

# Satya Harish Reddy Pulipelli

Boston, MA | [pulipelli.s@northeastern.edu](mailto:pulipelli.s@northeastern.edu) | [LinkedIn](#) | [Portfolio](#)

## EDUCATION

### Northeastern University, College of Social Sciences & Humanities

*M.S. in Economics - Data Science Concentration*

Boston, MA

- Courses: Algorithms, Advanced Econometrics, Econometric Modeling, Machine Learning & AI, Statistics
- GPA: 3.9/4.0

Sept 2024 – May 2026

### ICFAI Business School

*Bachelor of Economics*

Hyderabad, India

- Courses: Predictive Analytics, Machine Learning, Econometrics, Financial Economics
- GPA: 3.8/4.0 | Honors: Creative Head – Young Economists Society

Sept 2020 – May 2023

## TECHNICAL KNOWLEDGE

**Programming Languages:** Python, SQL, R | **Database Management:** MySQL, PostgreSQL, BigQuery, MS Access

**Statistical & Data Analysis:** Statistical Modeling, Predictive Modeling, Data Mining, Quantitative Analysis, ETL

**Tools & Visualization:** Tableau, R Studio, Stata, SPSS, EViews, SAS, Oracle, Advanced Excel, VBA, Jira

## PROFESSIONAL EXPERIENCE

### Data Science Fellow

Sept 2025 – Present

*Northeastern University, Registrar's Office*

Boston, MA

- Collaborated with University Registrar's Office to transform course scheduling from supply-based to demand-based models through network analysis and enrollment pattern modeling, identifying structural barriers in academic pathways affecting 6,364 courses across 936 undergraduate and graduate programs
- Engineered automated ETL pipeline to extract and structure data from 5,248 pages of unstructured PDF catalogs using Python, regex, and BeautifulSoup, generating zero-cost local processing solution that parsed 8,072 prerequisite relationships and comprehensive program requirements
- Developed Graph Neural Network with multi-head attention mechanism for bottleneck discovery and LSTM model for predicting student course-taking sequences, implementing PyTorch-based deep learning architecture to identify structural barriers and enrollment patterns across prerequisite networks
- Implemented sentence transformer models for semantic course analysis identifying 25,431 similar course pairs across 334 content clusters, and designed 8-tab interactive Plotly Dash dashboard analyzing 25,000+ enrollment records to reveal 36% of courses operating at or over capacity with 31.9% maintaining waitlists

### Business Development Analyst

Jun 2023 – Jul 2024

*Micro SaaS Capital*

Bengaluru, India

- Implemented data-driven strategies for market penetration and product diversification, enhancing company revenue by 20% through statistical modeling and optimized pricing algorithms
- Developed interactive data visualizations and Tableau dashboards to analyze key performance metrics, enabling stakeholders to identify revenue opportunities through data-driven decision making, resulting in 40% improvement

## ACADEMIC PROJECTS

### Economic Analysis of Crime Reporting Behavior | *Econometric Research & Design*

Mar 2025

- Applied research design methodology to analyze 28,821 crime victimization incidents from National Crime Victimization Survey using Linear Probability Models with interaction terms in Stata for quantitative analysis of socioeconomic disparities
- Implemented comprehensive ETL processes and data preprocessing using advanced Stata programming to generate interaction terms for complex econometric modeling and statistical analysis

### Economic Impact on Data Science Salaries | *Predictive Modeling Project*

Feb 2025

- Developed predictive modeling solutions using Multiple Linear Regression, Random Forest, and XGBoost for quantitative analysis of economic factors impacting data science salaries across countries and experience levels
- Implemented comprehensive ETL processes and data engineering techniques to integrate salary data with macroeconomic indicators using Python programming and R programming

### Development and Analysis of iPhone Quality Index | *Statistical Modeling Project*

Dec 2024

- Constructed a Comprehensive Quality Index incorporating 12 technical specifications weighted by consumer preferences for quantitative analysis of innovation patterns across 45 iPhone models (2007-2024)
- Applied statistical modeling and data mining techniques in R using libraries including dplyr, tidyr, and ggplot2 to calculate compound annual growth rates, demonstrating 14.25% annual improvement through predictive analytics
- Created advanced data visualizations and statistical models revealing distinct evolutionary phases through comprehensive data analysis