15 scientific inquiries.
- Conducted real-time testing to enhance AI performance and reasoning in chemistry.

Team Member - Dominos

SEPTEMBER 2018 - OCTOBER 2019

- Improved team communication, reducing order processing time by 15% during peak hours while maintaining 95%+ customer satisfaction.
- Analysed workflows and implemented solutions, increasing order throughput by 30+ orders per hour during high-demand periods.