

# Madhuri Malgareddy

madhurireddymalga@gmail.com, Houston, Texas 77098, (832)523-8503

[LinkedIn](#) | [GitHub](#) | [Portfolio](#)

## SUMMARY

ML Data Analyst with 3 years of experience at Amazon Inc., specializing in hypothesis testing and KPI analysis to enhance operational efficiency and improve ML model accuracy. Skilled in leveraging Tableau and statistical techniques to deliver insights and optimize workflows.

## EDUCATION

University of Houston	Master of Science in Business Analytics (CGPA 3.8)	2023-2025
Osmania University	MBA in Finance	2016-2018
Osmania University	Bachelors in Business	2012-2016

## SKILLS / CERTIFICATIONS

**Languages/platforms:** Python, SQL, R, Tableau, Excel, Data orchestration (Mage AI), Looker Studio, GCP, Big Query, AWS

**Modelling skills:** Regression Analysis, A/B testing, Large Language Models (LLMs), PCA, Random Forest, KNN, Time-series Forecasting

**Certifications:** Gen AI with LLM's (AWS & DeepLearning.ai), A/B Testing (Datacamp)

## EXPERIENCE

**ML Data Analyst – Amazon Inc, India**

**Jan 2021-Jul 2023**

- **Conducted Hypothesis Testing** to analyze the effect of updates in Standard Operating Procedures (SOPs) on the accuracy of tasks performed by ML Data Associates, achieving a 15% improvement in task accuracy and enabling data-driven workflow enhancements.
- **Developed 10+ Tableau reports** to analyze Key Performance Indicators (KPIs) such as task completion time and accuracy, driving 20% efficiency gains in data annotation processes.
- Performed high-quality **data Labelling to fine-tune Large Language Models (LLMs)** for Amazon Go's Just Walk Out (JWO) technology, ensuring precise labeling and enabling autonomous billing of products in stores
- **Monitored inter-annotator agreement** using Cohen's Kappa metric to evaluate data quality and consistency with ground truth data, providing actionable insights that contributed to a 20% reduction in discrepancies
- **Ensured the accuracy and reliability** of training datasets by performing detailed quality checks, identifying and correcting 100+ errors weekly, and maintaining consistent labeling standards for machine learning models.
- **Mentored 20+ new hires** in various annotation techniques and developed documentation to enable effective onboarding
- Enabled various strategies **by participating in 12+ Weekly Business Reviews** ensuring alignment with operational goals and enhancing data workflows

## ACADEMIC PROJECTS

**Market Research and Pricing Strategy Development for Electric Scooters and Bikes**

- Conducted a **Causal Factorial experiment** to analyze the impact of pricing strategies on willingness to pay (WTP) for electric bikes, designing a comprehensive questionnaire and effectively collecting survey data from 400 students

**Blood Stream Infections (CLABSI) prediction**

- Performed exploratory data analysis (EDA) on data provided by Texas Children Hospital and evaluated various ML models, including Neural Networks, Random Forests, Decision Trees, and KNN

**Timeseries Forecasting for Wilkins, a Zurn company**

- Forecasted the demand for plumbing and fire valve products using the HOLT model, HOLT-Winters model, ARIMA, and exponential smoothing among which HOLT-Winters demonstrated a superior performance with the lowest MAPE of 8.2%

**Insurance subscription prediction**

- Built a Random Forest model in Python to predict customer subscription to insurance policies, achieving 70.93% accuracy and a ROC-AUC score of 73.53%

## PERSONAL PROJECTS

**NYC Taxi Analysis:** Built a data pipeline for NYC Taxi Data on **GCP** and generated insights through **Looker Studio** Dashboard

**Amazon Sales Analysis:** Developed a **Tableau dashboard** to analyze the trends in Amazon sales data