

Shreyas Mahangade

+91-9322651611 | shreyasmahangade1@gmail.com | [linkedin.com/in/shreyas3011](https://www.linkedin.com/in/shreyas3011) | github.com/shreyas3011

PROFILE SUMMARY

Data-driven Analyst with a strong foundation in SQL, Python, and statistical analysis, skilled at transforming raw data into actionable business insights. Experienced in building interactive dashboards and reports, automating data pipelines, and applying exploratory data analysis (EDA) to support data-driven decision-making. Proficient in Power BI, Excel, and SQL, with an AI/ML background that strengthens predictive and diagnostic analysis.

EDUCATION

- Bachelor of Engineering (B.E) Artificial Intelligence and Data Science** 2022 – 2026
Savitribai Phule Pune University (SPPU), Pune /CGPA: 9.20/
- Higher Secondary Certificate (H.S.C) Science** 2021 – 2022
Maharashtra State Board, Kolhapur /76%/

EXPERIENCE

- Machine Learning Engineer Intern** Feb 2025 – Jul 2025
Design Techno SERVICES PVT LTD Pune, Maharashtra
- Collaborated with cross-functional team members to build and tune classification models improving accuracy by 18% through effective feature engineering and hyperparameter tuning with GridSearchCV.
 - Automated preprocessing workflows using Pandas and NumPy, reducing data preparation time by 25% and streamlining model evaluation through cross-validation.
 - Deployed a Streamlit app for real-time inference, enhancing accessibility and showcasing an end-to-end ML pipeline.
- AI/ML Intern** Jan 2026 – Mar 2026
Ecomomitychain AI (ECC) Remote
- Collaborated with an AI/ML team to design and implement knowledge graph-based systems for supply chain disruption analysis, modeling entities and relationships to enable intelligent decision-making.
 - Built and managed graph data pipelines using Neo4j, contributing to schema design, data transformation, and efficient querying of large-scale supply chain data.
 - Contributed to the development of AI-driven applications by performing data preprocessing and supporting machine learning workflows integrated with graph-based insights.

PROJECTS

- Banking Client Segmentation & Risk Analytics Dashboard** / Python, SQL, Power BI / [Project Link](#) /
- Engineered an end-to-end data pipeline cleaning 3,000+ customer records in Python (pandas), resolving missing values via flag-and-impute strategy, removing duplicate records, and handling outliers with IQR-based clipping — improving data integrity for analysis.
 - Wrote 65 SQL queries (MySQL) across 8 business domains using window functions, CTEs, and subqueries to analyze risk and income.
 - Conducted exploratory data analysis and correlation analysis across 40+ engineered features in a 17-section Jupyter notebook, uncovering a 0.63 correlation between customer risk rating and income that reframed how risk exposure was interpreted.
 - Built a 5-page Power BI dashboard with 13 DAX measures on a star-schema model, including drill through for advisor-level detail.
 - Delivered recommendations on cross-sell targeting and loyalty tier realignment, translating raw data into business action.
- Config-Driven Predictive Diagnostics Pipeline for APS Sensors** | Scikit-Learn, Docker, AWS (EC2/S3), Linux / [Project Link](#) /
- Built a modular, config-driven ML pipeline for APS component failure prediction using high-dimensional sensor data.
 - Applied KNN Imputer, RobustScaler, and SMOTE-Tomek to handle missing data, outliers, and class imbalance.
 - Optimized multiple models using a cost-sensitive metric to reduce high-risk false negatives.
 - Built Dockerized AWS deployments with GitHub Actions CI/CD for automated testing, deployment, and scalable cloud inference.
 - Achieved 99.6% accuracy using XGBoost while minimizing failure-related costs.

TECHNICAL SKILLS

- Languages:** Python, SQL
- BI & Visualization:** Power BI, Tableau, DAX, Power Query, Excel, Dashboard Development
- Data Analysis & Statistics:** Statistics, Exploratory Data Analysis (EDA), Data Cleaning, Data Visualization, KPI Analysis, Business Metrics, Pandas, NumPy, Matplotlib, SciPy
- Business Intelligence & Reporting:** Business Intelligence (BI), Reporting, Storytelling with Data, ETL, Data Modeling, Database Management, Forecasting, Trend Analysis, Root Cause Analysis
- Tools & Databases:** Git & GitHub, MySQL, Neo4j, PostgreSQL, Docker, AWS
- AI & Machine Learning:** TensorFlow, Scikit-Learn, LLMs, Generative AI, Deep Learning, NLP, AI Agents, RAG

CERTIFICATIONS

- Data Science & Machine Learning [EduTech Warje, Pune, 2025]
- Oracle Cloud Infrastructure 2025 Generative AI Professional (1Z0-1127-25) [Oracle University, 2025]