

# Dhanesh B B

Data Scientist, Machine Learning, Python & SQL  
Bengaluru, India, +91 80508 35879, [dhaneshbb5@gmail.com](mailto:dhaneshbb5@gmail.com)

LinkedIn: [linkedin.com/in/dhaneshbb](https://www.linkedin.com/in/dhaneshbb), GitHub: [github.com/dhaneshbb](https://github.com/dhaneshbb), Portfolio: [dhaneshbb.github.io](https://dhaneshbb.github.io)

## PROFESSIONAL SUMMARY

---

Data Scientist with economics background and expertise in end-to-end ML pipelines and large-scale data processing (58.4M+ records) using Python and SQL. Built classification and regression models for default prediction, lead scoring, and demand forecasting. Published two PyPI packages for automated EDA; IABAC and NASSCOM certified.

## SKILLS

---

**Languages:** Python, SQL

**Machine Learning:** scikit-learn, XGBoost, LightGBM, CatBoost, SHAP; Classification, Regression, Cross-validation, Hyperparameter Tuning, Feature Engineering

**Data:** Pandas, NumPy, SciPy, Statsmodels; EDA, Data Cleaning, ETL, Statistical Analysis

**Databases:** PostgreSQL, MySQL, SQLite

**Visualization:** Matplotlib, Seaborn, Plotly

**Tools:** Git, GitHub Actions, pytest, Jupyter, VS Code

## EXPERIENCE

---

### Data Science Consultant, Rubixe (Think Ahead Innovations Pvt. Ltd.)

Bengaluru, India, Sep 2024 - Mar 2025

- Executed POC projects and client assignments involving data cleaning, EDA, visualization, and ML model development.
- Solved real-world data science challenges; recognized by project teams and clients for delivering quality results.
- Applied theoretical concepts to practical business problems; committed to excellence with consistent, diligent delivery.

### Coordinator, Pantaloons (Aditya Birla Fashion and Retail Ltd.)

Bengaluru, India, Apr 2022 - Mar 2023

- Maintained 99% transaction accuracy across 100+ daily records; automated Excel reporting to reduce errors 15%.

## PROJECTS

---

### AutoCSV Profiler, PyPI Library, Feb 2025 - Present

PyPI: [pypi.org/autocsv-profiler](https://pypi.org/project/autocsv-profiler), GitHub: [github.com/dhaneshbb/autocsv-profiler](https://github.com/dhaneshbb/autocsv-profiler)

- Published Python CLI + API for statistical CSV profiling with chunked processing; cut manual EDA from hours to minutes.

### InsightfulPy, PyPI Library, Dec 2024 - Present

PyPI: [pypi.org/insightfulpy](https://pypi.org/project/insightfulpy), GitHub: [github.com/dhaneshbb/insightfulpy](https://github.com/dhaneshbb/insightfulpy)

- Published EDA toolkit with 30 reusable functions for summaries, data-quality checks, and visualization.

### Home Loan Default Prediction, Jun 2024 - Mar 2025

GitHub: [github.com/dhaneshbb/HomeLoanDef](https://github.com/dhaneshbb/HomeLoanDef)

- Trained XGBoost default-risk model on 58.4M records (7 datasets, 218 features); achieved 83.0% accuracy and 0.785 ROC-AUC with threshold tuning.
- Reduced memory usage by 68.5% via chunked processing and reduced false positives by 20% to improve risk flag precision.

### Sales Lead Scoring (FicZon), Feb 2025 - Mar 2025

GitHub: [github.com/dhaneshbb/FicZon-Sales-Effectiveness](https://github.com/dhaneshbb/FicZon-Sales-Effectiveness)

- Developed XGBoost lead-quality classifier on 7,420 leads; optimized threshold for high-potential capture (84.74% recall, 0.8106 ROC-AUC).
- Estimated 384 hours saved annually and 45% fewer junk-lead reviews via automated pre-categorization.

### Employee Performance Analysis (INX), Mar 2025 - Apr 2025

GitHub: [github.com/dhaneshbb/INX-Employee-Performance-Analysis](https://github.com/dhaneshbb/INX-Employee-Performance-Analysis)

- Created multiclass XGBoost model (3 classes) with 92.5% accuracy and 0.9756 ROC-AUC; used SHAP explainability to identify key drivers (satisfaction, salary hike, promotion delay).

### **Bike Rental Demand Forecasting, Jun 2024 - Mar 2025**

GitHub: [github.com/dhaneshbb/BikeRental](https://github.com/dhaneshbb/BikeRental)

- Built Ridge regression forecasting model on 731 daily records with temporal and weather features; achieved 83.2% R2 with 0.5% overfitting gap (vs XGBoost 9.4%).

**Additional Projects:** Automobile Price Prediction, Portuguese Bank Marketing, AutoCSV Profiler Suite, Multi-AI Chat Manager, SpectraTact

## **EDUCATION**

---

### **Master of Arts in Economics, Bangalore University**

Bengaluru, Aug 2021 - Nov 2023

CGPA: 7.15/10 (First Class Distinction)

Relevant coursework: Econometrics, Statistical Methods, Research Methodology

### **Bachelor of Arts in History, Economics, Political Science, Bangalore University**

Bengaluru, May 2018 - Aug 2021

CGPA: 7.36/10 (First Class Distinction)

## **CERTIFICATIONS**

---

### **Certified Data Scientist Program, DataMites**

Bengaluru, Jan 2024 - Mar 2025

120-hour program covering Python, ML, Deep Learning, and Big Data (Hadoop, Spark); 4 capstone projects with A/A+ grades.

### **Certified Data Scientist, IABAC (International Association of Business Analytics Certification)**

Apr 2025 - Apr 2028, Grade A

### **Certified Data Scientist, NASSCOM FutureSkills Prime**

Apr 2025, Gold Category (NSQF Level 5)

**Kaggle Certifications:** Python, SQL, Pandas, ML, Deep Learning, Feature Engineering, Time Series (+10 more)