

# NISHANT RAI

DATA SCIENCE | ANALYTICS

nishant\_rai123@hotmail.com   
<https://www.linkedin.com/in/nishant-rai-data/>   
+44-747-542-9317   
<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.

## SKILLS & TOOLS

**Programming:** SQL, Python (Base, Pandas, Numpy, Matplotlib, Scikit-Learn)  
**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

- Applied [K-means clustering](#) to grocery transaction data to segment customers into distinct "shopper types".
- Analysed shopper types to understand customer behaviour over time and improve targeting of promotions and content.
- Enhanced marketing effectiveness by enabling more relevant, targeted campaigns based on customer segments.

### Quantifying Sales Uplift With Casual Impact Analysis

- Used [Causal Impact Analysis](#) to assess the effect of the "Delivery Club" initiative on sales.
- Compared sales for Delivery Club members against a control group to identify the initiative's performance.
- Identified a [41.1% uplift](#) in sales for members, leading to refined marketing strategies and future initiatives.

### Predicting Customer Loyalty Scores

- Built a [machine learning](#) model to predict customer loyalty for a grocery retailer, aiming to enhance retention and marketing effectiveness.
- Handled data gaps and tested three [regression](#) models (Linear Regression, Decision Tree, Random Forest); [Random Forest](#) performed best (Adj.  $R^2 = 0.955$ ,  $R^2 = 0.925$ ).
- Enabled more accurate customer segmentation and informed targeted marketing strategies, improving decision-making across teams.

### A/B Testing for Delivery Club Mailer Campaign

- Performed [Chi-Square Test](#) for Independence on signup rates between low-cost and high-cost mailer groups.
- Aggregated campaign data into a 2x2 contingency table for analysis using Python (Pandas, SciPy).
- Discovered no significant difference in signup rates, providing insights for more targeted marketing decisions.

# NISHANT RAI

DATA SCIENCE | ANALYTICS

nishant\_rai123@hotmail.com   
<https://www.linkedin.com/in/nishant-rai-data/>   
+44-747-542-9317   
<https://nishantraai567.github.io/> 

## PROJECTS (Con't)

### E-Commerce Power BI Report

- Conducted [ETL](#) on a 120,000-record e-commerce dataset, extracting data from Azure Storage and SQL Database into Power BI.
- Built a [star schema](#) model and implemented 25 DAX measures to enhance reporting.
- Created a 4-page report with 40 visualizations, delivering insights that improved marketing and business strategies.

### Exploratory Data Analysis: Online Shopping in Retail

- Built a custom Python class with [SQLAlchemy](#) to extract and clean a 54,000-record retail dataset.
- Performed data preparation using [Pandas](#) (missing value imputation, outlier removal) to ensure accurate analysis.
- Analysed 2,000+ records, identifying trends that helped refine marketing strategies and improve customer targeting.

## 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.

## 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.

## 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\* (Chemistry), A\* (Biology), A\* (Mathematics), A (Further Mathematics - AS Level)