

# Naveen M

Junior Data Analyst (Data Cleaning | SQL | Python | Power BI | Tableau | Statistical Analysis)

📍 Palani, Tamil Nadu | ☎ +91 93447 91056 | 📩 iamnaveen1401@gmail.com | 💬 LinkedIn: naveen1401 | 🐾 GitHub: iamnaveen1401

## Professional Summary

Results-driven Junior Data Analyst with expertise in **data cleaning, SQL, Python, and business intelligence**. Strong background in **statistical analysis, machine learning, and data visualization** using **Power BI & Tableau**. Proven ability to transform raw data into actionable insights, optimize queries, and build predictive models. Certified in **Data Science & SQL** with hands-on experience in **ETL processes, exploratory data analysis (EDA), and dashboard development**.

## Technical Skills

**Programming:** Python (Pandas, NumPy, Scikit-learn), SQL (MySQL, PostgreSQL)

**Data Visualization:** Power BI (DAX, Power Query), Tableau, Matplotlib, Seaborn

**Database & Analytics:** Query Optimization, Joins, CTEs, Stored Procedures, ETL

**Machine Learning:** Regression, Classification, Feature Engineering, Model Evaluation

**Statistical Analysis:** Hypothesis Testing, A/B Testing, Descriptive Statistics

**Tools:** Excel (PivotTables, VLOOKUP), Jupyter Notebook, Git

## Professional Experience

Machine Learning Intern | TechCiti Software

Aug 2024 – Nov 2024

- Optimized **SQL queries**, improving report generation speed by **25%** and reducing runtime.
- Developed **predictive models** (Logistic Regression, Random Forest) with **85%+ accuracy** for customer behavior analysis.
- Automated **data preprocessing** for **10,000+ records** using **Pandas**, reducing manual effort by **30%**.
- Assisted in deploying **ML models** using **Flask** and tracked KPIs for performance improvement.

## Projects

Lung Cancer Prediction | Python, Scikit-learn [🔗](#)

- Built a **classification model (Logistic Regression)** with **87% accuracy** to predict cancer risk.
- Performed **EDA** to identify key features (smoking habits, chest pain) impacting predictions.
- Reduced **false negatives by 15%** through hyperparameter tuning, improving clinical relevance.

E-Commerce Customer Churn Analysis | SQL, Power BI [🔗](#)

- Analyzed **10,000+ customer records** to identify churn drivers (satisfaction score, cashback usage).
- Developed **15+ optimized SQL queries** and interactive **Power BI dashboards** for churn tracking.
- Provided insights leading to a **19% reduction in churn** for high-risk customer segments.

Smartphone Addiction Prediction | Python, Firebase, Scikit-learn

- Trained a Random Forest model with **89% accuracy** on behavioral data.
- Engineered **10+ features** and used **SHAP values** for model interpretability.
- Built a **full-stack dashboard** (Firebase Firestore + Python) for real-time monitoring.

## Certifications

Data Science & Machine Learning – Entri Elevate (Illinois Tech)

Aug 2024 – Jun 2025

Credential ID: 7KOHOII-GVIU | [Certificate Link](#)

- Covered data analysis, machine learning, statistics, MySQL, Python, Power BI, and Excel.
- Hands-on projects on real-world data with a focus on implementation and reporting.

SQL (Intermediate) – HackerRank

May 2025

Credential ID: 07e078373847 | [Certificate Link](#)

- Demonstrated proficiency in complex SQL operations including joins, subqueries, and set operations.
- Validated practical skills in working with relational database queries for data analysis.

## Education

Bachelor of Technology in Computer Technology

Sep 2021 – Apr 2025

Bannari Amman Institute of Technology

Erode, Tamil Nadu