#### Harshitha Marasu

https://www.linkedin.com/in/harshitha-marasu/ | Tampa, FL

### **EDUCATION**

EBOCATION	
University of South Florida	August 2023 – May 2025
Master of Science, Computer Science	GPA: 3.77
Jawaharlal Nehru Technological University	August 2019 – May 2023
Bachelor of Technology, Computer Science and Engineering	GPA: 3.46
WORK EXPERIENCE	
Reinvision Labs Pvt. Ltd.   Data Engineer   Hyderabad, India	July 2021 – June 2023

- Designed and optimized ETL pipelines for AR/VR interaction data using SQL, Apache Spark, and Pandas, achieving 99% data accuracy and improving data processing speed by 45%.
- Built predictive analytics models using Python, Tableau, and Scikit-learn, increasing user engagement by 25% through data-driven insights.
- Automated ETL workflows using Apache Airflow, AWS Glue, and CI/CD pipelines, reducing manual intervention by 40% and increasing operational efficiency.
- Designed real-time dashboards in Power BI and Tableau, enhancing data visibility and stakeholder decision-making, reducing report generation time by 60%.
- Optimized **cloud data processing** on **AWS Redshift and Snowflake**, improving **query performance by 50%** and reducing infrastructure costs by **30%**.
- Developed data ingestion frameworks integrating REST APIs, AWS Lambda, and Terraform, enabling seamless data integration from multiple sources and reducing data pipeline failures by 98%.

## **PROJECTS**

## **Optimizing ETL Pipelines for Large-Scale Ad Data Processing**

Technologies: SQL, Python, Apache Spark, AWS Redshift, Airflow, Terraform

- Created an automated ETL pipeline using Apache Spark and SQL, processing over 10TB of ad impression data per month, reducing latency by 35%.
- Implemented data quality checks and error logging mechanisms, improving data integrity and reducing ingestion failures by 98%.
- Integrated orchestration tools like **Airflow, automating ETL workflows** and reducing manual intervention by **50%**, improving overall efficiency.

# Cloud-Based Business Intelligence Dashboard for TV Ad Campaign Performance

Technologies: Power BI, Tableau, AWS QuickSight, SQL, Snowflake, Python

- Built an interactive dashboard to visualize **TV ad campaign performance**, reducing **data reporting lag by 50%** and enabling **real-time decision-making**.
- Leveraged Snowflake as a data warehouse, ensuring 99.9% uptime and improving query execution time by 40%.
- Implemented row-level security and user access controls, ensuring data governance compliance and preventing unauthorized access.

### Automating Data Ingestion and CI/CD for BI Workflows

Technologies: GitHub Actions, Jenkins, Terraform, AWS Lambda, Python

- Facilitated an end-to-end **CI/CD pipeline for data ingestion workflows**, using **GitHub** Actions and **Jenkins**, reducing deployment time by **60%**.
- Established infrastructure provisioning with **Terraform and AWS Lambda**, eliminating manual errors and improving system scalability.
- Executed test-driven development **(TDD) principles**, enhancing system reliability by **30%** and ensuring seamless integration of new data sources.

### **PROFESSIONAL SKILLS**

**Tools:** Microsoft Office, Advanced Excel, Microsoft Power BI, Tableau, Google Sheets, AWS, GCP, Azure, AWS Quick Sight. **Programming and Databases:** Python, Java, R, SQL (MySQL, PostgreSQL, Oracle, MongoDB), C, C++, C#, JavaScript, Shell Scripting, HTML5, CSS3, NoSQL, Airflow, ETL pipelines, Hadoop, Snowflake, Databricks, Airflow, Data warehousing, AWS Redshift, Spark **DevOps & CI/CD:** Git, GitHub Actions, Jenkins, Terraform, Docker, Kubernetes.

**Skills:** Data Analysis, Data Visualization, Statistical Analysis, Machine Learning, Predictive Analytics, Data Storytelling, A/B Testing, ROI Analysis.

**Data Processing & Analytics:** Pandas, NumPy, SciPy, TensorFlow, Keras, Scikit-learn, Seaborn, Matplotlib, Streamlit, PyCharm, DAX, Snowflake, Databricks.

## **CERTIFICATIONS**