**MEGHNA REDDI**

973.906.9462 | [mreddi2001@gmail.com](mailto:mreddi2001@gmail.com) | [/in/meghnareddi](https://www.linkedin.com/in/meghnareddi/) | [Github](https://github.com/meghnareddi) | Jersey City, NJ

**EDUCATION**

New Jersey Institute of Technology, Newark, NJ **–** Master of Science (M.S.) – Data Science, Concentration in Statistics **–** GPA: 3.8

Mahatma Gandhi Institute of Technology, Hyderabad, IND – Bachelor of Technology (B.Tech.) – Computer Science – GPA: 3.5

**TECHNICAL SKILLS**

Dashboard Links: [Elections Dashboard 1](https://public.tableau.com/app/profile/meghna.reddi/viz/Dashboard2_17366165974020/Dashboard1) , [Elections Dashboard 2](https://public.tableau.com/app/profile/meghna.reddi/viz/Dashboard1_17366149207680/PartyPerformanceDashboard)

Programming Languages**:** Python | R | SQL | MySQL | MATLAB

**Data Visualization & BI:** MS Power BI | Tableau | Looker | RStudio | Matplotlib | Seaborn

**Cloud & Big Data Technologies:** AWS | Microsoft Fabric | GCP | Apache Spark | Hadoop | Databricks | BigQuery | Oozie | HBase

**Database & Automation Tools:** Oracle | Microsoft SQL Server | Docker | Git | Power Automate | JIRA | Confluence | Agile | Scrum

**Certifications:** [AWS Certified Cloud Practitioner](https://www.credly.com/badges/e23cfc99-837d-4ca6-9d3e-d52e807b0ed4/public_url) | [Generative AI with Large Language Models](https://www.coursera.org/account/accomplishments/certificate/Q8Z6ZIMPZ9JS) | [Google Data Analytics Professional](https://www.credly.com/badges/cbae6254-face-4e98-83a8-11d6f49a6f45)

**PROFESSIONAL EXPERIENCE**

**New Jersey Equity in Commercialization Collective |** Data Analyst Researcher – Newark, NJ **01/2024 – 12/2024**

* Enhanced data quality by reducing FPR by 60% in university name identification using regex and fuzzy matching in Python and processing 13GB of USPTO’s PatEx data stored in a **Microsoft Fabric’s Data Lakehouse**
* **Identified over 3000 male & female inventors using Fabric’s Python notebook based on inventor names across 8 NJ institutions**
* Increased inventor visibility by 20% by building a Power BI dashboard through Dataflow Gen2, analyzing gender demographics
* Orchestrated a scalable ETL pipeline in Data Factory, streamlining data collection, organization, and transformation for evolving datasets
* Enabled strategic decision-making by **using T-SQL within SQL Analytics Endpoint to query patent data**, generating actionable insights to non-technical stakeholders of NJECC, supporting women inventors
* Improved record match rate by 11% by automating REST API calls in Data Factory to validate genders using public sources like LinkedIn and university web pages, ensuring higher accuracy for analysis

**New Jersey Institute of Technology |** Teaching Assistant – Newark, NJ **01/2024 – 12/2024**

* Accelerated assignment scores by 15% through 6 lab exercises in RStudio on Convolutional Neural Networks, Data Analysis and Large Language Models (LLMs), using real-world datasets to simulate industry challenges
* Boosted class average by 17% by delivering lessons in Python and MATLAB, providing feedback, and additional course materials to solidify foundational programming and analytical concepts
* Conducted a workshop on data visualization techniques and statistical modeling using Python and R, demonstrating best practices for data cleaning, wrangling, and interpreting complex datasets to derive actionable insights

**Zenoti India Pvt Ltd |** OperationsData Analyst – Hyderabad, IND **07/2022 – 07/2023**

* Improved customer training strategies by designing 10+ KPIs using DAX within Power Query to analyze undertrained customer accounts
* Increased biweekly course completions by 200% by delivering region-specific insights to managers, enabling targeted improvements
* Reduced churn in high-value key accounts by building an intuitive Power BI dashboard (Google Analytics connector) to track go-live activity, churn metrics, and operational KPIs
* Boosted quarterly revenue by $20,000 through data-driven retention strategies that prevented seasonal churn
* Automated weekly employee reports, saved 10+ hours by developing Power Automate flows, to extract, format using Python and HTML
* Bolstered customer onboarding rate by 30% by troubleshooting integration issues via JIRA, managing Northpass & Zenoti University ,and facilitating a seamless transition from legacy systems to Zenoti
* Collaborated with cross-functional Agile Scrum teams to refine user stories and align development with business goals, ensuring timely execution of testing and feature rollouts

**Infor India Pvt Ltd |** Data Solutions Consultant – Hyderabad, IND **04/2022 – 06/2022**

* **Reduced manual intervention by 30%** by optimizing data migration solutions in Infor LN using **SQL, Postman, and CRUD operations**, while leveraging **Excel VBA macros** to automate data extraction and transformation for reporting
* **Developed 10+ custom 4GL reports** in Infor LN, integrating **Pivot Tables and VLOOKUP** in Excel to consolidate and analyze supply chain data, reducing stockouts by **15%** and improving inventory visibility; **Customized ERP workflows** to meet client-specific requirements, collaborating with **R&D and customization teams** to implement tailored data-driven solutions

**ACADEMIC PROJECTS**

**Election Data Visualization and Analysis 10/2024 – 12/2024**

* **Transformed a 91,000-row U.S. election dataset into an interactive Tableau dashboard**, uncovering insights on voter turnout trends and party victory margins, enhancing understanding of voter engagement and performance
* **Designed dynamic complementary visualizations and filters in Tableau**, allowing personalized data exploration by state, gender, and year, tripling student engagement and attracting 500+ weekly dashboard views

**Time Series Forecasting on Divvy Bicycle Sharing System 10/2023 – 12/2023**

* **Developed an ETL pipeline** using **Spark jobs** to preprocess **10 years of ride-sharing data (25M+ records) from an S3 bucket**, enabling ride duration prediction and bike allocation optimization
* Performed exploratory data analysis (EDA) with Matplotlib, Seaborn, identifying trends in ride duration, user behavior, and station usage
* Trained and fine-tuned time series models (ARIMA, VAR, Prophet) reducing forecasting error by 20%, improving ride duration prediction accuracy for better resource management
* **Deployed models using Docker** and established **CI/CD pipelines** with GitHub Actions, ensuring **reproducibility, continuous integration, and automated deployment**

**AWS Cloud-Based Image Recognition System 10/2024 – 12/2024**

* **Optimized object detection and image captioning** by implementing parallel processing across **two EC2 instances**, improving workload distribution and achieving over 90% detection accuracy
* **Developed and deployed a scalable image recognition pipeline** using **AWS S3, SQS**, integrating **Amazon Rekognition** to reduce image processing time by **2 seconds per image**, enhancing efficiency and scalability

**Wine Quality Prediction using Spark and Docker Deployment on AWS 10/2024 – 12/2024**

* **Deployed a 4-node EMR cluster on AWS** to process **wine quality data**, enabling scalable model training and achieving an **F1 score of 0.7**7
* **Dockerized the application**, reducing deployment time by **40%** compared to traditional methods, ensuring faster, efficient model delivery
* **Developed a distributed machine learning pipeline in Spark**, implementing and evaluating **Random Forest, Decision Tree, and Linear Regression** to optimize training and model performance