

Roopkumar Das

Pune, India | +91 9766890255 | [Gmail](#) | [LinkedIn](#) | [Kaggle](#)

Proficient in Python, pandas, and matplotlib for data analysis and visualization, with basic knowledge of SQL and Excel. Strong analytical mindset with hands-on experience in data manipulation and insight generation.

TECHNICAL SKILLS

- | | | | |
|-------------------|----------------------|-----------------|---------------------|
| • Advanced Python | • Advanced Pandas | • Matplotlib | • Excel |
| • SQL | • SQLite | • MySQL | • Python Notebooks |
| • Data Modelling | • Data Visualization | • Data Cleaning | • Data Storytelling |

PROJECTS

B2B Lead Generation for Branded Merchandise | Python, Playwright, Google Maps Scraping | [Link](#)

- Scraped Google Maps to extract a dataset of cafés and restaurants for B2B branded merchandise lead generation.

Online Retail Store Analysis | Python, Pandas, Matplotlib, SQL (Basic) | [Link](#)

- Identified that wholesale transactions dominate non-UK markets, while the UK has a mix of retail and wholesale, contributing 85% of total revenue.
- Showed that trend analysis is unreliable due to a single-year dataset; instead, identified peak sales months and seasonal revenue patterns.
- Analyzed product-level sales, finding no single-product dependency, with revenue distributed across multiple items, ensuring business stability.
- Visualized insights with line plots (daily revenue trends), bar charts (monthly revenue, top products), and pie charts (regional revenue distribution) to support data-driven decisions.

Titanic Survival Prediction using RandomForestClassifier | Python, Numpy, Pandas, Sklearn | [Link](#)

- Performed EDA and handled missing values using statistical imputation (mode for Embarked, grouped mean for Fare, and Random Forest Regressor for Age).
- Engineered relevant features to improve classification performance for survival prediction.
- Tuned RandomForestClassifier parameters using GridSearchCV, optimizing hyperparameters for better accuracy.
- Achieved a 0.78708 score (~87.76 percentile), demonstrating strong model performance through iterative experimentation.

EDUCATION

Bachelor of Science in Data Science and Applications

01/2023 – 01/2027

Indian Institute of Technology, Madras