

# Vaishnavi Mujmer

✉ mujmervaisnavi@gmail.com | 📞 +91 81039 08421 | 🌐 im-vaishnavi | 📺 im-vaishnavi

## EDUCATION

### Integrated Master of Technology

Vellore Institute of Technology

2022 – 2027

CGPA : 8.68/10

## EXPERIENCE

### Data Analyst Intern

Inkgraph Techno Pvt. Ltd.

Oct 2025 – Dec 2025

- Engineered automated **Python-based data validation frameworks** and ETL pipelines to cleanse 50,000+ records, integrating **unit testing** to ensure data integrity and reducing processing time by **40%**.
- Designed and deployed **interactive Tableau dashboards** using complex calculated fields and parameters, migrating legacy reports to provide stakeholders with real-time, high-fidelity performance metrics.
- Developed **automated Python scripts** for outlier detection and trend analysis, streamlining the reporting lifecycle and increasing data accuracy by **25%** through continuous integration of new data sources.
- Optimized SQL queries and database schemas for **Tableau integration**, reducing dashboard load times and enhancing the scalability of client-facing analytics products.

## PROJECTS

### Customer Segmentation & Demand Forecasting

Tech stack : R, Tidyverse, ggplot2, Hypothesis Testing

- Analyzed a dataset of **1 million+ ride records** to isolate behavioral differences between user segments, applying **correlation analysis, ANOVA, and T-tests** to validate retention hypotheses.
- Uncovered distinct **temporal usage patterns**, identifying a 20% higher weekend usage rate among casual riders to drive a targeted, data-backed membership conversion strategy.
- Generated **5+ executive-level visualizations** using ggplot2, synthesizing complex mobility data into a clear narrative for stakeholder presentation and strategic planning.

### Heart Disease Prediction & Risk Analysis

Tech stack : Python, Scikit-learn, Pandas, XGBoost, Matplotlib

- Developed a robust **binary classification model** (Random Forest & XGBoost) to predict heart disease risk, optimizing for **Recall** to minimize false negatives in critical medical diagnoses.
- Conducted extensive **Exploratory Data Analysis (EDA)** and feature engineering on clinical parameters (e.g., cholesterol, BP), employing **SMOTE** to handle class imbalance and improve model generalization.
- Achieved **92% accuracy** after implementing **GridSearchCV** for hyperparameter tuning, and visualized feature importance to provide interpretable medical insights.

## TECHNICAL SKILLS

- Languages & Core** : Python (Pandas, NumPy, SciPy), R (Tidyverse), SQL
- Machine Learning** : Keras, Scikit-learn, Supervised, Unsupervised, ANNs
- Visualization & BI** : Tableau, Power BI, Matplotlib, Seaborn, ggplot2, Streamlit
- Tools & Platforms** : Git, Docker, Linux, Google BigQuery, PostgreSQL

## CERTIFICATIONS

- Google Advanced Data Analytics Professional Certificate** – Coursera 11/2025
- Machine Learning (ELITE Certification)** – NPTEL 10/2024