Villarasu S

Data Analyst

villarasu2003@gmail.com | 9042819493 | Dharmapuri, India | Linkedin | Github

PROFILE

Motivated entry-level data analyst with a strong foundation in Python, SQL, and Excel. Hands-on experience in basic data cleaning and visualization. Eager to apply and develop analytical skills to support business decision-making and learn in a professional environment.

SKILLS

Python ⏐ Excel ⏐ PostgreSQL ⏐ Power BI ⏐ Data Visualization ⏐ Exploratory Data Analysis (EDA) ⏐ Machine Learning ⏐ Statistics ⏐ Computer Vision (CV) ⏐ Streamlit ⏐ ChatGPT

PROJECTS

SecureCheck: Python—SQL Digital Ledger for Police Post
Key Skills: Python, PostgreSQL, Streamlit

- Built a Streamlit app with PostgreSQL backend to manage **10,000+ police stop records**, improving reporting speed by 30%.
- Designed dashboards to **track vehicles, violations, and demographics**, enabling faster insights.
- Integrated **predictive models and digital log submission features** for efficient record-keeping.
- Applied SQL queries (joins, aggregations) to identify violation trends and location hotspots.

AgriData Explorer
Key Skills: PostgerSQL,PowerBi,EDA

- Analyzed **1M+ agricultural records** in PostgreSQL and performed EDA to study **crop yields and patterns**.
- Developed **interactive Power BI dashboards** for state- and district-level insights, enabling **data-driven agricultural planning**.
- Automated **data cleaning and loading workflows**, reducing manual effort by **40%**.
- Identified key productivity drivers, supporting **regional agricultural decision-making**.

Employee Attrition Analysis & Prediction
Key Skills: Machine Learning, Data Preparation, Data Visualization, Python, PostgreSQL, Power BI

- Built ML models (Random Forest, Logistic Regression) to predict attrition with **82% accuracy**, analyzing data from over **5,000 employee records**
- Engineered features and handled missing values to improve model performance and insights.
- Created **Power BI dashboards** to visualize attrition trends and key HR metrics.
- Provided **actionable recommendations** to HR for employee retention strategies.

Multiple Disease Prediction
Key Skills: Python, Scikit-learn, Pandas, NumPy, Matplotlib, Seaborn, Streamlit

- Developed ML models to predict diabetes, heart disease, and Parkinson's disease using a patient dataset of **1,400+ samples**.
- Applied normalization, feature selection, and visualization to highlight **risk factors and patterns**.
- Built an **interactive Streamlit app** for real-time disease prediction and analysis.
- Enabled users to **assess health risk factors instantly**, improving accessibility of medical insights.

EDUCATION

- Master Data Science Program | **GUVI | May 2025 – Aug 2025**
- Bachelor of Commerce | Sri Vidhya Mandir Arts and Science College (Autonomous) | Jun 2021 – Mar 2024

CERTIFICATES

- ChatGPT for Everyone Certificate – GUVI