Roopkumar Das

Pune, India | Gmail | Personal Website | Linkedin | Github

I'm a data-driven problem solver who builds end-to-end solutions that deliver real business value, not just reports. I take initiative, work with messy real-world data, and ask the right questions to uncover what actually moves the needle. Reliable, proactive, and outcome-focused—I've already been doing the work you're hiring for.

Technical Skills

- Tableau
- Intermediate Excel
- Advanced Python
- SQL
- Pandas
- Matplotlib
- Linux
- Python Notebooks
- Data Cleaning
- Data Modeling
- Data Visualization
- Data Storytelling

Projects

Increasing Revenue for Offline Clothing Store through Demand Analysis | Python, Pandas, Matplotlib, Google Sheets | Blog Link

- Collected and organized 30 days of daily transaction data using Google Sheets to enable structured analysis for revenue improvement.
- Identified peak revenue days (Thursday and Sunday) and consistent weekly earnings of Rs. 25,000 to inform targeted staffing and stocking.
- Analyzed customer arrival times and quantity sold per day to align inventory with demand and reduce missed sales opportunities.
- Identified high-demand products—Underwear, Vest, T-Shirt, Socks, Half Pant, Towel, Shirt—using product-level and time-wise heatmaps to guide inventory focus.
- Found that higher-priced items face more price negotiation; suggested pricing strategy adjustments and emphasized domain knowledge for effective analysis.

Online Retail Gift Store Market Expansion Analysis | Python, Pandas, Matplotlib, Tableau, Presentation | Basic Analysis | Expansion Strategy Video Presentation

- Cleaned and prepared UK-based retail dataset like by resolving StockCode to Description mismatches and extracting relevant features.
- Found that the UK accounts for 80% of total revenue and quantity sold; peak sales months are September to December based on daily and monthly revenue analysis.
- Identified top secondary markets (mostly European countries near the UK) with 80–90% of customers being wholesalers responsible for 90-95% of revenue.
- Showed that 20% of products contribute to 80% of revenue and sales in each country-month combination, helping guide inventory planning during expansion.

Titanic Survival Prediction using RandomForestClassifier | Python, Numpy, Pandas, Sklearn | Kaggle Notebook

- 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

Indian Institute of Technology, Madras, Data Science and Applications — Jan 2023 – Jan 2027

- CGPA: 9.27/10
- **Current Coursework:** Data Science Foundations, Statistics & Probability, Machine Learning

Extra Curricular & Certifications

Tata Data Visualisation: Empowering Business with Effective Insights Job Simulation on Forage - April 2025

- Completed a simulation involving creating data visualizations for Tata Consultancy Services
- Prepared questions for a meeting with client senior leadership
- Created visuals for data analysis to help executives with effective decision making
- Certificate of this job simulation